DIGITAL ARCHIVE EXPO

ASTANA 08-12.06.2026

III АРХИВШІЛЕР КОНГРЕСІ · III CONGRESS OF ARCHIVISTS

«Digital Archive EXPO»
III Архившілер конгресінің материалдар жинағы

Сборник материалов
III Конгресса архивистов «Digital Archive EXPO»

Collection of materials III Congress of Archivists
«Digital Archive EXPO»

УДК 930.25

ББК 79.3
D33

Главный редактор
Директор Архива Президента Республики Казахстан, кандидат исторических наук
А.Х. Мустафина

Редакционная коллегия
Заместитель директора Архива Президента Республики Казахстан,
доктор исторических наук **К.Ш. Алимгазинов** (отв. редактор)
Руководитель Службы обеспечения сохранности и государственного учета **Г.О. Бекишева**
Руководитель Центра образовательных проектов и международного сотрудничества
М.Т. Жылгелдинов
Руководитель Службы цифровых технологий и инноваций **К.М. Кабарова**
Руководитель Службы организации и экспертизы архивных фондов **Г.Ж. Мукажанова**
Руководитель Пресс-службы **А.А. Сакурова**
Заведующий сектором научных публикаций Центра научных и прикладных исследований,
кандидат исторических наук **Е.С. Сыздыкова**
Руководитель Центра научных и прикладных исследований,
доктор исторических наук **Н.У. Шаяхмет**

Технический редактор
Старший научный сотрудник Центра научных и прикладных исследований
кандидат педагогических наук **Б.С. Балгазина**

«**Digital Archive EXPO»**. Сборник материалов III Конгресса архивистов. Астана: Архив Президента Республики Казахстан, 2026. 581 с.

ISBN 978-601-7149-30-7

Сборник материалов III Конгресса архивистов «Digital Archive EXPO» (Астана, 8-12 июня 2026 г.) включает доклады, аналитические материалы, экспертные выступления и практические разработки, представленные в рамках международного форума, посвященного современным тенденциям развития архивной отрасли в условиях цифровой трансформации и развития технологий искусственного интеллекта.

Издание отражает результаты работы пленарной сессии, тематических панельных дискуссий, международного семинара UNESCO, экспертных сессий. В центре внимания авторов находятся вопросы цифровизации архивного дела, применения искусственного интеллекта в управлении документальным наследием, обеспечения долговременной сохранности цифровых данных, развития инновационных архивных экосистем, совершенствования нормативно-методических подходов, формирования новых профессиональных компетенций архивистов и расширения международного сотрудничества.

Сборник адресован архивистам, специалистам в области документационного обеспечения управления, исследователям, преподавателям, IT-экспертам, представителям органов государственной власти, обучающимся профильных образовательных программ и всем, кто интересуется вопросами сохранения, управления и использования документального наследия в условиях цифровой трансформации общества.

УДК 930.25
ББК 79.3

ISBN 978-601-7149-30-7

© Архив Президента РК, 2026

КОНГРЕСТІҢ САЛТАНАТТЫ АШЫЛУЫ
Құттықтау сөздер

ТОРЖЕСТВЕННОЕ ОТКРЫТИЕ КОНГРЕССА
Приветственные слова

OFFICIAL OPENING CEREMONY OF THE CONGRESS
Welcome Addresses

Қазақстан Республикасының
Президенті

Астана қ., Ақорда 09 / 06 / 2026

**«Digital Archive EXPO»
тақырыбына арналған
III Архившілер конгресіне
қатысушыларға**

**Құрметті конгреске қатысушылар!
Қадірлі қонақтар!**

Баршаңызды III Архившілер конгресінің ашылуымен құттықтаймын!

Озық технологиялар мен жасанды интеллектінің жаппай енгізілуіне байланысты архив саласы дамудың мүлдем жаңа кезеңіне қадам басып келеді. Бүгінгі форум – соның айқын көрінісі. Сондықтан, алқалы жиынды Қазақстанның басты технологиялық хабы – «Alem.ai» халықаралық орталығында өткізудің символдық мәні зор.

Отандық архив ісінің мамандары атқарып жатқан жұмыс ЮНЕСКО-ның және Халықаралық архив кеңесінің (ICA) бүкіл адамзатқа ортақ құжаттамалық мұраны цифрландыру жөніндегі бастамаларымен толық үндеседі. Елімізде жарияланған Цифрландыру және жасанды интеллект жылы аясында «Ақылды архивтер» желісін қалыптастыру негізгі басымдықтың бірі ретінде айқындалды. Қазір менің тапсырмаммен «Ұлттық цифрлық архив» жүйесі әзірленіп жатыр.

Бұл – Қазақстанның құжаттамалық мұрасын сақтауға мол мүмкіндік беретін жоба. Мұндай бастамалар жаңа Конституциямыздың жасампаздық рухына толық сай келеді. Себебі, Ата заңда мәдениет пен білім, ғылым мен инновация құндылықтарының мән-маңызы айрықша екені нақты бекітілді.

Конгресс аз уақыт ішінде ғалымдар мен сарапшыларды, халықаралық ұйым өкілдерін ортақ іске жұмылдырып, тиімді диалог алаңы ретінде қалыптасты. Келелі жиын архив саласын жан-жақты дамытуға, осы бағытта заман талабына сай білікті мамандар даярлауға және болашағы зор цифрлық жобаларды табысты жүзеге асыруға септігін тигізеді деп сенемін.

Еңбектеріңіз жемісті болсын!

Қасым-Жомарт Тоқаев

Құрметті әріптестер, Конгресс қонақтары мен қатысушылары!

Қазақстан Республикасы Президенті Архивінің атынан сіздерді III Архившілер конгресінің ашылуымен шын жүректен құттықтаймын! Конгреске қош келдіңіздер!

Бүгінгі форумға келген еліміздің мемлекеттік билік органдары өкілдеріне, Сенат және Мәжіліс депутаттарына, елшілер корпусына, ғалымдар мен архившілер қауымына алғысымызды білдіреміз!

Бұл Конгресс алдыңғы I және II Конгрестердің идеологиялық жалғасы болып табылады. Бүгінгі Конгреске әлемнің 30 елінен 500-ден астам қатысушы тіркелді. Жалпы, біздің ортамызда кездейсоқ адам жоқ. Сондықтан, дәл осы жерде біздің кәсіби болмысымыздың жүрегі соғып тұр!

Ардақты зиялы қауым! Мұндай жаһандық деңгейге ішкі қуатты іргетассыз қол жеткізу мүмкін емес еді.

Мемлекет басшысының теңдессіз қамқорлығы мен жүйелі мемлекеттік қолдау – еліміздің барлық деңгейдегі архивтер желісін бүгінгі цифрлық орта біріктіріп отыр! Оны ерекше атап өткім келеді.

Биылғы Конгрестің әрбір элементі терең символдық мәнге ие.

Мемлекеттілік символы. Биыл бүкілхалықтық референдумда қабылданған жаңа Конституцияда тарихи-мәдени мұраны сақтау мемлекет қызметінің басты қағидаттарының бірі ретінде бекітілді.

Уақыт символы. Елімізде 2026 жыл – цифрландыру және жасанды интеллект белгісімен өтуде.

Біз үшін ол жай ғана тренд емес, бұл тарихи міндет.

Дата символы. Конгресс архившілердің кәсіби мерекесі – 9 маусым, Халықаралық архивтер күніне сәйкес келіп отыр.

Осы сәтті пайдалана отырып, архив мамандарын кәсіби мерекесімен шын жүректен құттықтаймын!

Сонымен бірге, бүгінгі алқалы жиын Біріккен Ұлттар Ұйымы, ЮНЕСКО және Халықаралық архивтер кеңесімен серіктестік рухында өтуде.

Құрметті әріптестер! Алдымызда бізді ауқымды бағдарлама күтіп тұр.

Біздің Конгрестің бірегейлігі - оның пәнаралық сипатында. Конгресс тұңғыш рет көрме мен сессиялардың гибридті форматында өтеді.

Жасанды интеллект технологиялары архившілерге қажет екені анық. Дегенмен, алгоритмдер тарихи дерексіз жұмыс істей алмайтыны белгілі. Сондықтан, біз бір-бірімізге қажетпіз.

Олай болса, цифрлық архивтерді дамытудың жаңа дәуірін бірге ашайық !

Назарларыңызға рақмет!

Әлия Құсайынқызы Мұстафина,
Қазақстан Республикасы Президенті Архивінің директоры,
тарих ғылымдарының кандидаты

Excellencies, Distinguished Guests, Dear Colleagues!

It is a great honour to join you today for the opening of the Third Congress of Archivists and the Digital Archive Expo. Allow me, first of all, to express my very sincere gratitude to the Archive of the President of the Republic of Kazakhstan and to the International Council on Archives for bringing together such a distinguished group of international experts. You have created a privileged platform for dialogue, cooperation and innovation at a time when our profession is undergoing profound change.

I would like to commend the Republic of Kazakhstan for its leadership in the preservation, management and sharing of knowledge and for its leading role in fostering dialogue and cooperation at regional and global levels. Within the United Nations system, Kazakhstan has established itself as a highly valued multilateral leader, which brings together Member States and the full United Nations family around critical challenges, such as sustainable development and water management – to mention just a few.

It is therefore particularly fitting that the Congress takes place here, in a country that understands so well both the importance of preserving history and the value of building bridges toward the future, always demonstrating the highest level of ambition.

Within the United Nations family, the UN Library and Archives Geneva has a very particular position and role. Our institution was founded as the Library of the League of Nations in 1919, based on a donation from John D. Rockefeller, to serve as a centre of research and an instrument of international understanding. Based on that heritage, our mission today is to preserve institutional memory and to provide access to authoritative knowledge in support of the work of the United Nations and the international community. We appreciate greatly that the dimension of institutional

memory within international organizations, as guardians of the heritage of the multilateral process, are integrated in the very timely discussions at the Congress.

We are living through a period of technological transformation, reshaping the way information is created, managed, preserved and accessed. Artificial intelligence offers new possibilities for processing and describing archival collections, enhancing discoverability and supporting more efficient records management.

Advances in digitization technologies are enabling the preservation of fragile materials while broadening access to documentary heritage across geographical boundaries. Digital platforms allow archives to reach new audiences and to support transparency, accountability, inclusion, education and research on an unprecedented scale. We experienced this first-hand at the UN Library & Archives Geneva with the ambitious Total Digital Access to the League of Nations Archives project – also called LONTAD – which successfully digitized the entire archives of the League of Nations. I will be describing this project in more detail in the Plenary Session later this morning.

These opportunities are accompanied by equally important responsibilities. As custodians of documentary heritage, we must ensure the authenticity and integrity of digital records. We must strengthen our capacity to protect collections against cyber threats and technological obsolescence. And, we must address the ethical implications of emerging technologies, ensuring that innovation remains aligned with the principles of trust, reliability and accountability.

In short, we must ensure that future generations inherit not only data, but genuine knowledge.

For international organizations, these responsibilities carry particular significance. Our archives document the history of international cooperation, diplomacy, humanitarian action, scientific advancement, human rights and sustainable development – and much more. At the UN Library & Archives Geneva, we are privileged to care for records that span more than a century of multilateral history. They provide an indispensable resource for policymakers, scholars and citizens seeking to understand the evolution of our world and the institutions that shape it. These repositories evidence our common journey as an international community. Protecting and preserving this heritage is therefore a responsibility that extends beyond any single institution or country. It is a collective responsibility that requires international cooperation, professional solidarity and readiness to innovate.

This Congress offers an invaluable opportunity to advance that collective effort. I look forward to a rich exchange and to learning from the incredible expertise assembled here. I am confident that the conversations held during this Congress will help to shape the future of archives and records management through sharing and through partnership. I wish the Congress every success and thank you once again for your warm hospitality.

Thank you very much.

Charlotte Lindberg Warakaulle,
Director of the UN Library and Archives Geneva

May I recognize all the distinguished guests gathered here, all the participants, ladies and gentlemen!

On behalf of UNESCO, I would like to express my gratitude for the invitation by the Archive of the President of the Republic of Kazakhstan to participate in the proceedings of this III International Congress of Archivists as a representative of UNESCO's Memory of the World Programme – a flagship programme dedicated to the identification, preservation, and accessibility of the richly diverse documentary heritage of the world.

I'm glad to underscore that Kazakhstan has an abundant documentary heritage, some of which is inscribed on the Memory of the World International Register. We expect much more.

Needless to say, memory institutions here and all over the world face different kinds of challenges, including the risk of loss of their holdings due to natural and human-caused disasters. This makes the theme of this Congress – Digital Archive EXPO – all the more important.

The advent of Artificial Intelligence, while holding a great deal of promise for the digital preservation of paper and audiovisual archives, remains a challenge for policymakers, memory institutions, and civil society.

We are here to share experiences to this end, to learn from each other and, more importantly, to show solidarity with one another in this ever-changing world that sometimes causes us to pause and ask tough questions.

In a world of misinformation and disinformation, archives and records remain a source of truth. But what is truth anyway in the era of Artificial Intelligence? Is investment in archives threatened as a result of all this? What can we do as archival

professionals and advocates? These are troubling questions, nevertheless, they are ones that must be asked.
As we enjoy this Expo, these questions will be at the back of your and my mind. My hope is that this Congress, as it has done so well in the past, will provide a soft landing for our minds to engage with these difficult issues and yet, in the spirit of shared humanity, networking, if you will, for shared solutions.
Thank you.

Fackson Banda,
Chief of Documentary Heritage Unit,
Communication and Information Sector, UNESCO

Excellencies, distinguished guests, dear friends! Сәлеметсіз бе, құрметті әріптестер мен достар!

It is a great privilege and honor for me to join you today at the opening of this Congress. To be together on this occasion is deeply meaningful because today is International Archives Day, a moment that we are fortunate to celebrate together in our shared commitment to archives and their enduring value.

The 9th of June is the date on which the International Council on Archives was founded in Paris in 1948, which means that today we are not just opening a Congress together; we are marking the 78th anniversary of an organization built on the belief that archives are essential to society –essential to justice, to memory, to identity, and to trust between citizens and the institutions that serve them. This is a day to celebrate, and there is no better place to do so than here, among colleagues who put this belief into everything they do.

Allow me to express my deepest gratitude to Director Aliya Mustafina of the Archive of the President of the Republic of Kazakhstan for bringing us together in this way. The Congress is an excellent example of international collaboration, bringing together archivists, experts, representatives of state archival institutions, the academic community, and international organizations. Initiatives such as this demonstrate the potential of Central Asia as an active participant in the international archival dialogue. Allow me also to express my sincere gratitude to FamilySearch for its generous support in helping organize this conference and make it happen. Thank you for your trust, your partnership, and your continued support.

Today also marks an important step because the International Council on Archives and the Archive of the President of the Republic of Kazakhstan are signing a Memorandum of Understanding as the foundation of a long-term partnership,

through which ICA commits to supporting the growth of this regional network, in which the archives of Central Asia take their rightful place as active, engaged, and influential members of the global archival family. ICA and the Archive of the President of the Republic of Kazakhstan will strengthen archival collaboration in Central Asia through joint initiatives. The archivists of this region bring the experience of preserving multilingual documentary heritage across vast geographies, including manuscripts, traditions, and modern digital systems, all in the same professional hands. This is the expertise the global community needs.

So today is an opportunity to learn from one another, to strengthen our connections, and to reaffirm the values that bring us together. Let us use this moment to inspire new ideas and renew our commitment to the important work we do every day.

Thank you for being part of this journey.

Josée Kirps,
President of the International Council on Archives

Құрметті архив саласының ардагерлері мен мамандары, қадірлі ғалымдар, әріптестер, қонақтар!

Баршаңызды бүгінгі «Диджитал Архив Экспо» аясында өтіп жатқан үшінші архившілер Конгресінің ашылуымен шын жүректен құттықтаймын. Бүгінгі Конгресс тек кәсіби пікір алмасу алаңы ғана емес, бұл – ұлттық жадымызды сақтау, тарихи мұрамызды келешек ұрпаққа жеткізу және архив ісін жаңа сапалық деңгейге көтеру жолындағы маңызды қадам. Архив – өткенді бүгінге жеткізетін, ал бүгінгі күннің шежіресін болашаққа аманаттайтын бірегей институт.

Архив құжаттарында халқымыздың тағдыры, мемлекетіміздің қалыптасу тарихы, ұлттық рухани және мәдени құндылықтарымыз сақталған. Сондықтан архив – тек қана құжаттарды жинақтау мен сақтау орны емес, ол – тарихи әділдікті қалпына келтірудің, ұлттық бірегейлікті нығайтудың және қоғамдық сананы жаңғыртудың маңызды құралы. Қазіргі таңда әлем цифрлық дәуірдің жаңа кезеңіне қадам басты.

Жасанды интеллект, үлкен деректер (Big Data), бұлтты технологиялар мен цифрлық платформалар мемлекеттік басқарудан бастап ғылым мен білімге дейінгі барлық саланы түбегейлі өзгертуде. Архив саласы да бұл үдерістен шет қалмауы тиіс. Бүгінгі көрме мен Конгресс архив ісін цифрлық трансформациялау аясындағы үздік тәжірибелерді таныстырып, жаңа технологияларды енгізу, құжаттардың қолжетімділігі мен сақталуын қамтамасыз ету, сондай-ақ халықаралық ынтымақтастықты дамыту мәселелерін талқылауға мүмкіндік береді.

Біздің басты мақсатымыз – архив қорларын цифрландыру арқылы азаматтар мен зерттеушілер үшін тарихи деректерге қолжетімділікті дамыту, құжаттардың

сақталу қауіпсіздігін күшейту және архив қызметтерін заман талабына сай жаңғырту. Бұл бағыттағы жұмыстар ұлттық мұрамызды сақтаумен қатар, ғылыми зерттеулердің дамуына, тарихи сананың қалыптасуына және мемлекеттік институттарға деген қоғамдық сенімнің артуына да ықпал етеді.

Бүгінгі таңда архившінің рөлі бұрынғыдан да маңызды. Архив маманы – тек қана құжат сақтаушы емес, ақпараттық қоғамның белсенді қатысушысы, цифрлық деректерді басқарушы және тарихи білімді насихаттаушы болып табылады. Сондықтан архив саласының кадрлық әлеуетін нығайту, жас мамандарды тарту және кәсіби біліктілікті арттыру – ортақ міндетіміз.

Құрметті әріптестер! Архив – ұлттың жады, ал ұлттық жады берік болса, мемлекеттің де болашағы баянды. Сондықтан біз өткенді сақтай отырып, болашақтың цифрлық архивін бүгіннен бастап қалыптастыруымыз қажет.

Үшінші архившілер конгресінің жұмысына табыс, мазмұнды пікірталастар мен нәтижелі бастамалар тілеймін! Баршаңызға зор денсаулық шығармашылық шабыт және архив ісін дамытудағы еңбектеріңізге толағай табыс тілеймін. Назарларыңызға рахмет.

Руслан Рүстемов,
Қазақстан Республикасы Парламенті Сенатының депутаты

Congratulatory Remarks at the Opening Ceremony of the 3rd Congress of Archivists!

I would like to extend my heartfelt congratulations on the grand opening of the Third Congress of Archivists today.

I am Furuya Ichiro, Senior Vice-President of the National Archives of Japan, and I have the honor of attending this congress representing my institution. I would like to offer a few words in greeting.

Exchanges between the Archive of the President of the Republic of Kazakhstan and the National Archives of Japan have steadily deepened in recent years. We were greatly pleased to welcome representatives from the Archive of the President to our institution in both 2024 and 2025; in particular, the visit by Director Aliya Mustafina herself in 2025 was a significant event that symbolized the trust between our two institutions.

Furthermore, in the fall of 2025, we held discussions between our two institutions on the occasion of the ICA Barcelona Congress, and in the winter of that same year, during the visit to Japan by President Kassym-Jomart Tokayev of the Republic of Kazakhstan, we were able to exchange a Memorandum of Cooperation between our two institutions. This memorandum was positioned as one of the outcome documents of the summit between the leaders of both nations, serving to demonstrate both domestically and internationally that cooperation in the field of archives is a key pillar of bilateral relations. In addition, at an online seminar for professionals hosted by our institution this past January, many colleagues from the Archive of the President participated, further accumulating tangible achievements in our cooperative activities.

For our two institutions, which have built such a close relationship, it is a great honor for me, representing the Japanese side, to visit the Republic of Kazakhstan for the first time, participate in this international academic conference, and have the opportunity to address you today.

The Republic of Kazakhstan is extremely active in participating in international initiatives across a wide range of fields, including the archival sector. Its dynamic progress is strongly reflected in both the scale of this congress and its ambitious themes. We fully share the view that it is of great significance to create a forum where experts from diverse regions of the world gather to share cutting-edge knowledge and practical experience, and we hope to benefit together from the outcomes of this congress.

In closing, I would like to express my heartfelt gratitude to the staff of the Archive of the President of the Republic of Kazakhstan for their tremendous efforts in organizing this congress. I sincerely hope that this congress will be a fruitful one for all participants and contribute to the further development of international cooperation in the archival field.

Furuya Ichiro,
Senior Vice-President of the National Archives of Japan

Уважаемый президент Международного совета архивов Жозе Кирпс! Уважаемые участники Конгресса! Уважаемые дамы и господа!

Разрешите от имени руководителей и сотрудников архивных служб и учреждений Республики Таджикистан выразить глубокую признательность и намерение на дальнейшее расширение профессионального сотрудничества в сфере сохранения документального исторического наследия.

Пользуясь случаем, хочу искренне поблагодарить Международный совет архивов, директора Республиканского государственного учреждения «Архив Президента Республики Казахстан», кандидата исторических наук уважаемую госпожу Мустафину Алию Хусаиновну и всех организаторов III Конгресса архивистов за Ваше приглашение принять участие в данном мероприятии.

Конгресс является международной платформой для профессионального диалога архивного сообщества Центральной Азии для обсуждения проблем по вопросам в сфере применения искусственного интеллекта в архивных учреждениях, а также разработки национальных стратегий внедрения искусственного интеллекта в архивах.

В регионе Центральной Азии первым важным документом в сфере искусственного интеллекта является принятая Стратегия развития искусственного интеллекта в Республике Таджикистан на период до 2040 года. Стратегия способствует развитию искусственного интеллекта и используется во всех сферах социально-экономического развития, учитывая процессы внедрения и использования передовых и современных технологий.

Уважаемые коллеги! После обретения независимости Республики Таджикистан по инициативе Основателя мира и национального единства, Лидера нации, Президента Республики Таджикистан уважаемого Эмомали

Рахмона в сфере развития архивного дела в республике произошли коренные изменения. Был принят Закон Республики Таджикистан «О Национальном архивном фонде и архивных учреждениях», а также ряд других нормативно-правовых документов в архивной сфере.

Главное архивное управление при Правительстве Республики Таджикистан является уполномоченным центральным органом государственной исполнительной власти, осуществляющим управление архивным делом в Таджикистане. Важной частью историко-культурного наследия таджикского народа является Национальный архивный фонд Республики Таджикистан.

В настоящее время работу с Национальным архивным фондом Республики Таджикистан ведет широкая сеть архивных учреждений республики, которая включает 13 государственных архива и их филиалов, а также 64 межведомственных архива по личному составу при исполнительных органах государственной власти, в архивохранилищах которых хранятся 38 312 фондов, и 11 599 825 дел.

Под руководством Лидера нации, Президента страны постепенно реализуется политика цифрового развития и внедрение технологий искусственного интеллекта. По инициативе Основателя мира и национального единства, Лидера нации, Президента Республики Таджикистан уважаемого Эмомали Рахмона Региональная Ассамблея Организации Объединенных Наций 25 июля 2025 года приняла резолюцию «Роль искусственного интеллекта в создании новых возможностей для устойчивого развития в Центральной Азии», по которому предусматривается создание Регионального центра искусственного интеллекта в г. Душанбе, который будет координировать совместные программы по внедрению соответствующих технологий в странах Центральной Азии.

Следует отметить, что 28 апреля 2026 года было принято Постановление Правительства Республики Таджикистан «О Плане действий по цифровизации архивов государственных органов Республики Таджикистан на 2026–2030 годы», целью которого является перевод архивов государственных органов в электронный формат, а архивы на бумажной основе в текущем делопроизводстве к 2030 году планируется полностью исключить.

Заканчивая свое выступление, выражаю надежду, что III Конгресс архивистов явится важным шагом для конструктивного обмена мнениями по актуальным вопросам в деле укрепления сотрудничества между архивными учреждениями стран в сфере развития и применения искусственного интеллекта и обмена опытом в архивном деле.

Хочу пожелать успехов в работе Конгресса. Благодарю за внимание!

Шарифзода Лутфия Шариф,
начальник Главного архивного управления
при Правительстве Республики Таджикистан

«DIGITAL ARCHIVE. ПАРАДИГМАНЫҢ АУЫСУЫ»
Пленарлық сессия

Пленарная сессия
«DIGITAL ARCHIVE: СМЕНА ПАРАДИГМЫ»

«DIGITAL ARCHIVE: A PARADIGM SHIFT»
Plenary Session

ARCHIVES IN THE AI ERA

Aliya Mustafina
Director of the Presidential Archives of the Republic of Kazakhstan,
Candidate of Historical Sciences, Kazakhstan
office@archive.president.kz

ЖИ ДӘУІРІНДЕГІ АРХИВТЕР

Әлия Құсайынқызы Мұстафина
Қазақстан Республикасы Президентінің Архиві директоры,
тарих ғылымдарының кандидаты, Қазақстан
office@archive.president.kz

АРХИВЫ В ЭПОХУ ИИ

Алия Хусаиновна Мустафина
Директор Архива Президента Республики Казахстан,
кандидат исторических наук, Казахстан
office@archive.president.kz

Dear colleagues!

Today, we are entering an era of dynamic Archives-ecosystems, that do not only preserve data but also generate value. For a couple of years now the world faces technological race of Artificial Intelligence. President of the Republic of Kazakhstan, Kasym Jomart Tokayev in his speech during the International technological forum “Digital Bridge 2023” in Astana has noted, that the Artificial Intelligence technology is as revolutionary as were in their due time electricity and the Internet.

2026 in Kazakhstan was declared as the year of Digitalization and Artificial Intelligence. The country has confirmed its commitment to the course towards comprehensive modernization and strengthening of own position in the global digital environment. By adopting the new Constitution (2026), our State on the highest level has secured Digital Rights of the citizens and the Responsibility to steward historical and cultural heritage, to safeguard the monuments of history and culture. And this is not just a legislative innovation, but a strategic step, defining the direction of our state's (nation's) development, a statement of the prioritization of digitalization of all economy sectors.

New starts to acquire features, structure, operating mode. Large scale state events became a prosperous highway for conceptual support, strategic planning, operational functioning including archival field as well.

Today, Archive of the President is more than 1,1 thousand fonds, including more than one and a half million archival storage units. This is total of 28 km of shelf length in Almaty and 40 km of archive racks in Astana! Those are records of socio-political history of the entire 20th Century. This is a documentary heritage of independent Kazakhstan beginning from 1991. Unprecedented was the acceptance within two years of 2022-2023 of the Archival fonds on the history of the tragic pages of 20th Century, that in the volume exceeds the entire 100 years of the political history documents.

Archive of the President keeps up with the times. Digital transformation is of a paramount importance in our activity and our development.

Dear colleagues!

Here we stand, at the threshold of the fundamental change of the archival paradigm. What awaits Archives? What is going on with our records? Where are we heading? Can the archives overcome the digital challenges? I think those questions trouble everyone present here.

The format of immersion into studied issues is well known to the European and international public. Exhibition activities also firmly entered archive practice. During more than one and half centuries of Archivistics (alt. archival science) an entire direction was formed – information dissemination (through exhibitions, projects et al.). And this is well known!

It wasn't by chance that we picked up the idea of EXPO. As I think is reflects internal state of current Archivistics (alt. archival science) – the desire to be heard, to draw the attention of the society to the issues of archives and to attempt to convey an idea through widely understandable forms, that the ocean of primarily digital information can be managed and can be comprehensible.

We already see a global alliance of international community of archivists, where ideas of the architecture of the digital future can be developed. And I am sure that the archives are the very lifeblood that ensures functioning of the State. It is in this projection that I see the significance of the III International Congress of Archivists "Digital Archive – EXPO" (DA-EXPO).

This endeavour of ours received the support of the International Council on Archives (ICA). We highly appreciate it.

This Congress is a logical continuation of the ideology of our two previous Congresses – "Open Society – Open Archive" (2022), and "Heritage Institutions on the Way to Sustainable Development of Society" (2024).

The locomotives of the ideas of our Congresses are the platforms of the International Congress on Archives in Abu-Dhabi (2023) and Barcelona (2025).

The III Congress will allow us to discuss a new role of archives – their role of the strategic center of the data verification, preventive informational hygiene and compliance with the Universal Declaration of Human Rights (1948) in the digital realities.

Right to information (RTI) – is a fundamental human right. Right to information contributes to stability of the functioning of State apparatus. Similar to how on the national level unified digital platforms for managing key resources are created, we position archive as a fundamental digital platform of historical memory, crucial for the stability of the State, effectiveness of the business and public awareness.

Points of growth

Where are points of growth of the archival field concentrated today?

Answer, that the first comes to mind – digitalization and transformation of the key archival processes that is driven by it: acquisition, storage and usage. However, in the last few years a fundamentally new variable emerged on the orbit of archival continuum – Artificial Intelligence (AI).

On one side, AI opens up wide opportunities in front of the archives: optimization of the document selection processes, automation of data input, recognition and analysis, substantial increase of the accuracy and speed of processing. Among the particular tasks that AI solves are archival appraisal, control of the retention period, archival description, intellectual search, remote user services though the chat-bots and digital assistants. This is just a tip of an iceberg though.

What remains out of our sight and what we would like to draw a particular attention to within the framework of the Congress, it is the issue of the deep transformation and re-evaluation of the archival worldview. Well-known Canadian researcher L. Duranti has rightly warned us: following the “archival fashion”, we risk to overlook the deeper meaning of the approaching changes – those changes that archivists should not only comprehend, but also accept as a new reality.

From strategy to pioneer solutions

Among seven priority development directions of the Archive of the President until 2035, digital transformation takes a central place. This is our aspiration to adapt best international archival practices to the conditions of a particular archival institution. It must be admitted that in Kazakhstan such experience is unique, not yet typical.

During the three days of the Congress participants will be able to get acquainted with our pioneer solutions for the AI implementation. Our undertaken methodology of moving from electronic archive towards intellectual archive is intended to solve the task of the formation of unified and integral archival ecosystem. Our three implemented solutions convincingly demonstrate: AI can and should be used! Opportunities, perspectives and issues that we faced are in fact typical for most of the archival institutions.

Risks and challenges

AI has immense potential. We got firmly convinced in it. AI can take upon itself routine functions of archives, allowing to reroute human potential to the tasks of the higher level and bring archive to a qualitatively new level. Along with this, it needs to

be clearly understood, that AI is a tool in the hands of the human being and its effectiveness is measured by how archives are using it.

AI's ability for self-learning, self-generation and adaptive algorithm construction raises all known risks that are more actively discussed among professional community. Self-preservation instinct pushes towards pushing the brakes on this process – to better prepare, to adapt, to wait out. Nevertheless, the Pandora's box is already open and the progress can't be halted. Archives should not be as an ostrich, burying the head in the sand, as the issues raised by the AI before Archival community are principled.

Will AI become the instrument of management and control? Can it replace the archivist's professional assessment? How to provide the ethical and legal validity of its application? Answers to those questions require not the passive observation, but a proactive professional engagement.

Archives as an infrastructure of trust

AI changes the whole positioning of the Archives in the society. Archive in no longer just a repository, it becomes an infrastructure of trust: environment, where data has verified origin. New role of the archive – role of the verifier of the primary sources for the court trials, journalist investigations and scientific publications acquires key meaning in the age of synthetic content distribution and disinformation.

Simultaneously the character of infrastructural challenges changes as well. If before archivists talked about the "limited archival shelf", today the focus shifts to the development of digital archival infrastructure and its energy security. The problem of preservation has acquired a systemic dependence on external factors, and this challenge is global.

In the context of total introduction of neural networks, the archival community has no right to be inert. AI will inevitably enter archival field, but without our active participation, we risk to lose control over the methodology, expertise, and, ultimately, over the function of the verification of historical truth. Paraphrasing the Latin wisdom: if you cannot resist the process, lead it.

Request for updating

As the practice of applying AI in the Archive of the President has shown, the regulatory and methodological base needs to be fundamentally updated. The technological revolution is a way ahead of the archival legislation, which comes forward as a tool for regulating new realities.

The archival community requires international standards developed broad consensus basis. This includes revision of existing and creation of new regulations, procedures and norms that take into account the specifics of the information technology. Unfortunately, the universal rules for the AI use in the archival field are yet to be formed.

That raises a legitimate question: as of today, is the community ready to adopt international or regional archival standard for the AI use?

Along with that, the question of the complimentary nature of the archival, I would say, consciousness arises. In the era of digitalization, we have an abundance of terms and categories of interdisciplinary nature. However, we have not reached a consensus on the meanings we put into those terms. For example, a national corpus of archival terms usually covers about 300 professional expressions. There are about 200

international terms that are incorporated in the country's archivistics (alt. archival science). Yet the number of categories used in the country's archival practice, but not regulated by the national legal acts – exceed 500. Think about that!

I will give you one example: IT specialists, historians, researchers, the public, the humanist scholars, engineers, politicians, all attach different meanings to the word "document" – that is the central definition of the archival studies (data, information, message, links, communication, information, source). The meaning that you and I, as archivists imbue it with is understood in a completely different way. Our specialists, studying this issue for two years, identified about 30 different interpretations. Do they match in their meanings? No! However, is it important what language the archivists speaks? It is important to agree on the meanings that we include and bring to our profession!

Personnel of the new generation

From here inevitably comes out the issue of training specialists of new generation equipped with competencies, previously absent from the professional archivist's standard. Digital technologies rapidly transformed the archivist training strategy. There is a growing demand for the staff that combine archival training with analytical, technical, engineering and digital skills, with the ability to manage data and interpret the results of automated processing. We are in need of "archival engineers" capable of designing and providing support of intellectual archival systems. This topic was strongly expressed at the venues of the ICA Barcelona Congress 2025, whenever new professional functions and areas of specialization were discussed.

At the same time, it is fundamentally important not to forget about Human intelligence, the bearer of which is the employee of the Archive. The cognitive abilities of the archivist retain their key importance, including, and above all, in the conditions of the active application of AI. The process of reassessment of the synthesis of two intellects – Human and artificial opens up the opportunity to build new professional qualifications framework, at the center of which – a Human being as the bearer of meanings, judgments and the final area of the responsibility.

The balance of openness and security

The implementation of the modern "Archives for society" paradigm inevitably confronts us with the dilemma of this responsibility – how to correlate unconditional public demand for openness with the need to protect sensitive information? How to observe the rights of the individual, ethical norms and interests of national security? Where do those boundaries pass? Are there certain "red lines" for the archives?

Archives started to understand responsibility more clearly. But where is that flawless, error-free mechanism, that allows archives to stay open for society while excluding destructive intervention into private life, state stability? The search for balance between the right of the society to know own history and the right of the individual and state to defend their legitimate interests is a key challenge of our time, requiring a response from the professional community. New facets of this challenge are connected with the solution of the problem of harmonious coexistence of the concept of "open archives" and the philosophy of the digital society. And from the position that the archivist will take, the future of the digital ecosystem of archives will be determined.

As I see it, today, more than ever, the International Code of Ethics of Archivists (1996) gains new relevance. Statement on Digital Rights (2026), officially signed by the ICA consolidated the central episteme of the progressive growth of archives, when archivist goes beyond the usual custodian of the past and becomes a mediator between the information (data), technologies and end user.

Zhumakhan Kuderin, one of the brightest representatives of the Kazakh intelligentsia of the beginning of the 20th century, defined in our Kazakh mentality the expression that is very close to the concept of "Ethics" in the meaning that is embedded in the International Code of Ethics of Archivists. The expression "aqyqatty uyat". He imbued these unambiguously interpreted two independent Kazakh words with the allegory of multi-tasking capability, revealing the meaning of the expression through socialization of the individuals in the variety of situations, thereby justifying the highest place of the Human in the hierarchy of the cognitive abilities of the living beings. Isn't that what is rooted in the anthropological shift for the Archives? Isn't that what we should think about, when we talk about transformation of the archives as a circulatory system of the documentary life support of the State.

Dear colleagues, participants and guests of the Congress!

Archives are on the threshold of the deepest transformation in the history of their existence. Digitalization, artificial intelligence, new requirements for the infrastructure, legislation, staff members and professional ethics are all our growth points! And only by joining our efforts we can create new practices that meet both global challenges as well as national interests of each country! This is the main message of our Congress.

I wish you productive work, meaningful discussion and fruitful cooperation!

DIGITIZING THE MULTILATERAL PROCESS: CHALLENGES AND OPPORTUNITIES FOR INTERNATIONAL ORGANIZATION

Charlotte Lindberg Warakaulle
Director of the United Nations Library and Archives in Geneva, Switzerland
charlottelindbergwarakaulle@gmail.com

КӨПЖАҚТЫ ҮДЕРІСТІ ЦИФРЛАНДЫРУ: ХАЛЫҚАРАЛЫҚ ҰЙЫМДАР ҮШІН СЫН-ТЕГЕУРІНДЕР МЕН МҮМКІНДІКТЕР

Шарлотта Линдберг Варакаулле
БҰҰ-ның Женевадағы кітапханасы мен архивтерінің директоры, Швейцария
charlottelindbergwarakaulle@gmail.com

ЦИФРОВИЗАЦИЯ МНОГОСТОРОННЕГО ПРОЦЕССА: ВЫЗОВЫ И ВОЗМОЖНОСТИ ДЛЯ МЕЖДУНАРОДНЫХ ОРГАНИЗАЦИЙ

Шарлотта Линдберг Варакаулле
Директор Библиотеки и архивов ООН в Женеве, Швейцария
charlottelindbergwarakaulle@gmail.com

United Nations Library & Archives Geneva

"A centre of international research and an instrument of international understanding"

- Founded in 1919 as the Library of the **League of Nations**
- 1924: donation by John D. Rockefeller for dedicated building
- Now serving the multilateral system as **UN Library & Archives Geneva** as part of the United Nations Office at Geneva
- Mission across three pillars
  - Library services
  - **Institutional memory (archives services)**
  - Cultural activities for Member States and management of artworks donated to the Organization (~ 1,300 in total)

Building B, home of the UN Library & Archives Geneva, part of the Palais des Nations, the seat of the United Nations Office at Geneva

United Nations archives Geneva

Archives: 5 categories

More than 12 linear kilometers (1870-present) with great wealth and diversity of collections

League of Nations Archives

UN Geneva Archives

Private Papers

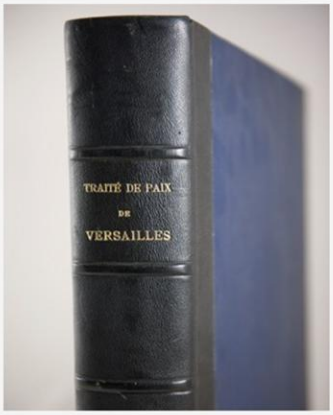
International Peace Movements

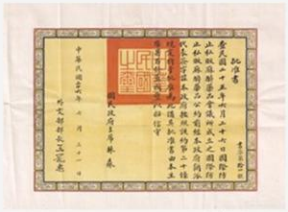
Additional Collections

United Nations archives Geneva

League of Nations Archives

- **2.7 linear kilometers** → ~15 million pages of unique documents dating 1919-1946
- **International importance** → recognized by UNESCO, with their inscription in the Memory of the World Register in 2009
- Extremely fragile and **heterogeneous material**
- Post-Cold War period bringing **increased interest of researchers in the history of multilateralism**, especially League of Nations → necessity to facilitate access to the archives of the League of Nations worldwide

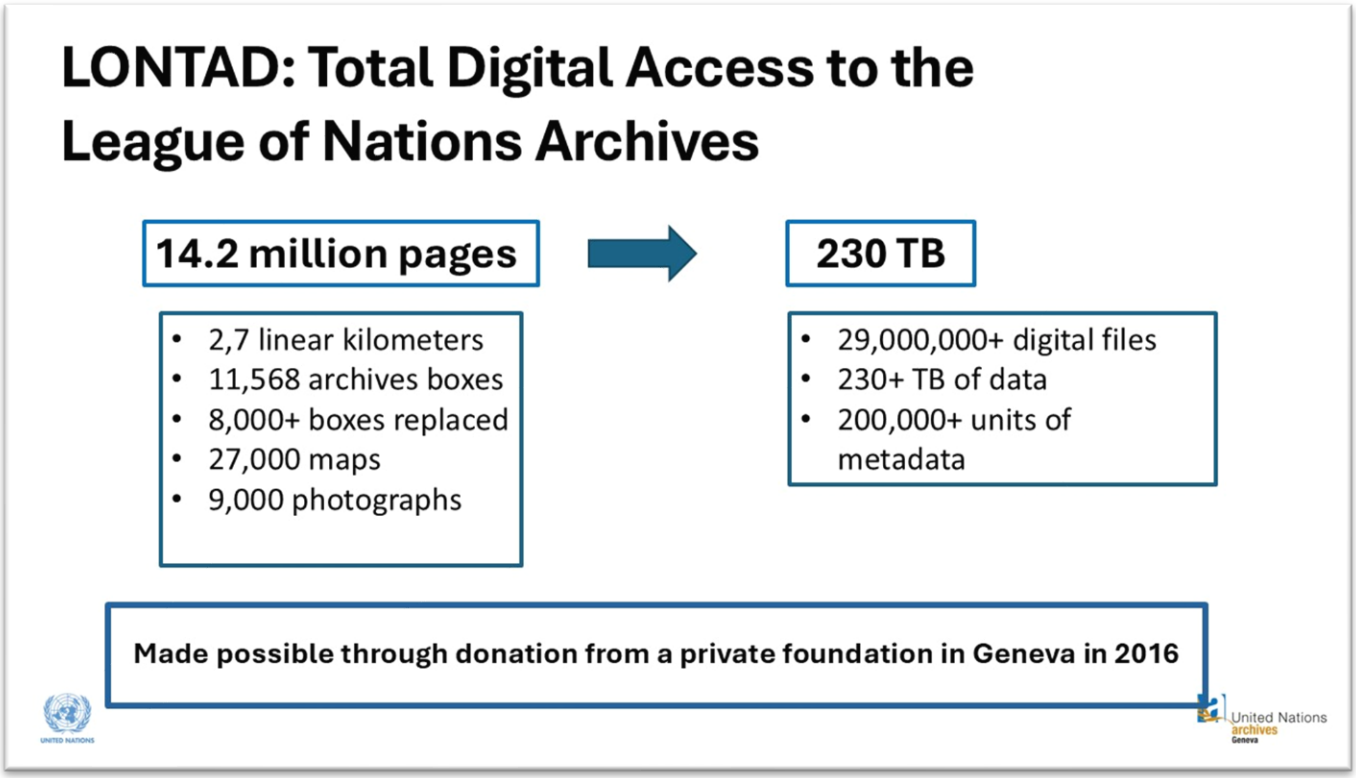


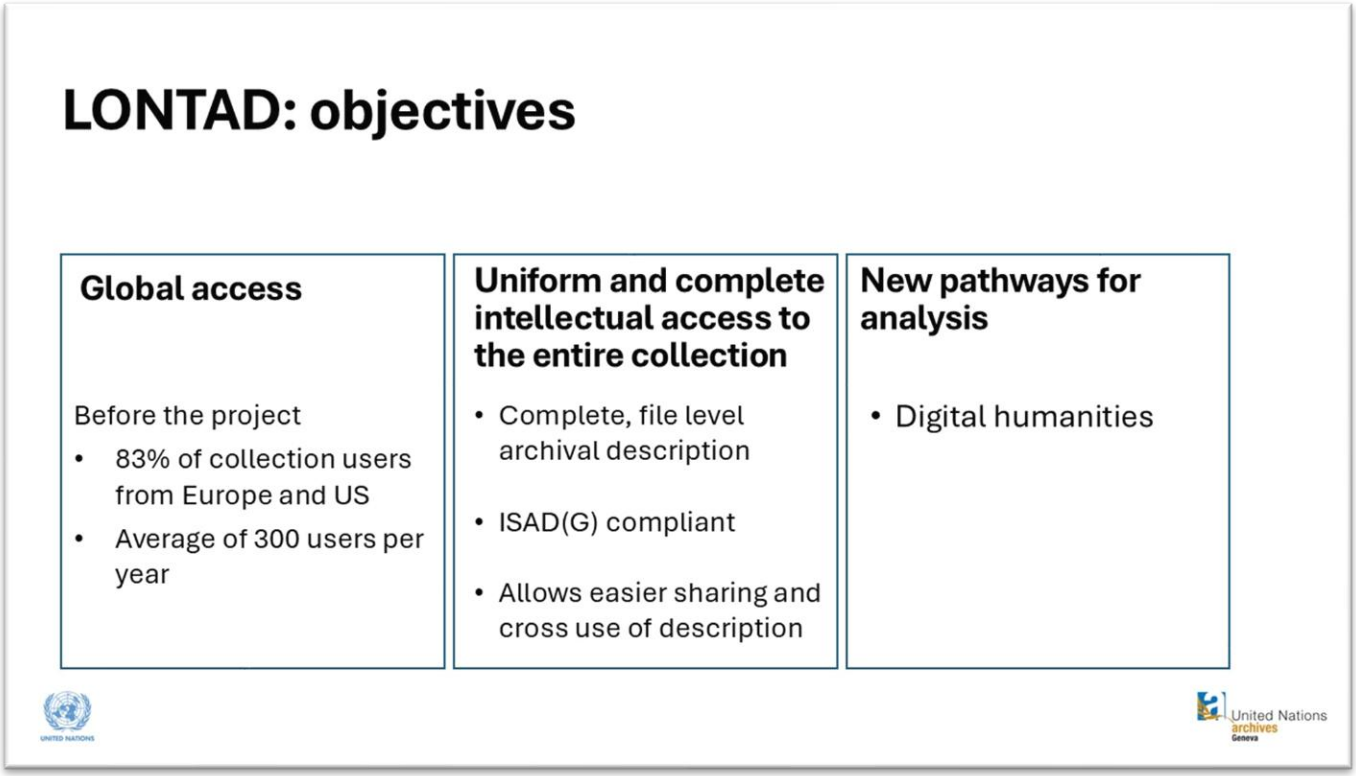


UNESCO
Organisation des Nations Unies pour l'éducation, la science et la culture

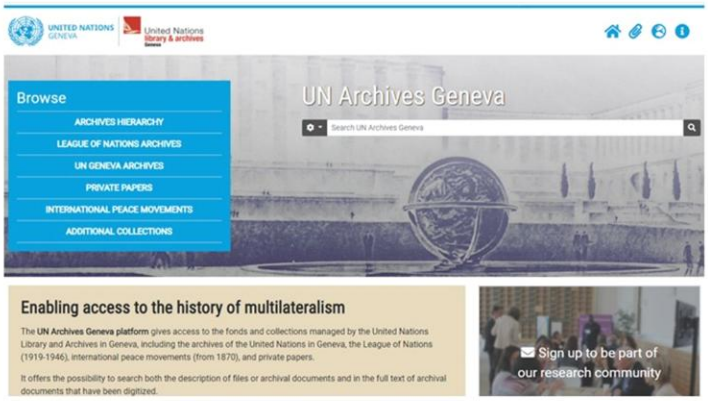
Archives de la Société des Nations 1919-1946
Inscrites au Registre en 2009
Mémoire du monde

United Nations archives Geneva


LONTAD: Total Digital Access to the League of Nations Archives
14.2 million pages
230 TB
2,7 linear kilometers
11,568 archives boxes
8,000+ boxes replaced
27,000 maps
9,000 photographs
29,000,000+ digital files
230+ TB of data
200,000+ units of metadata
Made possible through donation from a private foundation in Geneva in 2016
United Nations archives Geneva


LONTAD: objectives
Global access
Before the project
83% of collection users from Europe and US
Average of 300 users per year
Uniform and complete intellectual access to the entire collection
Complete, file level archival description
ISAD(G) compliant
Allows easier sharing and cross use of description
New pathways for analysis
Digital humanities
United Nations archives Geneva

Access: United Nations Archives Platform



AtoM : Access to Memory

https://archives.ungeneva.org/

Provides access to all LONTAD digitized content and other collection metadata

United Nations archives Geneva

LONTAD: reach and impact

- **2019: legacy stats (Pre-LONTAD and pre-Covid)**
  - In-person visits: 1,283 (~250-300 clients/y)
  - Reference requests (email, etc.): 1,266/y
  - Online catalogue use: 294,433 page views/y

- **Since AtoM implementation (6 Dec 2021 – 31 May 2026)**
  - **Users: 722,219**
  - **Page views: 4,084,629**
  - **Greatly increased geographical diversity**

United Nations archives Geneva

LONTAD: reach and impact

- Overall, users from > **200** countries and territories have consulted the platform
- From December 2021 until today, the platform has recorded **344 users from Kazakhstan**

☐	76	Azerbaijan	377 (0,01 %)
☐	77	Armenia	361 (0,01 %)
☐	78	Uganda	354 (0,01 %)
☐	79	Luxembourg	351 (0,01 %)
☐	80	Lithuania	346 (0,01 %)
☐	81	Kazakhstan	344 (0,01 %)
☐	82	Nepal	315 (<0,01 %)
☐	83	Syria	309 (<0,01 %)
☐	84	Senegal	307 (<0,01 %)
☐	85	Croatia	292 (<0,01 %)
☐	86	Bolivia	280 (<0,01 %)

United Nations archives Geneva

New pathways for research through partnership

Digital Humanities

research, teaching, and engineering practices combining informatics/technology with the arts, humanities, and social sciences

Hackathon, University of Geneva, 2019

CENTRE FOR DIGITAL HUMANITIES & MULTILATERALISM

Graduate Institute Geneva

VisuaLeague, University of Copenhagen, 2022

United Nations archives Geneva

Impact and reach through partnerships: focus on technology

Partnership with the CDHM from 2024

- **Enhancing the usability of a selection of LON Archives**
- **Expending the corpus to UN Archives**
- **Create a corpus combining archives from different organizations**

CENTRE FOR DIGITAL HUMANITIES & MULTILATERALISM

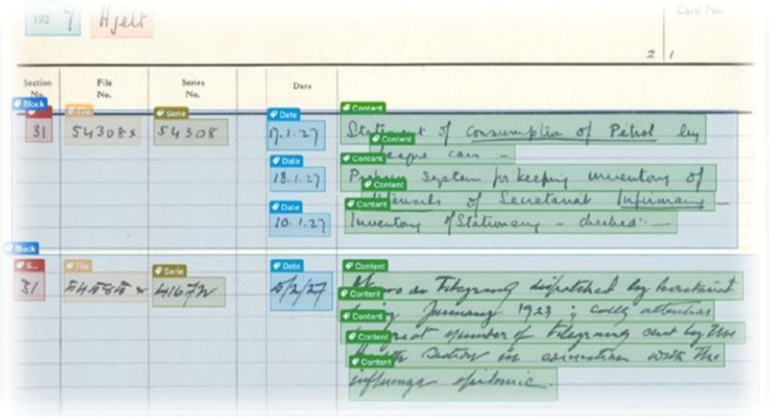
Graduate Institute Geneva

- Applying IIIF Protocol
- LON Registry Index Cards: Handwritten Text Recognition (HTR)
- Data Mining and Semantic Web: Minutes of Multilateralism

United Nations archives Geneva

New technology projects: handwritten text recognition (HTR)

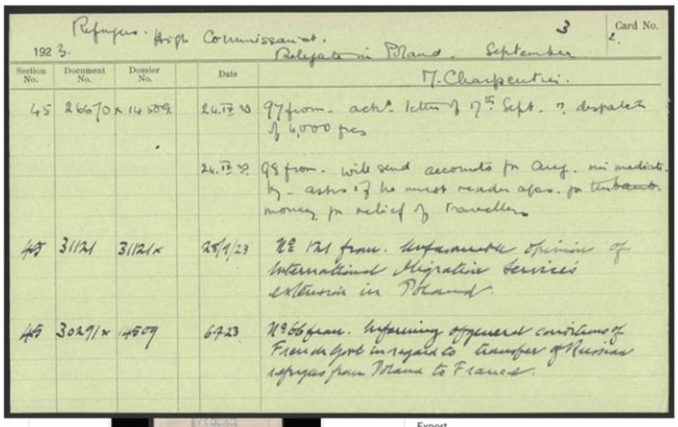
Registry Index Cards of the League of Nations → turning 250,000 multilingual handwritten index cards (FR/EN/DE) into structured digital data using HTR technology (AI model training)



United Nations archives Geneva

.....and datalinking

Matching reference codes with their corresponding file URLs on the platform



LON-RUSNAME-Refugees_High_Commissioner_in_Poland_00023	1923	Refugees. High Commissariat - Delegate in Poland. September - M. Charpentier	2

Export
Dublin Core 1.1 XML
EAD 2002 XML

Related subjects
Humanitarian Aid and Relief » Protection of and Assistance to Refugees and Displaced Persons » Refugees

Related genres
Paper File(s) / Document(s)

Physical storage
LON Box (375x270x150): R1749

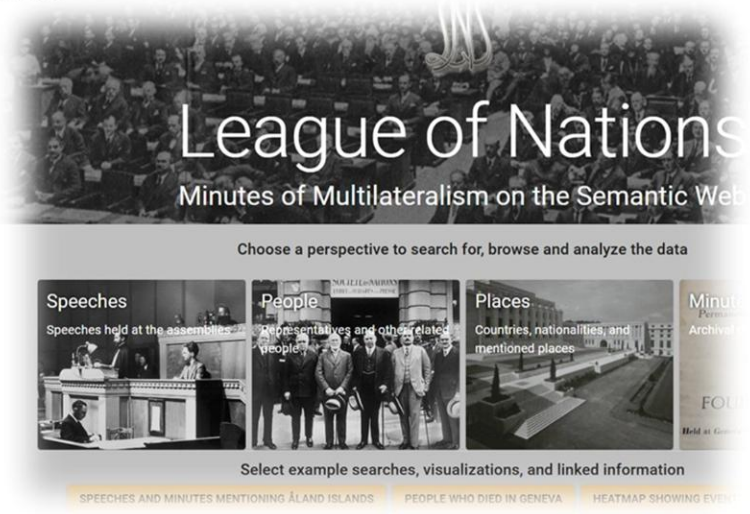
Identity area

Reference code	R1749/45/31121/31121
Title	Collaboration Between the High Commissariat for Refugees and the International Migration Service - Dossier concerning.
Date(s)	• 1923.09.21-1924.06.28 (Creation)
Level of description	File

United Nations archives Geneva

New technology projects: datamining of Minutes

A Linked Open Data Service and Semantic Portal to Study the Assembly Minutes and Prosopography of the League of Nations (1920-1946)



https://minutes.ldf.fi/en/

United Nations archives Geneva

Reflections: AI as an opportunity

- **New forms of research and engagement** → data mining and pattern analysis across vast collections (as shown in previous slide)
- **Operational efficiency**: support for internal workflows
  - automating repetitive tasks (metadata entry, redaction assistance)
  - supporting records management lifecycle processes
  - improving resource allocation through predictive analytics
- **Unlocking audiovisual and handwritten archives** → historically hard-to-use materials become accessible
  - Speech-to-text for audio/video archives
  - Handwriting recognition (HTR) for manuscripts
  - Computer vision for photographs (object, face, and scene recognition)

United Nations archives Geneva

Reflections: managing risk

AI also introduces challenges that archives must manage:

- **Bias and representation**: AI models may reflect historical or societal biases
- **Transparency and explainability:** decisions made by AI need to be auditable
- **Ethical use and privacy**: especially with sensitive or personal records
- **Skills gap**: archivists need new competencies in data and AI literacy

United Nations archives Geneva

CHALLENGES IN DIGITAL ARCHIVES COLLABORATION AND LEARNING CONTENT DEVELOPMENT BY THE NATIONAL ARCHIVES OF JAPAN

Furuya Ichiro
MPA, Columbia University (SIPA), BA in Law, University of Tokyo,
Senior Vice-President, National Archives of Japan
kokusai@archives

Abstract. This paper examines the National Archives of Japan's (NAJ) efforts to enhance digital archives and develop learning resources. Established in 1971, NAJ preserves and provides access to historically significant public records and archives, while increasingly focusing on digital access, interoperability, and outreach. The National Archives of Japan Digital Archive (NAJDA), launched in 2005, enables free online searching and viewing of digitized materials. By 2026, about 500,000 volumes (28% of holdings) and 40 million images are available. NAJDA also supports cross-searching with 29 institutions and promotes national standards for digital archives. The Japan Center for Asian Historical Records (JACAR) expands access by integrating materials from multiple institutions, offering about 34 million images. NAJ also contributes to *Japan Search*, a cross-sectoral platform linking over 300 databases and widening public access. Building on this foundation, NAJ launched its Learning Resources Program in 2025, providing videos, quizzes, and reference materials based on primary sources for secondary education. The program encourages learners in independent historical thinking. Looking ahead, NAJ aims to expand content, strengthen collaboration, and enhance multimedia learning, increasing the social and educational value of digital archives.

Keywords: Digital Archives; National Archives of Japan; NAJ Digital Archive (NAJDA); Japan Center for Asian Historical Records (JACAR); Japan Search; Digitization; Interoperability; Learning Resources Program; Primary Source Materials

**In this paper, "Digital Archive" refers to an online system that enables searching archival metadata and, where available, viewing digitized materials, rather than a digital repository.*

ЖАПОНИЯ ҰЛТТЫҚ АРХИВІНІҢ ЦИФРЛЫҚ АРХИВТЕР САЛАСЫНДАҒЫ ЫНТЫМАҚТАСТЫҒЫ ЖӘНЕ БІЛІМ БЕРУ КОНТЕНТІН ӘЗІРЛЕУДЕГІ МӘСЕЛЕЛЕРІ

Фуруя Ичиро
Колумбия университетінің (SIPA) мемлекеттік басқару магистрі (MPA),
Жапония Ұлттық архивінің аға вице-президенті
kokusai@archives

Аңдатпа. Мақалада Жапония Ұлттық архивінің (NAJ) цифрлық архивтерді дамыту және білім беру ресурстарын әзірлеу бағытындағы қызметі қарастырылады. 1971 жылы құрылған Жапония Ұлттық архиві тарихи маңызы бар мемлекеттік құжаттар мен архив материалдарын сақтауды және оларға қолжетімділікті қамтамасыз етеді, сонымен қатар цифрлық қолжетімділікті, өзара үйлесімділікті және қоғаммен өзара байланысты дамытуға ерекше көңіл бөледі. 2005 жылы іске қосылған Жапония Ұлттық архивінің цифрлық архиві (NAJDA) цифрланған материалдарды интернет арқылы тегін іздеуге және қарауға мүмкіндік береді. 2026 жылға қарай жүйеде шамамен 500 мың сақтау бірлігі (архив қорының 28%-ы) және 40 миллионнан астам цифрлық кескін қолжетімді болды. NAJDA сондай-ақ 29 мекеменің дерекқорлары бойынша бірыңғай іздеу мүмкіндігін қамтамасыз етіп, цифрлық архивтерге арналған ұлттық стандарттарды дамытуға ықпал етеді. Азияның тарихи құжаттары жөніндегі Жапония орталығы (JACAR) бірнеше мекеменің қорларын біріктіру арқылы қолжетімділікті кеңейтіп, шамамен 34 миллион кескінді ұсынады. Сонымен қатар, Жапония Ұлттық архиві 300-ден астам дерекқорды біріктіретін және мәдени мұраға қоғамдық қолжетімділікті кеңейтетін Japan Search ұлттық платформасының дамуына қатысады. Осы негізде 2025 жылы NAJ орта білім беру жүйесіне арналған алғашқы тарихи дереккөздерге негізделген бейнематериалдар, тесттер мен анықтамалық материалдарды қамтитын Learning Resources Program бағдарламасын іске қосты. Бағдарлама білім алушылардың тарихи ойлау қабілетін өздігінен дамытуға бағытталған. Алдағы уақытта NAJ білім беру контентін кеңейтуді, серіктестікті нығайтуды және мультимедиялық оқыту мүмкіндіктерін жетілдіруді жоспарлап отыр. Бұл цифрлық архивтердің қоғамдық және білім беру құндылығын одан әрі арттыруға ықпал етеді.

Түйін сөздер: цифрлық архивтер; Жапония Ұлттық архиві; Жапония Ұлттық архивінің цифрлық архиві (NAJDA); Азияның тарихи құжаттары жөніндегі Жапония орталығы (JACAR); Japan Search; цифрландыру; өзара үйлесімділік; Learning Resources Program; бастапқы тарихи дереккөздер

**Мақалада «цифрлық архив» ұғымы цифрлық репозиторийді емес, архивтік метадеректерді іздеуге және цифрланған материалдар қолжетімді болған жағдайда оларды қарауға мүмкіндік беретін онлайн жүйені білдіреді.*

ПРОБЛЕМЫ СОТРУДНИЧЕСТВА ЦИФРОВЫХ АРХИВОВ И РАЗРАБОТКИ ОБРАЗОВАТЕЛЬНОГО КОНТЕНТА НАЦИОНАЛЬНЫМ АРХИВОМ ЯПОНИИ

Фуруя Ичиро
Магистр государственного управления (MPA), Школа международных и общественных отношений Колумбийского университета (SIPA), старший вице-президент Национального архива Японии
kokusai@archives

Аннотация. В статье рассматривается деятельность Национального архива Японии, направленная на развитие цифровых архивов и создание образовательных ресурсов. Основанный в 1971 году, Национальный архив Японии осуществляет хранение и обеспечение доступа к исторически значимым государственным документам и архивным материалам, уделяя все большее внимание цифровому доступу, интероперабельности и популяризации архивного наследия.

Цифровой архив Национального архива Японии, запущенный в 2005 году, предоставляет пользователям возможность бесплатного поиска и просмотра оцифрованных архивных материалов в режиме онлайн. По состоянию на 2026 год, в системе доступны около 500 тыс. ед. хр. (28% архивных фондов) и свыше 40 млн цифровых изображений. Платформа NAJDA также поддерживает сквозной поиск по ресурсам 29 учреждений и способствует внедрению национальных стандартов в сфере цифровых архивов.

Японский центр азиатских исторических документов (Japan Center for Asian Historical Records, JACAR) расширяет доступ к архивным материалам за счет интеграции ресурсов нескольких учреждений, обеспечивая доступ примерно к 34 млн цифровых изображений. Кроме того, Национальный архив Японии участвует в развитии платформы *Japan Search* – межотраслевого цифрового ресурса, объединяющего более 300 баз данных и значительно расширяющего возможности общественного доступа к культурному и историческому наследию.

Опираясь на созданную цифровую инфраструктуру, в 2025 году Национальный архив Японии запустил *Программу образовательных ресурсов (Learning Resources Program)*, включающую видеоматериалы, викторины и справочные материалы, основанные на первоисточниках и предназначенные для использования в системе среднего образования. Программа направлена на развитие у обучающихся навыков самостоятельного исторического мышления.

В перспективе Национальный архив Японии планирует расширять содержание цифровых ресурсов, укреплять межведомственное сотрудничество и совершенствовать мультимедийные образовательные материалы, тем самым повышая общественную и образовательную ценность цифровых архивов.

Ключевые слова: цифровые архивы; Национальный архив Японии; Цифровой архив Национального архива Японии (NAJDA); Японский центр азиатских исторических документов (JACAR); Japan Search; цифровизация; интероперабельность; Программа образовательных ресурсов; материалы первоисточников

**В статье термин «цифровой архив» используется для обозначения онлайн-системы, обеспечивающей поиск архивных метаданных и, при наличии, просмотр оцифрованных архивных материалов, а не цифрового хранилища (репозитория).*

Introduction

The National Archives of Japan (NAJ) was established in 1971 with the mission of preserving and managing historically significant public records and making them widely available to the public. Its establishment enabled the centralize preservation and management of historical public records that had previously been retained under the responsibility of the respective agencies that created them. Since then, NAJ has served as a central institution for the preservation and utilization of historical public records and archives by properly preserving them and making them available for public access and exhibitions.

With the advancement of digital technology, archival institutions are now expected to take on new roles beyond simply preserving materials, such as improving access through digital environments, strengthening cooperation with other institutions, and utilizing their holdings for educational and learning activities. NAJ has been engaged in Digital Archives initiatives since the early 2000s, working to make materials available online and enhance search functionality. In recent years, NAJ has launched initiatives to produce and publish Learning Resources Program based on materials made available through Digital Archives, in an effort to make use of archival materials as educational resources.

This paper summarizes the Digital Archives initiatives promoted by NAJ and the expanding interoperability of these services, reports on the operation and challenges of its new Learning Resources Program, and discusses the potential for cooperation and educational applications of Digital Archives.

1. Expanding Interoperability of the National Archives of Japan Digital Archive

1.1 Overview of the National Archives of Japan

As of 2026, NAJ holds about 1.78 million volumes of materials, consisting primarily of public records transferred from government agencies. Of these, about 1.3 million

volumes are administrative records from the modern era. The collection also includes about 480,000 volumes of historical materials (the Cabinet Library) formerly held by shogunal families, temples and shrines, court nobles, and samurai families prior to the Edo period (1603–1868). The oldest material dates back to 908, and many of these historical materials are of great cultural value. In addition, materials such as 16th-century manuscripts of history texts originating in the 8th century have been preserved through changes in custodianship and natural disasters.

NAJ is an agency under the jurisdiction of the Cabinet Office. With a staff of about 200, it is not large in scale compared to national archives in other countries, but it is tackling various new challenges within its limited human and material resources. The main office is located in Kitanomaru Park, right in front of the Imperial Palace in the heart of Tokyo, the nation's capital. In addition to the main office, the Tsukuba Annex in Tsukuba Science City, about 70 km from Tokyo, was constructed in 1998. NAJ receives approximately 50,000 volumes of materials each year, and the shelves at both facilities are nearing capacity. Furthermore, since the existing facilities are unable to sufficiently fulfill the functions expected of a modern archives institution, NAJ is working to expand its institutional functions while preparing for a new facility.

1.2 The National Archives of Japan's Digital Archive Initiative

1.2.1 About the National Archives of Japan Digital Archive (NAJDA)[1]

The National Archives of Japan Digital Archive (NAJDA) is an online access system that began operation in 2005. The system allows users to search catalog information of holdings, and view digitized images of archival materials, with the aim of creating an environment in which public records can be accessed anytime, anywhere, by anyone, freely, and at no cost. As of the end of March 2026, about 500,000 volumes, corresponding to about 28% of total holdings, have been digitized, and about 40 million images are available via NAJDA. While the challenge of digitizing the vast volume of paper-based materials remains, NAJDA undergoes regular system updates (replacements). In the most recent update, accessibility and usability were enhanced through a redesigned and improved operability.

In addition, NAJDA features a cross-search function for catalog information provided by archives and other institutions throughout Japan, and as of March 2026, interoperability with 29 institutions has been established. Furthermore, through the development and dissemination of the Standard Specifications for Digital Archive Systems, NAJ supports national universities and local government archives in developing digital archive.

1.2.2 Japan Center for Asian Historical Records (JACAR)[2] *and Its Activities*

Established in 2001 and operated by NAJ, the Japan Center for Asian Historical Records (JACAR) aims to promote mutual understanding between Japan and its neighboring countries in Asia. JACAR provides public records and other records related to Japan's modern relations with neighboring Asian countries and other regions, referred to as "Asian historical records," using digital images via the Internet.

JACAR is based on materials held by three institutions: NAJ, the Diplomatic Archives of the Ministry of Foreign Affairs, and the Center for Military History at the National Institute for Defense Studies of the Ministry of Defense. It plays a vital role as a platform that provides integrated access to data from these three institutions. Today, the collection has been expanded to include materials dating up to the post-World

War II era, with about 34 million images now available to the public. In addition to multilingual support and enhanced subject-based search functions, the service seeks to improve accessibility and useability through links with archives and university libraries both in Japan and abroad.

1.3 Interoperability with Japan Search[3], Japan's Cross-Sectoral Platform

NAJDA and JACAR have helped advance historical research both in Japan and abroad by improving access to historical materials, while also providing opportunities for the general public to easily engage with these materials. To further expand access, NAJ participates in Japan Search, a national platform launched about five years ago.

Launched in 2020, Japan Search is a national portal that enables users cross-sectoral searching and use of metadata for digital content across a wide range of fields, including books, public records, cultural properties, art, and academic resources. The National Diet Library operates the system, which was developed with the cooperation of various institutions, including NAJ.

As of April 2026, the number of linked database has reached 311, and the number of participating organizations has reached 200. The Digital Archives Strategy 2026–2030, published by the government's Intellectual Property Strategy Headquarters in 2025, positions digital archives not merely as tools for preserving, passing on, and disseminating culture, but also as contributors to the creation of new value and the promotion of innovation through their use in tourism, regional revitalization, education and research, and business fields. The strategy also identifies the next five years seen as a period for strengthening cross-sectoral and regional interoperability centered Japan Search, and for building a foundation to promote digital archives nationwide. Within this framework, NAJ is positioned as the "core organization for promoting digital archives" in the field of public records and archives.

NAJDA data is also now searchable through Japan Search, thus expanding opportunities for users previously unfamiliar with NAJ to discover public records and archives through cross-sectoral searches.

2. National Archives of Japan's Initiatives on the Learning Resources Program

2.1 Activities at the Japan Center for Asian Historical Records (JACAR)

Since 2020, JACAR has been publishing "Asian Historical Records Learning" available online as learning resources for learners, utilizing modern Asian historical records provided through its digital archives.

This learning resources are organized to provide explanations of major historical events by era, and to allow users to view related digitized primary source materials, thereby contributing to a deeper understanding of history.

2.2 Operation and Challenges of the Learning Resources Program[4] at the National Archives of Japan

2.2.1 Launch of the Learning Resources Program in 2025

As part of a plan to strengthen learning support functions in anticipation of the opening of a new facility scheduled for 2030, NAJ has been producing and publishing "Learning Resources Program" on its website since 2025, based on materials made available through NAJDA.

The Learning Resources Program developed by NAJ aims to serve as an entry point linking school education with the use of archives, providing learning resources

that can be accessed without restrictions on location or time. In the future, NAJ envisions its use across a variety of educational levels and for diverse learning objectives.

Specifically, the Program combine multiple formats such as videos, quizzes, and reference materials, allowing learners to study historical events covered in textbooks through direct engagement with primary source materials. They are intended not merely as a supplement to exhibition commentary, but also as teaching materials that support users in interpreting historical records and thinking about them independently. Primarily using materials from the Meiji period (1868–1912) onward, NAJ is working to develop teaching materials for the history education at the secondary school level (junior high and high school students).

The themes of the Learning Resources Program are based on the 24 topics that make up the NAJ main office's permanent exhibition, "Japan's Modern History." In 2025, videos, quizzes, and reference materials were produced on the themes of "Universal Suffrage for Men" and "Women's Suffrage," while reference materials were also released on "Japan's First Railway: The Opening of the Shimbashi–Yokohama Line," "The Beginnings of the Modern Education System: From the School System to the Primary School Ordinance," and "Postwar Educational Reform: The Basic Act on Education." The video and quiz content uses dialogues between original characters and incorporates feedback from serving teachers to enhance its practicality at educational settings. On the reference materials pages, enhanced explanations are provided for several materials related to the topic, and transcriptions of the materials are also provided.

2.2.2 Enhancing the Learning Resources Program and Prospects from 2026 Onward

For online content scheduled for release in 2026, NAJ is working to further enrich the content by utilizing materials held by institutions other than NAJ and developing teaching materials that combines materials from multiple institutions. Additionally, efforts are underway to produce video and game content to deepen understanding of NAJ's operations and to clearly convey the importance of preserving and utilizing public records and archives.

NAJ plans to develop the Learning Resources Program covering all 24 topics in the permanent exhibition "Japan's Modern History" by 2030. The challenge will be to expand opportunities for users to engage with history in a multifaceted and layered manner through public records and archives by further enriching the resources and strengthening cooperation with other institutions.

Conclusion

This paper outlined the expanding interoperability of NAJ's Digital Archive initiatives and the operation of the Learning Resources Program on this foundation. By not only making holdings available through Digital Archive but also fostering connections across institutions and fields and utilizing them in educational settings, archives can become more open and accessible to society.

Going forward, NAJ remains committed to working with relevant organizations both in Japan and abroad to maximize the value of digital archives and lay the groundwork for supporting history education for future generations.

Notes

https://www.digital.archives.go.jp/?lang=en (accessed 25.05.2026).
https://www.jacar.go.jp/english/index.html (accessed 25.05.2026).
https://jpsearch.go.jp/en (accessed 25.05.2026).
https://www.archives.go.jp/learning/ (date of access: 25.05.2026) (in Japanese).

FAMILY HISTORY IN A DIGITAL AGE: ENDURING PURPOSE AND GLOBAL IMPACT

Mark Allen Bragg
Chair of the Board of Directors, FamilySearch International, USA

ОТБАСЫ ТАРИХЫ ЦИФРЛЫҚ ДӘУІРДЕ: ӨЗЕКТІЛІГІН ЖОҒАЛТПАЙТЫН ҚҰНДЫЛЫҒЫ ЖӘНЕ ЖАҺАНДЫҚ ӘСЕРІ

Марк Аллен Брэгг
FamilySearch International директорлар кеңесінің төрағасы, АҚШ

ИСТОРИЯ СЕМЬИ В ЦИФРОВУЮ ЭПОХУ: НЕПРЕХОДЯЩАЯ ЗНАЧИМОСТЬ И ГЛОБАЛЬНОЕ ВЛИЯНИЕ

Марк Аллен Брэгг
председатель Совета директоров FamilySearch International, США

Family History in a Digital Age:
Enduring Purpose and Global Impact
Mark A. Bragg
Chairman, FamilySearch International
FamilySearch

FamilySearch

Family history is enduring, even when the methods are changing

Access is an act of kindness

Collaboration has a multiplying effect

FamilySearch
Records are witnesses and memories of individuals, families, people, cultures and countries

ИННОВАЦИОННЫЕ ПРАКТИКИ ПРЕДСТАВЛЕНИЯ АРХИВНОЙ ИНФОРМАЦИИ: ВЗГЛЯД ИЗ КИТАЯ

Цюй Кай
Кандидат исторических наук, заместитель начальника
Департамента государственных и ведомственных архивов
Государственного архивного управления Китайской Народной Республики
qukai@naa.gov.cn

Аннотация. Архивные учреждения Китая последовательно внедряют инновационные подходы к представлению архивной информации и осуществляют интеллектуальную трансформацию архивной деятельности посредством интеграции цифровых, аналитических и технологий искусственного интеллекта. В статье, подготовленной на основе доклада, представленного на III Международном конгрессе архивистов, систематизированы инновационные практики комплексных архивов Китая, включая применение больших языковых моделей, мобильных интернет-сервисов, технологий виртуальной и дополненной реальности (VR/AR), интерактивных выставок и геймификации. Также рассматриваются институциональные механизмы развития отрасли, включая стратегическое планирование, нормативно-правовую базу, отраслевые стандарты и реализацию крупных государственных проектов, при этом отмечаются существующие региональные различия. Опыт Китая свидетельствует о том, что сочетание технологических инноваций и институциональной координации способствует повышению эффективности представления архивной информации, расширению общественного участия и развитию исторического просвещения.

Ключевые слова: архивная информация; цифровые архивы; инновационные практики; искусственный интеллект; большие языковые модели; виртуальная и дополненная реальность (VR/AR); геймификация; межархивное взаимодействие; электронные архивы; Китай

АРХИВТІК АҚПАРАТТЫ ҰСЫНУДЫҢ ИННОВАЦИЯЛЫҚ ТӘЖІРИБЕЛЕРІ: ҚЫТАЙ ТӘЖІРИБЕСІ

Цюй Кай
Тарих ғылымдарының кандидаты, Қытай Халық Республикасы Мемлекеттік архив басқармасының мемлекеттік және ведомстволық архивтер департаменті бастығының орынбасары
qukai@naa.gov.cn

Аңдатпа. Қытайдың архив мекемелері цифрлық, деректерге негізделген және жасанды интеллект технологияларын біріктіру арқылы архивтік ақпаратты ұсыну саласында қызмет көрсетуді жаңғырту мен интеллектуалдық трансформацияны жүйелі түрде жүзеге асырып келеді. Үшінші халықаралық архившілер конгресінде ұсынылған баяндама негізінде дайындалған бұл мақалада Қытайдың кешенді архивтерінде қолданылатын инновациялық тәжірибелер жүйелі түрде қарастырылады. Атап айтқанда, үлкен тілдік модельдерді пайдалану, мобильді интернет-қызметтер, виртуалды және толықтырылған шынайылық (VR/AR) технологиялары, интерактивті көрмелер және геймификация мәселелері қамтылған. Сонымен қатар стратегиялық жоспарлау, нормативтік-құқықтық реттеу, салалық стандарттар және ірі мемлекеттік жобалар сияқты институционалдық тетіктер сипатталып, өңірлер арасындағы даму айырмашылықтары да атап өтіледі. Қытай тәжірибесі технологиялық инновациялар мен институционалдық үйлестірудің үйлесуі архивтік ақпаратты ұсынудың тиімділігін арттыруға, қоғамның қатысуын кеңейтуге және тарихи ағартуды дамытуға ықпал ететінін көрсетеді.

Түйін сөздер: архивтік ақпарат; цифрлық архивтер; инновациялық тәжірибелер; жасанды интеллект; үлкен тілдік модельдер (LLM); виртуалды және толықтырылған шынайылық (VR/AR); геймификация; архиваралық ынтымақтастық; электрондық архивтер; Қытай

INNOVATIVE PRACTICES IN ARCHIVAL INFORMATION PRESENTATION: FROM A CHINESE PERSPECTIVE

Qu Kai
Candidate of Historical Sciences, Deputy Director of the Department of State and Institutional Archives, National Archives Administration of the People's Republic of China
qukai@naa.gov.cn

Abstract. Chinese archival authorities have been advancing service innovation and intelligent transformation in archival information presentation by integrating digital, data-driven, and AI technologies. Based on a presentation at the Third International Conference of Archival Workers, this paper systematically introduces innovative practices of comprehensive archives in China, including large-model applications, mobile internet services, VR/AR technologies, interactive exhibitions, and gamification. It also outlines institutional frameworks such as strategic planning, regulations, standards, and major projects, while acknowledging regional disparities. The Chinese experience demonstrates that technological innovation combined with institutional coordination effectively enhances presentation efficacy, public engagement, and historical education.

Keywords: archival information; digital archives; innovative practices; artificial intelligence; large language models (LLMs); virtual and augmented reality (VR/AR); gamification; inter-archival collaboration; electronic archives; China

Уважаемый председатель, эксперты и коллеги-архивисты, добрый день!

Для меня большая честь участвовать в III Международном конгрессе архивистов Казахстана и обмениваться опытом в области инновационных практик представления архивной информации.

Представление архивных документов и содержащейся в них информации является одним из ключевых направлений архивной деятельности. С начала XX века инновация в представлении архивной информации стали важной частью модернизации архивного дела в Китае. На протяжении более чем 20 лет мы последовательно развивали это направление и сегодня можем говорить о его реальных практических достижениях.

В китайских архивах динамично развиваются инновационные форматы представления архивной информации, сочетающие применение цифровых технологий, анализа данных и интеллектуальных методов с традиционными подходами, что выводит китайское архивное дело на путь модернизации. Архивы всех уровней в Китае оперативно размещают информацию через газеты и журналы, веб-сайты и социальные сети и другие средства массовой информации, проявляют инициативу по открытию своих архивных фондов, активизируют подготовку документальных публикаций и расширяют каналы доступа пользователей к открытой архивной информации. Создана общенациональная платформа поиска и использования архивных документов, к которой подключились все государственные архивы. Налаженное взаимодействие между архивными учреждениями обеспечивает возможность межархивного поиска, что упрощает доступ к документам из разных архивов и повышает общую эффективность архивного сервиса. Во многих городах возможность поиска и доступа к социально значимым архивным документам включена в перечень государственных услуг на региональных порталах «Одна сеть – все услуги». В результате граждане все острее ощущают удобство и комфорт от взаимодействия с архивными службами.

Инновации в представлении архивной информации осуществляются в форме иммерсивного и интерактивного погружения в виртуальную реальность с достоверным воспроизведением, основанным на архивных документах. Современные технологические средства, новые медиатехнологии и интеллектуальные методы активно используются для комплексного улучшения качества, выразительности и привлекательности выставочных проектов и позволяют сформировать более глубокую эмоциональную связь с аудиторией и достичь положительных результатов в историческом просвещении, передаче информации, эмоциональном вовлечении и побуждении к действию. Далее я представлю практические примеры и поделюсь опытом в этой области.

Практические примеры

В настоящий момент в Китае функционирует 3302 комплексных архива, играющих ключевую роль в развитии национального архивного дела. Ниже будут рассмотрены инновационные практики в сфере представления архивной информации, применяемые в этих учреждениях.

Применение больших языковых моделей

Пекинский городской архив и Пекинское городское архивное управление, опираясь на инфраструктуру компании Beijing Telecom, разработали большую языковую модель «Lantai Jinghua» для интеллектуальной архивной системы столицы. Данная система была официально введена в эксплуатацию в 2025 году и прошла модернизацию в 2026 году. Модель базируется на мультимодальном взаимодействии и технологии Agentic RAG (Agent-Oriented Retrieval Augmented Generation), глубоко интегрирована с государственными сервисными сценариями и формирует комплексную структуру сервисных возможностей по шести направлениям: научно обоснованное принятие решений и система вопросов-ответов по административным знаниям, генерация документации, разработка планов, составление хроник, формирование табличных данных, создание различных официальных документов. Это позволяет совершенствовать представление архивной информации, актуализировать ценность данных, расширять возможности научно обоснованного принятия решений, а также реализовать переход консультационно-аналитической поддержки от информатизации к интеллектуализации.

Удобные сервисы через мобильные интернет-мини-программы

Пекинский городской архив, опираясь на мини-программу «Jingtong», создал мобильную платформу архивных услуг, предоставляющую информационно-поисковый сервис через мобильные устройства. Система полного цикла обслуживания, включающая «очный запрос, межархивный запрос, онлайн-запрос и мобильный запрос», эффективно расширяет подходы к архивной информации и заметно повышает возможности и эффективность обслуживания граждан.

Виртуальная реальность и дополненная реальность (VR/AR)

Облачный выставочный зал с VR-технологией. Первый исторический архив Китая – центральный архив в прямом подчинении Государственного архивного управления КНР, хранящий документы династий Мин и Цин общим фондом более 10 млн единиц хранения. В этом архиве оборудовано крупномасштабное иммерсивное выставочное пространство с применением VR-технологией. Здесь посетители могут воочию увидеть слияние цифровых технологий с архивной гуманитаристикой, «путешествовать во времени и пространстве» с помощью технологии виртуальной реальности и знакомиться с документами, полностью погружаясь в историческую атмосферу.

Интерактивный документальный VR-фильм «Путешествие по древним литературам»

Первый исторический архив Китая провел совместно с партнерскими организациями выставку «Четыре великих открытия древних китайских документов в начале XX века». На основе архивных документов, музейных экспозиций и археологических находок был создан интерактивный документальный VR-фильм с элементами сторителлинга. Это дает пользователям возможность в режиме онлайн ознакомиться с уникальными документами и древними книгами: гадательными надписями на костях династии Шан, бамбуковыми дощечками династии Хань, дуньхуанскими рукописями, а также императорскими архивами династий Мин и Цин. В городском архиве Сюйчжоу

провинции Цзянсу создана метавселенная – виртуальный выставочный зал, где можно осматривать экспозицию со смартфона или компьютера и общаться друг с другом в реальном времени.

Дополненная реальность (AR)

AR-помощник «Прогулка по городу Шанхай», разработанный Шанхайским городским архивом, накладывает историческую информацию на реальные городские пейзажи. В основе программы лежат 12 карт г. Шанхай разных исторических периодов. Она включает в себя более 2300 оцифрованных документов, фотографий и видеозаписей, визуализирует прошлое и настоящее 133 достопримечательностей, демонстрируя славный путь развития Шанхая от небольшого городка в устье реки Янцзы до современного международного мегаполиса.

Интерактивные выставки

Интерактивная выставка акцентирует внимание на вовлеченности посетителей и исследовательском интересе, преобразуя традиционную модель «вы смотрите» в интерактивную модель «вы играете» с использованием современных технологий.

Нарратив на основе ГИС-карты

Платформа «Взгляд на Шанхай по документам», разработанная Шанхайским городским архивом на базе вековой картографической основы, позволяет изучать изменения облика города с помощью пространственно-временной карты. Общее количество обращений к платформе – почти полтора миллиона. Городской архив Хуаншань провинции Аньхой провел углубленное изучение собственных фондов в ответ на образовательные запросы, внедрил роботов-

рассказчиков на основе ИИ и представил исследовательские курсы «Путешествие по Хойчжоу по документам». С применением технологии голографической проекции, транслируемой посредством сети 5G, городской архив Куньшань представил цифровые 3D-модели 14 местных знаковых исторических личностей и сцен куньшаньской оперы. Интерактивное оборудование позволяет пользователю взаимодействовать с цифровыми реконструкциями образов древних мудрецов, словно путешествуя во времени.

Геймификация

Раскрытие секретов архивов в игровом формате стимулирует интерес молодежи и способствует ее активному вовлечению в изучение истории. Игровые квесты и иммерсивные перформансы: Архив провинции Фуцзянь представил интерактивную игру «Дом Цяопи: переулок Или», посвященную традиции писем-переводов китайских эмигрантов. Участникам предлагается исследовать выставочное пространство с помощью специального путеводителя и разгадывать тематические загадки. 73% участников проекта – это люди моложе 30 лет. Аналогичная программа в музее Цяопи предполагает решение задач посредством анализа аутентичных артефактов, таких как «письмо, пропитанное водой», что позволяет учащимся в течение трехчасового погружения прочувствовать глубокую связь семейных и патриотических ценностей.

Ролевые игры и настольные квесты: в рамках интеллектуальной собственности «Седьмой архив» (г. Сучжоу, провинция Цзянсу) разработаны детективная ролевая игра «Такого человека нет» и книга-головоломка «Седьмой архив». За три дня очных мероприятий проект привлек 2000 участников, а число просмотров коротких видео достигло более 1,058 млн. На 46-й сессии Комитета всемирного наследия ЮНЕСКО проект был удостоен награды «Звезда передового опыта» за инновационные подходы к образованию на основе культурного наследия.

Практический опыт и методы работы

Разработка инновационной стратегии: определение направления развития

В 2002 году был принят «План мероприятий по реализации информатизации архивного дела в Китае», в котором обоснована необходимость развития соответствующей инфраструктуры и ресурсной базы. В рамках интеграции межрегиональных архивных систем предусмотрено последовательное расширение возможностей совместного использования архивных ресурсов, а также обеспечение обществу доступного, оперативного и качественного сервиса.

В 2010 году было опубликовано «Руководство по созданию цифровых архивов», в котором поставлена задача по применению различных технологических средств для обеспечения многоуровневого совместного доступа к архивным информационным ресурсам. Это направлено на оперативное и качественное обслуживание различных категорий пользователей. В рамках системы предусмотрена реализация следующих функций: поиск документов, публикация ресурсов, информационный обмен, изучение и использование архивных материалов, а также взаимодействие в рабочих процессах.

«План развития архивного дела в КНР на 12-ю пятилетку» (2011 год) задал курс на создание комплексной платформы, нацеленной на обеспечение всестороннего архивного информационного обслуживания. В «Основных положениях плана развития архивного дела в КНР на 13-ю пятилетку» (2016 год) акцент был сделан на повышение качества и доступности государственных услуг, а также на внедрение инновационных форм обслуживания. «План на 14-ю пятилетку» (2020 год) вывел на новый уровень обработку и использование данных, предусматривая внедрение современных технологий управления знаниями, технологий искусственного интеллекта и цифровых гуманитарных наук. Утвержденный в текущем году «План на 15-ю пятилетку» определяет стратегическое направление в виде инициативы «Искусственный интеллект + архивы». Запланировано внедрение интеллектуальных инструментов, призванных повысить качество и эффективность доступа к документам, а также расширить возможности совместного использования ресурсов.

В апреле 2026 года Государственное архивное управление КНР опубликовало документ «Мнения о продвижении внедрения искусственного интеллекта в архивное дело». Документ определяет стратегию планомерного и упорядоченного внедрения технологий ИИ в сферу архивного дела и содействует интеграции ИИ с деятельностями архивов.

Совершенствование нормативно-правовой базы в сфере архивного дела: обеспечение правовой защиты

На законодательном уровне. В обновленной редакции «Закона КНР об архивном деле» предусмотрена отдельная статья, посвященная информатизации. В ней закреплены положения о создании и развитии платформ совместного использования информационных ресурсов, а также о содействии межрегиональному и межведомственному обмену данными и совместному использованию цифровых архивных ресурсов.

На уровне административных актов. В «Положениях о реализации Закона об архивном деле» конкретизирована необходимость обеспечения доступа общественности к архивным документам различными способами, повышения уровня совместного использования архивной информации, а также содействия межрегиональному, межуровневому и межведомственному обмену цифровыми архивными ресурсами и их использованию.

На уровне ведомственного регулирования. В утвержденном Государственным архивным управлением КНР Приказом № 19 документе «Правила раскрытия архивных документов в государственных архивах» закреплено требование об обеспечении доступа общественности к соответствующей архивной информации с использованием мультимедийных технологий и цифровых платформ.

На уровне отраслевых стандартов. Приняты следующие нормативные документы: «Технические нормы применения технологий COM и COLD в цифровых архивах», «Правила организации документальных выставок», «Правила экспертизы и использования архивных документов», «Правила предоставления архивных документов для использования». Данные документы устанавливают рекомендательные стандарты и формируют методическую основу для внедрения инновационных подходов к представлению архивной информации.

Реализация значимых государственных проектов: укрепление проектной поддержки

Национальный проект по обеспечению сохранности и развитию важных архивных документов. При финансовой поддержке проведены масштабные систематические работы с историческими документами. За последние 5 лет составлены описи 20 млн единиц хранения и выполнена оцифровка 80 млн страниц. Результаты проекта предоставлены общественности для использования через различные каналы.

Государственный проект «Документальное наследие Китая». Проведены масштабные мероприятия по распространению программы «Память мира в Китае». Организовано 6 отборов для включения в «Реестр документального наследия Китая», в который вошли 250 уникальных документальных объектов. В том числе 18 из них были включены в Международный реестр ЮНЕСКО «Память мира» и 17 – в Азиатско-Тихоокеанский реестр «Память мира».

Проект по укреплению фундамента архивной информатизации. В рамках проекта создается платформа для совместного использования архивных информационных ресурсов. Особое внимание уделяется комплексному содействию обмену архивными ресурсами на государственном и региональном уровнях. Разработана национальная система совместного доступа к архивной информации, которая базируется на национальном сайте поиска и использования архивных документов и позволяет предоставить сервисы через единую платформу.

Создание электронных архивов. Это стало ключевым этапом в реализации стратегии информатизации архивного дела. В период с 2020 по 2025 годы было создано 120 комплексных и 70 ведомственных электронных архивов, отвечающих современным требованиям надежности и доступности.

Проект «Искусственный интеллект + Архивы». В рамках данного проекта планируется стимулировать архивы, обладающие необходимыми условиями, к внедрению интеллектуальных решений в различные сценария архивной деятельности. Усиливается работа по распознаванию текста, аудиозаписей и изображений в ценных архивных материалах, совершенствуется управление архивными данными, создаются высококачественные наборы данных и разрабатываются специализированные большие языковые модели для архивной сферы. Особое внимание будет уделено обеспечению безопасности архивных данных, а также выявлению и контролю рисков, связанных со сбором и взаимосвязью данных.

Инновационное развитие – ключевой тренд в сфере представления архивной информации. Для повышения эффективности этого процесса архивным учреждениям важно систематически анализировать накопленный опыт. Следует выделить два аспекта. Во-первых, немаловажно сочетать традиционные форматы подачи с общественно значимыми событиями. Интеграция архивных материалов в контекст праздников и дней памяти привлекает внимание общества и повышает интерес к историческому наследию. Например, в 2019 году к 70-летию образования КНР Государственное архивное управление Китая опубликовало на официальном сайте цветную кинохронику, запечатлевшую церемонию образования КНР в 1949 году. Материал был ретранслирован рядом крупных

социальных сетей. В итоге общее количество просмотров в первые сутки достигло 320 млн, что демонстрирует высокую эффективность подобного подхода. Во-вторых, из-за обширной территории Китая уровень развития архивного дела в регионах существенно различается. В восточных регионах и городах субпровинциального значения инновационная деятельность в сфере представления архивной информации началась раньше и сейчас находится на передовом уровне. В то же время низовые архивы центральных и западных регионов сталкиваются с ограничениями: нехваткой финансирования и кадров из-за более низкого уровня экономического развития. В связи с этим особенно актуальна задача ускорения развития архивного дела в этих регионах.

Спасибо за внимание!

THE ECONOMICS OF DIGITAL ARCHIVING IN THE AGE OF AI

John Sheridan
Chief Digital and Information Officer at the National Archives of the United Kingdom, England
john.sheridan@nationalarchives.gov.uk

Abstract. This article considers the economics of digital archiving in the age of AI. It is a time of rapid technological change. The rise of Large Language Models is giving us highly capable, sophisticated, and adaptable AIs. Whilst archives' digital holdings are increasingly consumed by AI bots – a trend that may grow exponentially if unchecked – we should remain human centred institutions, concerned with records as evidence and collective public memory.

Yet, digital archives need AI. The adoption and use of AI by archives is both essential and urgent. The variety and scale of digital records, and the rapid growth of data volumes to process, are far beyond the capacity of our current methods for appraisal, selection, sensitivity review, preservation, archival description, and access management. AI is the only feasible hope for achieving the productivity growth we need in digital archiving.

This article reflects on lessons from previous revolutions in technology. It discusses Technology Diffusion Theory, Solow's Paradox, and Baumol's Cost Disease. Each of these tells a story about the arrival of new technology and the long road to unlocking the benefits. Economically, archives are a small sector with little market power. Ultimately, financial viability will strongly shape which AI tools become available to archives, while law, ethics, trust and institutional accountability will determine how AI can responsibly be deployed.

This article then identifies three practical priorities to advance AI in the archives sector: developing evaluation methods and evaluation data for AI use in archival work; decomposing archival tasks to maximise automation whilst ensuring human and machine responsibilities are assigned with care and understanding, and; improving the data that AI uses, in part by using AI to help us create and improve archival reference data. It concludes by arguing that all

progress depends on archives undertaking sustained experimentation – the test and learn approach – and sharing what we are learning as we progress.

Keywords: digital archives; economics; technology; AI; diffusion; Solow paradox; productivity; automation; test and learn

ЖАСАНДЫ ИНТЕЛЛЕКТ ДӘУІРІНДЕГІ ЦИФРЛЫҚ АРХИВТЕУ ЭКОНОМИКАСЫ

Джон Шеридан
Ұлыбритания Ұлттық архивінің цифрлық технологиялар және ақпарат жөніндегі директоры
john.sheridan@nationalarchives.gov.uk

Аңдатпа. Бұл мақалада жасанды интеллект (ЖИ) дәуіріндегі цифрлық архивтеудің экономикалық аспектілері қарастырылады. Қазіргі уақыттағы қарқынды технологиялық өзгерістермен сипатталады. Үлкен тілдік модельдердің (Large Language Models, LLM) дамуы жоғары өнімді, күрделі әрі бейімделгіш жасанды интеллект жүйелерінің қалыптасуына мүмкіндік берді. Цифрлық архив қорларын жасанды интеллект боттарының пайдалануы барған сайын артып келеді және бұл үрдіс тиісті бақылау болмаса, экспоненциалды сипат алуы мүмкін. Соған қарамастан архивтер құжаттарды ең алдымен дәлелдемелік дереккөз және қоғамның ұжымдық жады ретінде қарастыратын адамға бағдарланған институттар болып қала беруі тиіс.

Сонымен қатар цифрлық архивтердің жасанды интеллектті қолдануы аса қажет әрі шұғыл міндет болып табылады. Цифрлық құжаттардың алуан түрлілігі мен көлемі, сондай-ақ өңделуге тиіс деректер ағынының үздіксіз өсуі құжаттардың құндылығын сараптау, іріктеу, құпия ақпараттың болуын тексеру, ұзақ мерзімді сақталуын қамтамасыз ету, архивтік сипаттау және қолжетімділікті басқарудың қолданыстағы әдістерінің мүмкіндіктерінен әлдеқайда асып түседі. Сондықтан цифрлық архивтеу саласында қажетті еңбек өнімділігінің өсуіне қол жеткізудің іс жүзіндегі жалғыз тиімді жолы – жасанды интеллектті пайдалану.

Зерттеуде сондай-ақ өткен технологиялық революциялардан алынған сабақтар талданады. Технологиялардың таралу теориясы (Technology Diffusion Theory), Солоу парадоксы (Solow's Paradox) және Баумоль шығындар дерті (Baumol's Cost Disease) қарастырылып, жаңа технологиялардың пайда болуы мен олардың нақты пайдасын толық іске асыру арасында ұзақ уақыт өтетіні көрсетіледі. Экономикалық тұрғыдан архив саласы нарықтағы ықпалы шектеулі шағын сектор болып табылады. Сондықтан қаржылық орнықтылық архивтерде қандай жасанды интеллект құралдарының қолжетімді болатынын айқындайды, ал құқықтық нормалар, этикалық қағидаттар, қоғамдық сенім және институционалдық жауапкершілік оларды қауіпсіз әрі жауапты пайдаланудың шекарасын белгілейді.

Мақалада архив саласында жасанды интеллектті дамытуға бағытталған үш практикалық басымдық ұсынылады: архивтік қызметте ЖИ қолданудың тиімділігін бағалау әдістері мен бағалау деректерін әзірлеу; автоматтандыру деңгейін барынша арттыра отырып, адам мен машинаның міндеттерін ұқыпты әрі негізді бөлу мақсатында архивтік процестерді құрамдас бөліктерге жіктеу; сондай-ақ жасанды интеллект пайдаланатын деректердің сапасын арттыру, оның ішінде архивтік анықтамалық деректерді жасау және жетілдіру үшін ЖИ мүмкіндіктерін пайдалану.

Қорытындысында автор архивтердің тұрақты түрде тәжірибе жүргізуі, «сынақтан өткізу және үйрену» (test and learn) қағидатын ұстануы және жинақталған тәжірибе мен

алынған нәтижелерді өзара бөлісуі ғана осы саладағы тұрақты прогрестің негізі бола алатынын атап көрсетеді.

Түйін сөздер: цифрлық архивтер; экономика; технологиялар; жасанды интеллект; технологиялардың таралуы; Солоу парадоксы; еңбек өнімділігі; автоматтандыру; «сынақтан өткізу және үйрену» тәсілі

ЭКОНОМИКА ЦИФРОВОГО АРХИВИРОВАНИЯ В ЭПОХУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Джон Шеридан
Директор по цифровым технологиям и информации
Национального архива Великобритании
john.sheridan@nationalarchives.gov.uk

Аннотация. В статье рассматриваются экономические аспекты цифрового архивирования в эпоху искусственного интеллекта (ИИ). Современный этап характеризуется стремительными технологическими изменениями. Появление больших языковых моделей (Large Language Models, LLM) привело к созданию высокоэффективных, сложных и адаптивных систем искусственного интеллекта. Несмотря на то что цифровые архивные ресурсы все активнее используются ИИ-ботами, а при отсутствии соответствующего регулирования эта тенденция может приобрести экспоненциальный характер, архивы должны сохранять свою человекоориентированную миссию, рассматривая документы прежде всего как источник доказательной информации и коллективной общественной памяти.

Вместе с тем цифровые архивы объективно нуждаются в применении ИИ. Его внедрение и использование являются не только необходимыми, но и неотложными. Масштабы и разнообразие цифровых документов, а также стремительный рост объемов данных значительно превосходят возможности существующих методов экспертизы ценности, отбора, проверки на наличие конфиденциальной информации, обеспечения сохранности, архивного описания и организации доступа. Искусственный интеллект представляет собой практически единственную реальную возможность обеспечить необходимый рост производительности в сфере цифрового архивирования.

В статье также анализируются уроки предыдущих технологических революций. Рассматриваются теория распространения технологий (Technology Diffusion Theory), парадокс Солоу (Solow's Paradox) и болезнь издержек Баумоля (Baumol's Cost Disease). Каждая из этих концепций показывает, что внедрение новых технологий и получение от них ощутимых преимуществ требуют длительного времени. С экономической точки зрения архивная отрасль представляет собой сравнительно небольшой сектор с ограниченным рыночным влиянием. Поэтому финансовая устойчивость будет во многом определять, какие инструменты искусственного интеллекта станут доступны архивам, тогда как правовые нормы, этические принципы, общественное доверие и институциональная ответственность будут определять условия их безопасного и ответственного применения.

В заключительной части статьи определены три практических приоритета развития искусственного интеллекта в архивной сфере: разработка методов и данных для оценки эффективности применения ИИ в архивной деятельности; декомпозиция архивных процессов с целью максимальной автоматизации при четком и взвешенном распределении функций между человеком и машиной; повышение качества данных, используемых искусственным интеллектом, в том числе посредством применения самого ИИ для создания и совершенствования архивных справочных данных.

Автор приходит к выводу, что дальнейший прогресс возможен лишь при условии проведения архивами систематических экспериментов на основе принципа «тестировать и учиться» (test and learn), а также открытого обмена накопленным опытом и полученными результатами.

Ключевые слова: цифровые архивы; экономика; технологии; искусственный интеллект; распространение технологий; парадокс Солоу; производительность; автоматизация; подход «тестировать и учиться»

Introduction

We are living through a time of rapid technological change and advancement. The rise of Large Language Models is giving us highly capable, sophisticated, and adaptable AIs. This in turn is creating new opportunities and new challenges for archives and archivists.

We are not alone. AI adoption is advancing in software engineering, medicine, health care, education, public administration, culture, the creative industries, architecture, the law and so on. Digital archiving is a much smaller sector than all of those, with little market power or influence. We are a niche sector, comparatively, when it comes to technology solutions. Moreover, as the beneficiaries of our work include future generations, the preservation of digital archives of public records has many characteristics of an economic public good: its value is widely shared, and its benefits are not exhausted by use. This helps explain why archives are publicly funded in many jurisdictions. At the same time, the services and infrastructure through which records are accessed, including storage, compute, bandwidth and high-volume machine use, can impose real and uneven costs. The public funding model has consequences for the level of capital investment available to archives, and confidence levels that are needed for investment decisions. We are generally cautious about undertaking risky or speculative developments with the latest cutting-edge technology, using public money. We are understandably more cautious than venture capital backed start-ups, where there is the incentive of establishing first mover advantage, and the potential of large financial returns.

From the standpoint of The National Archives, the question is how might AI best be used for maximum benefit, within our risk appetite? What are the most important opportunities, and how do we mitigate the risks? How do we maximise the return on our investment? As the transaction costs reduce across record creation, transfer, preservation, description, and access, what does AI do to the economics of archival custodianship of digital records long term, in terms of our fixed and variable costs? What does AI mean for our work keeping evidence, enabling transparency and public accountability? Are there charging regimes, or commercial opportunities, that might help offset the costs and protect the taxpayer against poorly targeted consumption of public resources, such as charging AI bots and crawlers? And how do we govern our use of AI, as trustworthy custodians?

The role of archives

Archives are about, and for, humans. This was self-evident for most of our history. Today the archive's essential humanity has acquired fresh relevance, as our digital collections are increasingly processed, searched, summarised and interpreted by AI systems and tools. An archive is not simply a store of data for AI. We are institutions trusted with maintaining records as evidence, sustaining public memory, and

balancing the sometimes conflicting rights and interests concerning access to records, over time.

Records allow us to examine relations among individuals, communities and states, and to see how those relations have changed. As archivists, we do not settle the meaning of the evidence from the past; we preserve the evidential and contextual conditions under which others can investigate it. As AI lowers the cost of producing derivative accounts through prompts, records with secure provenance become more significant and valuable.

For this reason, the debate about AI matters to us in practical terms. If automated systems are to mediate access to records, assist in selection, or summarise material for users, then archival choices about context, provenance and accountability are part of the infrastructure through which those systems operate. We need to find ways of using AI that are consistent with our archival purpose and responsibilities.

Digital archiving and the responsibility of custodianship

At the heart of our purpose is custodianship. This is not limited to physical control over records. We have also to consider what Jenkinson called, ‘the moral defence of the record’. AI casts a fresh perspective on what this means.

Custodianship of digital records is active. It entails continual decisions about preservation, description and access, and about the rights and interests of people and communities, referred to in, or affected by, those decisions.

At The National Archives in the United Kingdom, our doctrine is to view digital preservation as an ongoing exercise of risk management over time. We identify risks to the survival, integrity and use of digital records, and mitigate those risks as best we can through preservation planning and action. We have developed models for digital preservation risk management and scenario planning, such as DiAGRAM. Systems,

tools, data, and standards are all necessary to doing this work, but they are subordinate to the records and to meeting the needs of users.

This practitioner view also reminds us that digital preservation is never a single act. It is a sequence of judgements and decisions over time. We can imagine AI assisting at several points in the decision making process. However, accountability for the decision making cannot be outsourced to an AI Agent. Custodianship is about taking responsibility for deciding what risks matter, which to then mitigate, and what type and level of assurance is acceptable. We can use AI to help do this work. But there is no meaningful way, at present, in which an AI Agent Archivist could be effectively held to account.

The practical difficulty is that our custodianship now takes place in an environment of accelerating technological change. The incentives that shape the wider technology industry determine the landscape. We may want, or prefer, long term stable storage solutions, durable formats and slowly changing platforms. In practice, because of the economics of the technology industry, we operate in an environment where we are navigating relentless change. We often have no choice about the technology systems we use. We just have to keep up with the market.

Archives' productivity challenge

The most immediate difficulty for digital archives is one of scale and scalability. Born-digital records require appraisal, selection, sensitivity review, description, transfer, preservation and access management. The demands on those processes exceeds the capacity of contemporary archival and information-management practices by at least an order of magnitude. This is not a new problem for archives. However, the increase in demand is exponential.

This scalability problem leads directly to the question of productivity. Greater automation is not a matter of preference but of viability. AI is the only plausible source of the productivity gains we now need. However, today's AI tools have not been devised, developed or configured, for archival work. This is a general-purpose technology, that has to be adapted and fitted to archival purposes and use. Such adaptation takes time and investment.

Stepping back, what are the opportunities to increase productivity using AI? In the wider economy there is vigorous debate about the scale of productivity growth that might be achieved. Some forecasts are markedly optimistic: Goldman Sachs Research has estimated that generative AI could raise global GDP by 7 percent and increase productivity growth by 1.5 percentage points over the next decade. Other accounts are more cautious, suggesting that the effects of AI may be much smaller, slower and distributed in ways that leave many organisations with little immediate benefit.

Why the difference in the estimates? Part of the explanation is that the world is a complicated place. AI capabilities need to be adapted to different contexts of use. This adaptation process takes time and money. There is an important difference between what is technically feasible innovation and what is economically viable. And economic viability is the key determining factor of what comes to pass.

Technology Diffusion Theory and Wardley Mapping

Technology Diffusion Theory helps explain how new innovations spread, moving from a few pioneers to general adoption over time. Economically, this process is driven

by costs, risk and benefits. Early on, new technologies are expensive and risky, so only a few businesses or consumers adopt them. As the technology matures, the costs reduce, risks decrease, and competition sees wider adoption. This applies to those innovations that are sufficiently useful to survive in the market. Many are not. The adoption process creates an "S-shaped" pattern over time. Technologies that prove useful, economically viable and supported by the necessary complementary practices tend to diffuse over time, until they mature, commoditise or are displaced by later innovations.

At The National Archives we use Wardley's Mapping method to help visualise technology diffusion in our context. The map shows how components of our systems evolve. The process is driven by competitive forces. Wardley maps manage the evolution along the horizontal axis not as a function of time, but as the outcome of economic competition. Thus the horizontal axis divides evolution into four distinct stages: Genesis, Custom-built, Product, and Commodity. This allows us to understand and plan our technology infrastructure by directly mapping it onto the economic realities of diffusion theory. For example, over time we see that digital preservation systems have become increasingly productised. Cloud computing has commoditised both compute and data storage. We have substantially benefitted from these changes in technology by understanding the process and accounting for it in our technology strategy.

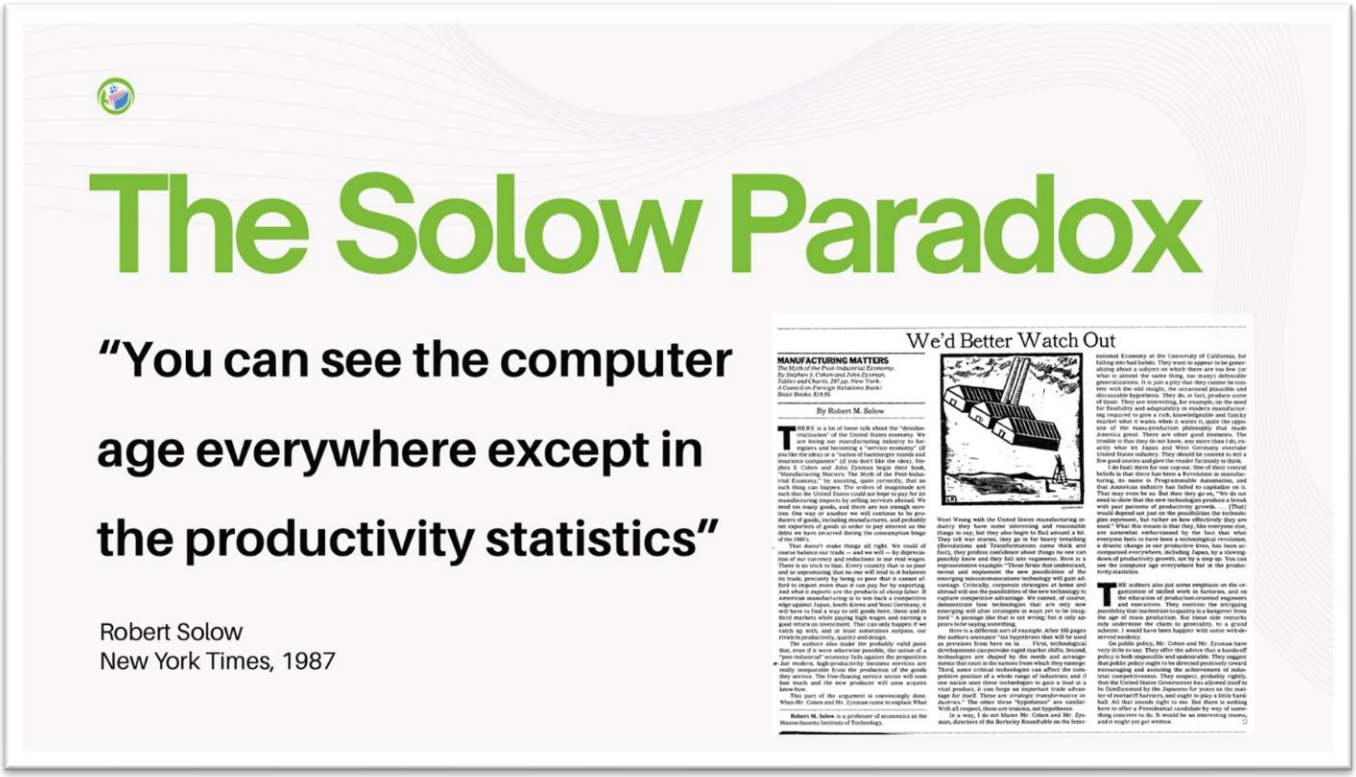
The productivity paradox of new technology

The Intel 4004 was released on 15 November 1971 and is widely recognised as the first commercially available microprocessor. The history of the microprocessor illustrates the diffusion process for arguably the most impactful innovation of the 20th century. Yet the microchip revolution did not immediately transform organisational productivity. It took time for cheaper processing power to diffuse into the Personal Computer of the 1980s, and the eventual transformation that was the general computerisation of the workplace. Over several decades of computerisation, microprocessors were slow to impact productivity (even with Moore's law helping to drive progress).

Robert Solow's 1987 observation in the New York Times, that "you can see the computer age everywhere except in productivity statistics" captures the issue. There are, at best, long lead times from initial technological innovation through to widespread adoption and then to measurable economic effect. We should expect the same with AI today – a pervasive new technology that may not deliver the productivity gains anticipated of it in the short term. This phenomenon is called the "Solow Paradox". This likely delay strengthens rather than weakens the case for action now: if the benefits depend on adaptation, governance, skills and redesigned processes, then archives need to begin that work before the scale problem becomes still harder to manage.

But why the lag before the gains? Four explanations are worth exploring. The first is simple overstatement: a new technology may simply turn out to be less transformative than its advocates claim. The second is mismeasurement: productivity is difficult to measure, so the gains may be hard to detect with standard metrics. This is especially true of technologies that unlock latent demand by lowering transaction costs. Now things that were too expensive to happen before the innovation, can be

done affordably. The third is distribution: productivity gains may accrue disproportionately to those who own or control the technology, rather than to institutions such as ours, where the benefits might be fully offset by various drawbacks. For example, a range of disbenefits from AI for archives has already emerged, including rapidly expanding levels of bot traffic, increasing software supply chain risks and cyber security threats, misinformed AI generated user enquiries that cite hallucinated records and so on. The fourth is institutional lag: organisations need time to adapt to new technology, to understand how to use it, and then to redevelop business processes – first putting in place underpinning governance, as well as having people with the right skills.



Baumol's cost disease and the hybrid archive

The next issue to grapple with is the uneven distribution of productivity gains across and between different sectors. Baumol posits that the economy has different types of sectors – high productivity sectors where innovations have the potential to achieve substantial gains; and low productivity sectors, where labour costs are proportionately high, the opportunities for automation low. Over time, productivity growth in high productivity sectors pushes up labour costs overall, making lower productivity sectors, proportionally increasingly expensive. Healthcare, education, law and cultural work often exhibit this cost pressure. The relative costs rise not necessarily because those sectors are inefficient, but because the quality of outcome depends on skilled human labour, which hitherto could not be automated. This phenomenon is now called Baumol's Cost Disease, and it explains, to quote from the cover of Baumol's book, "why computers get cheaper and health care doesn't".

Archives are a hybrid sector – with areas of both high and low potential for productivity gains. We can achieve substantial productivity gains in our digital

archiving work – as we must – but that needs to be supported by investment. Other elements of the work, including the operation of physical repositories, conservation, advising readers in the reading room, applying professional judgement in decision making with complex trade-offs and risks, has far less scope for automation. These human-intensive functions may therefore become relatively more costly for archives over time. With public funding constraints, the risk is that cost increases in less productive areas of the archive's work potentially narrows the funding available for AI enabled transformation, despite the opportunity for far greater productivity gains for the digital archive.

With investment in AI likely to be scarce in the archives sector, the practical question, therefore, is how to pursue automation using AI selectively, where it makes the most difference. Archives must consider carefully how to maximise the benefits of their investment in AI, accelerating the change process in those areas where investment makes most difference.

Advancing AI adoption in archives

We are seeing the most rapid advances in AI capabilities in economically powerful domains, such as software engineering, where the market is large, the financial returns substantial, and investment incentives strong. Like many others, archives are benefitting from these advances, using AI to help with our bespoke software development.

In smaller sectors such as archives, the same underlying technology may create the situation where AI enabled automation, that has become technically feasible, is not sufficiently financially viable to progress. The economic value of a tool that improves archival appraisal or sensitivity review may be very high in public value terms, but it may not generate the kind of private return that sustains a large product market. This is a core asymmetry: economically powerful sectors can turn technical possibility into products because the investment case is obvious. By contrast, archives need to assemble value through collaboration, shared methods, adaptation of general tools and careful prioritisation of investment.

This means that our path to adoption is different from that in larger sectors, both commercial and public. We should expect to borrow and adapt tools developed elsewhere first, to prove value through narrower tasks, and to invest selectively in the areas where the benefit is clearest. Our chances of success will be improved if we have a clear idea about how the change process will progress.

From point solutions to system change

In the book, "*Prediction Machines – the simple economics of Artificial Intelligence*", the authors set out to characterise the progression of AI. They chart a course from point solutions to system change. Point solutions improve or automate individual tasks while leaving the broader organisation of work largely intact. Application-level change assembles such improvements into products or workflows that alter outcomes across a process. System-level change goes further, making possible new patterns of work and, in some cases, displacing earlier arrangements altogether.

Most of our archival applications of AI remain at the point-solution stage. They include support for handwriting recognition, content summarisation, named entity extraction, metadata creation, and increasingly researcher support. In the near term,

we will be refining tools in those areas, but we should understand them as steps towards application-level and system-level change rather than as the destination.

The distinction is helpful for understanding where we are. Point solutions can create a strong impression of progress, unlock further investment, but leave the overall burden of work for the archive unchanged. A tool that helps identify named entities, summarise documents or suggest metadata may be useful, but if its outputs must all be manually checked, the gains may be modest. Application-level change begins when tools are integrated into redesigned workflows, with clear thresholds for confidence, escalation and human review. System-level change goes further still, affecting how records are prepared for transfer, how archival description is produced, and how users encounter records in context.

The larger gains for archives will come as AI is incorporated into redesigned systems for the creation, transfer, preservation and use of digital records. We have not yet reached that stage. Some of the tasks, such as sensitivity review, are difficult to automate fully. We are still in the midst of determining what can responsibly be automated, what requires human judgement, and how AI enabled systems can be deployed in ways that are compatible and sustain trust and accountability.

Three priorities for archival AI

Three areas for intervention follow, as we move from point solutions to system change.

The first is evaluation. We should deploy AI only where we have evidence about what it does well, what it does poorly and what risks it introduces. Evaluation is comparatively tractable in well bounded settings, such as testing a redaction assistant against defined criteria. It is much harder in open-ended settings, such as helping researchers decide which records to order in the reading room. We need to develop our knowledge and skills in evaluating AI, including the development of new evaluation methods and evaluation data designed to assess the efficacy of AI in archival work.

The second is task decomposition. We need to break archival processes into smaller components. This will aid introducing AI where AI can be shown to help, whilst enabling people to remain responsible for tasks that require human input or decision making. It is not simply a matter of inserting a tool into an existing workflow. It requires the redesign of processes for partial automation without treating human work as if it were machine work.

Task decomposition is a practical design discipline. It requires us to describe archival work more precisely. A process such as sensitivity review, for example, contains many smaller judgements: identifying personal data, recognising categories of harm, applying legal or policy tests, recording reasons and deciding whether a case should be escalated. AI may help with some of these steps and be unsuitable for others.

Decision making about acceptable error rate links this second priority to the first: evaluation. A tool used to suggest search terms to a researcher may tolerate a different level of imperfection from a tool used to support sensitivity review or access decisions. Accuracy is only one dimension. We also need to understand bias, explainability, consistency, auditability and the effort required to correct mistakes. In archival settings, a system that appears efficient but makes errors difficult to see may increase risk rather than reduce it.

The third is the reciprocal relationship between providing better data to AI, and using AI to create better data, including for AI use. Systems perform better when they are grounded in relevant, current and well-structured information. Retrieval-augmented generation and emerging Model Context Protocol based systems show the benefits of that approach. Excitingly, we can now use AI to help us to create, maintain, and link the data on which improved AI systems now depend.

For example, at The National Archives, we have used AI, in a carefully controlled way, to create a new dataset of powers and duties that are conferred on public authorities in legislation. We have then been able to use this data to improve the performance of Large Language Models when answering questions about the law.

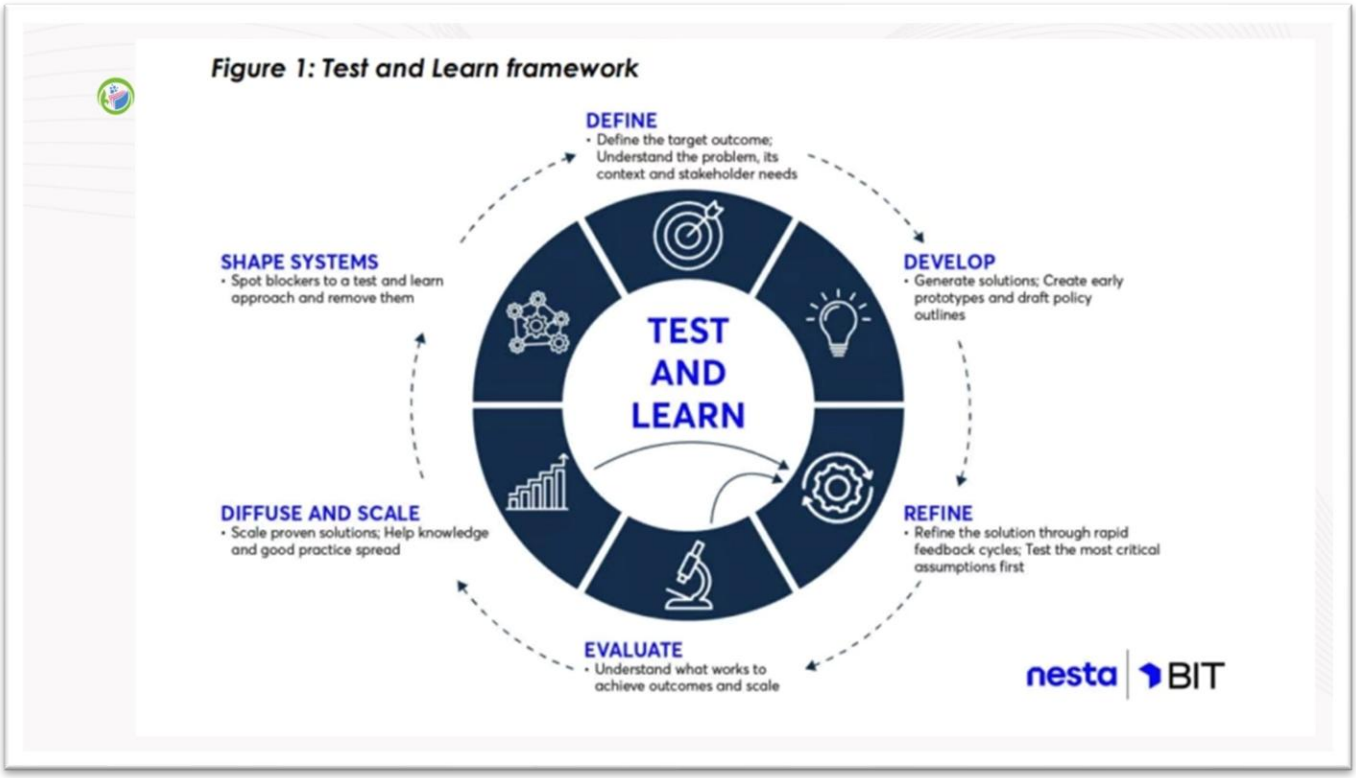
The same reciprocal logic applies to archival metadata. We should not imagine better data only as an input prepared in advance of automation. In practice, AI may help us normalise names, identify relationships between records, suggest dates, cluster related material, improve search terms and expose inconsistencies in existing descriptions. Those improvements to our data will then make later automation more reliable. The process requires clarity over data standards, as well as a thought through view of what level of error is acceptable for a given use.

Test, learn and share

There is, as yet, no settled engineering methodology for applying AI to archival work. Nor is there a mature doctrine on which we can rely. Early guidance, standards and research initiatives, such as the work of InterPARES Trust AI are of tremendous help, but we still have much to learn in practice.

The practical application of AI in archives, therefore, requires us to test, learn, refine, adapt and test again. We need to undertake bounded experiments, measure results, adjust methods and share evidence of both our successes and failure.

Sharing what we are learning is a fundamental part of this discipline. The archives sector is too small for every institution to repeat the same experiments in isolation. The market is too specialised and fragmented for us to assume that suitable products will arrive fully formed. Experimentation is not some preliminary stage before serious work begins. It is how the serious work of applying AI to archives must now be done.



Conclusion

AI presents archives with opportunities and challenges in equal measure. It is the most plausible route to the productivity gains needed to address the growth, complexity and variety of born-digital records, but it is not a solution in isolation. Its benefits will depend on organisational redesign, better archival data, careful evaluation, and clear decisions about where human judgement and institutional accountability must remain central.

The economics of technology suggests that adoption will be uneven. It will move quickly where markets are large, investment incentives are strong, and products can be generalised; it will move more slowly where public value is high but commercial returns are uncertain.

For archives, the task is therefore not to wait for fully formed solutions, nor to pursue automation indiscriminately, but to make deliberate choices about evaluation, task decomposition and the improvement of archival data. Above all, progress will depend on sustained experimentation: testing, learning and sharing evidence. The likely delay between technological possibility and realised productivity gain makes this work more urgent, not less. In the age of AI, the archive remains a human-centred institution, but one that must learn to use machines well if we are to continue preserving evidence, memory and accountability at digital scale.

References

Acemoglu, D. and Johnson, S. 2023. *Power and Progress: Our 1,000 Year Struggle Over Technology and Prosperity*.

Agrawal, A., Gans, J. and Goldfarb, A. eds. 2019. *The Economics of Artificial Intelligence: An Agenda*.

Agrawal, A., Gans, J. and Goldfarb, A. 2022. *Prediction Machines: The Simple Economics of Artificial Intelligence*.

Baumol, W. J. 2012. *The Cost Disease*.

Behavioural Insights Team (BIT). 2025. *AI and Human Behaviour – ADOPT Report*.

Duranti, L. and Rogers, C. eds. 2024. *Artificial Intelligence and Documentary Heritage*. SCEaR (Sub-Committee on Education and Research), UNESCO Memory of the World Newsletter Special Issue.

Eitel-Porter, R., Dongha, P. and Vogel, M. 2025. *Governing the Machine: How to Navigate the Risks of AI and Unlock Its True Potential*.

Ghosh, M. 2025. *Decoding User Readiness for Sustainable AI Adoption: A Behavioural Approach Through Technology Readiness Segmentation*.

Haynes, L. et al. 2012. *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials*. DOI: 10.2139/ssrn.2131581.

Jenkinson, H. 1922. *A Manual of Archive Administration*.

Maya, L. 2026. *The Age of Abundance: AI and the Future of Work, Wealth and Purpose*.

School of Archivology, National University of Córdoba. 2022. *ITrust AI: The New InterPARES Research Project*. School of Archival Studies Yearbook, No. 13. Published online June 9, 2022.

Solow, R. 1987. “We’d Better Watch Out.” *New York Times Book Review*, July 12, 1987, p. 36.

The National Archives (UK) and the University of Warwick. 2020. *DiAGRAM: The Digital Archives Graphical Risk Assessment Model*.

Vallor, S. 2025. *The AI Mirror: How to Reclaim Our Humanity in an Age of Machine Thinking*.

Wardley, S. 2015. *An Introduction to Wardley (Value Chain) Mapping*.

QUO VADIS, DIGITIZATION? FROM «DUST» TO «CAPITAL»

Piotr Zawilski
Master's Degree, Director of the National Digital Archives of Poland
piotr.zawilski@nac.gov.pl, https://orcid.org/0009-0007-2636-4119

Abstract. This article examines the evolution of the GLAM sector in Poland, with a particular focus on the National Digital Archives (NAC) and the szukajwarchiwach.gov.pl (SwA) portal. It traces the transformation from laborious microfilming to mass digitization at the turn of the 21st century, which made hundreds of millions of documents available in open access. Using SwA as a primary case study - which by 2026 hosts over 84 million scans - the author highlights the successes and inherent paradoxes of this process. Despite decades of effort and significant financial investment, digitized materials represent only about 2% of total archival holdings, raising critical questions about the efficiency and ultimate purpose of these projects.

A central theme of the analysis is the paradigm shift in user demographics occurring between 2024 and 2025. Traditional "protein-based" researchers are rapidly being outnumbered by crawlers and bots fueling Artificial Intelligence (AI) models. By 2026, archival collections have gained a new valuation as "digital gold" – high-quality, verified training data for Large Language Models and Handwritten Text Recognition projects. The text notes that machine-generated traffic now accounts for over 50% of web activity globally, and in some library services, it exceeds 90%. Furthermore, the author addresses the emerging economic and legal challenges. Public institutions must bear the rising costs of server maintenance, cybersecurity, and infrastructure to handle millions of "empty" machine queries, without receiving compensation from the AI industry. Meanwhile, EU regulations, such as the Open Data Directive, mandate that data be provided in machine-readable formats, fostering a "bot-first" environment.

The article concludes that modern digitization must satisfy the needs of both humans and algorithms. Scans serve as a foundation for credibility in an AI-dominated world, reinforcing the principle that information is only truly valuable when it is accessible and actionable for all users.

Keywords: mass digitization; GLAM sector; National Digital Archives; Artificial Intelligence; crawlers; training data; Large Language Models; Handwritten Text Recognition; Open Data; machine-readable formats; data traffic

QUO VADIS, ЦИФРЛАНДЫРУ?
«ШАҢНАН» ЦИФРЛЫҚ КАПИТАЛҒА

Петр Завильски
Магистр, Польша Ұлттық цифрлық архивтерінің директоры
piotr.zawilski@nac.gov.pl, https://orcid.org/0009-0007-2636-4119

Аңдатпа. Мақалада Польшадағы GLAM секторының эволюциясы қарастырылып, басты назар Польшаның Ұлттық цифрлық архивінің (NAC) және szukajwarchiwach.gov.pl (SwA) порталының қызметіне аударылады. XXI ғасырдың басындағы еңбек сыйымды микрофильмдеуден жүздеген миллион архивтік құжатты ашық қолжетімділікке шығаруға мүмкіндік берген жаппай цифрландыру кезеңіне көшу үдерісі талданады. 2026 жылға қарай 84 миллионнан астам цифрлық көшірмені қамтитын SwA порталы мысалында автор бұл үдерістің жетістіктерімен қатар ішкі қайшылықтарын да көрсетеді. Көпжылдық жұмыс пен елеулі қаржылық инвестицияларға қарамастан, цифрланған материалдар архив қорларының жалпы көлемінің небәрі шамамен 2 %-ын ғана құрайды. Бұл мұндай жобалардың тиімділігі мен түпкі мақсатына қатысты маңызды мәселелерді көтереді.

Зерттеудің негізгі тақырыбы – 2024–2025 жылдары орын алған пайдаланушылар құрамының түбегейлі өзгеруі. Дәстүрлі зерттеушілердің санын біртіндеп жасанды интеллект модельдерін оқытуға пайдаланылатын веб-краулерлер мен автоматтандырылған боттар басып озуда. 2026 жылға қарай архив қорлары жаңа құндылыққа ие болып, үлкен тілдік модельдерді (LLM) және қолжазба мәтіндерін тану (Handwritten Text Recognition, HTR) жобаларын оқытуға арналған жоғары сапалы әрі тексерілген деректер – «цифрлық алтын» ретінде бағалануда. Сонымен қатар, машиналық трафик бүгінде әлемдік веб-трафиктің 50 %-дан астамын құрайтыны, ал кейбір кітапхана қызметтерінде бұл көрсеткіш 90 %-дан асатыны атап өтіледі.

Автор сондай-ақ жаңа экономикалық және құқықтық мәселелерді талдайды. Мемлекеттік мекемелер серверлерді қолдауға, киберқауіпсіздікті қамтамасыз етуге және миллиондаған автоматтандырылған сұраныстарды өңдеуге арналған инфрақұрылымды дамытуға жұмсалатын шығындардың өсуін өз мойнына алады, алайда жасанды интеллект индустриясынан өтемақы алмайды. Сонымен қатар, Еуропалық Одақтың нормативтік құжаттары, оның ішінде Ашық деректер жөніндегі директива, ақпаратты машинамен оқылатын форматтарда ұсынуды талап етіп, ең алдымен автоматтандырылған жүйелерге бағдарланған ортаны қалыптастырады. Қорытындыда қазіргі цифрландыру үдерісі адамдардың да, алгоритмдердің де қажеттіліктерін қанағаттандыруы тиіс екендігі атап өтіледі. Архивтік құжаттардың цифрлық көшірмелері жасанды интеллект дәуіріндегі сенімділіктің негізіне айналып, ақпараттың барлық пайдаланушылар үшін қолжетімді әрі қолдануға жарамды болған жағдайда ғана шынайы құндылыққа ие болатынын дәлелдейді.

Түйін сөздер: жаппай цифрландыру; GLAM секторы; Ұлттық цифрлық архив; жасанды интеллект; веб-краулерлер; оқыту деректері; үлкен тілдік модельдер (LLM); қолжазба мәтіндерін тану (HTR); ашық деректер; машинамен оқылатын форматтар; деректер трафигі

QUO VADIS, ЦИФРОВИЗАЦИЯ? ОТ «ПЫЛИ» К ЦИФРОВОМУ КАПИТАЛУ

Петр Завильски
Магистр, директор Национальных цифровых архивов Польши
piotr.zawilski@nac.gov.pl, https://orcid.org/0009-0007-2636-4119

Аннотация. В статье рассматривается эволюция сектора GLAM в Польше с особым акцентом на деятельность Национального цифрового архива (NAC) и портала szukajwarchiwach.gov.pl (SwA). Прослеживается переход от трудоемкого микрофильмирования к массовой оцифровке на рубеже XXI века, благодаря которой сотни миллионов архивных документов стали доступны в открытом доступе. На примере портала SwA, который к 2026 году содержит более 84 млн цифровых копий, автор анализирует как достижения, так и внутренние противоречия данного процесса. Несмотря на десятилетия работы и значительные финансовые вложения, оцифрованные материалы составляют лишь около 2% общего объема архивных фондов, что ставит важные вопросы об эффективности и конечных целях подобных проектов.

Центральной темой исследования является смена парадигмы пользовательской аудитории, произошедшая в период 2024–2025 годов. Поисковые роботы и автоматизированные системы, обеспечивающие обучение моделей искусственного интеллекта, постепенно начинают численно превосходить традиционных исследователей. К 2026 году архивные коллекции приобретают новую ценность как «цифровое золото» – высококачественные и верифицированные обучающие данные для больших языковых моделей (LLM) и проектов по распознаванию рукописного текста (Handwritten Text Recognition, HTR). Отмечается, что машинный трафик уже составляет более 50% мировой веб-активности, а в некоторых библиотечных сервисах превышает 90%.

Кроме того, автор рассматривает возникающие экономические и правовые вызовы. Государственные учреждения вынуждены нести возрастающие расходы на обслуживание серверов, обеспечение кибербезопасности и развитие инфраструктуры для обработки миллионов автоматизированных запросов, не получая при этом компенсации со стороны индустрии искусственного интеллекта. Одновременно нормативные акты Европейского союза, включая Директиву об открытых данных, требуют предоставления информации в машиночитаемых форматах, формируя среду, ориентированную прежде всего на автоматизированные системы.

В заключение подчеркивается, что современная цифровизация должна учитывать потребности как людей, так и алгоритмов. Цифровые копии архивных документов становятся основой доверия в эпоху искусственного интеллекта, подтверждая принцип, согласно которому информация обладает реальной ценностью лишь тогда, когда она доступна и пригодна для использования всеми категориями пользователей.

Ключевые слова: массовая оцифровка; сектор GLAM; Национальный цифровой архив; искусственный интеллект; поисковые роботы; обучающие данные; большие языковые модели (LLM); распознавание рукописного текста (HTR); открытые данные; машиночитаемые форматы; трафик данных

"Un trésor qui ne se partage pas est un trésor qui n'existe pas."
– French proverb

The turn of the 20th and 21st centuries brought a technological revolution to GLAM sector institutions (Galleries, Libraries, Archives, Museums). Laborious microfilming and photocopying were replaced by mass digitization. Literally within a decade, hundreds of millions of scans of archival materials, books, and periodicals became available in open online access. Some fell into euphoria and self-satisfaction, saying: "Look, we are making gigantic collections available online. Please, help yourselves".

In Poland, one of the leaders in this field is the National Digital Archives (Narodowe Archiwum Cyfrowe – NAC), which since 2009 has been operating and constantly developing a popular internet portal forg researchers available at: szukajwarchiwach.gov.pl (SwA)[1*]. It presents information and scans from the resources of 122 institutions (state, university, diocesan, and museum archives, research institutions, associations, and trade unions)[2*]. The launch of mass digitization by Polish archives and the online publication of digital copies was intended to demonstrate their transformation into institutions that are open to citizens, socially engaged, and proactive in responding to current user needs. For this reason, archives prioritized for digitization those materials most frequently requested by users in reading rooms or through research inquiries (e.g., genealogical records, cartography, photographs, etc.).

The portal has firmly established itself as a tool for the workshops of both professional and amateur historians, genealogists, as well as students and publicists. What distinguishes SwA is that it has no designated maintenance (or closure) period - like archives, it is an "eternal" entity, even if its graphic design or tools for searching the shared resources change along the way.

[1*] It presents information and scans not only from the resources of state archives but also from a variety of other institutions that have expressed their willingness to participate.
[2*] As of April 1, 2026.

In SwA, which currently presents over 84 million scans[1*] and 17 million descriptions of archival materials (inventory units), we record from several hundred thousand to over a million visits per month (SwA does not require account creation or logging in). Data from Similarweb reports at the end of 2025 suggest that the site maintains a stable leading position in Poland in the "Ancestry and Genealogy" category. The average visit duration is approximately 6 minutes and 20 seconds. This is a very high result, indicating that users conduct in-depth research there. The average user views about 7–8 pages during a single session. Of course, many people use the service not to view scans, but to read the provided archival inventories.

Do these numbers seem impressive? Seemingly. 17 million descriptions of archival units and 84 million scans seem unimpressive to users. Despite these efforts, Archives face voices of criticism, as there is always someone impatient with the lack of **"**their**"** documents in digital access[2*]. We have the impression that most SwA users seem to no longer remember that until 2009, archival materials were available only in reading rooms.

The scans made available online constitute only a fraction of the total resources of archives and libraries. In the case of the network of Polish state archives, after nearly twenty years, we have barely approached 2%; for Polish libraries, it will be just over 3%. Furthermore, it turned out that statistically, 1 scan made available on the Internet by libraries averaged 0.5 views per year. In the case of Polish archives, the situation looks twice as good - there is 1 view per scan - although in past years, a single scan was viewed 3 or 4 times annually. How do Polish archives compare to other European institutions? Digitalization results in Poland can be compared with those achieved, for example, in France, whose archives hold a similarly sized collection of approximately 400 linear kilometers of records. The French balance sheet is currently lower than the Polish one: 16 million scans in total, 10 million of which are accessible online. This represents approximately 1% of the total French archival holdings (Herve 2023: 4).

Thus, we are dealing with a certain paradox in Poland. After 20 years of work by hundreds of people and spending a mountain of money, we have only about 2% of the total archive resources available on the Internet. Given the current rate of digitization work, the use of roughly 100 scanners in 32 state archives and the budget spent on external scanning services will result in the full digitization of state archive resources in approximately 600 years.

It seems that what Polish archives have actually achieved is awakening the boundless appetites of users, including those from abroad. For years, dedicating enormous strength and resources, they have been chasing the proverbial white rabbit without the slightest chance of catching it. Are archives therefore in a losing position in this pursuit?

Some began to wonder if all collections are really worth urgent digitization for the purpose of online access? A thesis was put forward that it is not worth for one hundred percent coverage of originals with scans. Presumably, in the interest of rational spending of public funds, an attempt was made to estimate the scope of digitization.

[1*] Every year, this figure increases by a further 5 to 8 million.
[2*] At the same time, it appears that users of platforms operated by other entities in the GLAM sector do not lodge such complaints. Explaining this phenomenon lies beyond the scope of this material.

In 2017, European archives determined that as much as 40% of collections do not need to be digitally copied; libraries calculated this parameter at 33% (Nauta 2017: 28).

So the question arose: why and for whom are we scanning, since we are only able to make a small part of the resources available on the Internet? Perhaps this last question was poorly formulated, and we should stop anthropocentric thinking? Not for whom, but for what purpose are we scanning?

Just 2-3 years ago in Europe, an attempt was made to address this problem. A discussion began on the purposefulness of mass digitization in archives, suggesting abandoning it in favor of "on-demand" digitization. Such solutions have already been introduced by Germany and Switzerland.

However, in Poland, as in the rest of the world, at the turn of 2024 and 2025, archives, libraries, and museums began to record significantly increased traffic on their portals. For example, the leading Polish portal Bazhum, which presents scans of scientific journals, suddenly began reporting a more than 30-fold increase in file downloads. What happened that files previously viewed for several minutes are suddenly "read" in 2 seconds? It is not crowds of enthusiasts who suddenly felt a passionate love for discovering cultural heritage. Data shows that the increased traffic in library services comes from the USA, China, and Singapore; for SwA mainly from the USA, Brazil, and Mexico—these must be crawlers, i.e., bots collecting data to feed AI models[1*]. Simultaneously, traffic from Poland in library services is falling to... 7th-8th place, and in SwA to 3rd place[2*]!

Online platforms and digital repositories, created for researchers, students, and enthusiasts in mind, are today overloaded by the AI industry. While in SwA this traffic is distributed more or less evenly, i.e., about 60% to 40% in favor of humans (in 2025, 77 million file views in SwA by 1.8 million unique users were accompanied by 54 million requests from bots), in library services, the advantage of crawlers is overwhelming. There are voices saying it may even be above 90% of all traffic. The explanation is simple: libraries primarily provide printed materials that are easy to recognize using OCR (Optical Character Recognition) technology. The services offered by libraries are the primary source of knowledge for LLM (Large Language Models) and indexing bots (OAI-SearchBot or Google-InspectionTool), which need high-quality scientific texts to better understand cultural and historical context. These bots download metadata and full texts in the form of PDF files.

Network traffic in SwA still maintains some balance, but this may soon change. We see the possibility of a threat in the near future regarding the overloading of our servers by traffic from research projects dealing with HTR (Handwritten Text Recognition). These bots can "click through" thousands of scans of vital records in a single day. Although the number of sessions may be smaller than in library services, the load on data links they generate is many times greater due to the size of the graphic files[1*].

Over the past several months, perhaps to the surprise of archives and libraries, a complete re-evaluation of their collections has taken place. From repositories of

[1*] Reports from HUMAN Security and Imperva show that in 2026, the Internet reached a tipping point where over 50% of all web traffic is generated by bots.
[2*] The measured traffic may be skewed due to the potential use of VPNs in other locations.
[1*] It is worth noting that SwA is constantly updated with new digital copies - in March 2026 alone, another 3 million scans were added. Each such update triggers a new wave of activity from bots indexing the fresh data.

hundreds of kilometers of 'paper' - valuable for cultural heritage but difficult to 'sell' in a business sense - they have transformed into a vast source of potential income. In 2026, archival and library collections gained a new, gigantic valuation as "clean training data". Unlike the "dirty" Internet full of spam, these collections are verified, high-quality sources of knowledge. Technology companies are willing to pay millions for licensed access to such databases to train AI models on historical texts, literature, and science, making these collections "digital gold". All of this inevitably leads to generative artificial intelligence, whose numerous emerging services benefit greatly from archival and library databases[2*].

On the other hand, thanks to tools such as ChatGPT (OpenAI), Gemini (Google), Grok, Deep Seek, Claude (Anthropic), and Copilot (Microsoft), recipients can reach resources they had no prior knowledge of - creating new connections, gaining clues for further research, and focusing on working with an object rather than finding it.

Institutions providing open online resources face new challenges, but perhaps also new opportunities. Culture and science institutions, primarily financed from public funds, are forced to bear additional, previously unplanned costs for the maintenance and operation of their repositories caused by the rapid increase in network traffic. Additional costs are related to cybersecurity (protection against malicious bots) and the purchase of additional servers to handle millions of "empty" machine queries. In the longer term, will this not discourage them from further digitization work?

As of now, in European Union countries, there seems to be no possibility of launching a system of fees or compensation for GLAM sector institutions providing their collections online. They therefore bear the costs without a chance to share in the profits generated by the AI industry. Since 2016, re-use regulations have been in force, i.e., regarding the re-use of public sector information. These regulations define the rules for further use, including commercial use, of data from offices, archives, or libraries. Since 2019, through Directive the so-called Open Data (Directive 2019/1024, 32), these have been expanded to include data provided online (in practice, this means there is now a much stronger legal basis to demand data from an archive or library in a form "for a bot" and not just "for viewing"). Data must be available in bulk, in machine-readable formats (e.g., CSV, JSON, XML). To sum up: the law encourages the use of bots, provided that you do not "ride" servers senseless. The Internet is becoming a "bot-first" environment. Most of the content created today is consumed not by human eyes, but by algorithms indexing and training new AI models.

Does this mean that in the near future "protein-based" users will cease to matter, and portals will be developed to satisfy the needs of AI? Or maybe there is no such threat? Both sides have equally strong positions, and their needs will be met at an equal level. It seems we can already provide answers to the questions: 1. whether, 2. what, and 3. why to scan? The answer is probably easier than it was 2-3 years ago. 1. Yes, scan! 2. Everything, with particular emphasis on the most valuable and sought-after materials. 3. Here the answer requires deepening. Every scan shared on the Internet is or will be valuable - if not for a human, then for AI, which will use it in the process of

[2*] Generative Artificial Intelligence (Generative AI) is a type of AI that creates new content - including text, images, code, audio, and video - by learning patterns from massive datasets. Unlike traditional AI, which analyzes existing information, this technology mimics human creativity by generating original outputs based on prompts.

learning, indexing, or creating new content. It is important that in this process it does not lose credibility and remains trustworthy.

Conclusion

The transition from "dust" to "capital" is a crucial moment in which the collections of archives and libraries can no longer be seen as static relics, but as valuable fuel for the digital economy. This transformation is driven by the explosive growth of AI models, which require huge, structured datasets for training and fine-tuning. Archives and libraries provide the "clean", validated data that is essential for training large language models (LLMs), understanding complex contexts, different languages, and nuanced human knowledge. As a result, what was once considered a commercially dormant paper is now a strategic asset sought by tech giants for its authenticity and depth. This new paradigm forces cultural institutions to rethink their role, balancing their law enforcement mission with their newfound status as providers of the highest quality training data. We must remember that: Archives are there to be used. Information only has value if it is available and can be used.

References

Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information. Official Journal of the European Union, L 172. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2019.172.01.0056.01.ENG&toc=OJ%3AL%3A2019%3A172%3ATOC (accessed 24.04.2026).

Herve 2023 – Hervé F. The National Archives of France balancing conservation and digitization. https://www.bundesarchiv.de/DE/Content/Downloads/Veranstaltungen/abstracts_internationale-archivkonferenz.pdf?__blob=publicationFile (accessed 24.04.2026).

Nauta 2017 – Nauta G.J., W. van den Heuvel W. van den, S. Teunisse S. Europeana DSI 2-access to digital resources of European heritage deliverable D4.4. Report on ENUMERATE core survey 4. https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/ENUMERATE/deliverables/DSI-2_Deliverable%20D4.4_Europeana_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf, (accessed 24.04.2026).

CURRENT STATUS AND CHALLENGES OF COLLECTING PRESIDENTIAL PRIVATE RECORDS AT THE PRESIDENTIAL ARCHIVES

Imjeong Ko
Archivist of the Archival Service Division, Presidential Archives of Korea,
M.A. History, Graduate School of Hanyang University, Korea
Koi567@korea.kr

Abstract. Although South Korea's institutional archiving began with the 1999 Public Records Management Act, dedicated presidential records management was formalized through the 2007 Act on the Management of Presidential Records and the subsequent founding of the Presidential Archives in 2008. Tasked with preserving the administrative legacies of five-year single-term presidencies, the Archives systematically collects official transfers alongside valuable private presidential records. These private collections, acquired through donations, oral history recordings, and overseas acquisitions, bridge critical historical gaps left by official documentation. Despite challenges including budget constraints, born-digital data management, and privacy concerns regarding public disclosure, private records offer profound qualitative insights into governing philosophies. Moving forward, the Archives aims to establish balanced disclosure policies and integrate emerging AI technologies to streamline processing and enhance semantic searchability for researchers.

Keywords: Presidential Archives; Private Records; Donation; Oral History

ПРЕЗИДЕНТ АРХИВІН ПРЕЗИДЕНТТІҢ ЖЕКЕ ҚҰЖАТТАРЫМЕН ТОЛЫҚТЫРУДЫҢ ҚАЗІРГІ ЖАЙ-КҮЙІ МЕН МӘСЕЛЕЛЕРІ

Имджон Ко
Корея Республикасы Президенттік архивінің Архивтік қызмет бөлімінің архившісі, тарих магистрі, Hanyang University Жоғары мектебі, Корея
Koi567@korea.kr

Аңдатпа. Корея Республикасында институционалдық архив ісінің негізі 1999 жылы қабылданған Мемлекеттік құжаттарды басқару туралы заңмен қаланса, президенттік құжаттарды басқарудың мамандандырылған жүйесі 2007 жылы қабылданған Президенттік құжаттарды басқару туралы заңның және 2008 жылы Президент архивінің құрылуының нәтижесінде қалыптасты. Президент архивіне бір ғана бес жылдық мерзімге сайланатын мемлекет басшыларының басқарушылық мұрасын сақтауға қатысты міндет жүктелген. Ресми түрде тапсырылатын құжаттармен қатар, Архив президенттердің құнды жеке құжаттарын да жүйелі түрде жинақтайды. Сыйға тарту, ауызша тарих жазбалары және шетелден сатып алынған материалдар арқылы толықтырылатын бұл қорлар ресми құжаттамада сақталмай қалған маңызды тарихи олқылықтардың орнын толтырады. Қаржыландырудың шектеулілігі, цифрлық ортада жасалған құжаттарды басқару, сондай-ақ архивтік ақпаратты жариялау кезіндегі құпиялылық мәселелері сияқты қиындықтарға қарамастан, жеке құжаттар мемлекеттік басқару философиясы мен саяси шешімдер қабылдау үдерісін тереңірек түсінуге мүмкіндік береді. Болашақта Президент архиві архивтік ақпаратқа қолжетімділікті қамтамасыз етудің теңгерімді саясатын қалыптастырып, құжаттарды өңдеуді автоматтандыру және зерттеушілерге арналған семантикалық іздеуді жетілдіру мақсатында жасанды интеллект технологияларын енгізуді жоспарлап отыр.

Түйін сөздер: Президент архиві; жеке құжаттар; сыйға тарту; ауызша тарих

СОВРЕМЕННОЕ СОСТОЯНИЕ И ПРОБЛЕМЫ КОМПЛЕКТОВАНИЯ АРХИВА ПРЕЗИДЕНТА ЛИЧНЫМИ ДОКУМЕНТАМИ ПРЕЗИДЕНТА

Имджон Ко

Архивист Отдела архивной службы Президентского архива Республики Корея, магистр истории, Высшая школа Hanyang University, Корея

Koi567@korea.kr

Аннотация. Хотя институциональная система архивного дела в Республике Корея была заложена с принятием в 1999 году Закона об управлении государственными документами, специализированная система управления документами президентов была сформирована после принятия в 2007 году Закона об управлении президентскими документами и создания в 2008 году Архива Президента. На Архив Президента возложена задача по сохранению документального наследия президентов, избираемых на единственный пятилетний срок полномочий. Помимо официальных документов, передаваемых в установленном порядке, Архив систематически комплектуется ценными личными документами президентов. Эти коллекции, поступающие в виде даров, записей устной истории и материалов, приобретаемых за рубежом, восполняют существенные пробелы, остающиеся в официальной документации. Несмотря на существующие проблемы, связанные с ограниченным финансированием, управлением цифровыми документами, созданными в электронной среде, а также вопросами конфиденциальности при обеспечении открытого доступа, личные документы позволяют значительно глубже понять философию государственного управления и особенности принятия политических решений. В перспективе Архив Президента намерен сформировать сбалансированную политику раскрытия архивной информации и внедрить технологии искусственного интеллекта для автоматизации обработки документов и совершенствования семантического поиска для исследователей.

Ключевые слова: Архив Президента; личные документы; дарение; устная история

Introduction

Although the Government of the Republic of Korea was established in 1948, systematic public records management was not officially initiated until the enactment of the Public Records Management Act in 1999. Prior to this legislation, a general societal tendency to undervalue records prevailed; however, this attitude shifted rapidly following the passage of the Act. Public institutions began employing professional archivists to manage institutional memory, while citizens increasingly sought to hold the government accountable by examining historical records. Consequently, electronic records management systems (ERMS) were widely adopted across public sectors in the early 2000s. This technological shift, coupled with growing public interest, heightened the demand for the systematic management of presidential records.

2. Mission of the Presidential Archives

Following the passage of the Presidential Records Act in 2007, the Presidential Archives was officially established in 2008. To ensure that all records generated during the course of a presidency are properly produced, preserved, and managed, the Act clearly delineates the legal scope of presidential records. Under Article 2 of the Act, presidential records are defined as 'any records and artifacts produced or received by the President, presidential advisory bodies, security agencies, and the presidential transition committee in connection with the performance of official duties.' Furthermore, the law mandates the collection of 'private presidential records' of national preservation value, as well as 'records produced by former presidents and their respective originating agencies.'

The founding of the Presidential Archives served as a institutional turning point, offering a solution to remedy the 40-year blank period in the nation's presidential record-keeping.

Currently operating under the Ministry of the Interior and Safety, the Presidential Archives receives and manages transferred records from presidential staff, security, and advisory bodies. The core functions of the Presidential Archives are categorized into three primary mandates:

First, the collection and transfer of presidential records;

Second, the establishment of a descriptive and classification system for presidential records

Third, the development of infrastructure for the preservation and restoration of presidential records to ultimately facilitate public access and services.

The most critical mission of the Presidential Archives involves the acquisition, preservation, and management of records upon the expiration of a president's term. South Korean presidents serve a single five-year term, and the law mandates that all presidential records be transferred to the Archives sequentially, beginning one year prior to the end of the tenure and concluding at the official termination of the term. Since 2008, the Archives has acquired and managed the records of five presidencies, currently housing approximately 50 million items of records. Table1 shows the quantity of records by type at the Presidential Archives:

Total	Non-electronic Documents	Electronic Document	Audiovisual Materials	Artifacts	Publications	Datasets, Web Records, etc.
50,150,096	1,992,112	3,240,685	7,489,066	43,334	34,342	37,350,557

***Table 1**. Quantity of Records by Type at the Presidential Archives*

3. Collection of Private Presidential Records: Donations

To date, the Presidential Archives has acquired private records primarily through two approaches: securing privately held documents from the civilian sector and conducting oral history interviews with individuals associated with past administrations.

The former method involves receiving donations from individuals or private organizations that possess president-related materials. The scope of this collection encompasses:

First, the private records of former presidents;

Second, historically significant records produced by presidential record-generating agencies prior to the enforcement of the Public Records Management Act;

Third, records held by individuals or organizations regarding the President's execution of duties, major historic events, or political and personal activities.

These materials are typically acquired by transferring legal ownership through formal donation agreements signed with the respective owners or organizations.

However, identifying records held by private individuals or organizations posed significant challenges. To address this issue, the Archives initiated intensive research

into the memoirs and publications of key figures – such as former senior presidential secretaries and cabinet ministers – to evaluate the likelihood of private holdings. Concurrently, the Archives launched public donation campaigns targeting citizens whose holdings were otherwise difficult to trace. This strategy aimed to incentivize voluntary public participation, thereby overcoming the limitations of relying solely on written and face-to-face surveys of key figures. Advertisements were placed in major daily and weekly newspapers, alongside online promotions conducted on the websites of prominent administrative agencies. Through these multifaceted initiatives, public awareness regarding the importance of archiving and interest in presidential records increased substantially.

Conversely, items were excluded from collection if their acquisition background, provenance, or ownership was ambiguous, or if the materials were deemed to lack national preservation value. Furthermore, to rigorously evaluate the historical value of materials, a Collection Review Committee composed of domain experts was convened to selectively acquire only high-value records. The Number of Donated Records by Presidential Terms is delineated in Table 2:

Presiden t	**1st~ 3rd**	**4th**	**5th~ 9th**	**10th**	**11th~ 12th**	**13th**	**14th**	**15th**	**16th**	**17th**	**18th**	**19th**
Holdings (Items)	21	-	4,214	29,977	44,031	7,853	49,826	2,996	4,338	19	4	9

***Table 2**. Number of Donated Records by Presidential Terms*

Major examples of donated records include documents directly contributed by former presidents, personal photographs, artifacts, and attire worn at official state events. Donations also include correspondence exchanged between renowned Korean scientists and the contemporary presidents, as well as professional journals maintained by administrative officials during their service at the Office of the President.

Importantly, some of these donated materials contained vital records regarding major national policies and pivotal events that had not been preserved in official public documents. By acquiring these records, the Archives has successfully begun filling the historical gap left prior to the enactment of the Act on the Management of Presidential Records. Moreover, by receiving personal holdings and mementos from former presidents and their bereaved families, the Archives secured critical primary source materials not only for researching presidential governance but also for life-history studies covering pre- and post-presidential periods. Furthermore, valuable artifacts reflecting diverse facets of the presidents—such as calligraphy and personal letters handwritten by former presidents, and photographs taken with ordinary citizens who shared distinct connections with them—have now been brought under systematic institutional management.

To incentivize donations and facilitate donor convenience, the Archives supports the preliminary processing of donated materials and provides high-fidelity replicas upon the donor's request. Additionally, a "Donors' Wall" was installed in the exhibition hall to honor the public-spirited intentions of contributors. The primary categories of

the donated records are summarized in Table 3:

Donor	Major Donation Items
Former Presidents	Documents from before and after the presidential and ministerial terms
	Presidential remarks and directives on briefing reports
	Event photographs and artifacts from the presidential term
Presidential Associates	Documents related to state administration
	Handwritten letters and booklets by the President
	Records of Cheong Wa Dae (Blue House) events during the presidential term
	Photographs of the President and the First Lady
	Local newspapers from the President's overseas visits
	Professional journals (work notebooks) of secretaries
General public	presidential calligraphy, photographs, videos, stamps, commemorative coins, books etc.

Table 3. *Major Types of Donated Presidential Materials*

Similar to officially transferred public records, these collected private donations are processed, registered, and preserved as permanent archival assets. When individuals or institutions file formal public information request access, the disclosable records are made accessible to the general public, thereby ensuring transparency and academic utility.

4. Collection of Private Presidential Records: Oral History

Another vital approach to securing presidential historical materials is oral history recording. This initiative was originally conceived to compensate for systemic gap within the Archives' official public holdings. Because the oral history process successfully unearths the nuanced context of historical events and brings to light facts absent from official documentation, the project has continued to this day.

Focusing primarily on high-ranking officials who served in presidential staff and advisory bodies across various administrations, the Archives recorded approximately 1,700 hours of oral histories from about 300 individuals between 2008 and 2025. This repository also encompasses select oral testimonies from former presidents, their families, and close acquaintances. Consequently, the Archives has been able to permanently preserve not only the public performance of the President as the chief executive but also their underlying life philosophies and human dimensions.

While interviews with officials from presidential staff agencies primarily focused on policy implementation throughout a president’s tenure, interviews from advisory body members concentrated on presidential philosophies and agenda-setting during the early stages of the presidency. The table 4 shows the volume of oral history records by president:

President	Total	1st~ 3rd	4th	5th~ 9th	10th	11th~ 12th	13th	14th	15th	16th	17th	18th
Interviewees	298	14	17	61	7	37	27	26	31	41	21	16
Hours	1,757	76	68.5	296	42.5	223	164	225	162.5	218	167.5	114

Table 4. *Volume of the Oral History Collection (categorized by the presidential term)*

As detailed above, oral histories from individuals associated with the presidency encompass comprehensive content that reveals not only official presidential governance but also the president's underlying philosophies and personal attributes. These materials will serve as crucial primary records for future presidential studies, adjacent academic disciplines, and broader research into contemporary society and culture.

5. Collection of Private Presidential Records: Acquisition of Overseas Records

In addition to donations and oral history recordings, the Presidential Archives has systematically collected president-related records held by overseas institutions to supplement its existing repository. The primary target nations include the United States (specifically the National Archives and Records Administration [NARA]), the United Kingdom, Australia, New Zealand, and France. The collection process generally entailed surveying comprehensive lists of Korean presidential records maintained by these foreign archival repositories. The records gathered thus far have been processed for public access, and finding aids have been compiled and distributed or made accessible online as digital e-books.

However, the acquisition of overseas records requires substantial time and budget, as it necessitates appointing locally residing experts as field investigators to survey holdings within foreign repositories. Consequently, the overseas record collection initiative is currently on hold. Once location information for significant records is systematically verified in the future, the Archives plans to resume the collection of presidential records held overseas. The major overseas institution for collected presidential records are summarized in Table 5:

Country	Institution
United States of America	National Archives and Records Administration Lyndon B. Johnson Presidential Library Jimmy Carter Presidential Library Ronald Reagan Presidential Library George H.W. Bush Presidential Library William J Clinton Presidential Library The George W. Bush Presidential Library
United Kingdom	The National Archives
Australia	National Archives
New Zealand	National Archives

Vietnam	National Archives
Taiwan	National Archives, Chinese Culture University

Table 5. *Major Overseas Institutions for collected Presidential Records*

Because the current personnel and budget of the Presidential Archives are heavily concentrated on officially transferred public records, tasks related to collecting private presidential materials tend to be relatively circumscribed. In particular, as diverse formats of born-digital records spanning from the past to the present (e.g., datasets, web records, etc.) are transferred, managing the entire volume of these accumulated collections requires substantial physical and human resources. Inevitably, this operational demand has led to a reduction in the administrative support allocated for private records collection.

Nevertheless, it is clear that a high proportion of private records possess substantial archival value, despite the small size of the collection. Consequently, these collected records offer critical interpretive clues for comprehensively understanding the history of a specific presidential term. This underscores precisely why the Archives continues to prioritize the collecting of private materials alongside official public records.

6. Utilization of Private Records and Future Challenges

A major impediment to utilizing presidential private records is that they frequently contain a high volume of personally identifiable information (PII), complicating public disclosure. Accordingly, donors often tend to request extended closure periods for their donated records. A similar pattern emerges in oral history records. Many interviewees express concern over potential defamation of third parties and are reluctant to compromise the public reputation they established during their careers through unfiltered oral records. Consequently, a significant number of interviewees request that their records be released exclusively posthumously.

Nonetheless, the Archives clearly recognizes its institutional mandate to establish reasonable evaluative criteria and disclose records at appropriate times. Therefore, the Archives plans to review disclosure protocols starting with records that carry a lower risk of personal defamation, thereby establishing a systematic and balanced disclosure policy for oral history records.

The institutional significance of private records within the Presidential Archives is as suggested above. Although presidents perform public duties, uncovering the core values and philosophies that shaped them, identifying the influential stakeholder groups, and analyzing how these variables manifested in policy remain an intriguing and meaningful scholarly endeavor. Furthermore, formal official documents often fail to capture the underlying, latent dimensions of major policies, making it difficult to comprehend a given historical issue holistically. This reality likely explains why not only the Presidential Archives of Korea but also other nations operating independent presidential libraries continue to aggressively collect private records and preserve even personal correspondences.

Currently, the Archives is exploring methodologies to efficiently utilize emerging technologies, such as Artificial Intelligence (AI), in records management. While AI and

decade-old private presidential records might initially appear to be an incongruous pairing, effectively leveraging AI can transform private records into a particularly potent archival asset. For instance, if advanced technology can extract meaningful semantic keywords from records to seamlessly map connections across all institutional holdings, it will highly streamline the integration of public and private records. Moreover, it could significantly shorten the complex processing stages required to service oral history records. Presidential private record content reconstructed through these technological frameworks will guide researchers into historical contexts with far more vivid and multifaceted narratives.

References

Act on the Management of Presidential Records, Law No. 8395 (Enacted on April 27, 2007).

Guidelines for the Collection and Management of Oral Histroy Records at the Presidential Archives, Guideline No. 10 (Enacted on November 9, 2012).

Imjeong Ko, et al. Records Transfer of the World's Highest Executive Agencies: Cases and Issues. 2020. Seonin.

Korea Legislation Research Institute. A Study on the Management System of Presidential Records. (Presidential Archives Commissioned Research Final Report) 2015.

National Archives of Korea. 40-Year History of the National Archives of Korea. 2009.

National Archives of Korea. http://www.archives.go.kr (accessed 5.15.2026).

Presidential Archives of Korea http://www.pa.go.kr (accessed 5.15.2026).

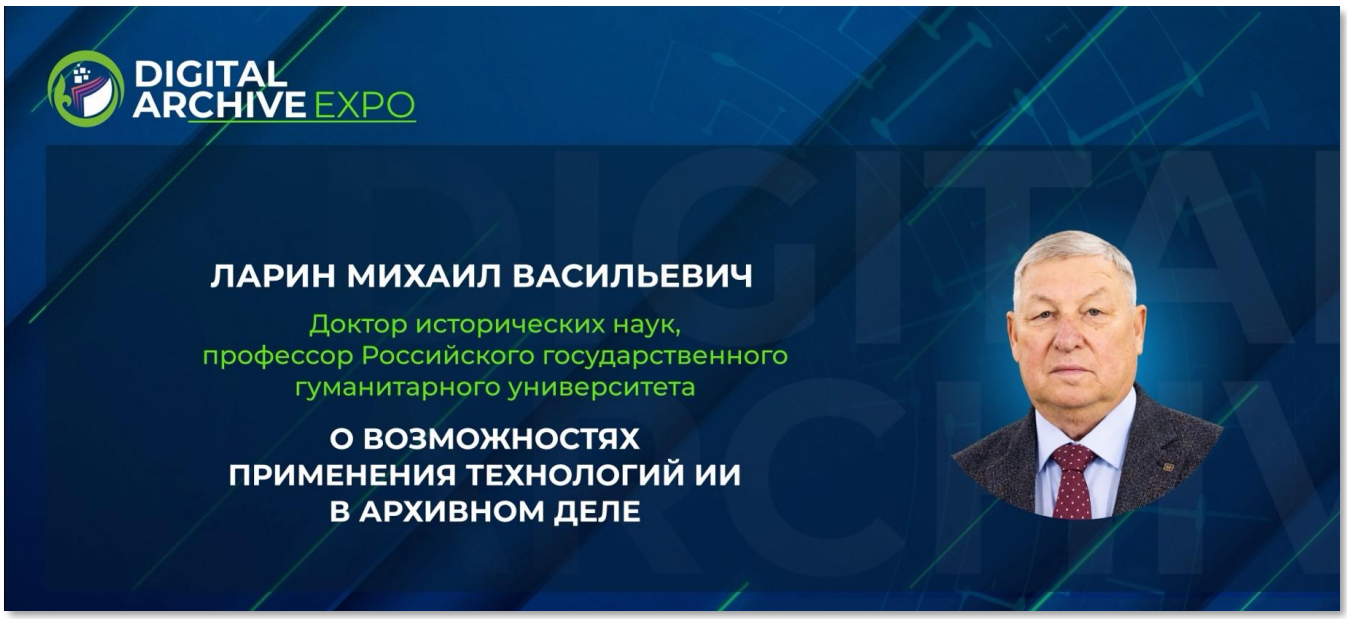
Public Agencies Records Management Act, Law No. 5709 (Enacted on January 29, 1999).

Regulations on the Collection of Presidential Records, Guideline No. 5 (Enacted on August 19, 2008).

«АРХИВТЕР ЖӘНЕ AI: ҚҰРАЛДАРЫ, ЭТИКА МЕН БОЛАШАҚ СТРАТЕГИЯЛАРЫ»
№ 1 сессия

Сессия № 1
«АРХИВЫ И AI: ИНСТРУМЕНТЫ, ЭТИКА И СТРАТЕГИИ БУДУЩЕГО»

«ARCHIVES AND AI:
TOOLS, ETHICS AND FUTURE STRATEGIES»
Session No. 1



О ВОЗМОЖНОСТЯХ ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В АРХИВНОМ ДЕЛЕ

Михаил Васильевич Ларин
Доктор исторических наук, профессор, заведующий кафедрой автоматизированных систем документационного обеспечения управления, Российский государственный гуманитарный университет, Россия
Larin.mv@rggu.ru

Аннотация. В стать рассматриваются возможности применения технологий искусственного интеллекта в архивном деле на современном этапе. Отмечается, что переход от классических экспертных систем к машинному обучению и большим языковым моделям открывает новые перспективы для распознавания архивных документов, их семантического анализа, автоматического извлечения метаданных, осуществления интеллектуального поиска и создания специализированных информационных систем и цифровых архивных помощников. Анализируются реализованные российские проекты, которые демонстрируют возможности применения технологий искусственного интеллекта при обработке, хранении, поиске и использовании архивных документов. Рассматриваются вопросы подготовки архивистов, связанные с формированием компетенций в сфере искусственного интеллекта. Отмечается, что успешное внедрение технологий искусственного интеллекта в архивное дело невозможно без междисциплинарной подготовки специалистов нового поколения. Подчеркивается, что использование технологий искусственного интеллекта в архивном деле сопряжено с рисками, которые включают утрату контекста, генерацию недостоверной информации, отсутствие наборов данных для создания специализированных моделей искусственного интеллекта. Делается вывод, что на современном этапе технологии искусственного интеллекта не заменяют архивиста, а служат инструментом расширения его профессиональных возможностей.

Ключевые слова: технологии искусственного интеллекта; большие языковые модели; архивное дело

АРХИВ ІСІНДЕ ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫН ҚОЛДАНУ МҮМКІНДІКТЕРІ ТУРАЛЫ

Михаил Васильевич Ларин
Тарих ғылымдарының докторы, профессор, құжаттамалық қамтамасыз етуді басқарудың автоматтандырылған жүйелері кафедрасының меңгерушісі,
Ресей мемлекеттік гуманитарлық университеті, Ресей
Larin.mv@rggu.ru

Аңдатпа. Мақалада қазіргі кезеңдегі архив ісінде жасанды интеллект (ЖИ) технологияларын қолдану мүмкіндіктері қарастырылады. Дәстүрлі сараптамалық жүйелерден машиналық оқыту мен ірі тілдік модельдерге көшу архивтік құжаттарды тану, оларды семантикалық талдау, метадеректерді автоматты түрде алу, интеллектуалды іздеуді жүзеге асыру, сондай-ақ мамандандырылған ақпараттық жүйелер мен цифрлық архивтік көмекшілерді әзірлеу бағытында жаңа мүмкіндіктерге жол ашатыны көрсетіледі. Архивтік құжаттарды өңдеу, сақтау, іздеу және пайдалану үдерістерінде жасанды интеллект технологияларын қолданудың тиімділігін айқындайтын Ресейде іске асырылған жобаларға талдау жасалады. Сонымен қатар архившілерді даярлау мәселелері қарастырылып, жасанды интеллект саласындағы кәсіби құзыреттерді қалыптастырудың маңыздылығы негізделеді. Архив ісіне жасанды интеллект технологияларын табысты енгізу жаңа буын мамандарын пәнаралық негізде даярлаумен тікелей байланысты екендігі атап көрсетіледі. Сонымен бірге архив ісінде жасанды интеллект технологияларын қолдану контекстің жоғалуы, сенімсіз ақпараттың генерациялануы, сондай-ақ мамандандырылған жасанды интеллект модельдерін әзірлеуге қажетті деректер жиынтықтарының жеткіліксіздігі сияқты тәуекелдермен ұштасатыны көрсетіледі. Зерттеу нәтижесінде қазіргі кезеңде жасанды интеллект технологиялары архившінің кәсіби қызметін алмастырмайтыны, керісінше оның кәсіби мүмкіндіктерін кеңейтетін тиімді құрал ретінде қызмет ететіні туралы қорытынды жасалады.

Түйін сөздер: жасанды интеллект технологиялары; ірі тілдік модельдер; архив ісі

ON THE POSSIBILITIES OF APPLYING ARTIFICIAL INTELLIGENCE TECHNOLOGIES IN ARCHIVAL SCIENCE

Mikhail V. Larin
Dr. (in history), professor, head of the Department of Automated Systems of Document Support for Management,
Russian State University for the Humanities, Russia
Larin.mv@rggu.ru

Abstract. The article examines the possibilities of applying artificial intelligence technologies in archival science at the current stage. It is noted that the transition from classical expert systems to machine learning and large language models opens up new prospects for the recognition of archival documents, their semantic analysis, automatic metadata extraction, intelligent search, and the creation of specialized information systems and digital archival assistants. An analysis is provided of implemented Russian projects that demonstrate the potential of using artificial intelligence technologies in the processing, storage, retrieval, and use of archival documents. The article also addresses issues of training

archivists related to the development of competencies in the field of artificial intelligence. It is emphasized that the successful implementation of artificial intelligence technologies in archival science is impossible without interdisciplinary training of a new generation of specialists. It is further stressed that the use of artificial intelligence technologies in archival science involves risks, including loss of context, generation of unreliable information, and the lack of datasets for creating specialized artificial intelligence models. The conclusion is drawn that at the current stage, artificial intelligence technologies do not replace the archivist but rather serve as a tool for expanding the professional capabilities of the specialist.

Keywords: Artificial intelligence technologies; large language models; archival science

Современный этап цифровой трансформации характеризуется активным внедрением технологий искусственного интеллекта практически во все сферы профессиональной деятельности. Архивное дело в данном случае не является исключением. Более того, именно архивная отрасль в силу своей специфики, а именно работы с большими массивами документов, метаданными, научно-справочным аппаратом, становится одной из наиболее перспективных сфер применения технологий искусственного интеллекта.

На рубеже веков начался переход к новой методике построения интеллектуальных систем. С учетом накопленного ранее опыта системы искусственного интеллекта строились преимущественно на символьном подходе и функционировали на основе заранее заданных правил, логического вывода и инженерии знаний. Именно таким образом создавались первые экспертные системы. Однако в дальнейшем происходит переход к системам машинного обучения, способным самостоятельно выявлять закономерности в данных и адаптироваться в процессе работы. Этот переход существенно расширил возможности применения технологий искусственного интеллекта в сфере обработки документов и информации.

Особое влияние на распространение технологий искусственного интеллекта на современном этапе оказало появление больших языковых моделей (large language model, LLM), например, ChatGPT, DeepSeek и др. Именно большие языковые модели стали наиболее заметным технологическим проявлением искусственного интеллекта в современных условиях. Для архивного дела это имеет принципиальное значение, поскольку архивы работают прежде всего с текстами, документами, историческим контекстом. Соответственно развитие больших языковых моделей создало возможности для осуществления интеллектуального поиска, автоматического составления описаний, генерации справочной информации, семантического анализа архивных документов и создания цифровых архивных помощников.

Современные исследования возможностей применения технологий искусственного интеллекта в архивном деле посвящены обеспечению сохранности электронных документов, совершенствованию алгоритмов поиска, вопросам экспертизы ценности, развитию научно-справочного аппарата в части составления внутренних описей дел, указателей, организации поиска на основе семантического анализа, извлечения метаданных (Кюнг 2022: 39). Существенное внимание уделяется вопросам семантического анализа архивных документов, поскольку именно семантическая обработка информации позволяет перейти от поиска по реквизитам к интеллектуальному поиску по смыслу и контексту.

Одним из наиболее значимых направлений является оцифровка архивных документов. Современная оцифровка представляет собой не просто создание электронных копий документов. Ключевое значение имеет перевод архивных документов в машиночитаемый вид. Именно здесь технологии искусственного интеллекта демонстрируют наибольшую эффективность. Они применяются для OCR (Optical Character Recognition) и HTR (Handwritten Recognition) распознавания, автоматического извлечения метаданных, анализа структуры документов и формирования полнотекстового поиска (Боброва 2023: 72).

Однако архивная сфера предъявляет особые требования к таким инструментам. Архивные документы отличаются неоднородностью, сложностью структуры, наличием рукописных текстов, специфической лексикой и орфографией, дефектами носителей и утратой части информации. Соответственно эффективность интеллектуальных систем напрямую зависит от качества обучающих наборов данных. В научных исследованиях справедливо подчеркивается, что отсутствие специализированных наборов данных для архивных документов способно привести к воспроизводству ошибок, потере контекста и искажению информации (Юмашева 2025: 104)/

Следовательно, внедрение технологий искусственного интеллекта в архивное дело требует разработки полноценной методической базы, включающей критерии качества, процедуры верификации, требования к обучающим выборкам, системы протоколирования действий моделей, а также оценку типовых ошибок и так называемых «галлюцинаций» больших языковых моделей.

На современном этапе развитие искусственного интеллекта в архивном деле характеризуется переходом от преимущественно теоретических исследований к прикладным разработкам, что находит непосредственное

отражение в архивном деле и демонстрирует возможности использования интеллектуальных технологий при обработке, хранении, поиске и использовании архивных документов. Именно поэтому особый интерес представляет анализ уже реализованных проектов.

Важным направлением является создание интеллектуальных поисковых систем. Наиболее показательным примером выступает семантическая поисковая система НИКА, внедренная в Государственном архиве Российской Федерации[1*]. Система использует технологии семантического анализа и российские языковые модели для обработки пользовательских запросов. Важно подчеркнуть, что речь идет уже не о традиционном поиске по реквизитам, а о поиске по смыслу и контексту. Система позволяет анализировать сложные запросы, формировать отчеты, объяснять результаты поиска и работать с ассоциативными связями между документами.

Другим важным направлением является создание специализированных архивных информационных систем. Показательным примером выступает система «Находка-Сфера», внедренная в Государственном архиве Тульской области[2*]. Система включает модули распознавания текста, интеллектуальной классификации, контекстной обработки запросов и электронного читального зала.

Подобные решения демонстрируют качественно новый уровень работы с архивной информацией, поскольку интеллектуальные системы начинают выполнять функции анализа и интерпретации данных.

[1*] Справка проекта «Семантическая поисковая система НИКА» // Государственный архив Российской Федерации. https://garf-nika.ru (дата обращения: 22.05.2026).

[2*] Информационная система «Находка-Сфера» // Реестр российского программного обеспечения. https://reestr.digital.gov.ru/reestr/366570/ (дата обращения: 22.05.2026).

Вместе с этим развивается направление роботизированных архивных комплексов. В 2023 году в Москве был открыт роботизированный архивный комплекс «Мосархив»[3*]. В данном случае технологии искусственного интеллекта используются не только для обработки информации, но и для автоматизации логистических процессов хранения и выдачи документов. Таким образом, искусственный интеллект начинает охватывать не только информационную, но и организационно-технологическую плоскость архивного дела.

Отдельный интерес представляет применение технологий искусственного интеллекта в экспертизе ценности документов. В данном случае искусственный интеллект рассматривается как интеллектуальный помощник архивиста. Существенное значение имеет эксперимент, проведенный Всероссийским научно-исследовательским институтом документоведения и архивного дела (ВНИИДАД) по использованию больших языковых моделей для предварительной оценки архивной ценности документов (Белов 2025: 1190). Результаты исследования показали высокую зависимость качества решений от полноты обучающих данных, качества разметки, разнообразия примеров и участия специалистов-архивистов. Таким образом, современные технологии пока не заменяют профессиональную экспертизу, а выступают инструментом поддержки принятия решений.

Особый интерес также представляет использование больших языковых моделей в архивном деле. В Республике Татарстан введен в эксплуатацию сервис «ГосПромт», ориентированный на безопасное использование генеративного искусственного интеллекта в государственном управлении[1*]. Сервис уже используется в Государственном комитете Республики Татарстан по архивному делу. Для архивной отрасли подобные решения открывают возможности автоматического составления описаний, подготовки справок, генерации поисковых подсказок и создания цифровых архивных помощников. При этом принципиально важно, чтобы подобные системы функционировали с учетом требований информационной безопасности и особенностей архивных данных.

Несмотря на значительный потенциал технологий искусственного интеллекта, архивная сфера также сталкивается с рядом серьезных методологических вызовов. Во-первых, существует риск утраты контекста архивных документов при автоматизированной обработке. Во-вторых, сохраняется проблема генерации недостоверной информации. В-третьих, возникает необходимость нормативного регулирования применения интеллектуальных систем в архивном деле. Также архивисты зачастую не обладают необходимыми компетенциями для эффективного применения технологий искусственного интеллекта в своей деятельности.

Следовательно, внедрение технологий искусственного интеллекта невозможно без подготовки специалистов нового поколения. Сегодня одним из ведущих центров подготовки документоведов и архивистов в России является Российский государственный гуманитарный университет (РГГУ). Особую роль

[3*] Главархив Москвы (ГБУ «Мосархив»). https://mosarchive.mos.ru/ (дата обращения: 22.05.2026).
[1*] В Татарстане запустили цифрового помощника в сфере госуправления. https://xn--80aapampemcchfmo7a3c9ehj.xn--p1ai/news/v-tatarstane-zapustili-tsifrovogo-pomoshchnika-v-sfere-gosupravleniya/ (дата обращения: 22.05.2026).

играет кафедра автоматизированных систем документационного обеспечения управления Историко-архивного института РГГУ, реализующая профиль «Интеллектуальные системы в управлении документами». Подготовка студентов включает изучение дисциплин документоведения и архивоведения, основ программирования, машинного обучения, обработки данных и проектирования интеллектуальных систем (Ларин 2024: 60). Таким образом, формируется междисциплинарная модель подготовки специалистов, способных работать в условиях цифровой трансформации архивного дела.

Подводя итог, можно сделать вывод, что современный этап развития архивного дела характеризуется активной интеллектуализацией процессов работы с документами и информацией. Технологии искусственного интеллекта уже применяются в оцифровке архивных документов, интеллектуальном поиске, экспертизе ценности, создании цифровых помощников, разработке роботизированных архивных комплексов. Все это невозможно без анализа научного наследия в области искусственного интеллекта, преемственности научных исследований, понимания взаимосвязи этого научного направления с управлением документами.

Однако дальнейшее развитие применения технологий искусственного интеллекта в архивном деле требует методической проработки, нормативного регулирования, формирования специализированных наборов данных и подготовки кадров нового поколения. На современном этапе технологии искусственного интеллекта следует рассматривать не как замену архивиста, а как инструмент расширения профессиональных возможностей специалиста. Именно сочетание профессиональной экспертизы архивиста и интеллектуальных технологий позволит обеспечить качественное развитие архивного дела в условиях цифровой трансформации, полноценно раскрыть возможности технологий искусственного интеллекта в этой сфере деятельности.

Литература

Белов 2025 – Белов И.И. Применение больших языковых моделей в экспертизе ценности документов. 2018–2025 гг. // Вестник архивиста. 2025. № 4. С. 1187-1201.

Боброва 2023 – Боброва Е.В. Перспективы применения технологии искусственного интеллекта в архивном деле: анализ опыта // Вестник ВНИИДАД. № 1. 2023. С. 67-78.

Кюнг 2022 – Кюнг П.А. Новации в системе научно-справочного аппарата государственных и муниципальных архивов // Вестник ВНИИДАД. 2022. № 3. С. 27-43.

Ларин 2024 – Ларин М.В. Новая образовательная программа «Интеллектуальные системы в управлении документами» // Управление документацией: прошлое, настоящее, будущее: материалы VII Международной научно-практической конференции, посвященной 90-летию со дня рождения профессора Т.В. Кузнецовой (23-24 марта 2023 г.). Москва: РГГУ 2024. С. 59-64.

Юмашева 2025 – Юмашева Ю.Ю. К вопросу о применении искусственного интеллекта в исторических исследованиях // Историческая информатика. 2025. № 1 (51). С. 95-121.

WHAT IS ARTIFICIAL INTELLIGENCE? WHY DOES IT MATTER FOR THE FUTURE OF DOCUMENTS AND ARCHIVES?

Sun Park
University College Dublin, Ad Astra Fellow / Assistant Professor in AI and Digital Cultural Heritage School of Information and Communication Studies, Ireland
sun.park@ucd.ie; https://orcid.org/0000-0001-9307-1171

Abstract. Artificial Intelligence (AI) is increasingly reshaping how documents are created, managed, preserved, searched, interpreted, and trusted. While archives have tended to remain discussed as sources for input data to train AI systems, this paper posits active roles of archivists and documentation professionals in affecting the entire AI lifecycle applied to a broader cultural context. This paper examines the meaning of AI, its relevance to digital document management and its implications for professional ethics of archivists. It argues that the transparency archivists owe their users cannot be satisfied by technological explainability alone. Drawing on the framework of cultural black boxes and cultural explainability, the paper contends that the most consequential opacity in AI-based archival work is not only algorithmic but cultural: the unexamined human decisions about what is collected, what is treated as representative, what is allowed to stand for an absent record, and whose contribution is acknowledged.

The paper recasts the risks of AI – fabrication, synthetic falsification, automated bias, and institutional dependence – as cultural black boxes and proposes a set of proactive roles through which archivists can implement cultural explainability of AI as curators of culturally core data, as stewards of representativeness of communities' records concern.

Keywords: Artificial Intelligence; digital archives; archival ethics; cultural black boxes; cultural explainability

ЖАСАНДЫ ИНТЕЛЛЕКТ ДЕГЕНІМІЗ НЕ? ОНЫҢ ҚҰЖАТТАР МЕН АРХИВТЕРДІҢ БОЛАШАҒЫ ҮШІН МАҢЫЗЫ ҚАНДАЙ?

Сан Парк

ЖИ және цифрлық мәдени мұра саласындағы Ad Astra бағдарламасының стипендиаты / ассистент-профессоры, Ақпараттық және коммуникациялық зерттеулер мектебі, Дублин университеттік колледжі, Ирландия
sun.park@ucd.ie; https://orcid.org/0000-0001-9307-1171

Аңдатпа. Жасанды интеллект (ЖИ) технологиялары құжаттарды жасау, басқару, сақтау, іздеу, интерпретациялау және олардың сенімділігін бағалау тәсілдерін түбегейлі өзгертуде. Архивтер көбіне жасанды интеллект жүйелерін оқытуға арналған бастапқы деректер көзі ретінде ғана қарастырылып келгенімен, бұл мақалада архившілер мен құжаттаманы басқару саласының мамандарының кең мәдени контексте қолданылатын жасанды интеллекттің бүкіл өмірлік циклiндегі белсенді рөлі негізделеді. Зерттеуде жасанды интеллект ұғымының мәні, оның цифрлық құжаттарды басқару жүйелері үшін маңызы және архившілердің кәсіби этикасына ықпалы қарастырылады. Автор архив пайдаланушылары алдындағы ашықтық қағидатын қамтамасыз ету тек технологиялық түсіндірмелілікпен шектелмеуі тиіс екенін дәлелдейді. Мәдени «қара жәшіктер» және мәдени түсіндірмелілік тұжырымдамаларына сүйене отырып, мақалада жасанды интеллект негізіндегі архивтік қызметтегі негізгі жабықтық алгоритмдердің өзінен гөрі мәдени сипаттағы факторлармен байланысты екендігі көрсетіледі. Атап айтқанда, қандай материалдардың жинақталатыны, қай деректердің репрезентативті деп танылатыны, қандай құжаттардың жоқ дереккөздердің орнына пайдаланылатыны және кімнің үлесі мойындалатыны жөніндегі жеткілікті деңгейде зерделенбеген адам қабылдайтын шешімдер ерекше маңызға ие.

Жасанды интеллектті қолданумен байланысты тәуекелдер, оның ішінде жалған ақпаратты генерациялау, синтетикалық бұрмалау, автоматтандырылған алгоритмдік бейтарапсыздық және институционалдық тәуелділік мәдени «қара жәшіктер» ретінде қарастырылады. Осыған байланысты архившілердің мәдени тұрғыдан маңызды деректердің кураторлары әрі қауымдастықтардың құжаттық мұрасының репрезентативтілігін қамтамасыз ететін кәсіби мамандар ретіндегі қызметі арқылы жасанды интеллекттің мәдени түсіндірмелілігін іске асыруға мүмкіндік беретін проактивті тәсілдер ұсынылады.

Түйін сөздер: жасанды интеллект; цифрлық архивтер; архивтік этика; мәдени қара жәшіктер; мәдени түсіндірмелілік

ЧТО ТАКОЕ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ? ПОЧЕМУ ОН ВАЖЕН ДЛЯ БУДУЩЕГО ДОКУМЕНТОВ И АРХИВОВ?

Сан Парк

Стипендиат программы Ad Astra / ассистент-профессор в области ИИ и цифрового культурного наследия, Школа информационных и коммуникационных исследований, Университетский колледж Дублина, Ирландия
sun.park@ucd.ie; https://orcid.org/0000-0001-9307-1171

Аннотация. Искусственный интеллект (ИИ) все в большей степени трансформирует процессы создания, управления, хранения, поиска, интерпретации

документов и формирования доверия к ним. Если ранее архивы преимущественно рассматривались как источники данных для обучения систем искусственного интеллекта, то в настоящей статье обосновывается активная роль архивистов и специалистов в области документационного обеспечения в управлении всем жизненным циклом технологий искусственного интеллекта в более широком культурном контексте. Рассматриваются сущность искусственного интеллекта, его значение для управления цифровыми документами и влияние на профессиональную этику архивистов. Обосновывается, что требование прозрачности, которое архивисты должны обеспечивать в отношении пользователей архивной информации, не может быть реализовано исключительно посредством технологической объяснимости алгоритмов. Опираясь на концепции культурных «черных ящиков» и культурной объяснимости, автор утверждает, что наиболее существенная непрозрачность в архивной деятельности, основанной на использовании искусственного интеллекта, имеет не только алгоритмическую, но и культурную природу. Она проявляется в неосмысленных человеческих решениях относительно того, какие материалы подлежат отбору, какие данные признаются репрезентативными, какие документы могут замещать отсутствующие источники и чей вклад получает официальное признание.

Риски применения искусственного интеллекта, включая фабрикацию информации, синтетическую фальсификацию, автоматизированную предвзятость и институциональную зависимость, рассматриваются как проявления культурных «черных ящиков». В статье предлагается комплекс проактивных профессиональных ролей архивистов, позволяющих обеспечивать культурную объяснимость искусственного интеллекта посредством их деятельности в качестве кураторов культурно значимых данных и хранителей репрезентативности документального наследия сообществ.

Ключевые слова: искусственный интеллект; цифровые архивы; архивная этика; культурные «черные ящики»; культурная объяснимость

Introduction

Archives are not merely repositories of old documents. They are institutions of memory, accountability, rights, and identity, helping people at present understand the past. The International Council on Archives describes its mission in relation to records, archives, and data in all formats, emphasising the importance of effective management and use across changing documentary environments (ICA, 2012). In this context, AI should be understood not as an external technological trend, but as a development that directly affects the archival mission. The emergence of AI has introduced a new set of questions about non-human-based story curation for archivists, documentation professionals and cultural heritage professionals. As AI becomes embedded in document production and archival processing, archivists must consider how it affects authenticity, provenance, appraisal, access, privacy and long-term preservation.

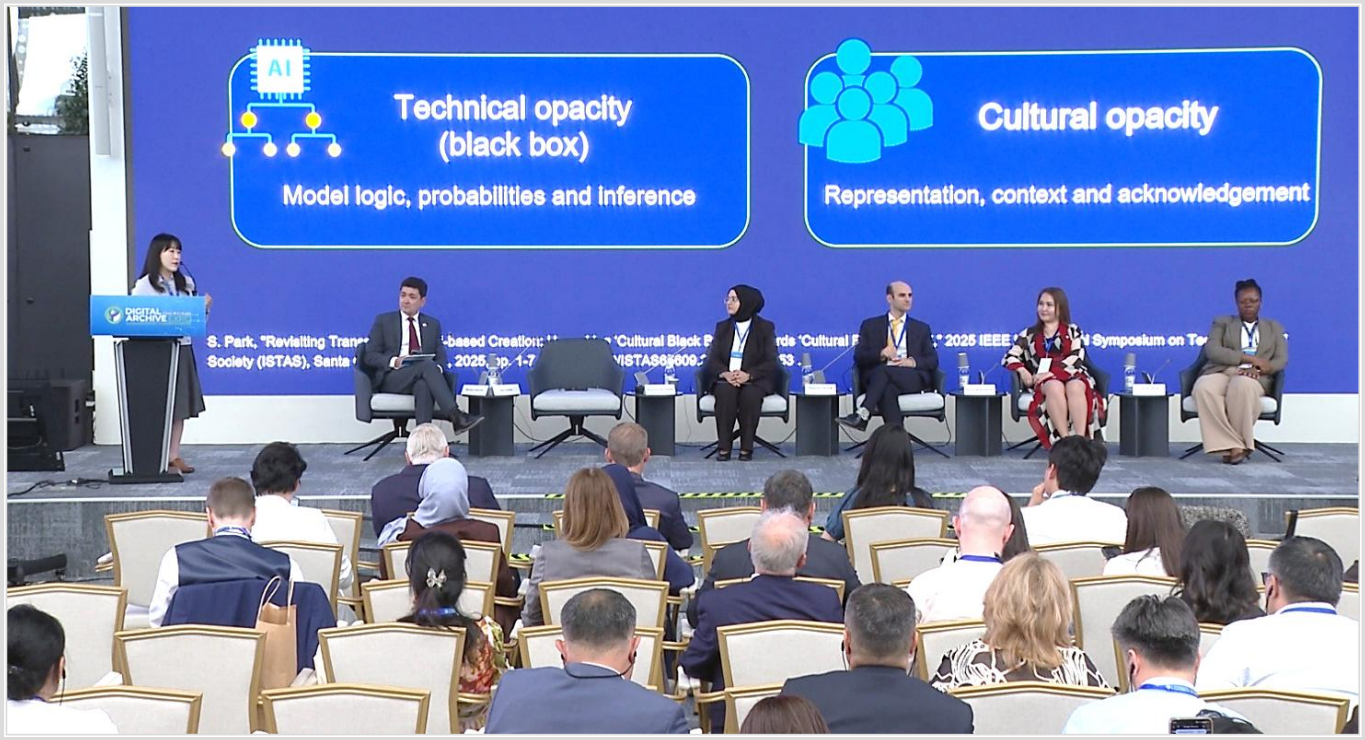
As the wider GLAM sector – galleries, libraries, archives, and museums – increasingly generate cultural content with AI rather than merely cataloguing it, a question arises that the profession cannot answer by technical means alone; when an archive uses such a system, how does it remain transparent not only about how the system works, but about why its inputs and outputs may be said to represent faithfully the people and pasts they depict? To address this question, this paper begins with summarising what AI is, along with its main functions used frequently in the archive field. Second, it discusses the black box problem with AI in relation to the archivist

profession and remit. Drawing on the concepts of cultural black boxes and cultural explainability (Park, 2025), it proposes a set of proactive roles through which the archivist profession can implement cultural explainability in interdisciplinary AI studies.

Understanding Artificial Intelligence in Documentation Context

AI is broadly defined as computational technology to perform intelligent tasks, such as pattern recognition, language processing, classification, prediction, problem-solving and content generation, which is often categorised into three tiers:

1. Artificial Intelligence: The overarching technology of algorithmic methodologies designed to process complex inputs and simulate intelligent analytical output.
2. Machine Learning: The statistical methodologies that train computational models on structured training datasets, enabling the software to dynamically establish underlying probabilistic patterns and generate predictions without explicit hard-coded instruction.
3. Deep Learning: Advanced deep multi-layered neural networks capable of evaluating highly unstructured data matrices.



In the conventional model of computing, a programmer specifies explicit rules: if a document bears a particular stamp, it is to be classified in a particular way. Machine learning inverts this relationship. Rather than being given rules, the system is given a large number of data and left to infer the underlying patterns itself. The systems that have lately attracted public attention, the so-called Large Language Models (LLMs), represent a particular instance of this approach. At their core they perform an operation that is, in principle, remarkably simple: having been trained upon an enormous body of text, they predict the most probable continuation of a given sequence of words. From

this elementary mechanism there emerge capabilities that appear far more sophisticated – the ability to summarise, to translate, to answer questions, and to generate description.

For records managers and archival scholars, an AI model can be best understood as a data processing system that excels at identifying latent correlations, parsing unstructured text, and executing classification tasks across multi-terabyte data repositories at speeds that manual processing cannot match. When trained or configured on large bodies of documents, they can learn to recognise recurring features explicitly indicated or implicitly embedded in documents, such as names, places, dates, topics, visual patterns, linguistic structures, or relationships between records. AI applications can help archivists identify patterns and priorities within massive collections. The United States National Archives and Records Administration has publicly identified AI use cases involving semantic search, metadata support, redaction assistance, and related archival and records functions (NARA, 2026). Such examples demonstrate that AI is already being explored by major archival institutions as a way to address the practical challenges of digital records management.

The main applications of AI in archives are summarised in Table 1. The first main application of AI in archives is automated processing for metadata extraction from unstructured mass documentation. In historical document management, deep neural network training has enabled remarkable advancements in Handwritten Text Recognition (HTR). Historical cursive, which previously required specialised expertise and hours of human transcription, can now be converted into machine-readable text. This systematically exposes hidden text layers to algorithmic index engines. Another AI application for documentation is to discover unindexed and unmonitored data. AI pipelines offer a viable solution through automated sensitivity scanning. By applying pattern-matching heuristics and semantic context analysis, machine learning models can audit large digital text repositories, flag potential privacy violations, and suggest automated, contextual redacting options. AI technologies are also used for semantic discovery and structural knowledge access. By leveraging natural language processing and vector embeddings, modern search interfaces can understand user queries based on conceptual intent rather than literal text matches.

Beyond the work of archivists and documentation managers, AI technology also impacts on public access to information. AI-enabled semantic search retrieves information based on meaning and context. A user searching for “displacement after flooding” may be directed to records using terms such as “evacuation,” “relocation,” “emergency housing,” or “temporary shelter.” This has significant implications for public access, especially in large digital collections. AI can also support multilingual access by translating archival descriptions or summarising documents in languages more accessible to users. It can generate plain-language summaries of complex administrative records, making archives more usable for students, journalists and non-specialist researchers. This does not eliminate the need for archivists, but it may change the way users encounter archival materials. The deeper question is therefore not simply whether AI can make archival processing faster. The more important question is whether AI would change the remit of archivists in the future.

Table 1. *Operational Comparison of Archival Methodologies*

Archival Dimension	Traditional Manual Practice	AI-Assisted Paradigm
Appraisal & Selection	Item-by-item manual review; highly subjective, slow, and prone to severe backlog accumulation.	Macro-appraisal via automated text classification and algorithmic pattern filtering across massive unstructured datasets.
Document Transcription	Manual transcription of historical manuscripts; high cost and low throughput.	Handwritten Text Recognition (HTR) utilizing deep learning models capable of cross-era reading.
Privacy & Security Auditing	Line-by-line visual auditing for Personally Identifiable Information (PII); high margin for human error.	Automated Named Entity Recognition (NER) capable of instantly scanning, flagging, and redacting private metadata.
Public Information Access	Strict reliance on rigid finding aids, hierarchical catalogues, and literal exact-keyword search.	Semantic and natural language cross- repository discovery; conceptual contextual matching.

Impacts of AI on archives

The functional advantages of AI are accompanied by technical and non-technical issues that demand careful archival oversight. Nearly every global ethical framework for AI affirms the principle of transparency (UNESCO, 2022; EU, 2024): those who deploy an AI system should be able to explain how it arrived at a conclusion. Within computer science this aspiration has produced the field of Explainable AI (XAI), whose techniques seek to open the so-called algorithmic black box - the deep models whose internal reasoning resists straightforward inspection. The transparency of the engineer is, in this sense, rooted in a technological contribution to the demonstration of algorithmic cause and effect.

Yet this conception of transparency, as Park (2025) observes, presumes that the opacity worth resolving is the opacity of AI systems, i.e., computation programmmes. For the cultural sector the technology-oriented assumption is insufficient. When an archive collaborates with technologists to generate cultural content - a reconstruction, a simulation, an enriched description - the archivist's distinctive responsibility is not to justify the algorithm but to justify why the data gathered and the model trained may be said to represent, with integrity, the human group or historical reality they concern. That justification resides in the negotiated, interpretive judgements of the people who built the project. The opacity of AI, in other words, is epistemic; different stakeholders confront different black boxes in the very same system, according to their knowledge and their role. I refer to such non-technological opacity of AI as cultural black box (Park, 2025), namely a set of opaque human decisions that researchers must uncover, so far as they are able, in order to legitimise the way they have designed and deployed a system to create cultural content. Where the technological black box concerns how the model computes, the cultural black box concerns what the people decided and

why. This paper argues that the archive's most pressing profession in relation to AI lies in revealing such cultural black boxes.

The first cultural black box concern in documentation and archives is issues of selection. AI systems used in archival environments depend on datasets. These datasets may include digitised records, catalogues, images, audiovisual records, oral histories, emails, administrative documents, or born-digital collections. However, archival datasets are never complete representations of the past. They reflect histories of survival, destruction, institutional power, colonial administration, bureaucratic practice, social inequality and archival appraisal. If AI systems are trained on archival data without explaining these political and social factors, they may treat historically partial collections as culturally complete datasets. For example, an AI model trained on colonial administrative records may learn a representation of a community shaped primarily by the language of colonial governance. A model trained on institutional records may marginalise community-based knowledge. A model trained on already well-described collections may further privilege records that have historically received more archival attention. The cultural black box here is not only "how did the algorithm work?" but also why these records were available, selected, digitised and used as the training data.

The second cultural black box concerns representation. Cultural researchers to justify how datasets and AI training represent the cultural identity of the human group addressed in an AI project (Park, 2025). This is highly relevant to archives, where records are often used to represent communities, institutions, nations and historical events. AI systems may create the impression that archival representation is comprehensive. Yet archives often contain silences. Some communities are over-documented by state surveillance, others are under-documented because their knowledge was oral, informal, marginalised, or deliberately excluded. Some people appear in archives only through records created by authorities, not through their own voices. When AI systems summarise, classify, visualise or retrieve archival records, they may reproduce these archival inequalities. A search system may make dominant records more visible while further obscuring marginalised voices. A summarisation system may compress contested histories into simplified narratives. A recommendation system may guide users toward the most frequently described or digitised materials. The cultural black box here concerns how AI transforms archival visibility into cultural authority. Archivists need to explain not only what the AI system has found, but also what the archive may not contain, what the data underrepresents and whose perspectives are absent.

The last cultural black box concerns context. A record becomes archival evidence because it is connected to provenance, custody and the time when it was created. AI systems, however, often process records as data objects. They may extract text, names, images, or patterns while detaching them from the archival context that gives them evidential meaning. This creates a risk of decontextualisation. A document retrieved through semantic search may appear relevant, but its administrative function, creator, record series, custodial history, and relationship to other records may be unclear. A generated summary may be linguistically accurate but archivally misleading if it omits provenance or uncertainty. Archivists need to make clear how AI outputs relate to provenance, original order, archival arrangement and the evidential status of records.

Cultural Explainability through proactive archival Practice

While the cultural black box issues specify the concerns of AI in archives, cultural explainability offers approaches to address some remedy. I defined cultural explainability as the procedural clarity that arises from uncovering those opaque human decisions (Park, 2025), i.e., AI transparency not about the machine's reasoning but about the people's. For the archive, this reframes a generative AI project not as a technical pipeline but as a documented sequence of interpretive choices, each of which can be questioned and revised. Crucially, cultural explainability cannot be achieved by any individual's good intentions or technical skill alone; it is a collective accomplishment, requiring that the relevant stakeholders agree upon, and record, the interpretive approach embedded in the system.

The archivist's first and most natural role in implementing cultural explainability is to document the project as it unfolds, capturing the interpretive decisions that would otherwise dissolve into the finished output. Before any AI model is trained, someone must decide what data could legitimately represent the subject at hand, what I term the culturally core data (Park, 2025), the fundamental elements that constitute a cultural practice or historical reality. This is a curatorial judgement of exactly what archivists are trained to make. They can articulate why certain materials are central and others peripheral and can make that reasoning explicit. Archivists are uniquely positioned to explain the context of records. They understand provenance, original order, custodial history, appraisal, description, and access conditions. When a period was chosen to stand for a centuries-long tradition, when a body of data was judged sufficient and collection was halted, when a missing record was treated as a null value rather than reconstructed; each such decision, and the reasoning behind it, belongs in the paradata.

Of course, no collection fully represents the reality it documents; and no archivist needs persuading of this. The archivist can be rather positioned to interrogate the representativeness of a dataset, asking which voices the available records privilege and which they omit, to flag imbalances among regions, periods, or communities, and to insist that the resulting limitations be declared rather than smoothed over by technical expedients such as treating absent evidence as interchangeable with present evidence. In this role the archivist ensures that an AI-generated representation is offered as an interpretation of the record, candidly bounded, rather than as the record itself.

The archivist is the institution's established keeper of documentation, who can extend such custodianship to AI workflows. An AI-generated cultural artefact without its paradata can be suspected as a document without provenance. These principles are essential for preventing AI systems from treating records as isolated datasets. In AI projects, archivists can ensure that archival context is included in model design, metadata enrichment, search interfaces, and user-facing explanations. They can ask whether AI outputs preserve the relationship between records and their creators, the entire collection and institutional histories. This role shifts the archivist from passive collection manager to active interpreter of data context to be embedded in an AI model.

Archivists can further design human-in-the-loop workflows that determine when AI-generated outputs require review, correction, rejection, or community consultation. High-risk materials may include records involving personal data, legal rights, children,

health, ethnicity, religion, political persecution, conflict, displacement, or traumatic events. In these cases, cultural explainability requires careful explanation of why the particular AI system was used, how outputs were checked, and what limitations remain. Archivists can define different levels of review depending on risk, sensitivity and public impact. Archivists can therefore provide guidance that encourages critical use of AI-mediated archival systems for non-archivists.

Finally, archivists can act as the bridge between the institution and the communities whose heritage is at stake, and as the keeper of an account of who contributed to the work. AI projects in archives often involve computer scientists, software vendors, institutional managers, legal teams, community representatives, researchers, and users. Archivists can act as mediators among these groups. It requires collective decisions about how culture is interpreted and represented. In archival settings, archivists can facilitate these collective decisions by translating between technical language, archival principles, and community concerns. They can help technical teams understand why provenance, sensitivity, historical terminology, and representational gaps matter. They can also help communities understand how AI tools may affect access, visibility, and interpretation. Much of the labour behind AI-based creation - the gathering, the transcription, the countless micro-tasks of curation - is routinely rendered invisible. The archivist, long accustomed to recording the chain of custody and the hands through which materials have passed, is the appropriate steward of an inclusive acknowledgement: one that names not only the engineers and curators but the source communities, donors and local knowledge-holders whose participation legitimises the project. Convening those communities in the decisions that concern them, and recording their role, is the surest defence against an archive that speaks about a people without speaking with them.

Conclusion

Artificial Intelligence represents a major structural shift in how human civilization records, manages, and preserves its institutional memory. Far from rendering the archivist obsolete, the expansion of automated data management underscores the critical need for professional archival oversight. While AI provides the processing capacity needed to manage (digital data), humans must provide the necessary ethical framework, historical context, and critical judgment. It offers significant opportunities for managing digital scale, improving access, enriching metadata, supporting discovery and assisting preservation workflows. The future of archives should not be imagined as AI replacing archivists. Rather, it should be understood as archivists shaping the responsible use of AI.

The central question is no longer whether AI will affect archives. It already does. The more important question is whether archivists will help determine how AI is used in relation to documents, memory, rights and public trust. The archivist profession should and will be evolving from a custodian of physical items into an active designer of AI trust systems. This requires direct collaboration with computer scientists and software engineers to ensure that core archival principles - specifically provenance, original order, and long-term authenticity - are explicitly programmed into the design of emerging technologies. By taking an active role in shaping these technologies, the archival community can harness AI to protect and preserve our increasingly complex digital heritage. In doing so, archivists can ensure that AI does

not merely accelerate archival work but strengthens archival accountability and cultural responsibility. The future of AI in archives should therefore be guided not only by computational efficiency, but by archival values that ensure interdisciplinary profession.

References

European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). Official Journal of the European Union. https://eur- lex.europa.eu/eli/reg/2024/1689/oj/eng. (accessed 28.05.2026).

International Council on Archives (2012) Principles of Access to Archives. https://www.ica.org/resource/principles-of-access-to-archives/ (accessed 28.05.2026).

NARA. U.S. National Archives and Records Administration. (2026). Inventory of NARA Artificial Intelligence (AI) Use Cases. https://www.archives.gov/ai?utm_source (accessed 28.05.2026).

Park, S. (2025) 'Revisiting Transparency in AI-based Creation: Unpacking "Cultural Black Boxes" towards "Cultural Explainability", 2025 IEEE International Symposium on Technology and Society (ISTAS). DOI: 10.1109/ISTAS65609.2025.11269653

UNESCO (2022) Recommendation on the Ethics of Artificial Intelligence. UNESCO. https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence (accessed 28.05.2026).

ARTIFICIAL INTELLIGENCE FOR ACCESSIBLE ARCHIVES: MAKING RECORDS SEARCHABLE FOR ALL

Stephen Filios
PhD, Senior Research Scientist in Handwriting Recognition, FamilySearch International, USA

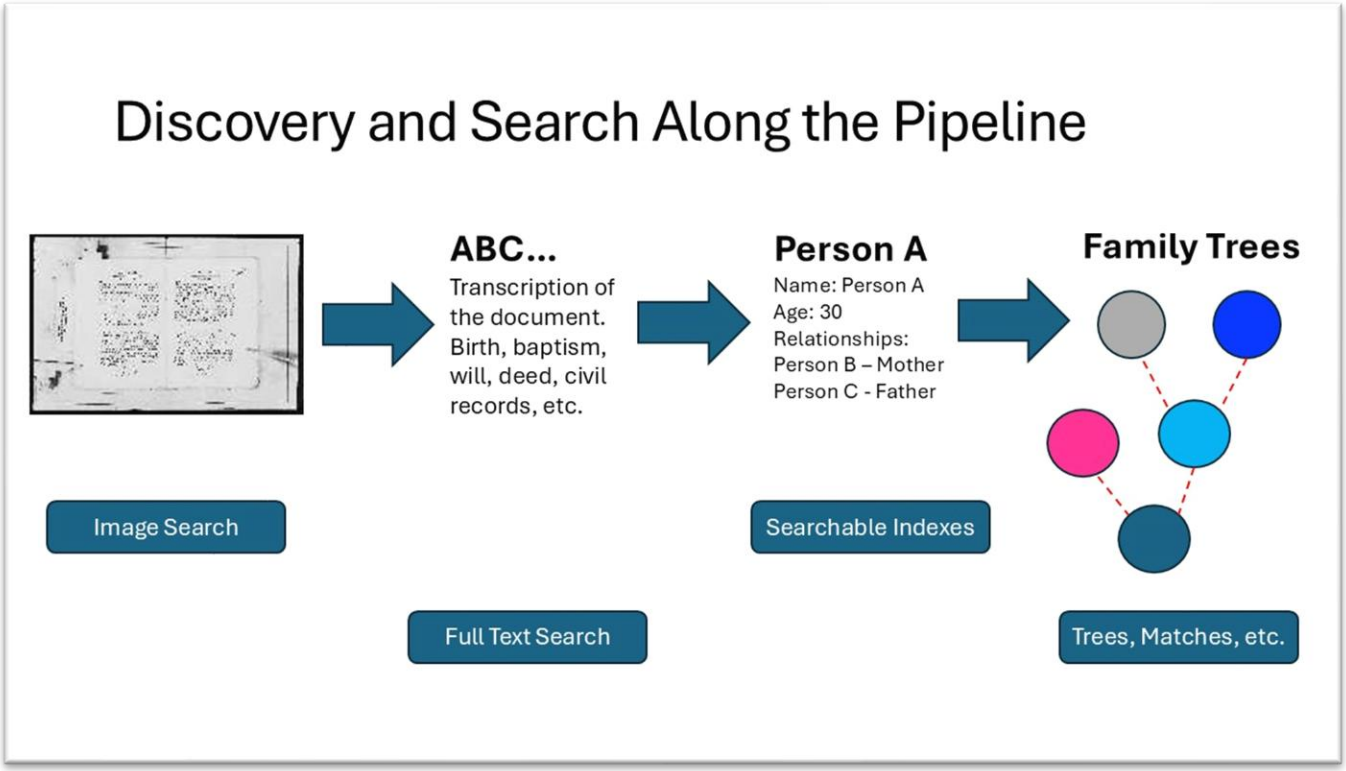
ҚОЛЖЕТІМДІ АРХИВТЕРГЕ АРНАЛҒАН ЖАСАНДЫ ИНТЕЛЛЕКТ: ҚҰЖАТТАРДЫ ІЗДЕУ МҮМКІНДІГІН БАРШАҒА БІРДЕЙ ҚАМТАМАСЫЗ ЕТУ

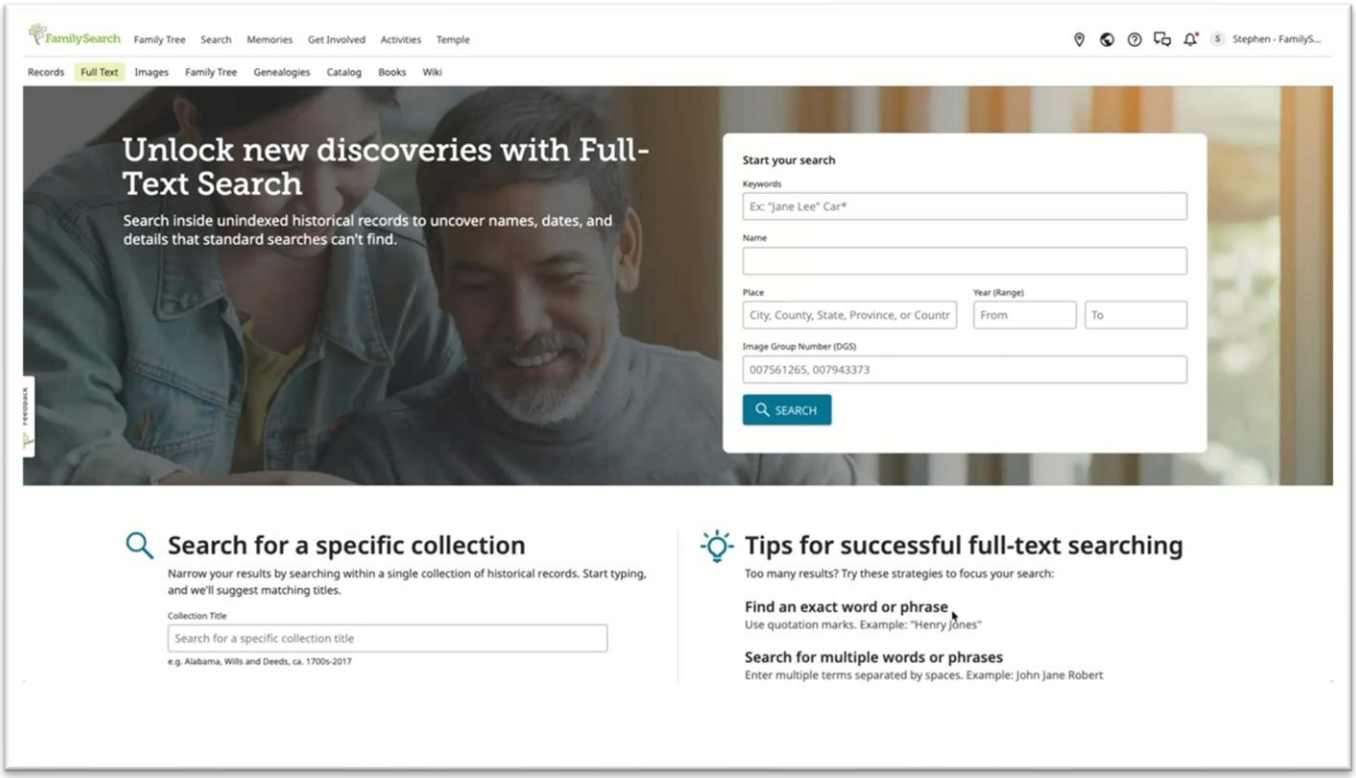
Стивен Филиос
PhD, FamilySearch International қолжазба мәтіндерін тану бойынша жетекші ғылыми қызметкері, АҚШ

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ДЛЯ ДОСТУПНЫХ АРХИВОВ: ОБЕСПЕЧЕНИЕ ВОЗМОЖНОСТИ ПОИСКА ЗАПИСЕЙ ДЛЯ ВСЕХ

Стивен Филиос
PhD, ведущий научный сотрудник по распознаванию рукописного текста FamilySearch International, США

DIGITAL ARCHIVE EXPO
ASTANA 08-12.06.2026
III АРХИВШІЛЕР КОНГРЕСІ · III CONGRESS OF ARCHIVISTS


Discovery and Search Along the Pipeline
ABC...
Transcription of the document. Birth, baptism, will, deed, civil records, etc.
Person A
Name: Person A
Age: 30
Relationships:
Person B – Mother
Person C - Father
Family Trees
Image Search
Searchable Indexes
Full Text Search
Trees, Matches, etc.



LLMs Can Improve Full-Text Results

Show Translation (English)

Numero 79 Bongiorni Anna Numero 80 **Figliossi** Anerina ATTI DI NASCITA . L' anno milleottocentonovantuno , addi ventisette di Luglio a ore po meridiane quattro e minuti Zero , nella Casa Comunale . Avanti di me Cuiffo Giuseppe Segretario delegato dal Sindaco con atto dellin dici Marzo mille ottocento ottantacinque debitamente approvato Ufficiale dello Stato Civile del Comune di Santonio a Trebbia , e comparso Bongiorni Francesco di anni trentasette Veterinario domiciliato in questo Comune , il quale mi ha dichiarato che alle ore ante meridiane due e minuti Zero , del di ventisei del corrente mese , nella casa posta in Sombardina al numero duecento sedici , dalla Manenti Maria massaia , sua moglie seco lui convivente è nato un bambino di sesso femminile che non mi presenta , e a cui da i nome di Anna Maria Maddalena A quanto sopra e a questo atto sono stati presenti quali testimoni Marente

ATTI DI NASCITA

Numero 79: Bongiorni Anna

L'anno milleottocentonovantuno, addi ventisette di Luglio a ore po meridiane quattro e minuti Zero, nella Casa Comunale. Avanti di me Cuiffo Giuseppe Segretario delegato dal Sindaco con atto dellin dici Marzo mille ottocento ottantacinque debitamente approvato Ufficiale dello Stato Civile del Comune di Santonio a Trebbia, e comparso Bongiorni Francesco di anni trentasette Veterinario domiciliato in questo Comune, il quale mi ha dichiarato che alle ore ante meridiane due e minuti Zero, del di ventisei del corrente mese, nella casa posta in Sombardina al numero duecento sedici, dalla Manenti Maria massaia, sua moglie seco lui convivente è nato un bambino di sesso femminile che non mi presenta, e a cui da i nome di Anna Maria Maddalena.

A quanto sopra e a questo atto sono stati presenti quali testimoni Marente Giuseppe, di anni ventinove Galtro Ferraio, e Dondi Andrea, di anni cinquantadue Pensionato, entrambi residenti in questo Comune. Il dichiarante è stato da me dispensato dal presentare il bambino suddetto per la gracilita del medesuno, essendomi io stesso in altro modo accertato della verita della nascita. Letto il presente atto a tutti gl'intervenuti lo hanno questi con me sottoscritto.

Bongiorni Francesco Manenti Giuseppe Il Segretario delegato Gallo

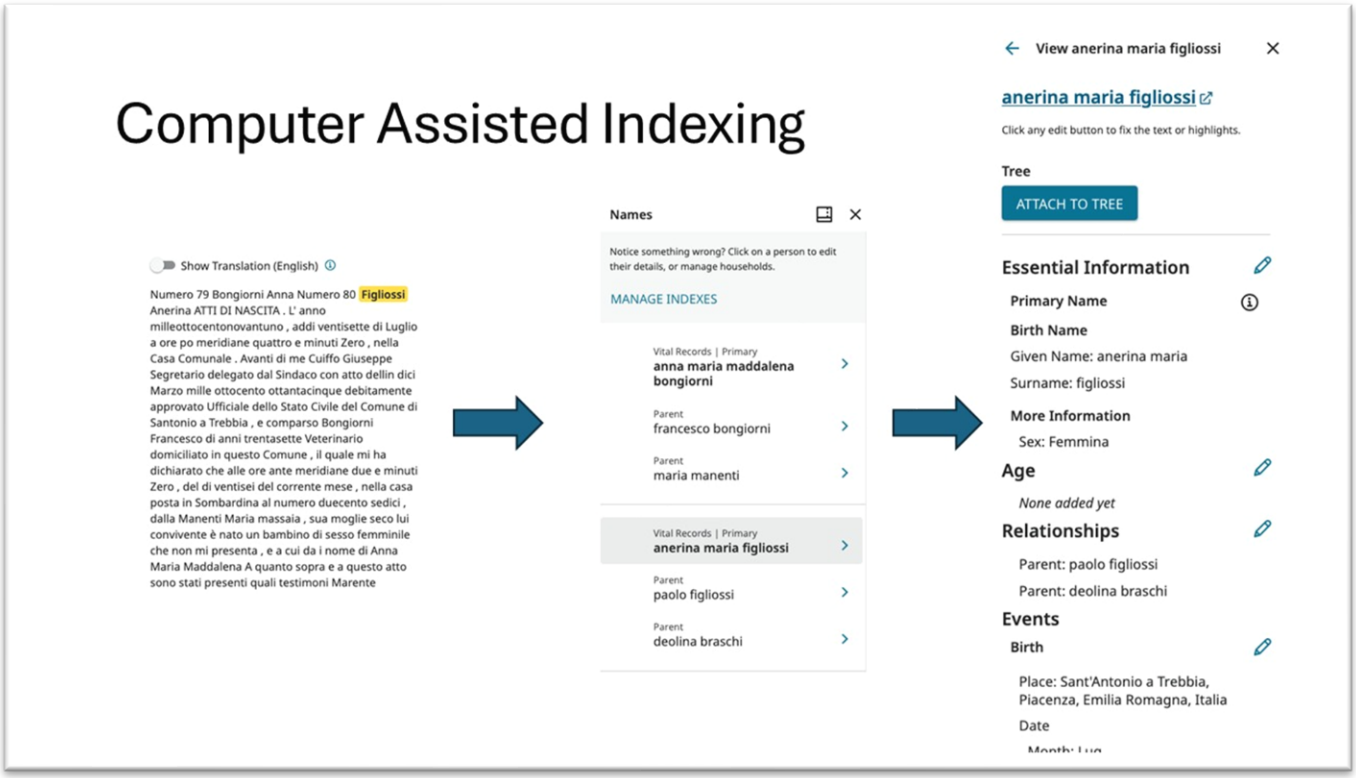
Numero 80: Figliossi Anerina

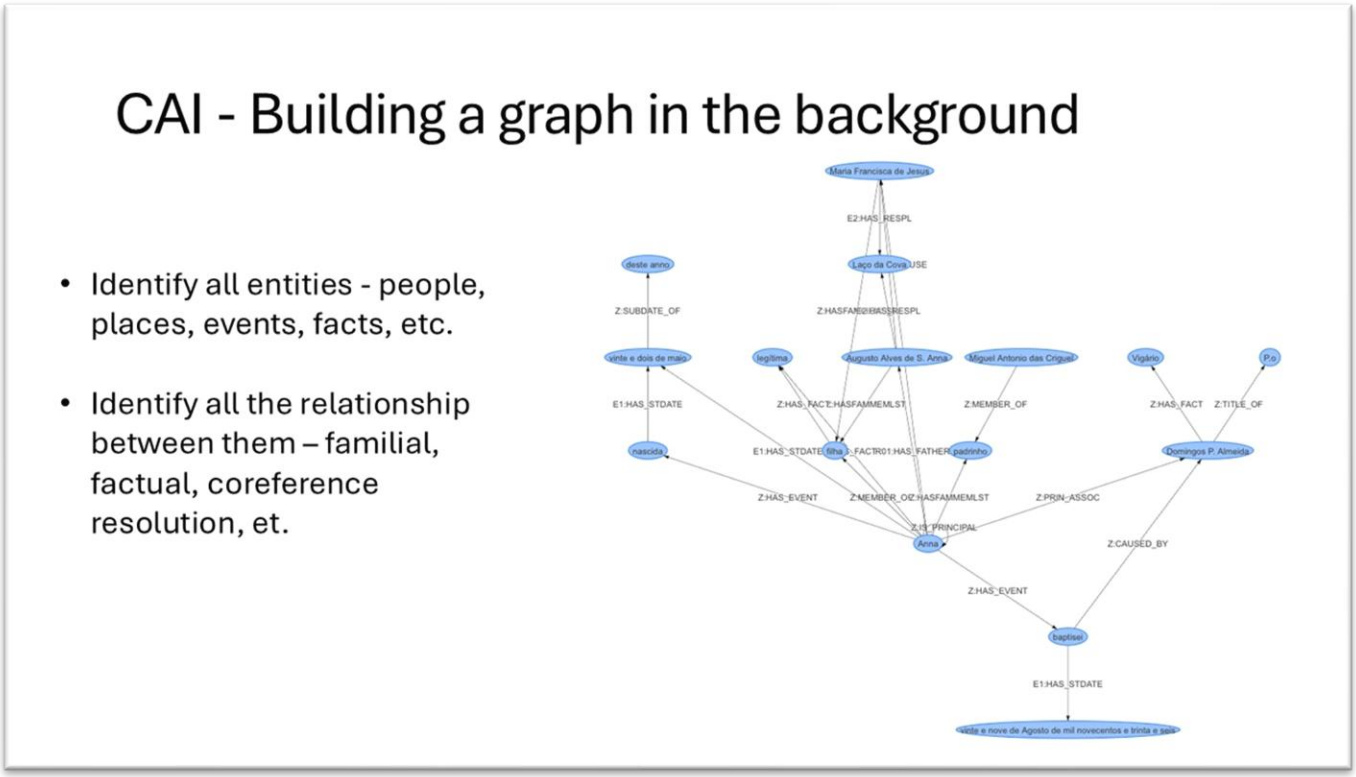
L'anno milleottocentonovantuno, addi trentuno di Luglio a ore po meridiane due e minuti quaranta nella Casa Comunale. Avanti di me Cuffo Giuseppe Segretario Delegato dal Sindaco con atto dell'indici Marzo milleottocento ottantacinque debitamente approvato Ufficiale dello Stato Civile del Comune di S. Antonio e Tubbia, e comparso Figliossi Paolo, di anni trenta Fittabile domiciliato in questo Comune, il quale mi ha dichiarato che alle ore epo meridiane due e minuti Peco, del di trenta del volgente mese, nella casa posta in Maganza al numero centoquarantuno, delle Bruschi Desolina, massaia, sua moglie seco Lui convivente è nato un bambino di sesso femminile che mi presenta, e a cui da i nomi di Anerina Maria.

(Continua a pagina destra) A quanto sopra e a questo atto sono stati presenti quali testimoni Sava Luigi, di anni sessantatre Amatore, e Sozzo Ercole, di anni cinquantasei Falegnome, entrambi residenti questo Comune. Letto il presente atto a tutti gl'intervenuti il solo Dichiarante lo ha con me sottoscietro, avendo i testimoni assento d'essere illetterati.

Figliossi Paola Il Segretario delegato

Nota a margine: Figliossi Averina, morta in PIACENZA 11 7-7-71 atto di morte del Comune di PIACENZA anno 97 P. 45. docunza, 15 FEB. 1972 Il crontioro


Computer Assisted Indexing
Show Translation (English)
Numero 79 Bongiorni Anna Numero 80 Figliossi Anerina ATTI DI NASCITA . L' anno milleottocentonovantuno , addi ventisette di Luglio a ore po meridiane quattro e minuti Zero , nella Casa Comunale . Avanti di me Cuiffo Giuseppe Segretario delegato dal Sindaco con atto dellin dici Marzo mille ottocento ottantacinque debitamente approvato Ufficiale dello Stato Civile del Comune di Santonio a Trebbia , e comparso Bongiorni Francesco di anni trentasette Veterinario domiciliato in questo Comune , il quale mi ha dichiarato che alle ore ante meridiane due e minuti Zero , del di ventisei del corrente mese , nella casa posta in Sombardina al numero duecento sedici , dalla Manenti Maria massaia , sua moglie seco lui convivente è nato un bambino di sesso femminile che non mi presenta , e a cui da i nome di Anna Maria Maddalena A quanto sopra e a questo atto sono stati presenti quali testimoni Marente
Names
Notice something wrong? Click on a person to edit their details, or manage households.
MANAGE INDEXES
Vital Records | Primary
anna maria maddalena bongiorni
Parent
francesco bongiorni
Parent
maria manenti
Vital Records | Primary
anerina maria figliossi
Parent
paolo figliossi
Parent
deolina braschi
View anerina maria figliossi
anerina maria figliossi
Click any edit button to fix the text or highlights.
Tree
ATTACH TO TREE
Essential Information
Primary Name
Birth Name
Given Name: anerina maria
Surname: figliossi
More Information
Sex: Femmina
Age
None added yet
Relationships
Parent: paolo figliossi
Parent: deolina braschi
Events
Birth
Place: Sant'Antonio a Trebbia, Piacenza, Emilia Romagna, Italia
Date


CAI - Building a graph in the background
• Identify all entities - people, places, events, facts, etc.
• Identify all the relationship between them – familial, factual, coreference resolution, et.
Maria Francisca de Jesus
E2:HAS_RESPL
deste anno
Laço da Cova
Z:SUBDATE_OF
vinte e dois de maio
legitima
Augusto Alves de S. Anna
Miguel Antonio das Criguel
Vigário
P.o
E1:HAS_STDATE
Z:MEMBER_OF
Z:HAS_FACT
Z:TITLE_OF
nascida
filha
padrinho
Domingos P. Almeida
Z:HAS_EVENT
Z:PRIN_ASSOC
Z:IS_PRINCIPAL
Anna
Z:CAUSED_BY
Z:HAS_EVENT
baptisei
E1:HAS_STDATE
vinte e nove de Agosto de mil novecentos e trinta e seis

Assisted Tree Building

- Connect the graphs across records, documents, family trees, and all disparate sources of data

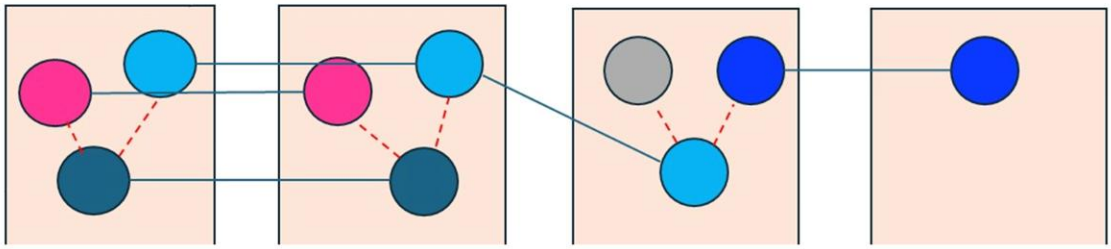
Birth record
Stephen Filios

Parents
Dad Filios
Mom Filios

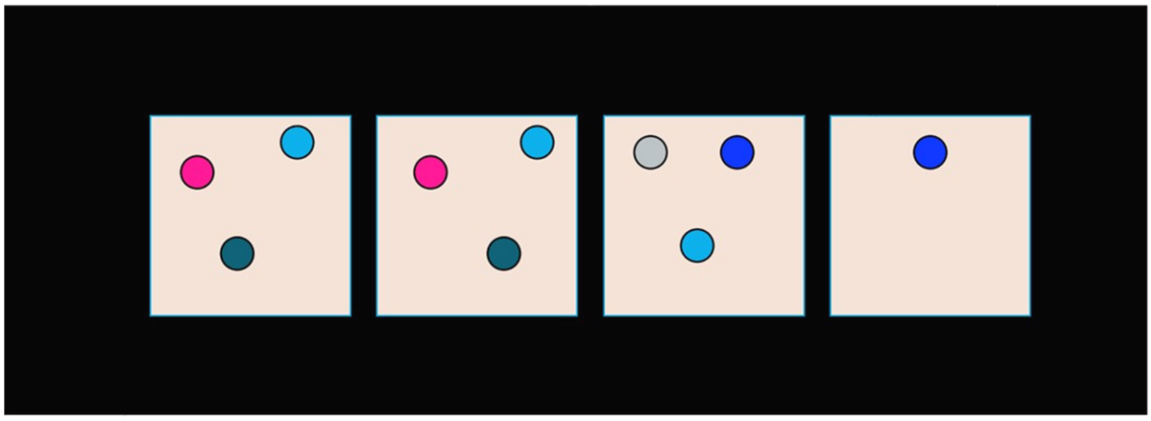
Census Record
Dad Filios
Mom Filios
Stephen Filios

Census Record
Grandpa Filios
Grandma Filios
Dad Filios

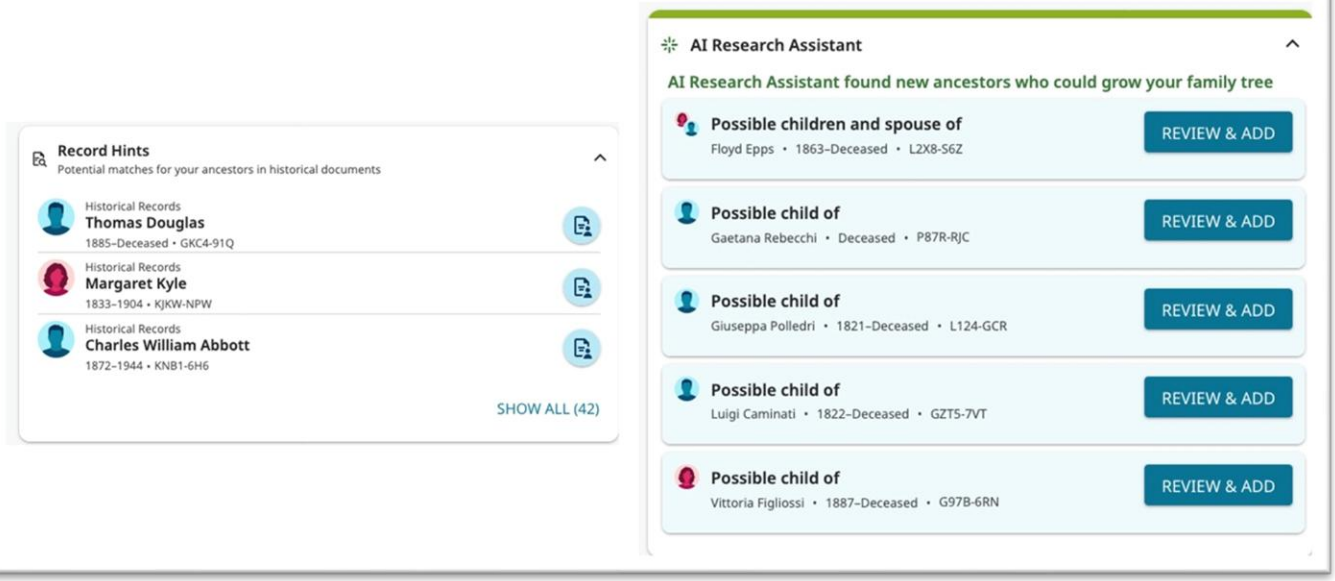
Immigration Record
Grandpa Filios



Personas Connected Across Sources

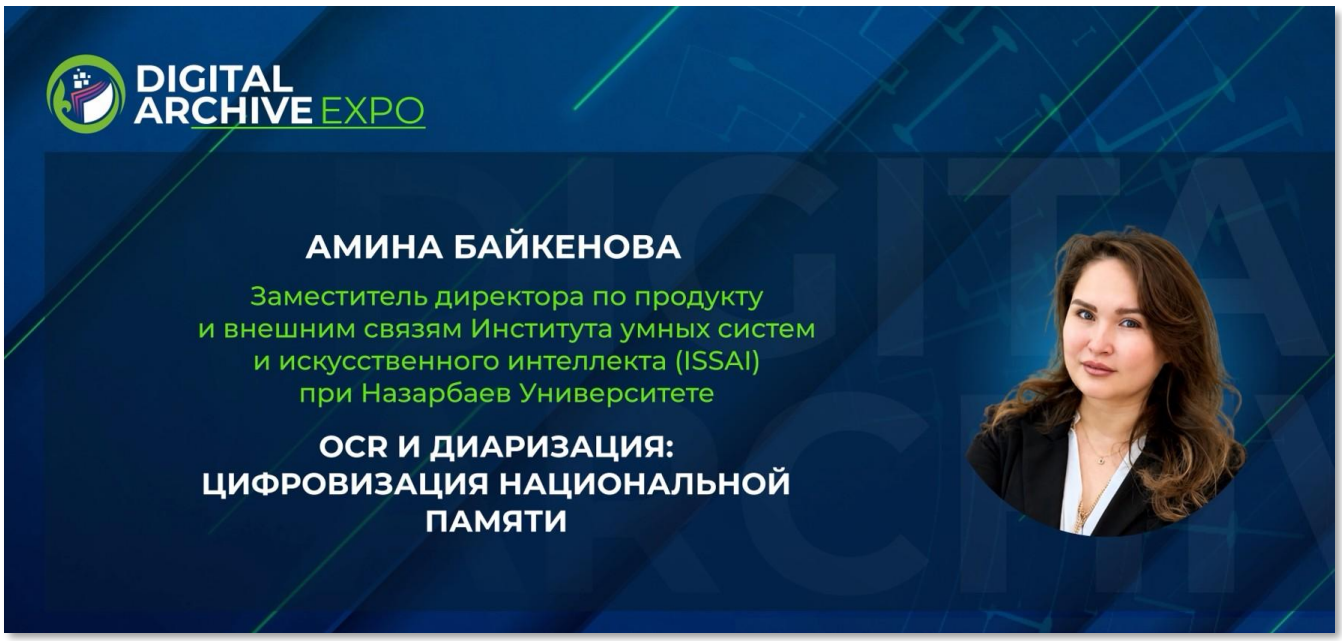


Humans in the Loop



How do we maintain data quality?

- Use neuro-symbolic Systems
  - Don’t rely solely on generative AI – combine it with ontologies, known schemas, and tools.
  - Enforce constraints on generations
- Provide grounding when using generative AI
  - Refer to specific parts of an image, specific examples
  - Provide transcriptions, or give gen AI material to start from rather than create de novo
- As much as possible, use only when you can verify the answers
- Keep humans in the loop where it makes sense



ЦИФРОВИЗАЦИЯ НАЦИОНАЛЬНОЙ ПАМЯТИ: МУЛЬТИСКРИПТОВЫЙ OCR И ДИАРИЗАЦИЯ ДЛЯ АРХИВНОГО НАСЛЕДИЯ

Амина Мурзабековна Байкенова
Заместитель директора по продуктам и внешним связям, магистр наук в области информатики, Институт умных систем и искусственного интеллекта (ISSAI), Назарбаев Университет, Казахстан
amina.baikenova@nu.edu.kz

Аннотация. Данная статья подготовлена на основе доклада для Архива Президента Республики Казахстан и посвящена вопросам цифровизации архивного наследия с применением передовых технологий искусственного интеллекта, разрабатываемых Институтом умных систем и искусственного интеллекта (ISSAI) Назарбаев Университета. В центре внимания находится мультискриптовое оптическое распознавание символов (OCR), которое критически необходимо для качественной обработки исторических документов на кириллице, латинице и арабской графике (төте жазу). Рассматривается успешное внедрение передовых моделей ИИ, таких как Oylan и MangiSoz, для обеспечения исторической преемственности, а также доступности культурного наследия для широкой общественности. Особое внимание в работе уделяется преодолению проблемы малоресурсности казахского языка в базах данных ИИ. Описывается использование синтетической генерации данных и трансферного обучения (Transfer Learning) с родственных тюркских языков, что позволяет значительно снизить частоту ошибок и достичь высокой точности распознавания уникальных рукописей. Кроме того, освещается технология диаризации речи, применяемая для аудиовизуальных архивов. Данная технология позволяет автоматизировать индексацию интервью и мемуаров с точным определением говорящего с учетом фонетических особенностей казахских диалектов и акцентов. Делается вывод, что комплексная цифровизация национальной памяти является стратегическим шагом для сохранения культурной идентичности стран Центральной Азии.

Ключевые слова: мультискриптовый OCR; диаризация речи; цифровизация архивов; искусственный интеллект; ISSAI; MangiSoz; Oylan; төте жазу; малоресурсные языки; трансферное обучение

ҰЛТТЫҚ ЖАДТЫ ЦИФРЛАНДЫРУ: АРХИВТІК МҰРАҒА АРНАЛҒАН МУЛЬТИСКРИПТТІК OCR ЖӘНЕ СӨЙЛЕУДІ ДИАРИЗАЦИЯЛАУ

Байкенова Амина Мұрзабекқызы
Өнімдер және сыртқы байланыстар жөніндегі директордың орынбасары, информатика саласы бойынша ғылым магистрі, Назарбаев Университетінің Ақылды жүйелер және жасанды интеллект институты (ISSAI), Қазақстан
amina.baikenova@nu.edu.kz

Аңдатпа. Мақала Қазақстан Республикасы Президентінің Архивіне арналған баяндама негізінде дайындалып, Назарбаев Университетінің Ақылды жүйелер және жасанды интеллект институтында (ISSAI) әзірленген озық жасанды интеллект технологияларын пайдалану арқылы архивтік мұраны цифрландыру мәселелеріне арналған. Зерттеудің негізгі бағыты тарихи құжаттарды кирилл, латын және араб графикасында (төте жазу) сапалы өңдеуді қамтамасыз ететін мультискриптті таңбаларды оптикалық тану (OCR) технологиясына бағытталған. Тарихи сабақтастықты сақтауды және мәдени мұраның қалың жұртшылық үшін қолжетімділігін қамтамасыз ету мақсатында Oylan және MangiSoz сияқты жасанды интеллект модельдерін тиімді қолдану тәжірибесі қарастырылады. Сонымен қатар жұмыста жасанды интеллект жүйелеріндегі қазақ тілінің ресурстық шектеулілігі мәселесін еңсеруге ерекше назар аударылады. Осы мақсатта синтетикалық деректерді генерациялау және түркі тілдес туыстас тілдер негізінде трансферлік оқыту (Transfer Learning) әдістерін қолдану сипатталып, олардың бірегей қолжазбаларды тану кезінде қателік деңгейін едәуір төмендетуге және жоғары дәлдікке қол жеткізуге мүмкіндік беретіні көрсетіледі. Бұдан бөлек, аудиовизуалды архивтерді өңдеуде қолданылатын сөйлеуді диаризациялау технологиясының мүмкіндіктері қарастырылады. Бұл технология сұхбаттар мен мемуарлық материалдарды автоматты түрде индекстеуге, сондай-ақ қазақ тілінің диалектілік және фонетикалық ерекшеліктерін ескере отырып, сөйлеушілерді дәл сәйкестендіруге мүмкіндік береді. Зерттеу нәтижесінде ұлттық жадты кешенді цифрландыру Орталық Азия елдерінің мәдени бірегейлігін сақтаудың маңызды стратегиялық бағыты болып табылатыны туралы қорытынды жасалады.

Түйін сөздер: мультискриптті таңбаларды оптикалық тану (OCR); сөйлеуді диаризациялау; архивтерді цифрландыру; жасанды интеллект; ISSAI; MangiSoz; Oylan; төте жазу; ресурсы шектеулі тілдер; трансферлік оқыту

DIGITALIZATION OF NATIONAL MEMORY: MULTISCRIPT OCR AND DIARIZATION FOR ARCHIVAL HERITAGE

Amina M. Baikenova
Deputy Director for Products and External Relations, Master of Science in Informatics, Institute of Smart Systems and Artificial Intelligence (ISSAI), Nazarbayev University, Kazakhstan
amina.baikenova@nu.edu.kz

Abstract. This article is based on a report for the Presidential Archive of the Republic of Kazakhstan and is devoted to the digitalization of archival heritage using advanced artificial intelligence technologies developed by the Institute of Smart Systems and Artificial Intelligence (ISSAI) at Nazarbayev University. The focus is on multi-script optical character recognition (OCR), which is critically necessary for processing historical documents in Cyrillic, Latin, and Arabic scripts (Tote zhazu). The implementation of AI models, such as Oylan and MangiSoz, is considered to ensure historical continuity and accessibility of cultural heritage. Special attention is paid to overcoming the low-resource problem of the Kazakh language in AI databases. The use of synthetic data generation and transfer learning from related Turkic languages is described, which significantly reduces the error rate and achieves high accuracy in recognizing unique manuscripts. In addition, speech diarization technology for audiovisual archives is highlighted. This technology automates the indexing of interviews and memoirs, accurately identifying the speaker while taking into account the phonetic features of Kazakh dialects. It is concluded that the comprehensive digitalization of national memory is a strategic step for preserving cultural identity.

Keywords: multiscript OCR; speech diarization; archive digitalization; artificial intelligence; ISSAI; MangiSoz; Oylan; Tote zhazu; low-resource languages; transfer learning

Уважаемые коллеги, руководство Архива Президента Республики Казахстан, дорогие гости и участники!

Для меня большая честь выступать сегодня в стенах учреждения, которое по праву можно назвать главным хранителем документальной истории нашей страны. Тема, которую мы сегодня затронем, находится на стыке гуманитарных традиций и передовых технологий. Мы поговорим о цифровизации национальной памяти и о том, как продукты Института умных систем и искусственного интеллекта (ISSAI) Назарбаев Университета способны трансформировать наше взаимодействие с историческим наследием (Fazli 2023: 12).

Мультискриптовый OCR для исторической преемственности

Архивное наследие Казахстана – это не просто набор текстов, это сложный лингвистический и визуальный ландшафт. Исторически сложилось так, что документальная база нашей страны охватывает сразу несколько графических систем: кириллицу, латиницу и арабскую графику, известную как төте жазу. Долгое время перевод таких документов в машиночитаемый формат требовал колоссальных затрат человеческих ресурсов.

Сегодня мы представляем решения, меняющие эту парадигму. Внедрение передовых моделей искусственного интеллекта, разработанных в ISSAI, таких как Oylan и MangiSoz, открывает новые горизонты. Эти модели способны к мультискриптовому распознаванию текстов (OCR). Это означает, что машина может с высокой точностью считывать исторические документы независимо от того, на какой графике они написаны, и мгновенно переводить их в современный цифровой формат.

Зачем это нужно? В первую очередь, для преодоления разрыва между историческими документами, хранящимися в фондах, и современным цифровым доступом исследователей, студентов и всех граждан. Мы получаем возможность оцифровать огромный массив ранее недоступных или трудночитаемых записей. Это делает архивы по-настоящему открытыми для народа и служит

надежным инструментом сохранения культурного наследия не только Казахстана, но и всех стран Центральной Азии (Алпысбаева 2020: 15).

Преодоление ограничений «малоресурсности»

На пути разработки таких систем мы столкнулись с серьезным вызовом. В мировом сообществе разработчиков ИИ казахский язык традиционно считается малоресурсным. Это означает, что в открытом доступе недостаточно качественных текстовых и аудиоданных для обучения нейросетей, в отличие от английского или испанского языков.

Как ISSAI решает эту проблему? Мы применяем инновационные подходы в машинном обучении. Во-первых, это синтетическая генерация данных – процесс, при котором алгоритмы создают миллионы искусственных, но грамматически и визуально корректных примеров текста, обучая систему распознавать сложные и редкие символы.

Во-вторых, мы активно используем Transfer Learning – перенос обучения. Обучив модель на данных крупных и родственных тюркских языков, мы переносим полученные «знания» на казахский язык. Этот подход позволяет значительно снизить частоту ошибок в распознавании символов (CER – Character Error Rate). В результате мы обеспечиваем высочайшую точность оцифровки уникальных архивных рукописей, даже когда обучающих выборок критически мало. Мы заставляем ИИ понимать контекст, стилистику и орфографию прошлых эпох.

Диаризация речи для устной истории

Однако архивы – это не только бумага. Аудиовизуальные фонды: записи парламентских сессий, исторические интервью, радиопередачи, мемуары – это живой голос нации. Работать с многочасовыми аудиозаписями вручную – тяжелейший труд для архивистов.

Здесь на помощь приходит технология диаризации речи. Диаризация – это процесс автоматического разделения аудиопотока на фрагменты по принципу «кто и когда говорит». Используя модели ISSAI, система самостоятельно анализирует аудиозапись, транскрибирует текст и маркирует его: «Спикер 1», «Спикер 2» и т.д.

Особенная гордость наших разработок заключается в учете фонетических особенностей казахского языка. Модели ISSAI натренированы таким образом, что они гарантируют точность при различении носителей разных диалектов и акцентов. Это революционизирует процесс каталогизации устной истории. То, на что раньше уходили недели ручной расшифровки, теперь автоматизируется, позволяя архивистам сосредоточиться на анализе и систематизации смыслов, а не на технической работе (Smith 2022: 104).

Заключение

Подводя итог, хочу подчеркнуть: интеграция технологий ISSAI – моделей Oylan, MangiSoz, систем OCR и диаризации – в работу Архива Президента РК и других архивных учреждений является историческим шагом. Цифровизация национальной памяти – это не просто сканирование документов. Это возвращение истории в активный оборот общества. Это гарантия того, что ни один голос прошлого не будет утерян из-за смены алфавитов или хрупкости бумаги.

Литература

Fazli 2023 – Fazli S., Алиев A. Advanced OCR models for low-resource Turkic languages. London: TechPress, 2023. 210 p.

Smith 2022 – Smith J. Speech diarization in historical audio archives: methods and practices. New York: Academic Books, 2022. 150 p.

Алпысбаева 2020 – Алпысбаева Н., Шайлазымов Б. Оцифровка архивных документов: методические рекомендации. Алматы: Қазақ университеті, 2020. 68 с.

Амандыкова 2011 – Амандыкова С.К. Декларация о государственном суверенитете – новая эпоха в конституционном развитии Казахстана.

https://articlekz.com/article/6118 (дата обращения: 08.08.2025).

FROM TRADİTİONAL ARCHİVES TO DATA CENTERS: THE CONTRİBUTİON OF ARTİFİCİAL INTELLİGENCE TO THE PROCESSİNG OF ARCHİVAL RECORDS AND RESEARCH PROCESSES

Beyzanur Altındiş
Archive Assistant Specialist, Directorate of State Archives
under the President of the Republic of Türkiye
beyzanur.altindis@devletarsivleri.gov.tr

Abstract. Today, information technologies play a decisive role at both discursive and practical levels in the production, reproduction, and circulation of knowledge. These technological capabilities, which support the theory-and-concept-oriented traditional approach of the social sciences and humanities, allow for a more systematic and efficient handling of data that is increasingly reaching massive proportions and, in some cases, carrying the risk of physical deterioration. In this context, the transition of the social sciences and humanities into a virtual "laboratory environment" makes it possible to reconstruct the processes of storage, archiving, classification, comparison, and analysis within a digital framework. At the core of this transformation lies Digital Humanities, an interdisciplinary field that not only provides new technical tools but also transforms the methodological and conceptual framework of historiography. Having its foundations laid in the text analysis studies of the 1940s, this field expanded with digital libraries and hypertext structures in the 1990s and has gained the quality of an independent discipline through institutionalization since the 2000s. Today, with the contribution of AI-supported analytical techniques, new horizons are emerging in the production and interpretation of historical knowledge.

Within this framework, tools such as Geographic Information Systems (GIS) and Social Network Analysis (SNA) offer new perspectives to scholarly work by making the spatial and relational dimensions of historical data visible. In particular, the analysis of big data sets through programming languages like Python, along with OCR/HTR (Optical/Handwritten Character Recognition) processes conducted via artificial intelligence applications, constitutes the current agenda of historical research. This study will exemplify how data obtained from

Ottoman archival documents can be processed by writing Python code with AI support and will discuss the inevitable impact of digital transformation on archives.

Keywords: Ottoman Arcvhives; Inheritance Records (Tereke); Imams; Data Analysis with Python; Artificial Intelligence (AI)

ДӘСТҮРЛІ АРХИВТЕРДЕН ДЕРЕКТЕР ОРТАЛЫҚТАРЫНА ДЕЙІН: АРХИВТІК ҚҰЖАТТАРДЫ ӨҢДЕУ МЕН ҒЫЛЫМИ-ЗЕРТТЕУ ҮДЕРІСТЕРІНДЕГІ ЖАСАНДЫ ИНТЕЛЛЕКТТІҢ ҮЛЕСІ

Бейзанур Алтындыш
Түркия Республикасы Президенті Әкімшілігінің
Мемлекеттік архивтер басқармасының қызметкері
beyzanur.altindis@devletarsivleri.gov.tr

Аңдатпа. Қазіргі таңда ақпараттық технологиялар білімді өндіру, қайта өндіру және тарату үдерістерінде теориялық та, практикалық та деңгейде шешуші рөл атқарады. Бұл технологиялық мүмкіндіктер әлеуметтік және гуманитарлық ғылымдардағы теориялық-концептуалдық дәстүрлі тәсілдерді қолдай отырып, көлемі үздіксіз артып келе жатқан әрі кейбір жағдайларда физикалық тұрғыдан бүліну қаупіне ұшырайтын деректерді неғұрлым жүйелі әрі тиімді өңдеуге мүмкіндік береді. Осы тұрғыдан алғанда, әлеуметтік және гуманитарлық ғылымдардың виртуалды «зертханалық ортаға» көшуі деректерді сақтау, архивтеу, жіктеу, салыстыру және талдау үдерістерін цифрлық кеңістікте қайта ұйымдастыруға жол ашады. Аталған трансформацияның өзегінде жаңа техникалық құралдарды ғана емес, сонымен қатар тарихнаманың әдіснамалық және тұжырымдамалық негіздерін де өзгертетін пәнаралық Цифрлық гуманитаристика (Digital Humanities) бағыты жатыр. Негізі XX ғасырдың 40-жылдарындағы мәтіндерді талдау зерттеулерінен бастау алған бұл бағыт 1990-жылдары цифрлық кітапханалар мен гипермәтіндік құрылымдардың дамуымен кеңейіп, 2000-жылдардан бастап институционалдану нәтижесінде дербес ғылыми пән ретінде қалыптасты. Қазіргі кезеңде жасанды интеллектке негізделген аналитикалық әдістердің дамуы тарихи білімді өндіру мен интерпретациялаудың жаңа көкжиектерін ашып отыр.

Осы тұрғыда географиялық ақпараттық жүйелер (GIS) мен әлеуметтік желілерді талдау (SNA) құралдары тарихи деректердің кеңістіктік және қатынастық өлшемдерін көрнекі түрде көрсету арқылы ғылыми зерттеулерге жаңа мүмкіндіктер береді. Әсіресе Python бағдарламалау тілі арқылы үлкен деректер жиынтықтарын талдау, сондай-ақ жасанды интеллект негізіндегі OCR/HTR (баспа және қолжазба мәтіндерін оптикалық тану) технологияларын қолдану қазіргі тарихи зерттеулердің негізгі бағыттарының біріне айналып отыр. Мақалада Осман архиві құжаттарынан алынған деректерді жасанды интеллекттің қолдауымен Python бағдарламалау тілі арқылы өңдеу тәсілдері мысал ретінде көрсетіліп, архив ісіндегі цифрлық трансформацияның ықпалы талданады.

Түйін сөздер: Осман архивтері; мұрагерлік істері (тереке); имамдар; Python көмегімен деректерді талдау; жасанды интеллект

ОТ ТРАДИЦИОННЫХ АРХИВОВ К ЦЕНТРАМ ОБРАБОТКИ ДАННЫХ: ВКЛАД ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОБРАБОТКУ АРХИВНЫХ ДОКУМЕНТОВ И ИССЛЕДОВАТЕЛЬСКИЕ ПРОЦЕССЫ

Бейзанур Алтындыш
Архивист Управления государственных архивов
при Администрации Президента Турецкой Республики
beyzanur.altindis@devletarsivleri.gov.tr

Аннотация. В настоящее время информационные технологии играют определяющую роль как на теоретическом, так и на практическом уровнях в процессах создания, воспроизводства и распространения знаний. Эти технологические возможности, дополняя традиционные теоретико-концептуальные подходы социальных и гуманитарных наук, обеспечивают более системную и эффективную обработку данных, объем которых постоянно возрастает и которые в ряде случаев подвержены риску физической утраты или повреждения. В этом контексте переход социальных и гуманитарных наук в виртуальную «лабораторную среду» позволяет по-новому организовать процессы хранения, архивирования, классификации, сопоставления и анализа данных в цифровом пространстве. В основе данной трансформации лежит междисциплинарное направление «Цифровая гуманитаристика» (Digital Humanities), которое не только предоставляет новые технологические инструменты, но и существенно трансформирует методологические и концептуальные основы исторической науки. Зародившись в исследованиях по анализу текстов в 1940-х годах, это направление получило дальнейшее развитие в 1990-х годах благодаря цифровым библиотекам и гипертекстовым технологиям, а с 2000-х годов, в результате институционализации, сформировалось как самостоятельная научная дисциплина. В настоящее время аналитические методы, основанные на технологиях искусственного интеллекта, открывают новые возможности для формирования и интерпретации исторического знания.

В этой связи такие инструменты, как географические информационные системы (GIS) и анализ социальных сетей (Social Network Analysis, SNA), позволяют выявлять пространственные и сетевые взаимосвязи исторических данных, расширяя возможности научных исследований. Особую актуальность приобретают анализ больших массивов данных с использованием языка программирования Python, а также применение технологий OCR/HTR (оптического распознавания печатного и рукописного текста), основанных на искусственном интеллекте. В статье на примере документов Османского архива демонстрируются возможности обработки архивных данных с использованием программного кода Python, созданного при поддержке искусственного интеллекта, а также рассматривается влияние процессов цифровой трансформации на современное архивное дело.

Ключевые слова: Османские архивы; наследственные дела (тереке); имамы; анализ данных с использованием Python; искусственный интеллект (ИИ)

Introduction

The proliferation of information technologies and the internet since the final quarter of the 20th century has fundamentally transformed scientific research processes, providing researchers with significant conveniences in areas such as catalog scanning, data storage, and access to information. However, the next-generation technological tools, software, and algorithms that have emerged in recent

years are transcending mere convenience to open new methodological horizons in the social sciences. These applications, which have rapidly established themselves in the literature under the heading of "Digital Humanities," currently constitute the primary agenda for many scholars.

One of the most promising application areas for this digital transformation is Ottoman historical studies. Millions of Ottoman documents and records are housed in numerous domestic and international institutions, most notably the Directorate of State Archives and the Directorate of Manuscrips of Türkiye. While this colossal heritage offers a rich source of data, it also brings certain structural challenges. The physical volume of archival material, the fragmentation of data across different institutions, and heterogeneous classification standards occasionally complicate research processes. Furthermore, the fact that digitized data is predominantly in "unstructured" form limits the ability to perform comprehensive, statistical, and comparative analyses on these texts. To overcome these limitations, digital tools such as Social Network Analysis (SNA), Geographic Information Systems (GIS), Python-based data mining, and Handwritten Text Recognition (HTR) have begun to be utilized more effectively in Ottoman historiography. In particular, leaps in artificial intelligence technology have increased the accessibility of these tools, enabling rapid and multi-layered analysis across extensive datasets.

Tools such as Gephi, UCINET, Pajek, and NodeXL provide vast opportunities for social network analysis. In the field of Ottoman historiography, the TÜBİTAK project titled «Professional and Intellectual Networks and Groupings of the High Ottoman Ulema (1470–1650)», completed by Abdurrahman Atçıl in 2020, represents an eminent example of social network analysis. Rather than interpreting the ulema through conventional institutional, hierarchical, and legal frameworks, this study employs an actor-oriented approach to analyze the complex networks of relationships among

these figures. Using Gephi, the study revealed the groupings and relationship patterns among the ulema serving in the *Sahn* Madrasas. Additionally, there are network analysis studies based directly on Ottoman archival documents. In his work, Atçıl emphasizes that ulema biographies provide a dataset particularly suitable for such tools. The greatest obstacle for researchers in utilizing these applications remains the creation of appropriate datasets and the challenges regarding their structuring[1*] (Atçıl, 2020).

Another field that has gained momentum in historical research with the integration of digital tools into the humanities is mapping studies conducted through Geographic Information Systems (GIS) in the light of historical documents. Although there are existing studies in this field regarding Ottoman history, Baş notes that the current literature remains limited and emphasizes the necessity of systematic databases for such methodological approaches (Baş, 2023: 49). Undeniably, the most strategic component in the construction of these databases consists of OCR and HTR (Handwritten Text Recognition) studies that provide automated transcription. Indeed, in response to the needs in this field, projects carried out by various initiatives such as IRCICA, Transkribus, BuildQ, and Osmanlıca.com have assumed a central position in the literature of digital archiving (Akça and Uğur, 2025: 363). While text analysis systems exhibit high accuracy rates in languages with discrete character structures, such as those using the Latin alphabet, they have yet to reach the expected proficiency in languages with cursive scripts like Ottoman Turkish. On the other hand, the momentum in artificial intelligence technologies in recent years presents a promising outlook, suggesting that these technical bottlenecks will soon be overcome (Akça and Uğur, 2025: 358-359).

Indeed, in recent years, archival institutions and scholars have been debating the role of artificial intelligence within archives. In this context, AI-powered semantic search algorithms integrated into digital document scanning systems process queried keywords not only morphologically but also in conjunction with word groups, root directories, and semantic equivalents (synonyms). This technological expansion aims to offer a new user experience that enables researchers to access required data in a much more holistic and comprehensive manner.

The utilization of specifically trained AI Agents in the processes of document summarization, indexing, and the determination of metadata parameters represents a strategic threshold for ensuring standardization and increasing operational efficiency in archival management. These technological capabilities are not limited to processing retrospective data; they also enable the autonomous management of Standard Filing Plans and Retention Schedules during the production phase of contemporary documents. This brings about numerous practical benefits, ranging from minimizing operational errors in archival management to maximizing data security. Beyond physical security, the protection of digital data against cyber threats through the principle of immutability provided by blockchain technology heralds a new horizon in the pool of practical benefits offered by modern archival science (Öztürk, 2021: 288-289).

[1*] Reference should also be made here to Zeynep Dörtok Abacı's 2015 project, "Being Too Connected or Powerful to Fail": Social Networks and Actors in Ottoman Society According to Sharia Court Records (1695-1700), and Betül İpşirli Argıt's article, "An Application of Social Network Analysis: Debt-Credit Relations of Ottoman Court Members in the Second Half of the 17th Century," in Digital Humanities and Ottoman Studies, ed. Yunus Uğur (Istanbul: VakıfBank Kültür Yayınları, 2023), 117-134.

This study aims to operationalize the methodological possibilities offered by these digital methods through 'Imam Inheritance Records' (Tereke Kayıtları) processed with Python and AI-supported analyses. In this regard, while exemplifying how raw data obtained from inheritance records are structured in a digital environment and analyzed with advanced algorithms, the process of archives evolving from static structures that merely 'preserve documents' into dynamic centers that directly 'produce processable data' for scientific research will be opened to methodological discussion.

1. Analytical Data in Ottoman Archives

The Ottoman archival heritage, by virtue of the formal structure of its constituent documents, offers a unique field for data mining within the digital humanities. However, not every document type is equally conducive to being transformed into an analytical dataset. In this context, "analytical data" pertains to the extent to which the information within the text can be standardized, categorized, and subjected to quantitative analysis. The record-keeping system developed by the Ottoman bureaucracy has prepared a highly suitable ground for converting documents – particularly those in specific forms—into structured data.

1.1. Principal Document Types Suitable for Data Mining

Documents with the highest potential for analytical data production are generally those that are list-based, possess a hierarchical order, and contain systematically recurring data entries. The formal structure of such documents methodologically facilitates the conversion of "raw data" into "structured data."

In this regard, the following document types within the Ottoman bureaucratic record-keeping system stand out for their potential to be transformed into digital databases:

a) Tahrir and Temettuât Registers (Cadastral and Tax Surveys): By classifying populations and heads of households according to geographical units, these records serve as rich data pools for demographic and socio-spatial analyses.

b) Ruznamçe and Muhasebe Registers (Daybooks and Accounting Ledgers): Containing the financial cycles of the state and various institutions, these registers function as financial datasets that allow for the quantitative modeling of economic trends.

c) Kassam and Sicil Registers (Inheritance/Estate Records - Tereke): These documents, which list the movable and immovable assets, debt-credit relations, and inheritance distributions of deceased individuals item by item, contain the most "micro" and "analytical" data of socio-economic history.

d) Temellük Records (Manuscript Ownership Records): Ownership notes on manuscripts substantiate the relational bonds between individuals, locations, and books, providing structured actor data for Social Network Analysis (SNA).

Since these records possess a systematic data structure kept within certain standards, they constitute the most suitable source group to be processed through Python-based data mining and AI-supported text analysis methods.

1.2. Inheritance Records (Tereke) as a Source of Raw Data

The suitability of inheritance records for analytical data production stems not only from the diversity of information they contain but also from their recording format—specifically, their formal and physical characteristics. Within the record-keeping

system developed by the Ottoman bureaucracy, *tereke* registers were generally maintained within a specific template, despite undergoing certain changes over the centuries. These records typically appear as inventory lists organized following a death and are usually found within the *sharia* court records (*kadı sicilleri*). Physically, a *tereke* record consists of three sections.

The first section detailed the identity of the deceased and their heirs. In addition to the name, title, profession, and social status of the deceased, the neighborhoods where they resided and passed away were recorded in detail. Furthermore, the identities of the heirs, their kinship relations to the deceased, their legal age status, and the guardians appointed for minors were explicitly stated. Information regarding whether the total assets of the deceased were sufficient to cover existing debts is also included here.

The second section, which constitutes the bulk of the record, is the inventory part containing the breakdown of the estate. In this section, every type of movable and immovable property; commercial commodities, tools of production and investment, books, household furniture, kitchenware, personal belongings, receivables, and cash assets were recorded in detail along with their quantities and monetary values. In this respect, the inventory section represents the richest data field for researchers. Indeed, this section allows economic historians to analyze components of wealth, material culture researchers to navigate through the objects of daily life, and historians of literacy to identify the books circulating in Ottoman society. The fact that each item was recorded with its market value (*raiç bedel*) enables the comparison of similar commodity items across different records, thus allowing for a quantitative analysis of price levels and consumption trends.

The third section is where expenditure and expense items were recorded. Although this is the shortest part, it is the most decisive in terms of legal and social networks. At this stage, after deducting expenses and debts from the total estate value, the remaining amount was distributed among the heirs according to the principles of Islamic inheritance law (*farā'id*). Despite its brevity, this final section contains immense numerical data suitable for analyzing research topics such as consumption habits, debt relations, and the effectiveness of cash waqfs within society. (Bozkurt, 2013; 202)

The aforementioned systematic structure of these records significantly facilitates the process of converting raw data into structured data. Consequently, at the stage of digital migration, these documents provide a fertile ground for producing datasets with high processability via algorithmic tools, suitable for standardization and big data analytics.

Tereke records appear as inventory lists prepared following a specific event (death). In the Ottoman Empire, the estates of individuals belonging to the *askeri* (military/administrative) class were recorded upon their death, partly because they were subject to certain taxes. This included various strata of officials such as janissaries, scribes, judges (*kadı*), madrasa professors (*müderris*), the *shaykh al-Islām*, and imams. Additionally, estates subject to inheritance disputes or those of individuals without heirs were also recorded in these registers.

Within the scope of this study, Imam Inheritance Records have been selected as a specific case study to operationalize the aforementioned analytical opportunities.

Due to their social status, the rich book lists and household items found in imams' estates offer a high-quality "data laboratory" for Python-based analysis at the intersection of intellectual capital and socio-economic profiles. In the following sections, the digitalization processes of these records and the results of algorithmic analyses conducted on this data will be discussed.

2. Transformation and Analysis of 18th-Century Imam Inheritance Records

The educational background and intellectual levels of imams in Ottoman society remain a contentious issue in the literature. While some researchers argue that the position of the imamate was a status passed from father to son without requiring high academic qualifications, (Beydilli,2018) others suggest that meritocracy was fundamental and that candidates were selected through specific examinations. (Doguş, 2021; 23) However, as both approaches lack a data pool broad enough to generalize these claims, these discourses have remained theoretical frameworks developed over a limited number of examples. At this juncture, the book lists inventoried in *tereke* records hold unique value for analyzing the intellectual world of imams, the nature of their relationship with knowledge, and their fields of interest on an empirical plane. Therefore, by moving beyond the restricted narratives of traditional historiography, it is possible to reinterpret the profile of Ottoman imams through digital methods and rewrite this profile using the resulting analytical data. Beyond individual life stories, this study aims to digitally reconstruct the collective mindset of a social group through large datasets.

2.1. Identification and Structuring of Data

During the research process, various court registers were examined to identify the register containing the highest concentration of imam *tereke* records. The Kısmet-i Askeriye Registers contain the majority of the estates belonging to the military/administrative (*askeri*) class. The number of these registers increased significantly, particularly from the second half of the 18th century onwards. Consequently, the scope of this research was limited to the second half of the 18th century, prioritizing registers from years marked by mass casualties such as plague epidemics, earthquakes, and fires. As a result of these screenings, 86 imam inheritance records were identified across 120 registers. In 75 of these records (87%), book entries were present. The total number of books owned by imams in these records is 1,617. While this figure represents a daunting heap for a historian to process manually, it is an operation of only a few seconds for digital tools.

Within the scope of this study, a dataset was created to analyze each book, along with its subject and price, across different contexts. Using Excel, a separate row was assigned to each book, recording which *tereke* it belonged to and its valuation. A "Subject" column was added to the Excel sheet to facilitate the classification of books by topic. The image below provides a view of the database.

Tablo 1. *Sample View of the Book Database*

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Sayı	Yazar	Konu	Eser Bilgisi	Eser Adı														
2	19	x	Mushaf	x	Kur'ân-ı Kerîm	125/64	x	127/84	x	129/35 (	141x	137/80	6000	137/80 (	x	138/59 (	x	139/52	(x
3	13	x	Mushaf	x	Kur'ân-ı Kerîm	421/4 (	x	421/35	x	421/51 (	1500(	421/51 (	4800	421/51 (	7700	427/92 (	25100	432/24	(x
4	13	Halebî	Fıkıh	İbrâhîm b. Muhammed b. İbrâhî	Mülteka'l-Ebhur	125/64	x	127/84	1020(	137/80 (	7120	150/37	2410	150/37 (	6200	410/39	6250	427/89	(465(
5	13	x	x	Yaygın olarak rastlanılan kitapla	Mecmua	139/52	x	150/37	1810(	150/37	3100	413/92 (	1900	413/92 (	4320	413/92 (	4400	421/51	(430(
6	11	x	x	x	Evrak	421/4	x	421/35	9100	427/92	2100	432/24	1560	432/54	4300	498/45	4010	498/45	(290(
7	10	x	x	Bu çalışmada da sıkça rastladığı	Evrak-ı perişân	125/64	127/84	129/35	8000	129/42	1950	137/80	1500	158/51	2400	161/1	2690	413/89	270(
8	10	Halebî	x	İbrâhîm b. Muhammed b. İbrâhî	İbrahim Halebi	137/80	26100	137/80	5000	150/37 (	1705	150/37 (	1440	410/39	2000	427/89 (	2660	427/89	(240(
9	9	Birgivi	x	Çeşitli sahalarda eser veren büyü	Birgivi	137/80	26100	421/4 (	3600	421/4 (E	1680	421/51 (	1120	421/35 (	1803	421/51 (	2400	427/89	(184(
10	8	Birgivî	İlmihal	x	et-Tarikatü'l-Muhammediyye	137/80	26100	150/37	3100	427/89	1600	427/92 (	1080	432/24 (	1620	496/39 (	1459	498/45	168(
11	8	Cezûlî	Evrad/Zikr	Şeyh Muhammed b. Süleyman e	Delail-i Şerif/Delaili Hayrat	137/80	19200	139/52	3000	421/35	1500	498/42	900	427/92	980	427/92	1020	496/39	849
12	8	x	x	Yaygın olarak rastlanılan kitapla	Mecmua	432/91	18500	498/42	2670	496/39 (	1200	496/39 (	860	421/35 (	910	421/35 (	980	498/45	750
13	7	Kudûrî	x	Müctehid Hanefî âlimi	Kudûrî	125/64	12800	129/42	2400	150/37(	1200	421/4 (K	850	496/39	900	421/35 (	810	498/45	550
14	7	Molla Câmi	x	Nakşibendî tarikatına mensup İr	Molla Câmi	127/84	11000	129/35	2300	413/89	1100	427/89	740	498/42 (	780	496/39	800	498/45	25(
15	7	Kâdî İyâz	Siyer	Kâdî İyâz'ın (ö. 544/1149) Peyg	Şifa-yı Kadı İyaz	125/64	9600	137/80	2100	427/89 (	1000	496/39 (	680	498/45	720	498/45	780	498/45	23(
16	7	Çatalcalı Ali Efe	Fıkıh	Şeyhülislâm Çatalcalı Ali Efend	Fetâvâ-yı Ali Efendi	150/37	9100	421/2	2040	427/92	1000	432/52 (	650	498/42	660	498/45	700	498/45	20(
17	7	x	Ayet	x	Enâm-ı Şerîf	158/51	8800	413/89	2040	432/52 (	810	432/52	600	491/10	420	498/45T	600	698/45	19(
18	7	x	Gramer	Arapça gramerinin kelime yapısı	Sarf Cümlesi	150/37	8300	410/39	2040	413/89	800	427/89 (	490	427/89 (	360	427/92	510	498/42	125
19	7	x	Gramer	Arapça gramerinin söz dizimi ku	Nahiv Cümlesi	125/64	8000	421/35	1960	129/35 (90	720	427/89	430	432/24	300	432/52 (	500	150/37	30
20	6	Yazıcıoğlu Mehı	?	Yazıcıoğlu Mehmed Efendi'nin (	Muhammediyye	150/37	7300	138/59	1820	155/63 (	600	203/141	429	427/89	300	498/45 (	450		
21	6	Radıyyüddin es-	Hadis	Radıyyüddin es-Sâgânî'nin (ö. 6	Meşârık	129/35	6550	137/80	1810	150/37 (	550	427/89 (	410	496/39 (	250	498/45	380		
22	6	x	x	TDV Nahiv maddesinde geçen a	Dinküz	127/87	6120	129/35	1710	137/80 (	500	427/89	400	421/35 (	240	498/45	270		
23	6	x	Hadis	Çeşitli konulara dair kırk hadisi	Hadis-i Erba'in	129/35	4970	137/80	1690	150/37	490	498/45	340	498/45	240	498/45 (	240		
24	6	x	Kıraat	x	Tecvid	150/37	4810	150/37	1590	427/89 (	480	427/89 (	330	498/42 (	240	498/45 t	140		

In a secondary Excel spreadsheet, comprehensive data such as the total wealth found in the imams' estates, the total number of books, the monetary value of these books, and the proportional share of books within the total wealth were recorded.

Tablo 2. *Sample View of the Estate Wealth and Book Value Dataset*

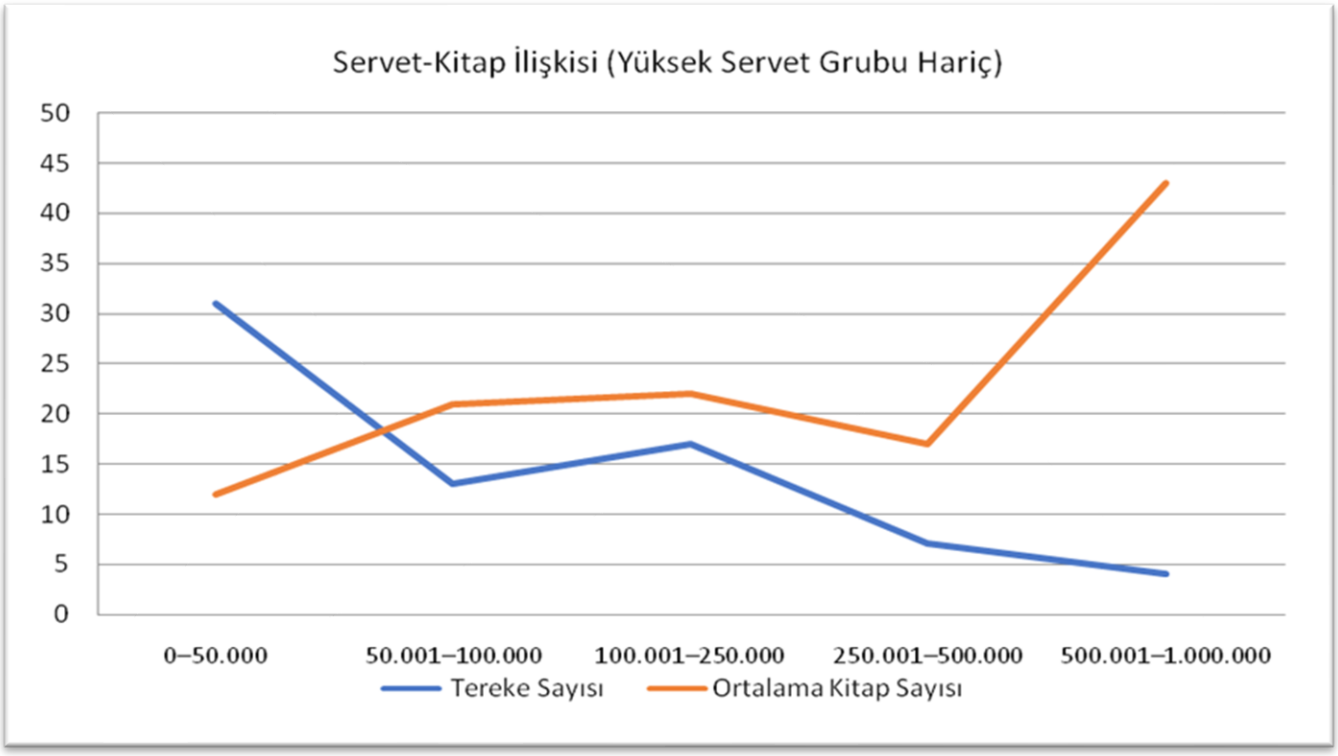
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Defter No	Jpeg No	Varak No	Meslek	Kitap	İsim	Ünvan	Meskun Mahalli	Görev Yaptığı Mahalle/	Toplam Servet	Kitap Sayısı	Kitapların Tutarı	Kitapların Servete Oran
2	54	10 b 2	54/8b-2	İmam	kitap var	Hüseyin b. Must	Efendi/el-Hac	x	Mahmud Paşa Camii İma	59,593	38	24,650	42
3	58	27 b 2	58/23b-2	İmam	kitap var	Abdurrahman	Efendi	İskender Ağa Mah	x	39,688	1	1,500	4
4	58	37 b 2	58/33b-2	İmam	kitap var	Süleyman b. Ma	Efendi	Hadika-i hassa şeh	Bağçe Bostanları İmamı	60,567	4	2,460	5
5	63	71 a 1	63/68a-1	İmam	kitap var	Abdurrahman	Efendi/el-Hac	Cağfer Ağa Mahal	Cağfer Ağa Mahallesi İm	174,440	1	800	1
6	79	49 b 3	79/47b-3	İmam	kitap var	Mustafa b. Mehı	Efendi/el-Hac	Hüseyin Ağa Mah	x	322,960	6	3,180	1
7	79	61 a 2/62	79/60a-2	İmam	kitap var					38015	3	420	2
8	82	69 b 4	82/66b-4	İmam/Attar	kitap var	Hasan b. el-Hac	Efendi			20,482	10	5,785	25
9	84	71 a 2	84/70a-2	İmam	kitap var	Ali b. Ali	Efendi/Hafiz/el-	Hobyar Mahallesi	Meydancık Camii	42,515	18	18,670	44
10	92	47 b 1	92/43b-1	İmam	kitap var	Osman b. Ali	Efendi	Tavaşi Süleyman /	Tavaşi Süleyman Ağa Ca	190,190	13	11,245	6
11	98	26 b 3	98/24b-3	İmam	kitap var	Mahmud Said b.	Efendi	Hekim Çelebi Tekl	Karabaş Mahallesi İmam	554,670	26	6,568	2
12	98	76 a 2	98/75a-2	İmam	kitap var	İbrahim b. Mehı	Efendi/Hafiz/el-	Keçeci Piri Mahall	Sultan Camii İmamı	53,859	51	36,299	68
13	104	57 a 2/58	104/54a-2	İmam/Sahaf	kitap var	Mehmed b. Abd	Efendi/el-Hac	Uzun Şüca Mahall	Peykhane İmamı	57,009	6	6,103	11
14	106	56 b 2	106/53b-2	İmam	kitap var	Yahya b. Mehm	Efendi/el-Hac	Molla Hüsrev Mal	x	150,833	26	8,101	6
15	106	64 b 2	106/61b-2	İmam	kitap var	Veliyüddin b Ab	Efendi	Mimar Sinan Mah	Hafiz Ahmed Paşa Camii	144,150	24	8,950	7
16	107	38 a 3	107/36a-3	İmam	kitap var	Osman b. Hasan	Efendi/Seyyid	Karaki Mahallesi	Çavuş Katibi Amiri Hazr	5,961	6	810	14
17	110	89 b 1	110/85b-1	İmam	kitap var	Salih b. Salih	Efendi	Asmalı Mescid Mı	x	141,179	37	23,009	17
18	112	31 b 5	112/28b-5	İmam	kitap var	Mehmed b. Mah	Efendi/el-Hac	Molla Hüsrev Mal	... İmamı	52,776	27	9,710	15
19	112	94 a 3/95	112/94a-3	İmam	kitap var	Mehmed b. Ali	Efendi	Arab Mahallesi	Arab Camii İmamı	593,699	8	3,560	1
20	115	53 b 4	115/57b-4	İmam	kitap var	Süleyman b. Ha	Efendi/el-Hac	Arabacılar Mahall	x	129,967	3	1,445	2
21	122	15 b 2	122/12b-2	İmam/Kağıd	kitap var	Mustafa b. Ali	Efendi/el-Hac	Keçi Hatun Mahal	Keyci Hatun Camii İman	250,060	20	5,665	3
22	125	64	125/64b-2	İmam	kitap var	Ahmed bin Meh	x	Molla Şeref mahal	x	338,298	18	4,796	2
23	127	87 b 2	127/84b-2	İmam	kitap var	Ahmed b. İbrahi	Efendi	istanbulda avrat pa	x	16,056	13	7,537	47
24	127	80 b 2	127/78b-2	İmam	kitap yok	El-Hac Ahmed b	Efendi/el-Hac	Balaban Ağa Mah	Balaban Ağa Mahallesi n	25,282	x	x	x
25	129	35	129/34a-2	İmam	kitap var	El-Hac Mehmed	Efendi/el-Hac	Silivrikapısı kapısı	Veledi Karabaş mahallesi	28,055	28	6,714	24
26	129	44 b 3	129/42b-3	İmam/Saray	kitap var	Mehmed b. Ali	x	Hadika-i hâssa-i şe	hadika-i mezburede iki .	13,657	16	1857	14

2.2. Data Analysis Process

Primary filtering was first performed on the generated Excel spreadsheet to isolate the inheritance records containing books as a distinct group. Instead of losing time with manual calculations, AI-supported Python scripts were utilized to observe the correlations between variables more rapidly and accurately.

2.2.1. The Book-Wealth Relationship

With the help of Python codes executed over the Excel data, a comparison was drawn between the total value of cash and household goods in the estates and the monetary value allocated to books. By sorting and grouping the numerical data in ascending order, the link between the imams' economic power and their library assets was substantiated.



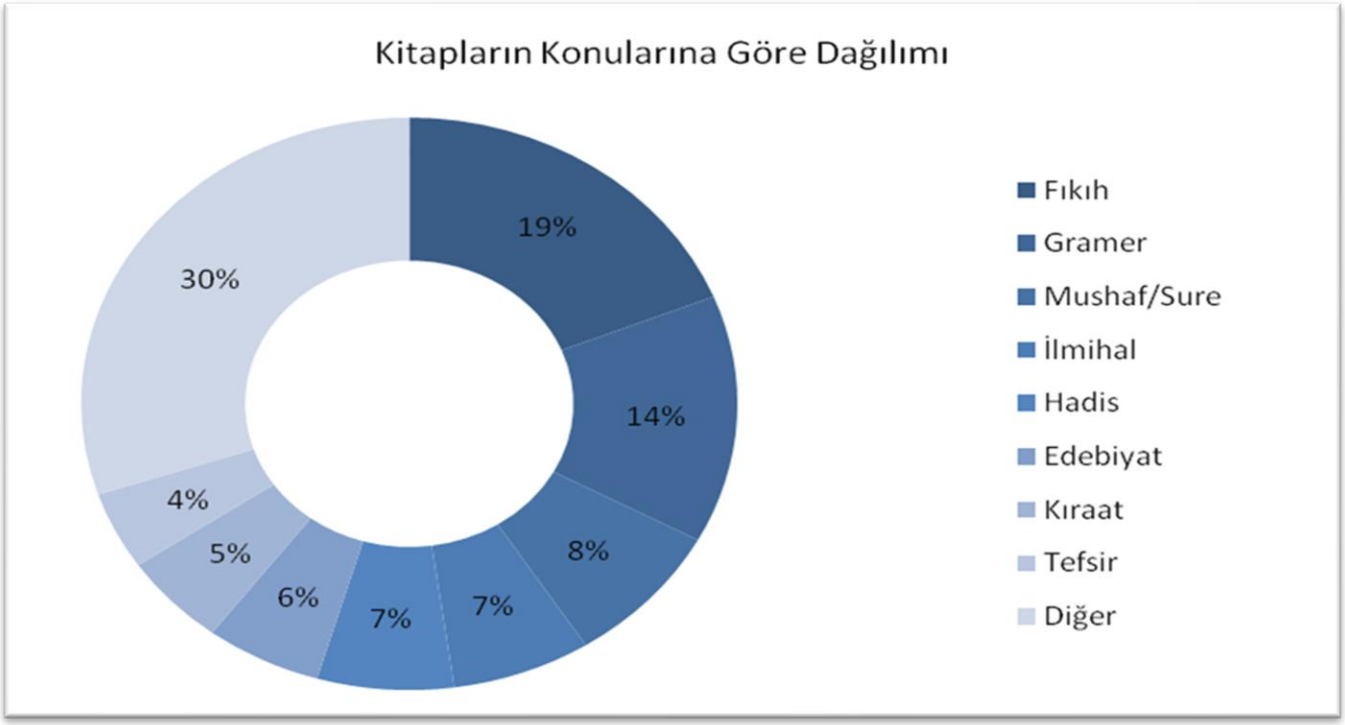
Shape 1. *Correlation Between Books and Wealth*

The relationship between wealth and book ownership among imams is visualized in the graph below using two distinct lines. The blue line represents the number of imam estates corresponding to specific wealth intervals, while the orange line indicates the average number of books per estate within the defined scale. This graph provides a more balanced curve by excluding outliers. While an upward trend in the number of books is observed as wealth increases, this rise becomes significantly more pronounced in estates valued at over 1,000,000 aspers (akçe).

The graph above, which includes the extensive collections of İmam-ı Sultani Mehmed Emin Efendi b. Salih b. Mustafa and Bezzâzistan İmamı el-Hac İshak b. Abdullah, demonstrates that the number of books increases dramatically in estates above a certain wealth threshold. This situation reveals that economic means are directly correlated with knowledge acquisition and book collecting. It is observed that in estates with a wealth exceeding 500,000 akçe, the number of books surpasses 100, and in some cases, exceeds 150.

2.2.2. Thematic Distribution of Books in the Estates

It was not possible to identify the subject of every single one of the 1,617 books in the dataset. Therefore, the analysis was conducted only on the 989 books whose subjects were determined. These books were classified under thematic headings such as jurisprudence (fiqh), literature, lexicography (lughat), and sufism (tasawwuf); subsequently, their proportions both within the overall total and on an estate-by-estate basis were calculated.



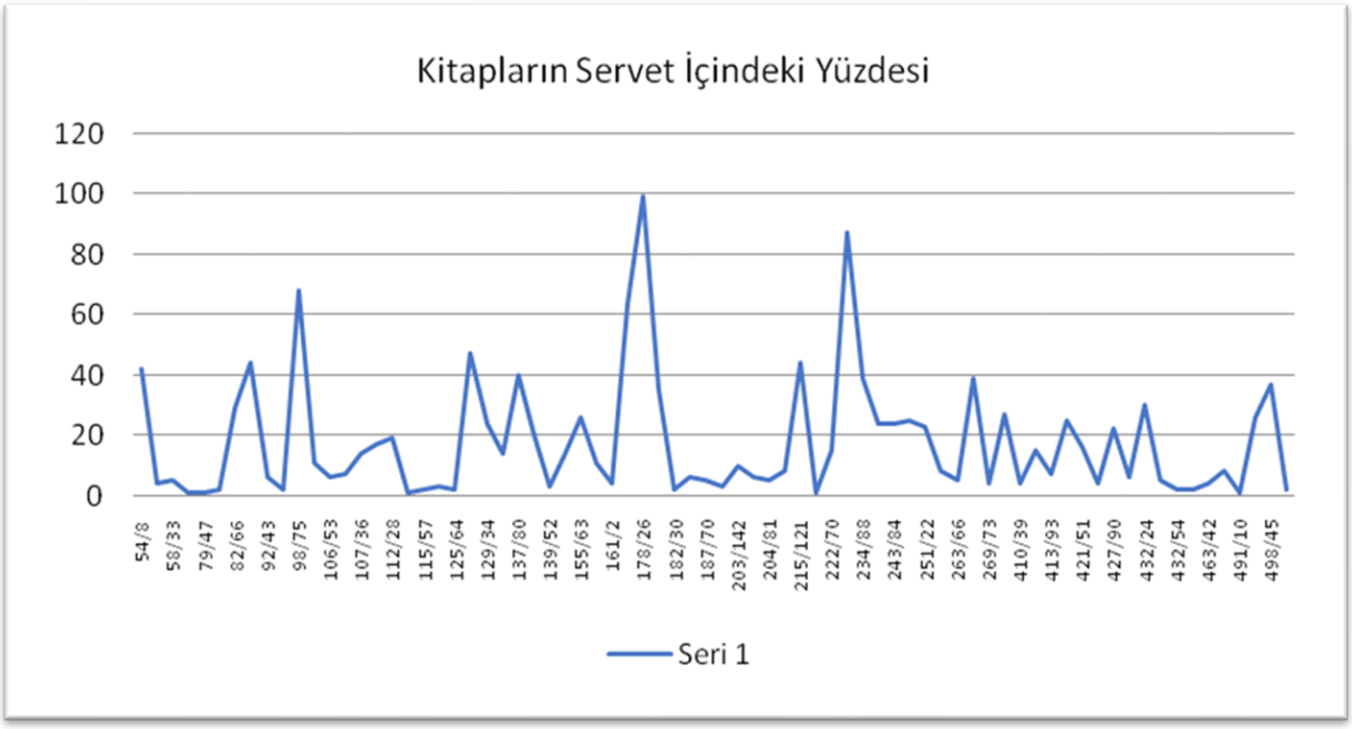
Shape 2. *The Subject Distribution of Books in Estate Inventories*

Works on jurisprudence (fiqh) occupy the largest share in the inheritance records of imams. A total of 199 jurisprudence books constitute 19,18% of the entire collection. Furthermore, the presence of jurisprudence books in 49 different imam estates clearly demonstrates the central position of this field in their professional practice. These are followed by books on Arabic grammar. The 123 identified grammar books are found in the estates of 41 different imams, corresponding to 14% of the total books. In third place are the Mushaf (copies of the Qur'an) and works related to the Holy Qur'an. In addition to 63 Mushafs, 12 *En'ām-ı Şerif*, 2 *Yasin-i Şerif*, and various *juz'* (sections of the Qur'an, such as the *Amme juz'*) were identified. While this group accounts for 8.01% of the total books, it holds the distinction of being the most frequently encountered type of work across the estates.

The fact that imams serve as the primary point of reference for the daily religious matters of the community explains why catechisms (ilmihal) also hold a significant place in the records. Indeed, a total of 67 catechism books constitute 7% of all books and are present in 38 different imam estates. In addition, literary works account for 7% (56 volumes) of the total and are observed in 28 estates. Works related to the science of recitation (qira'at) number 51, holding a share of 5.17%, and are recorded in 30 different estates.

2.2.3. The Percentage of Books within Total Wealth

Book prices were analyzed both within the overall total and on a thematic basis. Through these analyses, the share of libraries within the total wealth was calculated, thereby revealing the budgetary ratios allocated by imams to their collections.



Shape 3. *Books' Share of Total Wealth (%)*

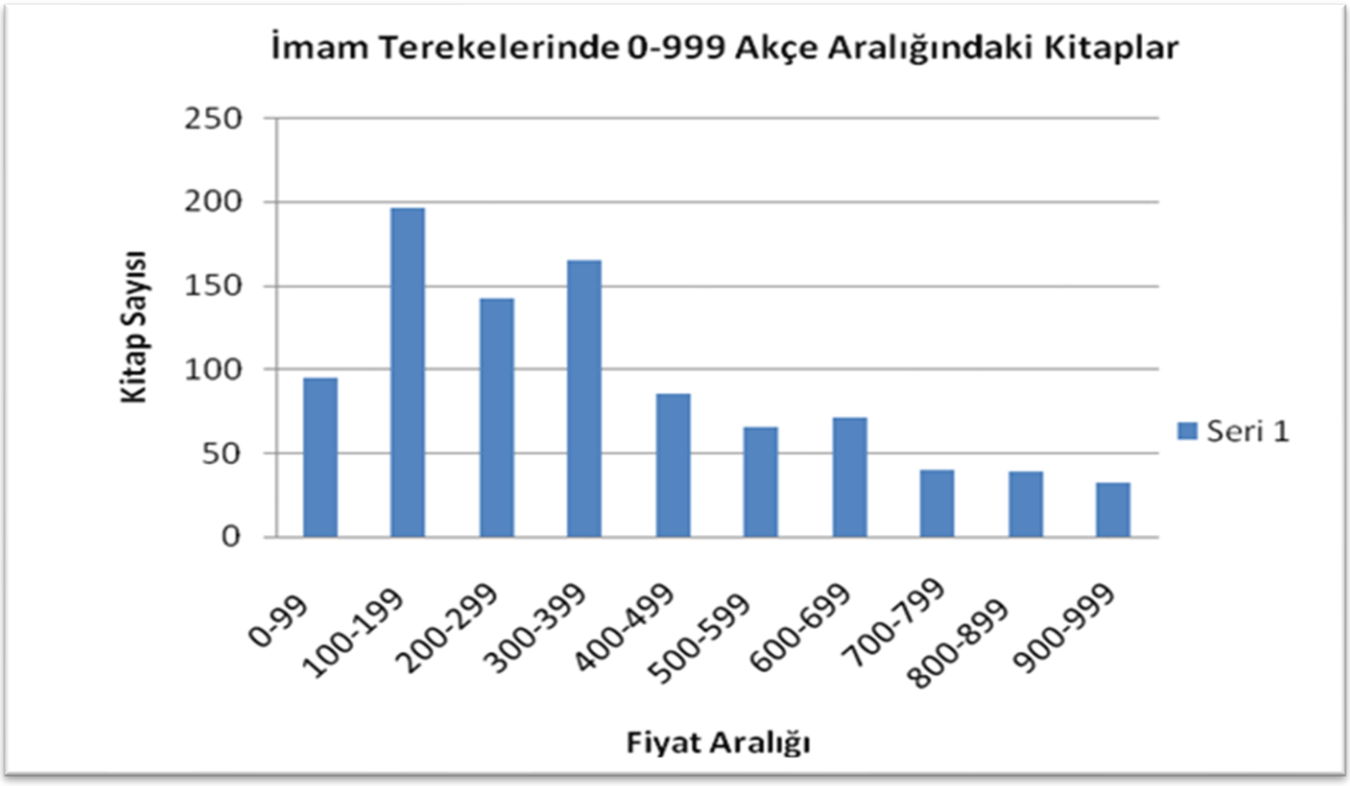
The sharp fluctuations in the graph above clearly demonstrate the significant disparities among imams regarding book ownership. This suggests that the opportunities for imams to acquire books and accumulate intellectual capital varied substantially based on individual, economic, and institutional factors. While some imams possessed extensive and valuable collections, others were only able to access a very limited number of books or works of lower monetary value. This situation indicates that imams did not constitute a homogeneous group within the socio-economic structure of the period; on the contrary, they possessed highly diverse socio-economic profiles[1].

2.2.4. Book Prices

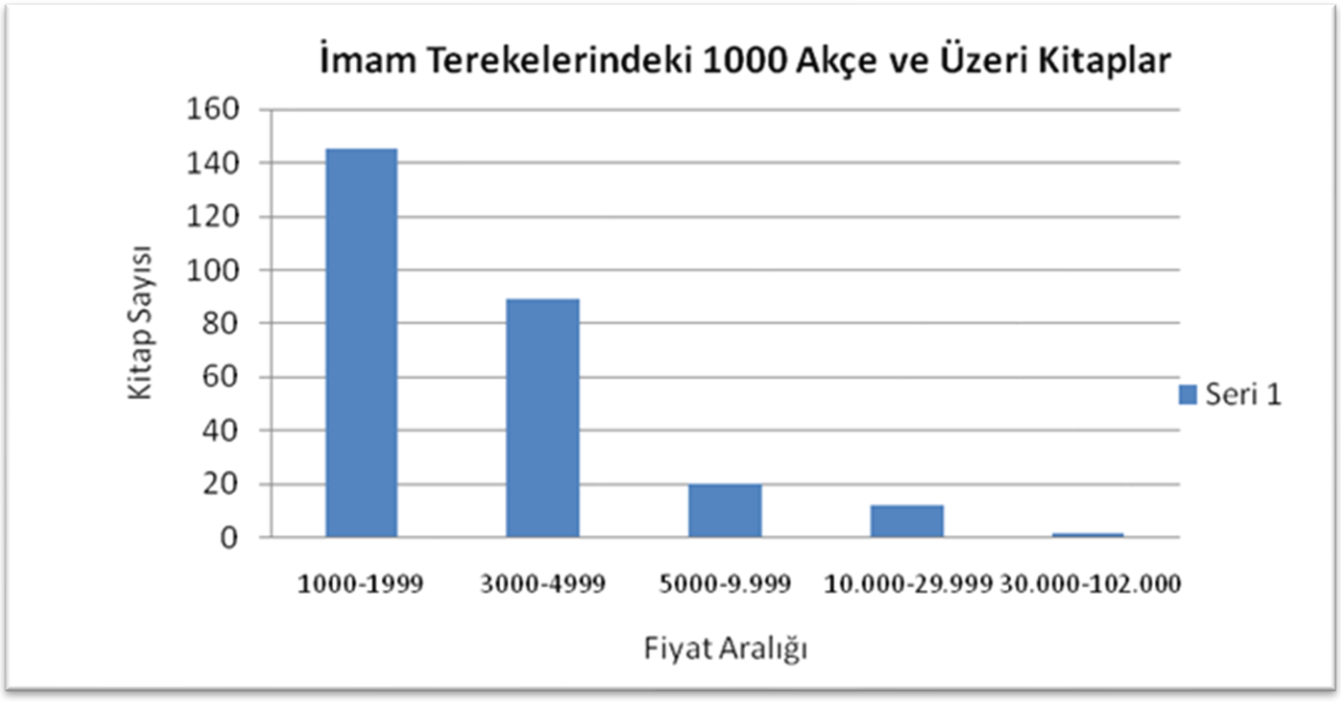
Book prices constitute the most labor-intensive and dense data group to process manually. A total of 1,617 price data points were sorted and automatically scaled using Python. These prices were analyzed separately on both a general and a thematic basis; however, only the general analyses are included in this paper.

According to the results obtained from the analysis, the books were visualized in two different graphs: those valued below and above 1,000 akçe. Based on the data, the vast majority of the books are valued at less than 500 akçe.

[1] To ensure the clarity of the graph, two inheritance records containing an exceptionally high number of books were excluded as outliers.



Şekil 4. *Book Prices Between 0-999 Akçe*



Shape 5. *Book Prices Above 1000 Akçe*

The graph above illustrates the numerical distribution of books valued at 1,000 akçe and above. As can be inferred, there are 145 books in the 1,000–1,999 akçe range, 89 books in the 2,000–5,000 akçe range, 20 books in the 5,000–10,000 akçe range, and

12 books in the 10,000–30,000 akçe range. Additionally, there is one book valued at 102,000 akçe.

Conclusion

This study demonstrates that the integration of artificial intelligence and Python-based analytical tools into Ottoman archival research transcends a mere technical upgrade; it represents a fundamental methodological shift. By transforming the "unstructured" data of 18th-century imam inheritance records into a structured digital dataset, it has been possible to analyze 1,617 books and their socio-economic contexts with a speed and precision unattainable by traditional manual methods.

The findings reveal that imams were not a socio-economically homogeneous group. The sharp fluctuations in book ownership and the correlation between total wealth and library size-particularly the dramatic increase in collections within estates valued over 1,000,000 akçe-highlight the diversity of their intellectual profiles. Furthermore, the thematic dominance of jurisprudence (fiqh) and Arabic grammar underlines the professional and practical scholarly foundations of the imamate in Ottoman society.

In conclusion, the transition from "traditional archives" to "archive-based data centers" allows for the reconstruction of collective mentalities and social networks through big data. This algorithmic approach not only preserves the historical record but also revitalizes it as a dynamic source of processable knowledge, providing a scalable model for future Digital Humanities projects in the field of Ottoman Studies.

Sources

T.C. Cumhurbaşkanlığı Devlet Arşivleri Başkanlığı (Presidency of the Republic of Türkiye, Directorate of State Archives)

Kısmet-i Askeriye Defterleri

54, 58, 63, 79, 79, 82, 84, 92, 98, 104, 106, 107, 110, 112, 115, 122, 125, 127, 129, 137, 138, 139, 143, 150, 155, 158, 161, 166, 178, 179, 181, 182, 185, 187, 192, 198, 198, 203, 204, 209, 215,222, 224, 234, 243, 251, 263, 263, 269, 410, 413, 413, 421, 427, 432, 463, 491, 496, 498, 631.

Lıterature

Abacı, D. Z. (2015). *"Batmayacak kadar bağlantılı ya da güçlü olmak": Kadı sicillerine göre Osmanlı toplumunda sosyal ağlar ve aktörler (1695-1700)* (Proje No. 112K270). TÜBİTAK.

Akça, S., & Uğur, Y. (2023). Osmanlıca metinlerin dijital/otomatik transkripsiyonu. Y. Uğur (Ed.), *Dijital beşerî bilimler ve Osmanlı çalışmaları* içinde (ss. 363-390). VakıfBank Kültür Yayınları.

Argıt, B. İ. (2023). Bir sosyal ağ analizi uygulaması: 17. yüzyılın ikinci yarısında Osmanlı saray mensuplarının borç alacak ilişkileri. Y. Uğur (Ed.), *Dijital beşerî bilimler ve Osmanlı çalışmaları* içinde (ss. 145-170). VakıfBank Kültür Yayınları.

Atçıl, A. (2020). *Yüksek Osmanlı ulemasının mesleki ve entelektüel ilişki ağları ve gruplaşmaları (1470-1650)* (Proje No. 116K404). TÜBİTAK.

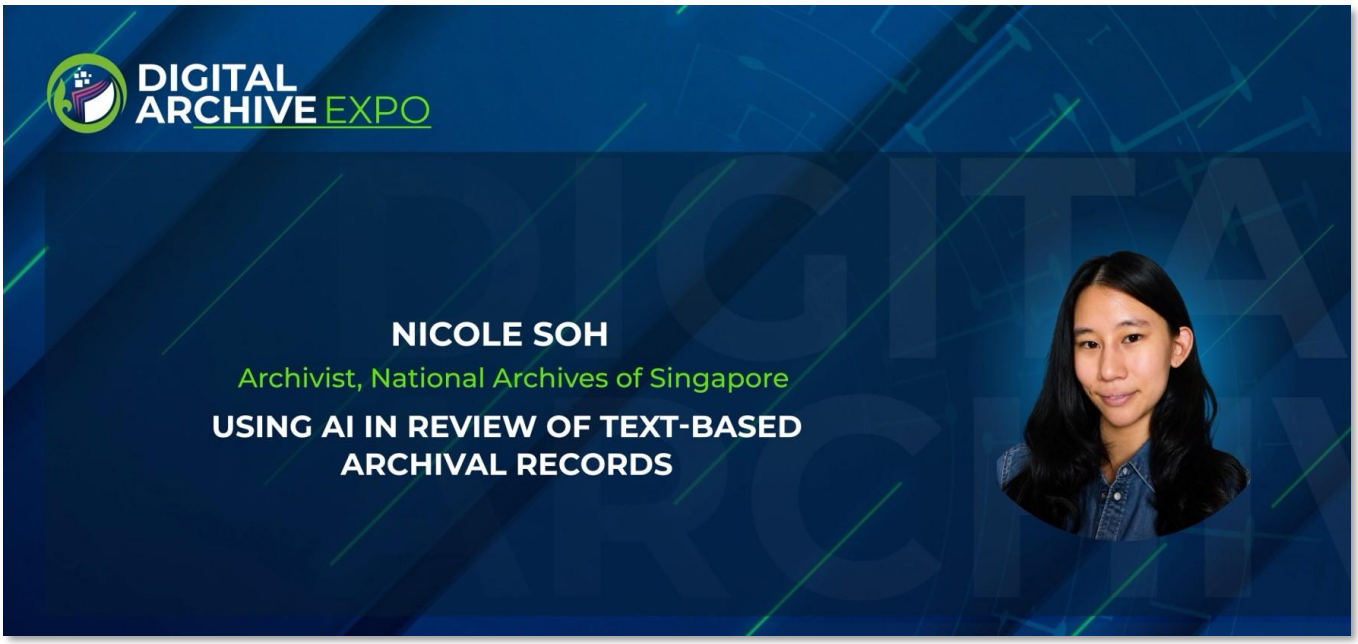
Baş, G. (2023). Dijital beşerî bilimler ve geçmişin haritalandırılması. Y. Uğur (Ed.), *Dijital beşerî bilimler ve Osmanlı çalışmaları* içinde (ss. 49-80). VakıfBank Kültür Yayınları.

Beydilli, K. (2018). *Bir imamın günlüğü*. Pınar Yayınları.

Bozkurt, F. (2021). Osmanlı dönemi tereke defterleri ve tereke çalışmaları. *Trakya Üniversitesi Edebiyat Fakültesi Dergisi*, 11(22), 202-230.

Dögüş, N. (2021). *Osmanlı Devleti'nde imamlık (19. yüzyıl İstanbul örneği)* [Doktora tezi, Marmara Üniversitesi]. YÖK Ulusal Tez Merkezi.

Öztürk, H. (2021). Arşivler ve yapay zeka. *Bilgi Yönetimi Dergisi*, 4(2), 285-305.



USING AI IN REVIEW OF TEXT-BASED ARCHIVAL RECORDS

Nicole Soh
Master of Science (Information Studies), Nanyang Technological University, Archivist (Records Management - Policy and Collection), National Archives of Singapore
nicole_soh@nlb.gov.sg

Abstract. The National Archives of Singapore (NAS) facilitates public access to archival records in its custody, ranging from government files, photographs and audiovisual recordings. At present, for public to access text-based government files, government agencies across Singapore have to review file contents first to assess any sensitivity, which can be time-consuming. This presentation outlines NAS' exploration of AI tools developed by the Government Technology Agency of Singapore (GovTech), the lead agency driving Singapore's Smart Nation initiative and public sector digital transformation. The GovTech AI tools used are AiSay2 and AIBots, AiSay2 is used to transcribe files into machine readable format and AIBots is used to perform automated content analysis, sensitivity assessment, and summarisation of the government files.

While NAS's testing has revealed time-saving benefits in reviewing files, there are limitations to how much AI can be used including how human intervention is still required to verify the AI's recommendations. As such, this illustrates AI's potential to transform the page-by-page review process of text-based records, offering practical solutions to longstanding challenges whilst highlighting the continued importance of professional human judgement when using emerging technology.

Keywords: National Archives of Singapore; AI; public access; sensitivity assessment; government records

МӘТІНДІК АРХИВТІК ҚҰЖАТТАРДЫ САРАПТАУДА ЖАСАНДЫ ИНТЕЛЛЕКТТІ ҚОЛДАНУ

Николь Со
Ғылым магистрі (ақпараттық зерттеулер), Наньян технологиялық университеті, архивист (құжаттарды басқару – саясат және қорларды толықтыру), Сингапур Ұлттық архивтері
nicole_soh@nlb.gov.sg

Аңдатпа. Сингапур Ұлттық архиві (NAS) өз қорында сақталған архивтік құжаттарға, оның ішінде мемлекеттік құжаттарға, фотоматериалдарға және аудиовизуалды жазбаларға жұртшылықтың қолжетімділігін қамтамасыз етеді. Қазіргі уақытта мәтіндік мемлекеттік құжаттарды пайдалануға ұсынбас бұрын, Сингапурдың мемлекеттік органдары олардың мазмұнын құпиялылық пен сезімталдық тұрғысынан алдын ала сараптап, бағалауы тиіс. Бұл үдеріс едәуір уақыт пен еңбек шығынын талап етеді. Мақалада Сингапур Үкіметінің технологиялар агенттігі (GovTech) әзірлеген жасанды интеллект құралдарын енгізу тәжірибесі қарастырылады. GovTech – «Ақылды ұлт» (Smart Nation) бастамасын және мемлекеттік сектордың цифрлық трансформациясын іске асыруға жауапты жетекші мемлекеттік ұйым. Зерттеу барысында қолданылған AiSay2 жүйесі архивтік құжаттарды машинада оқылатын форматқа түрлендіруге пайдаланылса, AIBots жүйесі мемлекеттік құжаттардың мазмұнын автоматты түрде талдауды, сезімтал ақпаратты бағалауды және қысқаша мазмұнын жасауды жүзеге асырады.

Жүргізілген сынақ нәтижелері жасанды интеллектті пайдалану құжаттарды сараптау уақытын едәуір қысқартатынын көрсетті. Сонымен қатар зерттеу нәтижелері жасанды интеллектті қолданудың белгілі бір шектеулері бар екенін де айқындады, өйткені оның ұсынымдарын тексеру және түпкілікті шешім қабылдау үшін маманның қатысуы әлі де қажет. Осылайша, жасанды интеллект мәтіндік архивтік құжаттарды парақтап сараптау үдерісін түбегейлі жаңғыртуға мүмкіндік беретін тиімді құрал ретінде қарастырылғанымен, жаңа технологияларды қолдану жағдайында кәсіби сарапшылық бағалаудың маңызы сақталатыны көрсетіледі.

Түйін сөздер: Сингапур Ұлттық архиві; жасанды интеллект; архивтік құжаттарға қолжетімділік; сезімтал ақпаратты бағалау; мемлекеттік құжаттар

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ ЭКСПЕРТИЗЕ ТЕКСТОВЫХ АРХИВНЫХ ДОКУМЕНТОВ

Николь Со
Магистр наук (информационные исследования), Наньянский технологический университет, архивист (управление документами – политика и комплектование фондов), Национальные архивы Сингапура
nicole_soh@nlb.gov.sg

Аннотация. Национальный архив Сингапура (NAS) обеспечивает общественный доступ к находящимся на его хранении архивным документам, включая государственные дела, фотографии и аудиовизуальные записи. В настоящее время для предоставления доступа к текстовым государственным документам государственные учреждения Сингапура должны предварительно проводить экспертизу содержания документов с целью выявления возможной чувствительной информации. Данный процесс может занимать значительное количество времени.

В статье рассматривается опыт изучения и применения инструментов искусственного интеллекта, разработанных Агентством правительственных технологий Сингапура (GovTech), которое является ведущей организацией в реализации инициативы «Умная нация» (Smart Nation) и цифровой трансформации государственного сектора. В исследовании анализируются два инструмента искусственного интеллекта GovTech – AiSay2 и AIBots. Система AiSay2 используется для преобразования документов в машиночитаемый формат путем транскрибирования, тогда как AIBots применяется для автоматизированного анализа содержания, оценки наличия чувствительной информации и подготовки кратких резюме государственных документов.

Результаты тестирования, проведенного Национальным архивом Сингапура, показали, что использование искусственного интеллекта позволяет сократить время, необходимое для экспертизы документов. Вместе с тем возможности применения ИИ имеют определенные ограничения, поскольку для проверки рекомендаций, сформированных искусственным интеллектом, и принятия окончательных решений по-прежнему требуется участие человека. Таким образом, искусственный интеллект обладает значительным потенциалом для преобразования процесса постраничной экспертизы текстовых документов, предлагая практические решения для давних проблем архивной практики, при этом подчеркивая сохраняющуюся важность профессионального человеческого суждения при использовании новых технологий.

Ключевые слова: Национальный архив Сингапура; искусственный интеллект; общественный доступ; оценка чувствительной информации; государственные документы

1. Introduction

The National Archives of Singapore (NAS) is the custodian of Singapore's collective memory. As an institution of the public service, NAS plays an important role in facilitating public access to archival records in its custody, ranging from text-based files to photographs and audiovisual recordings.

Today, access to government records can be requested by the public via NAS's Archives Online platform. Requestors may request for records for various reasons such as for research, genealogy or general interest in the nation's history especially stemming from recent high-profile exhibitions on Singapore's nation building years.

In February 2023, the Singapore government announced its decision to declassify and publicly release the Albatross file, a collection of Cabinet papers and handwritten notes kept by founding deputy prime minister Dr Goh Keng Swee documenting the top-level negotiations that preceded Singapore's separation from Malaysia in 1965 (Ong 2023). Two years later, in 2025, as a capstone exhibition for Singapore's 60th year of independence (SG60), The Albatross File: Singapore's Independence Declassified, featuring declassified documents from the file, opened at the National Library Building in December 2025, with 1,500 people registering for opening day (Koh 2025). Also in celebration of SG60, the National Museum of Singapore opened Once Upon A Tide, an exhibition tracing the nation's development through its waterways and drawing on artefacts, photographs, and documents spanning centuries of history (Tham 2025). Together, these exhibitions reflect a broader movement and emphasis where archival records play a role in presenting national narratives to Singaporeans.

From such publicity, NAS has also seen interest in text-based archival records grow as between 2020 and 2025 where public demand for text-based archival records

grew by two and a half times. Over the same period, the number of text-based records made available for online request expanded also from approximately 190,000 in 2019 to over 700,000 in 2026. Though the demand has surged, the review process which consists of the work of reading through files, assessing their sensitivities, and determining their declassification status has not kept pace with this growth. The result is a labour-intensive bottleneck that risks creating a backlog of public requests.

This paper documents NAS's exploration of artificial intelligence tools to speed up the labour-intensive process of text-based file reviews. It describes the tools used, the testing carried out, the findings from testing, and the lessons drawn from practice.

2. Background and Context

2.1. The National Archives of Singapore

NAS was established under the National Archives and Records Centre Act of 1967 and has operated since 1968 (National Archives and Records Centre Act 1967). Since 2013, NAS has been part of the National Library and Archives Board (NLB). As the collector and keeper of government records, NAS is responsible for the collection, preservation, management and access of Singapore's public and private archival holdings (National Archives of Singapore n.d.).

2.2. The Challenge of Text-Based File Review

The review of text-based archival files is one of the many resource-intensive tasks that reviewers across WOG undertake. A single large file can take between one and three working days to read through and assess, which is a significant investment of time, especially when multiplied across the thousands of files awaiting review. This is compounded by the fact that file review is rarely the core responsibility of the officers tasked with it. Reviewers are often subject matter experts or senior officers within their agencies, where declassification work sits in competition with their core work. The result is that reviews may be deferred due to limited manpower.

Another challenge for file review is the diversity of document types that can be found in archival records. Text-based files include handwritten records, typewriter-written documents, computer-generated files, and raw scans of varying quality. Each of these presents different challenges for review, handwritten documents require careful reading and interpretation, poor-quality scans may be partially illegible and the sheer volume of files to review driven by increasing public demand means that reviewers must find ways to work more efficiently without sacrificing the file review quality.

This is why NAS began exploring the potential of AI to support the file review process. This exploration started in tandem with Singapore's broader national commitment to AI adoption across the public sector, articulated recently in the National AI Strategy 2.0 (NAIS 2.0), and launched in 2023 (Smart Nation Singapore 2023). NAS's exploration of AI tools for text-based file reviews would tap on AI tools developed by the GovTech and made available to public agencies as part of Singapore's whole-of-government approach to digital transformation.

3. The AI Tools: AiSay2 and AIBots

NAS's approach to AI-assisted file review for text-based documents is structured as a two-step process. The first step addresses the challenge of document transcription while the second addresses the challenge of content assessment. They form a pipeline that takes a scanned archival file as its input and produces a sensitivity assessment as its output for a reviewer's final assessment.

The first step uses AiSay2, a GovTech AI Document Parser that performs AI-powered transcription of PDF image files (GovTech 2026a). By first performing image correction (such as contrast brightening, de-skewing, etc.) and then using a mix of OCR programs and AI models to fill in any gaps predictively, AiSay2 converts scanned

documents into machine-readable text. This transcription step is a key step, without it, different text types found in the files would make automated content assessment difficult. By converting scanned images into text, AiSay2 makes the content of a file accessible to downstream AI programs for content analysis.

The second step uses AIBots, a GovTech platform that enables public agencies to create and share customised AI assistants for specific use cases (GovTech 2026b). In NAS's case, the platform was used to build a declassification assistant, the bot is configured through customised instructional prompts that define its role specific to an agency's historical context and review framework, its reasoning process on what sensitivities to consider, and the format of its outputs. The bot is also enabled by Retrieval-Augmented Generation (RAG), meaning that it can draw on an uploaded knowledge base of reference documents, such as NAS's declassification guidelines, when formulating its assessments. This knowledge base is what allows the bot to produce outputs that are relevant to the agency's particular review framework rather than generic AI responses. Reviewers interact with the bot through a user-facing chat interface, which allows them to not only to receive the bot's initial assessment but to ask follow-up questions, seek clarification, and reason through the file's content in conversation with the AI.

4. Testing Process

For testing, the AIBot was configured specifically to review a particular agency's records. This meant that the instructional prompts were designed with a particular covering period and agency in mind. The instructional prompt also utilises various prompt engineering techniques, such as chain-of-thought prompting where the bot was instructed to take a step wise approach to first assess the quality of transcription of the file, and if the quality was above a certain threshold, to proceed to summarise the content of the file, identify any sensitive material, and recommend a declassification position based on NAS's established guidelines. To further calibrate the bot's responses in-context learning via shot-based prompting was also used to guide the AI on logical reasoning for more complex sensitivity criteria. The prompt attempts to define not only what the bot should do but how it should reason so that its output will be evidence-based and transparent to the reviewer.

The testing process followed the two-step process. Files were first processed through AiSay2 to produce transcribed text, which was then submitted to the AIBots for sensitivity assessment and summarisation. To mirror actual test cases, the range of files tested covered the range of handwritten records, typewriter-written documents, computer-generated files, scans of varying quality, and files of few and many pages. The AI's outputs were then compared and evaluated with the human reviewer's assessments and judgements previously made for the same file, eventually forming the findings outlined in the next section.

5. Strengths and Limitations

The testing process revealed a set of strengths and limitations which carry implications for how such AI tools should be understood and used.

On the strengths side, the most immediate benefit was speed of comprehension. The AI summaries enabled reviewers to develop a quicker understanding of a file's content, reducing the time spent reading through documents page-by-page before forming an initial assessment. The bot also demonstrated a useful consistency in its

application of sensitivity categories, assessing each file against the same framework without the variation that can arise with different human reviewers working under time pressure. The chat interface proved to be a meaningful feature where reviewers were able to ask the bot questions to clarify its reasoning, query specific passages, and reason in a more interactive way versus a static report. The AI also showed an ability to pick up on tone and language within files which might not be immediately apparent to a reviewer skimming through dense text.

The limitations, however, are equally important to understand. The quality of the AI's content assessment is directly dependent on the quality of the transcription produced in the first step. When AiSay2 struggles with a document due to poor scan quality, illegible handwriting, or faded ink, the downstream assessment by AIBot will be also be more unreliable, in such cases the bot would recommend full human intervention for the file review instead. The bot also demonstrated some tendency to hallucinate, generating plausible sounding but inaccurate content in its attempt to produce complete and coherent responses. Related to this, the bot's responses showed a degree of oversensitivity to strong vocabulary, flagging documents as potentially sensitive citing forceful language even where the content did not warrant concern. Such false positives could add to a reviewer's file review burden as extra verification and discernment would be needed. Finally, the use of AI for highly classified files presents a structural challenge as such AI applications would need to be hosted on local secured servers to meet the necessary security requirements, making them considerably more expensive to develop and sustain over the long term.

6. Conclusion

The NAS experience with GovTech AI tools, AiSay2 and AIBots, offers a practical account of what AI can and cannot do in the context of text-based file reviews. While the AI tools can deliver benefits such as speeding up comprehension and supporting consistent assessment. They are best understood as decision-support tools for reviewers rather than autonomous decision-makers, offering a first cut understanding the file but not replacing the quality of human judgement.

A less visible but meaningful benefit from the usage of AI tools is that configuring these tools and writing effective instructional prompts require agencies to surface and articulate the knowledge that experienced reviewers have long applied intuitively. The use of AI tools in file review bring to focus how important the institutional practices are rather than the technical capabilities of AI. How the archives and agencies across WOG structure and create their review frameworks and even more importantly how they train their staff in a rapidly evolving AI landscape matters more than the choice of AI tools to use.

Lastly, the human must remain in the loop, as the review of archival records is not just merely making historical records accessible, but is also tied to higher stakes such as transparency and public trust in the government and should not be delegated to machines without human oversight.

References

GovTech 2026a – Government Technology Agency of Singapore. AiSay2: AI Document Parser. https://aisay.ai.tech.gov.sg/ (accessed 01.05.2026).

GovTech 2026b – Government Technology Agency of Singapore. AIBots: Customise GenAI Chatbots with Your Documents. https://aibots.gov.sg (accessed 01.05.2026).

Koh 2025 – Koh M.L. Albatross File exhibition draws crowds on opening day. https://www.straitstimes.com/singapore/albatross-file-exhibition-draws-in-crowds-on-opening-day (accessed 01.05.2026).

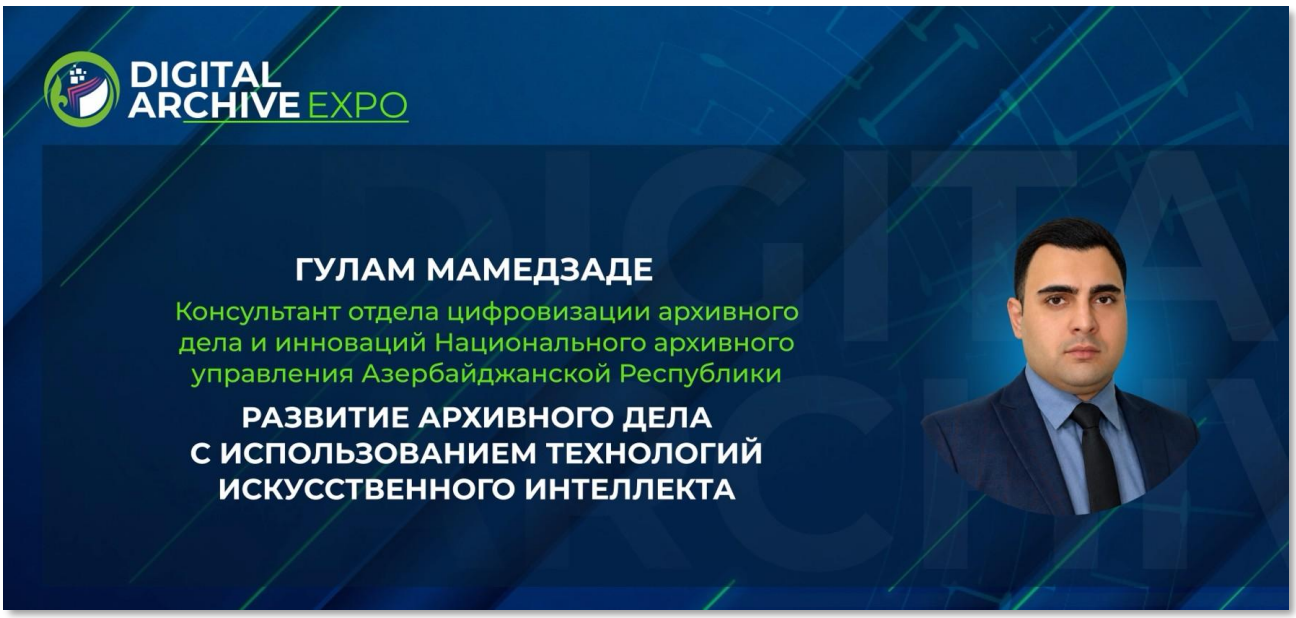
National Archives and Records Centre Act 1967 – National Archives and Records Centre Act, Cap. 193 (1967). Singapore.

National Archives of Singapore n.d. – National Archives of Singapore. Archives Online. National Library Board. https://www.nas.gov.sg/archivesonline/about-us/ (accessed 01.05.2026).

Ong 2023 – Ong J. 'Albatross file' with details of top-level talks before Singapore separated from Malaysia to be declassified, released publicly. https://www.channelnewsasia.com/singapore/albatross-file-declassified-talks-separation-malaysia-5590921 (accessed 01.05.2026).

Smart Nation Singapore 2023 – Smart Nation Singapore. National AI Strategy 2.0 (NAIS 2.0). Smart Nation and Digital Government Office. https://www.smartnation.gov.sg/initiatives/national-ai-strategy/ (accessed 27.04.2026).

Tham 2025 – Tham D. Sail up Singapore River and back in time at National Museum's exhibition Once Upon A Tide. https://www.straitstimes.com/life/arts/sail-up-the-singapore-river-and-back-in-time-at-national-museums-new-exhibition-once-upon-a-tide (accessed 01.05.2026).



РАЗВИТИЕ АРХИВНОГО ДЕЛА С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Гулам Мамедзаде
Консультант Отдела цифровизации и инноваций архивного дела,
Национальное архивное управление Азербайджанской Республики
q.memmedzade@milliarxiv.gov.az

Аннотация. Доклад посвящен анализу влияния цифровизации и технологий искусственного интеллекта (ИИ) на развитие современного архивного дела, а также рассмотрению международного опыта и перспектив внедрения ИИ в архивные системы Азербайджанской Республики. В работе раскрываются ключевые направления применения ИИ в архивах, включая автоматическую классификацию, распознавание текста (OCR), интеллектуальный поиск, оцифровку и сохранение документов, а также предиктивную аналитику. Отмечаются преимущества использования ИИ: повышение эффективности обработки документов, снижение ошибок, улучшение доступа к архивной информации и возможность работы с большими объемами данных. Рассматриваются примеры внедрения технологий ИИ в национальных архивах Великобритании, США, Франции, а также в проекте Transkribus.

Особое внимание уделено инициативам Азербайджанской Республики по развитию нормативно-правовой базы, цифровой инфраструктуры и институциональных механизмов внедрения ИИ, включая деятельность службы «ASAN» и создание единой электронной архивной информационной системы. Делается вывод о том, что Азербайджан находится на этапе активной цифровизации архивной сферы, формируя технологические и организационные предпосылки для будущего широкомасштабного внедрения технологий искусственного интеллекта в управление документами и развитие национальных архивов.

Ключевые слова: искусственный интеллект; OCR; Transkribus; архивные системы; электронное правительство; Цифровизация архивов; Электронные услуги; служба ASAN; Системы электронного документооборота

АРХИВ ІСІН ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫН ҚОЛДАНУ АРҚЫЛЫ ДАМЫТУ

Мамедзаде Гулам
Әзірбайжан Республикасы Ұлттық архив басқармасының архив ісін цифрландыру және инновациялар бөлімінің кеңесшіс
q.memmedzade@milliarxiv.gov.az

Аңдатпа. Баяндамада цифрландыру және жасанды интеллект (ЖИ) технологияларының қазіргі архив ісінің дамуына ықпалы талданып, сондай-ақ халықаралық тәжірибе мен Әзербайжан Республикасының архив жүйелеріне ЖИ технологияларын енгізу перспективалары қарастырылады. Зерттеуде архивтерде жасанды интеллектті қолданудың негізгі бағыттары, оның ішінде автоматты жіктеу, мәтінді тану (OCR), интеллектуалды іздеу, құжаттарды цифрландыру және сақтау, сондай-ақ болжамды талдау (predictive analytics) мүмкіндіктері ашып көрсетіледі. Жасанды интеллектті қолданудың артықшылықтары ретінде құжаттарды өңдеу тиімділігін арттыру, қателер санын азайту, архивтік ақпаратқа қолжетімділікті жақсарту және үлкен көлемдегі деректермен жұмыс істеу мүмкіндігі атап өтіледі. Ұлыбритания, АҚШ, Франция ұлттық архивтерінде, сондай-ақ Transkribus жобасы аясында жасанды интеллект технологияларын енгізу тәжірибесіне мысалдар келтіріледі.

Ерекше назар Әзербайжан Республикасында жасанды интеллектті енгізудің нормативтік-құқықтық базасын, цифрлық инфрақұрылымын және институционалдық тетіктерін дамыту бастамаларына аударылады. Атап айтқанда, «ASAN» қызметінің қызметі және бірыңғай электрондық архивтік ақпараттық жүйені құру үдерістері қарастырылады. Зерттеу нәтижесінде Әзербайжан архив саласын белсенді цифрландыру кезеңінде екендігі, сонымен қатар құжаттарды басқару және ұлттық архивтерді дамыту бағыттарында жасанды интеллект технологияларын ауқымды енгізуге қажетті технологиялық және ұйымдастырушылық алғышарттар қалыптасып келе жатқаны туралы қорытынды жасалады.

Түйін сөздер: жасанды интеллект; OCR; Transkribus; архивтік жүйелер; электрондық үкімет; архивтерді цифрландыру; электрондық қызметтер; ASAN қызметі; электрондық құжат айналымы жүйелері

DEVELOPMENT OF ARCHİVAL WORK USİNG ARTİFİCİAL İNTELLİGENCE TECHNOLOGİES

Gulam Mammadzade
Advisor, Department of Digitization and Innovation of Archival Work, National Archive Administration of the Republic of Azerbaijan
q.memmedzade@milliarxiv.gov.az

Abstract. This report analyzes the impact of digitalization and artificial intelligence (AI) technologies on the development of modern archival science, as well as examines international experience and prospects for implementing AI in the archival systems of the Republic of Azerbaijan. The paper explores key areas of AI application in archives, including automatic classification, text recognition (OCR), intelligent search, document digitization and preservation, and predictive analytics. The advantages of using AI are highlighted, including increased document processing efficiency, reduced errors, improved access to archival information, and the ability to work with large volumes of data. Examples of AI

implementation in the national archives of the UK, USA, and France, as well as in the Transkribus project, are considered.

Particular attention is paid to the initiatives of the Republic of Azerbaijan to develop the regulatory framework, digital infrastructure, and institutional mechanisms for the implementation of AI, including the activities of the ASAN service and the creation of a unified electronic archival information system. It is concluded that Azerbaijan is at the stage of active digitalization of the archival sphere, forming the technological and organizational prerequisites for the future large-scale implementation of artificial intelligence technologies in document management and the development of national archives.

Keywords: artificial intelligence; OCR; Transcribe; archive systems; electronic government; Digitization of archives; Electronic services; ASAN service; Electronic document management systems

Современное архивное дело переживает значительные изменения под влиянием цифровизации и внедрения технологий искусственного интеллекта (ИИ). Архивы традиционно выполняют функции хранения, систематизации и предоставления доступа к документам, однако рост объема информации требует новых подходов к управлению данными. Использование ИИ позволяет повысить эффективность архивной деятельности, автоматизировать процессы и улучшить доступ к информации.

Искусственный интеллект – это совокупность технологий, позволяющих компьютерным системам выполнять задачи, требующие человеческого интеллекта: распознавание текста, анализ данных, классификация информации и принятие решений. В архивной сфере ИИ применяется для обработки больших массивов данных и оптимизации работы с документами (Artificial Intelligence).

К основным направлениям применения ИИ в архивах можно отнести:

1. Автоматическая классификация и каталогизация

ИИ способен анализировать содержание документов и автоматически присваивать им категории, ключевые слова и метаданные. Это существенно сокращает время обработки архивных материалов.

2. Распознавание текста (OCR)

Технологии оптического распознавания символов позволяют переводить отсканированные документы в редактируемый цифровой формат. Это облегчает поиск и использование архивной информации.

3. Поиск и анализ информации

ИИ улучшает поисковые системы архивов, делая их более точными и быстрыми. Пользователи могут находить нужные документы по смыслу, а не только по ключевым словам.

4. Оцифровка и сохранение документов

ИИ помогает в восстановлении поврежденных документов, улучшении качества изображений и предотвращении потери данных.

5. Предиктивная аналитика

Системы ИИ могут прогнозировать, какие документы будут востребованы чаще, и оптимизировать их хранение и доступ.

Преимущества использования ИИ в архивном деле:

- ускорение обработки документов;

- снижение человеческого фактора и ошибок;
- повышение доступности архивов;
- улучшение качества хранения информации;
- возможность работы с большими объемами данных.

Несмотря на преимущества, существуют и определенные сложности:

- высокая стоимость внедрения технологий;
- необходимость обучения специалистов;
- риски утечки данных и киберугрозы;
- этические вопросы использования ИИ;
- возможные ошибки алгоритмов.

В будущем роль ИИ в архивном деле будет только расти. Ожидается развитие интеллектуальных архивных систем, полностью автоматизированных хранилищ и интеграция архивов с глобальными информационными сетями. Это позволит создать единое цифровое пространство знаний.

Интеграция технологий искусственного интеллекта в архивные системы: международный опыт и перспективы его внедрения в Азербайджанской Республике

1. Национальный архив Великобритании

В The National Archives применяются технологии ИИ для автоматической классификации документов и улучшения поиска. Используются алгоритмы анализа текста, которые помогают быстрее находить нужную информацию среди миллионов записей (The National Archives).

2. Национальный архив США

National Archives and Records Administration (NARA) внедряет ИИ для распознавания рукописных текстов и оцифровки исторических документов. Это особенно важно для старых архивов, где документы сложно читать вручную (National Archives).

3. Архивы Франции

Во французских архивах используются системы машинного обучения для автоматической обработки изображений и восстановления поврежденных документов, что помогает сохранять культурное наследие (International Council on Archives).

4. Проект «Transkribus»

Transkribus – это ведущая ИИ-платформа для распознавания, транскрибирования и поиска в исторических рукописных и печатных документах. Она разработана для преобразования изображений старинных рукописей в цифровой, машиночитаемый текст. В отличие от стандартного OCR, который работает в основном с печатным текстом, Transkribus использует нейронные сети для обучения почерку, что позволяет распознавать сложные соединенные рукописные тексты различных эпох. Эта платформа широко используется архивами и библиотеками в странах ЕС для распознавания рукописных текстов и создания цифровых архивов (Transkribus).

В Азербайджанской Республике формируется комплексный подход к внедрению технологий искусственного интеллекта в систему государственного управления, охватывающий стратегические, институциональные и технологические аспекты. Ключевым направлением является развитие стратегической и нормативно-правовой базы. В частности, утверждение «Стратегии искусственного интеллекта Азербайджанской Республики на 2025–2028 годы» определяет приоритеты государственной политики в данной сфере. Наряду с этим осуществляется разработка и внедрение национальных стандартов, регулирующих применение технологий ИИ (в Азербайджане началось принятие национальных стандартов по применению искусственного интеллекта).

Важным элементом является институционализация процессов внедрения ИИ. Так, предусматривается введение в государственных органах должностей заместителей руководителей, ответственных за внедрение и применение технологий искусственного интеллекта, что направлено на координацию цифровых инициатив и повышение эффективности управленческих процессов (Azərbaycan Respublikasının 2025–2028-ci illər üçün süni intellekt Strategiyası).

Существенное значение имеет использование ИИ-агентов в государственном управлении. Предполагается их применение для автоматизации рутинных операций, оптимизации документооборота, анализа больших массивов данных и поддержки принятия управленческих решений.

Одновременно реализуются меры по формированию современной цифровой инфраструктуры. В частности, разрабатывается концепция «Новой цифровой архитектуры Азербайджана», ориентированная на повышение прозрачности деятельности государственных органов, оптимизацию управленческих процессов и обеспечение интеграции информационных систем (Azərbaycanın yeni rəqəmsal arxitekturası).

Развитие технологий искусственного интеллекта также связано с расширением спектра цифровых государственных услуг. Внедрение интеллектуальных решений способствует ускорению обработки обращений граждан, снижению вероятности ошибок и повышению качества предоставляемых услуг.

Рассмотренные направления внедрения технологий искусственного интеллекта в государственных органах формируют основу для их практической реализации в различных секторах государственного управления. В частности, данные процессы получают развитие в архивной сфере, где цифровизация и интеллектуальные технологии начинают играть все более значимую роль. А именно:

Цифровизация архивов

Государственные архивы постепенно переходят к электронному документообороту. В Национальном архивном управлении Азербайджана осуществляется оцифровка документов и формирование электронных баз данных, что способствует повышению эффективности хранения, систематизации и поиска архивной информации.

Электронное правительство

В рамках платформы службы «ASAN» (предоставление государственных услуг по принципу «одного окна») осуществляется внедрение цифровых услуг, предусматривающих использование элементов интеллектуального анализа данных с целью обработки обращений граждан и обеспечения доступа к архивной информации.

В основе функционирования службы ASAN, находящейся в ведении Государственного агентства по услугам гражданам и социальным инновациям при Президенте Азербайджанской Республики, лежит институциональная модель, предусматривающая участие множества акторов с четким разграничением их компетенций, что направлено на предотвращение дублирования функций и повышение эффективности управления. Служба ASAN не осуществляет непосредственное предоставление услуг, а выступает в качестве организационно-инфраструктурной платформы, обеспечивающей взаимодействие государственных и частных структур в рамках специализированных центров обслуживания граждан. Государственное агентство формирует унифицированную систему стандартов и принципов оказания услуг, включающую прозрачность, эффективность, подотчетность, клиентоориентированность и комфорт, а также осуществляет мониторинг и оценку деятельности уполномоченных структур в соответствии с установленными критериями. Важным направлением деятельности является подготовка кадров и систематическая оценка уровня удовлетворенности получателей услуг.

Реализация инновационных управленческих решений в сочетании с неукоснительным соблюдением норм законодательства при оказании услуг населению – как в стационарных центрах, так и посредством мобильных сервисов – позволила сформировать институциональную среду, минимизирующую предпосылки для коррупции, административных барьеров и злоупотреблений должностными полномочиями. В результате обеспечивается

стабильно высокий уровень удовлетворенности граждан качеством предоставляемых услуг. Деятельность службы ASAN, ставшая примером успешной модернизации системы государственного сервиса, получила международное признание. В частности, она была отмечена Организацией Объединенных Наций и удостоена престижной награды в области государственной службы как одна из передовых практик предоставления государственных услуг (ASAN xidmət).

Автоматизация поиска документов

Создаются системы, позволяющие быстро находить архивные материалы через электронные каталоги. В перспективе возможно внедрение ИИ для семантического поиска.

Перспективы внедрения ИИ

Хотя внедрение ИИ находится на начальной стадии, ожидается:

- использование OCR для старых документов;
- автоматическая классификация архивных материалов;
- создание интеллектуальных поисковых систем;
- интеграция архивов с государственными цифровыми платформами.

Передача документов на постоянное хранение в государственные архивы Азербайджанской Республики регулируется Национальным архивным управлением и проходит этап активной цифровизации, направленной на создание единой электронной архивной информационной системы (Azərbaycan Respublikasında arxiv işinin inkişafına dair 2020-2025-ci illər üçün Dövlət Proqramı).

Основные положения и текущая ситуация

Единая система. В Азербайджанской Республике формируется единая электронная архивная информационная система для модернизации процесса управления документами. Архивные документы, проходящие процесс цифровизации, подлежат регистрации и научному описанию в рамках указанной информационной системы. Одновременно предусматривается, что передача документов на постоянное государственное хранение со стороны источников комплектования будет осуществляться посредством данной системы с обеспечением соответствующих процедур учета и контроля (Milli arxiv fondu haqqında).

Центр электронных документов. В Национальном архивном управлении создается инфраструктура для Центра электронных документов, который будет отвечать за прием и хранение документов в цифровом формате.

Системы электронного документооборота (СЭД), применяемые в деятельности Национального архивного управления, обеспечивают автоматизацию процессов работы с документами и интеграцию с информационными системами иных государственных структур Азербайджанской Республики, что способствует повышению эффективности межведомственного взаимодействия и управления документными потоками.

Электронные услуги. Следующие виды государственных услуг предоставляются в электронном виде: выдача архивных справок, выписок и копий архивных документов, предоставление разрешения на использование архивных документов для научных исследований и других целей, а также

предоставление информации о составе и содержании Национального архивного фонда Азербайджанской Республики (Elektron xidmətlər portalı).

Мировая практика свидетельствует о том, что применение технологий искусственного интеллекта способствует существенному повышению эффективности архивной деятельности и обеспечению сохранности информации в условиях цифровой эпохи. В Азербайджанской Республике уже реализуются процессы цифровизации архивной сферы, создающие предпосылки для возможного внедрения технологий искусственного интеллекта в будущем и повышения доступности, оперативности поиска и качества использования историко-документального и государственного информационного ресурса.

Литература

Artificial Intelligence – Artificial Intelligence: A Modern Approach, 4th US ed. https://aima.cs.berkeley.edu (дата обращения: 13.05.2026).

ASAN xidmət – 4. ASAN xidmət. https://www.asan.gov.az/about (дата обращения: 13.05.2026).

Azərbaycan Respublikasında arxiv işinin inkişafına dair 2020-2025-ci illər üçün Dövlət Proqramı – «Azərbaycan Respublikasında arxiv işinin inkişafına dair 2020-2025-ci illər üçün Dövlət Proqramı»nın təsdiq edilməsi haqqında. https://e-qanun.az/framework/44398 (дата обращения: 13.05.2026).

Azərbaycan Respublikasının 2025–2028-ci illər üçün süni intellekt Strategiyası – «Azərbaycan Respublikasının 2025–2028-ci illər üçün süni intellekt Strategiyası»nın təsdiq edilməsi haqqında Azərbaycan Respublikası Prezidentinin Sərəncamı. https://president.az/az/articles/view/68364 Elektron xidmətlər portalı – Elektron xidmətlər portalı. (Müraciətlər üzrə). http://milliarxiv.gov.az/az/elektron-xi-dmetler-portali (дата обращения: 13.05.2026).

Azərbaycanın yeni rəqəmsal arxitekturası – «Azərbaycanın yeni rəqəmsal arxitekturası» adlı vahid fəaliyyət planına həsr olunmuş müşavirədə İlham Əliyevin nitqi. https://president.az/az/articles/view/71588 (дата обращения: 13.05.2026).

International Council on Archives – https://www.ica.org/ (дата обращения: 13.05.2026).

Milli arxiv fondu haqqında – «Milli arxiv fondu haqqında» Azərbaycan Respublikasının Qanununda dəyişiklik edilməsi barədə Azərbaycan Respublikasının Qanunu. https://president.az/az/articles/view/56935 (дата обращения: 13.05.2026).

National Archives – https://www.archives.gov/ (дата обращения: 13.05.2026).

The National Archives – https://www.nationalarchives.gov.uk/ (дата обращения: 13.05.2026).

Transkribus – https://www.transkribus.org/ (дата обращения: 13.05.2026).

В Азербайджане началось принятие национальных стандартов по применению искусственного интеллекта – https://mincom.gov.az/ru/media-ru/novosti/v-azerbaidzane-nacalos-priniatie-nacionalnyx-standartov-po-primeneniiu-iskusstvennogo-intellekta (дата обращения: 13.05.2026).

RAISING AWARENESS OF THE RISKS POSED BY LARGE LANGUAGE MODELS FOR ARCHIVAL DESCRIPTION

Liah Shonhe
Sorbonne Abu Dhabi for Innovation and Research Institute (SAFIR), The Centre for Artificial Intelligence and Digital Policies, and History Department-Records Management & Archival Science, Sorbonne University Abu Dhabi (SUAD), United Arab Emirates
liah.shonhe@sorbonne.ae, https://orcid.org/0000-0003-2233-2164

Abstract. The proliferation of Large Language Models (LLMs) in the public sector promises significant gains in efficiency and improved service delivery. However, the use of LLMs, especially in low-resource settings, is often unguided and informal. While LLMs offer a practical means of addressing backlogs in archival processing, novice adopters may not be aware of the risks associated with using decontextualised models. This study employed a novice-driven experimental case study to evaluate three commercially available LLMs: DeepSeek, ChatGPT, and Gemini, across five archival records from the Botswana National Archives and Records Services (BNARS). Using a minimalist prompt with no domain-specific instructions, the outputs were assessed against eleven descriptive categories derived from archival principles.

The findings reveal systematic risks: hallucination of non-existent information that undermines provenance integrity; unsubstantiated temporal assertions that misrepresent chronological relationships; non-compliance with descriptive standards, thus preventing interoperability; variable vision capabilities introducing collection-specific biases; and inconsistent handling of sensitive historical language. No single LLM proved universally superior, indicating that domain-specific fine-tuning must consider multimodal and multi-agent capabilities.

The originality of this study lies in the development of an actionable decision flowchart for LLM adoption in archival functions.

The study concludes that safe LLM integration requires procedural rigour, shared resources, and human verification of every output before it enters the archival record, offering a realistic pathway between uncritical automation and technological abstinence.

Keywords: Archival Description, Large Language Models, Archives and Records Management, Algorithmic Bias, Professional Ethics, Generative AI

АРХИВТІК СИПАТТАУДА ІРІ ТІЛДІК МОДЕЛЬДЕРДІ ҚОЛДАНУ ТӘУЕКЕЛДЕРІ ТУРАЛЫ ХАБАРДАРЛЫҚТЫ АРТТЫРУ

Лиа Шонхе

Сорбонна Абу-Даби инновациялар және зерттеулер институты (SAFIR), Жасанды интеллект және цифрлық саясат орталығы, Тарих департаменті – құжаттарды басқару және архивтану, Сорбонна университеті Абу-Даби (SUAD), Әл- Біріккен Араб Әмірліктері

liah.shonhe@sorbonne.ae, https://orcid.org/0000-0003-2233-2164

Аңдатпа. Ірі тілдік модельдердің (Large Language Models, LLM) мемлекеттік секторда кеңінен қолданылуы жұмыс тиімділігін арттыруға және мемлекеттік қызметтердің сапасын жақсартуға елеулі мүмкіндік береді. Алайда, әсіресе ресурстары шектеулі жағдайларда, LLM технологияларын пайдалану көбіне жүйесіз әрі тиісті әдістемелік нұсқаулықтарсыз жүзеге асырылады. LLM архивтік құжаттарды өңдеудегі жинақталған жұмыстар көлемін азайтудың тиімді құралы бола алғанымен, тәжірибесі аз пайдаланушылар пәндік саланың ерекшеліктерін ескермейтін модельдерді қолданумен байланысты тәуекелдерді әрдайым бағалай бермейді. Осы зерттеуде LLM технологияларын алғаш рет қолданушылардың қатысуымен жүргізілген эксперименттік жағдайлық зерттеу барысында коммерциялық қолжетімді үш модель — DeepSeek, ChatGPT және Gemini — Ботсвана Ұлттық архиві мен құжаттамалар қызметінің (Botswana National Archives and Records Services, BNARS) бес архивтік құжаты негізінде бағаланды. Арнайы архивтанулық нұсқауларсыз берілген қарапайым сұраныс негізінде модельдердің нәтижелері архив ісінің қағидаттарына сәйкес әзірленген сипаттаманың он бір өлшемі бойынша талданды.

Зерттеу нәтижелері бірқатар жүйелі тәуекелдерді анықтады: архивтік құжаттардың шығу тегінің тұтастығына нұқсан келтіретін шындыққа сәйкес келмейтін мәліметтердің (галлюцинациялардың) жасалуы; хронологиялық байланыстарды бұрмалайтын дәлелсіз уақыттық тұжырымдар; халықаралық архивтік сипаттау стандарттарының талаптарына сәйкес келмеуі салдарынан ақпараттық жүйелердің өзара үйлесімділігінің (интероперабельділігінің) төмендеуі; визуалды материалдарды өңдеу мүмкіндіктеріндегі айырмашылықтардан туындайтын архивтік қорларға тән бұрмаланулар; сондай-ақ тарихи тұрғыдан сезімтал терминологияны бірізді өңдемеу.

Зерттелген тілдік модельдердің ешқайсысы әмбебап артықшылық көрсеткен жоқ. Бұл архив саласына арналған LLM модельдерін мультимодальды және көпагентті мүмкіндіктерді ескере отырып арнайы бейімдеудің қажеттілігін көрсетеді. Зерттеудің ғылыми жаңалығы архивтік функцияларда LLM технологияларын енгізуге арналған практикалық шешім қабылдау алгоритмін әзірлеуінде болып табылады.

Зерттеу қорытындысы бойынша, LLM технологияларын архив қызметіне қауіпсіз енгізу рәсімдердің қатаң сақталуын, ортақ әдістемелік ресурстарды пайдалануды және архивтік жүйеге енгізілер алдында әрбір нәтижені міндетті түрде маманның тексеруін талап етеді. Ұсынылған тәсіл автоматтандыруды сынсыз қолдану мен жаңа технологиялардан толық бас тартудың арасындағы тиімді әрі шынайы теңгерімді қамтамасыз етеді.

Түйін сөздер: архивтік сипаттама; ірі тілдік модельдер; архивтер және құжаттарды басқару; алгоритмдік бейтарапсыздық; кәсіби этика; генеративті жасанды интеллект

ПОВЫШЕНИЕ ОСВЕДОМЛЕННОСТИ О РИСКАХ ПРИМЕНЕНИЯ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ ПРИ АРХИВНОМ ОПИСАНИИ

Лиа Шонхе
Институт инноваций и исследований Сорбонны Абу-Даби (SAFIR),
Центр искусственного интеллекта и цифровой политики,
Департамент истории – управление документацией и архивоведение,
Университет Сорбонна Абу-Даби (SUAD), Объединенные Арабские Эмираты
liah.shonhe@sorbonne.ae, https://orcid.org/0000-0003-2233-2164

Аннотация. Широкое распространение больших языковых моделей (Large Language Models, LLM) в государственном секторе открывает значительные возможности для повышения эффективности деятельности и качества предоставления государственных услуг. Однако использование LLM, особенно в условиях ограниченности ресурсов, зачастую осуществляется неформально и без надлежащих методических рекомендаций. Несмотря на то, что LLM способны стать эффективным инструментом для сокращения накопившихся объемов архивной обработки документов, начинающие пользователи нередко не осознают рисков, связанных с применением моделей, не учитывающих контекст предметной области. В настоящем исследовании применен экспериментальный подход с участием пользователей, не имеющих опыта работы с LLM, для оценки трех коммерчески доступных моделей – DeepSeek, ChatGPT и Gemini – на примере пяти архивных документов, предоставленных Национальным архивом и службой документации Ботсваны (Botswana National Archives and Records Services, BNARS). Используя минималистичный запрос без специальных архивоведческих инструкций, авторы оценили результаты работы моделей по одиннадцати категориям описания, сформированным на основе принципов архивного дела.

Результаты исследования выявили ряд системных рисков, включая: генерацию недостоверной информации (галлюцинаций), нарушающую принцип происхождения архивных документов; необоснованные временные утверждения, искажающие хронологические связи; несоответствие международным стандартам архивного описания, препятствующее интероперабельности информационных систем; различия в возможностях обработки визуальных материалов, приводящие к специфическим искажениям при работе с архивными коллекциями; а также непоследовательную обработку исторически чувствительной терминологии. Ни одна из исследованных языковых моделей не продемонстрировала безусловного превосходства, что свидетельствует о необходимости специализированной адаптации LLM для архивной сферы с учетом мультимодальных возможностей и использования многоагентных архитектур.

Научная новизна исследования заключается в разработке практического алгоритма принятия решений по внедрению больших языковых моделей в архивную деятельность.

Автор приходит к выводу, что безопасная интеграция LLM в архивные процессы требует строгого соблюдения процедур, использования общих методических ресурсов и обязательной экспертной проверки каждого результата человеком до его включения в архивную систему. Предложенный подход представляет собой реалистичный

компромисс между некритичным использованием автоматизации и полным отказом от применения современных технологий.

Ключевые слова: архивное описание; крупные языковые модели; архивы и управление документацией; алгоритмическая предвзятость; профессиональная этика; генеративный искусственный интеллект

Introduction

The integration of Large Language Models (LLMs) into public sector operations has generated considerable optimism regarding the potential for enhanced efficiency, automation, and data-driven decision-making. Archives and records management (ARM), as a discipline fundamentally concerned with the preservation of authentic, reliable, and usable records over time, has not remained unaffected by this technological shift. Scholars and practitioners alike are exploring the application of LLMs to core archival functions including text classification, description, translation, and the analysis of large textual corpora to support appraisal and arrangement decisions (Svard and Seljan, 2024; Cherukuri *et al.*, 2025). Proponents argue that these tools offer a pathway to reduce mounting backlogs, accelerate descriptive workflows, and improve access to archival holdings (Zhang *et al.*, 2023).

Despite this enthusiasm, the adoption of LLMs (or AI in general) in ARM remains largely unregulated and, in many contexts, entirely informal, especially in low-resourced nations (Schellnack-Kelly and Modiba, 2024). From informal communications I had with African researchers in my network, it was evident that archival professionals are turning to publicly available LLMs, accessed through personal accounts, without institutional approval, formal policies, or technical safeguards. These practitioners are often motivated by genuine necessity: limited

budgets, understaffing, and overwhelming backlogs create pressure to adopt any tool that promises rapid results. These are novice users who are unaware that the uncritical use of off-the-shelf, unrefined, and decontextualized LLMs introduces profound risks to archival integrity. These models are known to produce hallucinations, replicate and amplify systemic biases present in their training data, and generate outputs that are factually inconsistent or culturally inappropriate (Gonzalez-Agirre *et al.*, 2024; Kim and Lee, 2026; Shujaa, Dorleon and Tang, 2026). When such outputs are accepted as authoritative, whether in descriptive metadata, finding aids, or preservation documentation, the archival record itself becomes compromised. Future researchers may encounter inaccurate information, misdescribed records, or systematically distorted representations of historical events. The consequences extend beyond immediate error correction; they threaten the very trustworthiness of archives as institutions of authentic memory.

The research problem, therefore, is not whether LLMs can be used in archives, clearly, they are already being used, but rather how their unmediated and unexamined application affects core archival functions and professional responsibilities. There exists a critical gap between the rapid, informal adoption of LLMs by practitioners and the development of empirical evidence regarding the specific risks, limitations, and appropriate conditions for their use in archival contexts, from a novice perspective, within an African context. While the literature has begun to address AI ethics and transparency in public administration (Lin and Zou, 2024; Pulijala, 2024; Svard and Seljan, 2024), however there is lack of evidence directly examining the outputs of unrefined, off-the-shelf LLMs when applied to authentic archival materials within an African context. Globally, several studies have documented, through controlled experimentation, the types and frequencies of errors, biases, and hallucinations that emerge when LLMs perform tasks traditionally reserved for trained archival professionals (Zhang *et al.*, 2023; Pandi *et al.*, 2024; Hollow *et al.*, 2025; Černý, Avramov and Pendse, 2026).

This study directly addresses this gap by presenting a novice experiment using DeepSeek, a publicly accessible LLM, applied to sample archival materials from Botswana. The experiment is deliberately designed to mirror the informal use patterns observed in practice: no fine-tuning, no contextual training, and no institutional customization. Instead, the LLM is prompted to perform archival description tasks as would any practitioner opening an account and uploading materials. By systematically analyzing the outputs, this study aims to demonstrate, concretely and transparently, the types of risks that emerge when decontextualized LLMs are deployed in archival settings. The findings are intended not to dismiss the potential of LLMs but to inform the development of realistic guidelines, ethical frameworks, and, where appropriate, training protocols that can mitigate harm while preserving the benefits of automation. The main objectives of the study are:

1. To demonstrate the types and frequencies of errors produced by off-the-shelf, unrefined LLMs when used for archival description by novice practitioners/adopters.
2. To analyse the impact of unrefined LLM use on core archival functions.
3. To propose a practical framework/Guideline for the safe adoption of LLMs in archival institutions operating within low-resource settings.

This research is conducted under the auspices of a collaborative project between SAFIR: Sorbonne Abu Dhabi for Innovation and Research Institute (Centre for Artificial Intelligence and Digital Policies) and ArchiLab at the Geneva School of Business Administration (HEG), funded by the Sorbonne University Abu Dhabi Internal Research Excellence Grant (IREG). The broader project, titled 'Sustainability and LLMs Implementation in the Public Sector: Concepts, Measurements and Piloting Tool', investigates the environmental, economic, and social dimensions of LLM adoption. This study contributes specifically to addressing the overarching project's first research question: In what ways has the integration of large language models (LLMs) within the public sector influenced the management of public records?

Related Work. What are LLMs?

Large Language Models (LLMs) represent a class of artificial intelligence models characterized by complex neural network architectures and an exceptionally high number of parameters, trained on vast and heterogeneous text corpora (Rillig *et al.*, 2023). These models are designed to understand, generate, and manipulate human language by learning statistical patterns, syntactic structures, and semantic relationships from the data on which they are trained (Patil and Gudivada, 2024). Unlike traditional rule-based or machine learning systems that require carefully curated and task-specific training data, LLMs operate as general-purpose engines capable of performing a wide range of language tasks without explicit task-specific programming (Zhai, 2024). LLMs are typically trained in two stages:

- *Pre-training*: self-supervised learning on huge amounts of unlabeled text to learn general language patterns (Patil and Gudivada, 2024).
- *Fine-tuning/adaptation*: additional training or instruction tuning on smaller, labeled or curated data to specialize or align with human preferences (Shao *et al.*, 2024).

The fundamental limitation of LLMs, however, lies in their design. They are, by their very nature, general-purpose models. They lack inherent domain-specific knowledge (Gonzalez-Agirre *et al.*, 2024), for instance, the specialized terminology, contextual understanding, and professional standards that govern archival science.

Application of LLMs in ARM Functions

The literature on LLM application to ARM functions reveals a field in productive tension. On one hand, researchers consistently demonstrate significant efficiency gains and, in specific tasks, performance improvements over traditional methods. On the other hand, the same studies repeatedly document accuracy limitations, bias vulnerabilities, and the non-negotiable necessity of human oversight. This section analyzes the research trends on application of LLMs across core archival functions.

Metadata Generation and Description: Metadata generation and archival description are frequently explored ARM functions in LLM literature due to the labour-intensive and costly nature of manual processes. Reusens, Adams and Baesens (2025) showed that Meta's Llama model can efficiently generate keywords from curator descriptions, offering notable advantages over manual methods. However, generative AI also introduces biases, hallucinations, over-generalizations, and outdated language—though these errors occur less often than commonly perceived, they remain too frequent to disregard. This prompts critical consideration of what error rate is acceptable in archival description, as mistakes can spread through finding aids,

search systems, and citations. Nair *et al.* (2024) found GPT-4o greatly reduced web archive metadata generation costs but human-curated metadata remained superior in quality. They concluded LLMs should supplement, not replace, human cataloguers. Groppe *et al.* (2025) further proposed a federated system combining multiple LLMs for archival metadata, which outperformed single-model approaches in quality and reliability. This indicates the future lies in coordinating multiple models rather than improving individual ones.

Zhang, Hou, Peng, Li, Hu, and Wang (2023) took a different approach with ArcGPT, the first general-purpose LLM explicitly tailored to the archival field through pre-training on massive archival domain data. Their benchmark, AMBLE, showed that domain-specific pre-training outperforms state-of-the-art general models. This finding is analytically crucial: the problem is not LLMs per se but the mismatch between general-purpose training and domain-specific requirements. Domain adaptation, whether through fine-tuning, federated orchestration, or specialized pre-training, is not optional but essential. Last but not least, Zhang *et al.* (2023) developed ArcGPT, an LLM pre-trained specifically for archives. Their benchmark showed that domain-specific models surpass general ones, emphasizing that adapting LLMs to specific fields through fine-tuning or specialized training is essential, not optional.

Transcription and Handwritten Text Recognition: Recent research by Humphries et al. (2024) found that multimodal LLMs outperformed specialized Handwritten Text Recognition (HTR) software like Transkribus, with lower Character and Word Error Rates (5.7–7% CER; 8.9–15.9% WER), marking a 14% and 32% improvement. When refining HTR outputs, LLMs reached near-human accuracy (CER 1.8%, WER 3.5%) and completed tasks much faster and at a fraction of the cost. This research challenges the belief that specialised tools always outperform general-purpose LLMs in domain tasks. For archives, using commercial LLMs with proper validation may be more effective than costly HTR software. Research established that bigger models don't always perform better; InternVL2.5-38B outperformed larger and proprietary models on historical German index cards (Vafaie *et al.*, 2025). This highlights that architecture, prompt engineering, and task-specific optimisation are more important than sheer model size, especially for institutions with limited resources.

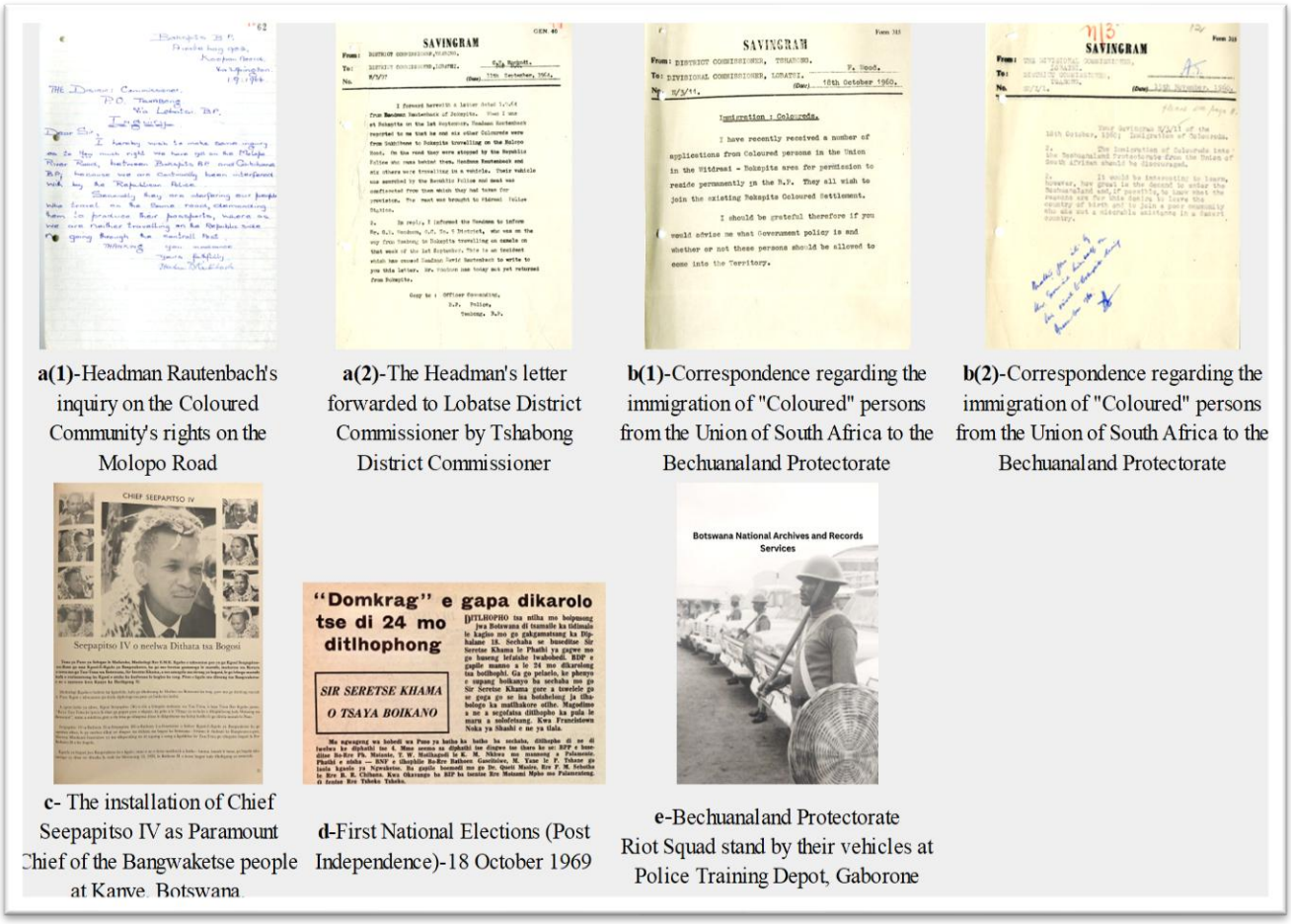
Information Extraction and Content Analysis: LLMs are used to turn unstructured archival texts into structured, analyzable data. For instance, Černý, Avramov, and Pendse (2026) created an eight-agent system for extracting intelligence from CIA records on the 1968 Prague Spring, comparing four LLMs: GPT-5 had the best quality (F1 score: 0.731); Claude Sonnet 4.5 excelled in cost and speed; Claude and Grok were highly stable; Mistral was best at reducing text. No model led in all areas, so choosing a model depends on the task, and institutions should be ready to evaluate and switch models as needed. Nguyen, Nguyen and Nguyen (2025) proposed a Retrieval-Augmented Generation (RAG)-based smart search for digital archives, combining vector and keyword retrieval with a router query engine, which outperformed traditional methods. They concluded that LLMs enhance, not replace, classic retrieval, supporting hybrid systems as the future of archival search. Additionally, graph-based RAG approaches reveal deeper content than Classic RAG in archival workflows (Fu *et al.*, 2025). Sun, Yang and Liu (2024) also leveraged LLM-RAG

to build knowledge graphs from dispersed oral history archives, showing LLMs can connect isolated records. These studies signal a shift toward knowledge-level archival access, with LLMs bridging disconnected historical narratives.

Personally Identifiable Information (PII) and Privacy Protection: LLMs have shown clear advantages over traditional methods in PII detection, requiring less training data and providing strong interpretability for audit purposes. Yang *et al.* (2023) found LLMs deliver satisfactory results without extensive datasets and offer understandable rationales for their decisions, facilitating archival review. Furthermore, Tharayil *et al.* (2024) demonstrated that combining heuristic extraction with LLM-based similarity search enhances speed and accuracy in processing archival documents. Overall, privacy protection tasks are well-suited to LLMs due to high error costs, clearly defined objectives, and their transparent decision-making.

Materials and Methods

This study employed a novice-driven experimental case study designed to simulate real-world conditions under which archival practitioners informally adopt off-the-shelf Large Language Models (LLMs) without institutional oversight, customisation, or structured prompt engineering. The experiment mimicked a low-resource scenario where a practitioner, acting as a novice user, accesses publicly available LLMs through personal accounts, provides no contextual training or domain-specific rules, and submits raw archival materials with minimal instructional prompting. Five archival records (Figure 1) were purposively selected from materials publicly posted by the Botswana National Archives and Records Services on their official Facebook page. Each record was presented to three commercially available LLMs: DeepSeek-V3.2, ChatGPT -5.3, and Google Gemini 3 Flash, using an identical, minimalist prompt. Figure 2 provides a sample of the prompt and corresponding model output. No further instructions, guidelines, test corpus, or contextual information were provided to the models, reflecting the unrefined, decontextualised use common among novice adopters. The models were not sensitized or commanded to align with the General International Standard Archival Description (ISAD(G)). The outputs were collected, transcribed, and evaluated across 11 categories derived from archival principles (see Figure 3). A comparative error analysis was then conducted to identify the error rates by each model. This methodology enables the demonstration of risks inherent in unguided LLM use while providing empirical evidence to inform professsional guidelines for LLM integration in archival practice.



a(1)-Headman Rautenbach's inquiry on the Coloured Community's rights on the Molopo Road

a(2)-The Headman's letter forwarded to Lobatse District Commissioner by Tshabong District Commissioner

b(1)-Correspondence regarding the immigration of "Coloured" persons from the Union of South Africa to the Bechuanaland Protectorate

b(2)-Correspondence regarding the immigration of "Coloured" persons from the Union of South Africa to the Bechuanaland Protectorate

c- The installation of Chief Seepapitso IV as Paramount Chief of the Bangwaketse people at Kanye, Botswana

d-First National Elections (Post Independence)-18 October 1969

e-Bechuanaland Protectorate Riot Squad stand by their vehicles at Police Training Depot, Gaborone

Figure 1. *Sample materials used for the experiment*

The archival description content generated by the three LLMs was assessed against 11 variables as presented on the framework in Figure 3. The variables were selected based on the archival principles and ISAD(G) elements. Quality assessment for each variable was assigned a rating based on the following index: Excellent = 0 errors across all 5 records, Good = 1 error across 5 records, Poor = 2-3 errors across 5 records, Very Poor = 4-5 errors across 5 records.

Prompt: "Archival description exercise: Act as an ARCHIVIST and describe the following records."

Reference code: [To be assigned, e.g., BPP/1964/01]

Title: Correspondence regarding the Molopo River Road rights and interference by the Republican Police, involving Headman David Rautenback of Bokspits.

Date(s): 1 September 1964 – [forwarding date not specified, but likely shortly after 1 September 1964]

Level of description: File (2 items)

Extent and medium: 2 sheets of paper; manuscript and typescript letters.

Creator(s): Heather Makin (on behalf of Headman David Rautenback, Bokspits B.P.)

[Forwarding officer – name not signed, but likely a District official, possibly from Tshabong]

Administrative / Biographical history:

Bokspits (also spelled Boksbits) was a settlement in the Bechuanaland Protectorate (present-day Botswana) near the border with South Africa. The Molopo River Road ran along the border area. The "Republican Police" refers to the South African police (Republic of South Africa). The Bechuanaland Protectorate was under British administration until independence in 1966. Headmen held local administrative authority over designated communities.

Scope and content:

The file consists of two related documents:

A letter dated 1 September 1964 from "Heather Makin" (signing on behalf of Headman David Rautenback of Bokspits) to the District Commissioner, P.O. Tshabong. The letter inquires about the rights of residents on the Molopo River Road between Bokspits and Gakhibane (also spelled Gaikibane), citing repeated interference by South African Republican Police, who demand passports and search travellers, even though the road does not cross into South Africa nor pass through a border control post.

A forwarding letter (unsigned, but from a district official) to the District Commissioner, enclosing the above letter. The forwarding officer reports that on 1 September 1964, Headman Rautenback and six other Coloureds were travelling from Gakhibane to Bokspits when Republican Police stopped their vehicle, searched it, and confiscated meat intended as provisions. The meat was taken to Widraai (likely Witdraai) Police Station (South Africa). The headman was advised to inform Mr. G.E. Woodman, O.C. No. 9 District, who was travelling by camel from Tsabong to Bokspits. A copy was sent to the Officer Commanding, E.P. Police, Tsabong, B.P.

Conditions governing access: No restrictions (assuming public records).

Conditions governing reproduction: Subject to archival copyright and preservation policies.

Language of material: English.

Physical characteristics: Original manuscripts; signatures present; minor fold marks and aging.

Finding aids: This description.

Related units of description: Likely other district correspondence from Tshabong (Tsabong) and border affairs between Bechuanaland Protectorate and South Africa, 1960–1966.

Notes: "Coloureds" is an historical racial classification used in the region; the archivist retains original terminology for contextual integrity. "Widraai" may be a variant spelling of Witdraai. "B.P." stands for Bechuanaland Protectorate.

Figure 2. *Sample archival description output from DeepSeek*
**This is a description of the material presented in Figure 1a (1-2).*

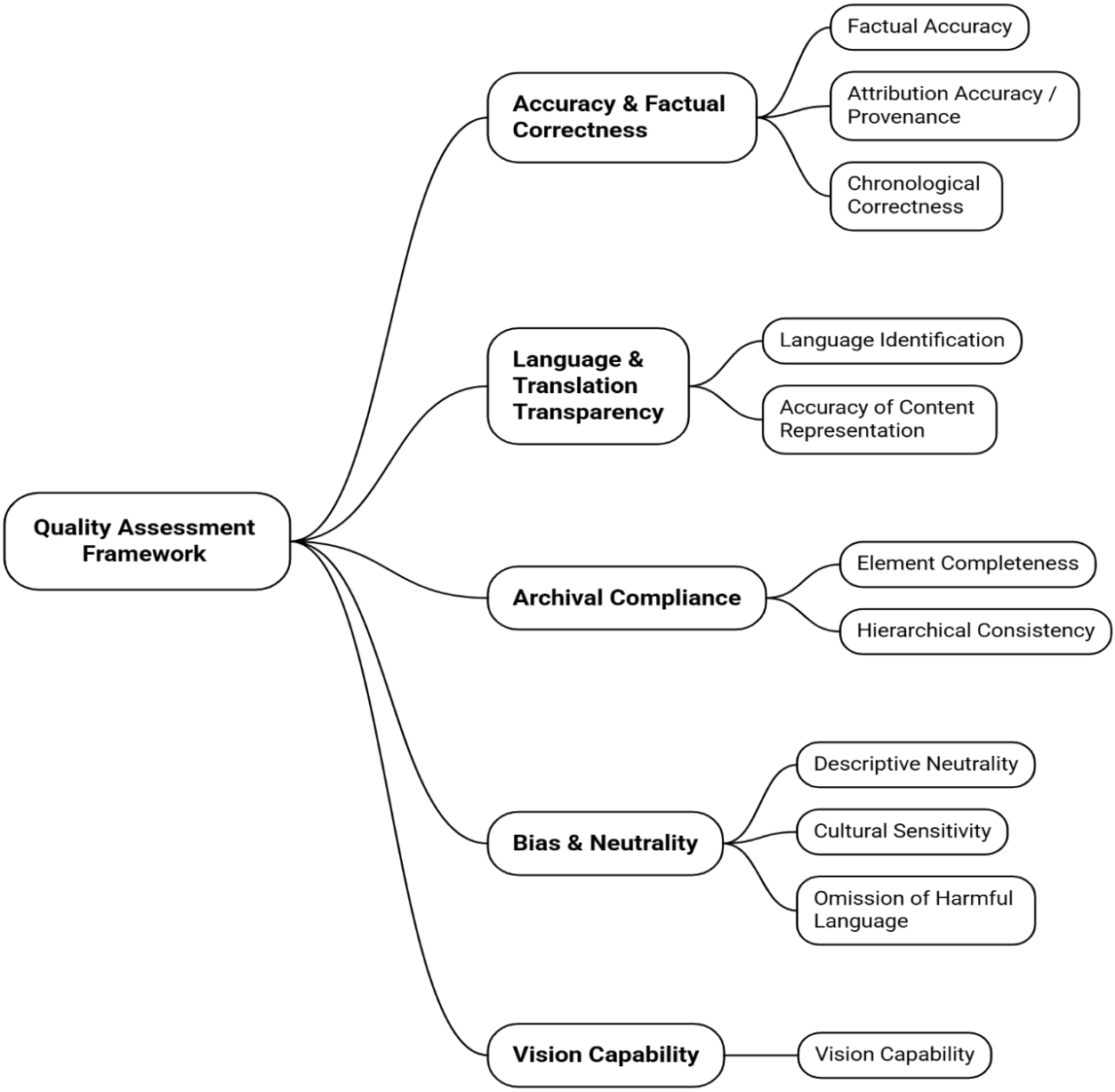


Figure 3. *Assessment Criteria for AI-Generated Content*

Results

The archival description exercise examined the outputs of three LLMs using the sample materials shown in Figure 1. The results in Table 2 show that ChatGPT and DeepSeek achieved identical accuracy (92.7%, 4 errors each out of 55 possible), while Gemini lagged substantially (74.5%, 14 errors). However, identical scores mask different risk profiles as depicted in Figure 4. ChatGPT's errors were minor and structural: missing level of description, inconsistent cultural sensitivity notes, and one factual conflation. DeepSeek's errors were more severe: hallucination and vision deficiency. Gemini's 14 errors were concentrated in chronological inaccuracies and missing mandatory ISAD(G) elements.

As shown on Table 1, the 7.3% error rate for DeepSeek and ChatGPT suggests high but not perfect reliability. Critically, error type matters more than error count. Gemini's 25.5% error rate exceeds acceptable thresholds for professional archival description, as the labour required for manual correction would outweigh any efficiency gains. ChatGPT offers the best balance of text accuracy, vision capability, and readability, requiring only minor manual verification. DeepSeek is viable for text-only collections with rigorous named-entity verification but is unsuitable for image records. Gemini is not recommended for archival description tasks.

Table 1. *Total Errors by Model (n=5)*

Category	**DeepSeek**	**Gemini**	**ChatGPT**
Factual Accuracy	2	2	1
Attribution Accuracy / Provenance	1	1	0
Chronological Correctness	0	3	0
Language Identification	0	0	0
Accuracy of Content Representation	0	0	0
Element Completeness	0	3	1
Hierarchical Consistency	0	3	1
Descriptive Neutrality	0	1	0
Cultural Sensitivity	0	1	1
Omission of Harmful Language	0	0	0
Vision Capability	1	0	0
Total Errors	*4*	*14*	*4*
Percentage of Errors	7.3%	25.5%	7.3%
Percentage of Accuracy	92.7%	74.5%	92.7%

**Note: Total possible errors = 11 categories × 5 records = 55.*

Discussion

The results of this novice-driven experiment reveal several interrelated risks that arise when off-the-shelf, unrefined LLMs are deployed for archival description without institutional oversight, or domain-specific fine-tuning. While the models achieved high accuracy rates on surface-level metrics (92.7% for DeepSeek and ChatGPT), the nature and distribution of errors raise profound concerns for archival integrity, professional ethics, and long-term researcher trust.

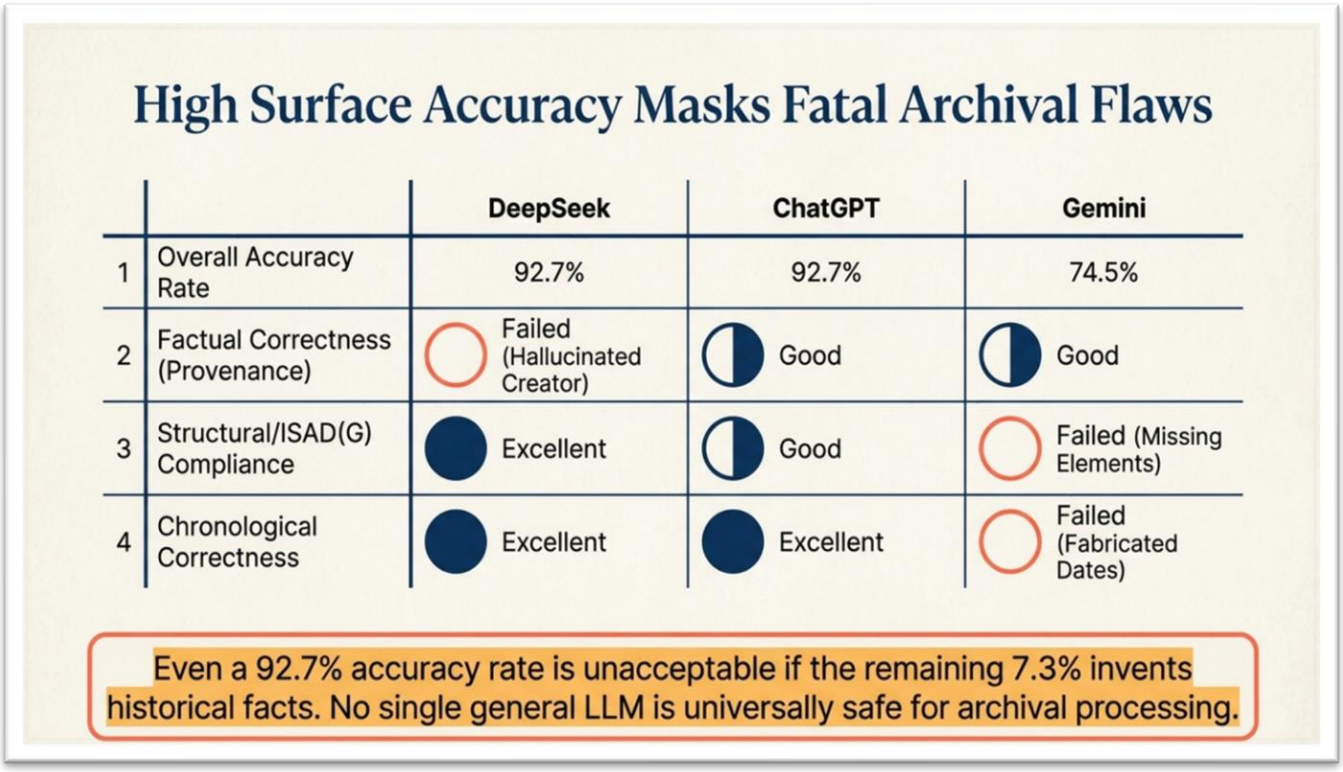


Figure 4. *Comparative Analysis from Three Models*

Hallucination and Fabricated Precision: DeepSeek's fabrication of 'Heather Makin', a non-existent person inserted into the provenance statement of the Bokspits headman's letter. This is a fundamental violation of archival authenticity. Provenance, the relationship between records and their creators, is a cornerstone of archival theory. When an LLM invents a creator, the descriptive record becomes a false historical artefact. Future researchers who encounter this description may waste time searching for a person who never existed, or worse, may cite this fabricated provenance as fact, perpetuating error across scholarly literature. These results corroborate with findings by Reusens, Adams and Baesens (2025). This LLM weakness directly attacks the chain of custody and evidential value of the record. Additionally, Černý, Avramov and Pendse (2026) observed that different LLMs exhibit systematic trade-offs in output quality and operational stability, but our findings suggest an additional dimension: the tendency to assert unsubstantiated dates varies significantly by model. Gemini's pattern of over-precision without justification represents a failure of archival rigour that cannot be remedied by simple verification, it requires the user to know, a priori, that the model fabricates dates.

Structural Non-Compliance and Interoperability Failure: Gemini's consistent use of non-standard field names and omission of mandatory elements (e.g. level of description) renders its outputs non-interoperable with archival systems. ISAD(G) was designed precisely to enable the exchange of descriptive information across repositories and systems. When LLMs produce descriptions that cannot be mapped to

archival standards (e.g. Encoded Archival Description), the efficiency gains of automation are negated. A practitioner who accepts Gemini's output would need to completely restructure each description before it could be ingested into an archival management system. Thus, the need to consider adopting pre-trained domain specific LLMs like ArcGPT which has been confirmed to outperform general models (Zhang *et al.*, 2023). Domain adaptation is not optional but essential.

Vision Capability Gaps Create Collection-Specific Vulnerabilities: DeepSeek's failure to describe the riot squad image's visible content (uniformed personnel, helmets, shields, vehicles) while ChatGPT and Gemini succeeded highlights a critical but often overlooked risk: model selection creates *collection-specific biases*. An archives with predominantly photographic or cartographic holdings that selects similar models like DeepSeek, based on its strong text performance would systematically under-describe its visual materials. Conversely, an archives prioritising non-English text records might find DeepSeek's Setswana transparency superior. This means that the choice of LLM introduces an unacknowledged archival appraisal decision, determining which materials receive adequate description, without professional deliberation. Thus multimodal LLMs are essential for archival work (Humphries *et al.*, 2024).

Cultural Sensitivity and Colonial Language: All three models kept the term 'Coloured' for accuracy, but only DeepSeek consistently added notes on its context. ChatGPT and Gemini flagged issues with the term occasionally. The Divisional Commissioner's phrase ("miserable existence in a desert country") was reproduced unflagged by ChatGPT, labelled as colonial bias by Gemini, and neutrally described by DeepSeek as reflecting administrative attitudes. This variation matters because archival description is never neutral; it shapes how future researchers interpret records (Schwartz and Cook, 2002). LLMs that fail to contextualise harmful language risk normalising it, while those that over-critique may introduce presentist interpretations that distort historical understanding and undermine contextual integrity. The archival profession has not yet established clear guidelines for LLM handling of sensitive historical terminology, and our results suggest urgent need for such standards.

The False Economy of Novice Adoption: The key risk exposed by this experiment is professional rather than technical. Using LLMs via personal accounts without institutional policies or technical safeguards can provide quick efficiency but threatens long-term archival quality. Descriptions with errors, like made-up creators, dates, or missing elements may go unchecked due to backlog and lack of human review. Novices may lack the skill to spot these mistakes. Groppe *et al.* (2025) proposed intelligence systems that unite multiple LLMs to improve quality and reliability, but such approaches require technical expertise beyond the novice user's capacity. Our experiment demonstrates that even the best-performing model still committed errors requiring human correction. Therefore, the efficiency argument for unguided LLM adoption is false: any time saved by automation must be reinvested in verification, and the verification burden would likely exceed the labour saved.

Implications for ARM Ethics and Principles. The findings challenge several core archival principles. *Respect des fonds;* maintaining records in their original

provenance and order, is undermined when LLMs invent creators or misattribute provenance. *Authenticity;* the trustworthiness of records as evidence, is compromised when descriptions contain hallucinated details that cannot be distinguished from factual content. *Accessibility;* the ability of users to find and understand records, is reduced when chronological errors misdirect searches or when non-compliant structures prevent system interoperability. Furthermore, the informal adoption of LLMs without institutional approval violates professional standards requiring transparency and accountability in recordkeeping processes.

Best Practices and Mitigation Strategies for Safe LLM Integration

Addressing the risks posed by LLMs requires practical, context-sensitive strategies, particularly for low-resourced nations where formal policies and technical infrastructure may be limited. The following recommendations focus on institutional actions rather than repeating risks already discussed.

Establish Mandatory Human-in-the-Loop Validation Protocols: No LLM output should enter the archival description pipeline without human verification. Institutions must implement tiered review processes: (1) automated flagging of uncertain outputs using confidence scores; (2) random sampling for low-risk or high-volume materials; (3) full verification for records with high research value or containing sensitive information. Critically, institutions should document acceptable error thresholds and establish correction workflows that feed back into prompt design. This does not require advanced technology, only procedural discipline.

Develop and Share Structured Prompt Templates for Novice Users: The minimalist prompt used in this experiment deliberately simulated novice practice and produced predictable errors. Institutions can mitigate this by developing open-access prompt templates that explicitly require: (1) statements of uncertainty when dates or attributions are unverified; (2) ISAD(G)-compliant output formats; (3) mandatory language identification for non-English materials; (4) prohibitions against inventing named entities or fabricating precision. These templates should be co-designed with archival professionals and distributed through professional networks, requiring no technical infrastructure.

Prioritise Local or Privacy-Preserving Deployment for Sensitive Records: Novice adopters using personal cloud accounts risk exposing sensitive archival data. For records containing personal or restricted information, institutions should deploy open-weight models locally where feasible. Where local infrastructure is lacking, institutions can: (1) establish data processing agreements with providers; (2) use on-device models that run on modest hardware; (3) develop shared regional computing hubs for resource-constrained archives. For purely descriptive tasks on unrestricted records, cloud-based models may be acceptable with appropriate verification protocols.

Implement Language and Cultural Sensitivity Protocols for Non-English Materials: The experiment revealed inconsistent handling of Setswana materials, with only one model providing explicit language transparency. Institutions serving multilingual or post-colonial collections should require prompts that: (1) mandate identification of original language before summarising; (2) flag historical terminology (e.g., "Coloured") with contextual notes explaining retention rationale; (3) disclose when

English summaries are translations, not original text. For indigenous or minoritised languages, institutions should explore small-scale fine-tuning or shared model repositories across regional networks.

Establish Institutional Governance for LLM Use: The most critical risk is procedural, novice practitioners using personal accounts without oversight. Even low-resourced institutions can implement incremental governance: (1) maintain a simple register of approved LLM use cases and models; (2) register official, institutional accounts on approved models e.g. ArcGPT; (3) prohibit personal account use for sensitive or restricted records; (4) require dual verification for all LLM-generated descriptions before publication or system ingestion; (5) create peer-review networks for sharing error reports, prompt templates, and mitigation experiences across institutions. These measures require no financial investment, only professional commitment.

Adopt Hybrid and Federated Approaches Where Possible: For institutions with some technical capacity, federated intelligence systems that unite multiple LLMs have been shown to outperform single-model solutions in metadata quality and reliability. Where fine-tuning is infeasible, Retrieval-Augmented Generation (RAG) offers a lower-cost alternative by grounding outputs in external knowledge bases of archival standards and examples. Even without these, institutions can implement a simple hybrid: use LLMs for specific, well-defined tasks where they have demonstrated reliability (e.g., PII detection) while maintaining traditional methods for provenance-sensitive description.

Proposed Decision Flowchart for LLM Adoption

Based on the study findings, the researcher urges practitioners and managers to assess their situation or circumstances before adopting an LLM in their work procedures. This can be done by using a decision flowchart similar to the one presented in Figure 5. This is a question-driven tool, to guide practitioners through a sequence of binary (YES/NO) decisions to determine the appropriate level of LLM adoption for a given archival task or institutional context. The flowchart begins with two critical gatekeeping questions: whether the record is sensitive or restricted (if YES, the answer is Level 0: Prohibit-No LLM use, revert to traditional methods), and whether the output can be verified by a human (if NO, again Level 0: Prohibit). Only when both conditions are satisfied does the flowchart proceed to assess institutional readiness: whether a formal LLM policy exists, whether verification protocols are documented, and whether models can be deployed locally for sensitive data. Based on these answers, the flowchart directs the user to Level 1-3, with clear prohibitions and safeguards embedded at each decision point. The flowchart also incorporates critical thresholds for progression, including the requirement that error rates remain consistently below 5% before advancing from Level 1 to Level 2, and that hallucination rates remain below 2% before advancing from Level 2 to Level 3. This decision tool is designed for low-resourced settings where formal policies may be absent but procedural discipline is possible, enabling practitioners to make defensible, risk-aware choices about LLM use without requiring technical expertise.

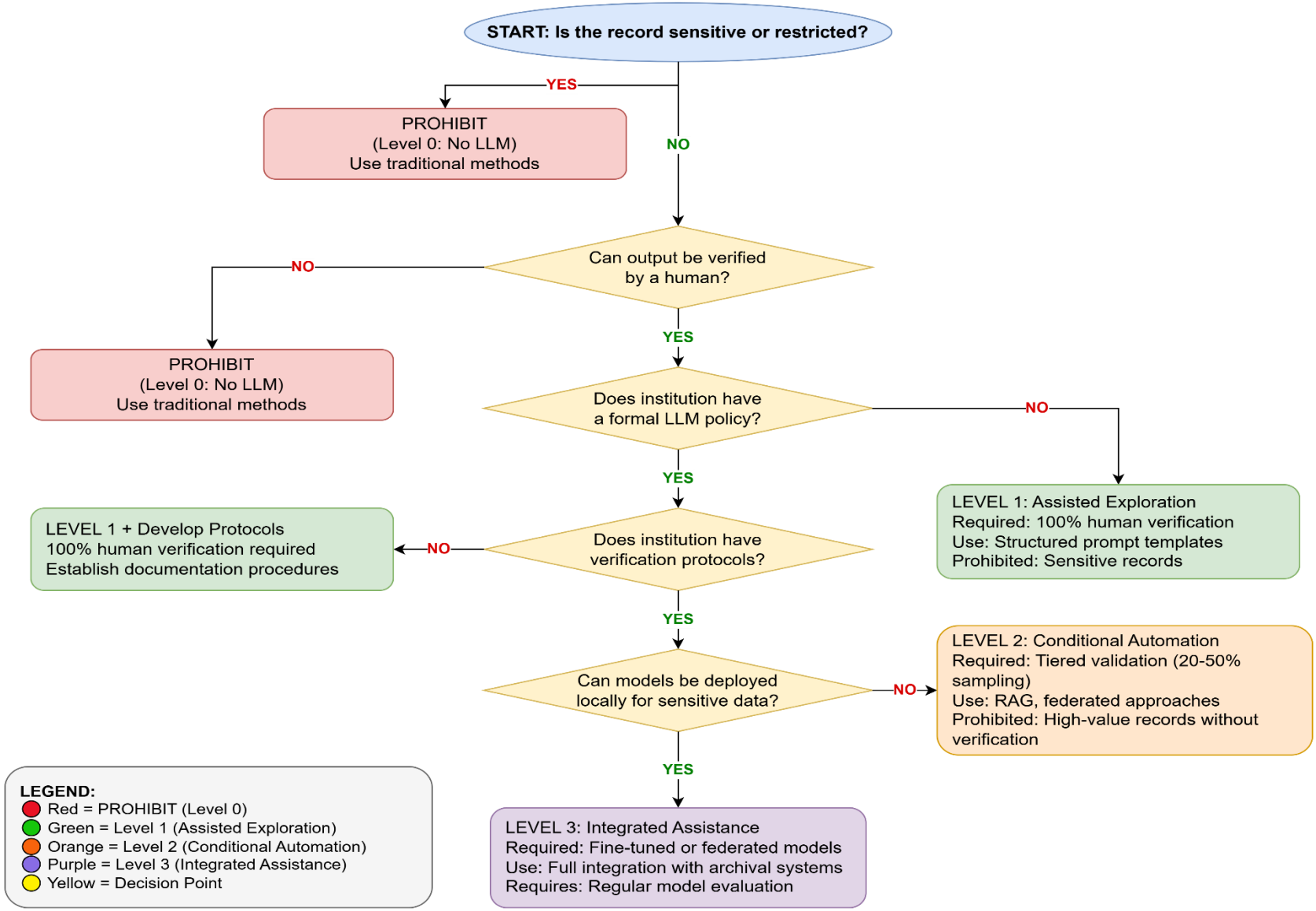


Figure 5. *Decision Flowchart for LLM Adoption in Archival Tasks Involving Sensitive Records*

Conclusion

This study makes an original contribution to archival science by moving beyond speculative discourse on LLMs to provide empirical evidence from a novice-driven experiment that deliberately simulated the informal, unguided use of off-the-shelf LLMs by practitioners operating without institutional oversight, domain-specific fine-tuning, or structured prompt engineering. The findings reveal that even high-performing models exhibit multiple categories of risk that directly threaten core archival principles. The qualitative nature of these errors poses a fundamental threat to archival authenticity, evidential value, and long-term trustworthiness, regardless of aggregate accuracy metrics. Importantly, no single LLM emerges as universally superior across all archival tasks; therefore, any decision to refine a model for archival use must consider multimodal capabilities alongside domain-specific fine-tuning to address the diverse range of record types and functions encountered in practice.

The originality of this study resides in its translation of these empirical findings into actionable professional guidance tailored for resource-constrained environments. It provides a decision flowchart as a practical, question-driven tool for practitioners and policymakers. This output demonstrates that safe LLM integration does not require expensive infrastructure but rather procedural rigour, and professional discipline. The most critical and universally accessible mitigation remains human verification of every LLM output before it becomes part of the archival record. This study equips the ARM community to make informed, risk-aware decisions about LLMs, avoiding both uncritical automation and technological abstinence.

References

Černý et al. 2026 – Černý, J., Avramov, K. and Pendse, L.R. 'A multi-stage agentic AI system for extracting information from large digital archives: case study on the Czechoslovak year 1968 in CIA's FOIA collection', The Electronic Library. 2026. Pp. 1–26. Available at: https://doi.org/10.1108/EL-06-2025-0272.

Cherukuri et al. 2025 – Cherukuri, K.S. et al. 'Large Language Models for Oral History Understanding with Text Classification and Sentiment Analysis'. 2025. arXiv. Available at: https://doi.org/10.48550/ARXIV.2508.06729

Fu et al. 2025 – Fu, Y. et al. (2025) 'Innovative practice of archival data development workflow in the AGI era: a case study of scientist archives project', Information Research an international electronic journal, 30 (iConf). 2025. Pp. 349–360. Available at: https://doi.org/10.47989/ir30iConf47335

Gonzalez-Agirre et al. 2024 – Gonzalez-Agirre, A. et al. 'Building a Data Infrastructure for a Mid-Resource Language: The Case of Catalan', in N. Calzolari et al. (eds) Jt. Int. Conf. Comput. Linguist., Lang. Resour. Eval., LREC-COLING - Main Conf. Proc. European Language Resources Association (ELRA). 2024. Pp. 2556–2566. Available at: https://www.scopus.com/inward/record.uri?eid=2-s2.0 85195920357&partnerID=40&md5=7a00578176c9fce86e66cf46f96bea7b

Groppe et al. 2025 – Groppe, J. et al. 'Automated Archival Descriptions with Federated Intelligence of LLMs'. arXiv. 2025. Available at: https://doi.org/10.48550/ARXIV.2504.05711

Hollow et al. 2025 – Hollow, A.G. et al. 'The AI-powered archivist: Harnessing generative artificial intelligence for streamlined archival description', Journal of Digital Media Management, 14(1). 2025. P. 6. Available at: https://doi.org/10.69554/ZMQC4019

Humphries et al. 2024 – Humphries, M. et al. 'Unlocking the Archives: Using Large Language Models to Transcribe Handwritten Historical Documents'. arXiv. 2024. Available at: https://doi.org/10.48550/ARXIV.2411.03340

Kim and Lee 2026 – Kim, H. and Lee, S.W. 'Sorry, it's my fault: Politeness, attribution, and anthropomorphism in managing generative AI hallucinations', International Journal of Information Management, 86. 2026. P. 102996. Available at: https://doi.org/10.1016/j.ijinfomgt.2025.102996

Lin and Zou 2024 – Lin, D. and Zou, R. 'Applications, Risk, Challenges, and Future Prospects of ChatGPT in Electronic Records Management', Journal of Artificial Intelligence Research, 1(1). 2024. P. 1–9. Available at: https://doi.org/10.70891/JAIR.2024.100021

Nair et al. 2024 – Nair, A. et al. 'Web Archives Metadata Generation with GPT-4o: Challenges and Insights'. arXiv. 2024. Available at: https://doi.org/10.48550/ARXIV.2411.05409

Nguyen et al. 2025 – Nguyen, H.D., Nguyen, T.-H.A. and Nguyen, T.B. 'A Proposed Large Language Model-Based Smart Search for Archive System', in W. Buntine et al. (eds) Information and Communication Technology. Singapore: Springer Nature Singapore (Communications in Computer and Information Science). 2025. Pp. 210–223. Available at: https://doi.org/10.1007/978-981-96-4285-4_18

Pandi et al. 2024 – Pandi, S.P. et al. 'Can LLMs categorize the specialized documents from web archives in a better way?', Proceedings of the 24th ACM/IEEE Joint Conference on Digital Libraries. JCDL '24: 24th ACM/IEEE Joint Conference on Digital Libraries, Hong Kong China: ACM. 2024. Pp. 1–11. Available at: https://doi.org/10.1145/3677389.3702591

Patil and Gudivada 2024 – Patil, R. and Gudivada, V. 'A Review of Current Trends, Techniques, and Challenges in Large Language Models (LLMs)', Applied Sciences, 14(5). 2024. P. 2074. Available at: https://doi.org/10.3390/app14052074

Pulijala 2024 – Pulijala, S. 'Artificial Intelligence in Governance: Opportunities, Challenges, and Ethical Implications for Public Administration', International Journal for Multidisciplinary Research, 6(6). 2024. P. 29990. Available at: https://doi.org/10.36948/ijfmr.2024.v06i06.29990

Reusens et al. 2025 – Reusens, M., Adams, A. and Baesens, B. 'Large Language Models to make museum archive collections more accessible', AI & SOCIETY, 40(6). 2025. Pp. 4485–4497. Available at: https://doi.org/10.1007/s00146-025-02227-8

Rillig et al. 2023 – 'Rillig, M.C. et al. 'Risks and Benefits of Large Language Models for the Environment', Environmental Science & Technology, 57(9). 2023. Pp. 3464–3466. Available at: https://doi.org/10.1021/acs.est.3c01106

Schellnack-Kelly and Modiba 2024 – Schellnack-Kelly, I. and Modiba, M. 'Developing smart archives in society 5.0: Leveraging artificial intelligence for managing audiovisual archives in Africa', Information Development. 1 OLIVERS YARD, 55 CITY ROAD, LONDON EC1Y 1SP, ENGLAND: SAGE PUBLICATIONS LTD. 2024. Available at: https://doi.org/10.1177/02666669241286224

Schwartz and Cook 2002 – Schwartz, J.M. and Cook, T. 'Archives, Records, and Power: The Making of Modern Memory', Archival Science, 2(1). 2022. Pp. 1–19. Available at: https://doi.org/10.1007/BF02435628.

Shao et al. 2024 – Shao, M. et al. 'Survey of Different Large Language Model Architectures: Trends, Benchmarks, and Challenges', IEEE Access, 12. 2024. Pp. 188664–188706. Available at: https://doi.org/10.1109/ACCESS.2024.3482107

Shujaa et al. 2026 – Shujaa, S., Dorleon, G. and Tang, A. 'How LLMs Handle Cultural Bias: Reactions to Asian Minority Historical Narratives', in S. Oh et al. (eds) Intelligence and Equity: Shaping the Future of Knowledge. 27th International Conference on Asian Digital Libraries, ICADL 2025, Metro Manila, Philippines, December 3-5, 2025, Singapore: Springer Nature Singapore (Lecture Notes in Computer Science). 2026. Pp. 39–52. Available at: https://doi.org/10.1007/978-981-95-4861-3_3

Sun et al. 2024 – Sun, Y., Yang, W. and Liu, Y. 'The Application of Constructing Knowledge Graph of Oral Historical Archives Resources Based on LLM-RAG', Proceedings of the 2024 8th International Conference on Information System and Data Mining. ICISDM 2024: 2024 the 8th International Conference on Information System and Data Mining, Los Angeles CA USA: ACM. 2024. Pp. 142–149. Available at: https://doi.org/10.1145/3686397.3686420

Svard and Seljan 2024 – Svard, P. and Seljan, S. 'The Use of Language Models (LLMs) in the Public Sector and The Impact on Public Records: A Case of Sweden and Croatia', Atlanti, 34(2). 2024. Pp. 53–72. Available at: https://doi.org/10.33700/2670-451X.34.2

Tharayil et al. 2024 – Tharayil, S.M. et al. 'Transforming Knowledge Management System with AI Technology for Document Archives', GOTECH. GOTECH, Dubai, UAE: SPE. 2024. Pp. D022S006R001. Available at: https://doi.org/10.2118/219313-MS

Vafaie et al. 2025 – Vafaie, M. et al. 'End-to-end Information Extraction from Archival Records with Multimodal Large Language Models', Proceedings of the 34th ACM International Conference on Information and Knowledge Management. CIKM '25: The 34th ACM International Conference on Information and Knowledge Management, Seoul Republic of Korea: ACM, pp. 6075–6083. 2025. Available at: https://doi.org/10.1145/3746252.3761503

Yang et al. 2023 – Yang, J. et al. 'Exploring the Application of Large Language Models in Detecting and Protecting Personally Identifiable Information in Archival Data: A Comprehensive Study*', 2023 IEEE International Conference on Big Data (BigData). 2023. IEEE International Conference on Big Data (BigData), Sorrento, Italy: IEEE. 2023. Pp. 2116–2123. Available at: https://doi.org/10.1109/BigData59044.2023.10386949

Zhai 2024 – Zhai, C. 'Large Language Models and Future of Information Retrieval: Opportunities and Challenges', Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR 2024: The 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, Washington DC USA: ACM. 2024. Pp. 481–490. Available at: https://doi.org/10.1145/3626772.3657848

Zhang et al. 2023 – Zhang, S. et al. 'ArcGPT: A Large Language Model Tailored for Real-world Archival Applications'. arXiv. 2023. Available at: https://doi.org/10.48550/ARXIV.2307.14852

«DIGITAL ARCHIVE: ЗАМАНАУИ ДИСКУРСТАР»
№ 2 сессия

Сессия № 2
«DIGITAL ARCHIVE:
СОВРЕМЕННЫЕ ДИСКУРСЫ»

«DIGITAL ARCHIVE: CONTEMPORARY DISCOURSES»
Session No. 2

THE ENCYCLOPEDİA OF THE TURKİC WORLD PROJECT AND THE CONSTRUCTİON OF DİGİTAL CULTURAL MEMORY

Ahmet Kılınç
President of Atatürk Research Center, Türkiye
ahmetkilinc@atam.gov.tr; ORCID: ORCID:0000-0002-8977-9588
Halit Aytuğ Tokur
Director of the Library at Atatürk Research Center, Türkiye
halitaytugtokur@gmail.com, ORCID:0000-0002-2257-3198

Abstract. This study comprehensively examines the origins, development, institutional infrastructure, scientific content production processes, and significance of the Encyclopedia of the Turkic World project-developed by the Atatürk Research Center (ATAM)—for the Turkic World within the context of digital transformation. Aiming to produce reliable, systematic, and interdisciplinary knowledge regarding the Turkic World, this project presents a "living encyclopedia" model suited to the requirements of the digital age. While numerous collaborations related to the Turkic World have been observed in recent years, this marks the first time an encyclopedia has been prepared in this field under academic and state supervision. The encyclopedia is unique in terms of both content and methodology. The study provides a detailed analysis of the encyclopedia's development process, content organization, editorial oversight mechanisms, and digital access features. Additionally, it examines the project's significance in terms of preserving cultural memory, providing data for artificial intelligence, disseminating academic knowledge, and fostering international scientific collaborations. The findings indicate that the Encyclopedia of the Turkic World is not merely a reference source; it is also a strategic information platform that reconstructs the shared historical and cultural heritage of the Turkic world in a digital environment. By translating the encyclopedia's entries into other Turkic and major Western languages, the project aims to expand its impact.

Keywords: Turkic World; Digital Encyclopedia; Cultural Memory; ATAM; Digital Humanities; Information Management; Artificial Intelligence

«ТҮРКІ ӘЛЕМІ ЭНЦИКЛОПЕДИЯСЫ» ЖОБАСЫ ЖӘНЕ ЦИФРЛЫҚ МӘДЕНИ ЖАДЫНЫ ҚАЛЫПТАСТЫРУ

Ахмет Кылынч
Ататүрік зерттеу орталығының президенті, Түркия
ahmetkilinc@atam.gov.tr; ORCID: 0000-0002-8977-9588
Халит Айтуғ Токур
Ататүрік зерттеу орталығы кітапханасының директоры, Түркия
halitaytugtokur@gmail.com; ORCID: 0000-0002-2257-3198

Аңдатпа. Зерттеуде цифрлық трансформация жағдайында Atatürk Araştırma Merkezi әзірлеп жатқан «Түркі әлемінің энциклопедиясы» жобасының қалыптасу алғышарттары, даму кезеңдері, институционалдық инфрақұрылымы, ғылыми мазмұнын қалыптастыру үдерістері және оның түркі әлемі үшін маңызы жан-жақты қарастырылады. Түркі әлеміне қатысты сенімді, жүйеленген және пәнаралық білім қалыптастыруды мақсат ететін бұл жоба цифрлық дәуір талаптарына сәйкес келетін «тірі энциклопедия» моделін ұсынады. Соңғы жылдары түркі әлеміне қатысты көптеген ынтымақтастық жобалары жүзеге асырылғанымен, бұл – аталған салада алғаш рет академиялық және мемлекеттік деңгейде әзірленіп отырған энциклопедиялық жоба.

Энциклопедия мазмұны мен әдіснамасы тұрғысынан бірегей сипатқа ие. Зерттеуде энциклопедияны әзірлеу үдерісі, мазмұнын ұйымдастыру қағидаттары, редакциялық сараптау тетіктері және цифрлық қолжетімділік мүмкіндіктері жан-жақты талданады. Сонымен қатар жоба мәдени жадты сақтау, жасанды интеллект жүйелеріне арналған деректер қорын қалыптастыру, ғылыми білімді тарату және халықаралық ғылыми ынтымақтастықты дамыту тұрғысынан бағаланады. Зерттеу нәтижелері «Түркі әлемінің энциклопедиясы» тек анықтамалық басылым ғана емес, сонымен бірге цифрлық ортада түркі әлемінің ортақ тарихи-мәдени мұрасын қайта жаңғыртатын стратегиялық ақпараттық платформа екенін көрсетеді. Энциклопедия мақалаларын түркі және жетекші батыс тілдеріне аудару арқылы жобаның халықаралық ықпалын кеңейту көзделіп отыр.

Түйін сөздер: Түркі әлемі; цифрлық энциклопедия; мәдени жад; ATAM; цифрлық гуманитаристика; ақпараттық менеджмент; жасанды интеллект

ПРОЕКТ «ЭНЦИКЛОПЕДИЯ ТЮРКСКОГО МИРА» И СТРОИТЕЛЬСТВО ЦИФРОКУЛЬТУРНОЙ ПАМЯТИ

Ахмет Кылынч
Президент Центра исследований Ататюрка, Турция
ahmetkilinc@atam.gov.tr; ORCID: 0000-0002-8977-9588
Халит Айтуг Токур
Директор библиотеки Центра исследований Ататюрка, Турция
halitaytugtokur@gmail.com; ORCID: 0000-0002-2257-3198

Аннотация. В настоящем исследовании комплексно рассматриваются предпосылки создания, этапы развития, институциональная инфраструктура, процессы формирования научного содержания и значение проекта «Энциклопедия тюркского мира», разрабатываемого Atatürk Araştırma Merkezi для тюркского мира в условиях цифровой трансформации. Проект, направленный на формирование достоверного,

систематизированного и междисциплинарного знания о тюркском мире, представляет собой модель «живой энциклопедии», соответствующую требованиям цифровой эпохи. Несмотря на то, что в последние годы реализовано множество инициатив, связанных с изучением тюркского мира, данный проект является первой энциклопедией в этой области, подготовленной под академическим и государственным руководством.

Энциклопедия отличается уникальностью как по своему содержанию, так и методологическим подходом. В исследовании анализируются процесс подготовки энциклопедии, принципы организации ее содержания, механизмы редакционного контроля и экспертной оценки, а также возможности цифрового доступа. Кроме того, рассматривается значение проекта с точки зрения сохранения культурной памяти, формирования массивов данных для систем искусственного интеллекта, распространения научного знания и развития международного научного сотрудничества. Полученные результаты свидетельствуют о том, что «Энциклопедия тюркского мира» представляет собой не только справочное издание, но и стратегическую информационную платформу, воссоздающую общее историко-культурное наследие тюркского мира в цифровой среде. Перевод энциклопедических статей на другие тюркские и ведущие западные языки позволит существенно расширить международное влияние проекта.

Ключевые слова: тюркский мир; цифровая энциклопедия; культурная память; ATAM; цифровая гуманитаристика; информационный менеджмент; искусственный интеллект

1. Introduction

Digitalization, one of the most defining phenomena of the century, has led to profound changes in the processes of information production, storage, and transmission. Particularly in the humanities, digital-based approaches are increasingly replacing traditional methods; processes such as data management, archiving, and access to information are being reshaped. Although the act of reading continued through printed publications until the beginning of this century, digital publications in our time have given rise to a new form of reading. The evolution occurring in the established patterns of reading has begun to profoundly affect the key actors, publishers, authors, librarians, and readers. (Odabaş 2017: 277)

Today, worldwide, old printed books are being scanned using digital scanners and converted into e-publications and e-books. Publications initially offered in PDF format have since begun to be produced in EPUB3 format as well. (Aytekin; Kılıç, 2022) At this point, it is possible to access digital publications of nearly every type where academic knowledge is made available. Digital publication formats are now also being used in monographic works.

In addition to the e-publications of papers presented at the symposia organized under the auspices of the Atatürk Research Center, monographic scientific works are also made available to readers as e-publications, subject to the author's consent, in accordance with the decisions of the Publications Committee and the Board of Directors. (https://atam.gov.tr/e-yayin/sehitler-birinci-dunya-savasinin-ve-milli-mucadelenin-kayiplari/(date of Access:29.05.2026) In fact, some publishers and periodicals have decided to continue their publishing activities exclusively through e-publications, without producing printed editions. As a matter of policy, the Atatürk Research Center publishes scientific works—comprising conference and symposium papers—exclusively in digital format and makes them available to readers.

.(https://atam.gov.tr/e-yayin/bagimsizliginin-30-yilinda-azerbaycan-turkiye-iliskileri-uluslararasi-sempozyumu-bildiri-metinleri/(date of Access:29.05.2026)

Encyclopedias were produced in ancient Rome and in Turkic-Islamic states as a structured and systematic source of knowledge, and have also been used to preserve information. (Secundus, 2022), (Kuteybe, 1963) Investments in education during the Republican era, the expansion of education and instruction, rising school enrollment, the opening of new universities, university and school libraries, and the spread of provincial and district public libraries beginning in the 1920s have increased interest in and the need for encyclopedias. (Meydan Larousse Büyük Lügat ve Ansiklopedi, 1999).

National and international encyclopedias have also been influenced by digital publishing, and these information sources – which were previously presented in volumes – have adopted a method of providing faster, article-based service through alphabetical indexes, category selection, and search engines. As a result, newly added entries are seamlessly incorporated into the encyclopedia's database, and constraints such as volume integrity and page limitations encountered in printed encyclopedias have been eliminated. (https://referenceworks.brill.com/display/package/eio (Encyclopaedia of Islam Online (English) (date of Access:29.05.2026); (https://ansiklopedi.tubitak.gov.tr/search) (date of Access:29.05.2026); (https://islamansiklopedisi.org.tr/) (date of Access:29.05.2026).

Aware of its historical responsibility, the Presidency of the Atatürk Research Center, within the context of the New World Order that emerged following the dissolution of the Soviet Union in 1991, is conducting comprehensive academic studies in addition to establishing legal, political, economic, and commercial ties with countries that have close historical, cultural, and geopolitical relations with Turkey.

The Organization of Turkic States, which held its first summit under the theme of "Economic and Commercial Cooperation" on October 20–21, 2011, in Almaty, has

convened 12 summits to date and has rapidly institutionalized itself, establishing units across various disciplines. These Turkic Cooperation Organizations include: Türksoy, Türkpa, the Turkic Investment Fund, the Turkic Academy, and the Turkic Culture and Heritage Foundation. (https://turkicstates.org/tr) (date of Access:29.05.2026); (https://turkicstates.org/tr/iliskili-kurumlar) (date of Access:29.05.2026); (Bıyıklı, 2016:.405) (date of Access:29.05.2026)

The Republic of Turkey has also been developing academic ties with Turkic States in recent years. The Orhun Exchange Program, which facilitates student and faculty exchanges between higher education institutions that are members of the Turkic Council's Union of Turkic Universities, is a significant initiative in this regard. (https://turkunib.org/tr/orhun)(date of Access:29.05.2026)

The Hoca Ahmet Yesevi University, the Turkey-Azerbaijan University, the Kyrgyzstan-Turkey Manas University, and the International University of Turkic States—established through a joint effort and shared responsibility among Turkic states—constitute the most beautiful and valuable examples of academic cooperation (https://www.ayu.edu.tr/, https://tau.edu.az/, https://www.manas.edu.kg/tr, https://turkunib.org/tr) (date of Access: 29.05.2026).

Undoubtedly, these academic collaborations have also influenced scholarly publications. The Atatürk Research Center, a constitutionally established scientific institution in Turkey, has decided to contribute to the unity of the Turkic world in an encyclopedic manner. Indeed, a living encyclopedia tailored to the needs of the Turkic world has emerged. The Encyclopedia of the Turkic World stands out as one of the most comprehensive examples of this transformation in Turkey. The project is not merely a tool for storing information; it also represents an information ecosystem that systematically presents the shared historical, cultural, and civilizational heritage of the Turkic world in a digital environment. This study aims to examine the institution behind the encyclopedia, its origins, significance, and development process from the

perspectives of digital cultural memory, academic production, and knowledge management.

2. The Atatürk Research Center and Academic Institutionalization

The Atatürk Research Center (ATAM) was established in 1983 in accordance with Article 134 of the Constitution of the Republic of Turkey and operates as a prestigious scientific institution with public legal personality. The institution's primary objective is to examine the intellectual system and principles of Mustafa Kemal Atatürk, as well as the history of the National Struggle and the Republic, using scientific methods, and to present the resulting knowledge and documents to the academic community.

Following the signing of the Mondros Armistice Agreement on October 30, 1918, the Turkish Nation entered one of the most dangerous and darkest periods in its history due to the subsequent occupations. Although the Turkish Nation, whose armies had been disbanded, soldiers demobilized, and weapons confiscated, had spent the previous decade exhausted on the battlefields, it was compelled to launch a new struggle for the independence of Anatolia. Emerging spontaneously across various regions of Anatolia as a militia force in response to unjust occupations and the denial of its existence, the Kuvayı Milliye became the symbol of the "National Struggle" and resistance. Under the command and leadership of Mustafa Kemal and his comrades-in-arms, the Kuvayı Milliye continued the battles and, with the transition to a regular army, attained the power to expel the enemy from its lands. The Turkish War of Independence was a total struggle that freed the nation from the brink of enslavement despite impossibilities, scarcity, and every kind of hardship. The war, waged under the principle of "Independence or Death!" to fulfill the principles of the Misak-ı Millî declaration, culminated in independence. (90. Yılında..., 2011); (Selek, 1968) Since its establishment, ATAM (Atatürk Research Center) has adopted an interdisciplinary approach and has conducted extensive research in the fields of history, sociology, political science, law, economics, diplomacy, and cultural studies.

The Atatürk Research Center Presidency is also an institution holding the status of an International Publisher. Our institution has published a total of 621 books, 39 of which are in foreign languages (English, German, French, Russian, Azerbaijani, Kazakh, Tatar, Kyrgyz, Uzbek, Bulgarian, Arabic, Persian, Urdu, and Turkmen). The publications of the Atatürk Research Center Presidency, consisting almost entirely of academic books, have been cataloged by libraries in many countries. In this context, having published over 630 works is a significant indicator of the institution's academic production capacity. (https://atam.gov.tr/wp-content/uploads/2025/12/ATAMin-Uluslararasi-Yayinevi-Niteligi-1.pdf) (date of Access:30.05.2026).

In 1984, pursuant to paragraphs (c) and (d) of Article 19 of Law No. 2786, the "Atatürk Special Library" and the "Atatürk Archive," focusing on Atatürk, the National Struggle, and the History of the Republic, were established within the Atatürk Research Center Presidency. The Atatürk Research Center Library holds publications related to Atatürk's life and work, the National Struggle period, the foundation of the Republic, events in Republican history, and Atatürk's principles and reforms. The library collection currently includes approximately 12,000 printed books (mostly in Turkish), 77 albums, and around 500 bound journals. Additionally, the collection contains theses and audio-visual materials (audio tapes, CDs, DVDs, etc.). The books in the collection are classified according to the Dewey Decimal Classification System.

Operating on a closed-stack system, the collection is organized using a numeric shelf number system that starts with the prefix "ATAM." While the majority of the collection is in Turkish, it exhibits great diversity in terms of publication language. In addition to Turkish publications, there are works in Russian, Kazakh, Bulgarian, Hungarian, Georgian, Greek, Ottoman Turkish, Arabic, English, French, Kyrgyz, Kazakh, and German. The collection also includes copies of Mustafa Kemal Atatürk's *Nutuk* (Speech) in its original (Ottoman Turkish) and modern Turkish editions, as well as translations into various languages (English, German, French, Kyrgyz, Kazakh, Russian, Persian) (https://atam.gov.tr/kutuphane/) (date of Access: 30.05.2026).

The Atatürk Supreme Council for Culture, Language and History, along with the institutions under its umbrella, carry out their duties—conducting, commissioning, supporting, and coordinating scientific research projects on the Turkish language, history, culture, and all aspects of Atatürk and his legacy within the integrity of the social and human sciences—in accordance with the Directive on Scientific Research and Translation Projects of the Atatürk Supreme Council for Culture, Language and History. As the Atatürk Research Center Presidency, we support comprehensive internal and external research projects as well as investment projects. (https://atam.gov.tr/projeler/)(date of Access:30.05.2026)

The Atatürk Research Center Presidency, operating under the Atatürk Supreme Council for Culture, Language and History, provides domestic and international scholarships to master's and doctoral students who participate in projects prepared on priority topics relevant to the country within its own fields of study or in the social and human sciences, and who determine their thesis topics within the framework of these projects. It also offers post-doctoral research scholarships under the conditions set forth in the relevant regulation. In accordance with the "Regulation on Education and Research Scholarships of the Atatürk Supreme Council for Culture, Language and History," the Atatürk Research Center Presidency continues to award master's and doctoral scholarships each year (https://atam.gov.tr/etkinlik-kategori/burslar/) (date of Access: 30.05.2026).

The *Atatürk Araştırma Merkezi Dergisi* (Journal of the Atatürk Research Center), published by the Atatürk Research Center, aims to share scientific research and studies concerning Turkish history, Atatürk, and the Republic of Turkey, as well as related information and documents, with the national and international academic community. The journal focuses on topics from the second half of the 19th century onward within the context of Turkish and world history. The *Journal of the Atatürk Research Center*, which publishes in the fields of social and human sciences, is a journal of modern history that emphasizes the political and social developments in the Ottoman Empire from the second half of the 19th century, as well as topics related to Atatürk and the history of the Republic of Turkey. It is a peer-reviewed journal that features original scientific studies presenting new and noteworthy perspectives on the subjects it addresses.

Furthermore, the journal's inclusion in international indexes and its open-access policy have transformed the institution into a scientific center that meets the requirements of the digital age (https://atamdergi.gov.tr/) (date of Access:30.05.2026).

The Atatürk Research Center Presidency is among the leading academic institutions in Turkey in the field of encyclopedias. The *Atatürk Encyclopedia*, which

was made available online as of October 29, 2020, currently contains nearly 1,500 entries. The digital publication of the *Atatürk Encyclopedia* not only ensures broad public access but also facilitates its continuous updating and the incorporation of requests, suggestions, and criticisms arising from new information and findings. The *Atatürk Encyclopedia* has so far received nearly 12 million visits (https://ataturkansiklopedisi.gov.tr/) (date of Access: 30.05.2026).

3. Emergence and Development of the Encyclopedia of the Turkic World

The Presidency of Atatürk Research Centre conducts numerous activities targeting the Turkic world. As a requirement of historical responsibility, within the context of the New World Order that emerged following the dissolution of the Soviet Union in 1991, comprehensive academic studies are also carried out regarding countries that have close historical, cultural, and geopolitical relations with Turkey.

In line with this, the “Turkic World, History and Cultural Geography Book Series” has been developed, focusing on the countries of Ukraine, Georgia, Azerbaijan, Bulgaria, and Kyrgyzstan. The primary aim of the series is to evaluate the political structures, geographical characteristics, economic systems, social transformations, and cultural heritage of these countries, taking into account their historical backgrounds, through a holistic and multidimensional approach.

The works particularly aim to emphasise the developments that have occurred since 1991. This approach is significant both for comprehending the transformations these countries have undergone since gaining independence and for analysing the current dynamics of the Turkic world. Following the volume on Ukraine, the second book of the series, titled *Georgia*, was published. The promotional and signing ceremony for this work took place on 9 July 2025 in Tbilisi, hosted by the Georgian National Academy of Sciences, with contributions from Sokhumi State University. The third book of the series, the edited volume titled *Azerbaijan*, has also been released among the institution’s publications (https://emagaza-atam.ayk.gov.tr/product/default/757/Azerbaycan/) (date of Access: 31.05.2026).

The Presidency of Atatürk Research Centre (ATAM) makes significant contributions to the academic world not only through its studies on Atatürk, the National Struggle, and the history of the Republic of Turkey, but also through its high-quality publications and multifaceted events focused on the Turkic world. Works such as The Cyprus Events Through the Language of Witnesses: 1955–1983, The Psychosocial Recent History of Azerbaijan, Painful Anniversaries of Turks and Kindred Communities, and among its recent publications, The Future of the Eastern Mediterranean and Turkey, are concrete examples of its dedication to the Turkic world. (https://atam.gov.tr/e-yayin/dogu-akdenizin-gelecegi-ve-turkiye-uluslararasi-sempozyumu/) (date of Access:31.05.2026) Furthermore, by organising noteworthy events such as the “International Symposium on Azerbaijan–Turkey Relations on the 30th Anniversary of Independence” (https://atam.gov.tr/e-yayin/bagimsizliginin-30-yilinda-azerbaycan-turkiye-iliskileri-uluslararasi-sempozyumu-bildiri-metinleri/) (date of Access: 31.05.2026) and the “Panel on the Two Shores of the Caucasus: Turkey-Georgia Relations from Past to Present,” ATAM creates academic and cultural sharing platforms, fostering internationally interactive environments aimed at understanding the historical and social development of contemporary Turkic communities.

In addition to these comprehensive activities, in 2013 ATAM signed a scientific cooperation protocol with the Institute of History of the Azerbaijan National Academy of Sciences, thereby formally embarking on the production of encyclopedic content specifically focused on the Turkic world. Following the protocol, works entitled *Common Turkic States* and *Common Turkic Notables* were planned; subsequently, this idea was expanded, and ATAM initiated an independent "Encyclopedia of the Turkic World" project. This digital encyclopedia, which aims to present collectively the states, figures, events, structures, and cultural heritage elements that have contributed to Turkic civilisation, was launched on 31 December 2024 with 137 entries. The new political and cultural relations that developed in the Turkic world after the dissolution of the Soviet Union were decisive in the emergence of the project. During this process, academic collaborations between Turkey and countries such as Azerbaijan, Kyrgyzstan, Georgia, and Ukraine increased, and the need for a common historiography grew stronger.

In order to provide a more effective contribution in the current conjuncture, the Presidency of Atatürk Research Centre established a Working Group within its own structure. Together with this Working Group, it was decided to bring previous accumulated knowledge to light in the form of an encyclopedia. Indeed, as indicated, the institution's Directorate of Scientific Studies, the relevant members of the ATAM Scientific Board, and the relevant specialists of the Senior Institutional Expert Cadre possess considerable experience, also derived from the Atatürk Encyclopedia. The Working Group consists of members of the ATAM Scientific Board, all of whom are leading scholars in their respective fields. After the Working Group decided to produce the encyclopedia, the process of commissioning entries began.

The main objectives of the encyclopedia can be summarised as follows:

- To engage with the Turkic world on a shared academic platform
- To ensure that data concerning the Turkic world is written by Turks
- To provide academic, controlled information to artificial intelligence
- To produce reliable information based on primary sources
- To establish a common perspective regarding Turkic civilisation
- To develop a continuously updatable digital information platform
- To make academic knowledge freely accessible to broad audiences

In this respect, the project undertakes both a scientific and a cultural mission.

4. Content Structure and Thematic Organisation

The Encyclopedia of the Turkic World is published online at https://turkdunyasiansiklopedisi.gov.tr. The primary objective of the work is to develop a common perspective on Turkic history, historical figures, and events within the Turkic world in the 21st century. The encyclopedia entries are prepared with the aim of providing information based on primary and reliable sources regarding Personalities, States, Places, as well as Events and Phenomena, in a manner that will contribute to the strengthening of shared sentiment and thought in the Turkic world.

The Encyclopedia of the Turkic World aims to include states established in geographies bearing the traces of Turkic civilisation, significant figures who have emerged, events that have taken place, the culture and civilisation that have been built, and in short, the rich historical heritage found within the Turkic geography and world.

Incoming texts are formally checked by the Article Tracking Officer (ATO) for compliance with spelling and writing rules. If no violation of the writing rules is found, the text is added to the agenda of the Encyclopedia of the Turkic World Commission (Working Group). If the Working Group deems the article worthy of publication in the encyclopedia based on its subject matter, it is then sent to a member of the Commission for review. Taking into account the report prepared by the Working Group member, the Commission decides whether the article will be published.

The Working Commission of the Encyclopedia of the Turkic World may request corrections, either formal or content-related, from article authors based on checks conducted on published encyclopedia articles and/or in response to questions and requests received from the public. If the corrections are not made by the author within the specified deadlines, the relevant articles will not be published until the corrections are completed. Furthermore, if one year has passed since the correction request and the corrections have still not been made by the author, the article will be commissioned to another author. This situation is considered accepted by the authors upon the approval of the article proposal.

The rights to process, reproduce, distribute, represent, and communicate to the public by means of sign, sound and/or image transmission devices arising from the prepared work belong to the Atatürk Research Centre. The language of publication of the Encyclopedia of the Turkic World is Turkish.

Care should be taken to upload the article to the system according to the general order specified below:

1. Title (If the article is about a person, dates of birth and death should be given after the title. If the article concerns an event, the date of occurrence should be written after the title.)
2. Definition/Abstract
3. Keywords(A minimum of 4 and a maximum of 8 keywords should be given in alphabetical order.)
4. Main Text
5.Bibliography
6. Article Image Upload
7. Article Word File Upload

The main text of the article must be entered into the "Article Text" section of the Dictionary and Encyclopedia Management System (SAYS) in an encyclopedic article format, presented as a composition and written in a easily readable manner. Excessive quotations should be avoided in the text, nor should the author's own expressions be incorporated into the quoted material. The writing should be handled as a unified whole within a thematic and chronological context. Sentences should be kept short, and the vocabulary used should be simple enough for today's reader to understand. Neither praise nor censure of the individuals about whom the article is written should be made. The aspects that justify the inclusion of persons and events as "articles" in the Encyclopedia of the Turkic World should be prioritised. Article proposals/recommendations should also be made through the "propose article" section by entering SAYS via the address https://giris.ayk.gov.tr/signin.

Authors submit their articles, prepared in accordance with the encyclopedia's publication principles, the spelling rules of the Turkish Language Association (TDK),

and the TDK's Contemporary Turkish Dictionary, through SAYS. Authors are deemed to have accepted that minor corrections (the rate of changes shall not, under any circumstances, exceed ¼ of the main text) that are not substantive in nature may be made by the Working Group and the experts responsible for the encyclopedia.
Any article visuals should also be uploaded to the system separately. Texts should consist of a maximum of 3,500 words. Depending on the significance of the subject matter covered by the article, this number may be increased or decreased by the Working Group. .(https://turkdunyasiansiklopedisi.gov.tr/hakkinda/k%C4%B1lavuz) (date of Access:31.05.2026) The Encyclopedia of the Turkic World is built upon a comprehensive content architecture. The encyclopedia articles are grouped under nine main categories:

- Turkish States

The Cossacks and the Kazakh Khanate: (https://turkdunyasiansiklopedisi.gov.tr/detay/8248/Kazaklar-Ve-Kazak-Hanl%C4%B1%C4%9F%C4%B1-) (date of Access:01.06.2026)

- Statesmen Nuri Killigil: (https://turkdunyasiansiklopedisi.gov.tr/detay/819/Nuri-Killigil) (date of Access:01.06.2026)

- History and Culture

Yargucu / Yargan / Yargıcı:Title given to judges in pre-Islamic Turkish states. (https://turkdunyasiansiklopedisi.gov.tr/detay/8089/Yargucu-/-Yargan-/Yarg%C4%B1c%C4%B1) (date of Access:01.06.2026)

- Turkish Communities

The Yakut Turks, a Turkic people living in Siberia. (https://turkdunyasiansiklopedisi.gov.tr/detay/1716/Yakut-T%C3%BCrkleri) (date of Access:01.06.2026)

- Language and Literature Abdulla Aripov (Abdulla Oripov), Uzbek Poet. (https://turkdunyasiansiklopedisi.gov.tr/detay/55/Abdulla-Aripov-(Abdulla-Oripov) (date of Access:01.06.2026)

-Geography and Space Bishkek, the Capital of Kyrgyzstan. (https://turkdunyasiansiklopedisi.gov.tr/detay/813/Bi%C5%9Fkek)(date of Access:01.06.2026)

- Science and Art Necmettin Hacıeminoğlu, Turkologist. (https://turkdunyasiansiklopedisi.gov.tr/detay/1621/Necmettin-Hac%C4%B1emino%C4%9Flu) (date of Access:01.06.2026)

- Intellectuals and Thinkers Alimcan Idrisi (1887–1959) – Siberian reformist Tatar scholar, religious leader, journalist, and writer. (https://turkdunyasiansiklopedisi.gov.tr/detay/6780/Alimcan-%C4%B0drisi) (date of Access:01.06.2026)
- Pioneering Women in Turkish History

Saray Khatun, Azerbaijan's First Female Diplomat. (https://turkdunyasiansiklopedisi.gov.tr/detay/1109/Saray-Xatun) (date of Access:01.06.2026) (https://turkdunyasiansiklopedisi.gov.tr/) (date of Access:01.06.2026).

The encyclopedia continues its publication with the aim and effort of covering not only the Turks of Türkiye but the entire Turkic world. This classification makes it

possible to address the Turkic world in a holistic manner, encompassing its historical, cultural, and geographical dimensions.

The following criteria are decisive in the selection and preparation of the entries:

- Scientific contribution value
- Reliability and source-based approach
- Inclusivity and representational power
- Originality that contributes to historiography

These criteria strengthen the academic quality of the encyclopedia.

5. Scientific Content Production and Peer Review Process

The production of the encyclopedia's content is carried out through the Dictionary and Encyclopedia Management System (DEMS). This system manages the stages of article proposal, writing process, peer review, and publication in an integrated manner. The system itself is digital; applications can be submitted to this open-access system from anywhere in the world. (https://says.ayk.gov.tr/says/itemPrinciplesPage) (date of Access:02.06.2026)

The editorial process consists of the following stages:

- Article proposal and justification
- Expert review within the institution
- Deliberation within the working group
- Invitation to field experts (noting that this invitation is signed by the chair)
- Peer review
- Revision: Technical and linguistic corrections
- Publication approval

Doctoral-level academic expertise is accepted as the fundamental criterion in the selection of authors. Furthermore, objectivity, use of plain language, and chronological integrity hold significant places among the writing principles.

6. Digitalization Process and Publication Model

The Encyclopedia of the Turkic World is a "born-digital" project, designed directly for the digital environment. It is not a "retro-digitized" project (commonly known as digitized). This model offers significant advantages over traditional printed encyclopedias. These advantages include:

- Continuously updatable content
- Article-based linking system
- Search engine functionality
- Multi-author structure

The platform's mobile-compatible and accessible structure enhances the user experience. Featuring a user-friendly interface, the application offers functions such as day/night mode, text enlargement, a favorites list, article history, and in-text links; it is made available free of charge and without advertisements on iOS and Android platforms. It aims to transmit the common values of the Turkic World to future generations. (https://turkdunyasiansiklopedisi.gov.tr/mobile) (date of Access:02.06.2026)

7. Quantitative Development and Performance Analysis

The encyclopedia has demonstrated significant growth in a short period:

- 2024: Launch with 137 articles
- 2026: 345 articles

- More than 89,000 accesses

These data indicate the project's rapid growth and its reach to a broad user base.

8. Cultural Memory and Academic Contribution

The Encyclopedia of the Turkic World plays a significant role in the construction of digital cultural memory.

The main reasons the project is defined as a "living encyclopedia" are:

- Its capacity for continuous updating,
- Its openness to user feedback,
- Its facilitation of the dissemination of scientific knowledge.

In this respect, the project transcends the classical understanding of an encyclopedia.

The encyclopedia also provides data to artificial intelligence. Regardless of the specific AI system, since the data on the platform are open to all, information can be retrieved and processed. This is important for the future vision, as the use of artificial intelligence is increasing rapidly day by day.

9. Future Perspective and Conclusion

The project's future development goals include:

- Multilingualism (Turkic dialects)
- Multimedia content
- AI-supported systems

These developments will further strengthen the encyclopedia's role in the digital knowledge ecosystem. The Encyclopedia of the Turkic World offers a comprehensive and significant model for the preservation of cultural heritage and the dissemination of academic knowledge in the digital age. The project is not merely an information source but also a living and growing digital platform that constructs the shared memory of the Turkic world. In this capacity, it carries strategic value from both academic and cultural perspectives.

References

atam.gov.tr (accessed 29.05.2026)

Aytekin, Ç. & Kılıç, M. C. (2022). Popüler Medya Türü Olarak Dijital Metin İçeriği Üretimi: Çok Sayıda Alkış Alan Makalelerin Anatomisi, Yeni Medya Elektronik Dergisi, 6 (2), 113-130 (accessed 29.05.2026).

Bıyıklı, Mustafa, “Türk Dünyası Ortak Kuruluşu Olarak Türk Akademisi Uluslararası Teşkilatı (TWESCO), Rolü ve İş Birlikleri”, Dumlupınar Üniversitesi Sosyal Bilimler Dergisi / Dumlupınar University Journal of Social Sciences Afro-Avrasya Özel Sayısı-Aralık 2016 / Special number of Afro-Eurasia-December 2016, s.405 (accessed 29.05.2026).

https://atamdergi.gov.tr/ (accessed 30.05.2026).

https://atam.gov.tr/e-yayin/bagimsizliginin-30-yilinda-azerbaycan-turkiye-iliskileri-uluslararasi-sempozyumu-bildiri-metinleri/ (accessed 29.05.2026).

https://atam.gov.tr/etkinlik-kategori/burslar/ (accessed 30.05.2026).

https://ansiklopedi.tubitak.gov.tr/search (accessed 29.05.2026).

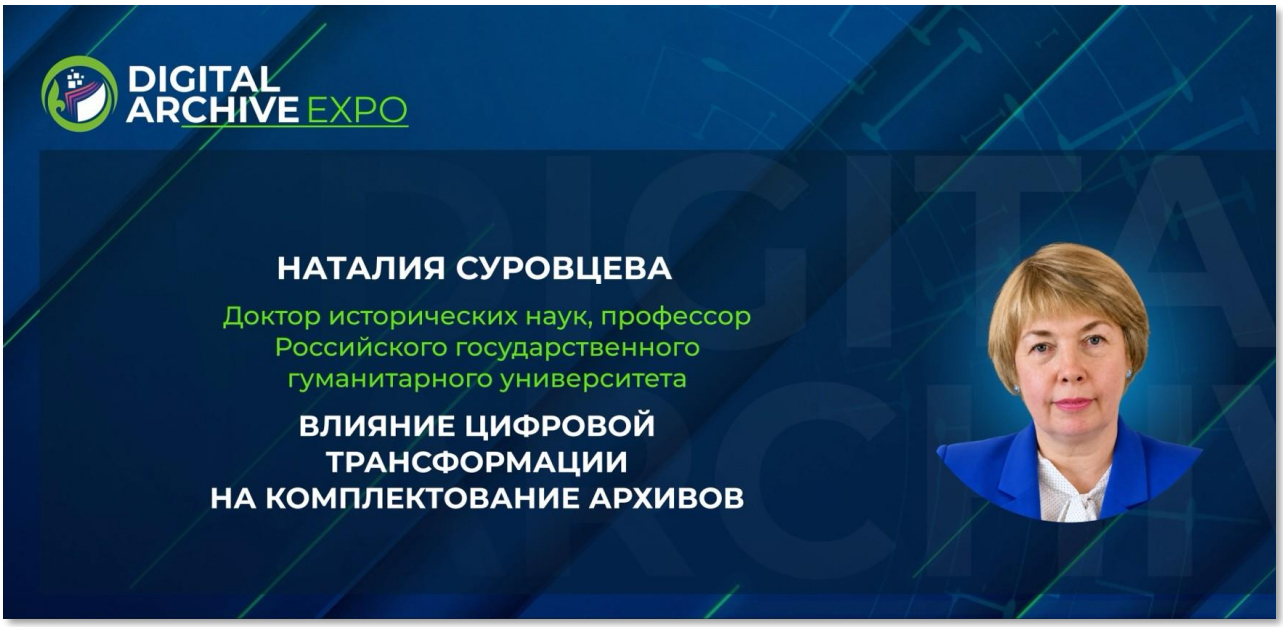
https://atam.gov.tr/kutuphane/ (accessed 30.05.2026).

https://atam.gov.tr/projeler/ (accessed 30.05.2026).

https://atam.gov.tr/e-yayin/bagimsizliginin-30-yilinda-azerbaycan-turkiye-iliskileri-uluslararasi-sempozyumu-bildiri-metinleri/(accessed 31.05.2026).

https://atam.gov.tr/e-yayin/dogu-akdenizin-gelecegi-ve-turkiye-uluslararasi-sempozyumu/ (accessed 31.05.2026).

https://atam.gov.tr/e-yayin/sehitler-birinci-dunya-savasinin-ve-milli-mucadelenin-kayiplari /(accessed 29.05.2026).
https://atam.gov.tr/wp-content/uploads/2025/12/ATAMin-Uluslararasi-Yayinevi-Niteligi-1.pdf (accessed 30.05.2026).
https://ataturkansiklopedisi.gov.tr/ (accessed 30.05.2026)
https://emagaza-atam.ayk.gov.tr/product/default/757/Azerbaycan/ (accessed 31.05.2026).
https://islamansiklopedisi.org.tr/(date of Access:29.05.2026).
https://turkdunyasiansiklopedisi.gov.tr/hakkinda/k%C4%B1lavuz (accessed 31.05.2026).
https://turkdunyasiansiklopedisi.gov.tr/ (accessed 29.05.2026).
https://turkdunyasiansiklopedisi.gov.tr/detay/55/Abdulla-Aripov-(Abdulla-Oripov) (accessed 01.06.2026).
https://turkunib.org/tr/orhun (accessed 29.05.2026).
https://turkdunyasiansiklopedisi.gov.tr/detay/8248/Kazaklar-Ve-Kazak-Hanl%C4%B1%C4%9F%C4%B1 (accessed 01.06.2026).
https://turkdunyasiansiklopedisi.gov.tr/detay/819/Nuri-Killigil (accessed 01.06.2026)
https://turkdunyasiansiklopedisi.gov.tr/detay/8089/Yargucu-/-Yargan-/Yarg%C4%B1c%C4%B1 (accessed 01.06.2026).
https://turkdunyasiansiklopedisi.gov.tr/detay/1716/Yakut-T%C3%BCrkleri (date of Access:01.06.2026).
https://turkdunyasiansiklopedisi.gov.tr/detay/813/Bi%C5%9Fkek (accessed 01.06.2026)
https://turkdunyasiansiklopedisi.gov.tr/detay/1621/Necmettin-Hac%C4%B1emino%C4%9Flu (accessed 01.06.2026).
https://turkdunyasiansiklopedisi.gov.tr/detay/6780/Alimcan-%C4%B0drisi (accessed 01.06.2026).
https://turkdunyasiansiklopedisi.gov.tr/detay/1109/Saray-Xatun (accessed 01.06.2026).
https://says.ayk.gov.tr/says/itemPrinciplesPage (accessed 02.06.2026).
https://turkdunyasiansiklopedisi.gov.tr/mobile (accessed 02.06.2026).
https://referenceworks.brill.com/display/package/eio (Encyclopaedia of Islam Online (English) (accessed 29.05.2026).
https://turkicstates.org/tr (accessed 29.05.2026)
https://turkicstates.org/tr/iliskili-kurumlar (accessed 29.05.2026)
https://www.ayu.edu.tr/,https://tau.edu.az/,https://www.manas.edu.kg/tr, https://turkunib.org/tr (accessed 29.05.2026).
Ibn Kuteybe, Uyunu'l Ahbar, Mısır, 1963.
Kuşcalı, A., & Kocaman, M. (2025). 1924'ten bugüne bilgi yurdu: Aksaray İl Halk Kütüphanesi. Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi, 15(1), 489-509 (accessed 29.05.2026)
Meydan Larousse Büyük Lügat ve Ansiklopedi, İstanbul, Milliyet, 1999.
Odabaş, H. (2017), s.277. Kitap Okumadan Dijital Okumaya: Okuma Kültüründe ve Davranışında Gözlemlenen Değişimler. *Bilişim Teknolojilerinin Bilgi Merkezlerine ve Hizmetlerine Etkileri* içinde (ss. 272-292). Yayına Hazırlayanlar Hüseyin Odabaş ve Mehmet Ali Akkaya. İstanbul: Hiperyayın (accessed 29.05.2026).
Secundus, Gaius Plinius, Doğa tarihi : XXVIII, XXIX ve XXX. kitap hayvan organlarına dayalı devalar ve tedaviler / Gaius Plinius Secundus ; Latinceden çeviren: Sema Sandalcı ; yayına hazırlayanlar: Taşkın Takış, Ufuk Coşkun, Ankara : Doğu Batı Yayınları, 2022.
Yılında Milli Mücadele Sempozyumları Bildiriler, Atatürk Araştırma Merkezi Yayınları, Ankara 2011; Sabahattin Selek, Anadolu İhtilali, Burçak Yayınevi, 4. Baskı, İstanbul 1968 (accessed 30.05.2026).



ВЛИЯНИЕ ЦИФРОВОЙ ТРАНСФОРМАЦИИ НА КОМПЛЕКТОВАНИЕ АРХИВОВ

Наталия Геннадиевна Суровцева
Доктор исторических наук, доцент, профессор кафедры автоматизированных систем документационного обеспечения управления, Российский государственный гуманитарный университет, Россия
nataliyasurovceva@yandex.ru, ORCID-0000-0002-2942-8372

Аннотация. В статье рассматривается цифровая трансформация государственного управления, которая ведет к отказу от использования традиционных текстовых документов, предназначенных для прочтения людьми, и замене их цифровым машиночитаемым документом, пригодным для автоматизированной обработки компьютером, т.е. содержащим особым образом структурированные данные. Данная ситуация приводит к сокращению масштабов документирования, создания документов, что может составлять угрозу для комплектования государственных архивов, поскольку, в соответствии с законодательством, на хранение в архивы поступают архивные документы. Одним из самых важных является вопрос о соотношении понятий «документ» и «данные» в сфере управления, их взаимосвязи и разграничении, условия и эффективность использования и передачи на архивное хранение.

Ключевые слова: цифровая трансформация; комплектование архива; цифровой документ; документирование

ЦИФРЛЫҚ ТРАНСФОРМАЦИЯНЫҢ АРХИВТЕРДІ ЖИНАҚТАУҒА ӘСЕРІ

Наталия Геннадьевна Суровцева
Тарих ғылымдарының докторы, доцент, құжаттамалық қамтамасыз етуді басқарудың автоматтандырылған жүйелері кафедрасының профессоры, Ресей мемлекеттік гуманитарлық университеті, Ресей
nataliyasurovceva@yandex.ru; ORCID: 0000-0002-2942-8372

Аңдатпа. Мақалада мемлекеттік басқаруды цифрлық трансформациялау үдерісі қарастырылады. Бұл үдеріс адамдардың оқуы үшін арналған дәстүрлі мәтіндік құжаттарды пайдаланудан біртіндеп бас тартып, олардың орнына компьютер арқылы автоматтандырылған өңдеуге жарамды, арнайы құрылымдалған деректерді қамтитын цифрлық машинамен оқылатын құжаттарды енгізуге алып келеді. Мұндай жағдай құжаттау көлемінің және құжаттарды жасау қарқынының қысқаруына себеп болып, мемлекеттік архивтерді толықтыру ісіне қауіп төндіруі мүмкін. Себебі қолданыстағы заңнамаға сәйкес архивтерге сақтауға тек архивтік құжаттар қабылданады. Осыған байланысты басқару саласындағы «құжат» және «деректер» ұғымдарының арақатынасы, олардың өзара байланысы мен аражігін айқындау, сондай-ақ оларды пайдалану және архивтік сақтауға беру шарттары мен тиімділігі ең өзекті мәселелердің бірі болып табылады.

Түйін сөздер: цифрлық трансформация; архивті толықтыру; цифрлық құжат; құжаттау

THE IMPACT OF DIGITAL TRANSFORMATION ON COMPLETING THE ARCHIVE

Nataliya G. Surovtseva
Doctor of Historical Sciences, Associate Professor, Professor of the Department of Automated Systems for Records Management Support, Russian State University for the Humanities, Russia
nataliyasurovceva@yandex.ru; ORCID: 0000-0002-2942-8372

Abstract. The article discusses the digital transformation of public administration, which leads to the abandonment of traditional text documents intended for human reading and their replacement with digital machine-readable documents suitable for automated processing by a computer, i.e., documents containing specially structured data. This situation leads to a reduction in the scale of documentation and the creation of documents, which may pose a threat to the collection of state archives, as archival documents are required to be stored in archives under the law. One of the most important issues is the relationship between the concepts of "document" and "data" in the field of management, their interconnection and differentiation, as well as the conditions and effectiveness of their use and archival storage.

Keywords: digital transformation; archiving; digital document; documentation

Сегодня цифровая трансформация является центральным вопросом в развитии архивов, архивного дела, использования архивных документов во всем мире. Тематика докладов и выступлений на заседаниях секций Конгресса лишь подтверждает эту ситуацию. Однако хочу обратить внимание уважаемых коллег на то обстоятельство, что все восторженные и оптимистичные выступления относительно огромных возможностей, которые открывают перед архивистами цифровые технологии возможны, уместны и применимы в условиях, когда речь идет о работе с уже сформированным архивным фондом. Действительно, развитие информационно-поисковых инструментов и использования архивных документов в условиях цифровизации поднимается на новый уровень. Но я не ошибусь, если отмечу, что в программе Конгресса практически отсутствуют доклады, связанные с исследованием перспектив комплектования архивов документами в этих условиях. Между тем, цифровая трансформация именно на

этот процесс может оказать весьма существенное и неоднозначное влияние. Рассмотрим в первую очередь, что мы понимаем под цифровой трансформацией.

Характерной чертой цифровой трансформации является изменение объекта, который, собственно, и трансформируется, не просто его совершенствование, оптимизация деятельности, а именно радикальное изменение. За счет чего происходит это изменение? Обратимся к определению цифровой трансформации.

Согласно Постановлению Правительства РФ от 10 октября 2020 года № 1646, цифровая трансформация – это совокупность действий, осуществляемых государственным органом, направленных на изменение (трансформацию) государственного управления и деятельности государственного органа по предоставлению им государственных услуг и исполнению государственных функций за счет использования данных в электронном виде и внедрения информационных технологий в свою деятельность (Постановление Правительства РФ 2020). Таким образом, переход на использование данных в сфере государственного управления является основным инструментом цифровой трансформации.

Вторым весьма значимым признаком цифровой трансформации является ее всеобъемлемость. Цифровой трансформации подвергаются все сферы деятельности, включая архивы и органы государственного управления, которые являются источниками комплектования государственных архивов. Именно поэтому цифровая трансформация государственного управления окажет непосредственное влияние на комплектование государственных архивов.

Дело в том, что в результате цифровой трансформации государственного управления должно быть сформировано единое информационное пространство взаимодействия. В Российской Федерации это реализуется сегодня в рамках государственного электронного документооборота (ГосЭДО). Но вот что интересно, в ГосЭДО нет документов в привычном нам виде, там происходит обращение файлов в формате разметки. Это формат передачи информации из одной информационной системы в другую, и далеко не всегда при этом мы можем увидеть документ в привычном нам виде. Больше того, на официальном сайте ГосЭДО прямо говорится, что в рамках данной модели электронного документооборота происходит переход от использования классических текстовых документов, предназначенных для прочтения людьми, к цифровым машиночитаемым документам, пригодным для автоматизированной обработки компьютером, т.е. содержащим особым образом структурированные данные. Создание таких машиночитаемых документов также можно поручить компьютерам, для чего необходимы специальные цифровые форматы и шаблоны электронных документов (ГИС «ГосЭДО» 2). Несмотря на то, что в нормативных документах эти файлы обозначены как форматы документов (Приказ Минцифры РФ 2024), очевидно, что признаками документа они не обладают.

Здесь возникает вопрос: может ли xsd-схема документа в ГосЭДО рассматриваться в качестве основного признака документа – его формуляра в цифровой форме? На наш взгляд, ответ должен быть отрицательным. Несмотря на то, что присутствует унификация формы для передачи информации, мы не можем быстро определить, о каком документе идет речь, как это позволяет

делать формуляр, xsd-схема не является признаком документа. Такой документ создается информационной системой для другой информационной системы, которая его с легкостью прочитает. Он является «машиносоздаваемым» и одновременно «машиночитаемым» документом, т.е. данными.

Если это и есть цифровая трансформация, то документ в привычном нам смысле исчезает вообще и говорить о его традиционных признаках не приходится. Если в условиях цифровой трансформации утрачивается формуляр документа, вопросы создания документа, разработки цифровых форматов уходят в компетенцию IT-специалистов. Унифицированные системы документации и унифицированные формы документов на основе формуляра, которые разрабатывались до сих пор, уже утрачивают свое значение.

Считаю, что в этих условиях особенно важно сохранить документ как безальтернативный способ взаимодействия субъектов в общественных отношениях. Можно передать информацию по запросу или в инициативном порядке в формате XML? Можно. Можно принять решение на основе данных, полученных в формате XML? Можно. Можно отдать приказ о выполнении этого решения в формате XML? Ответ очевиден: нет.

Таким образом, на первый взгляд кажется, что цифровая трансформация практически лишает нас документа. Такая ситуация является значительным вызовом нашему профессиональному сообществу, поскольку государственный архивный фонд комплектуется документами. Опорой для нас в этих условиях должно стать понимание того, что документ является феноменом социальным. Документ – это инструмент социального взаимодействия субъектов, и в таком взаимодействии он сохраняет свои функции и характеристики. В безлюдных технологиях документа нет. Именно этот аспект в современных условиях цифровой трансформации должен стать определяющим. Однако электронная среда бытования документа ставит новые вопросы, в т.ч. связанные с цифровой трансформацией. Одним из самых важных является вопрос о соотношении понятий «документ» и «данные» в сфере управления, их взаимосвязи и разграничении, условия и эффективность использования.

Очевидно, что цифровая трансформация ведет к сокращению документирования, не в смысле записи, фиксации информации, а в смысле создания документа, обладающего формуляром. Действительно, многие категории документов (плановые, отчетно-статистические и пр.) могут безболезненно перейти в электронную среду и передаваться как данные, особенно если целью этих документов является именно передача информации. Если же требуется удостоверение этой информации, то сначала будет создан документ. Он может быть бумажным, подписанным собственноручной подписью, или электронным, подписанным электронной подписью, но мы будем видеть оформленный документ.

Так что же сегодня в цифровой среде называется документом? Ответа на этот вопрос пока нет. По привычке мы называем документом те объекты в электронной среде, которые выглядят как документы или были такими раньше. В соответствии с Правилами обмена документами в электронном виде при организации информационного взаимодействия под документом в электронном виде понимается:

1) электронный документ, состав реквизитов которого определяется правилами делопроизводства в государственных органах, органах местного самоуправления, установленными Федеральным архивным агентством в соответствии с частью 2 статьи 11 Федерального закона «Об информации, информационных технологиях и о защите информации», либо

2) электронный документ, представленный в машиночитаемой форме в цифровых форматах, в том числе разработанных и утвержденных в соответствии с Положением об информационной системе «Единое информационное пространство процессов государственного управления» (Правила 2021).

Таким образом, теперь «цифровой документ» (в формате XML) признается равным по юридической силе привычным «текстовым документам» (в формате PDF). Пока цифровые форматы разработаны для десяти видов документов, но очевидно, что данная работа будет продолжена.

Если в результате цифровой трансформации системы государственного управления произойдет замена текстового документа данными, будет ли архив комплектоваться такими «документами»? Очевидно, что сама технология их отбора учета хранения, про использование я даже не говорю, будет соответствовать хранению данных в дата-центрах. Но это уже находится в компетенции IT-специалистов и к архивному хранению вряд ли будет иметь отношение. Следовательно, цифровая трансформация рано или поздно приведет к тому, что поток передающихся на хранение документов будет уменьшатся, превратится в ручеек, а со временем может и иссякнуть. Тогда архив лишится одной из своих основных функций. Какова реальная угроза этого процесса, нам необходимо оценить уже сегодня.

Литература

ГИС «ГосЭДО» – ГИС «ГосЭДО». Форматы документов. Электронный ресурс. https://gosedo.ru/форматы-документов/ (дата обращения 22.02.2026).

Постановление Правительства РФ 2020 – Постановление Правительства РФ от 10 октября 2020 г. № 1646 «О мерах по обеспечению эффективности мероприятий по использованию информационно-коммуникационных технологий, финансовое обеспечение которых осуществляется (планируется осуществлять) за счет средств федерального бюджета и бюджетов государственных внебюджетных фондов».

Правила 2021 – Правила обмена документами в электронном виде при организации информационного взаимодействия (утв. Постановлением Правительства Российской Федерации от 24 июля 2021 г. № 1264 с изм. и доп. от 25 февраля, 4 августа 2025 г.) п. 2. Официальный интернет-портал правовой информации (pravo.gov.ru) 29.07.2021 № 0001202107290003.

Приказ Минцифры РФ 2024 – Приказ Минцифры России от 11 июля 2024 г. № 610 «Об утверждении цифровых шаблонов и форматов документов, применяемых при работе с поручениями Правительства Российской Федерации в информационной системе обеспечения внутриведомственного и межведомственного документооборота и контроля исполнения поручений, в том числе с использованием облачных сервисов» // Официальный интернет-портал правовой информации (pravo.gov.ru) 16.08.2024. № 0001202408160016

BUILDING NATIONAL ARCHIVAL STRATEGIES IN THE AGE OF AI: THE CASE OF ARMENIA

Syuzanna Sahakyan
Master's Degree in International Relations, Leading Archivist
at the Department of Scientific Research and International Relations,
National Archives of Armenia, Armenia
syuzanna.sahakyan@armarchive.am, https://orcid.org/0009-0007-3654-165X

Abstract. Nowadays, digital transformation has become a priority for national archives as they face rapid technological change, a growing volume of digital records, and increasing security risks. This research examines how national archives can develop strategies and methods for the secure integration of Artificial Intelligence and digital technologies. It addresses primary obstacles such as cybersecurity, long-term digital preservation, data migration, and protection against disinformation, misinformation, and information warfare.

Using the National Archives of Armenia as a case study, the research explores how digitization and the country's adoption of the 'Digital First' strategy (it emphasizes prioritizing technological solutions over traditional methods) are transforming the operations, staff responsibilities, and organizational workflows of state institutions. The study analyzes existing business processes in the National Archives of Armenia, identifies strengths and weaknesses, and highlights organizational and technical challenges associated with mass digitization. Special attention is given to the impact of the Republic of Armenia's legal framework, particularly new laws on public information, the Information Systems Regulatory Body, cybersecurity, digital asset management as well as the need for staff training, workflow optimization, and ensuring public access to digital archival materials. These regulations play a vital role in shaping institutional practices and guiding the modernization of the National Archives of Armenia.

The study demonstrates that digital transformation in archives extends beyond the digitization of documents. It requires the development of well-structured systems, highly skilled staff, clear distribution of responsibilities, and the implementation of advanced

technologies to automate and optimize routine processes. The study highlights the critical role of strategic planning, institutional adaptation, and continuous professional development in achieving sustainable archival transformation in the digital age.

Keywords: AI; archives; Armenia; challenges; cybersecurity; digital transformation; digitization; legal framework; optimization; training

ЖИ ДӘУІРІНДЕ ҰЛТТЫҚ АРХИВТІК СТРАТЕГИЯЛАРДЫ ҚҰРУ: АРМЕНИЯ ТӘЖІРИБЕСІ

Сюзанна Саакян
Халықаралық қатынастар магистрі, Ғылыми зерттеулер және халықаралық байланыстар бөлімінің жетекші архившісі,
Арменияның Ұлттық архивтері, Армения
syuzanna.sahakyan@armarchive.am; https://orcid.org/0009-0007-3654-165X

Аңдатпа. Қазіргі таңда цифрлық трансформация ұлттық архивтер үшін басым бағыттардың біріне айналды, өйткені олар технологиялардың қарқынды дамуы, цифрлық құжаттар көлемінің үздіксіз артуы және ақпараттық қауіпсіздікке төнетін қатерлердің күшеюі жағдайында қызмет атқаруда. Бұл зерттеуде ұлттық архивтердің жасанды интеллект пен цифрлық технологияларды қауіпсіз енгізу стратегиялары мен тетіктерін әзірлеу мәселелері қарастырылады. Зерттеу киберқауіпсіздік, цифрлық құжаттарды ұзақ мерзімді сақтау, деректерді көшіру (миграциялау), сондай-ақ дезинформация, жалған ақпарат және ақпараттық соғыс қатерлерінен қорғау сияқты негізгі мәселелерді қамтиды.

Арменияның Ұлттық архиві мысалында жүргізілген зерттеу цифрландыру үдерісі мен елде қабылданған «Digital First» («Ең алдымен – цифрлық») стратегиясының мемлекеттік мекемелердің қызметіне, қызметкерлердің функционалдық міндеттеріне және ұйымдық жұмыс процестеріне қалай ықпал етіп жатқанын талдайды. Зерттеу барысында Арменияның Ұлттық архивіндегі қолданыстағы бизнес-процестер сараланып, олардың күшті және әлсіз жақтары айқындалады, сондай-ақ жаппай цифрландырумен байланысты ұйымдастырушылық және техникалық мәселелер көрсетіледі. Ерекше назар Армения Республикасының құқықтық базасының ықпалына, атап айтқанда жария ақпарат, ақпараттық жүйелерді реттеу органы, киберқауіпсіздік, цифрлық активтерді басқару туралы жаңа заңдарға, сондай-ақ қызметкерлердің біліктілігін арттыру, жұмыс процестерін оңтайландыру және архивтік цифрлық материалдарға қоғамның қолжетімділігін қамтамасыз ету мәселелеріне аударылады. Бұл нормативтік актілер Арменияның Ұлттық архивін жаңғыртуға бағытталған институционалдық тәжірибені қалыптастыруда және дамытуда маңызды рөл атқарады.

Зерттеу нәтижелері архивтердегі цифрлық трансформацияның тек құжаттарды цифрландырумен ғана шектелмейтінін көрсетеді. Ол тиімді ұйымдастырылған жүйелерді қалыптастыруды, жоғары білікті кадрларды даярлауды, жауапкершілікті және күнделікті жұмыс процестерін автоматтандыру мен оңтайландыруға арналған заманауи технологияларды енгізуді талап етеді. Сонымен қатар зерттеу цифрлық дәуір жағдайында архивтердің тұрақты дамуын қамтамасыз етуде стратегиялық жоспарлау, институционалдық бейімделу және қызметкерлердің үздіксіз кәсіби дамуы шешуші мәнге ие екенін дәлелдейді.

Түйін сөздер: жасанды интеллект; архивтер; Армения; сын-қатерлер; киберқауіпсіздік; цифрлық трансформация; цифрландыру; құқықтық база; оңтайландыру; кадрларды даярлау.

СОЗДАНИЕ НАЦИОНАЛЬНЫХ АРХИВНЫХ СТРАТЕГИЙ В ЭПОХУ ИИ: ОПЫТ АРМЕНИИ

Сюзанна Саакян
Магистр международных отношений, ведущий архивист
Отдела научных исследований и международных связей,
Национальные архивы Армении, Армения
syuzanna.sahakyan@armarchive.am; https://orcid.org/0009-0007-3654-165X

Аннотация. В настоящее время цифровая трансформация стала одним из приоритетных направлений развития национальных архивов, которые функционируют в условиях стремительного технологического прогресса, постоянного роста объема цифровых документов и усиления угроз информационной безопасности. В настоящем исследовании рассматриваются стратегии и механизмы безопасной интеграции технологий искусственного интеллекта и цифровых технологий в деятельность национальных архивов. Особое внимание уделяется таким ключевым проблемам, как кибербезопасность, долговременное хранение цифровых документов, миграция данных, а также защита от дезинформации, ложной информации и информационных войн.

На примере Национального архива Армении исследуется влияние процессов цифровизации и реализации государственной стратегии «Digital First» («Цифровые технологии прежде всего») на деятельность государственных учреждений, функциональные обязанности сотрудников и организацию рабочих процессов. В работе анализируются существующие бизнес-процессы Национального архива Армении, выявляются их сильные и слабые стороны, а также рассматриваются организационные и технические проблемы, возникающие в условиях массовой цифровизации. Особое внимание уделено влиянию нормативно-правовой базы Республики Армения, в частности новым законам о публичной информации, регулирующим органы информационных систем, кибербезопасности и управления цифровыми активами, а также вопросам подготовки кадров, оптимизации рабочих процессов и обеспечения общественного доступа к цифровым архивным материалам. Данные нормативные акты играют важную роль в формировании институциональной практики и модернизации Национального архива Армении.

Результаты исследования показывают, что цифровая трансформация архивов не ограничивается лишь оцифровкой документов. Она предполагает создание эффективно организованных систем, подготовку высококвалифицированных специалистов, четкое распределение ответственности и внедрение современных технологий для автоматизации и оптимизации рутинных процессов. Исследование подчеркивает решающее значение стратегического планирования, институциональной адаптации и непрерывного профессионального развития сотрудников для обеспечения устойчивой трансформации архивной отрасли в условиях цифровой эпохи.

Ключевые слова: искусственный интеллект; архивы; Армения; вызовы; кибербезопасность; цифровая трансформация; цифровизация; нормативно-правовая база; оптимизация; подготовка кадров

Parallel to the advancement of cutting-edge technologies, there is an imperative to modernize and undertake comprehensive measures to support digitalization of state institutions. To keep pace with rapid technological advancements, there are some key institutional reforms undertaken by the Republic

of Armenia (RA), which are having a huge impact on the digitalization journey of state institutions. One of these is a reorganization of the Ministry of High-Tech Industry on 8 May 2019 (currently the Ministry is responsible for developing and implementing state policy in a number of strategically important areas, including high tech industry, information technologies, information security, digitalization, etc.). The establishment of the Artificial Intelligence (AI) Virtual Institute as a strategic initiative by the Ministry of High-Tech Industry also has a direct impact on the digital transformation of all state institutions in Armenia, including the National Archives of Armenia (NAA). However, the most important initiative is the creation of the Digitalization Strategy of Armenia, the Action Plan for its implementation, and the performance indicators. According to the Strategy, *Digital transformation is the shift of business processes into efficient, integrated electronic formats. This transformation is expected to be achieved through the development of new technologies, cybersecurity frameworks, data policy, e-services, and e-governance systems within the public administration. It also includes the coordination of digitization processes, the establishment of unified standards and development of a digital environment, as well as the development and implementation of initiatives that promote the adoption of digital technologies in the private sector and encourage the use of electronic tools by society.*

The objective of this research is to analyze how digitalization, new technologies, the RA's adoption of the "Digitalization Strategy", and new laws on public information, the Information Systems Regulatory Body, cybersecurity, digital asset management are transforming the operations, staff responsibilities and organizational workflows of state institutions. The relevance of this research stems from the fact that the complete digitalization of national archives, alongside the professional development of archivists and the integration of new technologies into daily operations, has become a main present-day challenge.

The digital transformation of state institutions is a continuous process, not an end goal. Therefore, it is extremely important to pay attention to the digitalization as a continuous process when developing strategies for state institutions in an era of rapidly evolving digital technologies and the large-scale implementation of AI. Furthermore, it is strongly recommended that state institutions, particularly national archives, have a digital transformation strategy alongside their core strategy, which will help to quickly align the digital strategy with the state's policies, laws, and vision over time. To ensure the national archives' strategy develops alongside quick technological advancements and serves as a practical policy instrument, it must provide a comprehensive analysis of current challenges, identify high-priority interventions, and foresee the institutional requirements of the near future. Such details have been carefully considered and continue to remain a central focus throughout all stages of digitizing internal processes and shaping the digitalization strategy at the NAA. For example, a concept paper for the reform of Armenia's archival sector and the creation of a digital archive has recently been developed in collaboration with the World Bank and the Ministry of Justice. While strategic planning is not yet a global standard for national archives, with only 11.3% of institutions publicly disclosing such roadmaps (Martínez - Cardama 2021: 1-42), can be considered Armenia's move toward a formal "Digitization Concept" represents a significant step toward institutional transparency and AI readiness.

The successful implementation of the digitalization process requires the continuous execution of four key steps:

1. Assessment - Examining the current situation, identifying existing problems and discovering both opportunities and challenges.
2. Defining the vision - Establishing the desired future state and goals. This includes preparing the legislative and regulatory framework to align with the intended outcomes.
3. Creating a roadmap - Setting priorities and timelines for execution, identifying funding sources, distributing tasks, and clarifying the scope of responsibility for those involved.
4. Implementation - Realizing the vision based on the roadmap while accounting for the issues, opportunities, and challenges identified during the assessment phase.
5. Monitoring the progress of the program - Comparing final results against the initial vision. Establishing new priorities for the next cycle of digitalization.

Actually, digitalization cannot offer solutions for all the problems arising in the bureaucratic processes of state institutions, it can even have the opposite effect, leading to the collapse of an entire organization's operations. Consequently, when undertaking digitalization and formulating strategies for the organization's near future, it is important to be guided by international and national standards and criteria, while ensuring that each organization's digitalization model matches the specific characteristics of the sector, the institution, and the details of its internal work culture. The success of digitalization depends on three key components: 1) Technological infrastructure 2) Human resources 3) Institutional and legal frameworks.

As Couture and Grimard argue, archival managers must move beyond traditional roles to define strategic visions (Couture 2003: 71), a requirement that, in the age of AI, demands a hybrid expertise in both archival theory and IT, because the implementation of even the most advanced AI models, software solutions and digital platforms will have no practical value, if the workforce lacks the specialized knowledge to operate within this integrated ecosystem. The process of making systems usable for staff and migrating workflows to a digital environment often proves more time-consuming than the design, programming, and release of the digital solutions. Without strategic investment in human capital, technological adoption will fail to achieve the desired process optimization.

The National Archives of Armenia plays an important role in preserving documents of political, scientific, historical, and cultural significance, holding over 4.2 million archival documents. Its current development strategy focuses on modernizing archives through the integration of new technologies, e-governance tools and large-scale digitization. The NAA began its digital journey in 2007, investing in high-end scanners, launching an online platform for digital services, implementing an early booking system for citizens to appoint visits. A key part of this transformation was the modernization of the legal framework to meet the specific needs of electronic records. An important step was the adoption of the Republic of Armenia Government Decision No. 884-N on July 13, 2017 (It has been updated by later decisions, such as Decision No. 1572-N of November 14, 2019). The decision established regulations for the management, preservation, and use of electronic documents. It specifically regulates the management of electronic records within organizations, the acquisition of e-

documents into the National Archival Collection, the appraisal and archival preparation of electronic records, the creation and formatting of electronic inventories, the transfer of electronic records to NAA, the preservation and security of electronic documents both within organizations and NAA, and the use of electronic archival documents. Despite these early steps, the lack of a centralized state policy on new technologies made it difficult to implement an efficient digitalization strategy. Today, however, as the government adopts a digital transformation policy with clear legislative frameworks and timelines, the NAA has entered a new phase of development.

Based on the reports from the NAA, the number of pages digitized annually was 568 740 in 2022, 452 413 in 2023, 407 064 in 2024, and 420 581 in 2025. In the present era, the necessity extends beyond simply digitizing or scanning millions of existing paper-based documents, the implementation of advanced technologies demands operational optimization. Specifically, this requires the automation of various processes to ensure rapid and accurate execution, effectively transitioning routine manual tasks into the digital ones. Furthermore, it is critical to prepare national archives for the preservation of the large volumes of digital-born documentation currently being generated.

The NAA is now undergoing a significant transition, guided by a new Digitization Concept. Beyond improving workflows, this initiative focuses on updating archival laws, enhancing employees' current abilities to improve performance and prepare for new challenges, and attracting new expertise to ensure the sustainable digital future of Armenia’s historical records. The processes of service provision of the NAA, as in a number of state institutions of RA, are currently in a transition phase: on the one hand, online tools and digital business processes are implemented, on the other hand, traditional paper-based and manual procedures are maintained. This duality leads to the complexity of processes and problems of system optimization.

For instance, in the NAA, the first and most crucial step in conducting research is to find the correct fond and list, after which to order the whole package and study the available documents. The list is considered the primary reference tool in the archive, as it serves as the foundation for developing other finding aids within the archival search-and-reference system (Avagyan 2021: 187-260). Currently, this process requires a significant amount of time and manual searching, as not all documents have been digitized (scanned) yet. Even for documents already digitized, multi-filter search capabilities are not yet available. Additionally, another issue exists: difficulties arise when performing searches using Optical Character Recognition (OCR) software on scanned documents, which affects the quality of the search results. For example, some printed lists contain handwritten corrections or additions, which are impossible to search using OCR software. There are also lists that are poorly printed, which also makes it impossible to perform digital searches on the scanned documents.

The primary consequences of the development and application of new technologies are mainly visible in the labor market. Similar to the Industrial Revolution, the current technological revolution possesses the potential to displace millions of specialists from the labor market (Harari 2018: 27) unless professionals proactively acquire new digital competencies or undergo comprehensive retraining within their specific fields. The implementation of cybersecurity and digital asset laws has

triggered a high demand for technical experts across Armenia. However, a significant obstacle remains: a shortage of qualified specialists willing to enter the public sector. State institutions, including the NAA, face a challenge in competing with the private sector, where salaries are considerably higher. This inequality creates a talent gap, making it difficult for governmental bodies to recruit the expertise necessary to implement these new digital commands.

Regardless of the sector, whether it is a state institution or a large business corporation, organizations must account for the fact that emerging technologies are reshaping established labor models and traditional paradigms, giving them new meaning, functions, and standards.

New technologies are transforming the daily professional responsibilities and functions of the staff at the NAA in several ways:

- Currently, in addition to the physical preservation of records, archival staff are now responsible for the long-term digital preservation of archival assets.
- While archival staff traditionally assisted citizens with bureaucratic procedures and physical document navigation, their role has expanded to include providing technical support for using online platforms and utilizing various digital interfaces and hardware.
- Contemporary archival functions now encompass the curation of interactive online exhibitions. This involves using data visualization tools and platforms to enhance the public accessibility and appeal of archival collections.

Obviously, the role and fundamental functions of archivists are undergoing a major evolution. Archivists are transitioning from traditional archival work (which currently involves mostly manual tasks) to a new digital generation of archival practice. We can consider the following comparative analysis, in order to better understand the transformations occurring in daily operations:

Fig. 1. *Comparison of workflows in the archive before and after digitization*

Work Performance	Traditional workflow	Post-Digital Implementation workflow
Working with documents	• The establishment of optimal conditions for the physical preservation of documents, including the regulation of humidity and illumination levels • The restoration of degraded documents, disinfection	• The establishment of optimal conditions for the physical preservation of documents, including the regulation of humidity and illumination levels • The restoration of degraded documents, disinfection • Digital preservation • File format control • Adding metadata • Cloud backup • Migrating legacy formats to new formats, • Digitizing paper-based documents

Cataloging	• Manually adding descriptions • Creating card catalogs • Filling out information in handwritten registers	• Automated indexing • Entering information into databases. • Working with data management systems • Using Optical Character Recognition (OCR) software (to turn images of text into searchable files) • Tagging information using AI • Using metadata standards (like EAD, Dublin Core or PREMIS) to organize data
Access and Usage	• Physical presence of researchers in the reading room for studying the documents • Searching through paper-based guides • Filling out paper-based order forms to study documents • Registering orders in paper registers	• Online access to lists and guides • A wide range of search tools and filters • Ordering materials and tracking the order status online
Public Engagement	• Limited space at exhibitions, events • Limited resources and space in the reading room	• Digital reading rooms • Virtual exhibitions Social media coverage (promoting the archive on social media) • Online international projects • Voluntary cooperation with citizens (volunteering online programs)
Working tools	• Finding aids for documents • Microfilm readers • Paper registers • Scanners	• Automatic book scanners (for example, the Treventus Scan Robot with "touch-free" technology) • Scanners (at least 600 dpi) • Digital platforms • Cameras • Cloud storage systems (for example, Amazon AWS, Microsoft Azure, Google Cloud) and servers • Document and data management systems (Jordana S., 2022: 43-54) • AI systems for reading handwritten texts using Big Data (for example, Digilib, Arak 29, and Calfa)

Security and Privacy	• Physical control of documents • Monitoring by security guards • Restricted access rooms.	• Cybersecurity solutions • Security cameras • Digital access control for server rooms (using electronic keys or codes) • Privacy and access management in online systems (controlling who can see what in the database).
Collaboration	• Regional or national collaboration	• Global collaboration • Joint online exhibitions • Exchange of experience
The framework of training, educational programs and courses	• Archival science • Working with documents (restoration, processing, etc.) • Document analysis	• Digital and media literacy • Programming fundamentals • Data science • Machine learning • Data protection and cybersecurity • International standards (for example, ISO 19264-1:2021 Photography - Archiving systems - Imaging systems quality analysis, ISO 15489-1:2016, Information and documentation - Records management).

Digitalization processes create new challenges not only for citizens but also for state institutions, organizations and employees. Employees are required to undergo regular training to keep pace with the times, increase their digital literacy, and acquire new professional skills to meet modern demands. For example, the core directions for training archival professionals range from basic digital literacy to machine learning. While digital literacy helps integrate computer software more deeply into daily operations, machine learning which requires more fundamental IT knowledge can facilitate process optimization. Ultimately, this can lead to using AI for the transcription and analysis of handwritten documents, as well as the automated generation of descriptions and catalogs.

Creating “digital archivists” is currently one of the strategic priorities for the National Archives of Armenia, as archival science is not taught in higher education institutions, and existing materials and literature on the subject are extremely limited and increasingly outdated. Today, there is not a single department or employee in NAA whose daily work does not involve the use of modern technology. For years, the requirements for professional personnel in Armenia's archival sector were not high because of low salaries.

According to 2025 statistical data from the NAA, the staff age distribution is primarily concentrated in the middle-to-senior age brackets. The largest demographic is represented by the 45-54 age group at 32%, followed by those aged 35-44 at 25%. The 55-63 age group accounts for 18% of the workforce, while employees aged 64 and above make up 11%. Among younger age groups, the 25-34 age group accounts for 11%, while only 3% of staff are under the age of 25.

Currently, the sector faces a shortage of personnel who possess fundamental knowledge of archival science, mastery of modern technological tools, and communication and research skills. As a result, the NAA has divided personnel-related processes into two parts: first, retraining current staff to provide them with essential digital skills and second, recruiting specialists who already possess digital and research skills and are willing to learn the fundamentals of archival science.

Using the NAA as a case study, the research concludes that sustainable archival modernization requires the simultaneous development of technological infrastructure, human capital, and legal and institutional frameworks. It indicates that Armenia has made significant progress in establishing the foundations for digital transformation. However, there are some structural challenges and issues. These include limited search functionality in digital systems, OCR-related limitations, cybersecurity risks, and the absence of full system integration. The coexistence of traditional paper-based processes with new digital systems further creates operational inefficiencies and complicates workflow optimization. Another key finding is that technological advancement alone, including advanced software solutions and digital platforms, has limited practical impact without adequately trained personnel.

Hence archival activities are going beyond the scope of only physical work with documents and are turning into a digital process of information management, the shortage of professionals with combined expertise in archival science and digital technologies represents one of the main barriers to effective transformation. This strengthens the need for continuous training and the development of "digital archivists" capable of operating in hybrid technological environments. The research also confirms that digital transformation is not a one-time reform but an ongoing process requiring continuous monitoring and strategic planning. The success of this transformation depends on institutional flexibility. To ensure the sustainability of national heritage in the digital age, the national archival strategy should prioritize the modernization of educational and professional frameworks and implement targeted measures to retain specialized expertise within the public sector.

References

Avagyan H., Instructions on the compilation of archive lists. Instructions on the nomenclature of cases and its composition, Banber Scientific Collection of Archives of Armenia, N1 (130), 2021, Yerevan. Pp. 187-260.

Couture C., Grimard J., General and specific aspects of the administration of an archival service. The management of an archival center. Essays in honor of Robert Garon, Presses de l'Université du Québec, 2003. Pp. 57–82.

Harari Y. N., 21 Lessons for the 21st Century, Random House, New York, 2018. 400 p.

Jordana S., Zabukovšekb S., Klančnikc I., Document Management System – A Way to Digital Transformation, Naše Gospodarstvo / Our Economy, Volume 68 (2), 2022. Pp. 43-54.

Law of the RA on Cybersecurity, Adopted on December 4, 2025, https://www.arlis.am/hy/acts/218672 (accessed 10.04.2026).

Law of the RA on Public Information, https://www.arlis.am/hy/acts/218678, Adopted on 4 December 2025 (accessed 10.04.2026).

Law of the RA on the Information Systems Regulatory Body, Adopted on 4 December 2025, https://www.arlis.am/hy/acts/218686 (accessed 11.04.2026).

Martínez-Cardama S., Pacios A.R., National archives' priorities: an international overview. Archival Science, 22(1), 2021. Pp. 1–42.

RA Government Decision on approving the procedure for the implementation of the information technology assets registry creation and management system, Adopted on 12 February 2026, https://arlis.am/hy/acts/220912 (accessed 11.04.2026).

RA Government Decision, No. 183-L on approving the Digitalization Strategy of Armenia, the Action Plan for its implementation, and the performance indicators, Adopted on 11 February 2021, https://www.arlis.am/hy/acts/149957 (accessed 25.03.2026).

RA Government Decision, No. 884-N, Adopted on 13 July 2017, https://www.arlis.am/hy/acts/136332 (accessed 06.05.2026).

The National Archives of Armenia's strategy, https://armarchive.am/archive_development_strategy (accessed 22.03.2026).

The official website of Ministry of High-Tech Industry of the RA, https://hightech.gov.am/en/ministry/about (accessed 18.03.2026).

Treventus-Scan Robot, 2.0 MDS, https://www.treventus.com/ (accessed 13.03.2026).

ВНЕДРЕНИЕ ЦИФРОВЫХ ТЕХНОЛОГИЙ И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В АРХИВНОЕ ДЕЛО В РЕСПУБЛИКЕ ТАДЖИКИСТАН

Мухаббат Назарали Назарализода
Кандидат исторических наук, начальник отдела Главного архивного управления при Правительстве Республики Таджикистан
Муслима Садыковна Курбанова
Советник Главного архивного управления при Правительстве Республики Таджикистан
umumi@archive.tj

Аннотация. В статье рассматриваются вопросы внедрения цифровых технологий и искусственного интеллекта в архивное дело в Республике Таджикистан в годы независимости. Автор анализирует отдельные периоды развития цифровых технологий и искусственного интеллекта в Таджикистане. Излагаются перспективы внедрения цифровых технологий и искусственного интеллекта в архивное дело в Республике Таджикистан и возникающие при этом проблемы.

Ключевые слова: независимость; информационные технологии; цифровые технологии; искусственный интеллект; архивный документ; электронный документ

ТӘЖІКСТАН РЕСПУБЛИКАСЫНДАҒЫ АРХИВ ІСІНЕ ЦИФРЛЫҚ ТЕХНОЛОГИЯЛАР МЕН ЖАСАНДЫ ИНТЕЛЛЕКТТІ ЕНГІЗУ

Мұхаббат Назарали Назарализода
Тарих ғылымдарының кандидаты, Тәжікстан Республикасы Үкіметі жанындағы Бас архив басқармасы бөлімінің бастығы
Муслима Садыққызы Құрбанова
Тәжікстан Республикасы Үкіметі жанындағы Бас архив басқармасының кеңесшісі
umumi@archive.tj

Аңдатпа. Мақалада Тәуелсіздік жылдарындағы Тәжікстан Республикасының архив ісіне цифрлық технологиялар мен жасанды интеллектті енгізу мәселелері

қарастырылады. Автор Тәжікстан Республикасында цифрлық технологиялар мен жасанды интеллекттің даму кезеңдерін талдайды. Сонымен қатар Тәжікстан Республикасының архив ісіне цифрлық технологиялар мен жасанды интеллектті енгізудің келешегі және осы үдеріске байланысты туындайтын өзекті мәселелер сипатталады.

Түйін сөздер: тәуелсіздік; ақпараттық технологиялар; цифрлық технологиялар; жасанды интеллект; архивтік құжат; электрондық құжат

THE INTRODUCTION OF DIGITAL TECHNOLOGIES AND ARTIFICIAL INTELLIGENCE IN ARCHIVAL AFFAIRS IN THE REPUBLIC OF TAJIKISTAN

Muhabbat N. Nazaralizoda
Candidate of Historical Sciences, Head of Department, Main Archival Administration under the Government of the Republic of Tajikistan
Muslima S. Kurbanova
Adviser, Main Archival Administration under the Government of the Republic of Tajikistan
umumi@archive.tj

Annotation. The introduction of digital technologies and artificial intelligence in archival affairs in the Republic of Tajikistan**.** The article examines the implementation of digital technologies and artificial intelligence in archival affairs in the Republic of Tajikistan during the years of independence. The author analyzes individual periods of development of digital technologies and artificial intelligence in the Republic of Tajikistan. The article presents the prospects for the introduction of digital technologies and artificial intelligence into archival affairs in the Republic of Tajikistan and the problems that arise in this regard.

Keywords: independence; information technologies; digital technologies; artificial intelligence; archival document; electronic document

С приобретением независимости Республикой Таджикистан возросла роль информационных технологий, которые имеют огромное значение в обеспечении информационной безопасности каждого государства.

«История мировой цивилизации свидетельствует о том, что самосознание и национальная гордость являются цементирующей основой любой уважаемой себя нации», – отмечает Основатель мира и национального единства, Лидер нации, Президент Республики Таджикистан уважаемый Эмомали Рахмон в своей книге «Таджики в зеркале истории» (Рахмон 2020).

Информационное общество практически вошло в жизнь человечества, главной особенностью которого является свободный доступ к источникам информации.

Глобальная революция, в центре которой мы сегодня находимся, привела к потере ориентации относительно части и целого, и ежедневно мы ощущаем ее новые и болезненные влияния на политическую жизнь коллектива и психологическую жизнь индивида (Нойман 1998).

По инициативе Лидера нации, Основателя мира и национального единства, Президента Республики Таджикистан уважаемого Эмомали Рахмона

Региональная Ассамблея Организации Объединенных Наций 25 июля 2025 года приняла резолюцию «Роль искусственного интеллекта в создании новых возможностей для устойчивого развития в Центральной Азии», которая была предложена Президентом Республики Таджикистан в ходе общих дебатов на 79-й Генеральной Ассамблее ООН в сентябре 2024 года. В ней предусмотрено создание Регионального центра искусственного интеллекта в г. Душанбе, который будет координировать совместные программы по внедрению соответствующих технологий в странах Центральной Азии (Генассамблея ООН).

Правительством Республики Таджикистан было принято Постановление от 30 сентября 2022 года № 483 «О Стратегии развития искусственного интеллекта в Республике Таджикистан на период до 2040 года», которое явилось первым стратегическим документом подобного уровня в Центральной Азии (Холматзода и др. 2025).

Стратегия способствует развитию интеллектуального потенциала и используется во всех сферах социально-экономического развития, учитывая процессы внедрения и использования передовых и современных технологий.

Одной из основных целей Стратегии является обеспечение отраслей, в том числе и архивной отрасли, цифровыми информационными коммуникационными технологиями.

Внедрение технологий искусственного интеллекта во всех сферах человеческой деятельности изменит структуру экономики и облегчит работу людей.

Научная революция происходит тогда, когда существующие методы решения проблем перестают работать, а также осознается, что теперешние методы не могут даже в принципе объяснить изучаемый предмет (Пинкер 2000).

В настоящее время систему архивных учреждений Республики Таджикистан составляют Главное архивное управление при Правительстве Республики Таджикистан, 13 государственных архивов и их филиалов, 64 межведомственных архива по личному составу при исполнительных органах государственной власти, в архивохранилищах которых хранятся 38 312 фондов и 11 599 825 дел. Для дальнейшего развития архивной сферы в областях, городах и районах Республики Таджикистан осуществляют свою деятельность 64 архивных сектора исполнительных органов государственной власти.

Главное архивное управление при Правительстве Республики Таджикистан является центральным исполнительным органом государственной власти, обеспечивающим в пределах своих полномочий осуществление государственной политики в части надзора и обеспечения нормативно-методической базой порядка делопроизводства и архивного дела в государственных органах, предприятиях и других организациях независимо от форм собственности и подчинения, отмечено в Положении о Главном архивном управлении при Правительстве Республики Таджикистан (Постановление Правительства РТ 2023).

Следует отметить, что в архивохранилищах Центрального государственного архива Республики Таджикистан находятся особо ценные документы досоветского периода, так, например, документы Ходжентского уездного управления периода 1868 года (Нормативные акты 2005).

Закон Республики Таджикистан «О Национальном архивном фонде и архивных учреждениях» определил новую категорию – Национальный архивный фонд Республики Таджикистан, который является составной частью отечественного и мирового исторического и культурного наследия и находится под охраной государства.

Следует отметить, что неуклонное развитие науки, техники и технологий приводит к тому, что страны мира стараются внедрить цифровую систему деятельности отраслей.

В 2004 году в Таджикистане проходила первая общенациональная конференция «Аудиовизуальная индустрия и перспективы ее развития», которая содействовала становлению и развитию медиаиндустрии и аудиовизуального бизнеса в республике. Всем известно, что защита права интеллектуальной собственности также является одним из ключевых объектов, определяющих развитие медиаиндустрии (Internews Tajikistan).

Для развития технологий искусственного интеллекта и их использования необходимо наличие высокоскоростного Интернета с мощными процессорами, подготовка высококвалифицированных кадров, наличие надежных баз данных, привлечение молодежи к изучению технических наук и программирования.

Информационные технологии вошли в нашу повседневную жизнь и продолжают развиваться. Если раньше доступ к сети Интернет осуществлялся преимущественно с помощью персональных компьютеров, то сейчас подключиться к Интернету можно через мобильные устройства (Комилов и др. 2016).

В XXI веке с развитием новых технологий и появлением Интернета информация стала контролирующим органом общества. Правительство Республики Таджикистан уделяет большое внимание внедрению информационных технологий, построению цифровой экономики и эффективному развитию электронного правительства.

Цифровые технологии и глобализация меняют человеческую жизнь. Глобализация объединяет мировое сообщество, а цифровые технологии как основной механизм осуществляют эту связь (Комилов и др. 2026).

В целях укрепления основ цифровой экономики и развития информационно-коммуникационной инфраструктуры в Республике Таджикистан функции уполномоченного государственного органа возложены на Агентство инноваций и цифровых технологий при Правительстве Республики Таджикистан, которое было создано 27 марта 2024 года. Агентство смогло сконцентрировать работу министерств и ведомств, расширить процесс реализации электронного правительства. В республике внедрен единый портал государственных услуг.

Указом Президента Республики Таджикистан от 31 января 2020 года № 1445 2020–2040 годы объявлены двадцатилетием изучения и развития естественных, точных и математических наук.

Первая лаборатория искусственного интеллекта в Центральной Азии была запущена в Таджикистане в конце 2019 года при поддержке Исламского Банка Развития.

В августе 2021 года в г. Душанбе прошел первый Форум искусственного интеллекта с участием руководителей министерств и ведомств, представителей исследовательского сообщества, частного сектора, общественных и международных организаций и мировых технологических компаний.

Разработана система электронного Правительства между Аппаратом Председателя г. Душанбе, министерствами и ведомствами и исполнительными органами государственной власти в г. Душанбе.

Проблемы в развитии отрасли искусственного интеллекта в республике в архивной сфере:

- отсутствие специалистов в области искусственного интеллекта;
- нехватка источников для финансирования архивных проектов;
- нехватка учебной литературы в области искусственного интеллекта в архивной системе;
- низкий уровень использования современных технологий цифровизации;
- низкая скорость Интернета в городах и районах республики, где находятся архивные учреждения.

Целью отрасли искусственного интеллекта является обеспечение благоприятных условий для населения за счет внедрения ключевых элементов искусственного интеллекта.

Основными задачами является подготовка высококвалифицированных специалистов в области искусственного интеллекта, подготовка молодых ученых в этой области.

В республике принят Закон Республики Таджикистан «Об электронном документе и электронной подписи» от 15 марта 2023 года, который устанавливает правовые и организационные основы порядка оборота и использования электронных документов, а также создание и применение электронной подписи в процессе формирования и использования электронных документов (Закон РТ 2023).

Согласно статье 12 вышеуказанного Закона хранение электронных документов производится организациями, осуществляющими архивную деятельность, а также деятельность по хранению документированной информации в соответствии с законодательством Республики Таджикистан. То есть речь идет о нормативно-методическом обеспечении работы с электронными документами на всех этапах их жизненного цикла: от создания до сдачи и хранения в соответствующем архиве (Наумов 2020).

Работа по использованию архивных документов должна проводиться с применением информационных технологий. Наряду с традиционными формами использования все более широкое распространение получают разнообразные интернет-проекты (Ушакова 2020).

Для расширения инновационных процессов и использования цифровых технологий по предложению Лидера нации, Основателя мира и национального единства, Президента Республики Таджикистан уважаемого Эмомали Рахмона, 2025–2030 годы были объявлены годами развития цифровой экономики и инноваций (Digital-эволюция) Эмомали Рахмона, «2025–2030 годы были объявлены годами развития цифровой экономики и инноваций» (Digital-эволюция).

Цифровизация архивов начинает активное использование инструментов искусственного интеллекта. Искусственный интеллект позволит автоматизировать и улучшить процессы, связанные с хранением, поиском и анализом архивных данных. Сделан вывод о том, что технологии искусственного интеллекта способны значительно повысить эффективность работы архивов, обеспечивая ускоренную обработку документов и удобный доступ к информации.

В настоящее время в республике ощущается недостаток квалифицированных кадров в области информационных технологий в сфере архивной службы.

Следует отметить, что искусственный интеллект становится одним из ключевых направлений развития цифровой экономики и играет важную роль в совершенствовании государственного управления.

28 апреля 2026 года было принято Постановление Правительства Республики Таджикистан «О Плане действий по цифровизации архивов государственных органов Республики Таджикистан на 2026–2030 годы» (Постановление Правительства РТ).

Данный документ предусматривает министерствам и ведомствам, местным исполнительным органам государственной власти, органам самоуправления поселков и районов республики для обеспечения выполнения данного Плана действий разработать и утвердить внутриведомственные планы действий по цифровизации архивов на каждый год. Документ разработан с целью перевести архивные документы государственных органов в электронный формат и связать цифровые архивные данные с системой госуслуг. К 2030 году архивные документы на бумажной основе в текущем делопроизводстве будут полностью исключены.

Также планируется создание межведомственной комиссии по цифровизации архивов и утверждение стандартов цифровых архивных документов и внедрение единого программного обеспечения системы цифрового архива.

Идея создания платформы в архивной сфере «Цифровые технологии и искусственный интеллект в архивном деле в Республике Таджикистан» заключается в создании современной цифровой системы, которая объединяет обучение, хранение архивных документов, интеллектуальный поиск и автоматизацию архивной работы. Главная цель платформы – повысить эффективность архивной деятельности с помощью цифровых технологий и искусственного интеллекта.

Платформа должна помогать:

1) оцифровывать архивные документы: документы на бумажной основе переводятся в электронный формат для удобного хранения и защиты;

2) распознавать текст документов: с помощью OCR-технологий система может преобразовывать сканы и изображения в текст;

3) быстро организовывать поиск архивных документов: пользователь сможет находить документы по названию, дате, автору, теме и ключевым словам;

4) автоматически классифицировать документы: искусственный интеллект сможет распределять материалы по разделам – исторические; административные; научные; учебные и др.;

5) создавать электронную базу данных: все документы будут храниться в цифровом архиве с возможностью сортировки и анализа;

6) обучать пользователей: платформа может содержать учебные модули, лекции, презентации, тексты и электронную библиотеку;

7) проводить тестирование: система автоматически проверяет ответы, показывает результат в процентах и дает рекомендации;

8) защищать архивные данные: платформа должна включать резервное копирование, разграничение доступа и защиту информации.

Необходимо отметить, что создание такой платформы является актуальным, потому что современные архивы нуждаются в автоматизации, быстром доступе к информации и надежном хранении документов.

Искусственный интеллект позволяет ускорить работу архивистов, уменьшить количество ошибок и сделать архивные материалы доступными для науки, образования и управления.

Подводя итог, следует отметить, что платформа «Цифровые технологии и искусственный интеллект в архивном деле в Республике Таджикистан» – это современная электронная среда, которая соединяет архивное дело, обучение, цифровизацию и искусственный интеллект в единую систему. Поэтапно будет проводиться оцифровка приоритетных архивов республики. Также планируется цифровизация архивов в сфере интеллектуальной собственности и обучение архивных работников цифровым технологиям.

Для обеспечения международной информационной безопасности Республика Таджикистан сотрудничает с такими организациями, как Организация Объединенных Наций (ООН), Организация по безопасности и сотрудничеству в Европе (ОБСЕ) и др.

В настоящее время в Таджикистане активно внедряются цифровые технологии в государственное управление и осуществляется переход на электронный документооборот.

Главное архивное управление при Правительстве Республики Таджикистан проводит модернизацию системы хранения, учета и поиска документов Национального архивного фонда Республики Таджикистан.

Будущее архивного дела невозможно без применения современных цифровых технологий, это требование времени, с которым столкнулись все архивы мира (Артизов 2022).

III Международный конгресс архивистов «Digital Archive. EXPO» является важным шагом в укреплении сотрудничества государств – участников данного мероприятия в развитии искусственного интеллекта в архивной сфере, обмена опытом и внедрения современных технологий искусственного интеллекта в архивную службу.

Литература

Digital – эволюция – Служба связи при Правительстве Республики Таджикистан. Digital – эволюция: как цифровые технологии меняют жизнь. https://cs.gov.tj (дата обращения: 10.05.2026).

Internews Tajikistan – Internews Tajikistan. Аудиовизуальная индустрия и перспективы ее развития.

Артизов 2022 – Артизов А.Н. «Будущее архивного дела невозможно без применения современных цифровых технологий» // Отечественные архивы. 2022. № 2. С. 6.

Генассамблея ООН – Генассамблея ООН приняла резолюцию Таджикистана об искусственном интеллекте // СНГ Сегодня.

Закон РТ 2023 – Закон Республики Таджикистан «Об электронном документе и электронной подписи» от 15 марта 2023 года.

Комилов и др. 2016 – Комилов Ф. С., Рахмонов З. Ф. Технологияхои компютерӣ. Душанбе, 2016. С. 3.

Комилов и др. 2026 – Комилов Д. А., Кобилов М. З., Султонхайдаров Н. Х. Зеҳни сунъӣ, виҷдон ва маърифати ахлоқӣ дар замони муосир. Душанбе, 2026. С. 118.

Наумов 2020 – Наумов О.В. Цифровая трансформация Росархива // Отечественные архивы. 2020. № 6. С. 7.

Нойман 1998 – Нойман О. Происхождение и развитие сознания. «Рефл-бук», «Ваклер», 1998. С. 443.

Нормативные акты 2005 – Главное архивное управление при Правительстве Республики Таджикистан. Нормативные акты государства и Правительства Республики Таджикистан по архивному делу. Сборник документов. Душанбе, 2005. С. 23.

Пинкер 2000 – Пинкер С. Теория сознания. Москва, 2000. С. 269.

Постановление Правительства РТ – Постановление Правительства Республики Таджикистан «О Плане действий по цифровизации архивов государственных органов Республики Таджикистан на 2026–2030 годы».

Постановление Правительства РТ 2023 – Постановление Правительства Республики Таджикистан от 28 ноября 2023 года № 531. Положение о Главном архивном управлении при Правительстве Республики Таджикистан.

Рахмон 2020 – Рахмон Э. Таджики в зеркале истории. От Арийцев до Саманидов. Книга первая. Нур-Султан, 2020. С. 8.

Ушакова 2020 – Ушакова Е. М.. Электронные проекты архивных учреждений Удмуртской Республики // Отечественные архивы. 2020. № 5. С. 44.

Холматзода и др. 2025 – Холматзода А.Г., Сулаймонов Ф.С., Холиков Н.А. Искусственный интеллект дар рушди ҳуқуқ. Душанбе, 2025. С. 9.

FROM HUMAN INTELLIGENCE TO PROCESS AUTOMATION AND ARTIFICIAL INTELLIGENCE: TRANSITION AT THE ARCHIVE OF THE INSTITUTE OF NATIONAL REMEMBRANCE

Paweł Popławski
M. Ed., Chief Specialist, Archive of the Institute of National Remembrance, Digital Resources Department, Poland
pawel.poplawski@ipn.gov.pl; https://orcid.org/0009-0007-1247-3744

Abstract. The Archive of the Institute of National Remembrance operates on the basis of a law specific to the Archive. The rules governing its functioning, the disclosure of archival materials, and its publications are different from those applicable to state archives. Therefore, the development of digital tools in the IPN Archive must respect Polish legislation and the protection of information (both public and classified), as well as take into account not only the associated limitations and risks, but also possibilities and opportunities.

This presentation aims to show how a system combining these three aspects can be created: human intelligence, process automation, and artificial intelligence. This sequence may represent the stages of needs and opportunities in work organization and the work system.

The starting point is human intelligence, as well as showing what the needs of the IPN Archives have been over the last dozen or so years, what solutions were sought, what problems and challenges were faced, and where things stand today.

The next stage is the automation of processes that have been developed by humans. Automation is an important element in performing repetitive tasks; these can be programmed by humans, for example, creating digital working copies, OCR, processing document scans, combining fragments of digital copies into large-format documents, or generating reports and summaries of completed work. Thanks to process automation, human intelligence does not have to focus its attention on repetitive tasks, but can concentrate on further development opportunities.

The third stage in this system is the use of AI in everyday work, which is the IPN Archive strategy for the future. As a state budget institution, the IPN Archive is primarily guided by

rational spending, so decisions to introduce new, and often very costly solutions must eliminate all risks, which is why such changes are being introduced in stages. There are plans to create an offline system (for legal reasons) based primarily on language models supported by the Ministry of Digital Affairs; this will be based on the Polish Large Language Model (PLLuM) and use the capabilities of other open-source models. This system will learn from the digital resources in our specific environment (digital copies and metadata: titles and descriptions of archival units, etc.).

Keywords: archive, artificial intelligence, automation, digital archive, digital copy, digitization, human intelligence, metadata, new technologies, records management, system, workflow

АДАМ ИНТЕЛЛЕКТІНЕН ПРОЦЕСТЕРДІ АВТОМАТТАНДЫРУ МЕН ЖАСАНДЫ ИНТЕЛЛЕКТКЕ ДЕЙІН: ҰЛТТЫҚ ЕСТЕЛІК ИНСТИТУТЫ АРХИВІНДЕГІ ТРАНСФОРМАЦИЯ

Павел Поплавски
Педагогика магистрі (M. Ed.), Ұлттық жады институты архивінің Цифрлық ресурстар департаментінің бас маманы, Польша
pawel.poplawski@ipn.gov.pl; https://orcid.org/0009-0007-1247-3744

Аңдатпа. Ұлттық жады институтының архиві (IPN Archive) осы архивке арнайы қабылданған заң негізінде қызмет етеді. Сондықтан оның жұмыс істеу қағидаттары, архивтік материалдарды жария ету тәртібі және баспа қызметі мемлекеттік архивтерге қолданылатын ережелерден ерекшеленеді. Осыған байланысты IPN архивінде цифрлық құралдарды дамыту Польша заңнамасының талаптарына, ақпаратты (ашық және құпия ақпаратты) қорғау қағидаттарына сәйкес жүзеге асырылуы тиіс. Сонымен қатар бұл үдерісте туындайтын шектеулер мен тәуекелдер ғана емес, сонымен бірге жаңа мүмкіндіктер де ескерілуі қажет.

Баяндамада адами интеллект, процестерді автоматтандыру және жасанды интеллектті біріктіретін жүйені қалыптастырудың тәсілдері ұсынылады. Бұл үш құрамдас жұмыс үдерістерін ұйымдастыру мен дамыту қажеттіліктерінің және технологиялық мүмкіндіктердің кезеңдік эволюциясын көрсетеді.

Жүйенің бастапқы негізі – адами интеллект. Осы тұрғыда соңғы он жылдан астам уақыт ішінде IPN архивінің алдында тұрған қажеттіліктер, іздестірілген шешімдер, кездескен мәселелер мен сын-қатерлер, сондай-ақ қазіргі қол жеткізілген нәтижелер қарастырылады.

Келесі кезең – адам әзірлеген процестерді автоматтандыру. Автоматтандыру қайталанатын операцияларды орындаудың маңызды құралы болып табылады. Оларға цифрлық жұмыс көшірмелерін жасау, мәтінді оптикалық тану (OCR), құжаттардың цифрлық бейнелерін өңдеу, цифрлық көшірмелердің фрагменттерін біріктіріп үлкен форматтағы құжаттар құрастыру, сондай-ақ орындалған жұмыстар бойынша есептер мен қорытындыларды автоматты түрде қалыптастыру жатады. Процестерді автоматтандырудың арқасында адам қайталанатын жұмыстарға уақыт жұмсамай, өз назарын жаңа даму бағыттарына аудара алады.

Бұл жүйенің үшінші кезеңі – жасанды интеллектті күнделікті архив қызметіне енгізу. Бұл IPN архивінің болашақтағы стратегиялық бағыты болып табылады. Мемлекеттік бюджет есебінен қаржыландырылатын мекеме ретінде IPN архиві ең алдымен бюджет қаражатын ұтымды пайдалануды басшылыққа алады. Сондықтан жаңа әрі қымбат технологиялық шешімдерді енгізу барлық тәуекелдерді барынша азайтқан

жағдайда ғана кезең-кезеңімен жүзеге асырылады. Құқықтық талаптарға байланысты интернетке қосылмайтын (offline) жүйені құру жоспарланып отыр. Бұл жүйе негізінен Польшаның Цифрландыру министрлігі қолдайтын тілдік модельдерге сүйенеді және Польша Үлкен тілдік моделінің (PLLuM), сондай-ақ басқа да ашық бастапқы кодты модельдердің мүмкіндіктерін пайдаланады. Жүйе архивтің цифрлық ресурстары (цифрлық көшірмелер, архивтік сақтау бірліктерінің атаулары, сипаттамалары және басқа метадеректер) негізінде оқытылатын болады.

Түйін сөздер: архив; жасанды интеллект; автоматтандыру; цифрлық архив; цифрлық көшірме; цифрландыру; адами интеллект; метадеректер; жаңа технологиялар; құжаттаманы басқару; жүйе; жұмыс үдерісі

ОТ ЧЕЛОВЕЧЕСКОГО ИНТЕЛЛЕКТА К АВТОМАТИЗАЦИИ ПРОЦЕССОВ И ИСКУССТВЕННОМУ ИНТЕЛЛЕКТУ: ТРАНСФОРМАЦИЯ АРХИВА ИНСТИТУТА НАЦИОНАЛЬНОЙ ПАМЯТИ

Павел Поплавски
Магистр педагогики (M. Ed.), главный специалист Архива
Института национальной памяти, Департамент цифровых ресурсов, Польша
pawel.poplawski@ipn.gov.pl; https://orcid.org/0009-0007-1247-3744

Аннотация. Архив Института национальной памяти (ИНП) функционирует на основании специального закона, регулирующего деятельность данного архива. Правила его работы, порядок раскрытия архивных материалов и издательская деятельность отличаются от норм, применяемых к государственным архивам. В связи с этим развитие цифровых инструментов в Архиве ИНП должно осуществляться в соответствии с законодательством Польши и требованиями по защите информации (как открытой, так и конфиденциальной), а также учитывать не только существующие ограничения и риски, но и открывающиеся возможности.

Настоящий доклад посвящен созданию системы, объединяющей три взаимосвязанных компонента: человеческий интеллект, автоматизацию процессов и искусственный интеллект. Такая последовательность отражает этапы развития организации труда и системы управления рабочими процессами. Отправной точкой является человеческий интеллект, а также анализ потребностей Архива ИНП за последние более чем десять лет, рассмотрение применявшихся решений, возникавших проблем и вызовов, а также оценка современного состояния развития.

Следующим этапом выступает автоматизация процессов, разработанных человеком, которая является важным инструментом выполнения повторяющихся операций. К ним относятся создание цифровых рабочих копий, оптическое распознавание текста (OCR), обработка цифровых изображений документов, объединение фрагментов цифровых копий в крупноформатные документы, а также автоматическое формирование отчетов и сводок по выполненной работе. Благодаря автоматизации человеческий интеллект освобождается от выполнения рутинных операций и получает возможность сосредоточиться на решении более сложных задач и дальнейшем развитии.

Третьим этапом является использование искусственного интеллекта в повседневной деятельности архива, что рассматривается как стратегическое направление развития Архива ИНП. Будучи государственным бюджетным учреждением, Архив ИНП руководствуется принципом рационального расходования бюджетных средств. Поэтому внедрение новых зачастую дорогостоящих технологических решений

осуществляется поэтапно с обязательной минимизацией всех возможных рисков. В перспективе планируется создание автономной (offline) системы, что обусловлено требованиями законодательства. Основу системы составят языковые модели, поддерживаемые Министерством цифровизации Польши, включая Польскую большую языковую модель (PLLuM), а также возможности других моделей с открытым исходным кодом. Предполагается, что система будет обучаться на цифровых ресурсах архива – цифровых копиях документов и метаданных, включая названия и описания архивных единиц хранения.

Ключевые слова: архив; искусственный интеллект; автоматизация; цифровой архив; цифровая копия; цифровизация; человеческий интеллект; метаданные; новые технологии; управление документами; система; рабочий процесс

The Institute of National Remembrance – Commission for the Prosecution of Crimes against the Polish Nation was established in 1999, pursuant to the Act of 1998. Its primary tasks include: recording, collecting, storing, processing, securing, making accessible, and publishing the documents of state security agencies produced and collected between 1944 and 1990, as well as those of the security agencies of the Third Reich and the Union of Soviet Socialist Republics. In addition, the Archive of the Institute of National Remembrance collects materials obtained in the form of original documents, copies, or electronic media from institutions maintaining archives abroad, Polish émigré and Polish diaspora institutions, and private individuals. The Institute's collections are currently accessible to victims of repression, journalists, researchers, and public institutions. Every citizen has the right to access documents produced by the security agencies of the People's Republic of Poland that are stored at the Institute of National Remembrance, including documents concerning individuals holding public office.

To date, the IPN Archive has amassed over 94 linear kilometers of documents, which are stored at the central archive in Warsaw and at more than a dozen branches and regional offices throughout the country. The repository includes a rich array of supporting materials, comprising card indexes, archival and inventory registers, indexes, and lists. In addition, the IPN Archive also holds audiovisual materials, which include more than 39 million photographs, over 3,000 film recordings, and a similar number of audio recordings. An interesting part of this collection consists of printed materials (posters, flyers, leaflets, etc.), as well as collections of artifacts comprising items donated by private individuals, and those seized by state security agencies as evidence in proceedings, such as medals, coins, and seals.

The entire collection has been cataloged and preserved; any materials in poor physical condition are sent to the conservation lab on an ongoing basis. In order to protect the documents while also making them more accessible to the public, the collection is being digitized. As part of this planned process, archival materials are being scanned from files containing documents of particular historical significance, as well as materials at risk of deterioration due to frequent use, and archival materials requested by individuals making inquiries at reading rooms.

Information about the holdings of the IPN Archive can be searched online. The primary tool is the Digital Archive system – a database of archival documents held in the IPN Archive. The Digital Archive system provides information on all materials stored in the IPN Archive – case files, card indexes, reference aids, audiovisual

documentation, and spatial objects. The Digital Archive system also serves to manage the circulation of archival documentation; activities related to the collection, storage, conservation, digitization, and loan of files are recorded there.

The database is directly linked to the Central Digital Repository, which stores the digital copies created. The system is equipped with a range of search tools. The Digital Archive is now an indispensable tool for the daily work of archivists at the Institute of National Remembrance, as well as a valuable resource for those who use the Institute's collections.

Due to legal restrictions, the Digital Archive system is accessible only on the Institute's internal network. Consequently, to ensure widespread access to information about the Institute's holdings, the IPN Archive maintains the IPN Archival Inventory (inwentarz.ipn.gov.pl), which is an archival inventory that is continuously updated and is intended to eventually contain descriptions of all archival documents held by the Institute. To facilitate the interpretation of the published data, the Archival Inventory includes a glossary of terms and detailed explanations regarding the functioning of state security agencies.

This brief historical background and general overview of IPN Archive activities illustrate the challenges archivists have faced in their daily work for over twenty-five years. Although the core of the Archive has remained virtually unchanged over the years, these challenges have constantly evolved over time. However, if the Archive is viewed by archivists today in the same way it was viewed twenty-five years ago, this simply means that those years have been irretrievably lost. This article aims to present a completely different perspective, namely, using the last quarter-century as an example of continuous development and looking to the future with optimism, with plans for further uninterrupted growth based on a developed system combining three aspects: human intelligence (HI), robotic process automation (RPA), and artificial intelligence (AI). These elements do not have to be mutually exclusive; they can, however, work together in a planned and coordinated manner to accomplish more according to the principle of "smarter, faster, better."

The development of the IPN Archive is entirely due to human ingenuity. All the ideas, needs, and requirements originated among the archivists. They were also constantly seeking solutions to new problems using new digital tools. There are many examples of this kind. Let us focus, however, on a few key ones.

From the very beginning, the IPN Archive built its own database resources, which were intended to organize and facilitate day-to-day work. All units of the Archive (headquarters, branches, and regional offices) maintained somewhere between a few to a dozen registers for their own use, which at its peak led to the existence of several dozen databases. Consequently, there was a need to integrate them into a single central database, on the basis of which the Digital Archive system was built.

A second example is the storage of digital copies. The IPN Archive, along with its branches and regional offices, maintained their own digital repositories at their respective locations, based in part on tape arrays and hard disk drives. There were over a dozen such repositories. In 2012, following a decision by IPN management, a Central Digital Repository was established in Warsaw; data transfer between headquarters and branches took place via external drives transported by company vehicles. Access to requested digital copies was usually available the next day, following verification

and data migration. This solution had its advantages: it was relatively cheaper than others, reliable, and scalable; however, it had limitations that hindered the Archive's development in many areas. It was therefore decided to build a new Central Digital Repository based on an object storage array, to which digital copies from all of the dozen or so repositories were migrated. Currently, digital copies are added from headquarters and branches to the Central Digital Repository in real time, and access to them is immediate. By the end of 2025, the Central Digital Repository contained over 138 million digital files, requiring over 1,558 TB (1.5 PB) of disk space.

Another notable example of progress at the IPN Archive is the way documents are made available to interested parties. Initially, readers were provided with original documents in the research rooms, and when they requested copies, they received them as paper photocopies. As the years passed and the Polish parliament introduced amendments to the Act on the Institute of National Remembrance, expanding the group of people authorized to access IPN's archival holdings, interest surged. It was then decided to purchase the first scanners to begin providing digital copies, initially on CDs and DVDs, as well as on paper printouts. As recently as ten years ago, in the Archive research rooms, archival materials were made available to readers in the following proportions: 35 percent as digital copies and 65 percent as originals. By 2023, 75 percent were already being made available as digital copies and 25 percent as originals. The increase in the availability of archival materials in digital form was achieved, among other things, thanks to an internal cloud service, through which external users receive a link to download the archival materials they have requested. Only a user's first in-person visit is required to confirm their identity and authorization to access the archival materials.

Robotic Process Automation (RPA) is used in many areas of the IPN Archive. This paper will present three key examples. The first is directly related to the development of information technology and scanning equipment. Initially, the scanners used at the IPN Archive that had an automatic document feeder scanned only one side at a time, so a document printed on both sides had to be scanned twice, and many operations on the digital copies – such as rotating the image, straightening, and cropping – were performed manually. Another example is large-format documents (maps, plans, posters, billboards, etc.), which were scanned on scanners with a maximum scanning area of A1 (594 mm x 841 mm). Documents larger than this area had to be divided into parts, which in extreme cases led to a single document being split into more than 20 parts. Archivists often combined these parts into a single digital image using external software, which usually took several hours of additional work.

Currently, digitization efforts at the Archive are being developed in such a way as to make the widest possible use of commercially available technical solutions, while ensuring the economic viability of these efforts. The scanning equipment currently in use is capable of digitizing nearly 100% of the IPN Archives' highly diverse collection, which includes, among other things: bound paper case files, ledgers, large-format documents, photographs, reel microfilms (35 mm, 16 mm, 8 mm), microfiche (jacket and diazo), documents or photographs on glass slides, audio and video tapes, and 3D objects. New technology enables faster scanning of loose documents using scanners with an automatic document feeder, which currently allows for simultaneous scanning of both sides of a document, automatically straightens and crops to the

document's edges or leaves a set margin, and automatically rotates the digital copy to the correct text orientation. Large-format documents are currently scanned on wide-format flatbed scanners, which allow for the creation of a single digital image up to sixty inches (1,524 mm) wide and up to several meters long. Wider documents are scanned in a maximum of two or three sections, which significantly simplifies their subsequent merging and saves time.

Another example of the use of automated processes is the Digital Archive system and the Central Digital Repository it contains. When an archivist with image quality verifier permissions in the system approves the scan performed by the scanner operator, the system automatically downloads reference copies of that archival unit from the shared storage space to the repository, and then creates working copies in JPG format and a working copy in PDF format based on them. A single PDF file contains all the digital images from a given archival unit; the system automatically adds the first page to the PDF file as an "Information Sheet," which contains the following metadata retrieved from the system: name of the record holder, name of the archival collection, title of the archival unit, date range, former archival reference number, creator's reference number, microfilm reference number, accession reference number, and detailed conditions for accessing the documents. The system also automatically adds a header (with the archival unit reference number), a footer (with the page number), and an "End" page to this file.

Another example of process automation is the batch processing of edited digital images. The IPN Archive holds a fairly substantial collection of microfilms produced in the 1970s and 1980s; microfilming was a key method of document archiving at that time. For the most part, these microfilms were produced carelessly – they were overexposed, underexposed, or other elements next to the document itself were photographed. Additionally, the duration and method of storage often had a detrimental effect on the condition of these microfilms.

For the purposes of this article, we can cite one example. The Józef Piłsudski Institute in America has made available a collection of microfilms, containing a Polish-American newspaper published in New York from 1920, *Nowy Świat* (Polish Morning World), daily and Sunday issues, to the IPN Archive. This collection contains 15,635 archival items (i.e., issues of the newspaper). The plan was to scan all the microfilms and return them to the depositor as soon as possible, and then proceed with the editing of the digital copies. Manual editing yielded very good results, but one person could edit only about 50 digital files per day, which was extremely time-consuming. It was therefore decided to divide the entire collection, based on the quality of the digital copies, and then subject each set of similar copies to automatic editing by adjusting brightness, contrast, shadows, mid-tones, highlights, and sharpening, as well as straightening and cropping. Thanks to these automated processes, 85,000 digital files were edited over a period of about one month. Of course, with this type of operation, there was a certain error rate – in this case, less than one percent – which required manual re-editing.

Many manual processes are currently still in various stages of automation, such as the systematic assignment of correct file names, data transfer (currently, newly created digital copies are manually transferred to a buffer located between the local disk and the repository), full validation (currently, validation checks several parameters:

file formats, file names, and compression of TIFF reference files), and basic image processing (cropping, rotating, straightening, adjusting brightness and contrast levels). The goal is to create a tool within the Digital Archive system for all tasks performed during the digitization process.

The next step in the development of the IPN Archive will undoubtedly be the gradual implementation of artificial intelligence-based solutions. Since the IPN Archive operates on a budget provided by state authorities, these expenditures must be guided primarily by rationality, and decisions to implement new, often very costly, solutions must mitigate all risks; therefore, changes are introduced in stages, following a thorough examination of each solution in a test environment. Other limitations include statutory provisions, which means that archival resources – including digital ones – cannot be fully public.

Worth noting here are the provisions of the European Union's "AI Act," which aims to create and promote human-centered, trustworthy artificial intelligence that is technologically secure and resilient, ethically sound, and used in accordance with the law. Consequently, in the broader perspective, Polish law will also be aligned with these general principles, and in the next stage, entities such as the Institute of National Remembrance will have to implement a policy regarding the use of AI within the organization. The rules concerning the IPN Archive will certainly include the protection of personal data and the protection of state secrets (even though only unclassified documents are subject to digitization); prohibited practices, risks, and the required obligations for managing those risks will be described there (e.g., a clear definition of activities which may entail high risk and therefore must be supervised and approved by a human). An essential element in all of this will certainly be maintaining the public trust of those using the archives and of institutions collaborating with the IPN Archives (e.g., by clearly indicating that specific content was generated by AI).

Another issue is the selection of tools that can be integrated into the Digital Archive system. There are many commercial large language models (LLMs) on the market that are popular and highly advanced; however, they may pose too great a risk in the context of the IPN Archive. These risks may be associated with, among other things, changes to licensing agreements regarding the price of such solutions, changes to privacy policies (which could result in data leaks), shifts in the international political landscape, and infrastructure based on foreign servers, which may limit control over the data stored there.

These risks can be mitigated through the use of the Polish Large Language Model (PLLuM), which is supported by the Ministry of Digital Affairs and developed by Polish academic and research institutions as well as other organizations. This model certainly does not compete with the tech giants in the commercial market, but it is being developed with the intention of using it in Polish public administration. It is trained in Polish, which gives it a huge advantage over models trained in foreign languages, as it understands the specifics of the Polish language, its syntax, and its inflection. This model can also serve as the basis for a specialized model, such as one based on the language used in documents of Polish civil and military security agencies stored in the IPN Archives. The choice of PLLuM represents, in a sense, the technological sovereignty of the country and its state institutions.

An important element in the implementation of artificial intelligence solutions is Retrieval-Augmented Generation (RAG), which is a method of combining a language model with an existing database. It is essential that the database be well-organized, as this will ultimately yield better results. In the Digital Archive system, databases have been maintained in a systematic and unified manner for many years; the same applies to the Central Digital Repository, where, despite the enormous collection of digital copies (nearly 140 million files), finding a copy of a specific page is no problem, as every file in the repository has a unique name and is stored in a systematic location, making it easy to locate a specific file. With a database built in this way, the LLM will be able to train more effectively, based on both existing databases and the generated digital copies, which will consequently minimize or even eliminate the possibility of AI hallucinations. The only external sources for training the LLM could be trusted online sources created by the Institute of National Remembrance or other reliable sources.

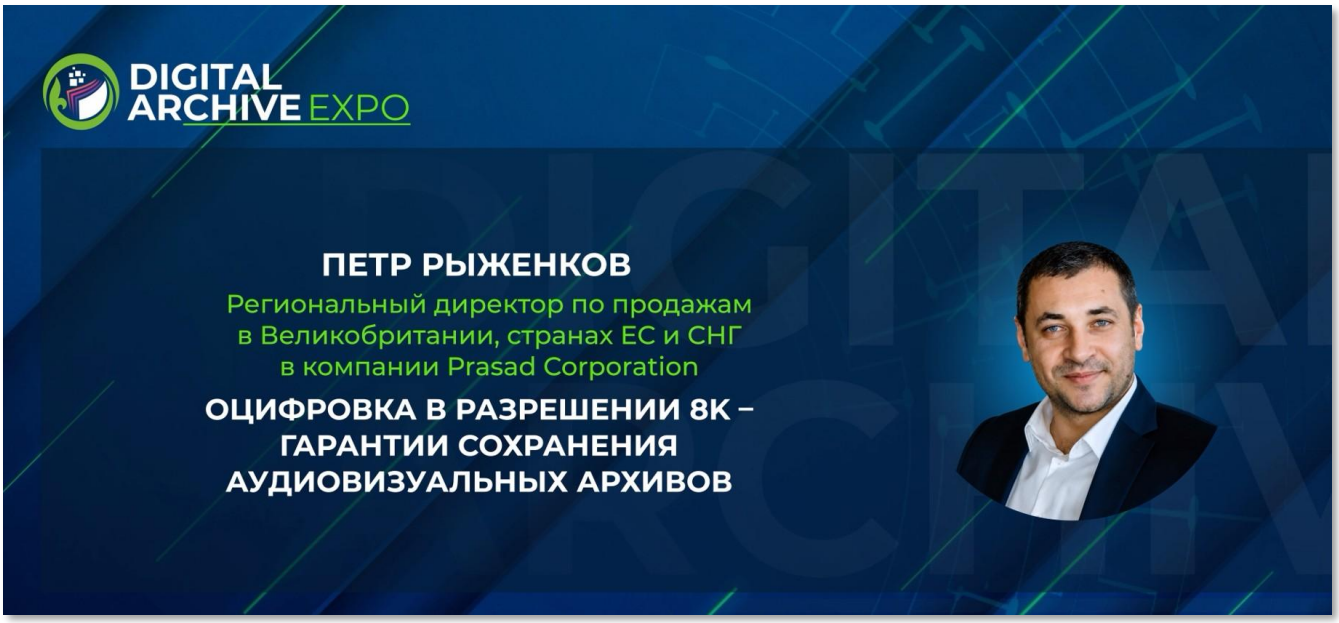
The IPN Archive test environment plays a crucial role in implementing such solutions within the Digital Archive system. As a separate infrastructure (servers, databases, etc.), it mirrors the production environment and is used to verify the functionality of new solutions before they are deployed in production. Finding errors in the test environment helps to prevent failures in the production environment.

When designing a strategy for implementing AI solutions in the Digital Archive system, the first step should be to specify the tasks that artificial intelligence could perform. Examples include transcribing audio and video recordings, transliterating other alphabets into the Latin alphabet, and generating descriptions of photographs (the latter would require human verification and approval). An important and planned change in the Digital Archive system is optical character recognition (OCR) in the digitized collection. OCR will be used for easy searching and creating descriptions within the system: at the level of archival units, personal indexes, and card index cards. All proposed descriptions will be verified by a human.

Another potential use could be the analysis of repetitive documents, such as index cards, from which individual fields could be automatically extracted, and then enter that data into a form; the only remaining step would be to verify that the form has been filled out correctly (currently, this data is entered manually by a person who transcribes it from a digital copy of the index card).

At the moment, each archival item has numerous attributes in the database, including the title and description of the archival item. Artificial intelligence, based on metadata and a digital copy of a given archival unit, could generate some kind of summary of the content. The long-term goal of using AI in the Digital Archive system is to enable searching for information at the document level, rather than at the metadata level as has been the case so far.

Artificial intelligence in the Digital Archive system does not have to take the form of an archivist's assistant or a chatbot that archivists use to solve problems or obtain information. Rather, artificial intelligence in the Digital Archive system should be a machine operating in the background, based on programmed algorithms and the principles of mathematics and statistics. It should reduce manual labor, automate repetitive tasks, optimize processes, increase efficiency, and minimize human error.



8K-ОЦИФРОВКА АУДИОВИЗУАЛЬНЫХ АРХИВОВ КАК СТРАТЕГИЯ СОХРАНЕНИЯ И МОНЕТИЗАЦИИ НАСЛЕДИЯ

Петр Рыженков
Международный эксперт по цифровой реставрации
и сохранению аудиовизуального наследия, Prasad Corporation, Индия
petr.ryzhenkov@prasadcorp.com, www.prasadcorp.com

Аннотация. В докладе рассматривается 8K-оцифровка аудиовизуальных архивов как стратегический инструмент сохранения культурного наследия и формирования повторно используемых цифровых активов. На основе предоставленных рабочих материалов и презентации анализируются риски физической деградации кинопленки, магнитных лент и иных аналоговых носителей, а также рост спроса на исторический контент со стороны телевидения, документального кино, цифровых музеев, образовательных платформ, OTT-сервисов и систем искусственного интеллекта. Особое внимание уделяется концепции Scan Once, согласно которой при первом обращении к оригинальному носителю создается мастер-файл максимально возможного качества, пригодный для реставрации, хранения, доступа, лицензирования и будущих технологических применений. Показано, что 8K следует рассматривать не как избыточный формат показа, а как архивный стандарт извлечения и сохранения информации. Расширенно анализируется потенциал Центральной Азии, где уникальные фонды кинохроники, национального кино, телевидения, документальных и этнографических материалов пока не охвачены едиными национальными программами массовой 8K-оцифровки. Отсутствие общих технических требований, долгосрочного финансирования, центров компетенций и подготовленных кадров приводит к фрагментарности цифровых коллекций и риску повторной оцифровки. Предлагается архитектура национальной программы, включающая инвентаризацию, приоритизацию, единый технический стандарт, устойчивое хранение, подготовку специалистов, метаданные, искусственный интеллект и механизмы доступа. Делается вывод о необходимости воспринимать цифровые архивы как элемент национальной цифровой инфраструктуры и культурного суверенитета.

Ключевые слова: аудиовизуальные архивы; 8K; цифровое сохранение; Scan Once; национальная программа; цифровая реставрация; искусственный интеллект; цифровой суверенитет; культурное наследие; Центральная Азия

МӘДЕНИ МҰРАНЫ САҚТАУ ЖӘНЕ КОММЕРЦИЯЛАНДЫРУ СТРАТЕГИЯСЫ РЕТІНДЕ АУДИОВИЗУАЛДЫ АРХИВТЕРДІ 8K ФОРМАТЫНДА ЦИФРЛАНДЫРУ

Петр Рыженков
Аудиовизуалды мұраны цифрлық реставрациялау және сақтау саласындағы халықаралық сарапшы, Prasad Corporation, Үндістан
petr.ryzhenkov@prasadcorp.com, www.prasadcorp.com

Аңдатпа. Баяндамада аудиовизуалды архивтерді 8K форматында цифрландыру мәдени мұраны сақтаудың және қайта пайдалануға болатын цифрлық активтерді қалыптастырудың стратегиялық құралы ретінде қарастырылады. Ұсынылған материалдар мен презентация негізінде кинотаспаның, магниттік таспалардың және басқа да аналогтық тасымалдағыштардың физикалық тозу қаупі, сондай-ақ тарихи контентке теледидар, деректі кино, цифрлық музейлер, білім беру платформалары, OTT-қызметтері мен жасанды интеллект жүйелері тарапынан сұраныстың артуы талданады. Scan Once тұжырымдамасына ерекше назар аударылады. Бұл тұжырымдамаға сәйкес бастапқы тасымалдағышқа алғаш рет қол жеткізген кезде реставрациялау, ұзақ мерзімді сақтау, қолжетімділікті қамтамасыз ету, лицензиялау және болашақ технологиялық мақсаттарда пайдалану үшін мүмкіндігінше ең жоғары сападағы мастер-файл жасалады. Зерттеу нәтижелері 8K технологиясын артық көрсетілім форматы ретінде емес, ақпаратты барынша толық алуға және ұзақ мерзімге сақтауға арналған архивтік стандарт ретінде қарастыру қажеттігін көрсетеді. Сонымен қатар, кинохроника, ұлттық кино, теледидар, деректі және этнографиялық материалдардың бірегей қорлары сақталғанымен, оларды жаппай 8K форматында цифрландырудың бірыңғай ұлттық бағдарламалары әлі қалыптаспаған Орталық Азияның әлеуеті жан-жақты талданады. Бірыңғай техникалық талаптардың, ұзақ мерзімді қаржыландырудың, құзыреттер орталықтарының және мамандардың болмауы цифрлық коллекциялардың бытыраңқылығына және қайта цифрландыру қаупіне алып келеді. Осыған байланысты түгендеуді, басымдықтарды айқындауды, бірыңғай техникалық стандартты, ұзақ мерзімді сақтауды, мамандарды даярлауды, метадеректерді, жасанды интеллектті және қолжетімділік тетіктерін қамтитын ұлттық бағдарламаның архитектурасы ұсынылады. Қорытындыда цифрлық архивтерді ұлттық цифрлық инфрақұрылымның және мәдени егемендіктің маңызды құрамдас бөлігі ретінде қарастыру қажеттігі негізделеді.

Түйін сөздер: аудиовизуалды архивтер; 8K; цифрлық сақтау; Scan Once; ұлттық бағдарлама; цифрлық реставрация; жасанды интеллект; цифрлық егемендік; мәдени мұра; Орталық Азия

8K DIGITIZATION OF AUDIOVISUAL ARCHIVES AS A STRATEGY FOR PRESERVATION AND MONETIZATION

Petr Ryzhenkov
International Expert in Digital Restoration and Preservation of Audiovisual Heritage, Prasad Corporation, India
petr.ryzhenkov@prasadcorp.com, www.prasadcorp.com

Abstract. The paper examines 8K digitization of audiovisual archives as a strategic instrument for preserving cultural heritage and creating reusable digital assets. Based exclusively on the supplied working papers and presentation materials, the study analyzes the physical degradation of film, magnetic tape and other analogue carriers, together with the growing demand for historical content from broadcasters, documentary producers, digital museums, educational platforms, OTT services and artificial intelligence systems. Particular attention is paid to the Scan Once concept, under which a master file of the highest possible quality is created during the first handling of the original carrier and then used for preservation, restoration, access, licensing and future technological applications. The paper argues that 8K should be regarded not as an excessive display format, but as an archival standard for extracting and preserving information. The Central Asian dimension is examined in greater depth: unique collections of newsreels, national cinema, television, documentary and ethnographic materials are not yet covered by unified national programmes for mass 8K digitization. The absence of common technical requirements, long-term funding, competence centres and trained personnel leads to fragmented digital collections and the risk of repeated digitization. The paper proposes an architecture for national programmes combining inventory, prioritization, a unified technical standard, sustainable storage, professional training, metadata, artificial intelligence and access mechanisms. It concludes that digital archives should be treated as an element of national digital infrastructure and cultural sovereignty.

Keywords: audiovisual archives; 8K; digital preservation; Scan Once; national programme; digital restoration; artificial intelligence; digital sovereignty; cultural heritage; Central Asia

Введение

Аудиовизуальное наследие является одной из наиболее насыщенных форм исторической памяти. Кинофильмы, телевизионные программы, документальная хроника, фонограммы и видеодокументы фиксируют не только событие, но и визуальную среду эпохи: лица, речь, архитектуру, одежду, производственные процессы, природные ландшафты и эмоциональный контекст. Поэтому такие материалы имеют культурную, научную, образовательную и экономическую ценность.

В XXI веке архивы перестают быть исключительно учреждениями хранения. Они постепенно превращаются в цифровые центры, обеспечивающие сохранение, описание, обработку, распространение и повторное использование контента. Рост интереса к документальному кино, историческим фильмам, редким хроникам, цифровым музеям и потоковым платформам создает новые требования к качеству архивных цифровых копий.

Одновременно физические носители продолжают стареть даже при соблюдении нормативных условий хранения. Кинопленка подвержена уксусному синдрому, усадке, механическим повреждениям, деформации и потере цветовой информации. Магнитные носители страдают от деградации связующего слоя, размагничивания и снижения стабильности воспроизведения. По этой причине цифровое сохранение уже не может рассматриваться как вспомогательная техническая задача, оно становится частью стратегии защиты национальной культурной памяти.

Для стран Центральной Азии данная проблема имеет особое значение. В регионе сохраняются значительные фонды кинохроники, телевизионных

программ, документальных фильмов, этнографических съемок, видеокассет, магнитных лент, микрофильмов, научных и образовательных материалов. Эти коллекции отражают формирование национальных государств, развитие промышленности, науки и культуры, исторические изменения городской и сельской среды, национальные языки и повседневную жизнь XX и начала XXI веков. Однако отсутствие комплексных программ массовой 8K-оцифровки, дефицит оборудования и недостаток специалистов ограничивают возможности долгосрочного сохранения и международного использования этих фондов.

Цель доклада состоит в том, чтобы на основе предоставленных источников обосновать значение 8K-оцифровки как стратегии сохранения и монетизации аудиовизуального наследия, раскрыть нереализованный потенциал Центральной Азии и предложить модель перехода от разрозненных проектов отдельных учреждений к национальным программам массовой оцифровки, основанным на принципе Scan Once.

Методология исследования

Исследование основано на сравнительном и системном анализе предоставленных материалов, посвященных цифровому сохранению аудиовизуального наследия, концепции Scan Once, международному опыту цифровой реставрации, перспективам искусственного интеллекта и особенностям развития архивов Центральной Азии. Использованы три рабочих текста докладов, подготовленных для Конгресса, а также презентационные материалы о 8K-оцифровке и практических возможностях компании Prasad Corporation.

Методологически анализ соединяет три уровня. Первый уровень связан с физическим состоянием аналоговых носителей и рисками утраты информации. Второй уровень относится к технологиям сканирования, цифровой реставрации, созданию архивного мастер-файла и подготовке производных версий. Третий уровень рассматривает институциональные, культурные и экономические условия использования цифровых коллекций, включая единые стандарты, государственные программы, лицензирование, телевещание, OTT-платформы, образование, цифровые музеи и системы искусственного интеллекта.

Отдельное внимание уделено результатам профессионального изучения архивной инфраструктуры Казахстана, Кыргызстана, Узбекистана и Таджикистана, упомянутого в предоставленных материалах. В настоящем докладе эти результаты используются не для ранжирования государств, а для выявления общих региональных проблем: неодинакового уровня цифровизации, отсутствия комплексных национальных программ, ограниченности высокопроизводительного оборудования, дефицита специалистов и слабой интеграции оцифрованных коллекций в цифровую экономику.

Такой подход позволяет рассматривать 8K не только как технический параметр изображения, но и как инструмент управления жизненным циклом архивного объекта: от физического носителя – к цифровому мастер-файлу, от мастер-файла – к реставрационной версии, от реставрации – к доступу, публикации, лицензированию и повторному использованию.

Концепция Scan Once как основа цифрового сохранения

Одним из ключевых принципов современной архивной цифровизации становится концепция Scan Once, т.е. «сканировать один раз». Ее смысл заключается в том, чтобы уже при первом обращении к оригинальному носителю создать цифровой мастер-файл максимально возможного качества. Такой мастер должен служить источником для всех последующих производных версий, включая копии для реставрации, телевещания, интернет-дистрибуции, кинопроката, образовательных сервисов и исследовательского использования.

Традиционная практика цифровизации часто была ориентирована на текущие стандарты распространения контента. В разные периоды такими стандартами становились SD, HD, 2K и 4K. Однако развитие технологий отображения и обработки изображения происходит быстрее, чем реализуются государственные программы архивной оцифровки. Поэтому материалы, оцифрованные в недостаточном разрешении, через 10-15 лет могут потребовать повторного сканирования.

Повторное сканирование связано с несколькими рисками. Во-первых, оно увеличивает механическую нагрузку на оригинал. Во-вторых, требует новых бюджетных расходов. В-третьих, к моменту повторного обращения физическое состояние пленки может ухудшиться настолько, что получить изображение высокого качества уже не удастся. Именно поэтому рациональная стратегия должна быть направлена не на многократное возвращение к оригиналу, а на создание максимально качественного цифрового архива с первого раза.

Архивный мастер-файл 8K в этой логике выполняет роль эталонного цифрового источника. Из него могут создаваться версии Super4K, 4K, 2K, HD и другие производные форматы без повторного контакта с оригинальной пленкой. Тем самым Scan Once переводит архивную работу из модели «создания цифровой копии для текущих нужд» в модель «формирования цифрового актива долгосрочного значения».

Почему именно 8K

Выбор 8K в контексте архивного сохранения определяется не стремлением получить максимально эффектное изображение для сегодняшнего просмотра. Основная задача состоит в максимально полном извлечении информации, содержащейся на оригинальном носителе. Архивный мастер должен сохранять не только визуальное содержание кадра, но и структуру кинозерна, микродетали изображения, особенности эмульсионного слоя, цветовые нюансы и динамический диапазон.

В предоставленных материалах подчеркивается, что 8K позволяет сформировать запас качества на десятилетия вперед. Если 4K уже стал массовым стандартом, то 8K постепенно входит в сферу вещания, потоковых сервисов, тестовых трансляций, IP-доставки UHD-контента и будущих сетей связи. Материал, оцифрованный сегодня в 2K или 4K, может оказаться недостаточным для новых платформ, тогда как 8K-мастер сохраняет более высокий потенциал адаптации.

Особое значение имеет отличие архивного сканирования от создания конечной пользовательской версии. Для публичного показа, телевидения или интернет-дистрибуции часто достаточно производного файла более низкого

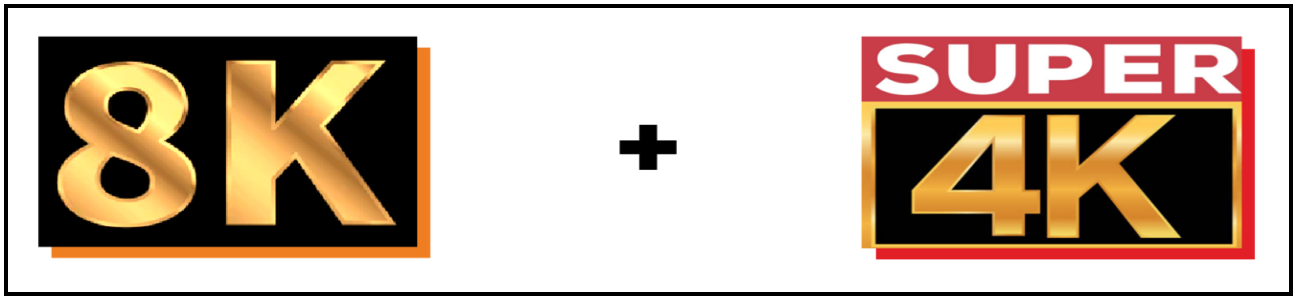
разрешения. Однако архивный мастер должен быть выше текущих стандартов распространения, поскольку он предназначен не только для сегодняшнего использования, но и для будущих технологических систем.

Таблица 1. *Сравнение архивного потенциала основных форматов сканирования*

Параметр	2K	4K	8K
Назначение	текущий доступ и просмотр	современное распространение	архивный мастер и будущие производные версии
Информационный запас	ограниченный	достаточный для многих текущих задач	максимальное сохранение деталей, зерна, цвета и динамического диапазона
Риск повторного сканирования	высокий	средний	минимальный при корректном хранении мастер-файла

Примечание. Таблица составлена по материалам, в которых 8K рассматривается как формат архивного мастер-файла, а 2K и 4K – как форматы текущего доступа и распространения (Рыженков, Prasad Corp).

Рис. 1. *Схема двойной стратегической ценности архивного мастер-файла 8K*



Разрешение [пиксели]		
12K	8K	4K
12288 x 9336	8192 × 6224	4096 x 3112

8K как двойная ценность: сохранение и монетизация

Государственные программы цифрового сохранения часто воспринимаются как расходы на защиту культурного наследия. Такой подход оправдан, но он не отражает всей стратегической ценности высококачественной оцифровки. Создание архивного мастер-файла 8K одновременно решает две взаимосвязанные задачи: сохраняет национальное аудиовизуальное наследие

для будущих поколений и формирует цифровой актив, который может использоваться для монетизации.

Первая задача связана с сохранением. Мастер 8K фиксирует визуальную информацию оригинального носителя в максимально полном объеме и снижает необходимость повторного обращения к пленке. Это особенно важно для старых и поврежденных материалов, поскольку каждый новый проход через сканер или воспроизводящее оборудование увеличивает риск механического повреждения.

Вторая задача связана с экономическим использованием. Высококачественный цифровой мастер может многократно использоваться для создания разных продуктов: реставрационной копии, телевизионного мастера, версии для OTT-платформы, фестивального DCP, образовательного фрагмента, музейной инсталляции или лицензируемого архивного клипа. Чем выше качество исходного файла, тем шире круг потенциальных потребителей и тем дольше жизненный цикл цифрового актива.

В презентационных материалах отмечается, что 8K рассматривается как инфраструктура будущего контента, а оцифрованный в 8K материал способен сохранять коммерческую пригодность в течение длительного периода. В рамках предоставленных источников эта логика напрямую связывает сохранение и монетизацию: архив не только защищает историческую память, но и создает основу для будущего дохода, международного обмена и новых культурных продуктов.

Искусственный интеллект и новые требования к архивным данным

Развитие искусственного интеллекта меняет представление о ценности архивных коллекций. Если раньше основными пользователями архивов были историки, журналисты, кинематографисты, исследователи и учреждения культуры, то сегодня к ним добавляются интеллектуальные цифровые системы. Алгоритмы компьютерного зрения, распознавания речи, автоматического описания, машинного перевода и тематической классификации нуждаются в больших массивах качественных данных.

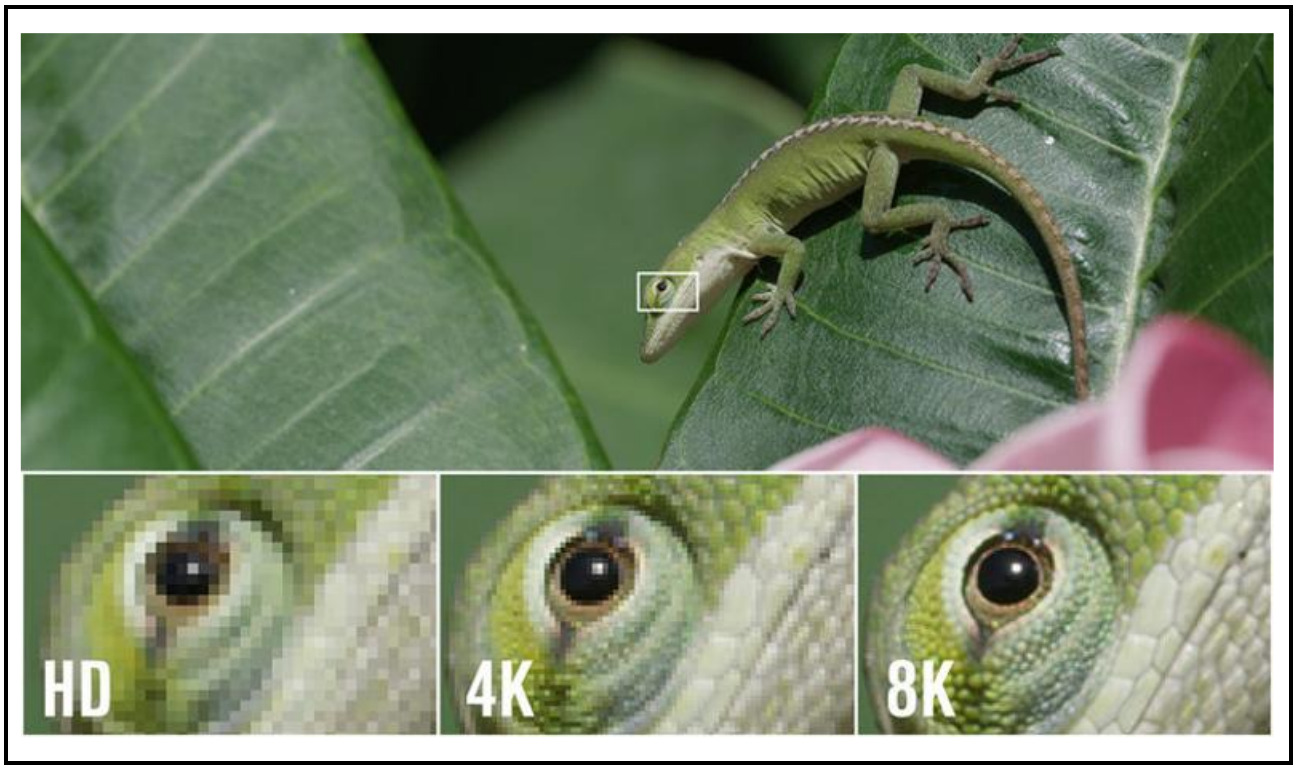
Аудиовизуальные архивы обладают особой ценностью для таких технологий, поскольку фиксируют не только содержание событий, но и визуальную среду эпохи: лица, одежду, архитектуру, транспорт, производственные процессы, ландшафты, национальные языки и бытовые детали. При низком разрешении теряются мелкие объекты, текстовые элементы, особенности лиц, детали фона и структуры изображения. Эти потери невозможно полностью компенсировать последующей обработкой.

Высококачественный цифровой архив может рассматриваться как национальный массив данных. На его основе возможны автоматическое распознавание речи, формирование субтитров, поиск объектов и персон, тематическая классификация, выявление географических локаций и создание расширенных метаданных. При этом качество результата ограничивается качеством исходного скана: сведения, не зафиксированные при оцифровке, не могут быть достоверно восстановлены позднее.

Особое значение имеет автоматическое создание метаданных. Многие аудиовизуальные фонды описаны неполно или представлены только в бумажных

каталогах. Искусственный интеллект способен ускорить распознавание объектов, персон, речи, локаций и исторических событий, но для этого архив должен иметь высококачественные цифровые исходники. Следовательно, 8K-оцифровка является не только мерой сохранения, но и подготовкой архивов к новой эпохе интеллектуального поиска и цифрового доступа.

Рис. 2. *Связь между качеством 8K-сканирования и возможностями искусственного интеллекта*



Потенциал аудиовизуального наследия Центральной Азии

Потенциал Центральной Азии определяется не только объемом сохранившихся фондов, но и уникальностью содержащейся в них информации. Кинохроника, художественные и документальные фильмы, телевизионные передачи, этнографические съемки, фонограммы, материалы научных и образовательных учреждений образуют визуальную и звуковую историю региона. В совокупности они документируют государственное строительство, индустриализацию, развитие науки и образования, изменение городов и сельских территорий, традиционные формы культуры, языковое многообразие и трансформацию повседневной жизни.

Особую ценность представляют материалы, связанные с историей Шелкового пути, советским кинонаследием, национальными кинематографиями, документальной хроникой XX века и уникальными языковыми архивами. Для международной аудитории такой контент может выполнять функцию первичного исторического источника, поскольку передает не только факт события, но и его визуальную, звуковую и эмоциональную среду. Именно эта многослойность делает аудиовизуальные коллекции

востребованными в документальном кино, цифровой гуманитаристике, музейных проектах и образовании.

В предоставленных материалах отмечается, что исторические и документальные материалы Центральной Азии пока крайне редко присутствуют на европейских телеканалах, международных стриминговых сервисах, крупных документальных платформах и фестивалях архивного кино. Такое положение не свидетельствует об отсутствии интереса к региону; напротив, оно указывает на недостаточную техническую и организационную готовность коллекций к международному использованию. Без высококачественного мастер-файла, профессиональной реставрации, метаданных, ясного правового статуса и подготовленных версий контент остается фактически невидимым для потенциального иностранного пользователя.

Оцифровка в 8K способна изменить эту ситуацию, поскольку создает единый исходный ресурс для разных сценариев. Один архивный мастер может служить основой для реставрации Super4K, телевизионного показа, фестивального DCP, образовательной версии, музейной экспозиции и лицензируемого фрагмента. В результате национальное наследие получает возможность многократного использования без нового обращения к оригинальному носителю.

Экономический потенциал региона не следует понимать только как прямую продажу архивных кадров. Он включает производство новых документальных фильмов, совместные международные проекты, создание тематических цифровых коллекций, развитие образовательных платформ, использование материалов телевещателями и OTT-сервисами, а также подготовку данных для систем искусственного интеллекта. Каждый из этих сценариев требует не только наличия контента, но и технической надежности, поисковой доступности и документированного происхождения цифрового файла.

Культурный потенциал связан с формированием международного образа государств региона. Профессионально восстановленные фильмы и хроника могут участвовать в фестивалях, демонстрироваться в кинотеатрах, включаться в программы культурных институтов и использоваться в международном образовании. Через такие проекты государства получают возможность рассказывать о собственной истории, продвигать национальные языки, искусство и традиции, а также расширять присутствие в мировом культурном пространстве.

Наконец, аудиовизуальные архивы имеют значение для цифрового суверенитета. Они содержат уникальное цифровое отражение исторической памяти государства и могут становиться национальными наборами данных для исследований и искусственного интеллекта. Утрата оригинала или создание низкокачественной цифровой копии ограничивает не только культурный доступ, но и будущие научные и технологические возможности. Поэтому сохранение архивов в максимальном качестве следует рассматривать как часть государственной информационной инфраструктуры.

Отсутствие единых национальных программ и стандартов

На момент подготовки доклада в предоставленных источниках не зафиксировано наличие в исследованных государствах Центральной Азии комплексной национальной программы массовой 8K-оцифровки,

охватывающей киноархивы, телерадиофонды и другие держатели аудиовизуального наследия. Цифровизация во многих случаях остается инициативой отдельных учреждений или ограничивается отдельными фильмами и коллекциями. Такой подход позволяет решать локальные задачи, но не формирует устойчивую систему национального масштаба.

Проблема состоит не только в отсутствии одного документа, содержащего термин «8K». Речь идет об отсутствии согласованной цепочки решений: государственного мандата, инвентаризации фондов, правил приоритизации, технического стандарта, многолетнего финансирования, инфраструктуры долговременного хранения, системы подготовки кадров и модели дальнейшего использования цифровых коллекций. Если хотя бы один из этих элементов не предусмотрен, массовая оцифровка превращается в набор несвязанных проектов.

Отсутствие единого технического стандарта означает, что разные учреждения могут выбирать неодинаковое разрешение, глубину цвета, цветовое пространство, форматы файлов, правила именования, структуру метаданных, процедуры контроля качества и резервирования. В результате цифровые копии могут быть достаточными для локального просмотра, но непригодными для профессиональной реставрации, объединенного поиска, международного обмена или повторного использования на новых платформах.

Фрагментарность особенно опасна для пленочных коллекций. Если материал один раз оцифрован только для текущего доступа, учреждение может считать задачу завершенной. Через несколько лет возникает потребность в более высоком качестве, однако оригинал к этому времени продолжает стареть. Повторное сканирование требует нового бюджета, нового цикла подготовки и дополнительного механического контакта с пленкой. Тем самым отсутствие стандарта прямо противоречит принципу Scan Once.

Неодинаковые подходы к метаданным создают отдельный барьер. Даже качественный мастер-файл не становится полноценным цифровым активом, если он не связан с описанием содержания, сведениями о носителе, техническими параметрами, правовым статусом и историей обработки. При разрозненных системах описания невозможно создать единый поисковый контур, а автоматические инструменты искусственного интеллекта не могут эффективно сопоставлять данные между учреждениями.

Отсутствие долгосрочного финансирования приводит к ориентации на разовые закупки оборудования. Однако сканер сам по себе не образует систему сохранения. Для устойчивой работы необходимы инспекция и подготовка носителей, очистка, сканирование, цветокоррекция, восстановление изображения и звука, контроль качества, создание метаданных, хранение мастер-файлов и регулярная проверка их целостности. Если бюджет предусматривает только один элемент цепочки, производственный процесс остается незавершенным.

Еще одно последствие разрозненности – невозможность объективно определить национальные приоритеты. Отдельные учреждения оцифровывают то, что соответствует их текущим задачам, тогда как наиболее уязвимые или уникальные материалы могут оставаться без внимания. Национальная

программа должна сопоставлять физический риск, культурную значимость, уникальность оригинала, востребованность и правовую готовность, чтобы распределять ресурсы на основе прозрачной приоритизации.

Таким образом, отсутствие программы массовой 8K-оцифровки создает двойной разрыв. Первый разрыв является технологическим: качество и совместимость цифровых материалов не соответствуют будущим требованиям. Второй разрыв – институциональный: государство не получает целостного представления о собственном аудиовизуальном наследии и не может управлять им как национальным цифровым ресурсом.

Таблица 2. *Последствия отсутствия единого национального стандарта 8K-оцифровки*[1*]

Область	Последствие разрозненного подхода	Требуемое системное решение
Качество сканирования	различные разрешения и параметры, ограниченная пригодность для реставрации	единые требования к архивному мастер-файлу и контролю качества
Хранение	файлы создаются без согласованного резервирования и проверки целостности	национальная политика долговременного хранения и миграции данных
Метаданные	несовместимые описания и невозможность объединенного поиска	единая структура архивных, технических и правовых метаданных
Финансирование	разовые закупки и незавершенные производственные цепочки	многолетняя программа с учетом полного жизненного цикла
Кадры	зависимость от отдельных специалистов и поставщиков	постоянная подготовка междисциплинарных команд
Использование	цифровые копии остаются локальными и мало востребованными	производные версии для образования, вещания, музеев и лицензирования

Технологический и кадровый разрыв

В предоставленных материалах одним из главных ограничений назван недостаток современных RGB-сканеров уровня 8K. Во многих учреждениях региона отсутствуют 8K-комплексы, продолжают использоваться устаревшие телесине-системы или решения с байеровскими сенсорами, а линии фотохимической подготовки пленки развиты недостаточно. Это затрудняет получение точной цветопередачи, полной детализации и стабильного результата при работе с поврежденными материалами.

Технологический разрыв нельзя сводить к сравнению разрешений. Полный цикл включает оценку состояния носителя, очистку, устранение деформаций и усадки, безопасную транспортировку пленки, настройку сканирования, обработку звука, реставрацию изображения, цветокоррекцию, контроль

[1*] Таблица представляет систематизацию проблем, сформулированных в предоставленных источниках (Рыженков).

качества и создание мастер-файла. При отсутствии одного из этапов даже высокое номинальное разрешение не гарантирует архивного качества.

Не менее важна инфраструктура хранения. 8K-мастер является крупным цифровым объектом, который требует надежного размещения, резервных копий, контроля целостности и планирования миграции данных. Поэтому решение о массовой 8K-оцифровке должно приниматься одновременно с проектированием цифрового хранилища. В противном случае государство получает дорогостоящие файлы без гарантии их долговременной сохранности.

Кадровый дефицит охватывает весь производственный цикл. Для работы необходимы инженеры по сканированию, специалисты по фотохимической подготовке, реставраторы изображения и звука, цветокорректоры, инженеры цифровых архивов, сотрудники контроля качества, специалисты по метаданным и эксперты по системам искусственного интеллекта. В предоставленных материалах подчеркивается, что в большинстве стран региона таких специалистов крайне мало.

Подготовка кадров должна быть непрерывной. Разовая демонстрация оборудования не заменяет обучение работе с разными типами пленки, поврежденными носителями, цветовыми пространствами, звуковыми дорожками, системами хранения и архивными метаданными. Необходимо формировать междисциплинарные команды, в которых архивист, инженер, реставратор, IT-специалист и специалист по правам работают в едином технологическом процессе.

Для Центральной Азии перспективной является модель национального или регионального центра компетенций. Такой центр может объединять дорогостоящее оборудование, контроль качества, методическую поддержку и обучение, обслуживая государственные архивы, кинофонды, телерадиокомпании, библиотеки, музеи и образовательные учреждения. При этом оригиналы и мастер-файлы остаются под контролем соответствующих государств и учреждений, а центр обеспечивает единый уровень технологии и документации.

Архитектура национальной программы массовой 8K-оцифровки

Переход к системной модели требует отдельной национальной программы в каждом государстве. Региональное сотрудничество может гармонизировать методику, однако ответственность за отбор фондов, правовой режим, финансирование и сохранность мастер-файлов должна быть закреплена на национальном уровне. Программа должна охватывать не только киноархив, но и все учреждения, где хранятся аудиовизуальные документы.

Первый элемент программы – единый государственный мандат и координационный механизм. Необходимо определить орган, который отвечает за общую политику, согласование стандартов, формирование приоритетов и мониторинг результатов. Без такого центра решения остаются распределенными между учреждениями, каждое из которых действует в пределах собственных задач и бюджета.

Второй элемент – национальная инвентаризация. До закупки оборудования и запуска массового сканирования требуется определить состав фондов, типы носителей, физическое состояние, объемы, наличие дублей, правовой статус и

уровень существующей цифровизации. Инвентаризация должна завершаться не только статистикой, но и ранжированием коллекций по срочности и значимости.

Третий элемент – единый технический стандарт. Он должен закреплять требования к разрешению архивного мастера, глубине цвета, цветовому пространству, форматам изображения и звука, структуре каталогов, контрольным суммам, резервным копиям, метаданным и производным версиям доступа. Стандарт должен разграничивать архивный мастер и пользовательскую копию: более компактная версия может использоваться для просмотра, но не должна заменять исходный файл сохранения.

Четвертый элемент – производственная инфраструктура. Государство может создать собственный национальный центр, распределенную сеть лабораторий или комбинированную модель с привлечением специализированных подрядчиков. Независимо от модели должны сохраняться единые требования к безопасности оригиналов, контролю качества, передаче файлов и документированию каждого этапа.

Пятый элемент – долговременное цифровое хранение. Массовая оцифровка должна быть связана с централизованной или согласованной инфраструктурой хранения, предусматривающей несколько копий, территориальное разделение, регулярную проверку целостности и плановую миграцию данных. Без этого программа создает цифровые копии, но не обеспечивает цифровое сохранение.

Шестой элемент – национальная программа подготовки специалистов. Она должна включать обучение операторов сканирования, реставраторов, цветокорректоров, специалистов по звуку, метаданным, IT-архитектуре, контролю качества и правам на контент. Важна передача знаний внутри учреждений, чтобы компетенции не зависели от одного сотрудника или поставщика оборудования.

Седьмой элемент – единая политика метаданных и интеллектуального доступа. Цифровые файлы необходимо связывать с архивными описаниями, технической информацией, правовым статусом и историей реставрационных вмешательств. После создания надежного массива данных можно внедрять автоматическое распознавание речи, объектов, лиц, локаций и тематических сцен, сохраняя контроль специалиста над результатами.

Восьмой элемент – модель использования и возврата общественной ценности. Программа должна заранее определять, какие производные версии создаются для исследователей, образования, телевидения, музеев, фестивалей, OTT-платформ и международного лицензирования. Такая модель позволяет показать, что бюджет направлен не только на хранение, но и на создание доступного культурного, научного и экономического ресурса.

Девятый элемент – долгосрочное финансирование и измеримые показатели. Массовая цифровизация не может быть завершена одним бюджетным циклом. Необходимы многолетние планы, в которых фиксируются объемы инвентаризации, подготовленных носителей, созданных 8K-мастеров, описанных объектов, подготовленных специалистов и опубликованных коллекций. Это позволяет оценивать не только количество файлов, но и качество сформированной системы.

***Таблица 3**. Основные компоненты национальной программы массовой 8K-оцифровки*[1*]

Компонент	Основное содержание	Ожидаемый результат
Управление	государственный мандат, координационный орган, распределение ответственности	единая политика и контроль реализации
Инвентаризация	учет фондов, состояния, прав и существующих цифровых копий	обоснованная приоритизация коллекций
Технический стандарт	параметры мастера, звук, метаданные, контроль качества и производные версии	совместимые и повторно используемые цифровые объекты
Инфраструктура	подготовка носителей, сканирование, реставрация, хранение и резервирование	завершенный производственный цикл
Кадры	обучение архивистов, инженеров, реставраторов, IT- и metadata-специалистов	устойчивая национальная компетенция
Доступ и использование	образование, наука, телевидение, музеи, фестивали, OTT и лицензирование	общественная и экономическая отдача
Мониторинг	многолетние показатели качества, объемов и использования	прозрачность и корректировка программы

Региональная координация и культурная дипломатия

Наличие отдельных национальных программ не исключает региональной координации. Напротив, согласованные подходы к терминологии, метаданным, контролю качества и подготовке специалистов позволят государствам Центральной Азии обмениваться опытом и реализовывать совместные проекты, не отказываясь от национального контроля над фондами. Региональная совместимость особенно важна для материалов, отражающих общую историю, трансграничные культурные процессы и взаимодействие народов.

Одним из практических направлений может стать общая методическая платформа, на которой публикуются рекомендации по приоритизации, шаблоны технической документации, требования к мастер-файлам и учебные материалы. Такая платформа не заменяет национальные стандарты, но облегчает их разработку и снижает риск того, что страны будут создавать взаимно несовместимые системы.

Региональный центр компетенций или сеть партнерских лабораторий может использоваться для обучения, сложной реставрации, независимого контроля качества и пилотных проектов. Для небольших учреждений это особенно важно, поскольку полный технологический цикл требует оборудования и специалистов, которые невозможно дублировать в каждом архиве. Разделение компетенций позволяет рациональнее использовать ресурсы при сохранении единых требований.

[1*] Компоненты сформулированы на основе практических рекомендаций и выводов предоставленных материалов (Рыженков).

Качественная реставрация национального кино и хроники создает основу для культурной дипломатии. Восстановленные материалы могут участвовать в фестивалях, международных ретроспективах, телевизионных проектах и цифровых выставках. Они способны расширить представление мировой аудитории о регионе, стимулировать новые документальные проекты и повысить интерес к национальным коллекциям.

Для международного продвижения недостаточно только высокого разрешения. Требуются профессиональная цветокоррекция, восстановление звука, субтитры, описание, юридическая ясность и версии, соответствующие требованиям кинотеатров, телевидения и OTT-платформ. Поэтому культурная дипломатия должна быть встроена в программу оцифровки на этапе планирования, а не рассматриваться как случайный результат уже выполненной технической работы.

В перспективе Центральная Азия может сформировать узнаваемое цифровое пространство аудиовизуального наследия, в котором национальные коллекции доступны для исследований, образования и международного сотрудничества. Такое пространство не требует объединения собственности или передачи оригиналов; оно требует совместимости описаний, технического качества и устойчивой институциональной связи между архивами.

Поэтапная дорожная карта для государств региона

Первый этап должен быть посвящен созданию основы. Он включает утверждение государственного мандата, инвентаризацию фондов, оценку физического состояния носителей, разработку национального технического стандарта и выбор приоритетной пилотной коллекции. Пилот необходим не как демонстрационный проект, а как проверка всей цепочки: от подготовки оригинала до хранения мастер-файла, метаданных и производной версии доступа.

На втором этапе формируется производственная способность. Государство определяет модель инфраструктуры, создает или выбирает центр компетенций, закупает оборудование либо формирует систему квалифицированного аутсорсинга, обучает специалистов и разворачивает долговременное цифровое хранилище. Массовое сканирование начинается только после того, как процедуры контроля качества и передачи данных проверены на пилотных материалах.

Третий этап связан с масштабированием и использованием. В программу включаются фонды разных учреждений, создается единый поисковый контур, внедряются расширенные метаданные и инструменты искусственного интеллекта. Параллельно готовятся коллекции для образования, телевидения, фестивалей, музеев и лицензирования, что позволяет обществу увидеть практический результат инвестиций.

Приоритетность должна определяться сочетанием нескольких критериев: физической уязвимости носителя, уникальности оригинала, культурной значимости, исторической срочности, востребованности и правовой готовности. Такой подход предотвращает ситуацию, когда массовая программа ориентируется только на наиболее известные фильмы, оставляя без внимания редкую хронику, этнографические записи и материалы национальных языков.

Ключевым показателем успеха является не количество приобретенных сканеров и не число созданных файлов само по себе. Успешная программа должна доказывать, что оригиналы обработаны безопасно, мастер-файлы соответствуют единому стандарту, данные надежно хранятся, коллекции описаны, сотрудники обучены, а производные версии реально используются исследователями, учреждениями культуры и обществом.

Таблица 4. *Предлагаемые этапы реализации национальной программы*[1*]

Этап	Ключевые действия	Контрольный результат
I. Подготовка	мандат, инвентаризация, приоритеты, стандарт, пилотная коллекция	проверенная технологическая и управленческая модель
II. Развертывание	центр компетенций, оборудование или аутсорсинг, обучение, хранилище	устойчивая производственная способность
III. Масштабирование	подключение фондодержателей, метаданные, ИИ, доступ и лицензирование	единый национальный цифровой фонд и общественное использование

Международный опыт и практические возможности реставрационной инфраструктуры

Предоставленные презентационные материалы показывают, что современные реставрационные центры работают не только как технические лаборатории, но и как комплексные сервисные платформы. В цепочку работ входят оценка состояния материалов, каталогизация, фотохимическая подготовка, оцифровка, восстановление изображения и звука, создание метаданных, контроль качества, визуальные эффекты, колоризация, подготовка вещательных и OTT-версий (Prasad).

Пример Prasad Corporation в предоставленных источниках используется как иллюстрация промышленного масштаба цифровой реставрации. Презентация указывает на основание компании в 1954 году, наличие более 1100 сотрудников, сохранение 25 000 фильмов, оцифровку 100 млн минут, обслуживание более 600 голливудских классических фильмов, более 20 лет опыта реставрации, 6200 часов восстановления изображения и более 400 реставраций для кинофестивалей.

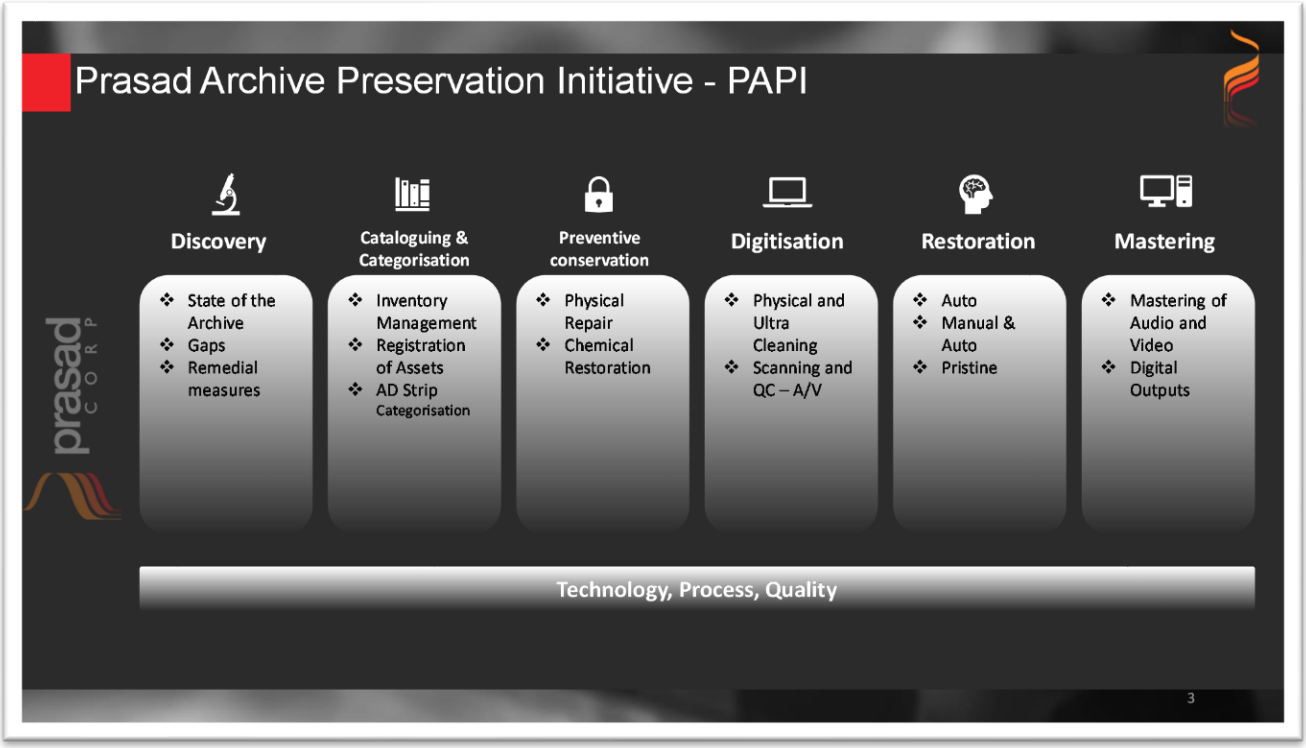
Для архивов Центральной Азии данный опыт важен не как готовая внешняя модель, которую можно механически перенести, а как ориентир для формирования собственной инфраструктуры. Национальным учреждениям необходимо развивать компетенции на всех этапах: от оценки физического состояния носителя до создания мастер-файла, от цифровой реставрации до публикации, лицензирования и интеллектуального поиска.

Практическая значимость международного опыта состоит также в понимании разделения ролей. Не каждое учреждение должно самостоятельно выполнять весь цикл работ. Возможны модели национального центра

[1*] Этапы являются аналитическим обобщением рекомендаций, содержащихся в предоставленных источниках.

компетенций, государственно-частного партнерства, аутсорсинга отдельных этапов, международной технической помощи и поэтапного формирования собственных лабораторий. Главное условие состоит в том, чтобы цифровизация была включена в долгосрочную стратегию, а не выполнялась как разовая закупка оборудования.

Рис. 4. *Комплексная цепочка работ: оценка, подготовка, оцифровка, реставрация, метаданные и контроль качества*



Организационно-экономическая модель внедрения 8K-оцифровки

Переход к 8K-оцифровке требует не только оборудования, но и управленческой модели. Первым условием является отделение архивного мастер-файла от производных версий. Мастер хранится как объект долговременного сохранения, а для практического доступа создаются более компактные копии. Такое разделение одновременно защищает качество источника и позволяет широко использовать материалы.

Вторым условием является понимание полной стоимости жизненного цикла. Планирование должно учитывать подготовку оригиналов, сканирование, реставрацию, контроль качества, метаданные, хранение, резервирование, обучение и последующую публикацию. Оценка только стоимости сканера или одного часа оцифровки не отражает реальной стоимости национальной программы.

Третьим условием является сочетание культурной миссии и модели использования. Государственные архивы не обязаны превращаться в коммерческие компании, однако они должны создавать условия, при которых качественно оцифрованный контент может законно и технически

использоваться телевидением, кинопроизводством, образованием, музеями, фестивалями и международными партнерами.

Четвертым условием является устойчивость. Программа должна продолжаться независимо от смены отдельных проектов и поставщиков. Для этого необходимы документированные процессы, открытая техническая документация, подготовленные сотрудники, регулярный аудит качества и бюджет на хранение данных. Только такая модель позволяет реализовать принцип «сканировать один раз – сохранять и использовать многократно».

Заключение

Анализ предоставленных источников показывает, что 8K-оцифровка аудиовизуальных архивов является не просто технологическим улучшением, а стратегией управления культурным наследием в условиях цифровой трансформации. Она позволяет одновременно снизить риски физической утраты оригинальных носителей, создать мастер-файл долговременного хранения, подготовить материалы к будущим форматам распространения и сформировать основу для экономического использования цифровых коллекций.

Концепция Scan Once меняет логику архивной работы. Вместо повторного сканирования одних и тех же материалов под очередной технологический стандарт она предлагает создать эталонный цифровой источник максимально возможного качества уже при первом обращении к оригиналу. Такой подход особенно важен для старых и поврежденных носителей, национального кинонаследия, исторической хроники, телевизионных архивов и редких этнографических материалов.

Для Центральной Азии основной проблемой является не отсутствие отдельных инициатив, а отсутствие целостной национальной системы в каждом государстве. Без государственного мандата, единого технического стандарта, долгосрочного финансирования, подготовленных специалистов, инфраструктуры хранения и общей политики метаданных отдельные проекты не складываются в массовую программу и не создают совместимый национальный цифровой фонд.

Одновременно регион обладает значительным нереализованным потенциалом. Его кинохроника, национальное кино, телевизионные коллекции, документальные и этнографические материалы могут стать ресурсом образования, науки, искусственного интеллекта, международного документального производства и культурной дипломатии. Для реализации этого потенциала качество оцифровки должно сочетаться с реставрацией, описанием, правовой подготовкой и доступом.

Каждому государству региона целесообразно сформировать собственную долгосрочную программу массовой 8K-оцифровки, а на региональном уровне согласовать методические подходы, обучение и совместимость метаданных. Такая модель позволит сохранить национальный контроль над наследием и одновременно создать условия для обмена опытом, совместных проектов и международного продвижения.

Искусственный интеллект дополнительно усиливает значение высококачественной оцифровки. Будущие системы автоматического описания, распознавания, поиска, реставрации и анализа будут требовать максимально

детализированных данных. Потери, возникшие на этапе низкоразрешенного сканирования, невозможно полностью восстановить позднее. Поэтому 8K следует рассматривать как подготовку архивов к новой эпохе интеллектуального доступа.

В этом смысле инвестиции в 8K-оцифровку являются не разовым расходом, а долгосрочным вложением в культурную память, цифровой суверенитет, научный потенциал, образование и международную узнаваемость государства. Формулу новой архивной политики можно выразить следующим образом: сканировать один раз, сохранять навсегда и использовать многократно.

AI-ENABLED ARCHIVE SOLUTION PRESENTED AT THE DIGITAL ARCHIVES EXPO, ASTANA, JUNE 2026

Kenneth Morris
BSc-Engg. Electrical, MS Management, MBA, CEO,
Chordify Inc, Menlo Park, United States
ken@chordify.com

Abstract. This report describes an AI-enabled Archiving solution as compared to the current solutions being used based on Optical Character Recognition or OCR. The proposed AI solution is based on the product Agent D. The article details the pros and cons of either technology and continues with a demo of the proposed solution. The Agent D platform solution is designed to turn document archives into Searchable Intelligence.
Keywords: Searchable Intelligence, AI-enabled, Vision AI, OCR, Semantic search, Entity Relationships

2026 Ж. МАУСЫМДА АСТАНАДА ӨТКЕН «DIGITAL ARCHIVES EXPO» КОНГРЕСІНДЕ ҰСЫНЫЛҒАН ЖАСАНДЫ ИНТЕЛЛЕКТКЕ НЕГІЗДЕЛГЕН АРХИВТІК ШЕШІМ

Кеннет Моррис
Менеджмент магистрі (MS Management), іскерлік басқару магистрі (MBA)
Chordify Inc. компаниясының бас атқарушы директоры (CEO),
Менло-Парк, АҚШ
ken@chordify.com

Аңдатпа. Баяндамада оптикалық таңбаларды тану (OCR) технологиясына негізделген қолданыстағы архивтік шешімдермен салыстырғанда, жасанды интеллект негізіндегі архивтеу шешімі сипатталады. Ұсынылып отырған шешім Agent D платформасына негізделген. Мақалада екі технологияның артықшылықтары мен

кемшіліктері талданып, ұсынылған шешімнің демонстрациясы ұсынылады. Agent D платформасы құжаттық архивтерді интеллектуалды іздеу мүмкіндігі бар ақпараттық жүйеге айналдыруға арналған.

Түйін сөздер: интеллектуалды іздеу, жасанды интеллект негізіндегі технологиялар, Vision AI, OCR, семантикалық іздеу, мәндер арасындағы байланыстар

АРХИВНОЕ РЕШЕНИЕ НА ОСНОВЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, ПРЕДСТАВЛЕННОЕ НА «DIGITAL ARCHIVES EXPO», АСТАНА, ИЮНЬ 2026 ГОДА

Кеннет Моррис
Магистр менеджмента (MS Management), магистр делового администрирования (MBA), генеральный директор Chordify Inc., Менло-Парк, США
ken@chordify.com

Аннотация. В докладе рассматривается решение для архивирования на основе технологий искусственного интеллекта и проводится его сравнительный анализ с существующими системами, основанными на технологии оптического распознавания символов (Optical Character Recognition, OCR). Предлагаемая система разработана на базе платформы Agent D. В работе подробно анализируются преимущества и ограничения обеих технологий, а также демонстрируются возможности платформы. Платформа Agent D предназначена для преобразования документных архивов в интеллектуальную поисковую среду, обеспечивающую семантический поиск и извлечение знаний из архивных документов.

Ключевые слова: интеллектуальный поиск, технологии на основе искусственного интеллекта, Vision AI, OCR, семантический поиск, взаимосвязи сущностей

Chordify Inc, a Silicon Valley-based technology solutions company, and its associates are presenting an AI-based Archive Solution for Searchable Intelligence.

Key user-facing features of the Agent D platform solution:

- Structured extraction of entities like organizations, people, dates, places.
- AI enabled summarization of documents for quick understanding without the need to peruse large documents.
- Traceability to the source document for easy verification.
- Multilingual semantic search across thousands of documents enabling access to global researchers.
- Gain insights using the interactive visual knowledge graph to identify relationships between entities.
- Archive administrators can manage the platform with Human-in-the-loop inputs to supplement AI.
    - For example, documents extracted and falling below a threshold can be flagged off to the Administrators for review and confirmation.
    - The AI then re-learns from these inputs, thus further enhancing the accuracy.

The 2-minute demo the highlights are as transcribed here below:

- Link to demo video: https://bit.ly/4viLdSU
- 0 to 10 seconds

- Once you log in, the dashboard gives you a summary view of the documents and Entities extracted.

• 10th second to 1 min 10 seconds

- You can proceed to check the “As is Documents”
- As you can see you can look at the original & the extracted version
- Also it shows the entities extracted such as “people” Organizations” etc.
- You can review the summary extract from AI or the full extracted document
- Also you may choose to flag it as “Verified”

• 1 min 12 seconds

- You can search for specific terms or phrases to see all relevant documents & dig deeper.

• 1 min 18 seconds to end

- You can also start by reviewing the entity maps, a very powerful tool
- For example you can check how an organization is related to other Organizations or people for example, gives you a visual bird's eye view.

Key technical features enabling the Agent D platform solution are:

- The platform leverages Vision AI rather than traditional OCR (optical character recognition).
- This approach leverages Deep learning and AI models rather than just pattern matching as in typical OCR, giving you much more context.
- This results in better ability to recognize a variety of distortions in document quality, handwritten text etc.
- Additionally, this approach results in a more structured Output facilitating easier classification, summarization & answering questions.
- Leverages knowledge graphs for entity mapping.
- The platform can be deployed in a secure single tenant architecture either using a Private Cloud or Secure Public Cloud, to ensure Data sovereignty & customization in compliance with Governance requirements.
- The platform demo is based on our initial discussions with the Archive of the Office of the President of Kazakhstan and can be customized further as per specific requirements.

Vision AI and OCR comparison

Vision AI vs OCR- some salient pointers: Vision AI and OCR both extract data from documents, but they differ significantly in how they handle real-world complexity. Understanding when to use each can have a major impact on accuracy, cost, and scalability.

• Vision AI delivers higher accuracy by understanding context, layout, and meaning, not just text.

• OCR works best for clean, consistent, high-volume documents with fixed formats.

On paper, both technologies promise the same outcome: turning documents into structured data. But in real workflows, the difference between these approaches becomes very clear, especially when formats vary, quality drops, or volume increases.

When comparing Vision AI and OCR, it helps to understand what each technology actually does. Both aim to extract data from documents, but they approach the problem very differently.

Traditional OCR (Optical Character Recognition) recognizes individual characters (A, B, C, 1, 2, 3), reads left-to-right and top-to-bottom, does not understand context or meaning, and often needs templates to know where fields are located. This is where OCR shows its limits. It can read text, but it does not understand what that text represents.

How OCR works:

1. Scan document and convert to pixels
2. Identify character shapes ("This looks like an A")
3. Convert shapes to text ("Invoice #12345")
4. Output raw, unstructured text

OCR is accurate for clean text but fragile when structure or layout changes.

Vision AI (Vision Language Models)

Vision AI has advanced features, and understands what it is reading, not just what the letters spell. It understands layout, structure, and meaning together, recognizes document types automatically (invoice, receipt, form), identifies relationships between elements, and adapts to format changes without constant retraining.

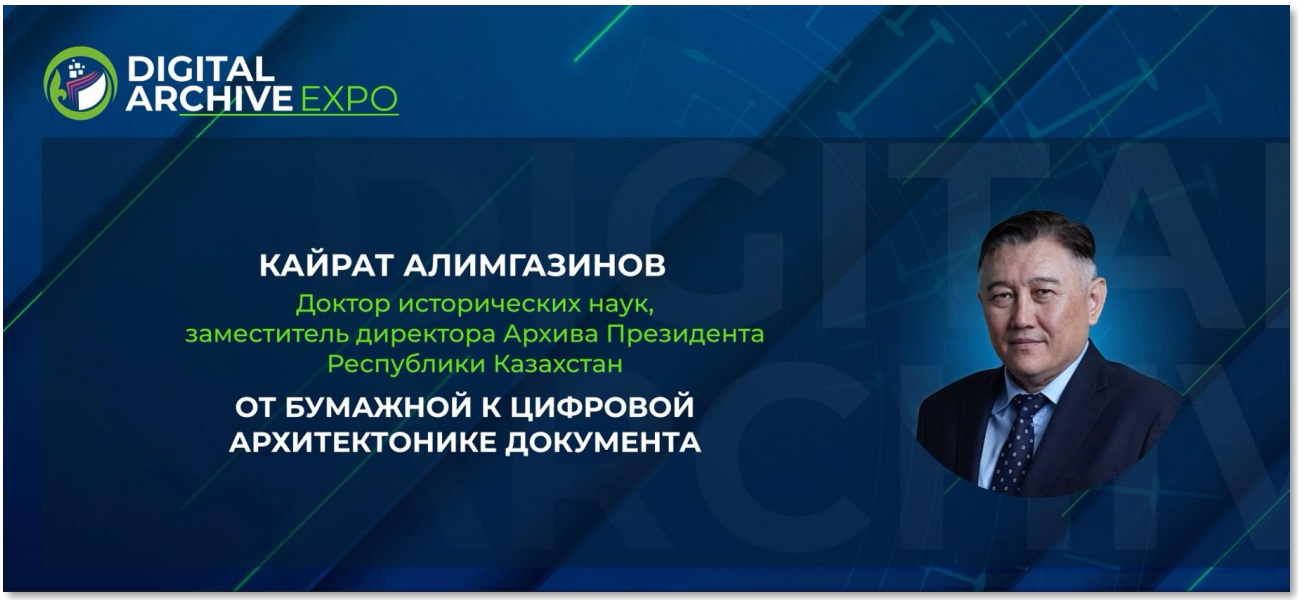
The key shift is this: Vision AI does not just read text. It interprets the document, using vision language models to process both text and visual structure simultaneously.

How Vision AI works:

1. Scan document and build a visual representation
2. Understand structure ("This is an invoice with a header, table, and totals section")
3. Extract with context ("Invoice #12345 is in the header, total is $1,234.56")
4. Output clean, structured, ready-to-use data

Implementation best practices:

To summarize, this demo gives a quick glimpse of the platform and can be customized as per your unique workflows and requirements. This can be deployed in a dedicated environment in the public cloud for quick start and migrated to your own sovereign cloud as needed. It is suggested to plan a rapid Initial Pilot Phase with minimal customization to get user feedback and then scale with further customization.



ОТ БУМАЖНОЙ К ЦИФРОВОЙ АРХИТЕКТОНИКЕ ДОКУМЕНТА

Кайрат Шакаримович Алимгазинов
Доктор исторических наук, заместитель директора
Архива Президента Республики Казахстан
alimgazinov@gmail.com

ҚАҒАЗ ҚҰЖАТТАН ҚҰЖАТТЫҢ ЦИФРЛЫҚ АРХИТЕКТОНИКАСЫНА

Қайрат Шәкәрімұлы Әлімғазинов
Тарих ғылымдарының докторы, Қазақстан Республикасы
Президенті Архиві директорының орынбасары
alimgazinov@gmail.com

FROM PAPER TO THE DIGITAL ARCHITECTURE OF THE DOCUMENT

Kairat Sh. Alimgazinov
Doctor of Historical Sciences, Deputy Director of the Archive
of the President of the Republic of Kazakhstan
alimgazinov@gmail.com

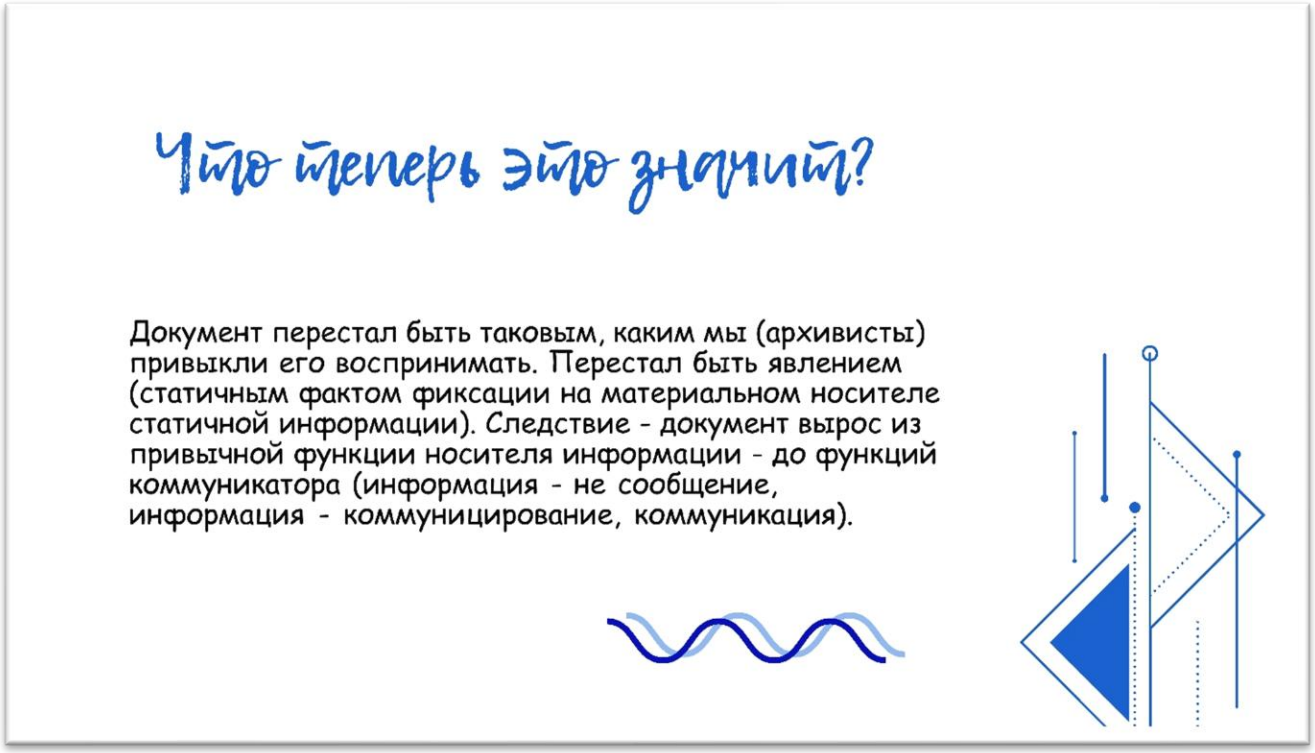
ОТ БУМАЖНОЙ К ЦИФРОВОЙ АРХИТЕКТОНИКЕ ДОКУМЕНТА

Алимгазинов Кайрат

Почему?

Потому что все чаще стали говорить не о документе как о явлении прошлой культуры и времени, а об информации, данных - как архивной материальной субстанции (читай - явлении культуры)

К чему это нас подвело?
Документ как материальный носитель прошлого и явление культуры перестал быть таковым. Таким образом, от структуры документа (который относится к организации и самому скелету текста, определяя его границы, иерархию и соотношение основных и второстепенных элементов) мы перешли к его архитектонике (соответствие текстовых блоков, семантических полей, элементов формы, что помогает закодировать смысл и принадлежность к определенному виду документа).
Архитектоника документа описывает общий концепт построения документа на разных уровнях организации его функций, определяет принципы взаимодействия и взаимозависимости компонентов документа (то есть реквизитов, которые включают способы отражения связей между отдельными элементами документа и порядок, в котором они следуют друг другу).


Что теперь это значит?
Документ перестал быть таковым, каким мы (архивисты) привыкли его воспринимать. Перестал быть явлением (статичным фактом фиксации на материальном носителе статичной информации). Следствие - документ вырос из привычной функции носителя информации - до функций коммуникатора (информация - не сообщение, информация - коммуницирование, коммуникация).

Суть?

Мы говорим о том, что на документ возложены новые функции, возникали новые обстоятельства его функционирования и его жизненного цикла

Какие?

Функции медиатора в информационном пространстве. Цифровом. С его нематериальной средой возникновения и функционирования. Цикличность (динамичность - как свойство изменчивости) и незамкнутость (потребность в верификации) жизненного цикла документа

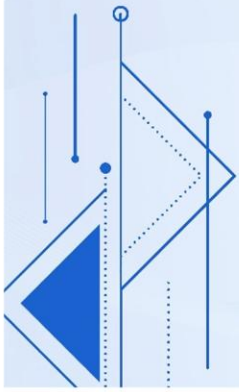
Что делать архивисту?

Менять отношение к документу! Рассматривать документ в единстве системы организации текста и его связей (Л. Дюранти), отойти от представления о документе с точки зрения организации в нем текста на основе статичных, неизменяемых во времени и пространстве критериев (реквизитов)

И тут мы говорим об архитектонике

...восприятии документа как последовательности текста с т.зр. организации в нем информационных связей (тематической, идейной и формальной)

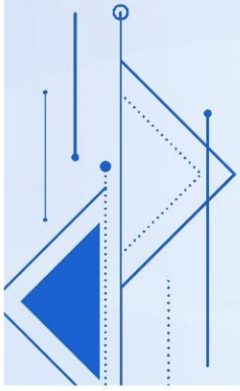
Что это значит в работе архивиста?



С т.зр. комплектования

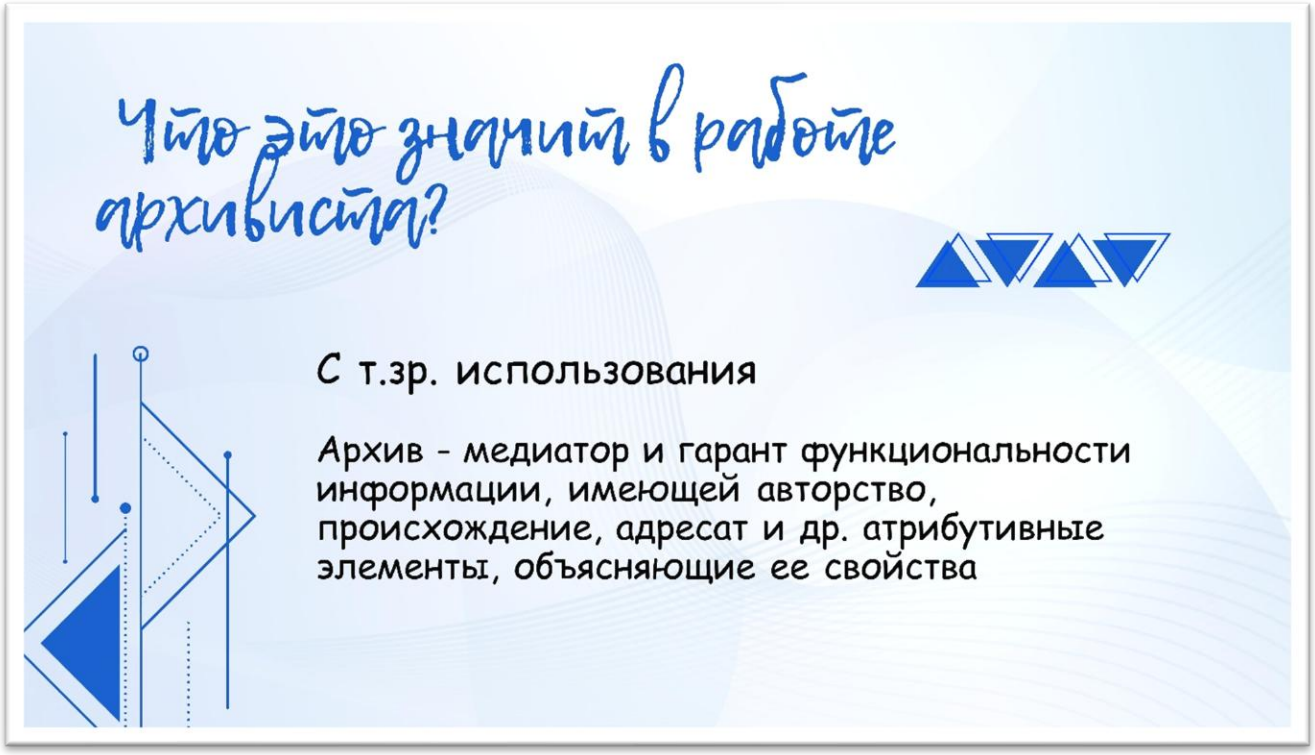
Расширить документ спектром категории «информация». Для документа было главнее - реквизитная часть (при решении вопроса комплектования). Для «информации» - главнее полнота функций, выполняемых ею. То есть следует отойти от принципа формирования архивного фонда по признаку происхождения (как основному) документа, к принципу формирования Архивного фонда по функционально-родовому критерию, характерному для анализа информации

Что это значит в работе архивиста?



С т.зр. сохранности

Наличие материального носителя - не панацея при принятии решения вопросов сохранности (т.ч. носители цифровой информации). Функциональность информации (документа - в традиционно-архивном представлении) обеспечивается наличием и постоянством верифицированных в нем информационных связей. То есть теперь в задачи сохранности должно входить как сохранение материального носителя информации, так и обеспечение постоянного хранения функциональных связей единицы информации, которые определяют выполнение ею своих функций


Что это значит в работе архивиста?
С т.зр. использования
Архив - медиатор и гарант функциональности информации, имеющей авторство, происхождение, адресат и др. атрибутивные элементы, объясняющие ее свойства

ФОТОГАЛЕРЕЯ

PHOTO GALLERY

Подписание Меморандума о взаимопонимании между Архивом Президента Республики Казахстан и Национальными архивами Великобритании Слева направо: директор по цифровым технологиям и информации Национальных архивов Великобритании Дж. Шеридан, директор Архива Президента Республики Казахстан А. Мустафина г. Астана, 8 июня 2026 г.

Президент Международного совета архивов Ж. Кирпс (слева) и директор Архива Президента Республики Казахстан А. Мустафина в экспозиционном зале Архива г. Астана, 8 июня 2026 г.

Встреча президента Международного совета архивов Ж. Кирпс
с руководителями архивных служб стран Центральной Азии
Слева направо: 1) заместитель директора Центрального государственного архива
Кыргызской Республики Н. Бейшеева; 2) советник Главного архивного управления
при Правительстве Республики Таджикистан М. Курбанова; 3) начальник
Главного архивного управления при Правительстве Республики Таджикистан Л. Шарифзода;
4) председатель Архивной службы при Министерстве цифрового развития и инновационных технологий
Кыргызской Республики Н. Базарбаева; 5) председатель Комитета архивов, документации
и книжного дела Министерства культуры и информации Республики Казахстан А. Курманова;
6) Президент Международного совета архивов Ж. Кирпс;
7) директор Архива Президента Республики Казахстан А. Мустафина;
8) директор Агентства «Узархив» У. Юсупов; 9) председатель Главного архивного управления
при Кабинете Министров Туркменистана Г. Ханбердиева;
10) начальник организационно-методического отдела Главного архивного управления
при Кабинете Министров Туркменистана М. Ахунова;
11) руководитель информационной службы Агентства «Узархив» А. Алиев
г. Астана, 8 июня 2026 г.

Подписание Меморандума о взаимопонимании между Архивом Президента Республики Казахстан и Агентством «Узархив» Республики Узбекистан
Слева направо: директор Агентства «Узархив» У. Юсупов,
директор Архива Президента Республики Казахстан А. Мустафина
г. Астана, 8 июня 2026 г.

Директор по цифровым технологиям и информации Национальных архивов Великобритании Дж. Шеридан (в центре) и руководитель отдела образования и информационно-просветительской работы в Национальном архиве Великобритании Э. Пейн (справа)
во время экскурсии по Архиву Президента Республики Казахстан
г. Астана, 8 июня 2026 г.

Открытие выставки «Digital Archive. EXPO»
Слева направо: президент Международного совета архивов Ж. Кирпс,
директор Архива Президента Республики Казахстан А. Мустафина,
заместитель председателя Комитета искусственного интеллекта
и развития инноваций МИИЦР РК С. Нуртазаев
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

На выставке «Digital Archive. EXPO»
Слева направо: директор Библиотеки и архивов ООН в Женеве Ш. Варакаулле,
президент Международного совета архивов Ж. Кирпс,
директор Архива Президента Республики Казахстан А. Мустафина
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

Интервью директора Архива Президента Республики Казахстан А. Мустафиной СМИ
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

Участники Конгресса на выставке «Digital Archive. EXPO»
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

Интервью президента Международного совета архивов Ж. Кирпс казахстанским СМИ
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

На торжественной церемонии открытия III Конгресса архивистов «Digital Archive. EXPO»
Слева направо: руководитель Отдела документального наследия сектора коммуникаций и информации ЮНЕСКО Ф. Банда, директор Архива Президента Республики Казахстан А. Мустафина, заведующий Общим отделом Администрации Президента Республики Казахстан Е. Адаев, директор Библиотеки и архивов ООН в Женеве Ш. Варакаулле, президент Международного совета архивов Ж. Кирпс
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

Церемония подписания Меморандума о сотрудничестве между Архивом Президента Республики Казахстан и Международным советом архивов
Слева направо: директор Архива Президента Республики Казахстан А. Мустафина, президент Международного совета архивов Ж. Кирпс
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

Участники III Конгресса архивистов «Digital Archive. EXPO»
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

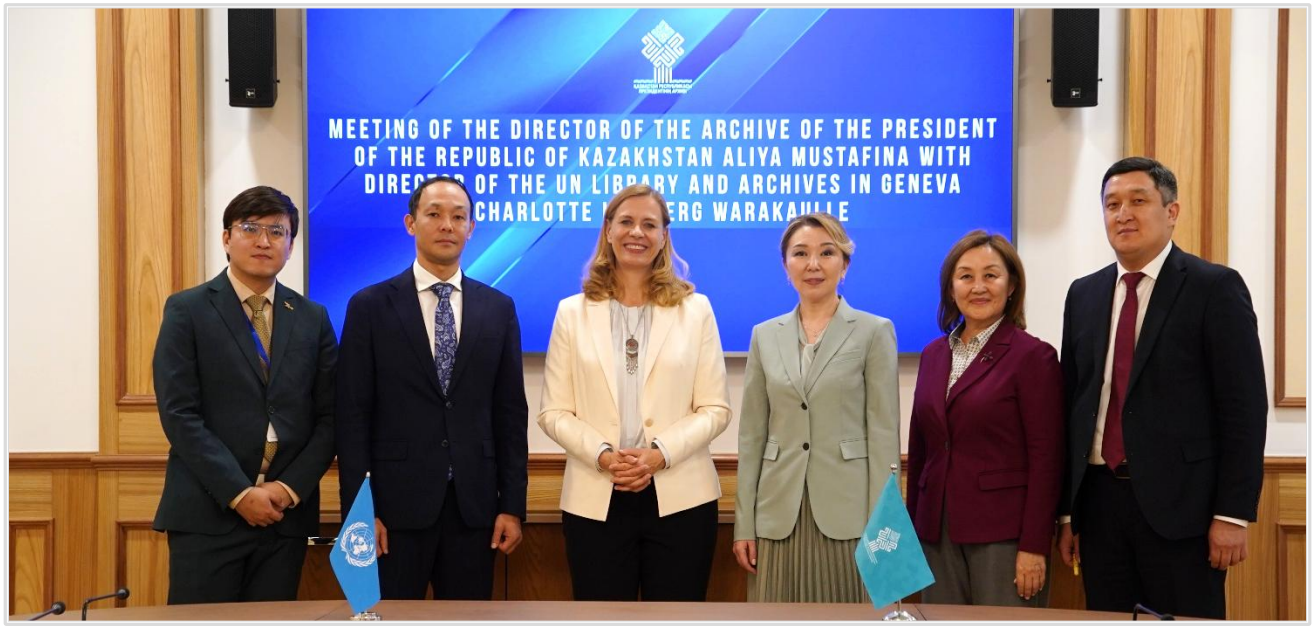
На пленарном заседании Конгресса
Слева направо: 1) руководитель Отдела документального наследия сектора коммуникаций и информации ЮНЕСКО Ф. Банда; 2) президент Международного совета архивов Ж. Кирпс, 3) директор Архива Президента Республики Казахстан А. Мустафина, 4) руководитель Библиотек и архивов ООН в Женеве Ш. Варакаулле, 5) заведующий Общим отделом Администрации Президента Республики Казахстан Е. Адаев, 6) депутат Сената Парламента Республики Казахстан Д. Кыдырали, 7) депутат Сената Парламента Республики Казахстан Р. Рустемов
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

Дружеское рукопожатие
Первый ряд слева направо: руководитель Отдела документального наследия сектора коммуникаций и информации ЮНЕСКО Ф. Банда, президент Международного совета архивов Ж. Кирпс
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

Участники 1-й панельной сессии «Архивы и AI: инструменты, этика и стратегии будущего»
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

В перерыве между сессиями
Слева направо: 1) руководитель Центра научных и прикладных исследований Архива Президента Республики Казахстан Н. Шаяхмет; 2) директор Архива Президента Республики Казахстан А. Мустафина; 3) исполнительный директор НИИ тюркологии и алтаистики при КазНУ им. аль-Фараби Н. Шаймерденова; 4) советник директора Архива Президента Республики Казахстан Р. Карибжанова; 5) ассистент-профессор Astana IT Unversity Д. Аманжолова; 6) ассистент-профессор Astana IT Unversity Ж. Маликова
г. Астана, Международный центр искусственного интеллекта ALEM.AI
9 июня 2026 г.

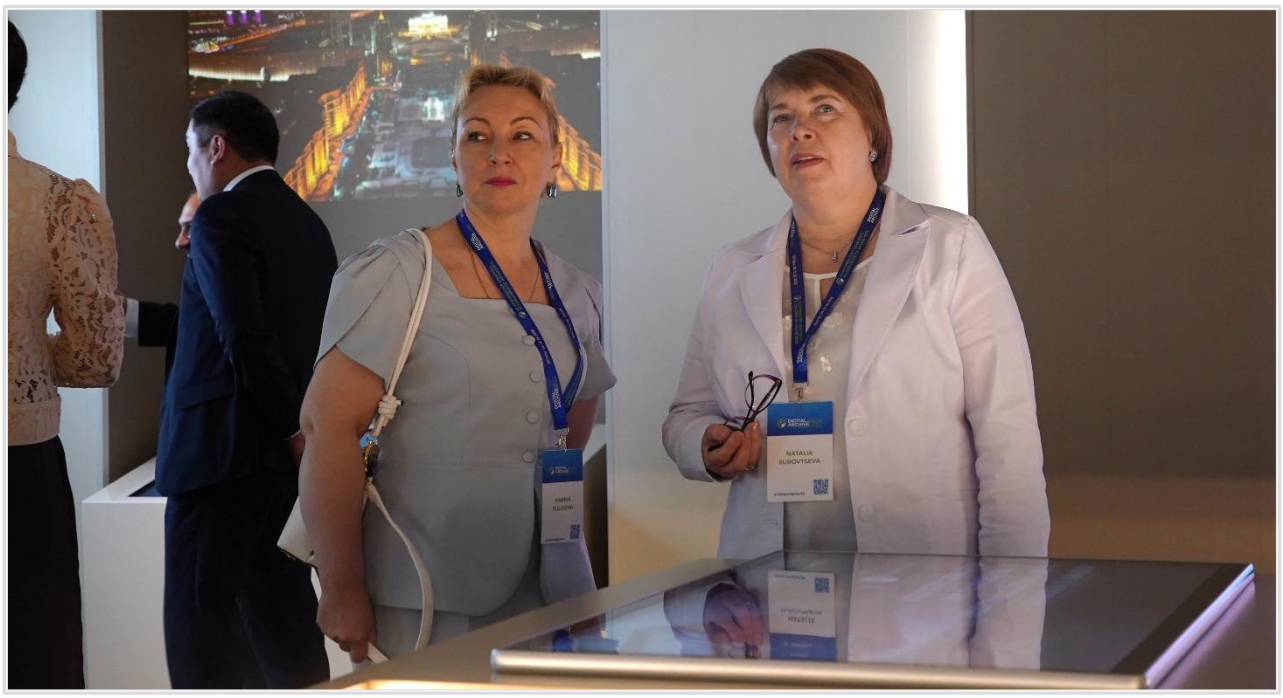
Спикеры, модераторы и участники 2-й панельной сессии
«Digital Archive: современные дискурсы»
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.



Встреча с руководством Архива Президента Республики Казахстан
Слева направо: 1) *старшие эксперты Центра образовательных проектов и международного сотрудничества А. Сагатов и 2) А. Кудайбергенов, 3) директор Библиотеки и архивов ООН в Женеве Ш. Варакаулле, 4) директор Архива Президента Республики Казахстан А. Мустафина, 5) советник директора Архива Президента Республики Казахстан Р. Карибжанова, 6) руководитель Центра образовательных проектов и международного сотрудничества М. Жылгелдинов*
г. Астана, 9 июня 2026 г.

Во время экскурсии по экспозиционному залу Архива Президента Республики Казахстан
Слева направо: директор Архива Президента Республики Казахстан А. Мустафина,
директор Библиотеки и архивов ООН в Женеве Ш. Варакаулле
г. Астана, 9 июня 2026 г.

В лабораторном корпусе Архива Президента Республики Казахстан
Слева направо: 1) руководитель Службы обеспечения сохранности и государственного учета
Г. Бекишева, 2) директор Архива Президента Республики Казахстан А. Мустафина,
3) старший эксперт Центра образовательных проектов и международного сотрудничества А. Сагатов,
4) директор Библиотеки и архивов ООН в Женеве Ш. Варакаулле
г. Астана, 9 июня 2026 г



Представитель Исполнительного комитета СНГ М. Кулагина (слева)
и профессор Российского государственного гуманитарного университета Н. Суровцева
в экспозиционном зале Архива Президента Республики Казахстан
г. Астана, 9 июня 2026 г.

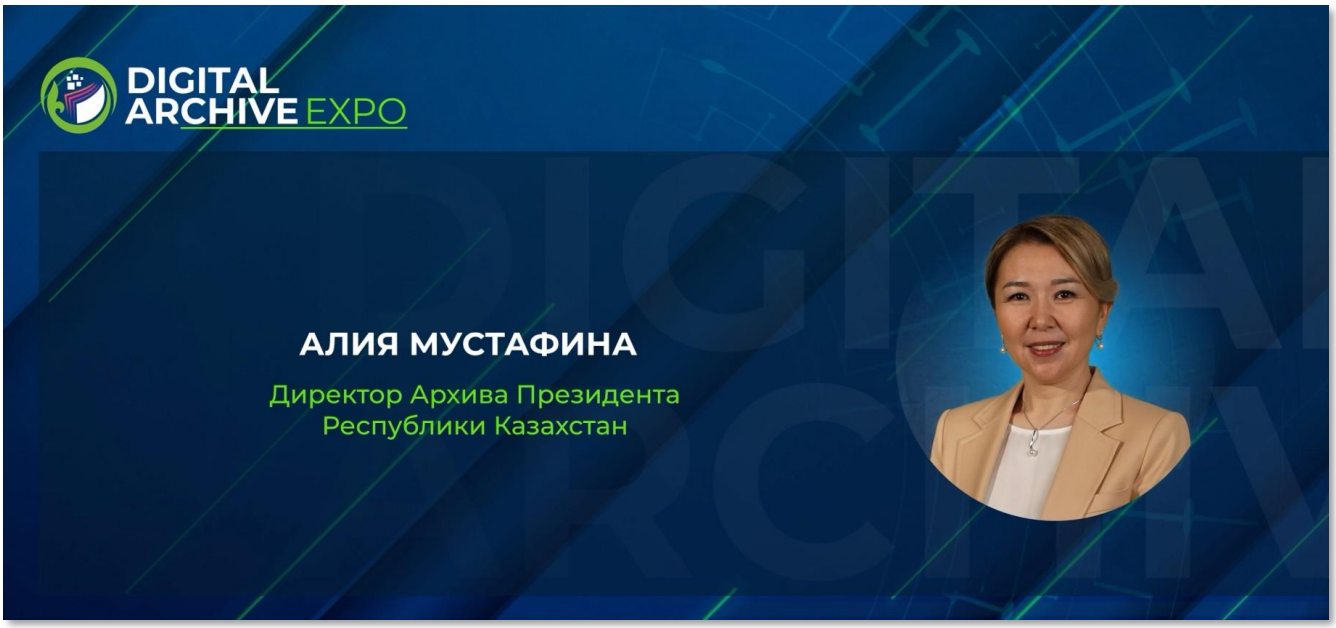
Участники Конгресса в читальном зале Архива Президента Республики Казахстан
г. Астана, 9 июня 2026 г.

Специалист по цифровым архивам Института Пилецкого (Польша) П. Лесиш в экспозиционном зале Архива Президента Республики Казахстан г. Астана, 9 июня 2026 г.

«ЖАСАНДЫ ИНТЕЛЛЕКТТІҢ
«ӘЛЕМ ЖАДЫ» ХАЛЫҚАРАЛЫҚ ТІЗІМІНЕ ӘСЕРІ»
UNESCO семинары

Семинар UNESCO
«ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ПОСЛЕДСТВИЯ
ДЛЯ МЕЖДУНАРОДНОГО РЕЕСТРА «ПАМЯТЬ МИРА»

«ARTIFICIAL INTELLIGENCE:
IMPLICATIONS FOR THE INTERNATIONAL MEMORY
OF THE WORLD REGISTER»
UNESCO Seminar



Құрметті әріптестер, ханымдар мен мырзалар!

Президенті Архиві мен UNESCO бірлесіп өткізіп отырған UNESCO-ның халықаралық семинарына қош келдіңіздер!

Естеріңізге сала кетейін, «Әлем жады» бағдарламасы бойынша алғашқы семинар 2024 жылы II Архившілер конгресі аясында өткізілген болатын. Семинар барысында Халықаралық тізілімге енгізу үшін жаңа номинациялар даярлау мәселелері бойынша мазмұнды пікір алмасу өтті. Қазақстанның бұл бағдарламаны іске асыруға белсенді қатысып отырғаны қуантады. Әріптестеріміздің үйлесімді жұмысының нәтижесінде бірқатар отандық құжаттық нысандар Халықаралық тізілімге енгізілді, бұл олардың тарихи маңыздылығының халықаралық деңгейде мойындалғанын білдіреді.

Бүгін біздің ортамызда құрметті қонақтарымыз:

- UNESCO-ның коммуникациялар және ақпарат секторының құжаттық мұра бөлімінің басшысы д-р Факсон Банда;

- Сингапур Ұлттық архивтері директорының аға орынбасары және «Әлем жады» бағдарламасының Халықаралық консультативтік комитеті жанындағы Сақтауды қамтамасыз ету жөніндегі кіші комитеттің төрағасы д-р Фанг Лай Ти;

- сондай-ақ Дублин университеттік колледжінің жасанды интеллект және цифрлық мәдени мұра кафедрасының доценті д-р Сан Парк.

Сіздерге үлкен ризашылығымызды білдіреміз.

Құрметті әріптестер!

Бүгінде цифрлық материалдар дәстүрлі қолжазбалар мен мәтіндік құжаттармен қатар архивтік сақтау және зерттеу нысандары ретінде тең дәрежеде орын алуда. Алайда жасанды интеллект технологияларының қарқынды дамуы құжаттық мұраны сақтау мәселелеріне қатысты қалыптасқан тәсілдерді өзгертіп отыр. Болашақта архивтік мұра тек цифрлық контентті ғана емес, жасанды интеллект көмегімен жасалған материалдарды да қамтитыны анық.

Сондықтан біз жасанды интеллект дәуіріндегі архивтердің рөлі қандай болмақ деген сұраққа жиі ораламыз. Бұл рөл тарихи жадыны сақтаумен тікелей байланысты бола ма, әлде цифрлық контентті алгоритмдік бақылау мәселелерімен ұштаса ма?

Қазіргі заманғы жасанды интеллект жүйелері ашық қолжетімді орасан зор деректер массивтері негізінде оқытылады. Алайда өңірлік тарихи мұраның едәуір бөлігі әлі күнге дейін цифрлық ортада қолжетімсіз күйінде қалып отыр.

Соның салдарынан ақпараттық вакуум пайда болып, оны біртіндеп жасанды түрде генерацияланған контент, қайталама интерпретациялар және кәсіби архивтік әрі ғылыми ортадан тыс жасалған материалдар толтыруда.

Тарихты цифрлық қабылдаудың бұрмалануының айқын мысалы ретінде Кеңес Одағының Батыры Әлия Молдағұлованың интернетте кең таралған «танымал» портретін атауға болады. Ол көбіне шынайы фотосурет ретінде қабылданады. Шын мәнінде, бұл – «Мергендер» («Снайперы», 1985) фильмінде Әлияның рөлін сомдаған актриса Айтуған Темірованың кадры. Ал Әлия Молдағұлованың түпнұсқа архивтік фотосуреттері бар болғанымен, олар көпшілікке әлдеқайда аз таныс. Бұл жағдай цифрлық орта тарихи дереккөздің орнын неғұрлым «ыңғайлы» әрі эмоциялық тұрғыдан әсерлі бейнемен алмастыра алатынын айқын көрсетеді.

Бүгінде пайдаланушылар дереккөздерді қосымша тексерместен, жасанды интеллекттің сенімді жауаптарына көбірек иек арта бастады. Соның нәтижесінде алгоритмдер қайталаған қателіктер мен дәлсіздіктер жылдам таралып, шынайы факт ретінде қабылдануы мүмкін.

Бүгін бізде UNESCO-ның «Әлем жады» бағдарламасының Халықаралық консультативтік комитеті жанындағы Сақтауды қамтамасыз ету жөніндегі кіші комитеттің жасанды интеллект ықпалын зерттеу жөніндегі жұмыс тобының жетекші сарапшылары және Қазақстан Республикасының UNESCO және ICESCO істері жөніндегі Ұлттық комиссиясының, «Әлем жады» бағдарламасының Ұлттық комитетінің өкілдерімен осы ықпалдың салдарын талқылауға мүмкіндік бар екеніне қуаныштымыз.

Әріптестеріміз өз жұмыстарының нәтижелерін ұсынып, біздің тәжірибемізбен, пікірлерімізбен танысатын болады. Біздің пікірталасымыз мазмұнды әрі нәтижелі болады деп сенемін.

А.Қ. Мұстафина,
Қазақстан Республикасы Президенті Архивінің директоры,
тарих ғылымдарының кандидаты

Құрметті ханымдар мен мырзалар, құрметті қатысушылар!

Қазақстан Республикасы ЮНЕСКО және ИСЕСКО істері жөніндегі Ұлттық комиссиясының атынан *«Жасанды интеллект: "Әлем жады" халықаралық тізіліміне әсері»* тақырыбындағы ЮНЕСКО семинарының ашылуымен шын жүректен құттықтап, зор ризашылығымызды білдіруге рұқсат етіңіздер.

Distinguished guests, Dear colleagues and friends,

On behalf of the National Commission of the Republic of Kazakhstan for UNESCO and ICESCO, it is a great pleasure to welcome you to this seminar dedicated to exploring the implications of Artificial Intelligence for the International Memory of the World Register.

At the outset, allow me to express my sincere appreciation to Professor Fackson Banda, Chief of the UNESCO Documentary Heritage Unit, for the visionary leadership and unwavering commitment to advancing the objectives of the Memory of the World Programme. Your efforts in fostering international dialogue on the future of documentary heritage in an era of rapid technological transformation are both timely and invaluable.

I would also like to extend my heartfelt gratitude to Ms. Aliya Mustafina, Director of the Presidential Archive of the Republic of Kazakhstan, for organizing this important seminar and creating a valuable platform for dialogue among experts, archivists, and researchers.

For Kazakhstan, documentary heritage represents a vital bridge between past, present and future generations.

The importance of preserving historical memory was reaffirmed by President Kassym-Jomart Tokayev during the Fourth Meeting of the National Qurultai in Burabai. In his address, he emphasized that our historical, cultural and intellectual heritage constitute the foundation of national identity and an invaluable legacy for future generations.

This commitment is reflected in Kazakhstan's active engagement with the UNESCO Memory of the World Programme. Most notably, for the first time in fourteen years, Kazakhstan celebrated a landmark achievement with the inscription of its manuscript "Khandar shejiresi" ("Genealogy of the Khans") on the UNESCO Memory of the World Register, representing a significant milestone in the international recognition of our documentary heritage.

Building on this achievement, our National Committee of the Memory of the World Programme has submitted a number of documentary heritage nominations to the Memory of the World Regional and International Registers, including "the Works of Abai", "the first Kazakh opera 'Kyz-Zhibek' by Yevgeny Brusilovsky", "the Collection of Mashkhur Jusup Kopeev's Printed Editions", and "the Issyk Inscription". These documentary treasures reflect the richness of Kazakhstan's cultural heritage and contribute to the shared documentary memory of humanity.

Today's discussion takes place at a pivotal moment. Artificial intelligence is reshaping the ways in which knowledge is created, preserved, accessed and transmitted.

As custodians of humanity's documentary legacy, we share a collective responsibility to ensure that technological innovation strengthens, rather than compromises the integrity, credibility and authenticity of documentary heritage. UNESCO's leadership in promoting an ethical, inclusive and human-centred approach to artificial intelligence is particularly valuable.

In this regard, allow me to extend our sincere appreciation to the UNESCO Regional Office in Almaty for its critical support in advancing the UNESCO Readiness Assessment Methodology (RAM) process in Kazakhstan and for its continued efforts in ensuring that technological progress remains aligned with the principles of cultural and documentary heritage preservation.

Kazakhstan remains firmly committed to working closely with UNESCO and other international partners to unlock the transformative potential of AI for the sake of safeguarding our shared documentary heritage, developing new avenues for research and innovation.

I wish you all productive and fruitful discussions.

Thank you for your attention.

Zhanar Shaimenova,
Secretary-General of the National Commission
of the Republic of Kazakhstan for UNESCO and ICESCO

ИСТОРИКО-ЛИНГВИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ ПИСЬМЕННЫХ ИСТОЧНИКОВ НА ОСНОВЕ АРХИВНЫХ КАРТОТЕК В ЦИФРОВУЮ ЭПОХУ

Нурсулу Жамалбековна Шаймерденова
Доктор филологических наук, профессор, исполнительный директор Научно-исследовательского института тюркологии и алтаистики, Казахский национальный университет им. аль-Фараби, Казахстан
shaimerdenova.altair@gmail.com; https://orcid.org/0000-0002-2830-8336

Аннотация. Статья посвящена комплексному анализу письменных источников в рамках историко-лингвистического подхода на материале архивных картотек в условиях цифровой эпохи. Цель исследования заключается в выявлении научного потенциала архивных картотек как систематизированных массивов лексикографических и грамматических данных, а также в определении возможностей их интеграции в цифровую среду. Рассматриваются особенности традиционных методов работы с картотеками, включая ручную каталогизацию и тематическую классификацию, а также современные подходы, основанные на цифровизации, разметке и создании электронных баз данных. Особое внимание уделяется картотекам словарей как уникальным источникам, отражающим динамику языковых процессов и позволяющим реконструировать историческую картину развития языка. Подчеркивается значение цифровых технологий для сохранения, систематизации и анализа архивных материалов. Делается вывод о том, что цифровизация картотек способствует расширению исследовательских возможностей, обеспечивая доступность данных и формируя основу для междисциплинарных исследований.

Ключевые слова: историко-лингвистическое исследование; письменные источники; архивные картотеки; цифровизация; языкознание; лексикография; корпусные технологии

ЦИФРЛЫҚ ДӘУІРДЕ АРХИВТІК КАРТОТЕКАЛАР НЕГІЗІНДЕ ЖАЗБА ДЕРЕККӨЗДЕРДІ ТАРИХИ-ЛИНГВИСТИКАЛЫҚ ЗЕРТТЕУ

Нұрсұлу Жамалбекқызы Шаймерденова
Филология ғылымдарының докторы, профессор, Түркология және алтаистика ғылыми-зерттеу институтының атқарушы директоры,
әл-Фараби ат. Қазақ ұлттық университеті, Қазақстан
shaimerdenova.altair@gmail.com; https://orcid.org/0000-0002-2830-8336

Аңдатпа. Мақала цифрлық дәуір жағдайында архивтік картотекалар материалдары негізінде тарихи-лингвистикалық тәсіл аясында жазба дереккөздерді кешенді талдауға арналған. Зерттеудің мақсаты – архивтік картотекалардың лексикографиялық және грамматикалық деректердің жүйелендірілген қоры ретіндегі ғылыми әлеуетін айқындау, сондай-ақ оларды цифрлық ортаға ықпалдастыру мүмкіндіктерін анықтау. Мақалада картотекалармен жұмыс істеудің дәстүрлі әдістері, соның ішінде қолмен каталогтау мен тақырыптық жіктеу, сондай-ақ цифрландыруға, таңбалауға және электрондық деректер базаларын құруға негізделген заманауи тәсілдер қарастырылады. Сөздік картотекаларына тілдік үдерістердің динамикасын бейнелейтін және тілдің тарихи даму көрінісін қайта қалпына келтіруге мүмкіндік беретін бірегей дереккөздер ретінде ерекше назар аударылады. Архивтік материалдарды сақтау, жүйелеу және талдау ісіндегі цифрлық технологиялардың маңызы айқындалады. Архивтік картотекаларды цифрландыру деректердің қолжетімділігін қамтамасыз етіп, зерттеу мүмкіндіктерін кеңейтетіні әрі пәнаралық зерттеулердің дамуына негіз қалайтыны туралы қорытынды жасалады.

Түйін сөздер: тарихи-лингвистикалық зерттеу; жазба дереккөздер; архивтік картотекалар; цифрландыру; тіл білімі; лексикография; корпустық технологиялар

HISTORICAL-LINGUISTIC STUDY OF WRITTEN SOURCES BASED ON ARCHIVAL CARD FILES IN THE DIGITAL AGE

Nursulu Zh. Shaimerdenova
Doctor of Philology, Professor, Executive Director
of the Scientific Research Institute of Turkology and Altaic Studies,
Al-Farabi Kazakh National University, Kazakhstan
shaimerdenova.altair@gmail.com; https://orcid.org/0000-0002-2830-8336

Abstract. The article presents a comprehensive analysis of written sources within the framework of a historical-linguistic approach based on archival card files in the digital age. The aim of the study is to identify the scientific potential of archival card systems as structured collections of lexicographic and grammatical data and to determine the possibilities of their integration into the digital environment. The paper examines traditional methods of working with card files, including manual cataloguing and thematic classification, as well as modern approaches based on digitization, annotation, and database creation. Special attention is given to dictionary card files as unique sources reflecting the dynamics of language development and enabling the reconstruction of historical linguistic processes. The importance of digital technologies for preserving, systematizing, and analyzing archival materials is emphasized. It is concluded that the digitization of card files expands research opportunities, ensures data accessibility, and provides a basis for interdisciplinary studies.

Keywords: historical linguistics; written sources; archival card files; digitization; linguistics; lexicography; corpus technologies

Введение

Историко-лингвистическое исследование письменных источников является одним из ключевых направлений гуманитарного знания. Письменные памятники фиксируют языковые процессы, отражая эволюцию лексики, грамматики и семантики. Работа в этом направлении ведется лингвистами и при создании различных корпусов, лексикографических источников как Казахстане, так и за рубежом (Atkins 2008; McEnery 2012). Особое значение в этом контексте приобретают архивные картотеки, представляющие собой систематизированные собрания языкового материала, накопленного в ходе многолетних исследований.

В условиях цифровой эпохи традиционные формы хранения и анализа данных подвергаются трансформации. Архивные картотеки, ранее существовавшие исключительно в бумажном виде, становятся объектом цифровизации, что открывает новые возможности для их использования в научной практике. Цифровизация требует хорошего знания методологии и методики работы, на что указывают работы исследователей, поэтому столь важны разрабатываемые учеными и педагогами методические рекомендации (Алпысбаева 2020: 68). В качестве примера можно привести картотеку «Словаря русского язык XI–XVII вв.», которая к настоящему времени полностью оцифрована и активно используется исследователями и читателями.

Архивные картотеки как источник исследования

Архивные картотеки представляют собой особый тип научно организованных источников, в которых концентрируется результат длительной лексикографической и текстологической (Лихачев 1997: 617).

В отличие от опубликованных словарей и корпусов текстов, картотека фиксирует не только итоговую интерпретацию языкового материала, но и сам процесс его научного осмысления: от первичной фиксации языкового факта до его классификации и семантической интерпретации. В этом смысле картотека выступает как «динамическая модель» языка, отражающая различные стадии анализа и позволяющая реконструировать эволюцию научных представлений о языковой системе.

Структурно картотека состоит из совокупности карточек, каждая из которых содержит комплексную информацию: лексему (или ее вариант), грамматические характеристики, контекст употребления, источник фиксации, а также возможные этимологические и семантические комментарии. Существенным является то, что карточка фиксирует слово не изолированно, а в живом контексте, что обеспечивает возможность изучения его функционально-семантических свойств, прагматических особенностей и стилистической маркированности. Таким образом, картотека сохраняет не только язык как систему, но и язык как процесс функционирования.

Научная значимость архивных картотек определяется рядом фундаментальных характеристик. Во-первых, это репрезентативность и полнота охвата языкового материала, достигаемая за счет включения данных из

различных письменных источников, относящихся к разным историческим периодам и жанровым традициям. Во-вторых, высокая степень точности фиксации контекстов обеспечивает достоверность последующего анализа и позволяет избегать редукции значения слова к абстрактной словарной дефиниции. В-третьих, наличие исследовательских комментариев придает картотеке характер интерпретированного источника, где отражается не только языковой факт, но и его научное осмысление в рамках определенной теоретической парадигмы.

Особую ценность картотеки приобретают в историко-лингвистических исследованиях, поскольку позволяют проследить динамику языковых изменений в длительной диахронической перспективе. Например, мое исследование, касающееся глосс в памятниках письменности, во многом построено на материалах картотеки и архивах, в которых отражаются тюркско-славянские языковые контакты (Шаймерденова 2018: 156).

На основе картотечных данных возможно выявление архаических элементов, установление этапов семантической эволюции слов, реконструкция утраченных или трансформированных значений. Кроме того, картотеки дают возможность фиксировать переходные явления, которые часто не находят отражения в нормативных словарях, но имеют принципиальное значение для понимания механизмов языкового развития.

В контексте тюркологических исследований архивные картотеки играют особенно важную роль, поскольку значительная часть языкового материала представлена разнородными письменными памятниками, диалектными записями и полевыми материалами. Картотека в данном случае выступает как интегративная система, объединяющая разнотипные данные и обеспечивающая их сопоставимость. Это позволяет не только реконструировать историю

отдельных языков, но и выявлять общие закономерности развития тюркских языков в сравнительно-историческом аспекте.

Таким образом, архивные картотеки следует рассматривать не просто как вспомогательный инструмент, а как самостоятельный источник научного знания, обладающий высокой степенью информативности и аналитического потенциала. Их использование позволяет существенно углубить историко-лингвистический анализ, обеспечивая переход от статического описания языковых фактов к их системному и динамическому осмыслению.

Междисциплинарный аспект

Современные исследования архивных картотек и письменных источников формируются в пространстве междисциплинарного взаимодействия, объединяющего методы и теоретические подходы лингвистики, истории, культурологии и информационных технологий. Такая интеграция обусловлена самой природой изучаемого материала: язык, зафиксированный в архивных источниках, выступает не только как система знаков, но и как носитель исторической памяти, отражающий социальные, культурные и когнитивные процессы.

В этом контексте использование цифровых инструментов существенно расширяет аналитические возможности исследователя. Применение методов обработки данных, корпусного анализа и семантического моделирования позволяет переходить от описания отдельных языковых фактов к выявлению устойчивых закономерностей, коррелирующих с историческими изменениями. Таким образом, языковой материал становится основой для реконструкции более широких культурно-исторических процессов, включая трансформации общественных структур, изменение коммуникативных практик и формирование ценностных систем.

Особую роль играет взаимодействие лингвистики с информационными технологиями, в рамках которого формируются новые исследовательские инструменты и подходы. Разработка специализированных баз данных, систем разметки и алгоритмов поиска обеспечивает возможность многомерного анализа архивных материалов. При этом исторический контекст, реконструируемый на основе текстов и картотек, приобретает дополнительную глубину за счет привлечения данных из смежных дисциплин.

Создание интегрированных цифровых платформ становится важнейшим направлением развития современной науки. Такие платформы обеспечивают объединение разнородных источников – текстовых корпусов, архивных картотек, рукописных материалов – в единую информационную среду. Это способствует не только сохранению и систематизации данных, но и формированию принципиально нового уровня их интерпретации.

Важным следствием цифровой интеграции является активизация международного научного сотрудничества. Доступ к унифицированным ресурсам и совместное использование данных создают условия для реализации коллективных исследовательских проектов, сопоставительных исследований и выработки общих методологических стандартов. В результате формируется глобальное научное пространство, в котором архивные материалы становятся

частью единого интеллектуального ресурса, доступного для комплексного и многоаспектного изучения.

Заключение

Архивные картотеки по праву занимают особое место в системе источников историко-лингвистического исследования, поскольку в них аккумулируется многолетний опыт фиксации и интерпретации языковых фактов. Их значение определяется не только объемом и разнообразием представленного материала, но и глубиной научной проработки, отражающей этапы становления лингвистической мысли.

Перевод картотечных массивов в цифровой формат существенно меняет исследовательскую оптику: данные, ранее доступные лишь в локальном и фрагментарном виде, становятся частью структурированного информационного пространства, пригодного для комплексного анализа. Это позволяет по-новому осмыслить динамику языковых процессов, выявлять скрытые закономерности и сопоставлять разнородные источники на принципиально ином уровне.

Сопряжение традиционных филологических методов с цифровыми инструментами формирует современную исследовательскую модель, в которой сохраняется преемственность научной школы и одновременно расширяется аналитический потенциал. В этих условиях архивные картотеки перестают быть исключительно вспомогательным ресурсом и приобретают статус ключевого элемента научной инфраструктуры.

Таким образом, их цифровая трансформация открывает перспективы не только для углубления историко-лингвистических исследований, но и для формирования единого научного пространства, обеспечивающего системное сохранение и интерпретацию языкового наследия.

Литература

Atkins 2008 – Atkins B.T.S., Rundell M. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press, 2008. 540 p.

McEnery 2012 – McEnery T., Hardie A. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press, 2012. 294 p.

Алпысбаева 2020 – Алпысбаева Н., Шайлазымов Б. Оцифровка архивных документов: методические рекомендации. Алматы: Қазақ университеті, 2020. 68 с.

Лихачев 1997 – Лихачев Д.С. Текстология: на материале русской литературы X–XVII вв. Санкт-Петербург: Наука, 1997. 617 с.

Шаймерденова, 2018 – Шаймерденова Н.Ж. Глоссы как феномен текста. Алматы: Флинта, 2018. 156 с.

ARTIFICIAL INTELLIGENCE IMPLICATIONS FOR UNESCO MEMORY OF THE WORLD REGISTER -VALORISATION OF AI-GENERATED WORKS AS DOCUMENTARY HERITAGE

Sun Park
University College Dublin, Ad Astra Fellow/Assistant Professor in AI and Digital Cultural Heritage School of Information and Communication Studies, Ireland
sun.park@ucd.ie; https://orcid.org/0000-0001-9307-1171

Abstract. The rapid uptake of AI across memory institutions – in the creation, description, and preservation of records – leads us to revisit the conceptual and procedural foundations of UNESCO Memory of the World (MoW) Programme. As part of the discussion of the UNESCO MoW AI Working Group, the idea of whether AI-generated works can be considered as UNESCO Documentary Heritage has been discussed. This presentation briefly shares how the values of AI-generated works can be associated with the MoW Programme, centring on the definition of a document, the nomination criteria and the Register nomination form. It concludes with reflecting on the shift from a memory ecology of retrieval to generation emerging from human-AI collective memory-making.

Keywords: artificial intelligence; UNESCO; Memory of the World; documentary heritage; AI-generated works

ЖАСАНДЫ ИНТЕЛЛЕКТТІҢ ЮНЕСКО-НЫҢ «ӘЛЕМ ЖАДЫ» ТІЗІЛІМІНЕ ЫҚПАЛЫ: ЖАСАНДЫ ИНТЕЛЛЕКТ ЖАСАҒАН ТУЫНДЫЛАРДЫ ҚҰЖАТТЫҚ МҰРА РЕТІНДЕ ҚҰНДЫЛАНДЫРУ

Сан Парк
Ad Astra бағдарламасының стипендиаты, Ақпараттық және коммуникациялық зерттеулер мектебінің жасанды интеллект және цифрлық мәдени мұра саласындағы ассистент-профессоры, Дублин университеттік колледжі, Ирландия
sun.park@ucd.ie; https://orcid.org/0000-0001-9307-1171

Аннотация. Жад институттарында – құжаттарды жасау, сипаттау және сақтау үдерістерінде – жасанды интеллектті (AI) кеңінен қолданудың қарқынды дамуы ЮНЕСКО-ның «Әлем жады» (Memory of the World, MoW) бағдарламасының тұжырымдамалық және рәсімдік негіздерін қайта қарастыру қажеттігін туындатады. ЮНЕСКО-ның MoW AI жұмыс тобы аясындағы талқылаулар барысында жасанды интеллект арқылы жасалған туындыларды ЮНЕСКО-ның құжаттық мұрасы ретінде қарастыру мүмкіндігі туралы мәселе көтерілді. Бұл баяндамада AI көмегімен жасалған жұмыстардың құндылықтарының MoW бағдарламасымен байланысы құжат ұғымының анықтамасы, номинациялау критерийлері және Тізілімге енгізуге арналған өтінім нысаны негізінде қысқаша қарастырылады. Қорытынды бөлімде адам мен жасанды интеллекттің бірлескен жады қалыптастыру үдерісі аясында іздеу негіздегі жады экологиясынан генерацияға негізделген жаңа жады экологиясына көшу мәселесі пайымдалады.

Түйін сөздер: жасанды интеллект; ЮНЕСКО; «Әлем жады» бағдарламасы; құжаттық мұра; жасанды интеллект арқылы жасалған туындылар

ВЛИЯНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА РЕЕСТР ПРОГРАММЫ ЮНЕСКО «ПАМЯТЬ МИРА»: ПРИЗНАНИЕ ПРОИЗВЕДЕНИЙ, СОЗДАННЫХ ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ, ДОКУМЕНТАЛЬНЫМ НАСЛЕДИЕМ

Сан Парк

Стипендиат программы Ad Astra, ассистент-профессор в области искусственного интеллекта и цифрового культурного наследия Школы информационных и коммуникационных исследований, Университетский колледж Дублина, Ирландия
sun.park@ucd.ie; ORCID: https://orcid.org/0000-0001-9307-1171

Аннотация. Стремительное внедрение искусственного интеллекта (ИИ) в деятельность учреждений памяти – в процессы создания, описания и сохранения документов – побуждает вновь обратиться к концептуальным и процедурным основам Программы ЮНЕСКО «Память мира» (Memory of the World, MoW). В рамках обсуждений Рабочей группы ЮНЕСКО MoW по искусственному интеллекту рассматривается вопрос о возможности признания произведений, созданных с помощью ИИ, документальным наследием ЮНЕСКО. В данной презентации кратко раскрывается, каким образом ценности произведений, созданных ИИ, могут соотноситься с целями и принципами Программы MoW, с акцентом на определение документа, критерии номинации и форму заявки для включения в Реестр. В заключении рассматривается переход от экологии памяти, основанной на поиске и извлечении информации, к экологии генеративной памяти, формирующейся в результате совместного создания памяти человеком и искусственным интеллектом.

Ключевые слова: искусственный интеллект; ЮНЕСКО; «Память мира»; документальное наследие; произведения, созданные искусственным интеллектом

Introduction

Artificial intelligence is no longer a speculative concern for memory institutions. It is already shaping the ways records are created, described, selected, preserved, interpreted, and accessed. Drawing on Park's (2024) doctoral dissertation "The Registration of AI-generated Works as UNESCO Heritage: Distributed Knowledge

Systems as Human-algorithmic Common Heritage of Mankind", the UNESCO MoW AI Working Group has discussed the question of how AI-generated works can be understood, nominated, evaluated, and preserved within the MoW framework. We focused on how to address uncertainty when historically significant AI-generated works can be considered documentary heritage and deemed preservation-worthy.

This timing problem frames a set of questions of the associated values of AI-generated works with UNESCO Documentary Heritage. Can AI-generated works be considered as UNESCO Documentary Heritage? How can the values of AI-generated works be described in the current Register nomination form? The remainder of this paper works through these questions in turn. The presentation concludes with reflecting on the shift from a memory ecology of retrieval to generation emerging from human–AI collective memory-making.

Definition and Scope of Documentary Heritage

The 2015 Recommendation concerning the preservation of, and access to, documentary heritage including in digital form defines a document in terms of informational content and a carrier: content that may comprise text, still or moving images, and sounds that can be copied or migrated; a carrier that may itself carry aesthetic, cultural, or technical value, with the relationship between content and carrier ranging from incidental to integral (UNESCO, 2015). Documentary heritage comprises documents or groups of documents of significant and enduring value whose loss would be a harmful impoverishment. These definitions remain sound for the world they describe.

Given that a 'carrier' can be a computational model and 'content' includes products of data processing, AI-generated works can be regarded as both digitised or born-digital documents, based on the 2025 MoW Companion (UNESCO, 2025). A digitised document normally involves converting an existing analogue object into

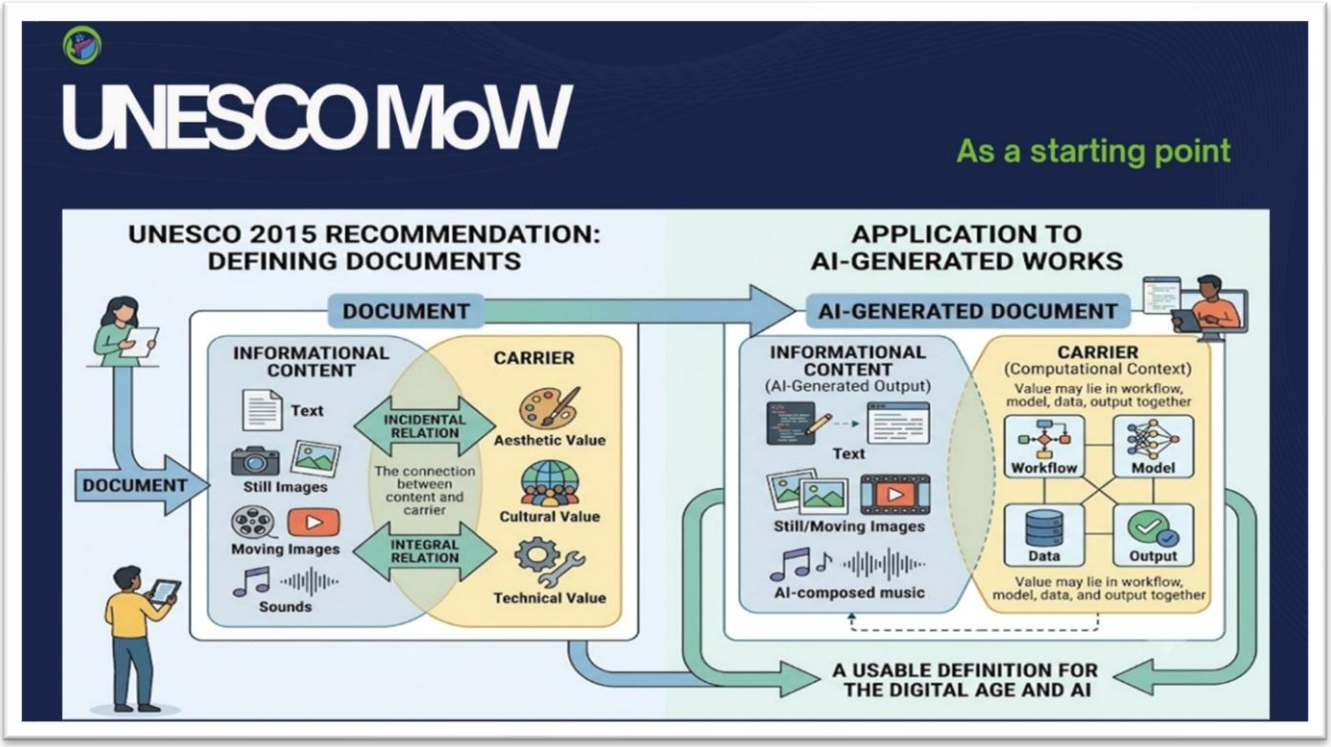
digital form whereas a born-digital document is created in digital form from the outset. AI-generated works can be either of them, containing computational generation, data selection, model behaviour, and iterative human-machine interaction. Their significance may therefore be technological and technical as well as aesthetic, historical or cultural. Such work includes at least one intellectual interactive process, at least one AI system, and at least one end-product. The heritage value as technical may lie partly in the workflow, the selected datasets and the model - as 'carrier' of human knowledge - while the socio-cultural value would be embedded in the human decisions and the socio-technical context that shaped the output - as 'content' of human culture.

Implications for Registration Criteria and Nomination Form

If AI-generated works are deemed documentary heritage, the MoW Programme needs a consistent basis for identifying their value. The question is not whether AI-generated works should automatically qualify for inscription, nor whether AI involvement should automatically disqualify them. Rather, the question is what evidence is needed for evaluators to judge authenticity, significance, rarity, community value and preservation feasibility.

The current MoW nomination framework already includes criteria that are relevant to AI-generated works. Three primary criteria are most directly engaged. *'Authenticity'* concerns whether a work is real, true, and uncorrupted from the original, with a reliable identity and provenance. *'Historical significance'* concerns what the heritage tells us about the history of the world. *'Social, community, or spiritual significance'* concerns its attachment to a specific existing community or incident. In terms of authenticity, some AI-generated works cannot be reproduced, because the generative nature of their algorithmic models makes each output, in effect, singular. Assessment should ask whether the work embodies a one-of-a-kind workflow that has not been repeated before. Nominators may need to explain how the work was produced, which parts were AI-generated, what AI systems were used, and how identity and provenance can be verified. Some AI-generated works may be non-repeatable because the generative workflow cannot be exactly duplicated.

For historical significance, technological evolution can be considered part of documentation and science history. Evaluators of AI-generated works should ask what socio-cultural and technological circumstances led to the choice of particular algorithms and datasets, and whether the work involved any distinctive developments in data curation, processing, or algorithm design. AI-generated works are associated with socio-cultural values too, especially when their data come from, or address local communities' cultural identity. In such cases, nominators should be able to explain whether relevant communities were involved at appropriate stages. This is especially important when datasets, cultural materials, or identity-related representations involve communities whose knowledge or experiences are being interpreted through AI systems, computer scientists, local communities, curators, archivists, and other collaborators.



Among the comparative criteria, *'uniqueness or rarity'* asks whether a work is the only one of its kind ever created, or one of few survivors of a larger group, and the framework also provides for inadmissible or contested nominations. The uniqueness of an AI-generated work can emerge from specific reasons why and how a given task was assigned to a particular AI system. Or work may be unique because it is the only surviving product of a larger number of generated outputs, because its model environment can no longer be reproduced. Nominators need to explain such relationships between the rarity of the final output and the rarity of the workflow.

There would be the possibility of joint nominations collaborated by data providers, AI model developers and a community across the public sector, private sector and civil society or individuals. Their contributions, copyrights of datasets and algorithmic models need to be detailed in the Register nomination form. Such joint nomination is followed by shared responsibility of preserving the AI-generated work as UNESCO Documentary Heritage. Institutions may need to preserve not only the final product but also datasets, model documentation, training methods, software environments, rights information and curatorial decisions. The concepts of meta-data (data about data) and para-data (data about workflow) may be helpful for custodians to preserve the processual information of the nominated AI-generated work.

Conclusion: AI and a New Memory Ecology

Hoskins (2024) describes a new memory ecology in which human and AI agents are deeply entangled - an entanglement that both enables and endangers human agency in the making, and the remixing, of individual and collective memory. It changes how individual and collective memory are made, remixed, retrieved and generated. For the MoW Programme, this means that AI should not be treated only as a technical tool. It is part of a wider transformation in the production, circulation

and preservation of memory. The MoW Programme could serve as a place to cover and address the necessary principles of valorising AI-generated works as digital documentary heritage, followed by the guidelines on long-term preservation responsibilities. The Programme can respond to AI in a way that is transparent, sustainable and aligned with its mandate, 'universal access to information'.

References

Hoskins, A. (2024). AI and memory. Memory, Mind & Media, 3, e18. https://doi.org/10.1017/mem.2024.16

Park, S. (2024). The registration of AI-generated works as UNESCO heritage: Distributed knowledge systems as human-algorithmic common heritage of mankind. PhD thesis, University of Warwick. https://wrap.warwick.ac.uk/id/eprint/190274/ (accessed 28.05.2026).

UNESCO. (2015). Recommendation concerning the preservation of, and access to, documentary heritage including in digital form. https://www.unesco.org/en/legal-affairs/recommendation-concerning-preservation-and-access-documentary-heritage-including-digital-form (accessed 28.05.2026).

UNESCO. (2021). Memory of the World: General Guidelines. https://unesdoc.unesco.org/ark:/48223/pf0000378405 (accessed 28.05.2026).

UNESCO. (2025). Memory of the World register companion. https://unesdoc.unesco.org/ark:/48223/pf0000393977 (accessed 28.05.2026).

PRESERVATION OF AI-GENERATED WORKS AS DOCUMENTARY HERITAGE

Lai Tee Phang
Senior Deputy Director and Senior Principal Archivist of the National Archives of Singapore (NAS)
PHANG_Lai_Tee@nlb.gov.sg

ЖИ ЖАСАҒАН ТУЫНДЫЛАРДЫ ҚҰЖАТТЫҚ МҰРА РЕТІНДЕ САҚТАУ

Лай Ти Фанг
Сингапур Ұлттық архивінің (NAS) аға директор орынбасары және бас архившісі
PHANG_Lai_Tee@nlb.gov.sg

СОХРАНЕНИЕ ПРОИЗВЕДЕНИЙ, СОЗДАННЫХ ИИ, КАК ДОКУМЕНТАЛЬНОГО НАСЛЕДИЯ

Лай Ти Фанг
Старший главный архивариус и старший заместитель директора Национального архива Сингапура (NAS)
PHANG_Lai_Tee@nlb.gov.sg

UNESCO Memory of the World

Preservation Sub-Committee

Was Sub-Committee on Technology, 1995 - 2019

- Oversees and coordinates all activities related to the preservation of Documentary Heritage, including in digital form
- Reviews developments in preservation and provides advice on analogue and digital preservation matters in response to referrals or requests from International Advisory Committee, its subsidiary bodies and the Secretariat, and from regional or national MoW committees
- Serves as an expert group responding to questions about storage and preservation of documents and their accessibility
- Recommends and undertakes preservation-related studies, produces publications and organizes training events
- Engages professional communities to identify policy gaps and issues, and advises UNESCO's policy work in the preservation and accessibility of Documentary Heritage

UNESCO MoW

Working Group on AI

Formed in May 2025 to study
- the impact of AI on documentary heritage
- how to manage the way AI affects the MoW programme

IAC/PSC Representation

Dr Phang Lai Tee
Chair, Preservation Sub-Committee (PSC) and Member, International Advisory Committee (IAC)
Catherine Bloch
Former Vice Chair, IAC
President of Mexican National Committee
Dr Natasa Milic Frayling
Member, PSC & Chair, PERSIST Research & Technology Working Group
Dr Victoria Lemieux
Representative, Sub-Committee on Education and Research
Professor of Archival Science
University of British Columbia, Canada

Academia

Dr Sun Park
Ad Astra Fellow/Assistant Professor in AI and Digital Cultural Heritage
School of Information and Communication Studies
University College Dublin, Ireland

AI Practitioners

Olle Green
Lead Machine Learning Engineer at North Rock Capital
Author of: *Bob The Robot: Exploring the Universe*

Boris Eldagsen
Artist, creator of the *PSEUDOMNESIA | The Electrician, "The picture that stopped the world"* (The Guardian)
LABASAD School of Art & Design, Barcelona
Deutscher Fotorat

Preliminary findings of study to date

- Current definition of document
- Current registration form
- Preservation
- Broader coverage for AI-generated document
- May need public-private joint nomination and additional information/documentation
- Lacks guidance for preservation of AI-generated document, e.g. meta/para-data
- Issue of reproducibility

unesco archives SINGAPORE

5 Categories for Preservation of AI-generated Works

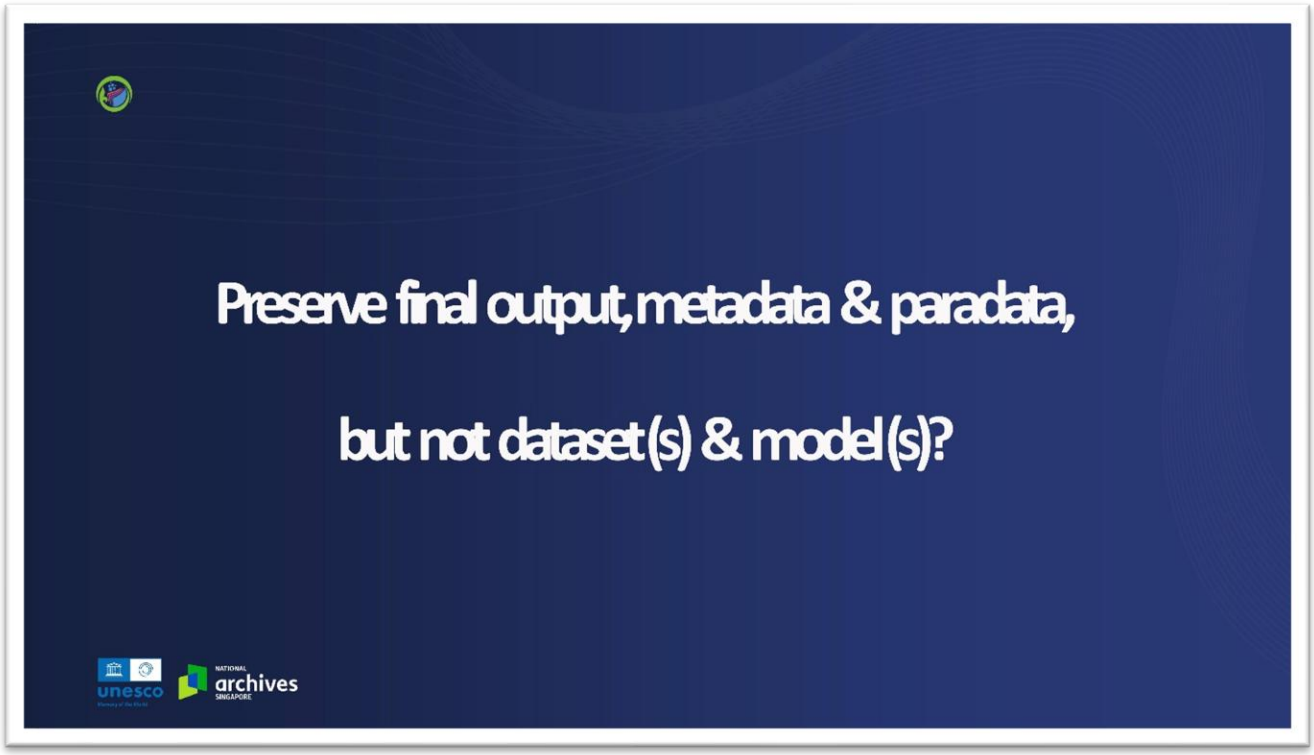
Descriptive metadata	Dataset(s), AI model(s), training methods, end product (the work), location of the work
Technical metadata	• How the datasets and algorithms used in an AI-generated work will be preserved • How particular devices or software and computing environment that (dis)play the work and perform the required content generation will be preserved
Paradata	Information about the procedure(s) and tools used to create and process datasets, along with information about the persons carrying out those procedures
Administrative metadata	Information about custodians, managers and any other individuals and organisations in charge of the above
Meta-metadata	How the afore-mentioned metadata will be managed by whom

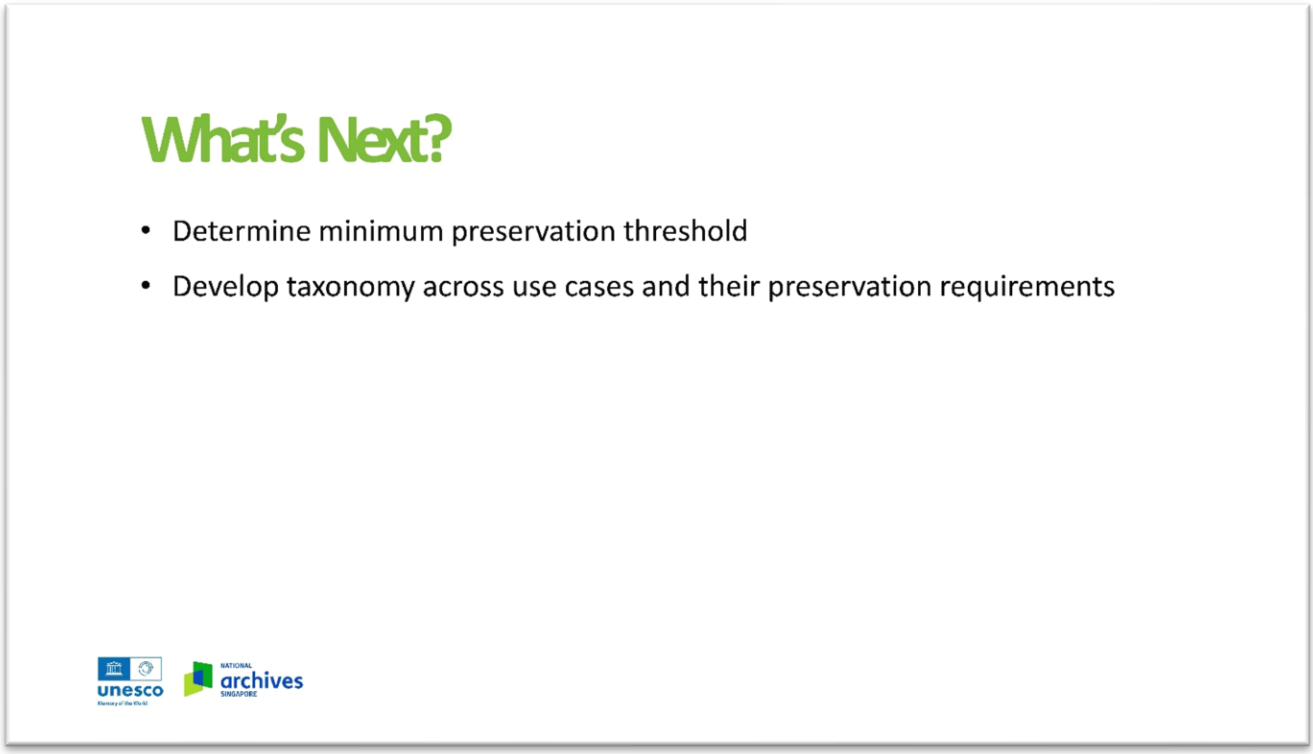
unesco archives SINGAPORE

Further Preservation Concerns

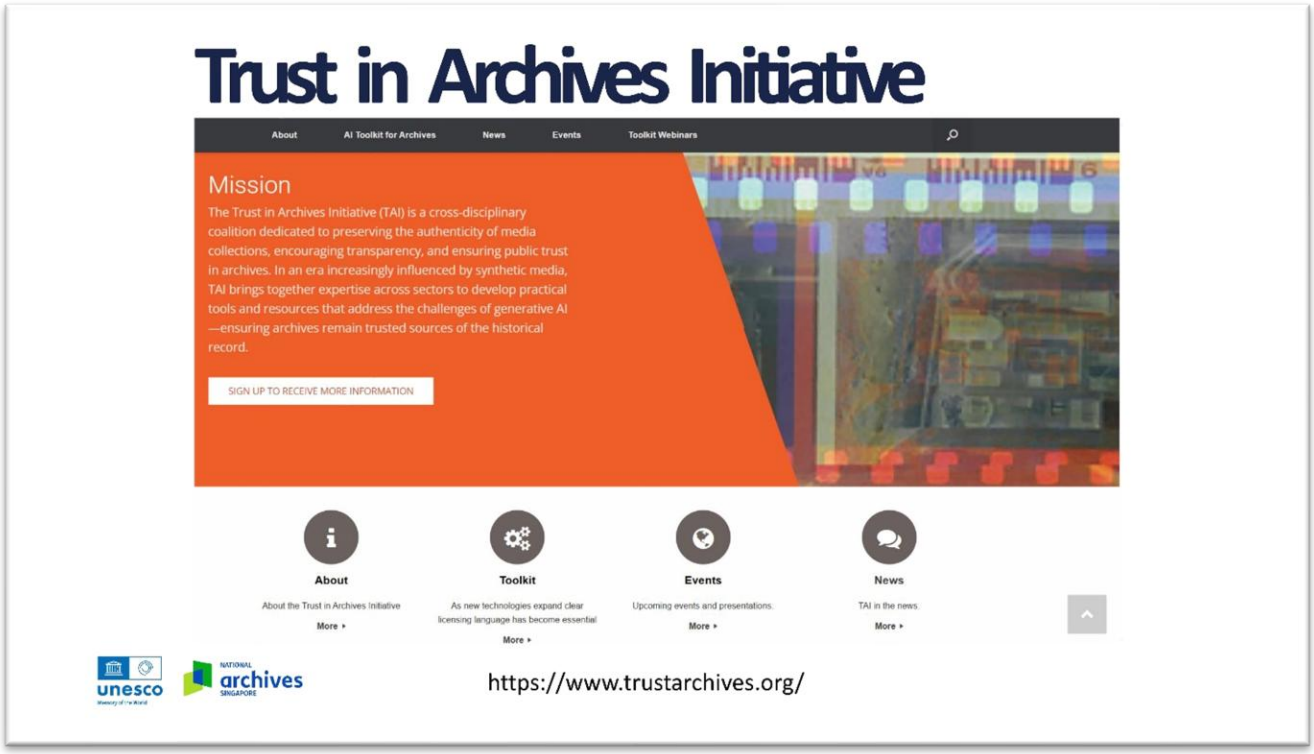
- An AI-generated work cannot be reproduced if one of its two core elements - dataset(s) and algorithmic model(s) – does not exist
- Even with the same dataset and model, a different AI-generated work can be created if the training method differs from one practitioner to another
- Reliability of metadata and paradata provided in submissions – no authoritative mechanism exists for verification
- Practicality of documenting and preserving datasets and models

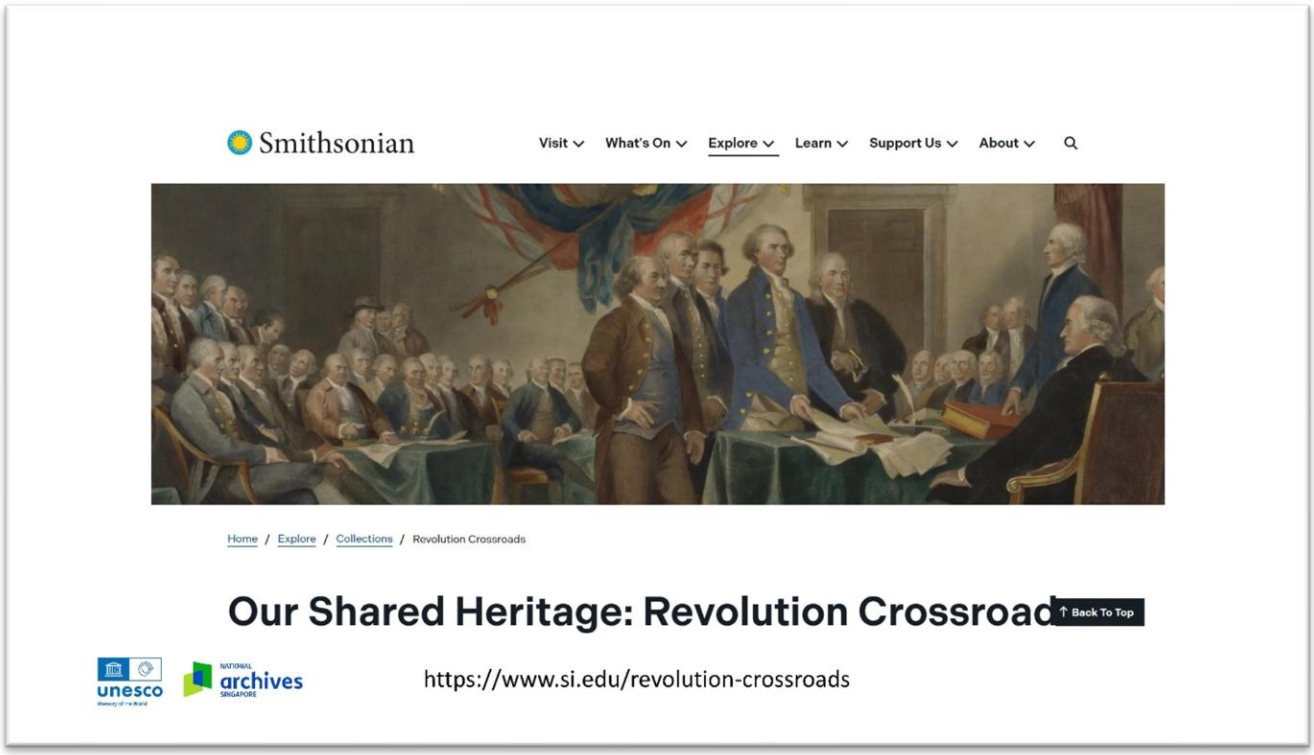
BALANCING THE IDEAL WITH THE PRACTICAL?

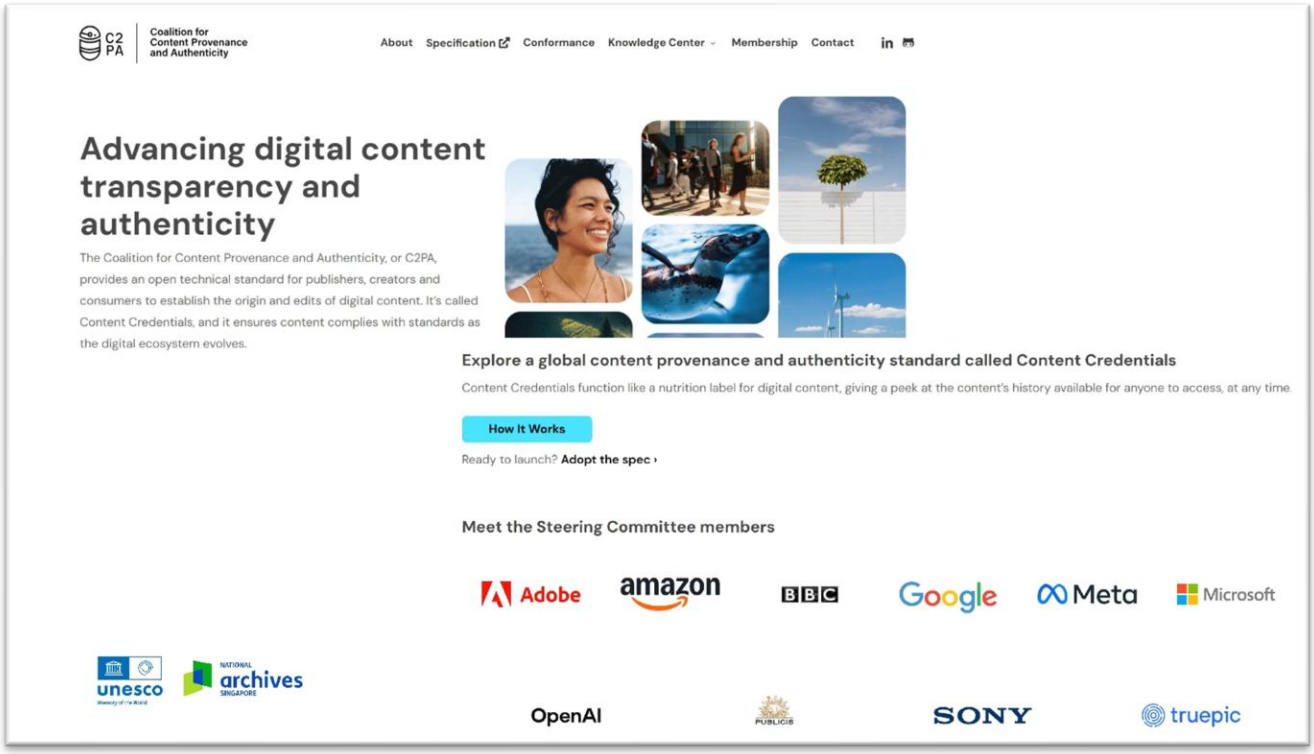

Preserve final output, metadata & paradata,
but not dataset(s) & model(s)?
unesco
NATIONAL archives SINGAPORE


What's Next?
• Determine minimum preservation threshold
• Develop taxonomy across use cases and their preservation requirements
unesco
NATIONAL archives SINGAPORE

Useful Resources
unesco
NATIONAL
archives
SINGAPORE


Trust in Archives Initiative
About
AI Toolkit for Archives
News
Events
Toolkit Webinars
Mission
The Trust in Archives Initiative (TAI) is a cross-disciplinary coalition dedicated to preserving the authenticity of media collections, encouraging transparency, and ensuring public trust in archives. In an era increasingly influenced by synthetic media, TAI brings together expertise across sectors to develop practical tools and resources that address the challenges of generative AI —ensuring archives remain trusted sources of the historical record.
SIGN UP TO RECEIVE MORE INFORMATION
About
About the Trust in Archives Initiative
More
Toolkit
As new technologies expand clear licensing language has become essential
More
Events
Upcoming events and presentations
More
News
TAI in the news
More
unesco
NATIONAL
archives
SINGAPORE
https://www.trustarchives.org/


Smithsonian
Visit
What's On
Explore
Learn
Support Us
About
Home / Explore / Collections / Revolution Crossroads
Our Shared Heritage: Revolution Crossroad
Back To Top
unesco
archives
https://www.si.edu/revolution-crossroads


C2PA
Coalition for Content Provenance and Authenticity
About
Specification
Conformance
Knowledge Center
Membership
Contact
Advancing digital content transparency and authenticity
The Coalition for Content Provenance and Authenticity, or C2PA, provides an open technical standard for publishers, creators and consumers to establish the origin and edits of digital content. It's called Content Credentials, and it ensures content complies with standards as the digital ecosystem evolves.
Explore a global content provenance and authenticity standard called Content Credentials
Content Credentials function like a nutrition label for digital content, giving a peek at the content's history available for anyone to access, at any time.
How It Works
Ready to launch? Adopt the spec ›
Meet the Steering Committee members
Adobe
amazon
BBC
Google
Meta
Microsoft
unesco
archives
OpenAI
PUBLICIS
SONY
truepic

«АРХИВТЕРДІҢ ИННОВАЦИЯЛЫҚ ЭКОЖҮЙЕЛЕРІ: ИНФРАҚҰРЫЛЫМ, ЫНТЫМАҚТАСТЫҚ ЖӘНЕ ТҰРАҚТЫЛЫҚ»
№ 3 сессия

Сессия № 3
«ИННОВАЦИОННЫЕ ЭКОСИСТЕМЫ АРХИВОВ: ИНФРАСТРУКТУРА, СОТРУДНИЧЕСТВО И УСТОЙЧИВОСТЬ»

«INNOVATIVE ARCHIVE ECOSYSTEMS: INFRASTRUCTURE, COLLABORATION AND SUSTAINABILITY»
Session No. 3

ARTIFICIAL INTELLIGENCE APPLICATIONS IN LIBRARY AND INFORMATION SCIENCE: AN APPLIED VISION FOR DEVELOPING THE SERVICES OF THE JORDANIAN NATIONAL LIBRARY, 2025

Firass Atallah Darabah
Director General Department of the National Library of Jordan
nl@nl.gov.jo

Abstract. This study examines the potential of artificial intelligence (AI) applications in advancing the services of national libraries, with particular emphasis on the Department of the Jordanian National Library (JNL) as the country's primary institutional custodian of the national memory, legal deposit, documentary archives, and the protection of Jordanian intellectual output. The research provides a comprehensive overview of the most significant AI application domains relevant to national library and archival environments, including intelligent automated cataloguing, subject analysis and automatic abstracting, optical character recognition (OCR) for Arabic-script documents, advanced semantic search systems, and applications for documentary management and preservation of the national narrative in digital space.

The study also addresses a range of challenges inherent in this digital transition, foremost among which are ethical concerns surrounding privacy and data protection, the limitations of large-scale AI language models in handling Arabic-language content, and the legal complexities surrounding intellectual property rights in digital environments. The principal contribution of this research lies in the development of a phased, context-sensitive implementation model specifically designed for the JNL, one that reflects the distinctive nature of national library and archival work while maintaining a careful balance between technological modernisation and the preservation of human professional judgment. The study concludes with practical recommendations advocating for the formulation of an institutional AI policy, the development of standardised Arabic datasets for cataloguing and

retrieval, and the integration of AI initiatives with existing service frameworks such as electronic legal deposit, the Jordanian Union Catalogue, and the Copyright Protection Office.

Keywords: Artificial Intelligence, National Libraries, Library and Information Science, Jordanian National Library, Intelligent Cataloguing, Digital Archiving, Natural Language Processing, Legal Deposit, Intellectual Property Rights, Data Governance

КІТАПХАНА ЖӘНЕ АҚПАРАТТЫҚ ҒЫЛЫМ САЛАСЫНДА ЖАСАНДЫ ИНТЕЛЛЕКТТІ ҚОЛДАНУ: 2025 ЖЫЛҒЫ ИОРДАНИЯ ҰЛТТЫҚ КІТАПХАНАСЫ ҚЫЗМЕТТЕРІН ДАМЫТУДЫҢ ҚОЛДАНБАЛЫ КӨЗҚАРАСЫ

Фирас Аталла Дарабах
Иордания Ұлттық кітапханасы департаментінің бас директоры
nl@nl.gov.jo

Аңдатпа. Бұл зерттеуде ұлттық кітапханалардың қызметтерін жетілдірудегі жасанды интеллект (ЖИ) технологияларын қолдану әлеуеті қарастырылады. Зерттеуде елдің ұлттық жадының негізгі институционалдық сақтаушысы, міндетті дана қорының, құжаттық архивтердің сақталуын қамтамасыз етуші және Иорданияның зияткерлік өнімдерін қорғаушы ұйым ретінде Иордания Ұлттық кітапханасы (JNL) департаментінің қызметіне ерекше назар аударылады. Зерттеу ұлттық кітапхана және архив ортасына қатысты жасанды интеллектті қолданудың негізгі бағыттарына жан-жақты шолу жасайды. Олардың қатарында интеллектуалды автоматтандырылған каталогтау, тақырыптық талдау және автоматты рефераттау, араб жазуындағы құжаттарға арналған оптикалық таңбаларды тану (OCR) технологиясы, жетілдірілген семантикалық іздеу жүйелері, сондай-ақ цифрлық кеңістіктегі құжаттарды басқару және ұлттық тарихи жадыны сақтауға арналған қолданбалар бар.

Зерттеуде цифрлық трансформация үдерісіне тән бірқатар мәселелер де қарастырылады. Олардың ішінде құпиялылық пен деректерді қорғауға қатысты этикалық мәселелер, араб тіліндегі мазмұнды өңдеу кезінде ірі тілдік модельдердің шектеулері, сондай-ақ цифрлық ортадағы зияткерлік меншік құқықтарына байланысты құқықтық күрделіліктер ерекше атап көрсетіледі. Зерттеудің негізгі ғылыми үлесі – Иордания Ұлттық кітапханасының ерекшеліктерін ескеретін, кезең-кезеңімен іске асыруға арналған және контекстке бейімделген модельді әзірлеуде жатыр. Бұл модель кітапхана-архив саласындағы кәсіби қызметтің өзіндік ерекшеліктерін сақтай отырып, технологиялық жаңғырту мен мамандардың кәсіби пайымын теңгерімді ұштастыруға бағытталған.

Зерттеу қорытындысында мекемелік жасанды интеллект саясатын әзірлеу, каталогтау және ақпарат іздеу үдерістері үшін стандартталған араб тіліндегі деректер жиындарын қалыптастыру, сондай-ақ ЖИ бастамаларын электрондық міндетті дана жүйесі, Иорданияның біріккен каталогы және Авторлық құқықтарды қорғау кеңсесі сияқты қолданыстағы қызметтік құрылымдармен ықпалдастыру жөніндегі практикалық ұсыныстар беріледі.

Түйін сөздер: жасанды интеллект; ұлттық кітапханалар; кітапхана және ақпараттық ғылым; Иордания Ұлттық кітапханасы; интеллектуалды каталогтау; цифрлық архивтеу; табиғи тілді өңдеу; міндетті дана; зияткерлік меншік құқықтары; деректерді басқару

ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОБЛАСТИ БИБЛИОТЕЧНОГО И ИНФОРМАЦИОННОГО ДЕЛА: ПРИКЛАДНОЙ ПОДХОД К РАЗВИТИЮ УСЛУГ НАЦИОНАЛЬНОЙ БИБЛИОТЕКИ ИОРДАНИИ В 2025 ГОДУ

Фирас Аталла Дарабах
Генеральный директор Департамента Национальной библиотеки Иордании
nl@nl.gov.jo

Аннотация. В исследовании рассматривается потенциал применения искусственного интеллекта (ИИ) для развития услуг национальных библиотек с особым акцентом на Департамент Национальной библиотеки Иордании (JNL) как ключевого институционального хранителя национальной памяти страны, обязательного экземпляра, документальных архивов и защитника интеллектуальной продукции Иордании. Представлен комплексный обзор основных направлений применения ИИ в национально-библиотечной и архивной среде, включая интеллектуальную автоматизированную каталогизацию, тематический анализ и автоматическое реферирование, технологию оптического распознавания символов (OCR) для документов на арабской письменности, современные системы семантического поиска, а также решения для управления документами и сохранения национального исторического нарратива в цифровом пространстве.

В исследовании также рассматривается ряд проблем, связанных с цифровой трансформацией, среди которых особое место занимают этические вопросы обеспечения конфиденциальности и защиты данных, ограничения крупных языковых моделей ИИ при работе с арабоязычным контентом, а также правовые сложности, связанные с защитой интеллектуальной собственности в цифровой среде. Основной научный вклад исследования заключается в разработке поэтапной модели внедрения ИИ, учитывающей специфику деятельности Национальной библиотеки Иордании. Данная модель направлена на обеспечение баланса между технологической модернизацией и сохранением профессионального человеческого суждения в библиотечной и архивной практике.

В заключении представлены практические рекомендации по формированию институциональной политики в области ИИ, созданию стандартизированных арабоязычных наборов данных для каталогизации и поиска информации, а также интеграции инициатив по внедрению ИИ с существующими сервисными системами, включая электронный обязательный экземпляр, Сводный каталог Иордании и Управление по защите авторских прав.

Ключевые слова: искусственный интеллект; национальные библиотеки; библиотечное и информационное дело; Национальная библиотека Иордании; интеллектуальная каталогизация; цифровое архивирование; обработка естественного языка; обязательный экземпляр; права интеллектуальной собственности; управление данными

1. Introduction

Hardly a day passes without the term "Artificial Intelligence" being raised in an institutional, academic, or industrial context. This overwhelming presence has forced itself upon knowledge institutions of all kinds, ranging from universities and research centres to national libraries and archives. Perhaps what distinguishes national libraries

from other information institutions is that their mission bridges the past, present, and future simultaneously; they are tasked with receiving, organising, and providing access to contemporary intellectual output, while concurrently preserving historical and heritage documents for generations yet unborn.

In this context, the Department of the Jordanian National Library stands at the forefront of national knowledge institutions, as it shoulders monumental responsibilities. These include the legal deposit of everything produced and published on Jordanian soil, the preservation of official and historical documents, the management of the Jordanian Union Catalogue, the operation of the Copyright Protection Office, and the provision of information services to researchers and government entities. This complex institutional composition makes the Jordanian National Library a fertile environment for applying AI technologies – not merely in imitation of Western libraries, but stemming from a genuine, verified national need.

The International Federation of Library Associations and Institutions (IFLA) asserts that AI represents a transformative opportunity for libraries, provided that library values and principles—such as freedom of access to information, privacy, and accountability – remain central as a governance framework governing any decision to adopt or deploy these technologies (IFLA, 2023). From this standpoint, this research proposes an integrated applied vision that answers a fundamental question: How can the Jordanian National Library employ artificial intelligence to serve its national mission without evading professional responsibility?

2. Research Problem, Questions, and Objectives

2.1 Research Problem

The main research problem can be formulated as follows: How can artificial intelligence applications be deployed to develop the services of the Department of the Jordanian National Library in a way that enhances the efficiency of national knowledge management, preserves Jordanian documents and their narrative, while ensuring compliance with applicable ethical and legal standards?

From this main problem, a set of sub-questions emerges which this research seeks to answer.

2.2 Research Questions

What are the prominent AI applications currently deployed globally in national libraries and archival institutions?

What are the fields most viable for implementation within the Department of the Jordanian National Library, given its current operational structure?

What are the ethical, legal, and technical challenges associated with employing AI in a national library environment?

How can a progressive national model for AI adoption be constructed to suit the specific nature of the Jordanian National Library?

2.3 Research Objectives

To review the precise scientific concept of AI applications within library and information science.

To monitor the key areas of utilization in national libraries at the global level.

To analyze the operational reality of the Department of the Jordanian National Library and the available development opportunities.

To propose a phased, gradual applied model for employing artificial intelligence.

To highlight the relevant legal and ethical dimensions, particularly in the fields of documentation, intellectual property, and data governance.

2.4 Significance of the Research

The significance of this research stems from three interconnected levels:

Theoretical Level: It provides a scientific framework for the relationship between AI and library and information science, taking into account the Arabic context and its linguistic and cultural specificities.

Applied Level: It leads to a model capable of actual adoption and experimentation within the Jordanian National Library.

Policy Level: It serves as an analytical reference for decision-makers in national information institutions seeking to formulate comprehensive digital strategies.

3. Conceptual Framework

3.1 The Concept of AI in the Context of Libraries

In library and information science literature, Artificial Intelligence is defined as a system of computing technologies capable of simulating human cognitive processes related to learning, reasoning, and decision-making, specifically within contexts involving information management, organisation, and access. Its deployment in libraries is not intended to replace specialists and librarians, but rather to serve as an intelligent support tool that opens horizons that could not be achieved by human effort alone amidst the accelerated explosion of digital content.

The core technologies constituting this field span four main axes:

1. Machine Learning (ML): Based on training models on data and extracting patterns from them.

2. Natural Language Processing (NLP): Allowing machines to understand and analyze human texts.

3. Computer Vision: Enabling systems to interpret images and visual documents.

4. Large Language Models (LLMs): Enhancing the machine's capacity for contextual understanding and coherent textual generation.

3.2 Library Science in the Digital Environment

Since the dawn of the third millennium, library and information science has been undergoing a profound structural transformation. This shift is not confined to digitizing collections and converting paper catalogues into electronic databases, but has extended to touch the core concepts of the profession itself.

The concept of "preservation" no longer means keeping a paper copy in a secure drawer; it now requires formulating long-term Digital Preservation policies that account for formatting obsolescence and software changes. Similarly, the concept of "access" has transcended its traditional boundaries to become a complex debate regarding priorities and the balance between the public's right to access and the author's right to protect their creation. In this intricate context, AI emerges to offer solutions to old problems while raising new questions. On one hand, it can accelerate cataloguing, improve search capabilities, and expand the scope of access; on the other hand, it calls for a meticulous review of established professional standards in cataloguing, classification, and preservation.

4. AI Applications in the Field of Libraries and Information

4.1 Intelligent Automated Cataloguing

Cataloguing remains one of the technical processes that most depletes the time and effort of library specialists, especially in light of the accelerated influx of publications and electronic publishing, which has opened the door to diverse materials not subject to traditional publishing controls. AI models trained on massive bibliographic databases have proven their ability to suggest subject headings with an accuracy level ranging between 75% and 90% in the best applications available today, according to research published by the IFLA Journal in its special issue on AI in 2025.
Crucially, intelligent cataloguing does not only accelerate output, but also contributes to standardizing subject processing and reducing discrepancies between different catalogues. This addresses a real problem faced by the Jordanian Union Catalogue when multiple entities supply it with bibliographic records.

4.2 Optical Character Recognition (OCR) for Arabic Documents

The digitization of historical documents and their conversion into searchable, indexable, and analysable text represents one of the most prominent practical applications of AI in national archival environments. Optical Character Recognition (OCR) technology has witnessed a qualitative leap with the emergence of deep learning models, particularly regarding Arabic scripts. Arabic text previously posed a major challenge due to the cursive nature of the characters, the diversity of calligraphy styles, and overlapping spacing. National libraries worldwide, including the Library of Congress and the National Archives (UK), utilize this technology to process their manuscripts and archives, resulting in millions of documents being released for public search after being frozen in boxes and on shelves.

4.3 Semantic Search and Intelligent Retrieval Systems

AI-driven semantic search goes beyond the limitations of classical keyword searching. It allows users to search using natural language (e.g., "Jordanian fictional works that dealt with the Nakba experience"), whereby the system analyzes the intent behind the query and matches it with relevant content, even if the terms do not match literally.

The National Library of Israel launched a pioneering experiment in this field with an AI-powered tool to search its repository of images and documents, allowing text-based and visual-similarity searches within the digital archive. The technical methodology of this model can serve as inspiration, while maintaining full consideration for the documentary specificity and the Jordanian national context.

4.4 Chatbots and User Services

Chatbots backed by LLMs have become remarkably adept at handling routine user inquiries, whether guiding them to specific sources, explaining legal deposit procedures, or answering general knowledge questions. This does not eliminate the role of human reference librarians, but rather liberates them from repetitive queries to focus on matters requiring deep expertise and specialized knowledge.

4.5 Management of Copyright and Intellectual Property

The World Intellectual Property Organization (WIPO) is working on initiatives that link artificial intelligence with the infrastructure of IP management. This reflects an accelerating global trend toward employing these technologies in examining similarities, tracking unauthorized use, and classifying disputes (WIPO, 2024). Given the mandates of the Copyright Protection Office under the Department of the Jordanian National Library, employing AI-based plagiarism detection and textual comparison tools emerges as a critically important strategic option.

4.6 Preservation of the National Narrative and Collective Memory Management

Artificial intelligence possesses unique capabilities in linking written documents with images, audio recordings, and historical events within an integrated semantic network. This enables the creation of interactive "National Memory Maps," which become a renewed educational and research asset for future generations. Furthermore, automated analysis algorithms help uncover documentation gaps in archives, alerting specialists to regions, events, and personalities that merit urgent documentation.

5. The Department of the Jordanian National Library as a Field of Application

5.1 Institutional Identity and Mission

In its official institutional description, the Department of the Jordanian National Library asserts that it represents the nation's memory, the state's library, and the primary reference for researchers and interested parties. It undertakes core duties centered around four main pillars:

Receiving legal deposits for everything published and produced on Jordanian soil.

Preserving and safeguarding the Jordanian documentary and intellectual heritage from loss.

Making its collections available to researchers and institutions.

Protecting copyright and regulating intellectual property.

The Department possesses a diverse service ecosystem including the "Wathiq" platform, the Jordanian Union Catalogue, and the Copyright Protection Office, alongside research assistance and reference guidance services.

5.2 Proposed Areas of Application

First: An Intelligent Legal Deposit System: AI can contribute to raising the efficiency of the legal deposit system through automated verification of the completeness of the deposited work's data, suggesting preliminary subject classifications, matching author and publisher data with reference databases, and detecting duplication in cases of re-deposits or multiple editions. This partial automation will make a tangible difference in processing speed and data accuracy without eliminating the final human review.

Second: Intelligent Cataloguing for Jordanian Intellectual Output: The National Library of Jordan possesses a massive repository of bibliographic records reflecting Jordanian intellectual production over decades. Applying intelligent cataloguing models to this repository can enhance the quality of subject headings, standardize the treatment of authors' names, and create automatic links between different editions of a single work.

*Third: Intelligent Documentary Archiving***:** This axis is one of the most critical in terms of long-term national impact. Digitizing historical Jordanian documents using advanced OCR, and subjecting them to automatic semantic analysis to extract names, places, and dates to build relational networks, will transform the archive from a mere storage repository into a dynamic research platform.

Fourth: Supporting the Copyright Protection Office: AI can perform preliminary similarity checks between submitted works and what is already registered in the national database or available on the internet. This saves specialists valuable time, enabling them to focus on complex dispute cases that require professional and legal judgment.

6. Ethical and Legal Dimensions and Challenges

6.1 Privacy and Data Protection Issues

AI systems in library environments raise serious questions regarding user privacy. Search data, usage patterns, and circulation records consists of a sensitive information profile that reveals individuals' interests, intellectual leanings, and perhaps political orientations. IFLA emphasizes that protecting the intellectual privacy of the user represents an unshakeable ethical principle in the library profession, one that must not be compromised under the pretext of technical improvement (IFLA Statement on Privacy, 2023). In the specific Jordanian context, official and archival documents demand heightened sensitivity, as some contain personal or institutional information that must not be subjected to automated processing without clear legal controls.

6.2 Algorithmic Bias and Arabic Language Representation

Large-scale AI models are predominantly subjected to intensive training on English data, making their performance in processing Arabic content sub-optimal compared to European languages. This disparity manifests in the accuracy of OCR for handwritten Arabic texts, the quality of Arabic subject heading suggestions, and the level of understanding of the broader Jordanian and Arabic cultural and historical context. This imposes a double duty on the Jordanian National Library: on one hand,

working to build high-quality, standardized Arabic datasets to train local models or improve the performance of available ones; on the other hand, reviewing AI outputs humanly before adoption to verify their thematic and linguistic accuracy.

6.3 Intellectual Property in the Age of AI

Generative AI models raise complex legal issues that international and national legislations have yet to resolve, particularly concerning the use of copyrighted works in training processes and the ownership of content generated by these models. The Jordanian National Library sits at the heart of this debate by virtue of its dual mandate as a custodian of protected works and the manager of the Copyright Protection Office.

6.4 Transparency and Accountability

Algorithmic decisions related to cataloguing, search result ranking, or work classification should not be an uninterpretable "black box" exempt from accountability. This principle demands that intelligent systems be designed to provide clear explanations for their outputs, allowing for human review and correction when necessary.

7. Proposed Applied Model for the Jordanian National Library

Based on the preceding analysis, this research proposes a phased, gradual applied model consisting of four stages, tailored to fit the operational reality of the JNL and the logical progression of technology adoption:

Stage 1: Institutional Preparation: This stage begins by forming a specialized smart transformation task force comprising library and information specialists, IT experts, and a legal consultant specializing in intellectual property. Key expected deliverables include: an institutional policy document for AI use, an audit report of existing systems to identify scalable priorities, and an initial literacy program for staff.

Stage 2: Low-Risk Core Applications: This stage targets applications that generate a tangible impact while carrying limited ethical and technical risks. These include: applying OCR to archival documents, activating duplication-detection tools in the legal deposit database, and piloting a chatbot for routine inquiries. This serves as a testing and learning phase to accumulate the institutional experience necessary to move forward.

*Stage 3: Advanced Applications***:** This stage involves a deeper embedding of AI into the core of technical operations. This is achieved by activating intelligent cataloguing and linking its outputs to the Jordanian Union Catalogue, launching advanced semantic search systems within databases, constructing semantic networks for archival documents to link them to the national narrative, and deploying AI tools to support the Copyright Protection Office.

Stage 4: Governance and Continuous Evaluation: This stage does not signify the end of development, but rather acts as the oversight and evaluative ecosystem accompanying the application to govern its direction. It includes defining measurable performance indicators, conducting periodic human reviews of smart system outputs, establishing ethical auditing mechanisms to detect bias or hallucination in results, and updating policies in light of emerging legislative and technical shifts.

8. Proposed Performance Indicators

The following table outlines a set of Key Performance Indicators (KPIs) proposed by the research to measure the success of AI projects in the Jordanian National Library:

Indicator	Target	Timeframe
Percentage of documents digitized via OCR annually	80% of new additions	Year One
Accuracy of OCR results in Arabic documents	>= 90%	Year One
Rate of duplication reduction in the Union Catalogue	>= 70%	Years One and Two
Processing time for legal deposits	Reduced by 50%	Year Two
Researcher satisfaction rate with semantic search	>= 85%	Year Two
Number of similarity detection cases in the Copyright Office	Annual performance indicator	Continuous

9. Results

The research yielded a set of findings that can be summarized as follows:

AI provides genuine transformative capabilities for national libraries; however, exploiting them requires a clear institutional framework to govern adoption and define priorities.

The Department of the Jordanian National Library possesses an integrated service ecosystem, rendering each unit a viable entry point for specific, high-impact AI applications.

Arabic documents and content present a real technical challenge for most off-the-shelf AI models, necessitating investment in building localized, standardized Arabic datasets.

The deployment of AI in the JNL cannot succeed without an ethical and legal policy establishing its boundaries and defining necessary areas of human oversight.

The proposed applied model reinforces the concept of "human reviews and machine assists" as a governing operational philosophy, rather than a mere rhetorical principle.

10. Recommendations

Draft a mini institutional AI strategy within the National Library that includes guiding principles, priorities, and governance and review mechanisms.

Begin with supportive applications that enhance professional decision-making without replacing it, particularly in cataloguing, archiving, and duplication detection.

Develop a clear ethical and legal policy defining the rules governing AI use in processing documents and user data.

Invest in building high-quality, standardized Arabic datasets for cataloguing and search, while encouraging national and regional cooperation to share them and expand their utilization.

Organically link AI projects with the existing service ecosystem in the Department: electronic legal deposit, the Jordanian Union Catalogue, the Copyright Protection Office, and the national document archive.

Design a specialized training program in "AI Literacy" directed at librarians and archivists, focusing on understanding, evaluating, and reviewing the outputs of smart tools.

Monitor international developments in AI legislation and intellectual property, continuously aligning internal policies accordingly.

Conclusion

The employment of artificial intelligence applications in the Department of the Jordanian National Library should not be viewed as a mere technical upgrade to keep pace with digital trends, but rather as a strategic institutional transformation where operational efficiency intersects with the national mission. The library that bears the responsibility of safeguarding Jordan's memory and the memory of future generations is called upon today, more than ever, to be at the forefront of national institutions handling this transformation with seriousness, depth, and a critical awareness.

What makes this phase a real opportunity rather than a threat is that AI does not eliminate the value of accumulated human expertise in library and archival institutions. Instead, it liberates that expertise to focus on what it is most worthy of: verification, documentation, professional judgment, and academic debate. However, the success of this transition is ultimately conditional upon a firm institutional will, meticulous planning, clear policies, and an equipped, qualified library professional who possesses the capacity to evaluate, guide, and correct the technology's path whenever it deviates.

References

Al-Zoubi, Muhammad. (2022). "Artificial Intelligence in Arabic Libraries: Reality and Expectations." Journal of Arabic Libraries and Information, 42(3), 15-38.

Cox, A. M., & Mazumdar, S. (2022). Defining artificial intelligence for librarians. Journal of Librarianship and Information Science, 54(4), 1062–1077. https://doi.org/10.1177/09610006211068247 (accessed 17.03.2026).

Department of the Jordanian National Library. (2024). Annual Report of the Department. Amman: Department of the Jordanian National Library. Retrieved from: https://www.nl.gov.jo (accessed 17.03.2026).

Hassan, Fatima Ali. (2021). "Employing Natural Language Processing in Automated Cataloguing of Arabic Content: An Analytical Study." Jordanian Journal of Libraries and Information, 18(1), 55-79.

IFLA Journal. (2025). Special Issue: Artificial Intelligence–Transforming Global Librarianship. IFLA Journal, 51(1). Retrieved from https://journals.sagepub.com/loi/ifl (accessed 17.03.2026).

International Federation of Library Associations and Institutions (IFLA). (2023). IFLA AI Entry Point for Libraries. The Hague: IFLA. Retrieved from https://www.ifla.org/ai (accessed 17.03.2026).

Machackova, A., & Drobíková, B. (2023). Implementing AI in national libraries: Challenges and opportunities for cataloguing. Library Hi Tech, 41(2), 480-496. https://doi.org/10.1108/LHT-06-2022-0279 (accessed 17.03.2026).

Manyika, J., et al. (2023). The economic potential of generative AI. McKinsey Global Institute.

National Library of Israel. (2024). Bina: AI-powered visual search tool. Retrieved from https://www.nli.org.il (accessed 17.03.2026).

Supreme Council for Information and Communications Technology – Jordan. (2023). National Digital Transformation Strategy 2023-2026. Amman.

UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000381137 (accessed 17.03.2026).

World Intellectual Property Organization (WIPO). (2024). AI Infrastructure Interchange Initiative (AIII). Geneva: WIPO. Retrieved from https://www.wipo.int/about-ip/en/artificial_intelligence (accessed 17.03.2026).

Yarrow, A., Clubb, B., & Draper, J. L. (2008). Public libraries, archives and museums: Trends in collaboration and cooperation. IFLA Professional Report No. 108$. The Hague: IFLA.

ЦИФРОВАЯ ЭКОСИСТЕМА АРХИВОВ УЗБЕКИСТАНА: ИНФРАСТРУКТУРА, ИИ И ГОСУДАРСТВЕННЫЕ СЕРВИСЫ

Улугбек Мухаммадиевич Юсупов
Директор Агентства «Узархив»
при Министерстве юстиции Республики Узбекистан
u.yusupov1985@gmail.com, https://orcid.org/0009-0005-8368-9644

Аннотация. В статье рассматриваются основные направления цифровой трансформации архивной сферы Республики Узбекистан. Особое внимание уделяется централизованной системе управления государственными архивами, внедрению единых информационных систем, развитию электронных государственных услуг и интеграции архивных информационных ресурсов с системами других государственных органов. Отмечается, что в настоящее время гражданам созданы условия для получения архивных услуг в онлайн-формате, включая получение электронных копий документов без личного посещения архива.

В статье также раскрываются ключевые положения Программы развития архивной сферы на 2026–2030 годы, утвержденной Президентом Республики Узбекистан. В рамках данной Программы предусмотрены масштабная цифровизация Национального архивного фонда, создание Единой национальной информационной системы, внедрение RFID-технологий, применение искусственного интеллекта при формировании метаданных, запуск новых электронных услуг, развитие информационной безопасности и повышение квалификации архивных работников. Кроме того, рассматриваются перспективные проекты, направленные на развитие генеалогических сервисов и укрепление нормативно-правовой и методической базы архивного дела.

Ключевые слова: архивное дело; цифровизация; электронные услуги; государственные архивы; Национальный архивный фонд; метаданные; искусственный интеллект; RFID; информационная безопасность; Узбекистан

ӨЗБЕКСТАН АРХИВТЕРІНІҢ ЦИФРЛЫҚ ЭКОЖҮЙЕСІ: ИНФРАҚҰРЫЛЫМ, ЖИ ЖӘНЕ МЕМЛЕКЕТТІК СЕРВИСТЕР

Ұлықбек Мұхаммадұлы Юсупов
Өзбекстан Республикасы Әділет министрлігі жанындағы
«Өзархив» агенттігінің директоры
u.yusupov1985@gmail.com, https://orcid.org/0009-0005-8368-9644

Аңдатпа. Мақалада Өзбекстан Республикасының архив саласындағы цифрлық трансформацияның негізгі бағыттары қарастырылады. Мемлекеттік архивтерді басқарудың орталықтандырылған жүйесін қалыптастыруға, бірыңғай ақпараттық жүйелерді енгізуге, электрондық мемлекеттік қызметтерді дамытуға және архивтік ақпараттық ресурстарды басқа да мемлекеттік органдардың ақпараттық жүйелерімен интеграциялауға ерекше назар аударылады. Қазіргі уақытта азаматтарға архив қызметтерін онлайн форматта алу үшін жағдай жасалғаны, оның ішінде архивке жеке бармай-ақ құжаттардың электрондық көшірмелерін алу мүмкіндігі қамтамасыз етілгені атап өтіледі.

Мақалада Өзбекстан Республикасы Президенті бекіткен 2026–2030 жылдарға арналған архив саласын дамыту бағдарламасының негізгі бағыттары ашып көрсетіледі. Бағдарлама аясында Ұлттық архив қорын ауқымды цифрландыру, Бірыңғай ұлттық ақпараттық жүйені құру, RFID-технологияларын енгізу, метадеректерді қалыптастыру кезінде жасанды интеллектті қолдану, жаңа электрондық қызметтерді іске қосу, ақпараттық қауіпсіздікті дамыту және архив қызметкерлерінің біліктілігін арттыру көзделген. Сонымен қатар мақалада генеалогиялық сервистерді дамытуға, архив ісінің нормативтік-құқықтық және әдістемелік базасын жетілдіруге бағытталған перспективалық жобалар қарастырылады.

Түйін сөздер: архив ісі; цифрландыру; электрондық қызметтер; мемлекеттік архивтер; Ұлттық архив қоры; метадеректер; жасанды интеллект; RFID; ақпараттық қауіпсіздік; Өзбекстан

DIGITAL ECOSYSTEM OF UZBEKISTAN'S ARCHIVES: INFRASTRUCTURE, AI, AND PUBLIC SERVICES

Ulugbek M. Yusupov
Director of the Agency "Uzarchive"
under the Ministry of Justice of the Republic of Uzbekistan
u.yusupov1985@gmail.com, https://orcid.org/0009-0005-8368-9644

Abstract. The article examines the main directions of digital transformation in the archival sector of the Republic of Uzbekistan. Special attention is paid to the centralized system of public archive management, the introduction of unified information systems, the development of electronic public services, and the integration of archival information resources with the systems of other government bodies. The article notes that citizens are now able to use archival services online, including receiving electronic copies of archival documents without visiting archives in person. The paper also outlines the key provisions of the Archival Sector Development Program for 2026–2030, approved by the President of the Republic of Uzbekistan. The Program provides for large-scale digitization of the National Archival Fund, the creation of a Unified National Archival Information System, the introduction of RFID technologies, the use of artificial intelligence in metadata generation, the launch of

new electronic services, the development of information security, and the improvement of archivists' professional skills. In addition, the article discusses promising projects aimed at developing genealogical services and strengthening the legal and methodological framework of archival work.

Keywords: archival work; digitization; electronic services; public archives; National Archival Fund; metadata; artificial intelligence; RFID; information security; Uzbekistan

Введение

В современных условиях цифровая трансформация архивной сферы является одним из важных направлений развития государственного управления, обеспечения сохранности документов и расширения доступа граждан к архивной информации. Архивы уже не ограничиваются традиционным хранением документов. Они становятся активными участниками цифровой экосистемы государства, обеспечивая электронное взаимодействие с гражданами, организациями и государственными органами.

Архивная система Республики Узбекистан имеет ряд особенностей. Все государственные архивы страны, включая районные и городские государственные архивы в регионах, входят в единую систему и подчиняются Агентству «Узархив» при Министерстве юстиции Республики Узбекистан. В настоящее время в системе насчитывается 104 государственных архива, из которых три являются национальными архивами. Такая централизованная модель управления позволяет проводить единую государственную политику в сфере архивного дела, внедрять общие цифровые решения и обеспечивать методическое сопровождение всех архивных учреждений.

Единая цифровая инфраструктура государственных архивов

Одним из ключевых результатов реформ последних лет стало объединение государственных архивов в единую информационную сеть. Все государственные архивы республики подключены к единым информационным системам, через которые предоставляются цифровые архивные услуги. Это позволяет обеспечить единый подход к приему, обработке и исполнению запросов, а также повысить прозрачность и оперативность архивного обслуживания.

Гражданам предоставлена возможность получать услуги государственных архивов в онлайн-формате посредством Единого портала интерактивных государственных услуг, а также через центры государственных услуг. Такой подход особенно важен для граждан, проживающих в отдаленных районах, а также для лиц, которым необходимо получить архивные сведения без личного посещения архива.

Информационные системы архивной сферы интегрированы с информационными системами других организаций. Обмен данными с государственными органами осуществляется посредством межведомственных информационных систем. Наибольший объем обмена данными приходится на взаимодействие с Пенсионным фондом. При оформлении пенсии уполномоченным организациям предоставлена возможность получать необходимые архивные сведения в онлайн-режиме. Это способствует сокращению бюрократических процедур, избавляет граждан от необходимости собирать дополнительные документы и обращаться в различные инстанции.

Развитие электронных архивных услуг

В настоящее время наиболее востребованными среди граждан являются десять архивных услуг, и все они полностью цифровизированы. Гражданам создана возможность получать электронные копии интересующих их документов без личного посещения архива. Это является важным шагом к повышению доступности архивной информации и улучшению качества государственных услуг.

Цифровизация услуг также оказывает положительное влияние на внутренние рабочие процессы архивов. Электронный прием запросов, автоматизированная обработка данных, межведомственная интеграция и возможность дистанционного предоставления результатов позволяют существенно сократить сроки исполнения запросов и повысить эффективность работы архивных учреждений.

При этом полноценная цифровизация архивных услуг невозможна только за счет сканирования документов. Для эффективного поиска, идентификации, учета и использования цифровых копий необходимо формирование качественных метаданных. Именно метаданные обеспечивают связь между цифровым образом документа, его содержанием, происхождением, фондовой принадлежностью и условиями использования.

Программа развития архивной сферы на 2026–2030 годы

В ноябре 2025 года Президентом Республики Узбекистан была утверждена Программа развития архивной сферы на 2026–2030 годы, в которой определены стратегические направления развития отрасли на последующие пять лет (Программа развития архивной сферы 2025).

В данной Программе предусмотрены 26 основных задач, направленных на развитие цифровой инфраструктуры и информационных систем архивной сферы, совершенствование материально-технической базы, повышение знаний и навыков работников архивов, а также развитие нормативно-правовой и методической базы. Для реализации этих задач определены 87 конкретных механизмов.

Одним из ключевых индикаторных показателей Программы является цифровизация 60% Национального архивного фонда к 2030 году. Кроме того, поставлена цель довести количество оцифрованных документов до более чем одного миллиарда листов. Наряду с этим предусматривается внесение в информационные системы более 12 млн записей метаданных. Также планируется полностью цифровизировать еще четыре архивные услуги.

Создание Единой национальной архивной информационной системы

В целях дальнейшего развития цифровой инфраструктуры предусматривается создание Единой национальной архивной информационной системы на базе трех действующих информационных систем. Такая система должна обеспечить более эффективное управление архивными процессами, унификацию учета документов, повышение качества электронных услуг и расширение возможностей межведомственного обмена данными.

Кроме того, в целях обеспечения безопасного хранения архивных документов и автоматизации их учета предусмотрено внедрение RFID-технологий. Использование таких технологий позволит повысить точность учета документов, усилить контроль за их движением и минимизировать риски, связанные с человеческим фактором.

Программа также предусматривает полную цифровизацию шести функций государственных архивов, включая регистрацию документов Национального архивного фонда, согласование описей дел и другие процессы согласования. Это позволит перевести важные административные и профессиональные процедуры в электронный формат и обеспечить их прозрачность.

Применение искусственного интеллекта и развитие новых сервисов

Одним из перспективных направлений является применение технологий искусственного интеллекта в рабочих процессах государственных архивов. В настоящее время такие технологии внедряются для формирования метаданных по отсканированным документам. Тестовые процессы уже начаты, и в дальнейшем создание метаданных будет осуществляться с применением инструментов искусственного интеллекта.

Использование искусственного интеллекта может значительно ускорить обработку больших массивов цифровых документов, повысить качество поиска, облегчить тематическую классификацию и расширить возможности доступа к архивной информации. Вместе с тем внедрение таких технологий требует методического контроля, проверки достоверности результатов и соблюдения требований информационной безопасности.

Отдельного внимания заслуживает проект «Шажара», направленный на развитие генеалогических сервисов. В августе 2026 года планируется запуск отдельного портала по созданию семейного родословного древа. Через данный портал граждане смогут создавать свои семейные деревья, а также направлять запросы в государственные архивы. Данный проект имеет большое общественное значение, поскольку способствует укреплению интереса граждан к истории своей семьи, родословной и национальной памяти.

В настоящее время также начата работа по цифровизации и запуску таких услуг, как составление семейного родословного древа, экспертиза документов, а также проведение аудита в сфере архивного дела. В ближайшие годы оказание данных услуг также будет осуществляться в онлайн-формате.

Информационная безопасность и материально-техническое развитие

Цифровая трансформация архивной сферы невозможна без укрепления информационной безопасности. Одной из важных задач стало внедрение Операционного центра информационной безопасности. С помощью данной системы обеспечивается постоянный мониторинг угроз, направленных на информационные системы, раннее выявление кибератак, а также своевременное обнаружение уязвимостей.

В настоящее время данная система полностью введена в эксплуатацию. Ее функционирование имеет особое значение, поскольку архивные информационные системы содержат большие объемы социально значимых данных и цифровых копий документов, требующих надежной защиты.

Наряду с цифровой инфраструктурой важное значение имеет укрепление материально-технической базы архивных учреждений. Оснащение архивов современным оборудованием, средствами сканирования, системами хранения данных и средствами защиты информации является необходимым условием для достижения целевых показателей цифровизации.

Повышение квалификации архивных работников

Цифровая трансформация требует не только технической модернизации, но и подготовки квалифицированных кадров. В Программе развития архивной сферы предусмотрен ряд задач, направленных на повышение профессиональных знаний и навыков работников архивов.

Одной из таких задач стало создание лабораторий «Электронный архив» в высших образовательных учреждениях, осуществляющих подготовку специалистов в сфере архивного дела. Такие лаборатории уже открыты в двух образовательных учреждениях. В них студенты могут работать со всеми информационными системами, используемыми в государственных архивах, а также выполнять процессы, связанные с цифровизацией документов.

Кроме того, в 2026 году предусмотрено повышение квалификации более 300 работников архивов в области информационных технологий. Также планируется проведение исследовательских работ в зарубежных архивах. Это позволит изучить международный опыт, внедрить современные подходы и повысить качество профессиональной подготовки архивистов.

Развитие нормативно-правовой и методической базы

Развитие архивной сферы требует постоянного совершенствования нормативно-правовой и методической базы. В рамках Программы ожидается утверждение в общей сложности шести нормативно-правовых актов. В частности, предусматривается принятие законов об установлении Дня работников архивов, а также о введении почетного звания «Заслуженный архивист Узбекистана».

В настоящее время проекты данных законов разработаны и внесены в Кабинет Министров. Их принятие будет способствовать повышению статуса профессии архивиста, укреплению кадрового потенциала и признанию вклада работников архивной сферы в сохранение исторической памяти страны.

Заключение

Опыт Республики Узбекистан показывает, что цифровая трансформация архивной сферы должна осуществляться комплексно. Она включает не только сканирование документов, но и развитие информационных систем, создание метаданных, внедрение новых технологий, интеграцию с государственными органами, обеспечение информационной безопасности, повышение квалификации кадров и совершенствование нормативно-правовой базы.

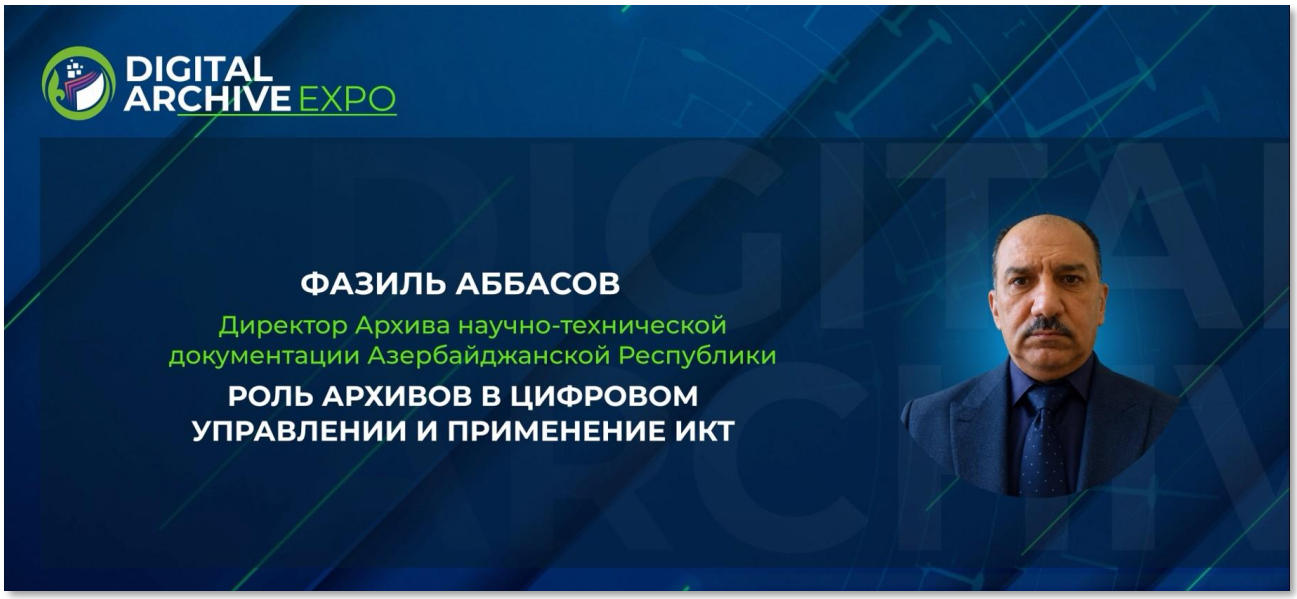
Реализация Программы развития архивной сферы на 2026–2030 годы позволит значительно расширить доступ граждан к архивной информации, повысить эффективность деятельности государственных архивов и обеспечить сохранность документов Национального архивного фонда в цифровой среде. В этом процессе архивы становятся не только хранителями исторической памяти, но и важными участниками цифрового развития государства.

Литература

Агентство «Узархив» 2026 – Агентство «Узархив» при Министерстве юстиции Республики Узбекистан. Официальный веб-сайт. https://archive.uz (дата обращения: 20.05.2026).

Закон Республики Узбекистан 2010 – Закон Республики Узбекистан «Об архивном деле». 15 июня 2010 года. Национальная база данных законодательства Республики Узбекистан.

Программа развития архивной сферы 2025 – Программа развития архивной сферы Республики Узбекистан на 2026–2030 годы, утвержденная Президентом Республики Узбекистан в ноябре 2025 года.



РОЛЬ АРХИВОВ В ЦИФРОВОМ УПРАВЛЕНИИ И ПРИМЕНЕНИЕ ИКТ

Фазиль Расимович Аббасов
Директор Архива научно-технической документации Азербайджанской Республики, заместитель начальника Национального архивного управления Азербайджанской Республики
ardetsa@milliarxiv.gov.az

Аннотация. В статье рассматривается влияние информационно-коммуникационных технологий на государственное управление и трансформацию архивного дела в Азербайджане. Автор анализирует нормативно-правовую базу, включая указы Президента и государственные программы, направленные на переход к «цифровому правительству». Особое внимание уделено изменению роли архивов, которые становятся активными участниками электронного документооборота с момента создания документа. Описываются практические результаты цифровизации: создание Единой электронной архивной информационной системы, оцифровка миллионов листов документов и интеграция с облачными платформами.

Ключевые слова: цифровое управление; информационно-коммуникационные технологии; электронный документооборот; архивное дело; Азербайджан; электронное правительство; оцифровка; национальный архивный фонд; информационная безопасность; интеграция систем; инновации; государственные услуги

ЦИФРЛЫҚ БАСҚАРУДАҒЫ АРХИВТЕРДІҢ РӨЛІ ЖӘНЕ АКТ ҚОЛДАНУ

Фазиль Расимұлы Аббасов
Әзербайжан Республикасының Ғылыми-техникалық құжаттама архивінің директоры, Ұлттық архив басқармасы басшысының орынбасары
ardetsa@milliarxiv.gov.a

Аңдатпа. Мақалада ақпараттық-коммуникациялық технологиялардың мемлекеттік басқаруға ықпалы және Әзербайжандағы архив ісінің трансформациясы қарастырылады. Автор «цифрлық үкіметке» көшуге бағытталған Президент жарлықтары мен мемлекеттік бағдарламаларды қоса алғанда, нормативтік-құқықтық базаны талдайды. Құжат жасалған сәттен бастап электрондық құжат айналымы үдерісінің белсенді қатысушысына айналып келе жатқан архивтердің рөлінің өзгеруіне ерекше назар аударылады. Мақалада цифрландырудың практикалық нәтижелері сипатталады: бірыңғай электрондық архивтік ақпараттық жүйені құру, миллиондаған құжат парақтарын цифрлық форматқа көшіру және бұлтты платформалармен интеграциялау үдерістері көрсетіледі.

Түйін сөздер: цифрлық басқару; ақпараттық-коммуникациялық технологиялар; электрондық құжат айналымы; архив ісі; Әзербайжан; электрондық үкімет; цифрландыру; Ұлттық архив қоры; ақпараттық қауіпсіздік; жүйелер интеграциясы; инновациялар; мемлекеттік қызметтер

THE ROLE OF ARCHIVES IN DIGITAL GOVERNANCE AND THE USE OF ICT

Fazil R. Abbasov
Director of the Archive of Scientific and Technical Documentation of the Republic of Azerbaijan, Deputy Head of the National Archival Administration of the Republic of Azerbaijan
ardetsa@milliarxiv.gov.az

Abstract. The article examines the impact of information and communication technologies on public administration and the transformation of archival affairs in Azerbaijan. The author analyzes the regulatory framework, including presidential decrees and state programs aimed at the transition to a “digital government”. Particular attention is paid to the changing role of archives, which are becoming active participants in electronic document management from the moment a document is created. The practical results of digitalization are described: the creation of the Unified Electronic Archive Information System, the digitalization of millions of sheets of documents, and integration with cloud platforms.

Keywords: digital governance; information and communication technologies; electronic document management; archival affairs; Azerbaijan; e-government; digitalization; national archival fund; information security; system integration; innovation; public services

Уважаемые коллеги, участники мероприятия!

От себя лично и от имени коллектива государственных архивов Азербайджанской Республики сердечно приветствую вас и выражаю искреннюю благодарность организаторам мероприятия, а также руководителю Архива Президента Республики Казахстан госпоже Алие Мустафиной за приглашение принять участие в столь значимом профессиональном форуме.

XXI век по праву называют веком информации. В современных условиях информационно-коммуникационные технологии стали одним из ключевых факторов развития государственного управления, экономики и общественных институтов. Для Азербайджана, как и для всех развивающихся стран, цифровое развитие является одним из приоритетных направлений государственной политики.

Основой этой политики стала утвержденная еще в 2003 году общенациональным лидером Гейдаром Алиевым «Национальная стратегия по информационно-коммуникационным технологиям для развития Азербайджанской Республики». Именно она заложила фундамент для дальнейшего формирования электронного правительства, развития цифровых сервисов и интеграции Азербайджана в мировое электронное пространство. В рамках реализации стратегии были приняты многочисленные государственные программы по развитию различных сфер, а также сформирована нормативно-правовая база, обеспечивающая развитие данной области в соответствии с международными стандартами.

За более чем двадцатилетний период в Азербайджане были реализованы масштабные государственные программы и проведены институциональные реформы. Все это позволило значительно расширить электронные государственные услуги, повысить прозрачность деятельности государственных органов и укрепить информационную безопасность.

Важным этапом стало подписание Указа Президента Азербайджанской Республики от 27 октября 2018 года «О мероприятиях по электронному управлению внутренними процессами в государственных органах». Этот документ определил новые приоритеты цифрового управления и перехода к электронному сопровождению внутренних процессов – делопроизводства, контроля исполнения документов, архивного дела, кадрового и финансового учета.

В феврале текущего года был утвержден «План действий по ускорению цифрового развития Азербайджанской Республики на 2026–2028 годы». В нем

особое внимание уделено цифровизации государственных услуг, развитию инновационной среды, поддержке научных исследований и внедрению технологий искусственного интеллекта. Одной из ключевых задач остается переход к принципам «безбумажного правительства».

В этих условиях существенно возрастает роль архивов и специалистов, работающих в сфере управления документами.

Как известно, основой управленческой деятельности является информация, а процесс управления включает ее сбор, обработку и передачу. Информация, возникающая в процессе управления, при наличии необходимых реквизитов превращается в документ независимо от носителя, на котором она зафиксирована. Документ как носитель социальной информации создается обществом для практического использования и сохранения исторической памяти. Основным критерием эффективности управления является своевременное получение достоверной информации и принятие на ее основе соответствующих решений с их последующим документированием.

Развитие общества объективно сопровождается увеличением объема информации и документов, являющихся ее носителями. Документы, создаваемые в процессе управления, отражают социальные реалии общества и формируют представление о происходящих процессах. Являясь носителями социальной памяти, они сохраняются как часть культурно-исторического наследия народа и организуются для научного использования, когда обществу становится необходима содержащаяся в них информация (Ларин 1998).

Сегодня управленческие документы составляют основную часть Национального архивного фонда Азербайджанской Республики. Их ценность определяется не количеством, а достоверностью содержащейся информации, способностью фиксировать события общественной жизни и сохранять их для будущих поколений.

Переход к цифровому управлению меняет сам подход к работе с документами. Если раньше документ прежде всего воспринимался как материальный объект, то сегодня на первый план выходит сама информация, ее сохранность, доступность и возможность многократного использования.

Современные технологии позволяют обеспечивать оперативный поиск документов, быстрый доступ к информации и ее сохранение без потери качества. Вместе с тем цифровая среда предъявляет новые требования к архивному хранению, информационной безопасности и устойчивости электронных систем.

Электронный документ сегодня существует не только как носитель информации, но и как средство коммуникации. Поэтому архивные службы уже не могут участвовать в процессе только на завершающем этапе хранения документов.

Таким образом, современный архив перестает быть исключительно местом хранения документов. Архивные службы становятся полноценными участниками электронного документооборота – от создания документа до его долговременного архивного хранения.

Новые технологические условия требуют участия архивистов в разработке информационных систем, определении стандартов хранения, проведении экспертизы ценности документов и обеспечении их долгосрочной доступности.

Внедрение новых технологических систем и переход к цифровому управлению требуют повышения профессионализма сотрудников, работающих с документами. Особое значение приобретают экспертиза ценности документов, отбор достоверной информации и оптимизация архивного фонда. Важное значение имеют рациональное распределение документов, а также обеспечение их эффективного хранения и использования.

Электронная обработка внутренних управленческих процессов государственных органов, переход на электронный документооборот и интеграция систем межведомственного электронного документооборота требуют не только организации процессов управления и архивирования документов на новых технологических основах, но и унификации управленческих документов, установления единых стандартов и норм для управления информационными ресурсами и системами. Структура информационных систем и информационных ресурсов государственных органов должна обеспечивать возможность повторного использования информации после передачи документов на государственное хранение.

В Указе Президента Азербайджанской Республики от 14 марта 2018 года «О развитии электронного правительства и мерах по переходу к цифровому правительству» поставлена задача по формированию, ведению, интеграции и архивированию государственных информационных ресурсов и систем. Такой системный подход к управлению документами, подтвержденный международным стандартом ISO 15489, обеспечивает надежное хранение документов и возможность их использования в качестве подтверждения деятельности организаций.

Внедрение информационно-коммуникационных технологий и переход к электронному документообороту позволяют решать принципиально новые задачи в сфере управления документами, которые ранее невозможно было эффективно реализовать с помощью традиционных технологий. Электронные документы по своему содержанию и функциям тесно связаны с бумажными документами, образуя единый информационный комплекс. Однако подходы к их управлению в цифровую эпоху существенно меняются.

Если в традиционной системе документ рассматривался прежде всего как материальный объект, то в современных условиях основное внимание уделяется самой информации, ее сохранности, доступности и возможности многократного использования.

Переход к цифровому управлению обеспечивает оперативный поиск информации, быстрый доступ к документам и возможность многократного воспроизведения данных без потери качества. Вместе с тем цифровая среда предъявляет новые требования к архивному хранению, устойчивости информационных систем и обеспечению информационной безопасности.

Все это свидетельствует о том, что электронный документ, формирующийся в современной информационной среде, выполняет двойную функцию – является одновременно хранителем информации и средством коммуникации.

Таким образом, внедрение цифрового управления и переход к электронному документообороту приводят к существенному изменению роли архивов и специалистов, работающих в сфере управления документами. Сегодня архивные

службы становятся активными участниками процессов (Гасанов 2018) формирования, систематизации, хранения и обеспечения доступности информации на всех этапах ее жизненного цикла.

Новые технологические системы требуют применения современных подходов к оценке, описанию и хранению документов уже на стадии разработки систем управления документами, а не только на завершающем этапе их жизненного цикла. В этих условиях участие архивных служб становится необходимым на всех этапах работы с документами – от их создания до организации долговременного архивного хранения.

Переход к цифровому управлению существенно меняет роль архивов. Сегодня архивные службы перестают быть исключительно хранилищами документов и становятся полноценными участниками системы электронного документооборота – от создания документа до его долговременного архивного хранения. При этом необходимо учитывать, что законодательство в области архивного дела предусматривает передачу документов на государственное хранение только после их длительного хранения в организациях. В связи с этим особое значение приобретают вопросы устаревания форматов электронных документов, сложности их повторного копирования, а также связанные с этим экономические затраты.

В современных условиях перед государственными архивными службами стоят важные задачи по управлению электронными документами. К ним относятся участие в создании электронных систем, разработке нормативов, стандартов и процедур, обеспечивающих формирование, хранение и доступность документов архивного фонда; проведение экспертизы ценности электронных документов; осуществление интеллектуального контроля за электронными архивными ресурсами; а также определение требований к приему электронных документов на государственное хранение для обеспечения их долгосрочной читаемости и доступности (Расулов и др. 2019).

Распоряжением главы государства от 12 февраля 2020 года была утверждена «Государственная программа по развитию архивного дела в Азербайджанской Республике на 2020–2025 годы», которая определила важнейшие стратегические задачи в данной сфере. В рамках реализации Государственной программы предусмотрены модернизация архивного дела с применением современных информационно-коммуникационных технологий, оцифровка документов Национального архивного фонда, создание Единой электронной архивной информационной системы, ее интеграция с государственными информационными системами, а также перевод архивных услуг в электронный формат.

В настоящее время в республике успешно функционирует Единая электронная архивная информационная система. В нее внесены первичные сведения почти о 16 тыс. архивных фондов, находящихся на государственном хранении, и обеспечен онлайн-доступ к соответствующей информации (Национальное архивное управление). Продолжается перевод архивных описей и каталогов в электронный формат. Созданы цифровые копии более 10 млн листов архивных документов, интегрированных в систему.

Обеспечена интеграция электронной архивной информационной системы с платформами «Электронного правительства» и «Правительственного облака». Автоматизированы процессы учета и регистрации документов Национального архивного фонда, сформирован электронный реестр архивных фондов. За последние годы на государственное хранение принято свыше 14 тыс. аудиовизуальных электронных документов. Показательно, что уже 47% обращений граждан, связанных с предоставлением архивной информации, обрабатываются в электронном формате.

Постановлением Кабинета Министров Азербайджанской Республики от 9 февраля 2024 года № 80, которым утверждены «Правила приема на постоянное государственное хранение, обеспечения сохранности, использования электронных документов Национального архивного фонда и оцифровки архивных документов», сформирована современная нормативно-правовая база, регулирующая деятельность в данной сфере.

Следует особо подчеркнуть, что внедрение информационно-коммуникационных технологий в архивное дело не только повысило эффективность работы, но и существенно изменило профессиональный подход сотрудников, усилив их ответственность. В современных условиях вопросы обеспечения безопасности информационных ресурсов, сохранности и целостности данных, защиты от внешнего вмешательства, а также принятия эффективных мер кибербезопасности приобретают первостепенное значение.

Надежное хранение государственных информационных ресурсов и обеспечение их доступности для повторного использования являются одними из ключевых факторов устойчивого развития государства. Именно поэтому в условиях перехода к цифровому правительству особое значение приобретают систематизация документов, создаваемых в информационных системах, формирование электронных папок в соответствии со значимостью документов, обеспечение возможности их повторного использования, а также создание защищенных информационных систем, устойчивых к несанкционированному вмешательству (Правила 2024).

В условиях цифрового управления одной из важнейших задач архивистов становится сохранение информации, создаваемой в различных сферах общественной деятельности, как части исторической памяти современности, а также ее надежная передача будущим поколениям.

Архивы Азербайджана, обладая богатым историческим опытом, успешно адаптируются к требованиям цифровой эпохи. В условиях цифрового управления они продолжают выполнять свою главную миссию – сохранение исторической памяти и обеспечение преемственности документального наследия для будущих поколений.

Литература

Гасанов 2018 – Гасанов Г.А. Роль архивов в цифровом управлении // Республика. 2018. 25 ноября. С. 6.

Ларин 1998 – Ларин М.В. Управление документацией и новые информационные технологии. Москва, 1998.

Национальное архивное управление – Национальное архивное управление. www.millarxiv.gov.az (дата обращения: 20.05.2026).

Правила 2024 – Правила приема, хранения, использования электронных документов, входящих в Национальный архивный фонд, а также цифровизации архивных документов: Постановление Кабинета Министров Азербайджанской Республики № 80 от 9 февраля 2024 года.

Расулов и др. 2019 – Расулов Э.А., Гасанов Г.А. Архивное дело в Азербайджанской Республике. Баку, 2019.

NATIONAL ARCHIVES MANAGEMENT AND INTEGRATION: THE BAYS PROJECT

Muhammet Ahmet Tokdemir
President of Directorate of State Archives under the President of the Republic of Türkiye
Mahmet.Tokdemir@Devletarsivleri.Gov.Tr
Selis Yüksel Koyun
Directorate of State Archives under the President of the Republic of Türkiye, Assistant Specialist
Selisnur.Yukselkoyun@Devletarsivleri.Gov.Tr
Ömer Faruk Kurtuluş
Directorate of State Archives under the President of the Republic of Türkiye, Assistant Specialist
Ofaruk.Kurtulus@Devletarsivleri.Gov.Tr

Abstrsact. Technological advancements and the resulting changes in requirements have not only transformed document management processes in the public sector but have also necessitated a transformation in archival practices. In this context, an assessment of the information systems and technological infrastructure used in archival operations by the Directorate of State Archives of the Republic of Türkiye, in terms of current requirements, has revealed a need for an archival management system that ensures integration with the relevant institutions. This study examines the Integrated Archive Management System Project (BAYS), carried out in collaboration between the Directorate of the State Archives and the Secretariat of Defence Industries; it explores the project's origins, development, and intended outcomes.

Keywords: archive management system, digital preservation, digital continuity, data transfer

ҰЛТТЫҚ АРХИВТЕРДІ БАСҚАРУ ЖӘНЕ ИНТЕГРАЦИЯЛАУ: BAYS ЖОБАСЫ

Мұхаммет Ахмет Токдемир
Түркия Республикасы Президенті Әкімшілігі жанындағы
Мемлекеттік архивтер басқармасының президенті
Mahmet.Tokdemir@Devletarsivleri.Gov.Tr
Селис Юксель Коюн
Түркия Республикасы Президенті Әкімшілігі жанындағы
Мемлекеттік архивтер басқармасының маман көмекшісі
Selisnur.Yukselkoyun@Devletarsivleri.Gov.Tr
Өмер Фарук Куртулуш
Түркия Республикасы Президенті Әкімшілігі жанындағы
Мемлекеттік архивтер басқармасының маман көмекшісі
Ofaruk.Kurtulus@Devletarsivleri.Gov.Tr

Аңдатпа. Технологиялық жетістіктер және соған байланысты өзгеріп отырған талаптар мемлекеттік сектордағы құжаттарды басқару үдерістерін ғана өзгертіп қоймай, архив ісі тәжірибесінің де жаңғыруын қажет етті. Осы тұрғыда Түркия Республикасының Мемлекеттік архивтер басқармасы жүзеге асыратын архивтік үдерістерде қолданылатын ақпараттық жүйелер мен технологиялық инфрақұрылымға қазіргі талаптар тұрғысынан жүргізілген талдау тиісті мекемелермен интеграцияны қамтамасыз ететін архивтерді басқару жүйесін енгізу қажеттігін көрсетті. Зерттеуде Мемлекеттік архивтер басқармасы мен Қорғаныс өнеркәсібі жөніндегі хатшылықтың бірлескен жұмысы нәтижесінде жүзеге асырылып жатқан Интеграцияланған архивтерді басқару жүйесі (BAYS) жобасы қарастырылады. Мақалада жобаның пайда болу алғышарттары, даму кезеңдері және күтілетін нәтижелері талданады.

Түйін сөздер: архивтерді басқару жүйесі; цифрлық сақтау; цифрлық сабақтастық; деректерді тасымалдау

УПРАВЛЕНИЕ И ИНТЕГРАЦИЯ НАЦИОНАЛЬНЫХ АРХИВОВ: ПРОЕКТ BAYS

Мухаммет Ахмет Токдемир
Президент Управления государственных архивов
при Администрации Президента Турецкой Республики
Mahmet.Tokdemir@Devletarsivleri.Gov.Tr
Селис Юксель Коюн
Помощник специалиста Управления государственных архивов
при Администрации Президента Турецкой Республики
Selisnur.Yukselkoyun@Devletarsivleri.Gov.Tr
Омер Фарук Куртулуш
Помощник специалиста Управления государственных архивов
при Администрации Президента Турецкой Республики
Ofaruk.Kurtulus@Devletarsivleri.Gov.Tr

Аннотация. Технологические достижения и связанные с ними изменения требований не только преобразовали процессы управления документами в государственном секторе, но и обусловили необходимость трансформации архивной практики. В этом контексте оценка информационных систем и технологической инфраструктуры, используемых в архивных процессах Управлением государственных архивов Турецкой Республики, с точки зрения современных требований выявила необходимость создания системы управления архивами, обеспечивающей интеграцию с соответствующими учреждениями. В исследовании рассматривается Проект интегрированной системы управления архивами (BAYS), реализуемый в сотрудничестве между Управлением государственных архивов и Секретариатом оборонной промышленности. В статье анализируются предпосылки возникновения проекта, этапы его развития и ожидаемые результаты.

Ключевые слова: система управления архивами; цифровое хранение; цифровая преемственность; передача данных

Transformation in Archival Information Systems in the Context of Technological Advancements

The growing need for information and the ability to process large amounts of data quickly are also increasing the demand for and expectations of archives. Meeting these expectations through routine practices or individual efforts alone is insufficient. Consequently, the importance of conducting archival operations in accordance with our country's archival policies and professional standards, while utilizing technological capabilities, is growing steadily. In line with the policies adopted to reduce bureaucracy and streamline procedures, the shift of institutional operations –along with their data and documents – to electronic platforms has led to the conduct of most procedures through information systems. With the active and widespread use of integrated information systems, it is anticipated that a planned phase-out of document production will begin in the near future. In this context, archives and archivists must redefine their roles to encompass their competencies, technological capabilities, and relationships with other disciplines (Rukancı et al. 2021: VI).

This transformation process necessitates fundamental changes not only at the administrative and political levels but also in technical infrastructure and service delivery models. Indeed, technological advancements and the evolving user expectations that accompany them make it essential to modernize archives' technical services and IT infrastructure. Today, archives are no longer just institutions that store information and documents; they are defined as dynamic information centers that manage, provide access to, and aim to preserve information produced in the digital environment over the long term. Accordingly, the need has arisen to update existing archival information systems within the framework of archival practices and standards adopted at the national and international levels. Digital continuity is concerned with ensuring that information remains complete, accessible, and therefore usable in accordance with business needs (Mater 2023). In this context, archival institutions must ensure digital continuity through information systems they establish in accordance with technical and professional standards. Indeed, the literature emphasizes that when digitization and digital preservation practices in cultural memory institutions are not addressed within the framework of institutional policies, technical infrastructure, and sustainable planning – in line with long-term

preservation and digital continuity objectives – significant risks regarding permanence and access arise. (Çakmak et al. 2017).

The current rapid advancements in information and communication technologies have shifted archives away from the traditional “storage-centric” approach, transforming them into dynamic information centers that manage and provide access to information generated in digital environments. In the literature, this process is described as an “archival transformation” in which professional knowledge and concepts are integrated with technological capabilities. (Rukancı et al. 2021), By placing archival science – the “profession of standards” at the heart of this transformation, it emphasizes the need to structure the entire process, from the creation of a document to its disposal, within a comprehensive framework based on national and international standards such as TS 13298.

Integrated Archive Management System Project (BAYS)

As the national archive of the Republic of Türkiye, the Directorate of State Archives launched archive automation initiatives in 1993 to manage the operations of its main service units. The system transitioned to a web-based platform in 2004, and since 2005, existing data has been made available to researchers via the internet. However, due to changing requirements driven by technological advancements, the existing systems proved insufficient, and there was a need to enhance integration among systems, institutions, and activities..

With the transition to the Presidential Government System in Türkiye in 2018, the institutional status of the Directorate of the State Archives was redefined, making it an establishment affiliated with the Presidency As a result of this structural transformation, Presidential Decree No. 11 was issued, significantly expanding and diversifying the scope of duties and responsibilities of the Directorate of State Archives.

In this context, the Directorate of the State Archives is responsible for: establishing an information security policy; ensuring the flow of data, data sharing, and coordination between systems – both online and offline – at the national and international levels in matters falling within its scope of responsibility; monitor developments in information technology, collaborate with other public institutions and organizations regarding the use of information technology hardware and software, establish and coordinate the State Archival Network, and make information and documents available for research, to establish the procedures and principles regarding the transfer of archival documents in electronic format to be transferred to the Directorate of State Archives, to define the necessary document management and core archival functions in electronic document management and archiving systems, and to provide guidance and coordination for the implementation of innovative techniques, particularly microfilming and digitization activities..

Within the State Archives Directorate, various independent information systems had been developed and were in use since the 2000s to address different aspects of archival services. These included the Ottoman Archives Automation System, which enabled the cataloging of Ottoman-era archival documents and made them available for use in research rooms, the Document Search and Cataloging System, which facilitates the identification and classification of Republican-era archival documents as well as enabling web-based access to Ottoman documents, the Institutional Information System, which manages the transfer and handover processes of archival documents and conducts training, oversight, and document management activities for public institutions and organizations, and the Digitization System, where digitization efforts are carried out, stand out as key components..

However, the fact that these systems operate on different platforms and within a decentralized structure has highlighted the need for an integrated approach to data integrity, process management, and service delivery. Consequently, in line with the new duties and responsibilities assigned to the Directorate of State Archives under Presidential Decree No. 11, the goal is to ensure that the services provided to researchers and obligated public institutions are made more effective, faster, and sustainable.

In this context, as mentioned in the 2024 annual report of the Directorate of State Archives, the Integrated Archive Management System (BAYS) Project was developed to analyze and consolidate existing information systems from both software and hardware perspectives and to integrate them under a unified system architecture; it was implemented on September 24, 2019, following approval by the Directorate. The BAYS Project should be regarded as a strategic transformation tool designed to standardize institutional business processes, ensure data integrity, and strengthen user-centric service delivery by bringing together systems with different functions under a unified structure.

Between 2016 and 2019, as part of the National e-Government Strategy and Action Plan, domestic and open-source software was prioritized under the initiative to promote the use of open-source software (OSS) for Priority Information Systems (E2.2.3) The use of OSS also aims to reduce dependence on software vendors. Since the institutions required to transfer archival documents to the Directorate of State Archives include the Ministry of Foreign Affairs, the Ministry of Interior, the Ministry of

National Defense, and the General Staff all of which handle a high volume of classified documents – it was decided to implement the Integrated Archive Management System (BAYS) Project in collaboration with the Secretariat of Defence Industries.

The project has been planned in phases to facilitate project management and development. In the first phase, the goal is to achieve greater effectiveness and efficiency in critical areas such as data integrity and data quality by utilizing current technological infrastructure, adhering to international archival standards, and ensuring integration across institutions, services, and locations. In other phases, the project involves the transfer of electronic records created from the outset (which qualify as archival records) to the Directorate of the State Archives, as well as digital preservation and digital sustainability activities. Additionally, the integration of artificial intelligence and deep learning applications into the BAYS project to enhance services is a key objective in these phases.

The BAYS project adopted an Agile approach for software development and project management, with the goal of completing system development through continuous feedback. As part of the project, risk assessments were conducted, resource planning was carried out, and scope/cost evaluations were performed. Weekly project review meetings and demos were held throughout the project. “Integration,” which was also the starting point of the project, was prioritized, and integration tests to facilitate inter-organizational collaboration and coordination were planned.

Under the project, the work processes carried out by the Directorate of State Archives as part of its activities and services were reviewed, and a requirements analysis was conducted by creating workflow diagrams. As a result of the requirements analysis, it was decided that activities would be managed using modules within the BAYS system to ensure the effective delivery of services. The system consists of various functional modules, including Storage and Document Tracking, Classification Services, Digitization, Research Services, Restoration, Institutional Information System, and Publications and Promotion, and full integration has been achieved among these modules.

The BAYS Project aims to transform the institutional website of the Directorate of State Archives into a more functional, accessible, and user-centered platform in line with digital transformation goals. As part of this initiative, the website is slated to be redesigned with an approach that facilitates access to information, addresses diverse user needs, and enhances the efficiency of service delivery. Consequently, the goal is to present archival services in a more comprehensive and systematic manner within the digital environment. Throughout this process, the website is intended to be transformed into a modern service platform guided by fundamental principles such as enhancing the user experience, sustainability, and information security.

The analysis also revealed that there was no archival data dictionary available for defining archival metadata. In this context, the relevant process was initiated at the Directorate of State Archives as part of the National Data Dictionary (UVS) Project, which was launched in 2019 to compile a national data inventory and ensure terminological consistency. Additionally, as a result of the current status analysis conducted under the BAYS Project, technical terms were standardized and definitions were established to accurately address the entity space within the information system.

In this context, metadata used to describe archival records in relevant public institutions was requested. As a result of the work carried out, UVS was established as a national data standard for digital archiving in Turkey. Thanks to this infrastructure, the authority directories and dictionaries to be used in archival description will be accessible via BAYS. This capability offered by the system contributes to the development of the institutional administrative structure while also enabling researchers to quickly access accurate information through multifaceted queries.

The project's system architecture has been designed to utilize international archival description standards in the identification of archival records. Accordingly, the standards used in archival description processes are addressed below within a comprehensive framework. ISAAR (International Standard Archival Authority Record for Corporate Bodies, Persons, and Families) enables the detailed description of the organizational structures of archival holders, document relationships, and the types and characteristics of archival records, regardless of the methods by which these records were acquired or created. CAAIS (Canadian Archival Accession Information Standard) is used to standardize information related to the accession process in the description of archival records. ISDIAH (International Standard for Describing Institutions with Archival Holdings), on the other hand, provides a framework for describing public institutions, legal entities, and private individuals that hold archival records. The ISDF (International Standard for Describing Functions) standard ensures the definition of archival functions to enable a proper understanding of the archival creation process and, in particular, the provenance of the record. This standard enables the explanation of the archival context and the systematic tracking of the document's creation process by identifying the institutional functions and activities that led to the production of archival records.

In the process of making archival records accessible and preparing them for long-term preservation, it is necessary to identify and organize materials with varying content and physical characteristics. At this stage, records are structured and described using metadata standards, taking into account elements such as title, physical characteristics, date, and arrangement. In this context, internationally recognized standards such as ISAD(G) (International Standard for Archival Description (General), MARC21 (Machine Readable Cataloguing), EAD (Encoded Archival Description), Dublin Core (DC), and RAD (Rules for Archival Description) are utilized; thereby ensuring both the physical and intellectual control of archival materials and enhancing accessibility and interoperability.

Under the project, as part of institutional relations, the processing and management of physical and born-digital archival records to be transferred to the State Archives Directorate will be tracked through the BAYS system. A responsible public institution will report the inventory of the archive documents to be transferred through this system; the goal is for the transfer request for these documents to be evaluated by the relevant unit of the Directorate of the State Archives and approved. Additionally, regarding organizational and functional changes occurring within public institutions and organizations (such as name changes, mergers/separations, or the transfer of functions), the goal is to establish relationships through BAYS so that these connections can be tracked, thereby ensuring archival continuity at the institutional level.Obligees will be able to access the system as external users and submit requests

to the Directorate of State Archives via the system regarding audits, training, consultation and guidance, weeding and disposal lists, filing plans, and requests for the identification of archival records. Additionally, the system will facilitate the receipt, registration, and verification of documents received in physical form and the transfer of their digital copies within BAYS. Reports containing statistical data on the services provided to obligated entities can be retrieved through the system.

The digitization activities carried out under the BAYS project aim to manage the transfer of archival documents to a digital format through a comprehensive and auditable workflow. This process is initiated on local computers while adhering to the physical hierarchy of the documents (Box/File/Folder); the resulting digital files are transmitted to central servers via scheduled batch processes to minimize network load after being approved by the Digitization Control Operator. To ensure the sustainability of the technological infrastructure, system performance is optimized by implementing "converter server" support as needed for access copies (JPEG) in addition to the master copies in high-resolution TIFF format. Digitized documents can be saved in various formats (TIFF, JPEG, PNG, PDF/A) or converted later. To enhance the project's operational efficiency, traditional manual task assignments are being replaced by online task assignment capabilities. Additionally, through the "Unique Identifier" system used during the physical sorting phase, document metadata is automatically transferred to the system during scanning, aiming to standardize and accelerate data entry. By using the "reference code" – a mandatory field in metadata standards – digital images will be identified, and these identification numbers will be linked to archival identifiers. Digitization will be performed through a separate desktop/web application to ensure integration.

Due to the fact that current services are provided through different software systems and lack integration, resulting in the operation of distinct user management procedures for each system, it has been quite challenging to control role and authorization activities. In this context, the BAYS project aims to utilize the LDAP (Lightweight Directory Access Protocol) and Microsoft Authenticator infrastructure. This will simplify user access management and enable it to be managed through a single system.

The services to be provided to researchers will include a mobile-compatible version of BAYS. Researchers will be notified via email and SMS through the system. The research module will offer language support in English, French, and Arabic.

According to the 2024 Annual Report of the Directorate of State Archives, the analysis and software development processes for Versions 1, 2, 3, and 4 of Phase 1– developed as part of the Integrated Archive Management System (BAYS) Project– have been completed. Accordingly, the provisional acceptance process was carried out on October 3, 2025. Furthermore, testing, validation, and improvement efforts for the application's various modules are ongoing, with the goal of enhancing system performance and expanding its functional scope based on user feedback. As of now, the BAYS Project has completed a significant phase in its transformation into a scalable and sustainable archival information system infrastructure capable of managing institutional business processes within an integrated framework. In line with the 2024–2028 Strategic Plan of the Directorate of State Archives, the BAYS Project plans to provide training and course opportunities for software personnel to

effectively manage the process of standardizing software used in institutional operations; accordingly, the preparation of work plans related to these standardization activities and the systematic execution of implementation processes are targeted.

Conclusion

The digital transformation of today has transformed archival institutions from mere repositories of documents into dynamic centers that manage digital continuity and access to information. In this process of change, the Integrated Archival Management System (BAYS) Project, implemented by the Directorate of State Archives, represents a strategic step in the preservation and management of institutional memory using modern technologies. Designed with international archival classification standards in mind to address new technological needs, the BAYS Project has established an integrated system architecture in accordance with principles of integration and information security.

Furthermore, BAYS is not merely a technical transformation tool; it also stands out as an application that directly aligns with the objectives outlined in the Directorate of the State Archives' 2026 Annual Performance Program, namely digital transformation, the strengthening of information systems, and the provision of fast and secure access to archival documents through a single system. In this regard, the project serves as a concrete outcome of the archival policies established at the national level.

The data infrastructure established under Phase 1 has been designed to enable the effective management of electronic document transfer processes and the integration of AI-based applications into the system. The development of electronic document transfer processes using AI and OCR (Optical Character Recognition) technologies –which are planned for integration into the system in future phases – will optimize digital continuity not only in terms of preservation but also in terms of access to and use of information. As a result, the BAYS Project, with its modular structure, integration capabilities, user-friendly web interface, and robust technological infrastructure, constitutes a sustainable and comprehensive model for the preservation and management of Türkiye's digital heritage.

References

Devlet Arşivleri Başkanlığı Hakkında Cumhurbaşkanlığı Kararnamesi, T.C. Resmi Gazete (16 Temmuz 2018 – 30480).

Mater 2023 – Mater, E.C. (2023). Dijital Çağda Doğuştan Dijital Belgelerin Belge ve Arşiv Yönetim Süreçleri. Hiperyayın.

Rukancı et al. 2021 – Rukancı, F., Anameriç, H. and Başar, A. (2021). Arşiv ve arşivcilik: Kuram, strateji ve uygulamalar. T.C. Cumhurbaşkanlığı Devlet Arşivleri Başkanlığı.

Türkiye Cumhuriyeti Cumhurbaşkanlığı Devlet Arşivleri Başkanlığı 2024 Yılı Faaliyet Raporu, Ankara: Strateji Geliştirme Dairesi Başkanlığı, 2025.

Türkiye Cumhuriyeti Cumhurbaşkanlığı Devlet Arşivleri Başkanlığı 2024-2028 Dönemi Stratejik Planı, Ankara: Strateji Geliştirme Dairesi Başkanlığı, 2024.

Türkiye Cumhuriyeti Cumhurbaşkanlığı Devlet Arşivleri Başkanlığı. (2026). 2026 yılıPperformans Programı, Ankara: Strateji Geliştirme Dairesi Başkanlığı, 2026.

Çakmak et al. 2017 – Çakmak, T. and Yılmaz, B. (2017). Bellek Kurumlarında Dijitalleştirme ve Dijital Koruma: Türkiye'deki Uygulamaların Analizi. Bilgi Dünyası, 18(1), 49–91. https://doi.org/10.15612/BD.2017.580 (accessed 12.04.2026).

DIGITAL LONG-TERM RECORDS PRESERVATION AND MIGRATION STRATEGIES

Torsten Kux
Regional Manager, FamilySearch International, Germany
kuxto@familysearch.org

Abstract. Digital storage has significantly transformed the preservation of archival materials by providing compact, accessible, and cost-efficient solutions. Nevertheless, the long-term sustainability of digital data remains a major challenge due to media degradation, technological obsolescence, and data format incompatibility. This report analyzes the risks associated with digital preservation and outlines strategies to mitigate these risks effectively. Based on the example of FamilySearch as a global repository, this paper highlights the importance of redundancy, geographic distribution, format migration, and strong security measures. Furthermore, it explores the role of dedicated preservation facilities, such as the Granite Mountain Records Vault, which combine physical and digital safeguards to ensure durability and integrity. The report also addresses operational aspects including cost control, adherence to standards, and metadata management. The findings demonstrate that long-term digital preservation requires continuous adaptation, proactive planning, and collaboration across institutions to ensure that cultural and historical records remain accessible for future generations.

Keywords: Digital preservation; archival records; data migration; long-term storage; redundancy; format migration; digital archives; FamilySearch; risk mitigation; distributed systems

АРХИВТІК ҚҰЖАТТАРДЫ ҰЗАҚ МЕРЗІМДІ ЦИФРЛЫҚ САҚТАУ ЖӘНЕ КӨШІРУ СТРАТЕГИЯЛАРЫ

Торстен Кукс
Өңірлік менеджер, FamilySearch International, Германия
kuxto@familysearch.org

Аңдатпа. Цифрлық сақтау архивтік материалдарды сақтауда айтарлықтай өзгерістер әкеліп, ықшам, қолжетімді және шығынды азайтатын шешімдерді ұсынды. Дегенмен, цифрлық деректердің ұзақ мерзімді сақталуын қамтамасыз ету тасымалдаушы құралдардың тозуы, технологиялардың ескіруі және деректер форматтарының үйлесімсіздігі сияқты факторларға байланысты маңызды мәселелердің бірі болып қала береді. Мақалада цифрлық сақтауға байланысты тәуекелдер талданып, оларды тиімді төмендету стратегиялары қарастырылады. Әлемдік цифрлық архивтеу орталығы ретінде FamilySearch мысалы негізінде зерттеуде деректердің резервтік көшірмелерін жасау, оларды географиялық тұрғыдан әртүрлі орындарда сақтау, форматтарды көшіру және ақпараттық қауіпсіздіктің сенімді шараларын қамтамасыз етудің маңыздылығы көрсетіледі. Сонымен қатар ұзақ мерзімді сақтаудың тұрақтылығы мен тұтастығын қамтамасыз ету үшін физикалық және цифрлық қауіпсіздік тетіктерін үйлестіретін Granite Mountain Records Vault сияқты арнайы сақтау нысандарының рөлі қарастырылады. Баяндамада шығындарды бақылау, стандарттарды сақтау және метадеректерді басқару сияқты ұйымдастырушылық аспектілер де қамтылады. Зерттеу нәтижелері көрсеткендей, цифрлық мұраны ұзақ мерзімді сақтау үшін үздіксіз бейімделу, алдын ала жоспарлау және мекемелер арасындағы ынтымақтастық қажет. Бұл болашақ ұрпақ үшін мәдени және тарихи деректердің қолжетімділігін қамтамасыз етуге мүмкіндік береді.

Түйін сөздер: цифрлық сақтау; архивтік құжаттар; деректерді көшіру (миграциялау); ұзақ мерзімді сақтау; резервтік сақтау; форматтарды көшіру; цифрлық архивтер; FamilySearch; тәуекелдерді азайту; таратылған жүйелер

СТРАТЕГИИ ДОЛГОСРОЧНОГО ЦИФРОВОГО ХРАНЕНИЯ И МИГРАЦИИ АРХИВНЫХ ДОКУМЕНТОВ

Торстен Кукс
FamilySearch International ұйымының аймақтық менеджері, Германия
kuxto@familysearch.org

Аннотация. Цифровое хранение значительно изменило подходы к сохранению архивных материалов, предоставив компактные, доступные и экономически эффективные решения. Вместе с тем обеспечение долгосрочной сохранности цифровых данных остается одной из ключевых проблем в связи с деградацией носителей информации, устареванием технологий и несовместимостью форматов данных.

В настоящем докладе анализируются риски, связанные с цифровым сохранением, а также рассматриваются эффективные стратегии их минимизации. На примере FamilySearch как глобального хранилища цифровых данных в статье подчеркивается важность резервирования информации, ее географического распределения, миграции форматов и применения надежных мер информационной безопасности.

Кроме того, рассматривается роль специализированных объектов хранения, таких как Granite Mountain Records Vault, которые сочетают физические и цифровые меры защиты для обеспечения долговечности и целостности архивных данных.

В докладе также затрагиваются организационные аспекты, включая контроль затрат, соблюдение стандартов и управление метаданными. Результаты исследования показывают, что долгосрочное цифровое сохранение требует постоянной адаптации, упреждающего планирования и сотрудничества между учреждениями для обеспечения доступности культурных и исторических документов для будущих поколений.

Ключевые слова: цифровое сохранение; архивные документы; миграция данных; долгосрочное хранение; резервирование; миграция форматов; цифровые архивы; FamilySearch; снижение рисков; распределенные системы

Over time, data storage technologies have undergone significant changes and development. Initially, digital storage relied on media such as floppy disks, which were eventually replaced by hard drives, solid-state drives, and cloud-based solutions. Despite these technological advancements, the lifespan of many digital storage media remains limited, and long-term preservation must account for technological obsolescence, changing file formats, and evolving user preferences. The OAIS (Open Archival Information System) reference model identifies long-term preservation as a continuing organizational responsibility rather than a one-time storage activity (ISO 14721: 2025). Therefore, preserving digital data for centuries or even millennia require comprehensive strategies that go far beyond conventional storage approaches. FamilySearch has addressed this challenge by developing both physical and digital preservation systems, including highly secure facilities such as the Granite Mountain Records Vault.

Digital storage provides numerous advantages that make it highly desirable for archival purposes. It allows for compact storage, lossless duplication, and easy accessibility. Furthermore, digital data can be searched efficiently, maintaining high quality and readability while also enabling greater production capacities at lower costs. However, these benefits are accompanied by challenges, particularly concerning long-term preservation. The Digital Preservation Coalition emphasizes that digital preservation is an ongoing process requiring active management of storage systems, metadata, formats, and access mechanisms to ensure continued usability over time.

Digital data can be lost due to a variety of factors. Media failure is one of the most common causes, as storage devices deteriorate over time. Additionally, data may be lost through physical destruction or simple misplacement of storage media. Other risks include accidental or malicious deletion. Technological obsolescence poses a further challenge, as outdated hardware and software may no longer support existing data formats. According to the NDSA Levels of Digital Preservation framework, risks to digital content include media failure, integrity problems, inadequate security controls, and the inability to render obsolete file formats.

To ensure long-term preservation, organizations must adopt a combination of technical and organizational strategies. These include the use of reliable storage media and the creation of redundant copies stored across geographically separated sites. The NDSA Levels of Digital Preservation specifically recommend geographic redundancy, integrity monitoring, information security controls, and format management as core preservation practices. Similarly, the OAIS Reference Model identifies preservation planning, archival storage, data management, and access functions as essential components of a trustworthy digital archive. Automation in media handling and the establishment of dedicated preservation facilities further enhance data security. Regular validation of stored data helps to ensure integrity over time. In addition, organizations must plan for continuous media refresh and format migration to address technological change.

Beyond technical solutions, several additional aspects must be considered. Cost control plays an important role in ensuring the sustainability of preservation efforts. Maintaining the provenance and authenticity of data is also essential. The Library of Congress highlights the importance of preservation metadata, sustainable file formats, and documentation of relationships between digital objects to ensure future accessibility and authenticity. Organizations must adhere to established standards and keep data formats as simple as possible. Preservation programs must also balance technical requirements with organizational factors such as sustainability, governance, and cost management, as emphasized in the Digital Preservation Handbook.

FamilySearch operates one of the largest digital preservation systems worldwide. The organization ingests large volumes of data daily and maintains extensive archival collections across multiple locations. These records are preserved using a variety of methods, including digital capture, scanning of microfilm, and partner data ingestion. Storage is distributed across geographically separated sites, ensuring resilience and reliability. The Granite Mountain Records Vault serves as a central component of this system. The facility was established to preserve genealogical records of long-term value and safeguards billions of images on microfilm, microfiche, and digital media while supporting record recovery and digitization initiatives.

The security of preserved data is ensured through both physical and digital measures. Access to preservation facilities is strictly controlled, and these facilities are physically separated from other operational environments. On the digital level, data is protected through restricted network access, encryption, and robust authentication systems. In addition, distributed archive systems store copies in geographically separated locations, thereby reducing the risk of catastrophic loss caused by local disasters and infrastructure failures. Geographic distribution is a widely recognized preservation principle promoted by the NDSA framework.

FamilySearch offers considerable advantages to its partners. Organizations benefit from access to advanced preservation infrastructure without having to bear the associated costs. Furthermore, partners can obtain preservation copies of their data when needed. This significantly reduces the risks associated with long-term digital storage and supports sustainable archival practices.

Long-term digital preservation is a complex and ongoing process that requires continuous adaptation to technological change. By implementing strategies such as redundancy, geographic distribution, and format migration, organizations can significantly reduce the risk of data loss. International frameworks such as OAIS and the NDSA Levels of Digital Preservation provide valuable guidance for building trustworthy and sustainable preservation programs capable of maintaining access to digital records over extended periods. The example of FamilySearch demonstrates how large-scale preservation systems can ensure the long-term accessibility and security of archival data. Ultimately, successful digital preservation depends on a combination of technological solutions, strategic planning, and institutional cooperation.

References

Digital Preservation Coalition. (2015). Digital Preservation Handbook (2nd Edition). Available at: https://www.dpconline.org/handbook

International Organization for Standardization. (2025). ISO 14721:2025 Space Data System Practices—Reference Model for an Open Archival Information System (OAIS). Available at: http://www.oais.info/

Library of Congress. Digital Preservation at the Library of Congress. Available at: https://www.loc.gov/preservation/digital/

National Digital Stewardship Alliance. (2026). Levels of Digital Preservation Version 2.1. Available at: https://www.ndsa.org/publications/levels-of-digital-preservation/

The Church of Jesus Christ of Latter-day Saints. (2021). Granite Mountain Records Vault. Available at: https://news-bb.churchofjesuschrist.org/article/granite-mountain-records-vault

AN INTEGRATED SYSTEM FOR THE PERMANENT PROTECTION OF ARCHIVAL HERITAGE

Irene Ferrari
Head of Export Department, Makros SRL, Italy
export@blockfire.it

Abstract. The preservation of original archival heritage represents a fundamental responsibility linked to cultural identity and collective memory. In a context increasingly influenced by climate change and digital transformation, the protection of physical "original" materials – such as manuscripts and artistic sketches – requires a shift from passive containment to active environmental control.

This paper presents Bacteria Blockfire, an advanced evolution of the patented Blockfire® system, integrating mechatronics, predictive modelling, and molecular-based biological protection to ensure long-term conservation of cultural assets. The system is structured around three complementary technological components.

The ACS (Aerobiological Control System), developed in collaboration with the Italian National Research Council (CNR), applies a predictive approach based on thermo-hygrometric monitoring and species-specific models, including *Eurotium halophilicum*. This enables the system to anticipate fungal germination by approximately 11 days compared to conventional dataloggers.

At the structural level, the system incorporates a passive thermal barrier with natural micro-ventilation. Fire resistance tests conducted according to UNI EN 1363-1 (ISO 834) demonstrate that internal temperatures remain around 78°C even after 120 minutes of exposure to external temperatures of 1000°C, preserving the molecular stability of cellulose.

Finally, the system integrates carvacrol-based co-crystals that release volatile organic compounds (VOCs) capable of inhibiting fungal growth and repelling insects, without altering the chemical properties of the materials. These materials exhibit adaptive behaviour, increasing their activity under environmental conditions that favour biodeterioration.

By integrating environmental monitoring, predictive modelling, and material innovation, Bacteria Blockfire enables a transition from passive storage to active heritage asset management.

Keywords: preventive conservation; archival heritage; biodeterioration; fungal growth; isopleth model; Eurotium halophilicum; co-crystals; VOCs; predictive monitoring

АРХИВТІК МҰРАНЫҢ ТҰРАҚТЫ САҚТАЛУЫН ҚАМТАМАСЫЗ ЕТУДІҢ ИНТЕГРАЦИЯЛАНҒАН ЖҮЙЕСІ

Ирен Феррари
Makros SRL компаниясының экспорт бөлімінің басшысы, Италия
export@blockfire.it

Аңдатпа. Архивтік мұраның түпнұсқаларын сақтау мәдени бірегейлік пен ұжымдық тарихи жадпен байланысты маңызды міндеттердің бірі болып табылады. Климаттың өзгеруі мен цифрлық трансформация үдерістері күшейіп отырған қазіргі жағдайда қолжазбалар мен көркем эскиздер сияқты физикалық «түпнұсқа» материалдарды қорғау жай ғана сақтаудан қоршаған ортаны белсенді басқару тәсіліне көшуді талап етеді.

Мақалада мәдени құндылықтарды ұзақ мерзімді сақтау мақсатында мехатроника, болжамды модельдеу және молекулалық негіздегі биологиялық қорғау технологияларын біріктіретін, патенттелген Blockfire® жүйесінің жетілдірілген нұсқасы болып табылатын Bacteria Blockfire жүйесі ұсынылады. Жүйе өзара толықтыратын үш негізгі технологиялық компоненттен тұрады.

Италияның Ұлттық зерттеу кеңесімен (CNR) бірлесіп әзірленген ACS (аэробиологиялық бақылау жүйесі) термогигрометриялық мониторинг пен түрге тән модельдерге, соның ішінде Eurotium halophilicum микроорганизміне негізделген болжамды тәсілді қолданады. Бұл жүйе дәстүрлі деректер тіркеушілерімен салыстырғанда саңырауқұлақтардың өну қаупін шамамен 11 күн бұрын анықтауға мүмкіндік береді.

Құрылымдық деңгейде жүйеге табиғи микро-желдетуі бар пассивті жылу тосқауылы енгізілген. UNI EN 1363-1 (ISO 834) стандартына сәйкес жүргізілген өртке төзімділік сынақтары сыртқы температура 1000°C болған 120 минуттық әсерден кейін де ішкі температура шамамен 78°C деңгейінде сақталатынын көрсетеді. Бұл целлюлозаның молекулалық тұрақтылығын қорғауға мүмкіндік береді.

Сонымен қатар жүйеге карвакрол негізіндегі ко-кристалдар енгізілген, олар ұшпа органикалық қосылыстарды (VOC) бөлу арқылы материалдардың химиялық қасиеттерін өзгертпей, саңырауқұлақтардың дамуын тежейді және жәндіктерді қайтарады. Бұл материалдар қоршаған орта жағдайларына бейімделу қасиетіне ие, яғни биодеструкцияға қолайлы жағдайларда олардың белсенділігі артады.

Қоршаған ортаны мониторингтеу, болжамды модельдеу және материалдық инновацияларды біріктіру арқылы Bacteria Blockfire жүйесі пассивті сақтау тәсілінен архивтік мұраны белсенді басқару моделіне көшуге мүмкіндік береді.

Түйін сөздер: превентивті сақтау; архивтік мұра; биодеструкция; саңырауқұлақтардың дамуы; изоплет моделі; *Eurotium halophilicum*; сокристалдар; VOC (ұшпа органикалық қосылыстар); болжамды мониторинг

ИНТЕГРИРОВАННАЯ СИСТЕМА ОБЕСПЕЧЕНИЯ ПОСТОЯННОЙ СОХРАННОСТИ АРХИВНОГО НАСЛЕДИЯ

Ирен Феррари
Руководитель экспортного отдела компании Makros SRL, Италия
export@blockfire.it

Аннотация. Сохранение оригинального архивного наследия является одной из важнейших задач, связанных с культурной идентичностью и коллективной памятью общества. В условиях усиления влияния изменения климата и цифровой трансформации защита физических «оригинальных» материалов – таких как рукописи и художественные эскизы — требует перехода от пассивного хранения к активному управлению параметрами окружающей среды.

В статье представлен Bacteria Blockfire – усовершенствованная версия запатентованной системы Blockfire®, объединяющая мехатронику, прогнозное моделирование и молекулярно-ориентированную биологическую защиту для обеспечения долгосрочного сохранения культурных ценностей. Система построена на основе трех взаимодополняющих технологических компонентов.

ACS (Aerobiological Control System – система аэробиологического контроля), разработанная в сотрудничестве с Национальным исследовательским советом Италии (CNR), использует прогнозный подход, основанный на термогигрометрическом мониторинге и моделях, учитывающих особенности отдельных видов микроорганизмов, включая *Eurotium halophilicum*. Это позволяет системе прогнозировать начало прорастания грибков примерно на 11 дней раньше по сравнению с традиционными регистраторами данных.

На структурном уровне система включает пассивный тепловой барьер с естественной микровентиляцией. Испытания на огнестойкость, проведенные в соответствии со стандартом UNI EN 1363-1 (ISO 834), показали, что даже после 120 минут воздействия внешней температуры в 1000°C внутренняя температура сохраняется на уровне около 78°C, обеспечивая сохранение молекулярной стабильности целлюлозы.

Кроме того, система включает сокристаллы на основе карвакрола, которые выделяют летучие органические соединения (ЛОС), способные подавлять развитие грибков и отпугивать насекомых без изменения химических свойств материалов. Эти материалы обладают адаптивными свойствами, повышая свою активность в условиях окружающей среды, благоприятных для биодеструкции.

Интеграция экологического мониторинга, прогнозного моделирования и инновационных материалов позволяет системе Bacteria Blockfire обеспечить переход от пассивного хранения к активному управлению объектами культурного наследия.

Ключевые слова: превентивная консервация; архивное наследие; биодеструкция; развитие грибков; изоплетная модель; *Eurotium halophilicum*; сокристаллы; ЛОС (летучие органические соединения); прогнозный мониторинг

Introduction. What Does It Mean to Preserve Today?

The preservation of archival heritage, with particular reference to the so-called “original” – whether a manuscript, legal document, or artistic sketch – represents a critical challenge in an era characterized by the digitization of industrial and creative processes and the increasing relevance of cultural enterprises.

Digitization ensures the permanence of content, enhances accessibility, and supports cultural dissemination even in the absence of original materials – materials

that, once lost, are inherently non-reproducible. At best, they may be reconstructed on the basis of supporting sources.

However, digitization and preservation should not be considered alternative or competing processes; rather, they are complementary. Digital media play a fundamental role in the dissemination of knowledge, while original materials are intrinsically linked to collective identity, history, belonging, and memory.

Poor conservation practices, material degradation, and inadequate environmental conditions can severely affect paper-based, wooden, and artistic materials, which embody the identity of nations and communities as repositories of collective memory. The impact is particularly significant for institutions such as libraries, courts, public administrations, hospitals, and museum storage facilities, especially when materials have not been digitized.

This raises a fundamental question: can time – and the damage associated with it – be effectively controlled?

According to the approach developed by Makros, the answer is affirmative, provided that preservation is understood not merely as containment, but as the maintenance of materials within an appropriate internal and external environment – controlled, healthy, and capable of preventing or governing bacterial and fungal proliferation.

Within this framework, the Bacteria Blockfire archival system has been developed to protect materials from environmental damage, increasingly associated with climate change, a priority issue for global policy strategies.

Preventive conservation strategies are widely recognized as the most effective approach for reducing long-term degradation risks in archival environments, particularly when based on environmental control and early detection of biodeterioration processes.

A Step Back

Bacteria Blockfire represents the natural evolution of Blockfire, a patented and internationally established archival system designed to protect materials from fire damage.

Its core innovation lies in a passive thermal barrier. Fire resistance tests conducted according to the UNI EN 1363-1 standard fire curve (ISO 834) demonstrate that, while conventional systems reach temperatures close to the auto-ignition threshold of paper (approximately 230°C), the Blockfire system maintains an internal temperature of approximately 78°C even after 120 minutes of exposure to external temperatures of 1000°C.

Thermodiffractometric analyses conducted by CNR-ICCOM confirm that, under these conditions, the molecular structure of cellulose remains intact.

In this integrated configuration, Bacteria Blockfire provides protection against both thermal and biological degradation processes.

Bacteria Blockfire

The system combines predictive monitoring, passive biological defense, and extreme physical protection within a unified framework.

Its methodology is based on a data-driven approach, beginning with an environmental assessment phase carried out through a “Survey Kit” designed to identify critical conditions without interrupting routine institutional activities.

At the core of the system lies the ACS (Aerobiological Control System), in which environmental data are processed and interpreted in real time, enabling operators to anticipate biological risk conditions.

Integrated directly into shelving systems, ACS continuously monitors relative humidity and temperature in order to feed predictive models of fungal growth. These models have been developed in collaboration with the Italian National Research Council (CNR).

To Protect and Preserve Organic Materials, It Is Essential to Adopt a Preventive Approach to Conservation. What Does This Mean?

It means intervening before biodeterioration processes are triggered, actively inhibiting or excluding degradation agents.

Microbial proliferation in organic materials is typically latent and asymptomatic in its early stages. By the time microbial growth becomes visible, it is often already extensively developed within the material, requiring invasive sanitization interventions that may negatively affect both the environment and the integrity of the materials themselves.

What Degrading Agents Are We Talking About? Moulds, for Example...

Among the main agents of degradation affecting organic collections stored in indoor environments, moulds play a significant role.

These organisms are capable of colonizing complex substrates even under relatively low relative humidity conditions. As pioneer species, preventing their proliferation represents an effective strategy for halting biological succession at its earliest stage.

In order to prevent mould growth in confined environments and on organic materials, it is essential to understand their ecological requirements, particularly their tolerance thresholds with respect to temperature and relative humidity. Monitoring and controlling these parameters directly influences the ability of fungal spores to germinate.

How Can This Be Explained More Clearly?

The germination rate of fungal spores can be described using the isopleth diagram.

The isopleth system is a predictive model that describes spore germination as a function of temperature and relative humidity, represented through contour lines within a temperature–humidity diagram.

This model allows the estimation of both spore germination rates and mycelial growth, taking into account the interaction between environmental parameters and substrate characteristics.

Each fungal species is characterized by a Lowest Isopleth for Mould (LIM), which defines the minimum temperature and relative humidity required for germination. Below this threshold, fungal activity – in terms of both growth and germination – is negligible.

Sedlbauer (2001: 1) provided a framework for predicting fungal proliferation on building, organic, and mineral materials using isopleth systems based on species typical of European climates.

Subsequently, the Hioki Wireless Fungal Logger was developed as a thermo-hygrometric datalogger capable of calculating germination risk based on hourly environmental data. However, such systems do not account for extremophilic species capable of colonizing archival materials, whose LIM thresholds are lower than those considered in standard models.

The Software

The ACS software integrated into Makros mobile shelving systems incorporates an updated reference model based on the xerophilic species *Eurotium halophilicum*. This species is particularly critical in archival environments as it is capable of germinating at lower relative humidity levels than those considered in standard predictive models, thus representing an early indicator of hidden biological risk, widely present in European conservation environments across various organic substrates.

The predictive model implemented within ACS is based on isopleth systems, integrating temperature, relative humidity, and exposure time in order to estimate fungal germination probability under dynamic environmental conditions (Sedlbauer 2001: 1).

Experimental validation conducted in archival environments demonstrated that the ACS system is capable of detecting the onset of fungal germination up to 11 days earlier than conventional commercial dataloggers.

Within an Integrated Pest Management (IPM) framework, this predictive capability is essential for ensuring timely intervention and reducing reliance on chemical disinfection methods.

Scientific Innovations. Passive Biological Defense Through Co-crystals

This approach introduces a key scientific innovation: passive biological defense through carvacrol-based co-crystals.

These solid systems release volatile organic compounds (VOCs) capable of inhibiting the proliferation of xerophilic fungi, such as *Eurotium halophilicum*, and acting as repellents against paper-damaging insects, while ensuring full compatibility with cellulose materials, as validated by studies published in *International Biodeterioration & Biodegradation*.

The effectiveness of this approach has been demonstrated through an 18-month experimental study conducted at Roma Tre University, Italy. The results show that co-crystals act as powerful passive bio-inhibitors. This approach is consistent with recent research on sustainable conservation strategies based on controlled-release systems for volatile bioactive compounds (Mazzeo 2019: 1). Experimental results showed a reduction in detectable fungal activity exceeding 80% under controlled conditions following the application of co-crystal systems (Roma Tre University 2024: 1).

Even in the absence of active diffusion mechanisms, the gradual release of natural volatile terpenes leads to a significant reduction in fungal activity. In many cases, contamination becomes no longer detectable within a few months of application.

Furthermore, the study confirms that the integrated ACS system provides an early-warning advantage of approximately 11 days compared to conventional dataloggers, enabling preventive intervention prior to germination.

Some Examples of Biological Risk Mitigation Measures

In the event of biological risk, one possible strategy involves the controlled release of terpenes for the management of biodeteriogens affecting archival and multi-material assets.

Terpenes, natural compounds derived from essential oils, are well known for their antifungal, antibacterial, and insect-repellent properties. However, their high volatility and liquid state at room temperature make their direct application difficult.

To overcome these limitations, co-crystals incorporating terpenes such as carvacrol, thymol, and eugenol have been developed.

Co-crystals are solid crystalline structures composed of two or more neutral organic molecules – an active compound and a co-former – held together by weak intermolecular forces such as hydrogen bonds or van der Waals interactions.

Within this structure, terpenes are released slowly and consistently, ensuring the maintenance of effective biocidal concentrations over extended periods.

These materials are easy to handle, transport, and dose, eliminating the need for solvents or dispersion systems such as sprays or diffusers.

Moreover, co-crystals act as adaptive or “intelligent” materials: the release of VOCs is enhanced by increases in temperature and relative humidity. As a result, the system intensifies its protective action precisely under environmental conditions that favour microbial growth.

A carvacrol-based co-crystal synthesized with phenazine has demonstrated high efficacy against fungal species such as *Alternaria alternata*, *Aspergillus spp.*,

Cladosporium spp., and *Trichoderma orientale*, as well as significant repellent activity against *Lasioderma serricorne*.

Its effectiveness has also been validated under quasi-real conditions, using inoculated book prototypes exposed to extreme humidity (98% RH) for 28 days. While untreated samples showed visible fungal proliferation, treated samples exhibited no detectable contamination, confirming the ability of vapours to penetrate between pages.

Material compatibility tests conducted on Whatman (pure cellulose) and Kraft (high lignin content) papers – both new and artificially aged – demonstrated that exposure to vapours does not induce degradation.

No significant changes in pH were observed, preventing acidification risks. Colorimetric analysis showed negligible variation ($\Delta E < 1–2$), not perceptible to the human eye. Mechanical testing confirmed that cellulose strength was not compromised, with any reductions attributable solely to humidity or ageing conditions.

FT-IR analysis revealed no chemical alteration of cellulose, and XRD analysis confirmed that crystallinity remained unchanged. No persistent residues were detected after one month.

Conclusions and Impact. The Value of Integrity

The Bacteria Blockfire system redefines conservation standards by offering a scalable, non-invasive technology that can be integrated into both existing infrastructures and new installations.

This approach significantly enhances long-term, verifiable, and responsible conservation, preserving the integrity of archival materials.

Protecting original materials is not only a technical necessity but also a scientific commitment to the future and a fundamental component of modern heritage governance.

Effective preservation means preventing material loss, avoiding emergency interventions, optimizing resource allocation, and promoting a culture of conservation among professionals.

By integrating IoT-based monitoring through ACS with advances in materials science, the system provides archivists with the tools to transition from passive storage to active heritage asset management.

Acknowledgments

The authors acknowledge the project SCRIBA (Sviluppo di CRIstalli contro i biodeteriogeni di Biblioteche ed Archivi), supported by the European Social Fund Plus (ESF+) under the PR Tuscany FSE+ 2021–2027 Programme, funded by the European Union and Regione Toscana within the framework of the "Bando ricerca anno 2024". The project is part of Giovanisì, the Regione Toscana programme for youth autonomy.

A. Ienco, L. Della Maggiora.

Consiglio Nazionale delle Ricerche – Istituto di Chimica dei Composti Organometallici, Italy, Sesto Fiorentino.

References

Mazzeo 2019 – Mazzeo P.P., Carraro C., Monica A., et al. Designing a palette of cocrystals based on essential oil constituents for agricultural applications. *ACS Sustainable Chemistry & Engineering*, 7, 2019. Pp. 17929–17940. URL: https://doi.org/10.1021/acssuschemeng.9b04576 (accessed 15.04.2026).

Menicucci 2022 – Menicucci F., Palagano E., Michelozzi M., et al. Effects of trapped-into-solids volatile organic compounds on paper biodeteriogens. *International Biodeterioration & Biodegradation*, 174, 2022. 105469. URL: https://doi.org/10.1016/j.ibiod.2022.105469

Menicucci 2023 – Menicucci F., Palagano E., Michelozzi M., Ienco A. Essential oils for the conservation of paper items. *Molecules*, 28, 2023. 5003. URL: https://doi.org/10.3390/molecules28135003 (accessed 15.04.2026).

Menicucci 2024 – Menicucci F., Pizzo B., Salvadori B., et al. Antifungal activity of carvacrol-based solids and their effects on Whatman and Kraft paper. *International Biodeterioration & Biodegradation*, 195, 2024. 105894. URL: https://doi.org/10.1016/j.ibiod.2024.105894 (accessed 15.04.2026).

Montanari 2024 – Montanari M. Experimental evaluation of fungal proliferation in archival environments. Roma Tre University.

Pinzari 2006 – Pinzari F., Pasquariello G., De Mico A. Biodeterioration of paper: A SEM study of fungal spoilage reproduced under controlled conditions. *Macromolecular Symposia*, 238, 2006. Pp. 57–66.

Sedlbauer 2001 – Sedlbauer K. Prediction of mould fungus formation on the surface of and inside building components. PhD Dissertation, University of Stuttgart.

US2025341448A1 – Makros S.r.l. Fire-resistant archival containment system. United States Patent Application, 2025.

SECURING DIGITAL ARCHIVES AGAINST DATA LOSS IN THE DISINFORMATION ERA

Radosław Łoboda
IT systems administrator, National Digital Archives, Poland
radoslaw.loboda@nac.gov.pl

Abstract. Archives have long served as the ultimate reference point for verifying historical claims, yet in the current environment of political tensions, armed conflicts, and generative artificial intelligence, the integrity of archival records faces unprecedented pressure. False narratives built from fragments of real events circulate widely, and text produced by large language models introduces fabricated details that are difficult to distinguish from genuine sources. The credibility of digital archives depends not only on the authenticity of their content but also on the resilience of the infrastructure that stores and delivers it. This paper examines the intersection of disinformation, digital preservation, and cybersecurity from the perspective of institutions responsible for maintaining authoritative records. The first part discusses the role of archives as instruments of fact verification in an era of manipulated information. The second part outlines the scale of digitised and born-digital holdings at major institutions, reviews the financial and organisational effort behind digitisation, and catalogues the threats that endanger these collections, from ransomware campaigns and supply chain compromises to distributed denial-of-service attacks and malicious insider activity. The third part describes the layered technical defences available to archival institutions, including firewalls and network monitoring, intrusion detection platforms, access control architectures, and data safeguarding through air-gapped environments, write-once storage, geographically distributed backups, and portable recovery infrastructure. The discussion is intended for archival professionals, technical staff, and institutional decision-makers who need a coherent picture of why protecting the source of truth is as much a cybersecurity problem as it is an archival one.

Keywords: Digital archives; born-digital records; disinformation; generative AI; fact verification; data integrity; backup; ransomware; cybersecurity; incident response

ЖАЛҒАН АҚПАРАТ ТАРАЛҒАН КЕЗЕҢДЕ ЦИФРЛЫҚ АРХИВТЕРДІҢ ҚАУІПСІЗДІГІН ҚАМТАМАСЫЗ ЕТУ ЖӘНЕ ДЕРЕКТЕРДІҢ ЖОЙЫЛУЫНА ЖОЛ БЕРМЕУ

Радослав Лобода
Ақпараттық жүйелердің әкімшісі, Ұлттық цифрлық архивтер, Польша
radoslaw.loboda@nac.gov.pl

Аңдатпа. Архивтер ұзақ уақыт бойы тарихи деректердің шынайылығын тексерудің негізгі қайнар көзі қызметін атқарып келеді. Алайда қазіргі саяси шиеленістер, қарулы қақтығыстар және генеративті жасанды интеллект кеңінен қолданылып отырған жағдайда архивтік жазбалардың тұтастығы бұрын-соңды болмаған сын-қатерлерге тап болуда.

Нақты оқиғалардың жекелеген бөліктерінен құрастырылған жалған ақпараттық баяндаулар кең таралып, ал ірі тілдік модельдер арқылы жасалған мәтіндер шынайы дереккөздерден ажырату қиын болатын ойдан шығарылған мәліметтерді қамтуы мүмкін. Цифрлық архивтердің сенімділігі тек олардың мазмұнының түпнұсқалығына ғана емес, сонымен қатар ақпаратты сақтайтын және пайдаланушыларға ұсынатын инфрақұрылымның тұрақтылығына да байланысты.

Мақалада беделді әрі ресми деректерді сақтауға жауапты мекемелер тұрғысынан жалған ақпарат тарату, цифрлық сақтау және киберқауіпсіздік мәселелерінің өзара байланысы қарастырылады.

Зерттеудің бірінші бөлімінде ақпаратты бұрмалау дәуіріндегі архивтердің тарихи фактілерді тексеру құралы ретіндегі рөлі талданады. Екінші бөлімде ірі архив мекемелеріндегі цифрландырылған және бастапқыдан цифрлық форматта жасалған (born-digital) қорлардың көлемі сипатталып, цифрландыруға жұмсалатын қаржылық және ұйымдастырушылық ресурстар қарастырылады. Сонымен қатар бұл қорларға төнетін қауіптер: бопсалаушы бағдарламалар (ransomware) шабуылдары, жеткізу тізбегіндегі осалдықтар, таратылған қызмет көрсетуден бас тарту (DDoS) шабуылдары және ішкі қауіптер жүйеленеді. Үшінші бөлімде архив мекемелерінде қолданылатын көпдеңгейлі техникалық қорғаныс шаралары сипатталады. Олардың қатарында желілік экрандар (firewall) мен желіні бақылау жүйелері, шабуылдарды анықтау платформалары, қолжетімділікті басқару архитектуралары, сондай-ақ оқшауланған орта (air-gapped environments), бір рет жазылатын сақтау технологиялары (write-once storage), географиялық тұрғыдан бөлінген резервтік көшірмелер және апатты қалпына келтіруге арналған мобильді инфрақұрылым арқылы деректерді қорғау тетіктері бар.

Мақала архив саласының мамандарына, техникалық қызметкерлерге және институционалдық шешім қабылдаушыларға арналған. Зерттеу «ақиқат дереккөзін» қорғау тек архивтік міндет қана емес, сонымен қатар киберқауіпсіздік мәселесі екенін түсінуге бағытталған тұтас көзқарас қалыптастырады.

Түйін сөздер: цифрлық архивтер; бастапқыдан цифрлық форматтағы құжаттар; жалған ақпарат; генеративті жасанды интеллект; фактілерді тексеру; деректер тұтастығы; резервтік көшіру; ransomware (бопсалаушы бағдарламалар); киберқауіпсіздік; инциденттерге жауап беру

ЗАЩИТА ЦИФРОВЫХ АРХИВОВ ОТ ПОТЕРИ ДАННЫХ В ЭПОХУ ДЕЗИНФОРМАЦИИ

Радослав Лобода
Администратор информационных систем,
Национальные цифровые архивы, Польша
radoslaw.loboda@nac.gov.pl

Аннотация. Архивы на протяжении многих лет являлись основным источником проверки исторических утверждений и подтверждения достоверности фактов. Однако в современных условиях политической напряженности, вооруженных конфликтов и широкого распространения генеративного искусственного интеллекта целостность архивных записей сталкивается с беспрецедентными вызовами.

Ложные информационные нарративы, созданные на основе фрагментов реальных событий, получают широкое распространение, а тексты, сформированные большими языковыми моделями, могут содержать вымышленные детали, которые трудно отличить от подлинных источников. Достоверность цифровых архивов зависит не только от подлинности их содержания, но и от устойчивости инфраструктуры, обеспечивающей хранение и предоставление доступа к данным.

В статье рассматривается взаимосвязь дезинформации, цифрового сохранения и кибербезопасности с точки зрения учреждений, ответственных за хранение официальных и авторитетных документов. В первой части анализируется роль архивов как инструмента проверки фактов в эпоху информационных манипуляций. Во второй части рассматриваются масштабы цифровых и изначально цифровых (born-digital) архивных фондов в ведущих учреждениях, оцениваются финансовые и организационные ресурсы, необходимые для цифровизации, а также систематизируются угрозы, представляющие опасность для этих коллекций: атаки с использованием программ-вымогателей (ransomware), компрометация цепочек поставок, распределенные атаки типа «отказ в обслуживании» (DDoS) и вредоносная деятельность внутренних пользователей. В третьей части описываются многоуровневые технические меры защиты, доступные архивным учреждениям, включая межсетевые экраны (firewall) и системы мониторинга сетевой активности, платформы обнаружения вторжений, архитектуры управления доступом, а также механизмы защиты данных посредством изолированных сред (air-gapped environments), технологий однократной записи (write-once storage), географически распределенного резервного копирования и мобильной инфраструктуры восстановления после аварийных ситуаций.

Материал предназначен для специалистов архивного дела, технических сотрудников и руководителей учреждений, которым необходимо комплексное понимание того, почему защита источника достоверной информации является не только архивной задачей, но и важнейшей задачей кибербезопасности.

Ключевые слова: цифровые архивы; изначально цифровые документы (born-digital records); дезинформация; генеративный искусственный интеллект; проверка фактов; целостность данных; резервное копирование; программы-вымогатели (ransomware); кибербезопасность; реагирование на инциденты

Archives as instruments of verification

During periods of political instability and armed conflict, information is routinely distorted. Fragments of genuine events are reassembled into misleading accounts, context is stripped away, and alternative interpretations are presented as established fact. This is not a new phenomenon, but the speed and scale at which manipulated content now travels are without precedent. Social media platforms amplify partial truths within hours, and corrections, when they arrive, rarely reach the same audience. Because the permanent loss of or tampering with digital records would destroy our very baseline for verifying reality, the stakes of this problem extend well beyond institutional IT (Weigle 2023).

Generative artificial intelligence has added another layer of difficulty. Large language models produce fluent text that may contain events that never occurred or attribute statements to people who never made them. The outputs are convincing precisely because they mimic the structure of credible writing, making it harder for non-specialists to separate fabrication from reality. As Yu et al. demonstrate in their survey of fake AI-generated content, the range of modalities and the sophistication of generative techniques now extend well beyond text to include images, audio, and video, all of which can be deployed to construct false historical narratives (Yu 2024). The rising popularity of these tools adds an unpredictable dimension to the future survival of the public domain archive.

When someone claims that a particular event took place, or that a document proves a certain lineage, nationality, or property right, archives are the place where such claims can be checked against original, unaltered records. Census returns, birth registers, military service files, immigration logs, and diplomatic correspondence exist in archival custody for exactly this reason. The importance of presenting these materials in their original, unmodified form cannot be overstated. As Pacheco et al. argue, in the digital environment integrity is not a stable concept but a property of records that must be verified through continuous documentation of any changes over time (Pacheco 2023: 630). The presumption of authenticity that once rested on physical characteristics of paper documents now depends on metadata, provenance chains, and transparent custodial histories.

Consider a straightforward example: disputed biographical facts about individuals whose family histories span several countries and multiple political regimes. Without access to original civil registry entries or church records held in national archives, competing narratives cannot be adjudicated. The same applies to territorial claims, where historical maps and treaty texts held in archival institutions serve as primary evidence. When those records are digitised and made freely available, any person anywhere in the world can consult them directly, reducing dependence on secondary accounts that may carry bias.

The information society has been moving steadily toward digitising everything. The United States National Archives and Records Administration has digitised over 500 million pages available in its online catalogue (NARA 2026). The Internet Archive, a non-profit digital library, has archived more than one trillion web pages and over 99 petabytes of unique data as of late 2025 (Internet Archive 2025). These numbers reflect a broad consensus that digital access is a public good, but they also create an obligation: if society depends on digital copies as the authoritative version of the

historical record, those copies must be protected with the same seriousness once reserved for fireproof vaults and climate-controlled reading rooms.

The scale of the resource and the threats it faces

The volume of digitised material held by major institutions is vast. Table 1 summarises key figures from two of the largest archival operations in the world.

Table 1. *Selected holdings of major archival institutions*

Institution	Metric	Value
NARA (USA), as of April 2026	Paper holdings	13.5 billion pages
	Photographs	41 million
	Maps, charts, drawings	10 million
	Electronic records	approx. 33 billion records (approx. 1 PB)
	Digitised pages online	over 500 million
	New electronic accessions (2024)	463 TB
	Staff	2,384
	Facilities	40, across 17 states and DC
Internet Archive (USA), as of late 2025	Web pages archived	over 1 trillion
	Unique data	99 PB
	Total with redundancy	over 212 PB
	Daily web page captures	approx. 498 million
	Texts in collection	over 47 million
	Videos	over 13 million

Sources: NARA 2026; Internet Archive 2025.

As of April 2026, NARA operates 40 facilities in 17 states plus the District of Columbia, with permanent archival holdings of 13.5 billion pieces of paper, and in 2024 alone it accessioned 88,498 cubic feet of analogue records and 463 terabytes of electronic records. These figures only cover a single national institution. Multiply them by the dozens of national archives, regional repositories, and specialised collections across Europe, Asia, and Latin America, and the aggregate is staggering.

Digitisation itself represents a massive investment of labour, funding, and infrastructure. As part of its latest strategic plan, NARA intended to move a total of 500 million pages online by September 2026, backed by 120 million dollars in digitisation

appropriations since fiscal year 2020. The agency built a dedicated Mass Digitization Center in College Park, Maryland, at a cost of 20 million dollars, equipped with highspeed conveyor belt scanners capable of processing up to 314 pages per minute (NARA 2026). Each digitised page must be scanned, quality-checked, assigned metadata, and ingested into a cataloguing system. The cost of maintaining this pipeline over years is substantial, and it does not end once the page is online; the digital copy must be stored, replicated, migrated to newer formats, and kept accessible indefinitely.

The COVID-19 pandemic accelerated this trajectory. When physical reading rooms closed and researchers could no longer travel to consult original documents, institutions that had invested in digitisation found themselves better positioned to serve their users. Those that had not were forced to accelerate their programmes under difficult conditions. As Schafer and Winters note, the pandemic period highlighted both the value and the fragility of web-based archival access (Schafer 2021: 135).

Born-digital records present a different set of challenges. Government agencies, courts, and private organisations now produce enormous quantities of electronic documents every day. Selecting which materials to preserve requires careful appraisal, because storage is not unlimited and not every digital file warrant permanent retention. Increasingly, archivists must also contend with synthetic content: text, images, and video generated by artificial intelligence systems that can be difficult to distinguish from authentic records without specialised analysis. Access control is another concern, as digital holdings may be harvested without authorisation and used as training data for machine learning models, raising both legal and ethical questions about consent and sensitive personal information. By early 2026, at least 241 websites, including major news organisations such as The Guardian and The New York Times, had begun blocking the Internet Archive's crawlers, citing concerns that the Wayback Machine provides unfettered access to content that AI companies scrape for training data (BleepingComputer 2024).

The dangers facing archives can be grouped into several broad categories. Physical threats remain relevant: fires, floods, earthquakes, and armed conflicts can destroy both original materials and the data centres that hold digital copies.

Political threats involve deliberate interference with content. In late 2024, reports emerged that the leadership of NARA had directed the alteration of museum displays to de-emphasise certain aspects of United States history, including the removal of images related to civil rights leaders and information about the forced incarceration of Japanese Americans during the Second World War. The Congressional Asian Pacific American Caucus, the Congressional Hispanic Caucus, and the Congressional Black Caucus issued a joint statement condemning the reported changes (Densho 2024). Whether one agrees with the characterisation of those decisions or not, the episode illustrates how archival presentation can become a site of political contest.

Criminal threats directed at archives for financial gain constitute perhaps the most acute danger. Modern cybercriminals do not merely seek to disrupt operations or steal data; in many documented cases their goal is the complete destruction of institutional data, either as leverage to compel ransom payment or as retaliation when payment is refused. Ransomware is the most prominent mechanism for this kind of

attack. On the evening of 25 October 2023, affiliates of the Rhysida ransomware group gained initial access to the British Library's network through a Terminal Services server deployed in 2020 that lacked multi-factor authentication. Within minutes, the attackers began moving laterally through the infrastructure. The ransomware was formally identified on 28 October. By that time, approximately 600 gigabytes of data had been exfiltrated, including staff passport scans, personnel files, and 22 internal databases. The attackers also deliberately destroyed servers to inhibit recovery and forensic analysis. Rhysida demanded 20 bitcoins, approximately 596,000 GBP, and when the Library refused, publicly released the stolen data. The Library had to use about 40 percent of its financial reserves, around 6 to 7 million GBP, to recover. Legacy systems and a complex network topology allowed attackers wider access than a more modern architecture would have permitted (British Library 2024).

Distributed denial-of-service attacks represent a distinct category of threat that targets availability rather than confidentiality. By flooding an institution's public-facing servers with enormous volumes of fraudulent traffic, attackers can render archival portals inaccessible to legitimate users for hours or days. The Internet Archive experienced exactly this in October 2024, when a hacktivist group calling itself SN_BLACKMETA launched sustained DDoS attacks that took the site offline for approximately four days. Simultaneously and separately, cybercriminals stole a user authentication database containing 31 million records. In a subsequent incident the same month, attackers exploited unrotated Zendesk API tokens linked to exposed GitLab authentication credentials to access over 800,000 support tickets dating back to 2018. This chain of events illustrates how a single institution can face compounding attacks across multiple vectors in a short period (BleepingComputer 2024).

Software supply chain vulnerabilities add another dimension of risk. Modern archival systems depend on complex stacks of open-source and commercial software components. If any link in that chain is compromised, whether a code library, an authentication token, or an update mechanism, attackers gain a trusted pathway into the institution's infrastructure. The Internet Archive's Zendesk breach, traced to a GitLab authentication token that had been exposed and left unrotated for nearly two years, is a direct example of supply chain failure in the archival domain. More broadly, the near-compromise of the XZ Utils compression library in March 2024, the result of a patient, multi-year social engineering campaign, demonstrated how even foundational open-source components can be weaponised.

Technological obsolescence is a quieter but persistent threat. Hardware and software capable of reading older formats disappear over time, and storage media degrade physically through processes sometimes called bit rot. Magnetic tapes lose signal strength, optical discs delaminate, and hard drives develop bad sectors. Without regular migration to current formats and media, digital collections can become unreadable within a generation.

Digital threats, understood broadly, encompass every pathway by which data in a data centre can be corrupted, altered, or destroyed. These include hardware and software failures leading to data loss; insider threats and human error, such as an administrator accidentally deleting a database, a misconfigured script removing files, or deliberate sabotage by a staff member with privileged access; and criminal intrusions that follow a typical escalation pattern of first demanding ransom, then

threatening to publish confidential data publicly, and finally destroying it altogether. According to the Verizon Data Breach Investigations Report, 60 percent of breaches involve a human element, with stolen credentials serving as the most common initial access vector in 22 percent of all cases (Verizon 2025).

Detection, response, and methods of protection

Protecting digital archives requires a layered approach that begins with controlling who can access what, continues with monitoring systems for suspicious activity, and ends with ensuring that recoverable copies of the data exist even in the worst-case scenario. Proactive defence must begin at the foundational IT level, ensuring that the infrastructure serving archival content to the public remains both highly secure and consistently accessible.

Access control is the foundation. Privileged Access Management systems restrict the use of administrative accounts so that high-level credentials are checked out only when needed and every action is logged. Role-based access ensures that staff members can reach only the data relevant to their responsibilities. Zero-trust architecture takes this further by assuming that no user or device should be trusted by default, even inside the organisational network; every access request is verified against identity, device posture, and context before being granted. As Alexander and Wang observe in their review of cybersecurity data sources and practices, the effectiveness of these controls depends on continuous monitoring and the integration of multiple data streams into a coherent security posture (Alexander 2024: 3).

Archival institutions must move beyond basic precautions to implement a layered security strategy at the network perimeter and within the network itself. Rigorous firewall configurations, including next-generation firewalls with intrusion prevention capabilities, control inbound and outbound traffic and block communication with known malicious hosts. Continuous network scanning, whether through dedicated vulnerability scanners or through the network monitoring features of NDR platforms, identifies misconfigured services, unpatched systems, and anomalous traffic patterns before attackers can exploit them. Active endpoint protection, including anti-malware agents and Endpoint Detection and Response tools deployed on every server and workstation, provides the final barrier against malicious software that bypasses network-level controls.

Attackers bypass these controls through well-known methods: aggressive and sophisticated phishing schemes, including voice phishing that impersonates IT support or management through real telephone calls; malware delivered through compromised websites or email attachments; exploitation of unpatched software vulnerabilities; abuse of excessive privileges granted to service accounts; internet-facing services that were left exposed unnecessarily; and social engineering that manipulates human trust rather than technical systems.

Once inside a network, attackers often move laterally, exploring systems and escalating privileges over a period of days or weeks before taking action. In 2025, the global median dwell time, the period from initial intrusion to detection, rose to 14 days, an increase from 11 days in 2024, reversing several years of improvement. The rise was driven largely by long-term espionage operations and insider placement schemes with median dwell times exceeding 122 days. Where organisations detected intrusions internally, the median fell to 9 days, while those relying on external notification faced

a median of 25 days (Mandiant 2026). Separately, incident response data from managed detection providers showed an overall median of just 2 days when active monitoring was in place, with ransomware cases detected in a median of 3 days and nonransomware cases in a single day (Sophos 2025). These figures demonstrate that continuous, professional monitoring dramatically compresses the window available to attackers, whereas organisations without such capability remain exposed for far longer. At the other end of the timeline, hand-offs between access brokers and ransomware operators now happen in as little as 22 seconds (Mandiant 2026).

Detecting suspicious activity relies on several complementary systems. A Security Information and Event Management (SIEM) platform aggregates logs from servers, firewalls, endpoints, and applications into a single searchable repository, applies correlation rules to identify patterns that may indicate an attack, and generates alerts for human analysts. As Măcăneață notes, SIEM systems serve as the central nervous system of organisational cybersecurity, but their effectiveness depends critically on the quality of rule tuning and the reduction of false positives (Măcăneață 2024: 16). Security Orchestration, Automation, and Response (SOAR) tools can automate parts of the response workflow, such as isolating a compromised host or blocking a suspicious network address. A Security Operations Centre (SOC) is the team, whether internal or contracted from an external provider, that monitors these alerts around the clock and decides how to respond. At a higher integration level, Extended Detection and Response (XDR) platforms combine telemetry from endpoints, networks, cloud services, and identity systems into a unified detection and response capability.

These systems are powerful but costly. Deploying a SIEM requires not only licences but also significant effort to connect data sources, tune correlation rules, and reduce false positives to a manageable level. The volume of logs generated by a midsized organisation can easily reach several terabytes per day, and retention policies that keep logs for months or years create substantial additional storage demands. A medium archive or library generating 500 gigabytes of logs per day would accumulate roughly 180 terabytes per year of raw log data alone. Staffing a SOC requires trained analysts available on a 24/7 schedule, or the organisation must purchase managed detection services from a specialist provider. Either path represents a recurring cost that competes with the institution's primary mission of preservation and access.

Backup strategies are the last line of defence when prevention and detection fail. The widely cited 3-2-1 rule calls for keeping at least three copies of data, on at least two different types of media, with at least one copy stored offsite. For archival institutions, this baseline should be treated as a minimum. Guaranteeing the survival of the public record requires robust, multi-tiered data and backup architectures that combine online, nearline, and offline solutions.

An air-gapped backup environment is physically disconnected from any network, meaning that ransomware spreading through connected systems cannot reach it. Data is transferred to the air-gapped environment on a schedule, typically by physically moving media or through a controlled, one-directional data transfer mechanism. Immutable storage such as Write Once Read Many (WORM) ensures that once data is written, it cannot be overwritten, encrypted or deleted, even by an administrator. This property is critical for archives because it provides a technical guarantee of

immutability that mirrors the archival principle of preserving records in their original state.

Modern backup practice also includes analysis of entropy, structure and statistical monitoring of backup data to detect anomalies that may indicate encryption by ransomware before the encrypted data overwrites healthy copies. Signature-based scanning of backup images for known malware patterns adds another layer of assurance that a restoration will not reintroduce the threat. Analysing data access patterns with Machine Learning/AI can also hint on unusual activity within backup systems.

Magnetic tape remains a practical medium for large-scale archival backup. Current LTO-9 cartridges cost approximately five dollars per terabyte, compared to over twenty dollars per terabyte for high-capacity hard drives and substantially more for cloud cold storage over a ten-year period. A tape cartridge stored in proper conditions can retain data for decades. However, tape requires maintenance: cartridges should be periodically rewound to prevent sticking, test restores should be performed regularly to confirm readability, and when tape generations change, data must be migrated to the newer format before drives capable of reading the older cartridges become unavailable. Physical transport of tape cartridges to geographically separated vaults is a well-established practice for disaster recovery.

To maintain uninterrupted public access even during a catastrophic failure or targeted strike, institutional data centres should maintain copies of their holdings across distinct geographical locations. Object storage platforms support techniques such as sharding, which distributes pieces of a file across multiple nodes, and erasure coding, which adds calculated redundancy so that the original data can be reconstructed even if some nodes fail. A minimum of three geographically separated locations is generally recommended for this model. Object storage also supports granular access controls and can enforce WORM policies at the storage layer. A significant benefit is the ability to store metadata about a record along with the data itself, which can then be used to efficiently search through entire set to gather related records.

For extreme scenarios requiring rapid deployment or the physical relocation of networks, portable ruggedised infrastructure offers an additional layer of resilience. Compact all-flash storage arrays and virtualisation servers, loaded with copies of the most critical data and the software needed to serve it, can be housed in reinforced cases or small cabinets, loaded onto a vehicle, and transported to any location with power and network connectivity. In a scenario where the primary data centre is destroyed or inaccessible, such a portable unit can restore essential services within hours rather than days. This concept borrows from military and emergency-response logistics and is particularly relevant for institutions in regions exposed to armed conflict or natural disaster.

Conclusion

No single technology is sufficient on its own. Effective protection of digital archives requires that access controls, monitoring systems, backup strategies, and recovery infrastructure work together as a coordinated whole, supported by trained personnel who understand both the archival mission and the technical risks. The cost of implementing these measures is significant, but as the British Library experience

demonstrates, the cost of recovery after a successful attack, approximately 7 million GBP in that case, almost always exceeds the cost of prevention (British Library 2024). Even the most comprehensive digital defences, however, have a structural blind spot: a patient, well-placed insider with legitimate privileges can corrupt records through authorised channels at a pace below detection thresholds, and those corrupted records will be faithfully replicated to every backup tier, including air-gapped and immutable storage, because the writes pass through the legitimate ingest pipeline. The only defence against such a scenario is external, independent verification of content against physically preserved originals held entirely outside the digital chain of custody. Bridging the gap between traditional digital preservation and extensive cybersecurity protocols is no longer just an administrative task; it is a societal obligation. Archival leadership must act immediately to establish layered physical and digital security, ensuring that uncompromised historical resources remain reliably available to counter modern falsehoods.

References

Alexander 2024 – Alexander C. A., Wang L. Cybersecurity data sources and practices. Journal of Computer Networks, 2024, Vol. 12, No. 1, pp. 1–6. https://pubs.sciepub.com/jcn/12/1/1/index.html (accessed 26.04.2026).

BleepingComputer 2024 – Cimpanu C. Internet Archive hacked, data breach impacts 31 million users. BleepingComputer, 10 October 2024. https://www.bleepingcomputer.com/news/security/internet-archive-hacked-data-breachimpacts-31-million-users/ (accessed 20.04.2026).

British Library 2024 – British Library. British Library cyber incident review. London, March 2024. https://www.bl.uk/home/british-library-cyber-incident-review-8-march-2024.pdf (accessed 20.04.2026).

Densho 2024 – Densho. National Archives: whitewashing ugly histories of Japanese American incarceration. 4 November 2024. https://densho.org/catalyst/national-archives-whitewashing-uglyhistories-japanese-american-incarceration/ (accessed 20.04.2026).

Internet Archive 2025 – Internet Archive. About the Internet Archive. https://archive.org/about/ (accessed 20.04.2026).

Măcăneață 2024 – Măcăneață C. Overview of security information and event management systems. Informatica Economică, 2024, Vol. 28, No. 1, pp. 15–24. https://www.revistaie.ase.ro/content/109/02%20-%20macaneata.pdf (accessed 26.04.2026).

Mandiant 2026 – Mandiant. M-Trends 2026 Special Report. Google Cloud, March 2026. https://cloud.google.com/security/resources/m-trends (accessed 22.04.2026).

NARA 2026 – National Archives and Records Administration. About the National Archives of the United States. https://www.archives.gov/about (accessed 22.04.2026).

Pacheco 2023 – Pacheco A., Guardado Da Silva C., Vieira De Freitas M. C. A metadata model for authenticity in digital archival descriptions. Archival Science, 2023, Vol. 23, pp. 629–673. https://d-nb.info/1315497565/34 (accessed 20.04.2026).

Schafer 2021 – Schafer V., Winters J. The values of web archives. International Journal of Digital Humanities, 2021, Vol. 2, pp. 129– 144. https://pmc.ncbi.nlm.nih.gov/articles/PMC8190571/ (accessed 20.04.2026).

Sophos 2025 – Sophos. The 2025 Sophos Active Adversary Report: It Takes Two. 2 April 2025. https://www.sophos.com/en-us/blog/2025-sophos-active-adversary-report (accessed 22.04.2026).

Verizon 2025 – Verizon. Data Breach Investigations Report 2025. https://www.verizon.com/business/resources/reports/dbir/ (accessed 22.04.2026).

Weigle 2023 – Weigle M. C. The use of web archives in disinformation research. arXiv, 2023. https://arxiv.org/abs/2306.10004 (accessed 26.04.2026).

Yu 2024 – Yu X., Wang Y., Chen Y., Tao Z., Xi D., Song S., Niu S., Li Z. Fake artificial intelligence generated contents (FAIGC): a survey of theories, detection methods, and opportunities. arXiv, 2024. https://arxiv.org/abs/2405.00711 (accessed 26.04.2026).

СОВРЕМЕННЫЕ ПРОБЛЕМЫ АРХИВОВ ГОСУДАРСТВЕННЫХ СТРУКТУР И БИЗНЕСА И ПУТИ ИХ РЕШЕНИЯ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ CANON

Жандаулет Жумагалиев
Менеджер по работе с корпоративными клиентами (Account Manager B2B), Представительство Canon Central and Eastern Europe в Казахстане
zhandaulet.zhumagaliyev@canon.kz

Аннотация. Современные архивы, как государственных учреждений, так и коммерческих организаций, сталкиваются с рядом системных вызовов: ростом объемов бумажных фондов, рисками физического износа документов, сложностью поиска информации, а также ужесточением регуляторных требований. В докладе рассматриваются ключевые проблемы архивного дела и предлагается комплексный технологический подход компании Canon для их решения. Решение охватывает полный жизненный цикл документа: от оцифровки бумажных оригиналов с помощью сканирующих устройств Canon – через интеллектуальное извлечение данных и OCR-обработку посредством Scan2x – до централизованного хранения, управления доступом, автоматизации бизнес-процессов и аналитики в системе электронного документооборота Therefore. Отдельно рассматривается инструмент Readiris PDF как профессиональное средство редактирования, конвертации и защиты электронных документов. Интеграция трех решений позволяет организациям не просто оцифровать архив, но преобразовать его в эффективную, удобную для поиска и соответствующую нормативным требованиям систему управления документами и знаниями.

Ключевые слова: цифровой архив; управление документооборотом; оцифровка; OCR; Scan2x; Therefore; Readiris PDF; Canon; управление электронными документами; автоматизация процессов; жизненный цикл документа

МЕМЛЕКЕТТІК ҚҰРЫЛЫМДАР МЕН БИЗНЕС АРХИВТЕРІНІҢ ЗАМАНАУИ МӘСЕЛЕЛЕРІ ЖӘНЕ ОЛАРДЫ CANON ТЕХНОЛОГИЯЛАРЫН ҚОЛДАНУ АРҚЫЛЫ ШЕШУ ЖОЛДАРЫ

Жандаулет Жұмағалиев
Корпоративтік клиенттермен жұмыс жөніндегі менеджер (Account Manager B2B), Canon Central and Eastern Europe компаниясының Қазақстандағы өкілдігі
zhandaulet.zhumagaliyev@canon.kz

Аңдатпа. Қазіргі заманғы архивтер – мемлекеттік мекемелердің де, коммерциялық ұйымдардың да архивтері – бірқатар жүйелі сын-қатерлермен бетпе-бет келуде: қағаз қорлары көлемінің ұлғаюы, құжаттардың физикалық тозу қаупі, ақпаратты іздеудің күрделілігі, сондай-ақ нормативтік талаптардың қатаңдатылуы.

Мақалада архив ісіндегі негізгі мәселелер қарастырылып, оларды шешуге бағытталған Canon компаниясының кешенді технологиялық тәсілі ұсынылады. Бұл шешім құжаттың толық өмірлік циклін қамтиды: Canon сканерлеу құрылғылары арқылы қағаз түпнұсқаларды цифрландырудан бастап, Scan2x жүйесі арқылы деректерді интеллектуалды түрде шығару және OCR өңдеу, одан әрі Therefore электрондық құжат айналымы жүйесінде орталықтандырылған сақтау, қолжетімділікті басқару, бизнес-процестерді автоматтандыру және аналитика жүргізуге дейінгі үдерістерді қамтиды.

Сонымен қатар электрондық құжаттарды өңдеу, түрлендіру және қорғауға арналған кәсіби құрал ретінде Readiris PDF бағдарламасының мүмкіндіктері жеке қарастырылады.

Үш шешімнің интеграциясы ұйымдарға архивтерді жай ғана цифрландыруға емес, сонымен қатар ақпаратты іздеуге ыңғайлы, тиімді және нормативтік талаптарға сәйкес келетін құжаттар мен білімді басқарудың заманауи жүйесін қалыптастыруға мүмкіндік береді.

Түйін сөздер: цифрлық архив; құжат айналымын басқару; цифрландыру; OCR; Scan2x; Therefore; Readiris PDF; Canon; электрондық құжаттарды басқару; үдерістерді автоматтандыру; құжаттың өмірлік циклі

MODERN CHALLENGES OF PUBLIC AND BUSINESS ARCHIVES AND THEIR SOLUTIONS USING CANON TECHNOLOGIES

Zhandaulet Zhumagaliyev
Account Manager B2B, Canon Central and Eastern Europe, Representative Office in Kazakhstan
zhandaulet.zhumagaliyev@canon.kz

Abstract. Contemporary archives - both in public institutions and private enterprises - face a range of systemic challenges: growing volumes of paper-based records, risks of physical deterioration, difficulties in information retrieval, and increasingly complex regulatory requirements. This paper examines the core problems encountered in modern archival practice and presents a comprehensive technological framework developed by Canon to address them. The proposed solution covers the full document lifecycle: from digitisation of paper originals using high-performance Canon scanning devices, through intelligent data extraction and OCR processing via Scan2x, to centralised storage, access control, workflow automation, and analytics within the Therefore document management system. The paper

further describes Readiris PDF as a professional tool for editing, converting, and securing electronic documents. The integration of these three solutions enables organisations not only to digitise their archives but to transform them into efficient, searchable, and compliant knowledge management systems. The authors argue that digital transformation of archives is no longer a matter of simple scanning but requires a holistic approach to document governance aligned with current legal and regulatory standards.

Keywords: digital archive; document management; digitisation; OCR; Scan2x; Therefore; Readiris PDF; Canon; electronic records management; workflow automation; document lifecycle

1. Современные проблемы архивов

Сегодня архивы сталкиваются с рядом системных вызовов, затрагивающих как государственные учреждения, так и коммерческие организации. Ключевая задача архива состоит не только в сохранении документов, но и в обеспечении их доступности, целостности и соответствия нормативным требованиям на протяжении длительного времени.

Рост объема бумажных фондов

Физические архивы исчерпывают доступные помещения, а их расширение сопряжено со значительными затратами. Ограниченность пространства влечет трудности масштабирования, ограниченный доступ к документам и трудоемкость учета. Ручное управление бумажными фондами требует значительных временных и кадровых ресурсов.

Риски утраты и физического износа

Бумажные носители подвержены естественному старению: бумага со временем разрушается, чернила выцветают, а частое использование документов приводит к их механическому повреждению. Утрата уникальных исторических материалов зачастую невосполнима.

Сложность поиска и доступа к информации

Даже при наличии документов в архиве они не всегда доступны в нужный момент. Поиск конкретного дела в неструктурированных фондах может занимать часы и требует высокой квалификации архивиста. Отсутствие единой системы учета ведет к дублированию данных и ошибкам.

Юридические требования и регуляторная политика

Современное законодательство предъявляет строгие требования к управлению документами: необходимость контроля версионности, ведение истории изменений, управление правами доступа и строгое соблюдение сроков хранения. Несоответствие нормативным требованиям влечет юридические и финансовые риски для организации. Именно поэтому цифровизация архива сегодня – это не просто сканирование, а комплексный подход к управлению документами на протяжении всего их жизненного цикла.

2. Архитектура решения Canon: полный цикл работы с документами

Решение Canon охватывает весь жизненный цикл документа – от бумажного носителя до аналитики и автоматизации процессов – и реализуется в четыре последовательных этапа.

Этап 1. Преобразование бумажных документов в электронные

Сканирующие устройства Canon обеспечивают быстрое и точное сканирование документов любой сложности и объема, включая ветхие, нестандартного формата или многостраничные материалы. Высокое разрешение и надежная подача листов гарантируют качество оцифровки.

Этап 2. Преобразование электронных документов в данные

На данном этапе осуществляется извлечение контента и метаданных из отсканированных материалов. Процесс включает: извлечение текста (полностью или частично); формирование индексов и метаданных; преобразование форматов; редактирование; перевод; иные функции обработки данных. Автоматическое распознавание ключевых реквизитов документа существенно сокращает трудозатраты на ручной ввод данных.

Этап 3. Хранение, поиск и управление

Система управления документооборотом (СЭД) позволяет: организовать централизованное хранение электронных документов; обеспечить быстрый полнотекстовый поиск; управлять правами доступа; контролировать версии документов; автоматизировать бизнес-процессы; формировать отчеты и аналитику.

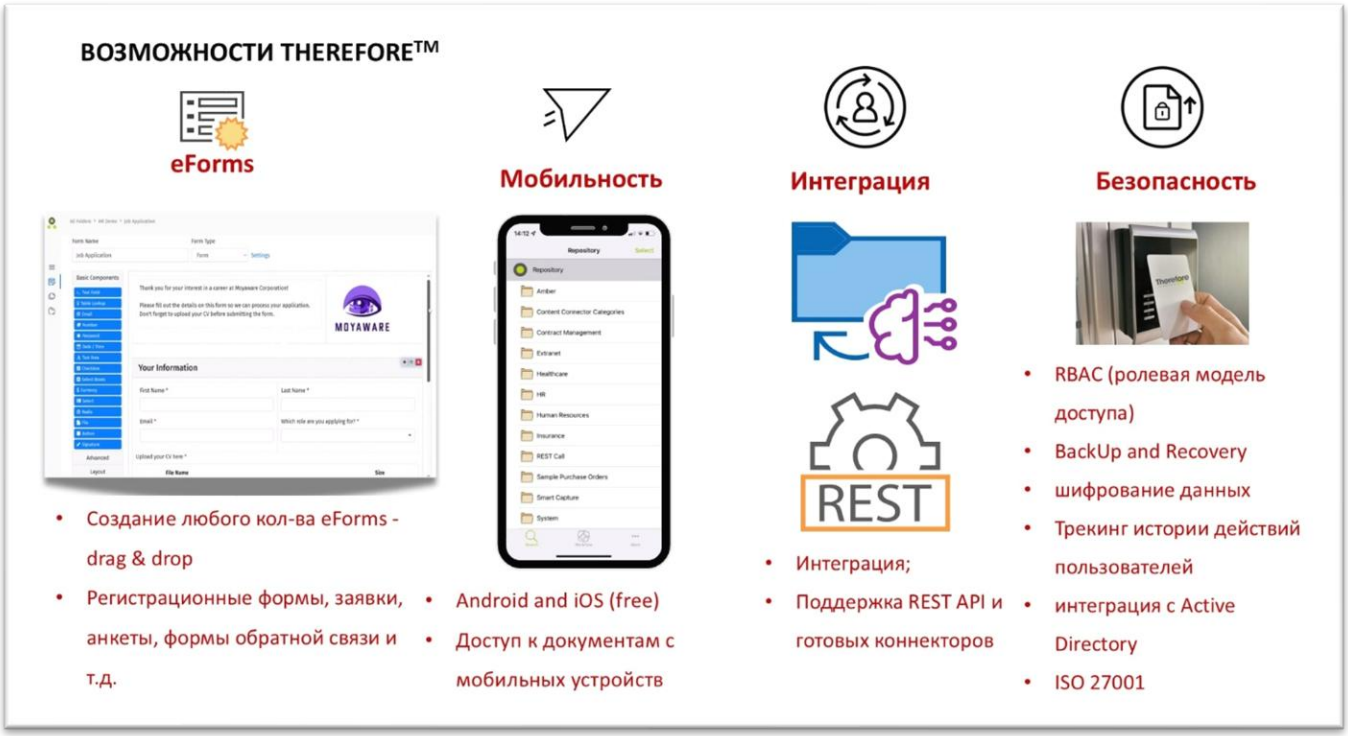
Этап 4. Редактирование и преобразование документов

Работа с форматами PDF, TIFF, JPEG, PNG предусматривает OCR-распознавание, редактирование, конвертацию в форматы MS Office, ePub, MP3 и другие, а также добавление электронной подписи и средств защиты информации.

3. Описание решений

Scan2x

Scan2x представляет собой интеллектуальную систему извлечения данных с использованием расширенных возможностей искусственного интеллекта. Система обеспечивает автоматическое распознавание текста (OCR), включая рукописный ввод, на более чем 130 языках с точностью свыше 99%. Гибкая система профилей позволяет настроить автоматическую маршрутизацию документов в нужные папки или системы хранения без участия оператора. Scan2x поддерживает пакетную обработку с автоматическим разделением документов по разделительным листам, штрихкодам или содержимому, что особенно востребовано при массовом переводе бумажных архивов в цифровой формат.



Therefore™

Therefore – система управления документооборотом, обеспечивающая централизованное хранение и мгновенный полнотекстовый поиск среди миллионов документов; гибкое управление правами доступа с полным журналом аудита; визуальный конструктор рабочих процессов для автоматизации согласования, утверждения и контроля сроков хранения; сбор статистики и аналитику по работе с документами. Система функционирует как в облачной, так и в локальной или гибридной инфраструктуре и интегрируется с

Microsoft SharePoint, SAP, Microsoft 365, а также с отечественными СЭД и системами ЭДО.

Readiris PDF

Readiris PDF – полноценный профессиональный инструмент для работы с PDF-документами, включающий: создание и редактирование документов, конвертацию и объединение файлов, автоматическое разделение и переименование; добавление электронной подписи; добавление парольной защиты; редактирование (замазывание) персональных данных в соответствии с требованиями GDPR; работу с неограниченным количеством документов.

Заключение

Для создания, модернизации и эффективного управления цифровым архивом Canon предлагает комплексный подход, включающий: широкую линейку сканирующих устройств для решения задач любой сложности; решение Scan2x для интеллектуального извлечения данных и OCR (включая рукописный текст); систему Therefore для управления документооборотом, хранения, поиска и автоматизации процессов; инструмент Readiris PDF для удобной и эффективной работы с электронными документами.

Совместное применение трех решений Canon позволяет организациям не только оцифровать архив, но и превратить его в эффективную цифровую систему управления знаниями и документами, полностью отвечающую современным требованиям законодательства и международным стандартам архивного дела.

Литература

Canon Therefore™ Online: Product Overview and Documentation. https://www.therefore.net

Scan2x Intelligent Capture Solution. Canon Europe Ltd. https://www.canon-europe.com/business/products/software/avantech-scan2x/

ISO 15489-1:2016 – Information and documentation - Records management - Part 1: Concepts and principles. Geneva: ISO, 2016.

Readiris PDF 23 Business: Professional PDF Management. IRIS Group / Canon. https://iriscorporate.com/softwares/readiris-pdf/

Алпысбаева Н., Шайлазымов Б. Оцифровка архивных документов: методические рекомендации. Алматы: Қазақ университеті, 2020. 68 с.

INTELLIGENT ARCHIVAL INFRASTRUCTURE MANAGEMENT SYSTEM

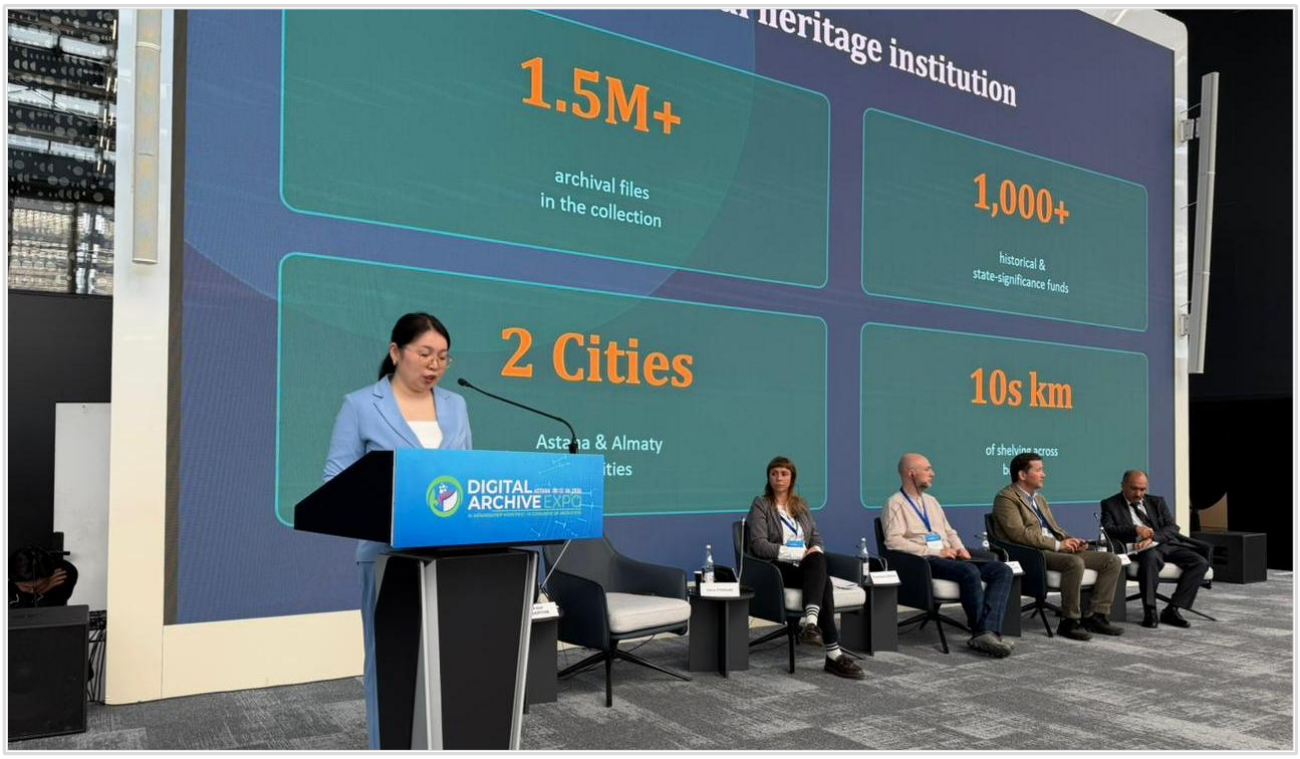
Kuralai Kabarova
Head of the Digital Technologies and Innovation Department,
Archive of the President of the Republic of Kazakhstan

АРХИВТІК ИНФРАҚҰРЫЛЫМДЫ БАСҚАРУДЫҢ ЗИЯТКЕРЛІК ЖҮЙЕСІ

Құралай Қабарова
Цифрлық технологиялар және инновациялар қызметінің басшысы,
Қазақстан Республикасы Президентінің Архиві

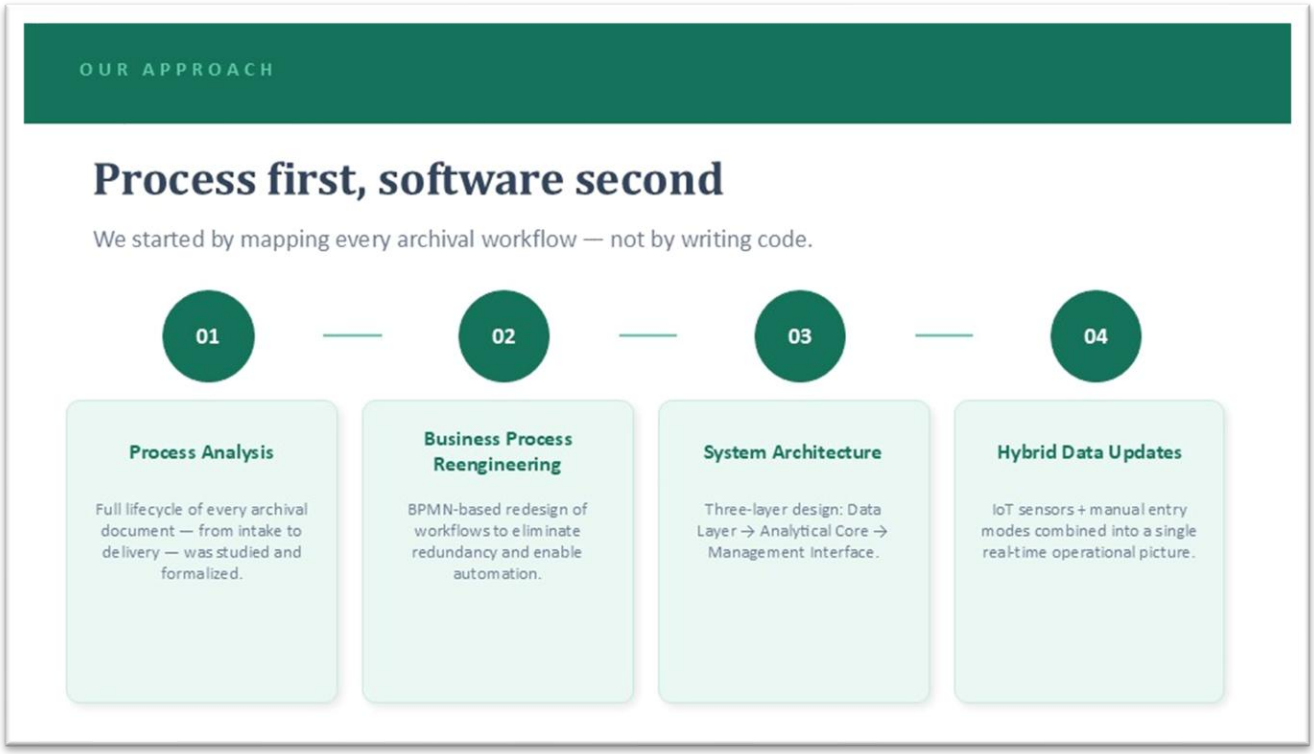
ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ АРХИВНОЙ ИНФРАСТРУКТУРОЙ

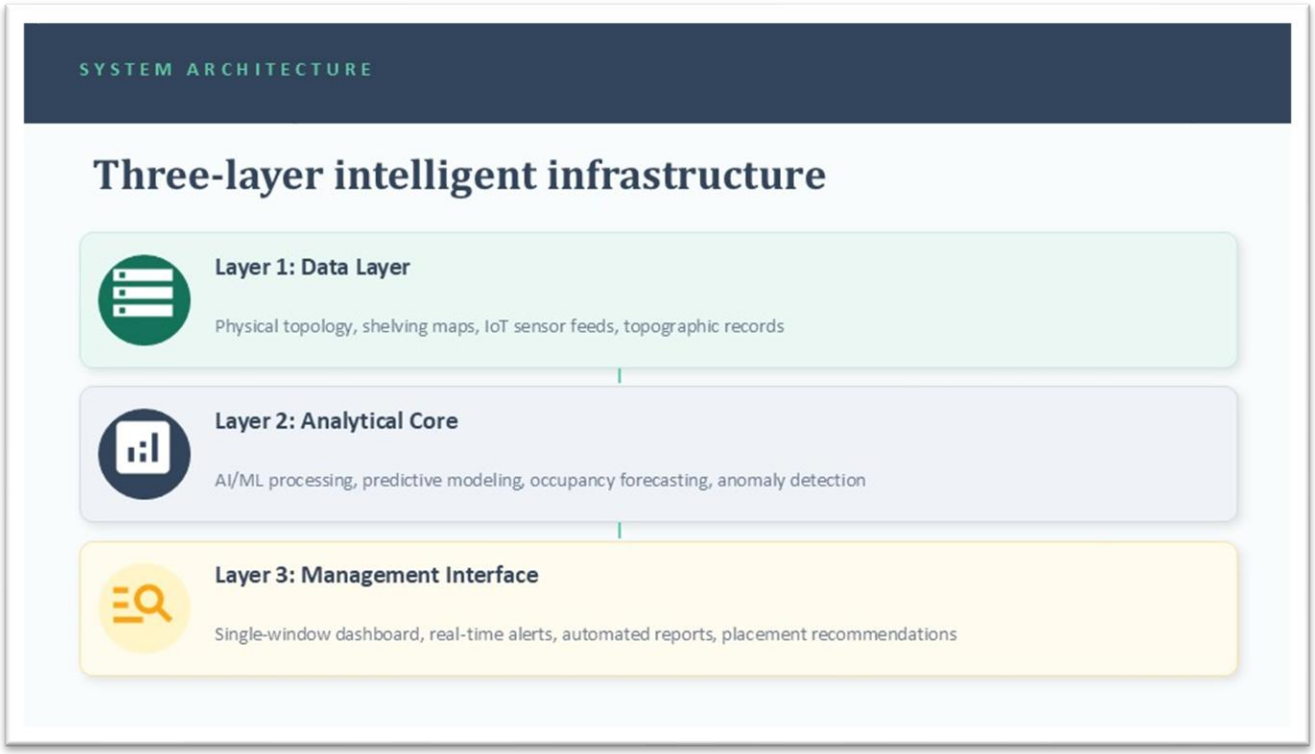
Куралай Кабарова
Руководитель Службы цифровых технологий и инноваций,
Архив Президента Республики Казахстан


heritage institution
1.5M+
archival files
in the collection
1,000+
historical &
state-significance funds
2 Cities
Astana & Almaty
10s km
DIGITAL ARCHIVE EXPO

THE CHALLENGE
Archives are more than documents
Complex Infrastructure
Tens of kilometers of shelving, specialized storage zones, engineering systems, and climate control — all interconnected.
Fragmented Data
Information scattered across separate registers, journals, and isolated systems — requiring manual cross-referencing.
Reactive Management
Decisions based on yesterday's figures. No unified view of current occupancy, conditions, or upcoming risks.

THE ARCHIVE AT SCALE
Managing a national heritage institution
1.5M+
archival files
in the collection
1,000+
historical &
state-significance funds
2 Cities
Astana & Almaty
facilities
10s km
of shelving across
both sites


OUR APPROACH
Process first, software second
We started by mapping every archival workflow — not by writing code.
01
02
03
04
Process Analysis
Full lifecycle of every archival document — from intake to delivery — was studied and formalized.
Business Process Reengineering
BPMN-based redesign of workflows to eliminate redundancy and enable automation.
System Architecture
Three-layer design: Data Layer → Analytical Core → Management Interface.
Hybrid Data Updates
IoT sensors + manual entry modes combined into a single real-time operational picture.


SYSTEM ARCHITECTURE
Three-layer intelligent infrastructure
Layer 1: Data Layer
Physical topology, shelving maps, IoT sensor feeds, topographic records
Layer 2: Analytical Core
AI/ML processing, predictive modeling, occupancy forecasting, anomaly detection
Layer 3: Management Interface
Single-window dashboard, real-time alerts, automated reports, placement recommendations

PILOT RESULTS
Measurable impact from day one
BEFORE
Hours
AFTER
→ Seconds
Locating free storage space for new accessions
BEFORE
Days
AFTER
→ Minutes
Generating a full warehouse passport document
BEFORE
~85%
AFTER
→ ~100%
Accuracy of topographic placement records
BEFORE
Siloed
AFTER
→ One window
Access to all operational data sources

THE VISION

Towards the Archive Digital Twin

Digital Twin technology is already transforming manufacturing, logistics, and smart cities. Archives are next.

- See not just current state — but future trajectories
- Forecast: when will storage zones reach capacity?
- Anticipate risks before they become incidents
- Shift from reactive to proactive management

From Reactive to Proactive

Reactive	→	**Proactive**
Document-centric	→	**Infrastructure-aware**
Yesterday's data	→	**Real-time + forecast**

THE COMPETITIVE ADVANTAGE OF THE 21ST CENTURY ARCHIVE

Not how many documents you've digitized —
but how precisely you can predict your own infrastructure.

AI Geotopography is our first step toward that future.

If the international archival community begins this conversation today,
the greatest result of this project will not be the technology itself —
but its contribution to shaping a new direction for the profession.

Thank you for your attention

«ЖЕТІЛДІРІЛГЕН ТАРИХ: ҚҰЖАТТАРМЕН ЖҰМЫС ІСТЕУ БАРЫСЫНДА БІЛІМ АЛУШЫЛАРДЫҢ НЕҒҰРЛЫМ ТИІМДІ ОЙЛАУЫН ДАМЫТУ ҮШІН ЖАСАНДЫ ИНТЕЛЛЕКТТІ ПАЙДАЛАНУ»
семинары

Семинар
«УЛУЧШЕННАЯ ИСТОРИЯ: ИСПОЛЬЗОВАНИЕ ИИ, ЧТОБЫ ПОМОЧЬ УЧАЩИМСЯ РАЗВИТЬ БОЛЕЕ ЭФФЕКТИВНОЕ МЫШЛЕНИЕ ПРИ РАБОТЕ С ДОКУМЕНТАМИ»

«BETTER HISTORY: USING AI TO DEVELOP BETTER HISTORICAL THINKING BY STUDENTS»
Seminar

BETTER HISTORY: USING AI TO DEVELOP BETTER HISTORICAL THINKING BY STUDENTS

Andrew Payne
Head of Education and Outreach at the National Archives of the United Kingdom
Andrew.Payne@nationalarchives.gov.uk

Abstract. This study examines the UK National Archives' experience in using Artificial Intelligence (AI) to enhance historical thinking among school students. According to OFSTED reports, pupils frequently hold misconceptions regarding historical enquiry methodologies. To address this, AI systems based on Microsoft Copilot and Anthropic Claude were piloted. The findings demonstrate that AI effectively assists students in analyzing extensive document collections and provides scaffolded feedback without directly giving answers. However, limitations such as AI hallucinations, user-pleasing bias, and failure to consistently redirect students to primary sources were identified. Ultimately, AI cannot replace teachers but serves as a powerful supportive educational tool.

Keywords: artificial intelligence; historical thinking; archival documents; education; methodology; students; enquiry; history; analysis; feedback; limitations; teacher

ЖЕТІЛДІРІЛГЕН ТАРИХ: ОҚУШЫЛАРДЫҢ ТАРИХИ ОЙЛАУЫН ЖЕТІЛДІРУ ҮШІН ЖИ ПАЙДАЛАНУ

Эндрю Пейн
Ұлыбритания Ұлттық архивінің білім беру және ағартушылық қызмет бөлімінің басшысы
Andrew.Payne@nationalarchives.gov.uk

Аңдатпа. Зерттеуде Ұлыбританияның Ұлттық архивінде мектеп оқушыларының тарихи ойлау қабілетін дамыту үшін жасанды интеллектті (ЖИ) қолдану тәжірибесі

қарастырылады. OFSTED есептеріне сәйкес, оқушыларда тарихи зерттеу әдіснамасы бойынша қате түсініктер жиі кездеседі. Мәселені шешу үшін Microsoft Copilot және Anthropic Claude негізіндегі ЖИ жүйелері сыналды. Эксперимент нәтижесі ЖИ-дің құжаттармен жұмыс істеу тиімділігін арттырып, оқушыларды бағыттай алатынын көрсетті. Дегенмен, ЖИ жүйелерінің логикалық қателіктер жіберуі және дереккөздерге сілтеме жасалмай, пайдаланушыға бейімделуге тырысуы сияқты кемшіліктері анықталды. ЖИ мұғалімді алмастыра алмайды, бірақ қосымша құрал ретінде маңызды рөл атқарады.

Түйін сөздер: жасанды интеллект, тарихи ойлау, архивтік құжаттар, білім беру, әдіснама, оқушылар, зерттеу, тарих, талдау, кері байланыс, шектеулер, мұғалім

УЛУЧШЕННАЯ ИСТОРИЯ: ИСПОЛЬЗОВАНИЕ ИИ ДЛЯ РАЗВИТИЯ ИСТОРИЧЕСКОГО МЫШЛЕНИЯ УЧАЩИХСЯ

Эндрю Пейн
Руководитель отдела образовательной и просветительской деятельности Национального архива Великобритании
Andrew.Payne@nationalarchives.gov.uk

Аннотация. В данном исследовании рассматривается опыт Национального архива Великобритании по внедрению искусственного интеллекта (ИИ) для развития исторического мышления у школьников. Согласно отчетам OFSTED, учащиеся часто имеют неверные представления о методологии исторических исследований. Для решения проблемы были протестированы ИИ-системы на базе Microsoft Copilot и Anthropic Claude. Результаты показали, что ИИ помогает эффективно анализировать массивы документов и направляет учащихся без предоставления готовых ответов. Однако выявлены и недостатки: галлюцинации ИИ, чрезмерное стремление угодить пользователю и игнорирование первоисточников в дискуссии. ИИ не заменяет учителя, но служит мощным вспомогательным инструментом.

Ключевые слова: искусственный интеллект, историческое мышление, архивные документы, образование, методология, учащиеся, исследования, история, анализ, обратная связь, ограничения, учитель

The Education and Outreach Department at the National Archives of the United Kingdom has recently been investigating whether Artificial Intelligence can help school students to become better at historical thinking. Many teachers are concerned that AI may be reducing the thinking that students do in their assignments and coursework, and this paper shares the team's experience of working with students using collections of documents from our archive.

In the United Kingdom, and in other jurisdictions too, there has been a lot of discussion about what is the most effective methodology for teaching history. Over the past 15 years there has been a greater emphasis on the importance of subject knowledge and trying to help students remember a huge amount of information which they are then tested on in their external exams at the age of 16 and 18. Archives, of course, are full of information so we are able to support students, and teachers, with subject knowledge, but we believe we can do more than that. We believe that knowing and understanding about archives is essential for history students because they are the foundation of the subject.

When working with students we talk about the archive as the laboratory of the historian, because it is the primary location for historians to do their research. Students know about laboratories from their science lessons and they understand that this is where scientists do their research and experiments and make new scientific discoveries, and the same is true for historians and archives. More than anywhere else, archives are where historians discover new history and uncover things that were not known before, or provide new insights and understanding about things previously only partially known about. To understand history properly it is essential that students understand how historians construct their accounts about the past. And yet we know that many students have little or no real understanding about the process of historical enquiry.

GOV.UK

Home > Education, training and skills
> Inspections and performance of education providers
> Subject report series: history

Ofsted
raising standards
improving lives

Research and analysis

Rich encounters with the past: history subject report

Published 13 July 2023

Applies to England

Contents

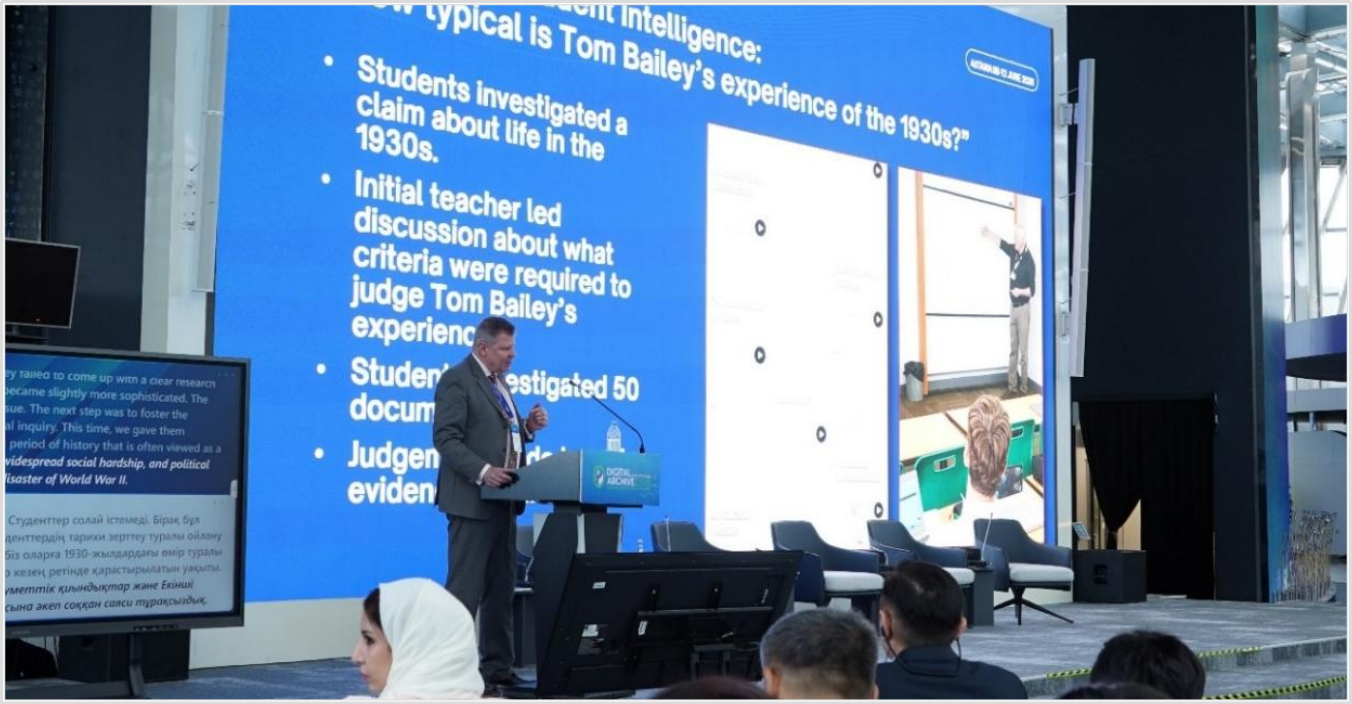
Context
Terms used in this report
Main findings
Discussion of findings
Recommendations
Primary
Secondary
Appendix

Figure 1. *Rich encounters with the past: History subject report. OFSTED*

"Rich Encounters With The Past" is a recent report produced by the school inspection service in the UK. OFSTED is the Office for Standards in Education, and every few years they do a deep dive into the teaching of specific subjects in schools at primary (elementary) and secondary (high school) level (OFSTED (Office for Standards in Education). 2023). The report on the teaching of history in schools in 2023 identified

a key finding that: “In most schools, pupils had misconceptions about how historians and others study the past and construct their accounts.”

This is understandable given the huge emphasis that has been placed on subject knowledge at the expense of disciplinary knowledge. The focus has been so much on teaching information to students that there has not been time for teachers to help develop an understanding of how history is done. The result is that many students do not truly understand the subject or how historians know about the past, which undermines their performance in the external examinations, which still ask questions about enquiry methodology, primary sources and historical interpretations.



As a result, OFSTED has now recommended that: “Schools should ensure pupils develop meaningful and increasingly complex knowledge about how historians and others study the past and construct their accounts.” This provides a great opportunity for archives to step in and support teachers and students with developing their understanding of how history is constructed. Archives are enormously powerful for education in a number of unique ways.

Firstly, they are repositories of activity. The collections that are held in archives are created in the process of people doing things. Whether they are the archives of state and government, or those of institutions, organisations and individuals, the documents which they hold represent the material output of the historical activity that historians seek to understand. They are a portal to that activity, and they provide a unique insight into the motives and beliefs, as well as the actions, of people in the past. It is essential that students understand how archival collections are created, and that they represent only a fraction of past activity and are often the result of intent on the part of the creators and the archivists who select what to keep. Understanding that the historical record is partial and not fully representative of groups of people is

important if historians are avoid merely replicating the norms and power structures of past societies in the histories that they construct.

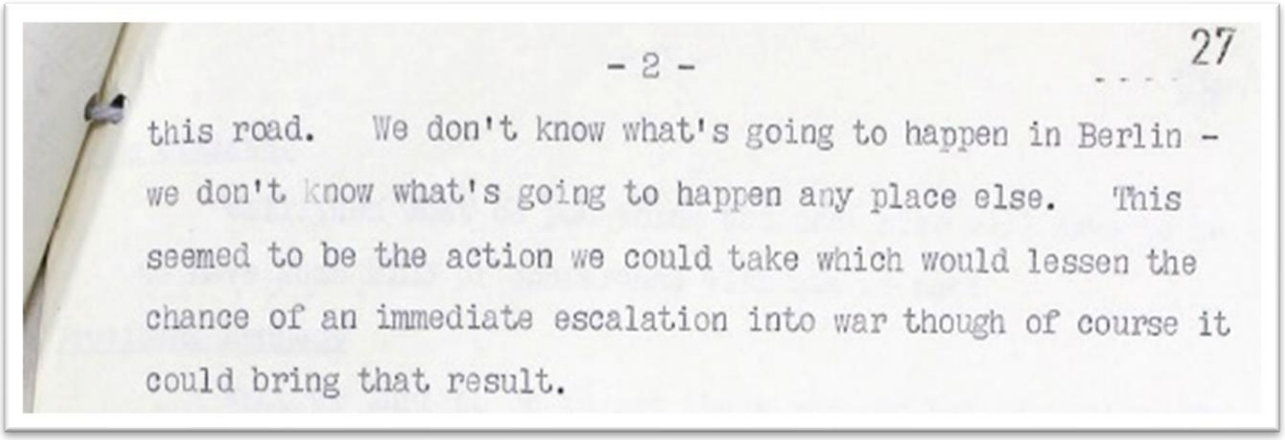
Secondly, archives offer a unique way of physically, intellectually and emotionally connecting with the past. Helping students to understand HOW the past is known is fundamental to their understanding of WHAT is known about the past. One particular workshop that The National Archives offers for 5 to 7 year olds about the Great Fire of London, which destroyed much of the medieval city in 1666. The students already know the intimate details about the fire: where it started; why it spread; how it was finally extinguished; and what the outcomes were; long before they come to the workshop. However, they are challenged to think about how these details are known. The workshop takes five key facts that they know and investigates the evidence that exists for those facts. During the workshop the students look at the Hearth Tax return, an inventive way that 17th century government assessed people's income based on the number of fireplaces in their houses. The bigger the house the greater the number of fireplaces, and so the more tax that had to be paid. The document they investigate includes the return for Pudding Lane where the fire began and shows Thomas Farinor's bakeshop with its 5 hearths and 1 oven listed, the source of the fire. The excitement that a 17th century tax document can elicit in 6 year olds as they see Thomas Farinor's name in the Hearth Tax return is truly wonderful.

***Figure 2**. Students inspecting the Hearth Tax return for Pudding Lane 1666, during a school workshop*

Thirdly, documents transport us to another present, rather than back, in time. History is not about the past, but rather it about the present (and the future) of people who were alive in the past. The documents that are created are produced in real time

as people are living through events. They do not have the benefit of hindsight and do not know how the story ends. In a transcript of the telephone conversation between JFK, John Kennedy, President of the United States and Harold Macmillan, British Prime Minister on the night of 23 October 1962 at the height of the Cuban Missile Crisis, Kennedy is reveals the full extent of what is happening and he, like everyone else, has no way of knowing how things will end. We know the resolution was peaceful but Kennedy did not. “We don’t know what’s going to happen in Berlin – we don’t know what’s going to happen any place else. This seemed to be action we could take that would lessen the chance of an immediate escalation into war though of course it could bring that result”.

Archives are a powerful reminder that whilst we read history backwards, the events happen forwards and nobody in the past knows the outcome/



- 2 - 27

this road. We don't know what's going to happen in Berlin - we don't know what's going to happen any place else. This seemed to be the action we could take which would lessen the chance of an immediate escalation into war though of course it could bring that result.

Figure 3. *Transcript of telephone conversation between Kennedy and Macmillan, 23 October 1962. TNA Cat Ref: (PREM11/3689)*

Finally, archives are powerful because they make students think. They challenge the preconceptions that they have about the past and they make them use the existing knowledge that they have about a topic to decode and understand what a document is telling them. This process of thinking is essential because it is the holy grail of education. If students are thinking, then they are learning. Daniel Willingham is Professor of Cognitive Psychology at the University of Virginia, and one of the leading experts on how the brain works. He has informed our understanding of how memories are formed and how long term and short term memory interacts and he is very clear on how long term memories are created; and that means learning. “Your memory is not the product of what you want to remember or what you try to remember; it’s the product of what you think about.” So, to help students learn they must be made to think, and archival documents are brilliant at making students think, if they are used effectively. That means using them in an authentic way to provide evidence to answer a historical enquiry question.

This is why The National Archives (UK) has had an education service for 25 years, working to curate the collection for teachers and students to use in the classroom, as well as to enable the delivery of workshops and events programmed for schools. The education section of the website has hundreds of different resources for students and

thousands of documents for teachers to use in their own resources. There are many different collections of documents curated on particular topics and periods for students to use in their research, particularly for those undertaking Advanced level History, which is the pre-university level studied by 16 to 18 year olds. Many of these collections can be challenging to use in the classroom with students so The National Archives (UK) has been thinking about how to support their use of these.

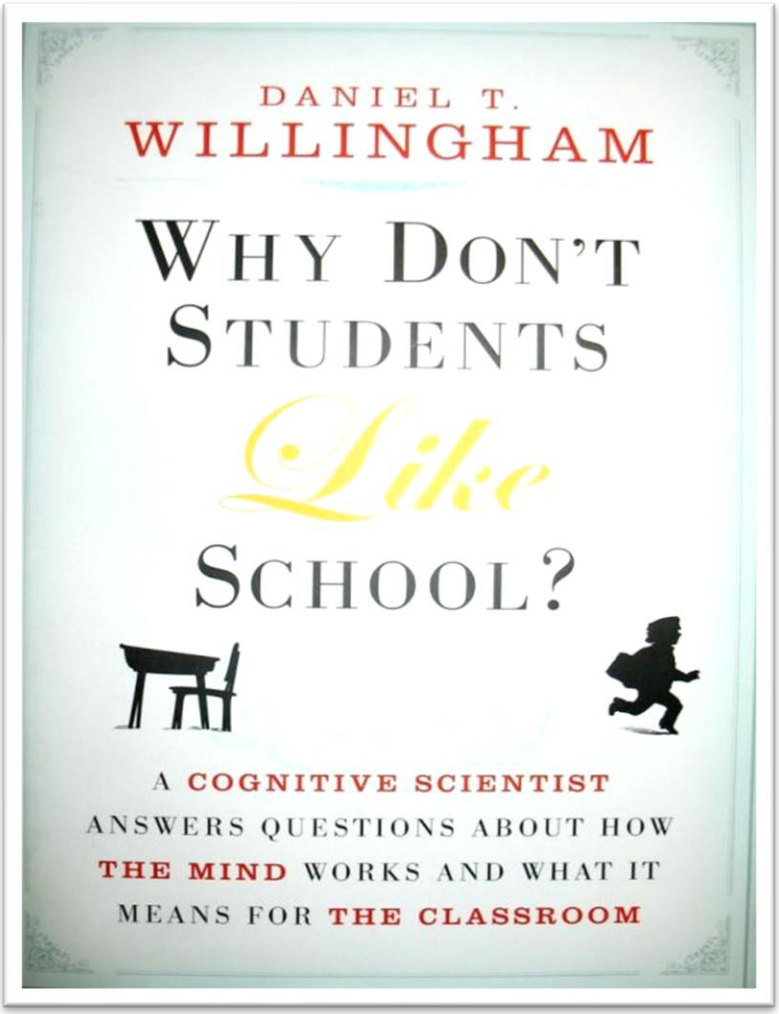


Figure 4. *"Why don't students like school?" Daniel T. Willingham 2009*

This is where AI could help, but the use of AI in education has a mixed reception. While it is seen as a solution to excessive workload and a means of quickly making learning resources and planning lessons, it is criticised by many for undermining student thinking and creating inaccurate and unethical material and experiences. However much of the criticism about AI in education stems from the fact that it is not used to address specific subject orientated problems. If AI is used indiscriminately and without real pedagogic thought then the results are inevitably going to be at best variable, and at worst damaging. This project has identified a specific problem that students have with History which we believe AI can not only help but actually enable better learning and outcomes.

Can AI can enable students to be better historical thinkers so they understand and evaluate historical interpretations by using and assessing historical sources? These are two requirements of the examination specification for Advanced level History which is sat by 18 year olds at the end of their school career prior to university. Exam boards write a report about the overall performance of each year's cohort of candidates and chief examiners consistently identify the two greatest weaknesses amongst students to be:

- Understanding and evaluating historical interpretations.
- Using and assessing a range of historical sources.

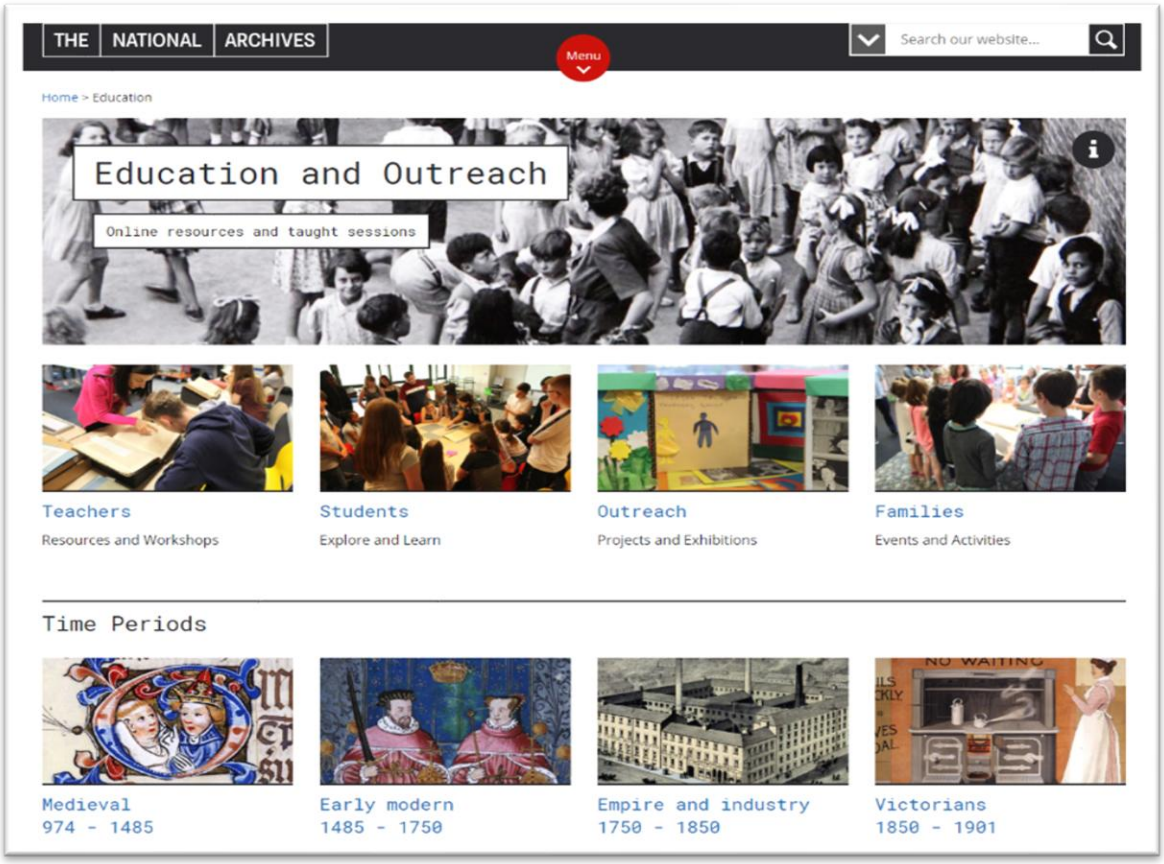


Figure 5. *The National Archives Education.* www.nationalarchives.gov.uk

Our aim was to test whether it would be possible to support students to become better at responding to these assessment criteria by developing their thinking about historical enquiries and their use of evidence from documents.

In July 2025 Andrew Payne worked with a group of Year 12 (Grade 11) students at Skegness Grammar School in Lincolnshire with the support of the History and Information Technology (IT) teams at the David Ross Education Trust (DRET) that runs the school. As part of their examination syllabus the students had been studying British History from 1906 to 1957, looking at the social, economic and political development of the country during the first half of the Twentieth Century. The students had nearly completed their first year of study so had good subject knowledge and the IT team at DRET were keen to test out the development of a new AI system for the school. Access to the class was provided for a whole day (approximately 5 hours

of lessons) which would allow progressively ambitious levels of activity to be undertaken which, it was hoped, would enable the development of their historical thinking.

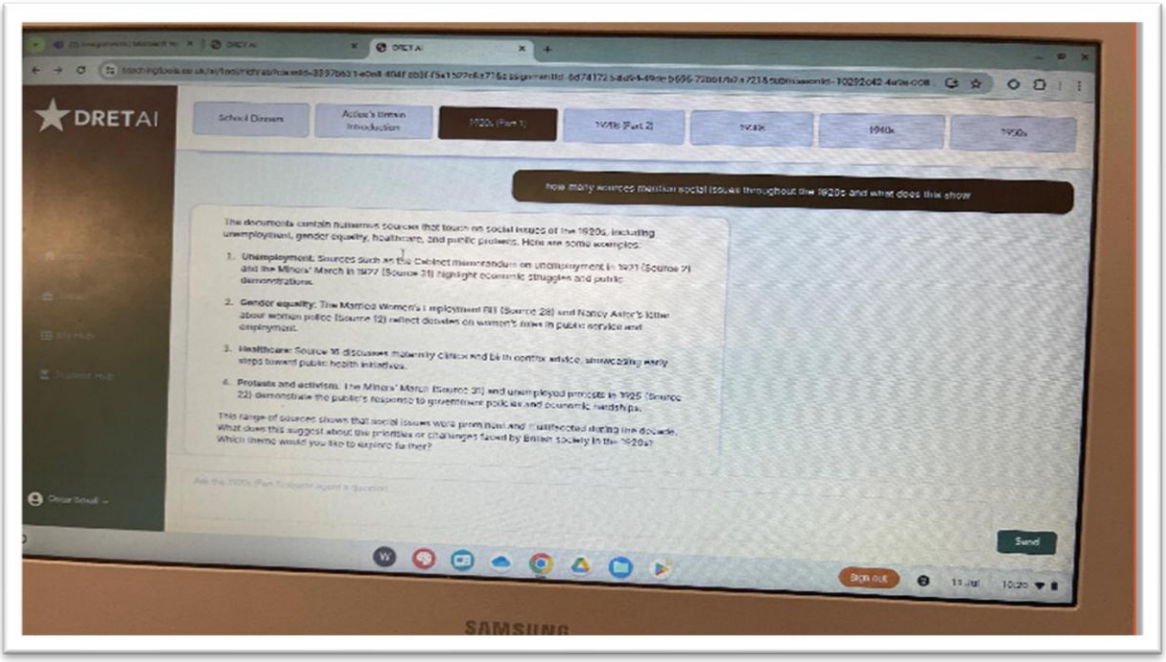


Figure 6. *Close up of the student interface with each of the collections accessible via the tabs at the top*

The students used Google Chromebooks running DRET AI which is an adapted version of MS Copilot (Microsoft). The system is a walled garden, meaning it cannot access the wider internet and can only access material that was provide. 5 collections of documents were loaded as PDF files covering the early 1900s, 1920s, 1930s 1940s and 1950s, with approximately 50 documents per collection (250 documents in total).

Most of the documents were text, often several pages long. There were also some photographs and some diagrams. The plan was to get the students to write prompts for the AI in the form of questions about the documents and for the AI then to respond with information that it found in the documents. This would enable the students to investigate the large number of documents without having to directly read all of them. The AI would only answer in response to the students questions so if they asked poor questions it would give poor answers. AI would not do the thinking. That would be the job of the student.

The session began simply looking at just a handful of documents about the introduction of free school meals for children in Bradford, an industrial city in the north of England, in 1907. This included some reports about the pilot project, photographs of the school children that were fed and a beautiful chart showing the impact of the free meals on the weight of the children. The chart shows how their weight increases rapidly during term time when they are fed at school and drops during the school holidays when they are at home and do not receive a meal. How would the group of students react to these documents? What kinds of questions would they ask? Would

they formulate any kind of enquiry or set any criteria for gathering evidence. This allowed a bench mark for the students own ability to be set before any intervention occurred.

Perhaps as was to be expected the students asked very generic, low level questions without much thought or planning about the purpose of the enquiry and no criteria set for the evidence they were seeking. For example, they asked questions like:

- What is source 1?
- When was this source created?
- Who created this source?

These are very low level "source analysis" questions which would expected from younger students but it showed that even these "advanced" level students were not really thinking historically about how to make use of the documents as sources of evidence. However, this was good as it set a benchmark for their level of thinking at the start of the day.

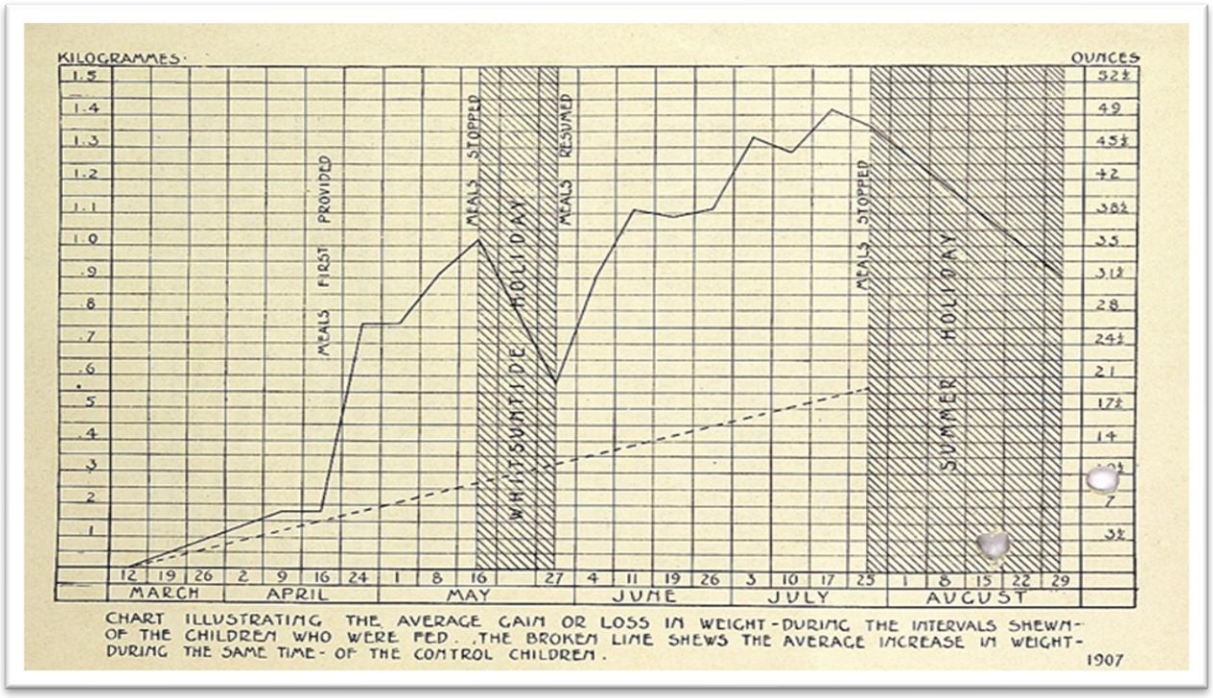


Figure 7. *Chart from Bradford School Dinners Report, 1907 TNA Cat Ref: ED 50_8*

The students then progressed to looking at a collection of 50 documents about the 1920s. Once again, no specific enquiry question was set, but it was suggested they focus on one area of life during that decade; either social, economic or political. We wanted to see if they would formulate any type of enquiry question or establish any criteria for gathering evidence. One change that was made was to provide the AI with a framework for analysing the documents when responding to the students' questions. This is based on a framework that students are already familiar with so it should help them think about the questions they ask.

RICHRAB is an acronym to help students think about what value a documents has as a source of evidence for the following:

R – Reasons: does it tell us about why something happened?

I – Impacts: does it tell us about the impact of an event or action?

C – Changes: does it help us understand what changes were happening?

H – How things worked: does it provide insight into how a society, institution or organisation operated?

R – Reactions: does it tell us how people reacted to events or changes?

A – Attitudes: does it reveal the attitudes that people had towards events or other groups of people?

B – Beliefs: does it reveal anything about the beliefs that people held or their understanding of events at the time?

We help them remember this with the picture of a Rich Rabbit!

But would they use it to guide their questioning of the documents?

Once they had gathered their responses from the AI we wanted them to write a paragraph to explain what they had discovered. Again they have previously been taught the idea of a CASE paragraph. This consists of:

C – Claim: make a historical claim about an event or period or person.

A – Argument: explain your argument for this case.

SE – Supporting Evidence: provide evidence to prove that your claim is valid.

So how would they do when confronted with a collection of 50 documents but no guidance about formulating an enquiry or gathering evidence?

While the AI provided more analytical responses the students still asked low level questions that did not really elicit much evidence and they did not formulate any clear enquiry question which they wanted to answer. So while the AI got a little more intelligent the students did not!

The next stage was to develop the students' thinking about historical enquiry. They were given a claim about life in the 1930s; a period of history that is frequently seen as one of economic collapse, great social hardship and political instability leading to the eventual disaster of the Second World War. The claim was based on a statement from the author's grandfather, Thomas Bailey, who always said it was the best decade of his life. It is a claim that challenges the traditional perspective, so students were set the question: "How typical is Tom Bailey's experience of the 1930s?". Before the students engaged with the documents their teacher led a discussion about what criteria would be required to judge his experience. Students suggested it would be necessary to know about levels of employment and economic activity, standards of living, how homogonous or diverse the population was and whether there was any political unrest in Tom Bailey's town compared to other parts of the country. The students were beginning to think about the criteria they would need to base their question prompts on for the AI. They were then given access to 50 documents about life in the 1930s to gather evidence from using the AI, and they had to write a CASE paragraph at the end.

This time the students' questions were much more focused on evidence gathering. They asked much more thoughtful questions and also asked follow-on questions about the response that the AI gave them. For example one student said:

"List the documents exploring the standard of living in the 1930s."

The AI then listed the documents that provided evidence for this criterion. The student then followed up with another question:

"Explore how document 19 (Going Short) reveals the standard of living in Britain in the 1930s, including the limitations of the source."

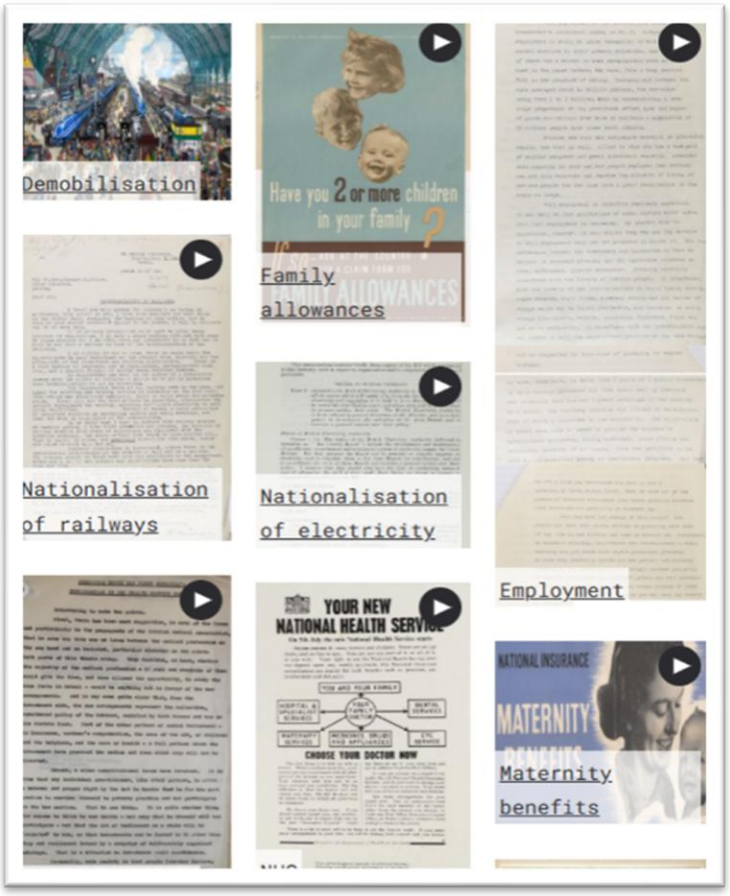


Figure 8. *A small selection of the document collection for the 1940s activity*

The AI then provided detailed evidence from document 19 about various aspects of living standards with direct quotes from the document.

Students then gathered this evidence together and wrote their own response assessing Tom Bailey's claim. The AI then assessed how effectively these responses were, using a rubric based on A level marking schema, and provided feedback on:

- The strength of argument
- Areas that could be improved
- A response level grade based on the examination rubric.

In the afternoon session the group were set a final challenge to test a historian's interpretation about the Labour Government of 1945 to 1951. This was a socialist government that created our welfare state and National Health Service, reformed the education system and undertook a programme of nationalisation of the major industries and other key parts of the economy.

We began by giving the students an interpretation by the historian Dr Robert Crowcroft who wrote a short assessment of the successes and failures of the government. Students read through a printed copy of this and highlighted the various claims made, as we can see in the picture. This was to ensure they had a good understanding of Crowcroft's argument. They then tested specific claims by setting

the AI prompts to investigate the collection of documents to see if there was evidence to support or refute the claims that were made. The students then wrote an assessment of Crowcroft's claim and asked the AI to evaluate and provide feedback on their response.

As a first pilot study it was anticipated there to be problems with the technology and so a lot of technical support was available in the room during the day. Without this the whole thing would have broken early on and it would have been a wasted day. Specific problems included:

- DRET AI initially struggled to work with so many documents because we had not trained it on them first. This meant the system was not powerful enough at first to cope.
- The number of available tokens for the AI needed to be increased (provide it with more power) so that it could read all of the documents and remember what it had read.
- This helped initially in the morning with simple questions.
- But… as students' questions became more demanding the AI began to fail again. It simply could not remember all of the information from the documents so it began to hallucinate.
- AI gave incorrect answers and students began to doubt the accuracy of some responses. They challenged the AI more and more, which was a good thing.
- This encouraged them to ask to see the original documents, so that they could read them for themselves and check if what the AI was saying was factually accurate.

Here is an example of where the AI began to forget where it had read something and so it just guessed about the documents, but the student challenged the AI.

Student: “Which source is the letter from the railway worker?”

AI: “The letter for the railway worker to Attlee is in Source 3…”

Student: “That is incorrect.”

AI: “The letter from the railway worker to Prime Minister Attlee is found in Source 25….”

The outcomes of this initial pilot study were very promising:

- Students really enjoyed the experience despite the challenges.
- They rapidly learned that AI needs clear instructions.
- This developed their understanding of how to structure historical enquiries by establishing criteria for gathering evidence.
- They actively managed the AI to engage with large numbers of documents and gather evidence.
- They became more effective at asking better questions and constructing substantiated historical interpretations.
- They developed a better understanding of the limits of AI and the need to check AI-generated responses.

Initial conclusions from the day were as follows:

- AI has huge potential to support students in their independent use of large numbers of original documents.
- AI needs a framework to evaluate documents effectively (RICHRAB).
- Students need teacher input to develop effective criteria and questions to instruct the AI.
- AI in the classroom requires a LOT of technical support to help overcome operating issues.
- Creating a dedicated Large Language Model (LLM) trained on documents from the collection in advance of the classroom activity would overcome the limitations encountered by a lack of tokens (AI memory capacity).

At the end of the day students were asked to complete an evaluation and five of them were taken aside to conduct short interviews. The final question was the key point of the whole investigation.

“Have you got better at thinking about the questions you need to ask of documents?”

The 5 students provided answers of varying detail.

Oscar said: “Yes!”. When if he would like to elaborate on that he replied: “No!” A typical 17 year old boy’s response.

Freddie and Wesley gave more explanation: “I’ve learned you have to give it [AI] a really nailed down version of what you want to know.” said Freddie, and Wesley commented: “By the end I was asking more detailed questions.”

The girls were more effusive, and both were very positive about the experience. “It helped in developing my thinking… and using a greater number of documents. AI helps in understanding what to ask the documents.” said Fatima. And Imogen stated: “It makes using documents less intimidating… I’ve realised that you have to be specific about what questions to ask.”

So it was a positive experience and most importantly it helped students to think more deeply about the process of historical enquiry and working with documents to gather evidence.

We were pleased with the results of the initial pilot day and decided it was worth trying to develop the platform further, especially as the power of AI has increased significantly over the past year. There were 5 main issues that we wanted to address in the second phase of the project:

- Students felt constrained by the RICHRAB framework. Even though they found it helpful initially they felt it would hold them back in the future.
- Students wanted to see original documents, especially as they did not trust the AI fully and they wanted to be able to check the accuracy of its responses.
- Insufficient tokens (memory capacity) caused hallucinations, so the AI needed to be trained on the documents in advance to create a Large Language Model (LLM)
- Lots of IT support was required because the original DRET AI platform was not very powerful and not very stable.
- Lots of teacher input and support was required to provide guidance and feedback. We hoped it would be possible to automate more of the guidance and provide feedback more quickly.

A new approach was taken with the development of a completely new tool using Anthropic's Claude (Anthropic). This was populated with the documents in advance to create an LLM which would reduce the number of tokens required. Both short extracts and full transcripts of the documents were added, plus the ability to view the original documents if required. An AI Tutor interface was developed to guide students in finding evidence and to provide feedback without providing the answers. In addition, we developed three specific activities for students to choose from:

- RICHRAB categoriser – where they could decide on what type of evidence a document may yield based on the RICHRAB acronym.
- Test the statement – where we provided a number of different statements about the topic under investigation and students had to investigate these using the documents and decide if they was true or false.
- Enquiry investigation – students were provided with a claim about the topic and they had to define the criteria and gather evidence before writing their response. This was similar to the previous activity in stage 1 where they evaluated the historical interpretation.

This second phase also investigated whether AI could help younger students work with more challenging documents, so two groups of students were chosen: the original group from phase 1 who were now in Year 13 (Grade 12); and a new younger group who were in Year 10 (Grade 9). This latter group were studying the reign of Elizabeth I (1558 – 1603) so the documents used had archaic language which would be another test for the AI, as well as for the students. The Test the Statement activity was the main focus of this enquiry which required them to select a statement and test whether the claim it made was accurate or not by gathering evidence from the documents. Once again, they were required to write a CASE response with the evidence they had discovered.

Some of the statements that they were asked to assess included:

Statement 1 – Elizabeth wavered in her Protestant faith.

Statement 2 – Elizabeth's relationship with Robert Dudley was purely professional

Statement 3 – Elizabeth always made quick decisive military decisions when England was threatened.

Students could then select from 21 documents to investigate for evidence relating to their chosen statement. They could ask AI to help them understand the document and to gather evidence but the AI would not provide a final answer about the accuracy of the statement. That would always be for the student themselves to write, based on the evidence they had gathered.

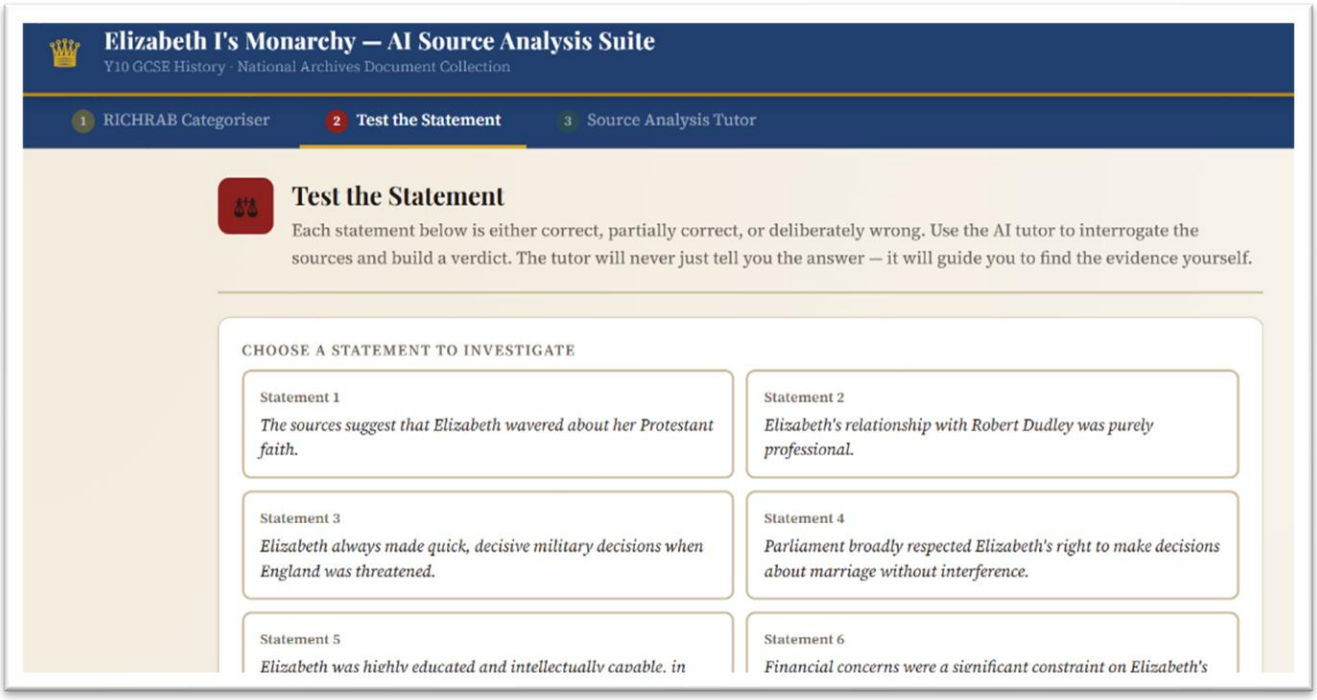


Figure 9. *Screenshot of Test the Statement interface*

The platform initially showed them an extract from the document but they could also see a full transcript and an image of the original document.

Students constructed paragraphs to test the claims by selecting the sources they thought would help them most. Here the student has written about the relationship between Elizabeth I and her cousin, Mary, Queen of Scots, based on a number of letters between them. “Her pleading tone indicates her compassion towards her cousin and how she believes that she may have a possibility of persuading Mary away from her death penalty”.

The AI then responds with an assessment and suggestions for improvement: “That’s a well-structured paragraph… consider briefly reiterating the complexity of Elizabeth’s feelings beyond just respect… mention there were undelying political motives that could complicate their relationship.”

Here the AI provided feedback with suggestions for improvements and encouraged more detailed responses…

“You might want to clarify that the phrase “upon good consideration”… emphasises her cautious nature in decision-making.”

Students responded to feedback rooted in familiar approaches to source analysis. Here the student writes a CASE response:

Claim – Elizabeth's relationship with Robert Dudley was not purely professional.

Argument – Evidence shows they had a deeper attachment.

Supporting Evidence – In Leicester's (Robert Dudley's) last letter to Elizabeth he says that Elizazbeth's health is "the chiefest thing in the world that I do pray for". This suggests that Leicester cares deeply for her… additionally, Elizabeth kept this letter in a casket by her bed until her death" (Last letter).

It helped students to refine their thinking… Here the student contradicts themselves by stating that Elizabeth was ruthless and unsentimental but then provides evidence of her empathy and emotional connection. The AI suggests a small revision in the concluding sentence "Therefore, while some sources may suggest that Elizabeth was ruthless, many others point to her caring nature and willingness to form close relationships with those around her." It then explains that: "This revision keeps your ideas intact but clarifies the contrast you've pointed out."

It encouraged students back to the documents… Here the AI acknowledges the student's claim that it is "balderdash" (rubbish) to call Elizabeth a ruthless tyrant but notes that they have not provided any supporting evidence. "That's a strong opening!" says the AI, "You've clearly stated your position against the statement while highlighting Elizabeth's respectful and affectionate attitudes. As you continue, consider including specific examples or sources that support your view, such as reference to her letters with Leicester or her concern for Mary, Queen of Scots. This will strengthen your argument further."

The AI also ensured students considered their line of argument in relation to their chosen claim. Here it suggests the student needs to cite specific evidence and add a concluding argument.

Most importantly it provided clear and useful scaffolds but didn't do the thinking for them. Here the student asks how to write a CASE paragraph about Leicester's affection for Elizabeth. The AI offers a suggestion for the structure without actually writing the paragraph for the student.

"To frame a CASE paragraph effectively, start with a clear topic sentence that states your main point… Then, integrate the evidence, such as the quote… and explain its significance…. Lastly, conclude by linking back to the statement." It then asks the student to expand on this in their own words.

The students gave positive feedback about the experience.

Jack – "AI in history can help you but not give you the answers out right."

Kobe – "It made me think more deeply and as a result achieving a better CASE paragraph."

Isabelle – "It helps deepen my understanding… making me think about what I am writing despite my previous reservations."

Lexi – "The AI encourages you to question different parts of the source and find evidence to support your claims, and when you find something good the AI praises you which is nice. It helped to steer me in the right direction without telling me the answers which encouraged me to actually think about it."

But they also had some complaints…

Joshua – "AI was obsessed with RICHRAB!"

Keanno – "This is where the AI was not very useful as it didn't help me [with]… what to analyse about it."

Sami – “The AI is unable to translate the texts as it focuses too much on attempting to achieve an answer from the person, making it difficult to continue analysis.”

Several students were also disappointed that the AI would not write the answer for them which they clearly expected it to do based on their experience of using AI platforms for homework.

The class’s History teacher, Laura Howey, gave positive feedback about how she felt the students interacted with the AI and the documents.

- Students were engaging with historical processes to see that historians construct and test their claims. They were thinking historically!
- They were using original documents to find evidence to support or challenge claims.
- They received instant feedback which was well structured and designed with our school systems in mind (CASE and RICHRAB).
- They produced some strong arguments to challenge the claims.
- They enjoyed working with documents.

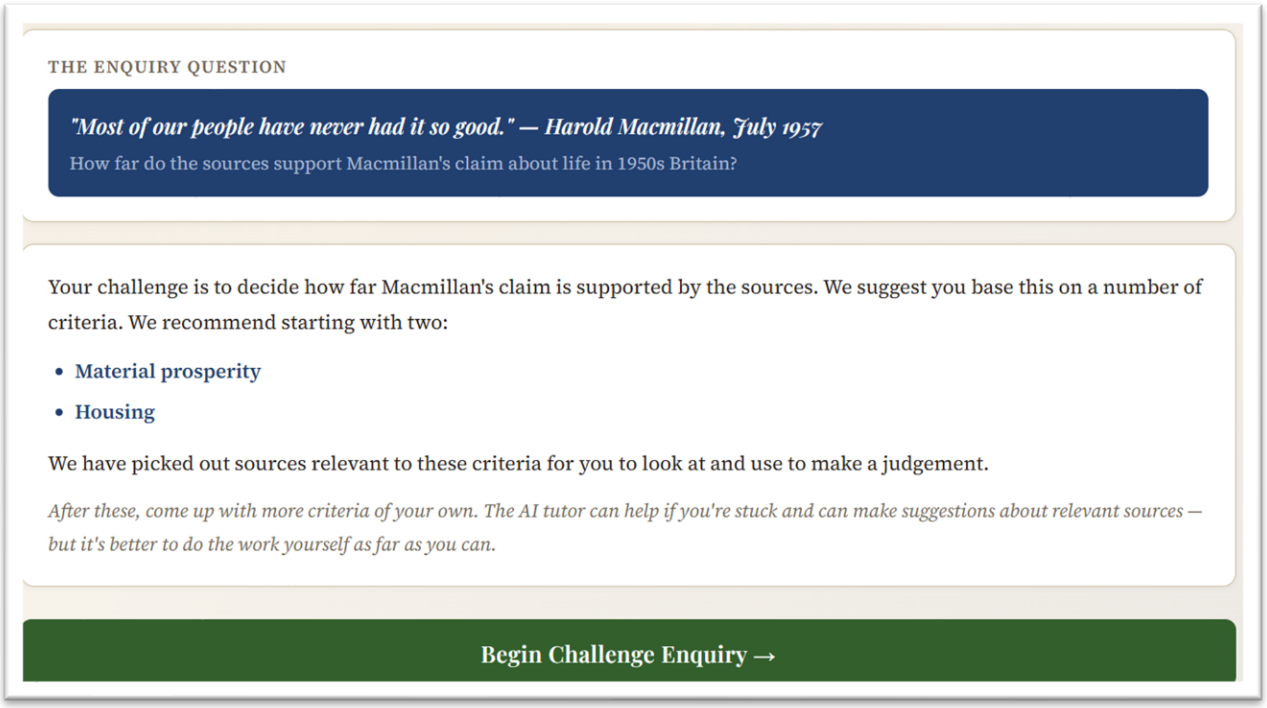


Figure 60. *The Enquiry Question with suggestions for two criteria to gather evidence about*

The final phase of the project was to return to the A level students that had done the initial sessions in 2025. It was an opportunity to see how they would compare this new system with the original DRET AI version that we had used with them before. This time the focus was the Macmillan government of the 1950s and provided a collection of documents to help them investigate the claim that in the 1950s “The people of Britain have never had it so good.” The AI also proposed two criteria which students should look for evidence about and then they could suggest further criteria to investigate.

The Enquiry Question interface showed the statement made by Harold Macmillan, Prime Minister, in July 1957 where he claimed: “Most of our people have never had it so good”.

It then suggested two criteria to start the investigation

- Material prosperity
- Housing

For each of the criteria the AI Tutor preselected some documents to look at and explained that it could help and make suggestions but not provide the answers.

Once again, it provided feedback to help students reflect on their work. This student writes about the housing situation and how there is a contrast shown between different sources of evidence. The AI responds at the bottom by saying: “In summary, you’ve compared the optimistic view of housing in Lansbury against the challenges highlighted in the “Housing Shortages” source. How would you summarise the overall perspective on Housing based on the sources you’ve examined? Do you think they collectively support or challenge Macmillan’s claim, considering the complexities you’ve discussed?”

Students engaged well with the AI Tutor and found it an enjoyable interface that responded positively to them. However, we began to note a problem as their discussions developed.

Increasingly both the student and the AI Tutor got into a cycle of discussion that did not refer to the documents.

When the student made a claim but failed to provide any supporting evidence the AI should have asked them to find a document that provided the evidence. However, we noticed that this was not the case. Here the student states that there was consensus between the two main political parties about economic policy in the 1950s but gives no supporting evidence. “Butskellism highlights a post-war consensus of economic polices,” states the student. “That’s a great insight!” replies The AI Tutor. “Butskellism indeed reflects the post-war consensus and suggests that there was a level of satisfaction with economic policies.

”However, it failed to refer the student back to the documents to get specific evidence to support their claim.

Both students and AI completely ignored directly relevant documents and the responses by the AI Tutor increasingly became a parroting back of what student had said. A recognised fundamental problem with the nature of AI is that it often seeks to please the user rather than challenge them.

Overall though the feedback from students was positive although they did raise some criticisms.

Lucus was positive about the support the AI Tutor gave in using the documents: “The AI is good at encouraging educated and expanded answers, alongside motivating for more developed answers. It is effective in its ability to analyse the sources and make these sources more easily understood.”

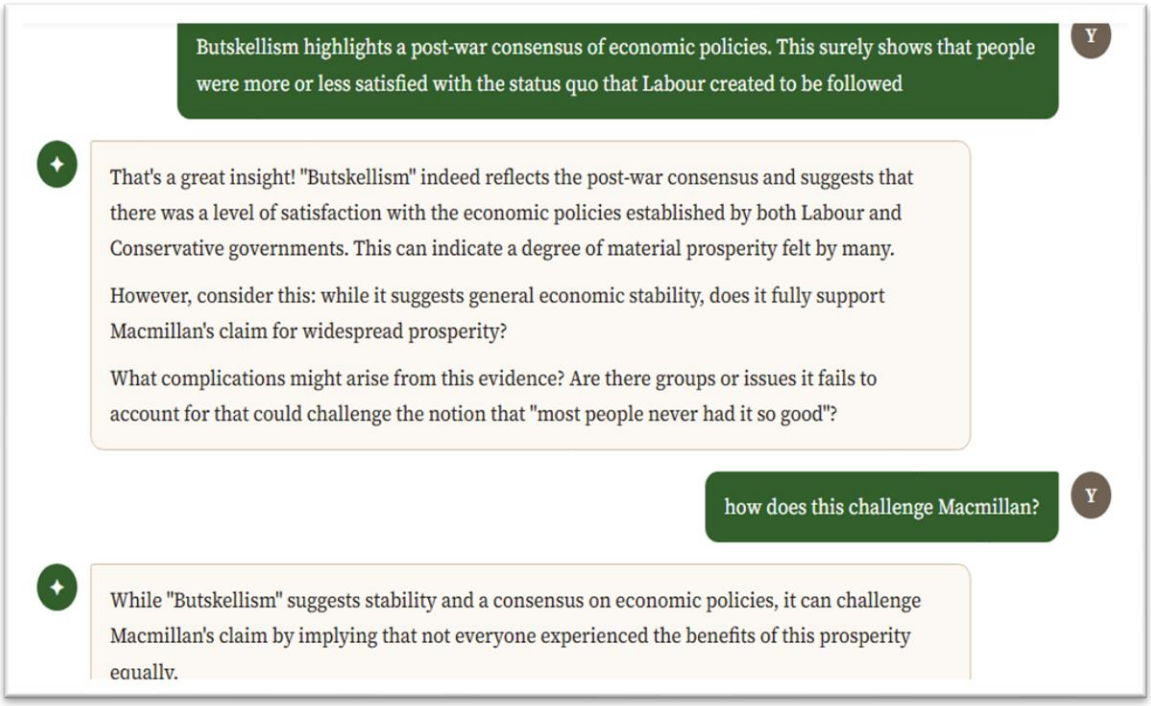


Figure 11. *Student and AI Tutor ignore the documents and develop a feedback loop of their own*

Caerla noted that while it had improved on the previous version she felt it was too structured: “It worked better… I liked the fact that it was structured, but it felt a little too structured, it should let you be more wide ranging to let you explore more deeply.”

Freddy identified a lack of proper challenge by the AI Tutor: “The only criticism could be that it did not argue about my points and may have taken them for granted. It offered very little opposition.”

Alfie also identified the problem of the AI Tutor failing to direct the students back to the documents: “It focused more on asking follow up questions than on getting you to analyse the sources.”

So, to conclude, these pilot studies have been really interesting and we have learned a lot.

- AI can definitely help students develop their historical thinking. They were much more aware of how to construct an enquiry, the need to set criteria for evidence gathering and how to ask effective prompt questions.
- It is essential to identify an issue specific to History education that it can help with i.e. developing historical thinking about sources and interpretations. This keeps everyone focused on the purpose of using AI rather than using it simply in a generic way to re-present information that students already have.
- AI is Artifical but it is NOT Intelligent!
    - It can help students analyse large numbers of documents quickly in response to effective prompt questions.
    - It can provide feedback and support to encourage reflective thinking.
    - It knows but it cannot understand or feel.

    - It can provide evidence but no historical empathy.
- AI can support students but it cannot replace their teachers.

And finally, it was inadvertently shown that not all change is progress. While the CLAUDE based platform definitely provided feedback to the students and was more robust, it did not consistently do what was intended, which was to refer the students to the documents and help them to analyse them for evidence. Developing an effective platform is an iterative process and needs to be tested in the classroom in order to identify unintended consequences. We shall continue to develop this and will report on progress in the future.

- To download the resource scan the QR Code or use the Bitly link.
- This will open a shared Google Drive Folder with two html files.
- Right click on each file and select download.
- Click on one of the html files in your download folder to make it run in a browser.
-

References

Anthropic 2026 – Anthropic. Claude Models and Capabilities: Documentation / Anthropic. 2026. URL: https://docs.anthropic.com/claude/reference (accessed 14.07.2026).

Crowcroft Robert – Crowcroft Robert. *Historical assessment of the Labour Government 1945-1951*.

Hearth Tax returns – National Archives Documents referenced: Hearth Tax returns (17th Century).

Last letter – National Archives Documents referenced: Last letter from Robert Dudley, Earl of Leicester, to Elizabeth I.

Letters – National Archives Documents referenced: Letters between Elizabeth I and Mary, Queen of Scots.

Macmillan Harold 1957 – Macmillan Harold. *Speech at Bedford, 20 July 1957*. ("Most of our people have never had it so good").

Microsoft – Microsoft. Microsoft Copilot Documentation and Educational Use Cases / Microsoft. 2025. URL: https://learn.microsoft.com/en-us/copilot/ (accessed 14.07.2026).

OFSTED 2023 – (Office for Standards in Education) – OFSTED (Office for Standards in Education). History subject report: Research and analysis / OFSTED. London: Government of the United Kingdom, 2023. URL: https://www.gov.uk/government/publications/subject-report-series-history/history-subject-report (accessed 14.07.2026).

The National Archives (UK) – The National Archives (UK). Education Service and Online Resources. URL: nationalarchives.gov.uk/education (accessed 14.07.2026).

Transcript of telephone conversation 1962 – National Archives Documents referenced: Transcript of telephone conversation between JFK and Harold Macmillan (23 October 1962).

Willingham 2009 – Willingham, D. T. (2009). Why Don't Students Like School? A Cognitive Scientist Answers Questions About How the Mind Works and What It Means for the Classroom. Wiley. URL: https://books.google.kz/books?id=8SDs8LZl41EC&printsec=frontcover&hl=ru#v=onepage&q&f=false (accessed 14.07.2026)

«АРХИВ БОЛАШАҒЫ:
ЖАҢА ҚҰЗЫРЕТТЕР МЕН КӨШБАСШЫЛЫҚ»
№ 4 сессия

Сессия № 4
«АРХИВНОЕ БУДУЩЕЕ:
НОВЫЕ КОМПЕТЕНЦИИ И ЛИДЕРСТВО»

«THE ARCHIVAL FUTURE:
NEW COMPETENCIES AND LEADERSHIP»
Session No. 4

AI-SANA: ТРАМПЛИН НОВЫХ КОМПЕТЕНЦИЙ

Динара Ринатовна Щеглова
Вице-министр науки и высшего образования Республики Казахстан
d.almasova@sci.gov.kz

Аннотация. Цифровая трансформация экономики и ускоренное развитие технологий искусственного интеллекта формируют новые требования к системе подготовки кадров и содержанию профессиональных компетенций. В Республике Казахстан развитие искусственного интеллекта рассматривается как стратегический государственный приоритет, охватывающий модернизацию высшего образования, научных исследований и цифровой инфраструктуры.

Особую роль в формировании национальной AI-экосистемы играет программа AI-Sana, направленная на развитие компетенций в области искусственного интеллекта, поддержку студенческого технологического предпринимательства и создание прикладных решений для приоритетных отраслей экономики. Программа способствует переходу от базового освоения AI-инструментов к практической разработке цифровых продуктов и отраслевых решений.

В условиях цифровизации особую актуальность приобретает внедрение технологий искусственного интеллекта в библиотечное и архивное дело. Развитие цифрового архивоведения, интеллектуальных систем поиска, анализа и обработки архивных данных, автоматизация библиотечных процессов и управление знаниями открывают новые возможности для повышения эффективности работы с информационными ресурсами.

Применение искусственного интеллекта в архивной отрасли требует подготовки специалистов нового поколения, обладающих междисциплинарными знаниями на стыке архивоведения, информационных технологий и анализа данных. Программа AI-Sana выступает важным инструментом формирования таких компетенций и создает основу для устойчивой цифровой трансформации архивной сферы.

Ключевые слова: искусственный интеллект, Национальная AI-экосистема; программа «AI-Sana»; обучение; развитие компетенций; драйвер; цифровая трансформация; межвузовский стандарт; инновационные программы; подготовка

высококвалифицированных специалистов; прикладные решения; разработка проектов; ИИ в отраслях

AI-SANA: ЖАҢА ҚҰЗЫРЕТТЕРДІ ДАМЫТУДЫҢ НЕГІЗІ

Динара Ринатовна Щеглова
Қазақстан Республикасы ғылым және жоғары білім вице-министрі
d.almasova@sci.gov.kz

Аңдатпа. Экономиканың цифрлық трансформациясы және жасанды интеллект технологияларының жедел дамуы кадрлар даярлау жүйесіне және кәсіби құзыреттер мазмұнына жаңа талаптар қалыптастыруда. Қазақстан Республикасында жасанды интеллектті дамыту жоғары білім беруді, ғылыми зерттеулерді және цифрлық инфрақұрылымды жаңғыртуды қамтитын стратегиялық мемлекеттік басымдық ретінде қарастырылады.

Ұлттық AI-экожүйені қалыптастыруда жасанды интеллект саласындағы құзыреттерді дамытуға, студенттердің технологиялық кәсіпкерлігін қолдауға және экономиканың басым салалары үшін қолданбалы шешімдер әзірлеуге бағытталған AI-Sana бағдарламасы ерекше рөл атқарады. Бағдарлама AI құралдарын базалық меңгеруден цифрлық өнімдер мен салалық шешімдерді практикалық әзірлеуге көшуге ықпал етеді.

Цифрландыру жағдайында жасанды интеллект технологияларын кітапхана және архив ісіне енгізу ерекше өзектілікке ие болуда. Цифрлық архивтанудың дамуы, архивтік деректерді іздеу, талдау және өңдеудің интеллектуалды жүйелерін енгізу, кітапханалық үдерістерді автоматтандыру және білімді басқару ақпараттық ресурстармен жұмыс тиімділігін арттырудың жаңа мүмкіндіктерін ашады.

Архив саласында жасанды интеллектті қолдану архивтану, ақпараттық технологиялар және деректерді талдау бағыттары тоғысындағы пәнаралық білімдерге ие жаңа буын мамандарын даярлауды талап етеді. AI-Sana бағдарламасы осындай құзыреттерді қалыптастырудың маңызды құралы ретінде архив саласының тұрақты цифрлық трансформациясына негіз қалайды.

Түйін сөздер: жасанды интеллект; Ұлттық AI-экожүйе; «AI-Sana» бағдарламасы; оқыту; құзыреттерді дамыту; драйвер; цифрлық трансформация; жоғары оқу орындары арасындағы стандарт; инновациялық бағдарламалар; жоғары білікті мамандар даярлау; қолданбалы шешімдер; жобаларды әзірлеу; салалардағы жасанды интеллект

AI-SANA: A SPRINGBOARD FOR NEW COMPETENCIES

Dinara R. Chsheglova
Vice-Minister of Science and Higher Education of the Republic of Kazakhstan
d.almasova@sci.gov.kz

Abstract. The digital transformation of the economy and the accelerated development of artificial intelligence technologies are creating new requirements for workforce training systems and the content of professional competencies. In the Republic of Kazakhstan, the development of artificial intelligence is regarded as a strategic national priority, encompassing the modernization of higher education, scientific research, and digital infrastructure.

The AI-Sana program plays a particularly important role in shaping the national AI ecosystem. The program is aimed at developing competencies in artificial intelligence,

supporting student technological entrepreneurship, and creating applied solutions for priority sectors of the economy. It facilitates the transition from basic AI literacy to the practical development of digital products and industry-specific solutions.

In the context of digitalization, the implementation of artificial intelligence technologies in library and archival studies is becoming increasingly relevant. The development of digital archival studies, intelligent systems for searching, analyzing, and processing archival data, the automation of library processes, and knowledge management create new opportunities for improving the efficiency of working with information resources.

The application of artificial intelligence in the archival sector requires training a new generation of specialists possessing interdisciplinary knowledge at the intersection of archival studies, information technology, and data analytics. The AI-Sana program serves as an important instrument for developing such competencies and establishes the foundation for the sustainable digital transformation of the archival sector.

Keywords: artificial Intelligence; National AI Ecosystem; AI-Sana Program; Education; Competency Development; Driving Force; Digital Transformation; Inter-University Standard; Innovative Programs; Training of Highly Qualified Specialists; Applied Solutions; Project Development; AI in Industries

Введение

Цифровая трансформация охватывает все сферы общественной жизни, включая архивное и библиотечное дело. Развитие искусственного интеллекта (ИИ) оказывает существенное влияние на рынок труда, содержание профессиональных компетенций и подходы к подготовке кадров. В современных условиях востребованными становятся специалисты, обладающие междисциплинарными знаниями на стыке отраслевой экспертизы, цифровых технологий и анализа данных.

В Казахстане развитие искусственного интеллекта рассматривается как один из ключевых факторов формирования новой экономики, основанной на знаниях, данных и инновациях. В соответствии с поручением Главы государства, 2026 год объявлен Годом цифровизации и искусственного интеллекта, что отражает стратегический курс страны на ускоренную технологическую трансформацию.

Государственная политика Казахстана в сфере искусственного интеллекта

В Казахстане последовательно формируется системная государственная политика в сфере искусственного интеллекта. Принят Закон «Об искусственном интеллекте», вступил в силу Цифровой кодекс, ведется разработка стратегии Digital Qazaqstan. Созданы Министерство искусственного интеллекта и цифрового развития, Цифровой штаб при Премьер-Министре, а также Совет по развитию искусственного интеллекта при Президенте Республики Казахстан.

Особая роль в реализации национальной AI-повестки отводится системе высшего образования и науки. Университеты становятся ключевой площадкой подготовки кадров нового поколения, развития прикладных исследований и формирования национальной AI-экосистемы.

С 2025 года навыки применения искусственного интеллекта внедрены во все образовательные программы организаций высшего образования. AI-дисциплины стали обязательными для всех студентов независимо от специальности. Утвержден межвузовский стандарт применения ИИ, приняты

методические рекомендации по использованию генеративного искусственного интеллекта в образовательном процессе, а также введены требования по обязательному повышению квалификации преподавателей в сфере ИИ.

На сегодняшний день 95 университетов включили AI-дисциплины в учебные планы, 30 вузов реализуют 42 образовательные программы в сфере искусственного интеллекта, а более половины профессорско-преподавательского состава уже прошли базовую AI-подготовку.

AI-Sana как инструмент формирования новых компетенций

Одним из ключевых ответов Казахстана на вызовы AI-экономики стала программа AI-Sana, реализуемая Министерством науки и высшего образования Республики Казахстан по поручению Главы государства.

Основной целью программы является формирование нового поколения специалистов, способных не только использовать технологии искусственного интеллекта, но и создавать собственные прикладные решения для экономики страны. Программа выстроена как комплексная экосистема подготовки AI-кадров и развития студенческого технологического предпринимательства.

На первом этапе программы сотни тысяч студентов прошли массовое обучение базовым AI-компетенциям с использованием международных образовательных платформ. Второй этап программы ориентирован на практическое применение знаний, разработку цифровых продуктов и решение реальных отраслевых задач.

Обучение строится на прикладных кейсах и проектной работе. В программу интегрированы отраслевые кейсы по приоритетным секторам экономики, а участники получают навыки работы с машинным обучением, большими

языковыми моделями, API, инструментами Python и платформами для разработки AI-решений.

Важной особенностью AI-Sana является переход от базовой цифровой грамотности к формированию предпринимательского мышления, ориентированного на создание инновационных продуктов и технологических стартапов.

Подготовка кадров для библиотечного и архивного дела

В условиях цифровизации особую значимость приобретает развитие компетенций в области библиотечного дела, обработки информации и архивного дела.

На сегодняшний день в Казахстане 12 вузов осуществляют подготовку специалистов по данному направлению по 36 образовательным программам. В системе подготовки активно развиваются новые и инновационные программы, включая «Цифровое архивоведение и документоведение», «Библиотечные информационные системы», «Архивоведение и библиотечно-информационная деятельность».

Современные образовательные программы ориентированы на подготовку высококвалифицированных специалистов, владеющих навыками работы со справочно-библиографическими системами, цифровыми инструментами поиска информации, методами обработки данных и автоматизированными системами управления документами.

Таким образом, подготовка кадров в архивной сфере постепенно выходит за рамки традиционного архивоведения и требует освоения новых цифровых компетенций.

Искусственный интеллект в архивном и библиотечном деле

Внедрение технологий искусственного интеллекта в архивное и библиотечное дело становится одним из ключевых направлений цифровой трансформации отрасли. AI-технологии позволяют существенно повысить эффективность поиска, систематизации, анализа и обработки больших массивов данных.

Особую значимость приобретают интеллектуальные системы распознавания документов, семантического поиска, автоматического индексирования, извлечения сущностей и анализа пользовательских запросов. Это позволяет не только ускорить обработку архивных фондов, но и существенно повысить качество работы с информационными ресурсами.

Использование искусственного интеллекта открывает возможности для создания интеллектуальных архивных систем, способных анализировать исторические документы, прогнозировать пользовательские запросы и формировать персонализированные рекомендации.

Одновременно AI способствует развитию knowledge management-подходов, при которых архивы становятся не только хранилищами данных, но и активными центрами интеллектуального управления знаниями.

Практические AI-решения в архивной сфере

В Казахстане уже реализуются прикладные AI-решения в архивной и библиотечной сферах. Одним из примеров является проект Dulaty Science, разработанный Таразским региональным университетом им. М.Х. Дулати. Решение представляет собой интеллектуального помощника для глубокого анализа научных публикаций и архивных данных с интеграцией научных баз данных.

В Казахстанско-Американском свободном университете создан KAFU AI Librarian – библиотечный AI-ассистент, автоматизирующий поиск научной литературы и оформление библиографических ссылок в соответствии с установленными стандартами.

Кокшетауский университет им. Шокана Уалиханова реализовал цифровой репозиторий редких книг с применением технологий искусственного интеллекта. Система позволяет анализировать востребованность фондов и читательский спрос.

Перспективным решением является платформа Jangyr, разработанная Алматинским университетом энергетики и связи им. Г. Даукеева. Платформа предназначена для интеллектуального поиска и интерпретации нормативных документов и регламентов.

Данные кейсы демонстрируют, что университеты Казахстана переходят от подготовки специалистов к созданию собственных прикладных AI-продуктов для реального сектора.

Заключение

Развитие искусственного интеллекта, образования и науки в Казахстане рассматривается как стратегический государственный приоритет. Подготовка специалистов нового поколения требует формирования междисциплинарных компетенций, объединяющих отраслевые знания, цифровые навыки и способность работать с интеллектуальными системами.

Программа AI-Sana выступает важным инструментом формирования таких компетенций, обеспечивая переход от базового освоения технологий к созданию прикладных AI-решений.

Для архивной отрасли это особенно актуально, поскольку цифровая трансформация архивов требует новых профессиональных ролей и компетенций. Искусственный интеллект становится не просто технологическим инструментом, а основой качественно нового подхода к управлению информацией и знаниями, создавая фундамент для устойчивого развития архивной сферы в цифровую эпоху.

Литература

Закон Республики Казахстан от 17 ноября 2025 года № 230-VIII ЗРК «Об искусственном интеллекте». https://adilet.zan.kz/rus/docs/Z2500000230 (дата обращения: 17.05.2026).

Кодекс Республики Казахстан от 9 января 2026 года № 255-VIII ЗРК: Цифровой Кодекс Республики Казахстан. https://adilet.zan.kz/rus/docs/K2600000255 (дата обращения: 17.05.2026).

Об объявлении Года цифровизации и искусственного интеллекта. https://www.akorda.kz/ru/ob-obyavlenii-goda-cifrovizacii-i-iskusstvennogo-intellekta-601222 (дата обращения: 16.05.2026).

THE UK NATIONAL EXPERIENCE IN ARCHIVIST TRAINING

Andrew Payne
Head of Education and Outreach at the National Archives of the United Kingdom
Andrew.Payne@nationalarchives.gov.uk

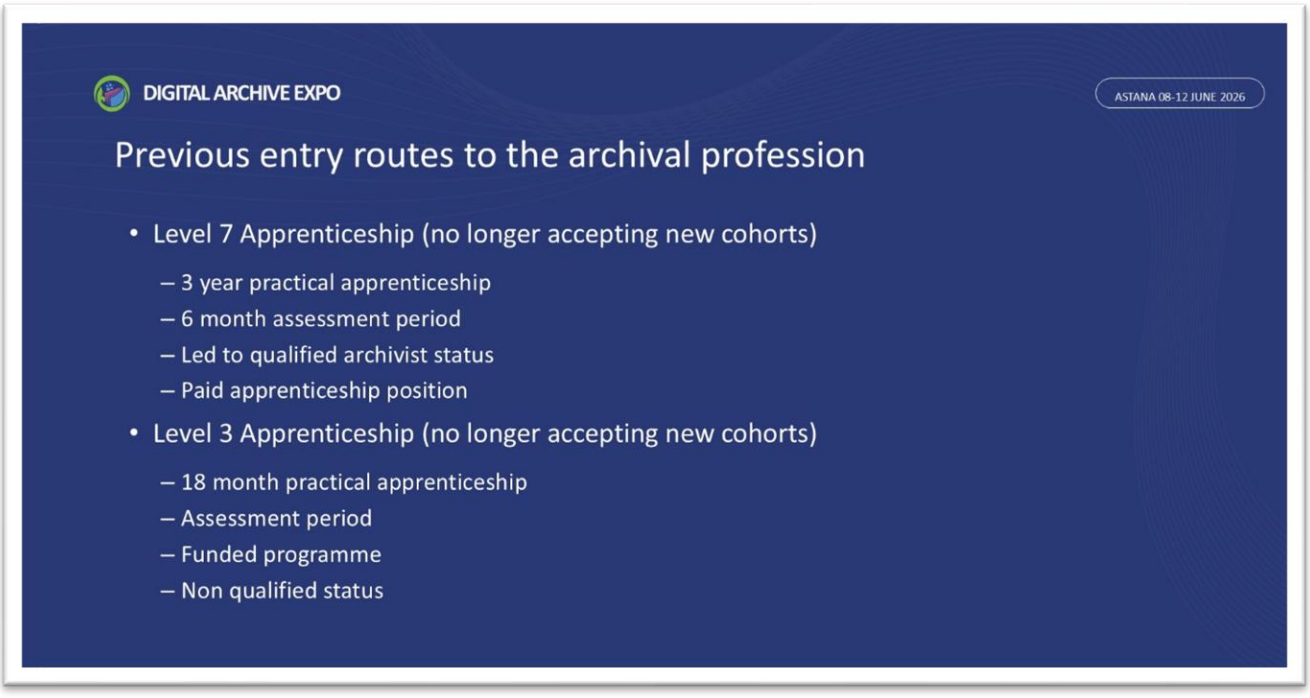
ҰЛЫБРИТАНИЯНЫҢ АРХИВШІЛЕРДІ ДАЯРЛАУДАҒЫ ҰЛТТЫҚ ТӘЖІРИБЕСІ

Эндрю Пейн
Ұлыбритания Ұлттық архивінің білім беру және ағартушылық қызмет бөлімінің басшысы
Andrew.Payne@nationalarchives.gov.uk

НАЦИОНАЛЬНЫЙ ОПЫТ ВЕЛИКОБРИТАНИИ В ПОДГОТОВКЕ АРХИВИСТОВ

Эндрю Пейн
Руководитель отдела образовательной и просветительской деятельности Национального архива Великобритании
Andrew.Payne@nationalarchives.gov.uk

Good afternoon, I am Andrew Payne, Head of Education and Outreach as The National Archives of the UK where I have responsibility for our service for schools, teachers, families, young people and communities. I would like to share some thoughts on how we can attract those who will become the next generation of archivists.

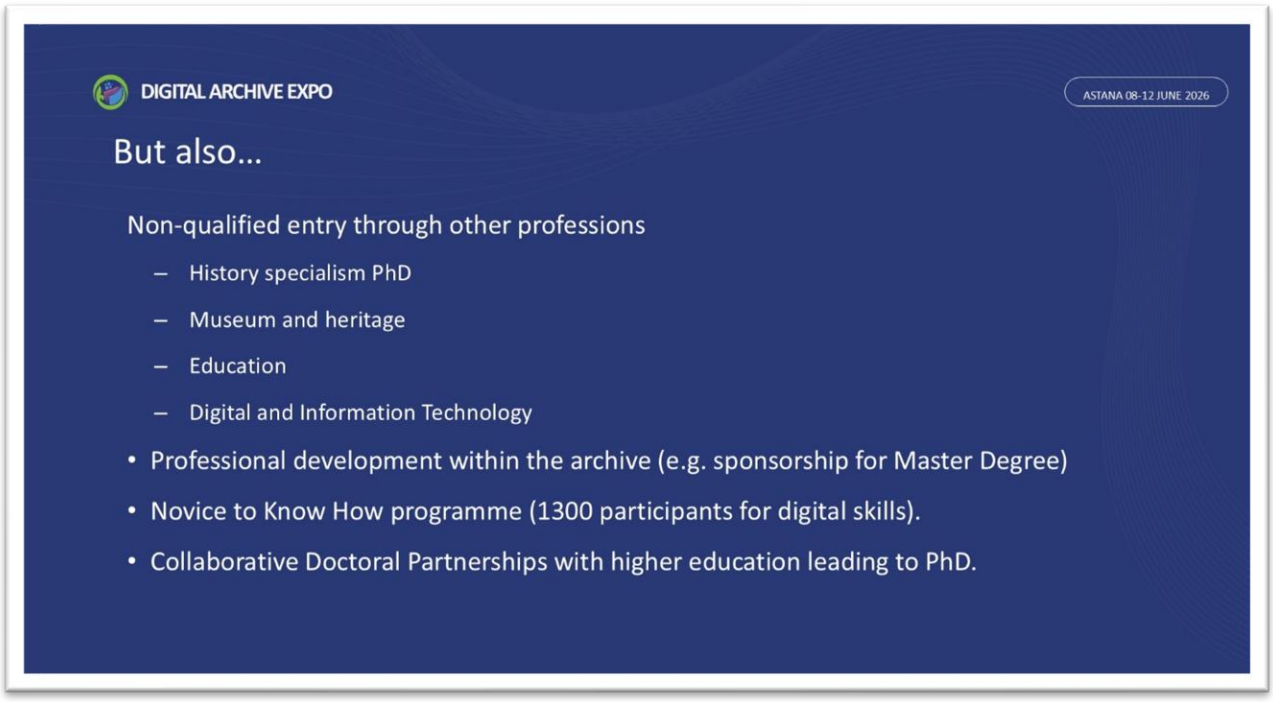


The main entry route to the archival profession in the UK is through a master's degree. There are 6 universities across the UK and Ireland that offer courses. Students can complete these in 1 year as a full-time course or take 2 or 3 years as a part-time student. This qualifies them fully as an archivist but the fees are expensive and often have to be paid up front.

Until recently there were paid apprenticeships available that enabled students to work within an archive while learning their craft. Level 7 apprenticeships lasted 3 years with 6 months assessment at the end and led to fully qualified status. It had the advantage of being a paid programme so provided an accessible route especially for those who needed income as they trained.

There was a lower level of apprenticeship - Level 3 was an 18 month programme which was funded but did not lead to full qualifaction. However, it provided a starting point for students to decide if they would enjoy the world of archives.

For a variety of reasons however, both of these routes have now closed but existing cohorts will complete their programmes.



However, many people enter the archives in a non-qualified route from other professions. I am not a qualified archivist and yet I have just completed 20 years of service at The National Archives. My back ground is in education, I was a high school history teacher. This is an example of how people join the archives from a variety of other professions including History specialists, those from the museum and heritage sector, and digital and information technology backgrounds. This reflects the wide variety of activity within the archives that we now have, beyond tradition archival skills. For some it may then provide an opportunity to qualify as an archivist by being sponsored for a Master's degree. Indeed, the rich diversity of our backgrounds at The National Archives is one of our strengths and what I believe will attract the next generation to work with us.

In addition our Novice to Know How programme has allowed 1300 participants to develop a deeper understanding of digital archiving while our collaborative doctoral partnerships with academic institutions allows candidates to pursue a doctoral degree in a related discipline.

You can find more details about qualifications and professional development, and the salary scales that are recommended in the UK for archivists, at the Archives and Records Administration website at this link.

However, the key issue is not how do people become archivists, it is why would they want to become an archivist. Entry requirements, programmes of training, qualifications and salary are all vital but these are not the things that ultimately attract people to become archivists.

In short, if people don't know about archives they won't think about becoming archivists. Let me illustrate the point with a short story.

In 2008 at the height of the global financial crisis I went with my department to a show of stand up comedy for our Christmas outing. We were sat in the very front row and immediately behind us were a group from the Bank of England. When the compere of the show came out he asked where the group from the Bank of England were and he proceeded to make fun of them. To be honest he ripped into them like a wolf rips into a flock of sheep. I am convinced there was blood on the floor by the time he had finished. And of course everyone laughed because we all have experiences of banks and so we could understand the jokes he made. But then he rather cruelly turned to me and said:

"You think it's funny do you?"

"Yes," I replied.

"Oh so where do you work then?" he asked.

"The National Archives," I said.

And look of horror came over his face.

"Where?"

"The National Archives," I repeated.

"What's that?" he asked looking desperate.

"It's the official archive of the UK Government." I stated.

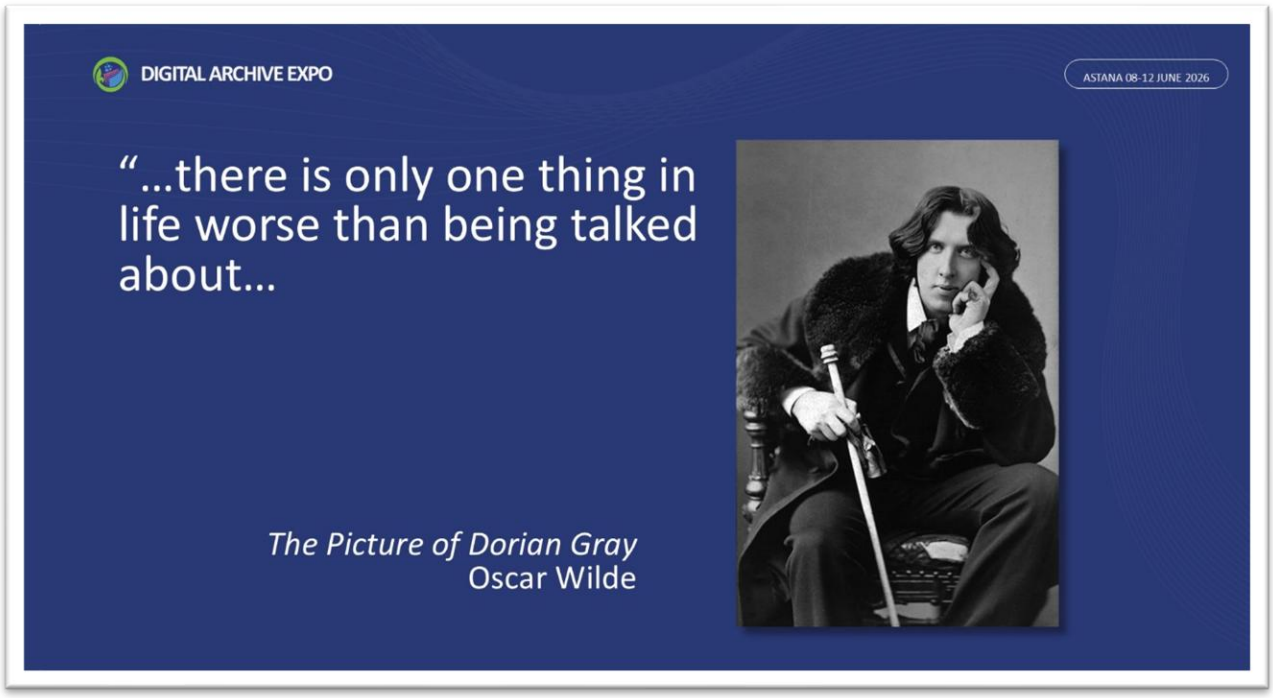
He looked around the room clearly trying to think of some response.

"I know what you're all thinking" he said to the room "let's hear your best archives joke." Then dejectedly he turned away and admitted "No. I have nothing!"
And moved to someone else.

To emphasise this, I found a blog pose from August 2000 asking:

"Why is it that so little archival humour has surfaced? … Why are there no archivist stand-up comics?… Is the lack of archival humour a compliment to the profession or is it a serious case of neglect?"

My point is that humour only works if we have a shared experience and very few people share the experience of archives. If they don't know about us they won't visit us, or make jokes about us and they certainly won't dream of becoming archivists.

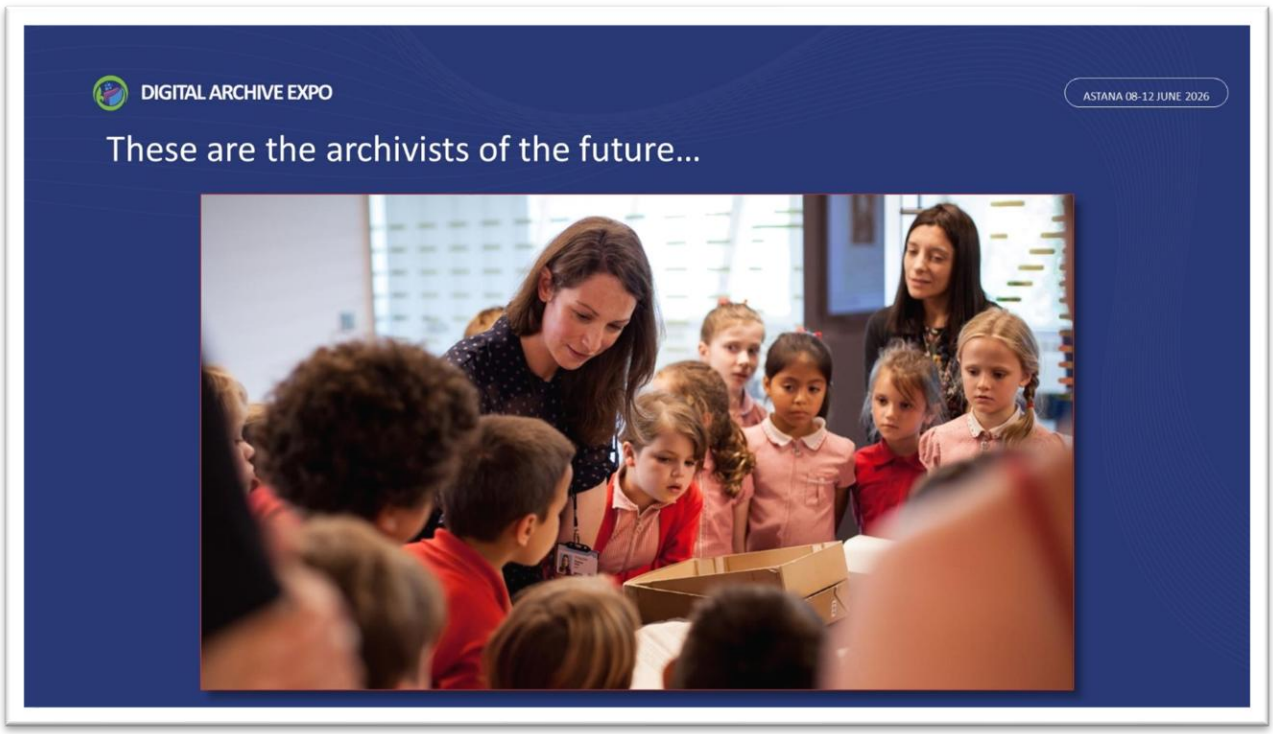


As Oscar Wilde states in *The Picture of Dorian Gray*

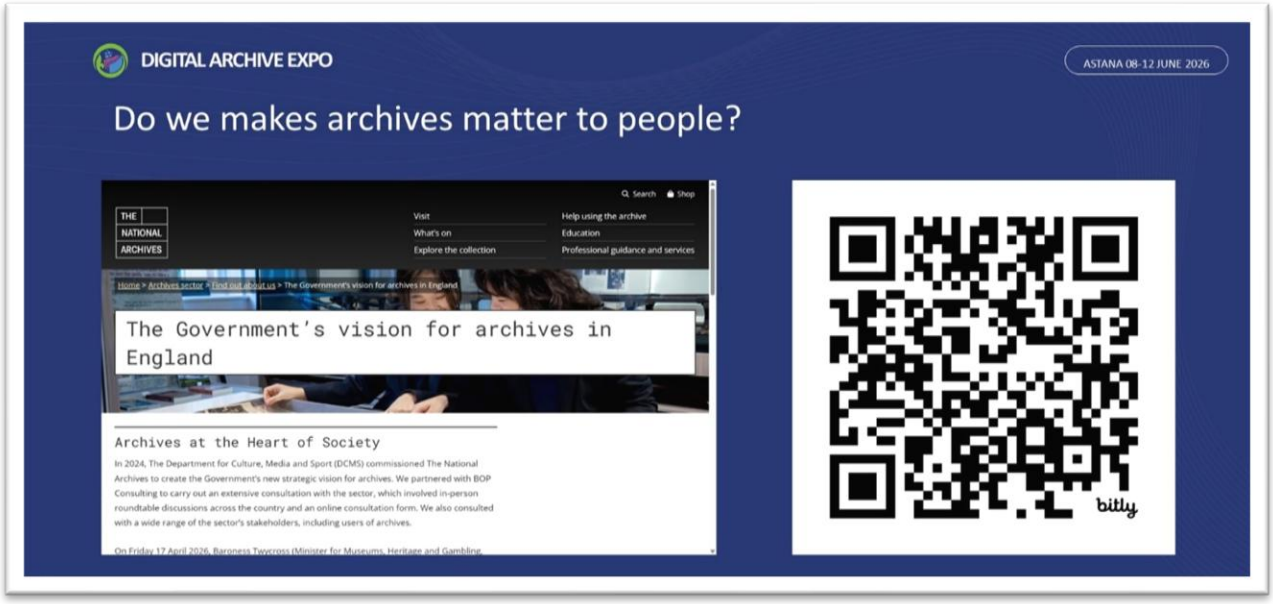
"There is only one thing in life worse than being talked about..." "...and that is *not* being talked about."

Our parent in government is the Department of Culture, Media and Sport and in 2016 it highlighted the challenge for archives in their survey of public participation in culture. In their *Taking Part* report for that year, they found that: “74 out of 100 adults visited a museum, gallery or historic location”.

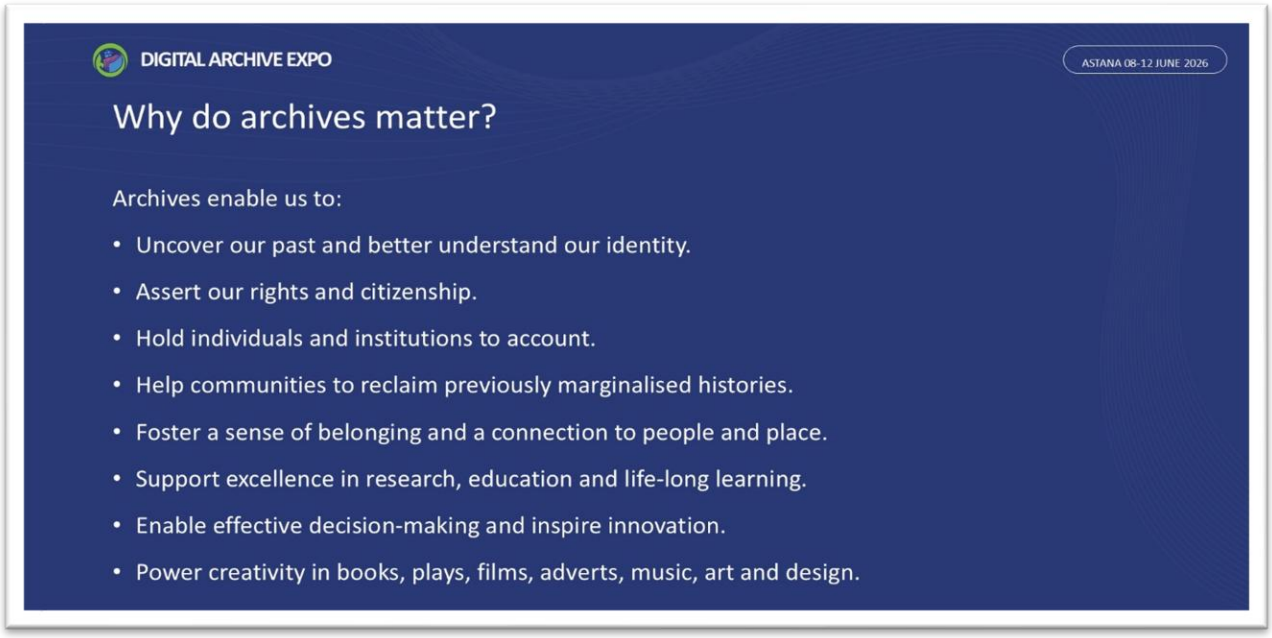
However, only 3 out of 100 adults visited an archive.



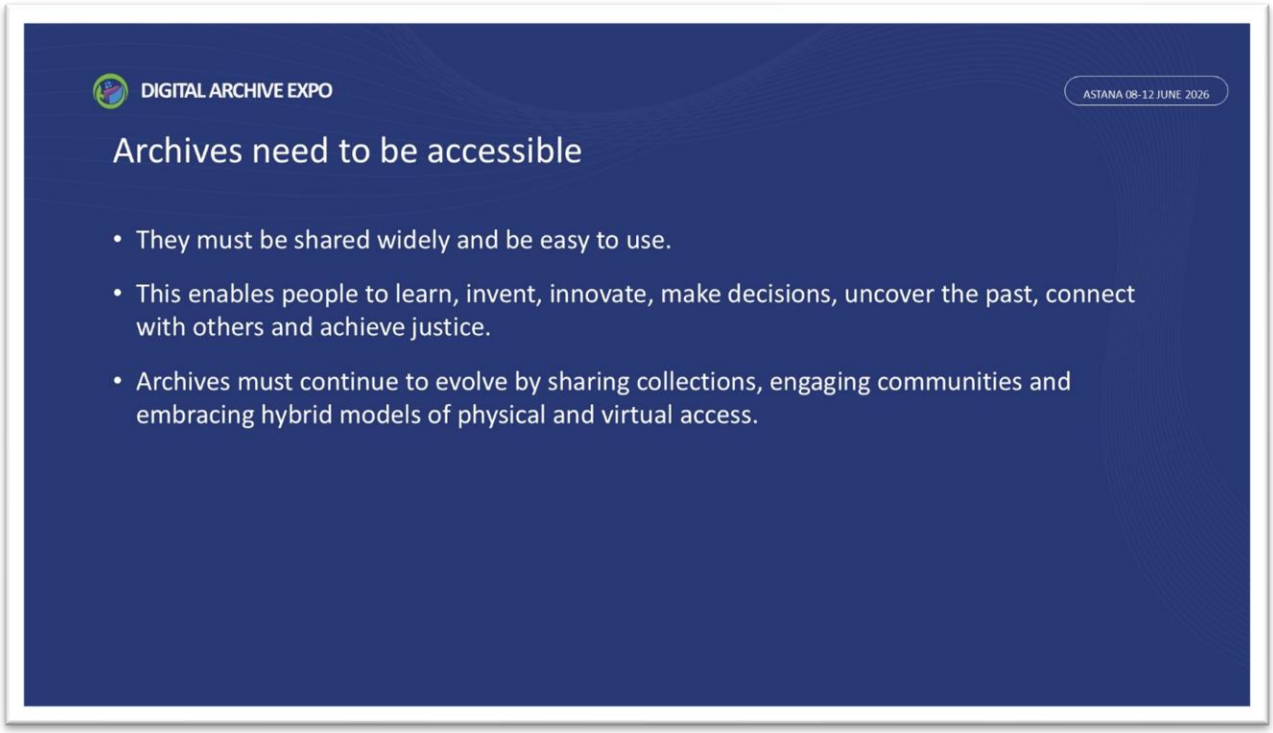
We need to make archives visible so that people know, enjoy and care about them in the same way they know, enjoy and care about museums, galleries and other cultural and heritage institutions. I believe this starts with education and requires a programme of active engagement and outreach. Workshops for school students; family days; exhibitions that actively engage audiences; public events and talks; and community outreach programmes. We have to care about our audiences or they will not care about us.



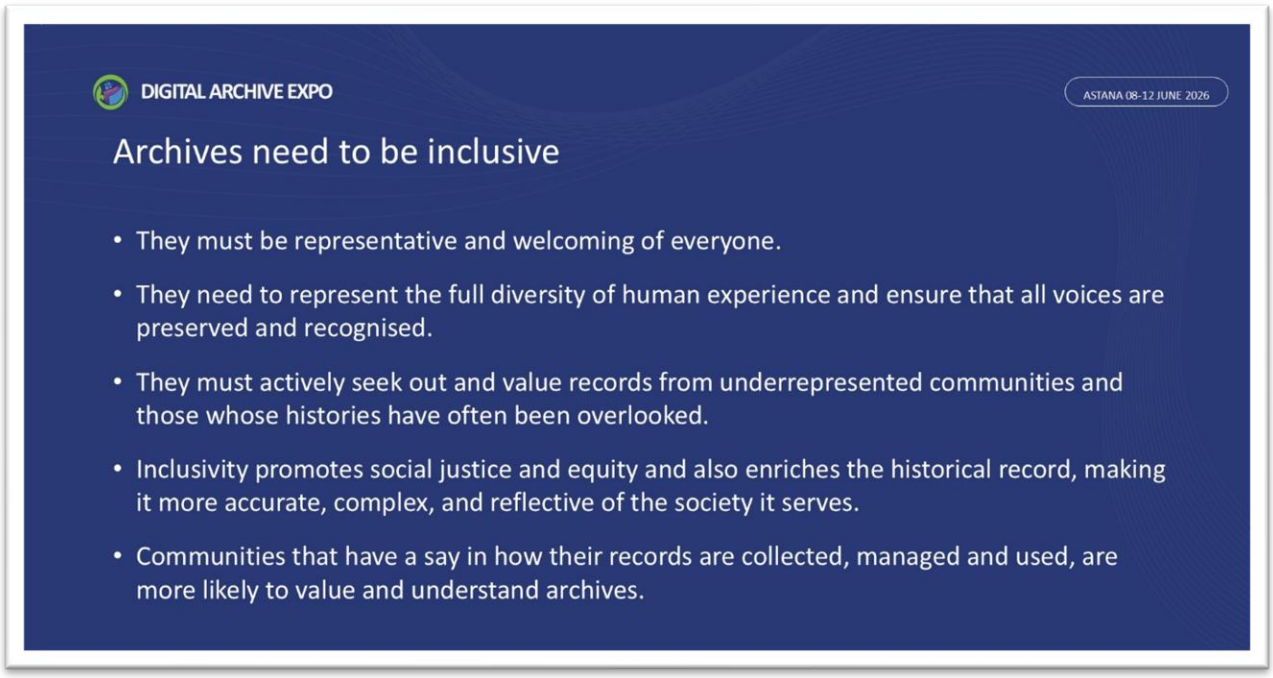
We have to make archives matter to people. This is why our strategic vision for the archives sector is called Archives at the Heart of Society. You can access the full document at the QR code here, but essentially it challenges the sector to make archives matter to people. This means actively working to show people that we have something very important for them.



• Archives matter because they enable us to:
• Uncover our past and better understand our identity.
• Assert our rights and citizenship.
• Hold individuals and institutions to account.
• Help communities to reclaim previously marginalised histories.
• Foster a sense of belonging and a connection to people and place.
• Support excellence in research, education and life-long learning.
• Enable effective decision-making and inspire innovation.
• Power creativity in books, plays, films, adverts, music, art and design.



Archives need to be accessible
• They must be shared widely and be easy to use.
• This enables people to learn, invent, innovate, make decisions, uncover the past, connect with others and achieve justice.
• Archives must continue to evolve by sharing collections, engaging communities and embracing hybrid models of physical and virtual access.



Archives need to be inclusive.

• They must be representative and welcoming of everyone.

• They need to represent the full diversity of human experience and ensure that all voices are preserved and recognised.

• They must actively seek out and value records from underrepresented communities and those whose histories have often been overlooked.

• Inclusivity promotes social justice and equity and also enriches the historical record, making it more accurate, complex, and reflective of the society it serves.

• Communities that have a say in how their records are collected, managed and used, are more likely to value and understand archives.

Archives need to be sustainable.

• They need proper funding and conditions to remain integral in the future.

• They need to promote themselves to their audiences and to their funders

• They must be environmentally sustainable too. They need to consider and reduce their environmental impact.

• Quite simply, no planet means no archives.

The challenge for the future of archives lies in attracting new entrants to the profession, and the solution lies in developing the skills and leadership that recognises that in order to draw people in we need to reach out to them. Online catalogues, digitisation, online presentation and the possibilities that AI provide are all vital and yes, we need to make archives digital...

... but we also need to make them human.

БОЛАШАҚ АРХИВШІЛЕР: БІЛІМ, ҚҰЗЫРЕТТЕР ЖӘНЕ ТЕХНОЛОГИЯЛАР

Қалима Төлеубайқызы Түйенбаева
Педагогика ғылымдарының кандидаты,
әл-Фараби ат. Қазақ ұлттық университеті, Қазақстан
tkt.kalima@gmail.com, https://orcid.org/0000-0001-8230-7740

Аңдатпа. Мақалада цифрлық трансформация жағдайында архивші кәсібінің өзгеруі қарастырылады. Цифрлық құжаттың табиғаты, оның метадеректермен, контекстпен, құқықтық режиммен және ұзақ мерзімді сақтаумен байланысы талданады. Архивтік ақпараттық жүйелер, электрондық архивтер, жасанды интеллект, үлкен деректер, машиналық оқыту, блокчейн, киберқауіпсіздік және цифрлық сақтау технологиялары архив жұмысының мазмұнын кеңейтетіні көрсетіледі. Автор болашақ архившінің кәсіби моделін архивтік дәлелділік, цифрлық басқару, деректер мен алгоритмдерді түсіну, құқықтық-этикалық жауапкершілік, ғылыми маңыздылық, коммуникация және басқару бағыттары арқылы сипаттайды. Сонымен қатар болашақ архившінің кәсіби құзыреттер картасы ұсынылып, білім беру, қайта даярлау, кәсіби сертификаттау және институционалдық әріптестікті дамыту мәселелері негізделеді. Мақалада технология архивтік сараптаманы алмастырмайтыны, керісінше архившінің кәсіби жауапкершілігін күшейтетіні тұжырымдалады.

Түйін сөздер: архивші, цифрлық құжат, метадеректер, архивтік ақпараттық жүйе, жасанды интеллект, цифрлық сақтау, кәсіби құзыреттер, цифрлық трансформация

БУДУЩИЕ АРХИВИСТЫ: ОБРАЗОВАНИЕ, КОМПЕТЕНЦИИ И ТЕХНОЛОГИИ

Калима Толеубаевна Туенбаева
Кандидат педагогических наук,
Казахский национальный университет им. аль-Фараби, Казахстан
tkt.kalima@gmail.com; https://orcid.org/0000-0001-8230-7740

Аннотация. В статье рассматривается трансформация профессии архивиста в условиях цифровой трансформации. Анализируется природа цифрового документа, его связь с метаданными, контекстом, правовым режимом и долгосрочным цифровым хранением. Показано, что архивные информационные системы, электронные архивы, искусственный интеллект, большие данные, машинное обучение, блокчейн, кибербезопасность и технологии цифрового хранения существенно расширяют содержание архивной деятельности. Автор раскрывает профессиональную модель архивиста будущего через такие направления, как архивная доказательность, цифровое управление, понимание данных и алгоритмов, правовая и этическая ответственность, научная значимость, коммуникация и управление. Также предлагается карта профессиональных компетенций будущего архивиста и обосновывается необходимость развития образования, переподготовки, профессиональной сертификации и институционального партнерства. В статье делается вывод о том, что технологии не заменяют архивную экспертизу, а усиливают профессиональную ответственность архивиста.

Ключевые слова: архивист, цифровой документ, метаданные, архивная информационная система, искусственный интеллект, цифровое хранение, профессиональные компетенции, цифровая трансформация

FUTURE ARCHIVISTS: EDUCATION, COMPETENCIES AND TECHNOLOGIES

Kalima T. Tuenbayeva
Candidate of Pedagogical Sciences,
al-Farabi Kazakh National University, Kazakhstan
tkt.kalima@gmail.com; https://orcid.org/0000-0001-8230-7740

Abstract. The article examines the transformation of the archivist's profession in the context of digital transformation. It analyzes the nature of the digital record and its relationship with metadata, context, legal regimes, and long-term digital preservation. The article shows that archival information systems, electronic archives, artificial intelligence, big data, machine learning, blockchain, cybersecurity, and digital preservation technologies expand the scope and complexity of archival work. The author conceptualizes the professional model of the future archivist through archival evidentiality, digital management, understanding of data and algorithms, legal and ethical responsibility, scholarly value, communication, and management. The article also presents a map of professional competencies for future archivists and substantiates the need to develop education, retraining, professional certification, and institutional partnership. It concludes that technology does not replace archival expertise; rather, it strengthens the professional responsibility of the archivist.

Keywords: archivist, digital record, metadata, archival information system, artificial intelligence, digital preservation, professional competencies, digital transformation

Кіріспе

Цифрлық трансформация архив саласының кәсіби логикасын, архившінің жауапкершілік аймағын және болашақ мамандарды даярлау жүйесін қайта қарауды талап етеді. Бұл мәселені тек архивтерді цифрландыру, ақпараттық жүйелерді енгізу немесе жасанды интеллектті қолдану шеңберінде қарастыру жеткіліксіз. Негізгі мәселе — цифрлық ортада құжаттың табиғаты, дәлелдік

құндылығы, құқықтық режимі және қоғамдық сенімді қамтамасыз ету тәсілдері өзгеріп отырғанында.

Қазіргі жағдайда құжаттар цифрлық ортада жасалады, ақпараттық жүйелерде өңделеді, деректермен және метадеректермен бірге сақталады, құқықтық режимдерге, технологиялық платформаларға және ақпараттық қауіпсіздік талаптарына тәуелді болады. Сондықтан архив жұмысы тек сақтау мен сипаттаудан ғана тұрмайды; ол контексті, дәлелділікті, қолжетімділікті, қауіпсіздікті және қоғамдық сенімді кәсіби басқаруға айналып келеді.

Мақаланың мақсаты – цифрлық трансформация жағдайында архившінің кәсіби рөлі мен технологиялық құзыреттілігін қайта қарау, болашақ архившінің кәсіби моделін және негізгі құзыреттерін айқындау, сондай-ақ білім беру, қайта даярлау, кәсіби сертификаттау және институционалдық әріптестікті жаңарту қажеттілігін негіздеу.

Мақаланың негізгі тезисі мынаған негізделеді: болашақ архивші цифрлық архивтік ортаны сенімді, басқарылатын, құқықтық тұрғыдан реттелген, ғылыми маңызы бар және қоғамдық жауапкершілігі жоғары ортаға айналдыратын кәсіби маман.

Цифрлық құжат және технологиялық орта

Архивші кәсібінің өзгеруі ең алдымен құжат табиғатының өзгеруімен байланысты. Дәстүрлі архивтік ортада құжат қағаз тасымалдағышпен, іс жүргізу тәртібімен, қорлық құрылыммен және физикалық сақтаумен ұштасса, цифрлық ортада ол деректерден, метадеректерден, техникалық сипаттамалардан, құқықтық режимдерден және өзге цифрлық объектілермен байланыстардан тұратын күрделі ақпараттық объектіге айналады. Сондықтан цифрлық құжатты тек файл ретінде қабылдау оның архивтік мәнін толық ашпайды.

Халықаралық тәжірибеде born-digital records ұғымы қолданылады. Бұл – қағаз түпнұсқасы болмай, бастапқыдан цифрлық ортада жасалған құжаттар. Мұндай құжаттарға электрондық басқарушылық құжаттар, деректер базалары, электрондық пошта, веб-ресурстар, цифрлық фотосуреттер, аудио және бейнематериалдар жатады. Олардың архивтік құндылығы тек мазмұнымен емес, шығу тегімен, жасалу жағдайымен, өзгерістер тарихымен, сақталған жүйесімен және құқықтық режимімен айқындалады.

Цифрлық құжаттың дәлелдік құндылығы мазмұн, метадеректер, контекст, шығу тегі, өзгерістер тарихы, құқықтық мәртебе, техникалық сипаттамалар, түпнұсқалықты растау және сақтау шарттары арқылы қамтамасыз етіледі. Құжаттарды басқару, яғни records management, құжаттың жасалуынан бастап сақталуына, пайдаланылуына және қайта пайдаланылуына дейінгі бүкіл өмірлік циклін ескеруді талап етеді (ISO 2016). Бұл тұрғыдан құжатты үздіксіз басқару моделі де маңызды, өйткені ол құжатты архивке тапсырылғаннан кейін ғана емес, оның пайда болу, қолданылу, сақталу және қайта пайдаланылу кезеңдерінің тұтастығында қарастырады (Upward 1996).

Ұзақ мерзімді цифрлық сақтау мәселесі Open Archival Information System — ашық архивтік ақпараттық жүйе моделі ретінде белгілі OAIS тұжырымдамасында жүйелі түрде қарастырылады (ISO 2025). Бұл модель цифрлық объектінің болашақта оқылатын, түсінікті және дәлелді күйде сақталуын қамтамасыз етуге бағытталған. Ұзақ мерзімді сақтау тек файлды сақтау емес; ол форматтардың

ескіруі, деректердің жоғалуы, жүйелердің жаңартылуы және қолжетімділіктің үздіксіздігі сияқты тәуекелдерді басқарумен байланысты (Rosenthal et al. 2005).

Қағаз ортада құжаттың контексті қор, тізімдеме, іс, тіркеу деректері және іс жүргізу байланыстары арқылы сақталды. Ал цифрлық ортада мұндай байланыстар өздігінен сақталмайды. Сондықтан олар метадеректер, идентификаторлар, әрекеттер журналдары, қолжетімділік құқықтары және сақтау регламенттері арқылы бекітілуі тиіс. Архивтік сипаттаудың халықаралық стандарттары мен контекстік модельдері құжатты жеке объект ретінде емес, оның шығу тегімен, қызметтік ортасымен және басқа құжаттармен байланысы арқылы қарастырудың маңызын көрсетеді (ICA 2000; ICA 2023). Цифрлық сақтауға арналған метадеректер, яғни preservation metadata, цифрлық объектінің сақталу тарихын, техникалық сипаттамаларын, құқықтық шектеулерін және өзгерістерін тіркеуге мүмкіндік береді (PREMIS Editorial Committee 2015).

Цифрлық ортада түпнұсқалық мәселесі де күрделене түседі. Түпнұсқалық тек файлдың өзгермегенін дәлелдеумен шектелмейді; ол құжаттың шығу тегі, жасалу жағдайы, контексті және дәлелдік күшін сақтаумен байланысты (Duranti et al. 2022). Сондықтан электрондық қолтаңба, файл тұтастығын тексеру, әрекеттер журналдары немесе блокчейн сияқты құралдар сенімді техникалық жағынан қолдағанымен, архивтік сенімнің негізін контекст, метадерек және кәсіби бағалау құрайды.

Цифрлық құжаттың табиғаты өзгерген сайын архив жұмысының технологиялық ортасы да кеңейеді. Қазіргі архивші архивтік ақпараттық жүйелермен, электрондық архивтермен, электрондық құжат айналымы жүйелерімен, цифрлық репозиторийлермен, деректер базаларымен, бағдарламалық қамтамасыз етумен және лицензиялық платформалармен жұмыс істейді. Бұл жүйелер құжатты тіркеуге, сипаттауға, іздеуге, сақтауға және қолжетімділікті басқаруға мүмкіндік береді. Алайда олардың архивтік логикаға сәйкес жұмыс істеуі өздігінен қамтамасыз етілмейді. Қандай метадеректер қажет екенін, қор мен іс құрылымы қалай сақталатынын, цифрлық объект құжатпен қалай байланысатынын және қолжетімділік қалай реттелетінін кәсіби тұрғыдан архивші анықтайды.

Сканерлеу, цифрлау және мәтінді тану технологиялары қағаз, қолжазба, аудио және бейнематериалдарды цифрлық ортаға енгізуге мүмкіндік береді. Бірақ цифрлық көшірме жасау құжаттың архивтік құндылығын автоматты түрде айқындамайды. Құжаттың тарихи контексті, дереккөздік ерекшелігі, дәлелдік маңызы және түпнұсқамен байланысы кәсіби сараптама арқылы анықталады.

Жасанды интеллект архив жұмысындағы маңызды технологиялардың біріне айналып отыр. Ол мәтінді тануға, бастапқы сипаттама жасауға, түйін сөздер ұсынуға, құжаттарды топтастыруға, ұқсас объектілерді табуға және семантикалық іздеуді дамытуға көмектеседі. Алайда жасанды интеллект нәтижесі дайын архивтік шешім емес, кәсіби тексеруді қажет ететін бастапқы нәтиже ғана. Жасанды интеллект жүйелерін жауапты қолдану халықаралық тәжірибеде тәуекелдерді басқару, сенімділік, түсіндірілетіндік және адами бақылау қағидаларымен байланыстырылып қарастырылады (NIST 2023).

Үлкен деректер мен машиналық оқыту электрондық хат алмасу, веб-архивтер, деректер базалары, аудиовизуалды коллекциялар, корпоративтік

жүйелер және ғылыми деректер сияқты ірі әрі әртекті массивтермен жұмыс істеуге мүмкіндік береді. Машиналық оқыту құжаттарды сұрыптауға, жіктеуге және байланыстарды анықтауға көмектескенімен, алгоритм нәтижесі ықтималды сипатта болады. Сондықтан автоматтандырылған жіктеу архивтік сараптаманы алмастырмайды; ол контекст, деректер сапасы, құқықтық режим және дәлелдік құндылық тұрғысынан кәсіби бағалауды қажет етеді.

Цифрлық архивтер үшін киберқауіпсіздік пен криптографиялық қорғау да маңызды. Қолжетімділікті басқару, пайдаланушыны аутентификациялау, деректерді шифрлау, резервтік көшіру, әрекеттерді журналдау және ақпараттық қауіпсіздік саясаты цифрлық құжаттардың қорғалуын қамтамасыз етеді. Ақпараттық қауіпсіздік талаптары цифрлық ортада деректердің құпиялылығы, тұтастығы және қолжетімділігі қағидаларымен тығыз байланысты (ISO/IEC 2022).

Блокчейн және таратылған реестр технологиялары цифрлық құжаттың шығу тегі мен өзгерістер тарихын техникалық тұрғыдан бекітуге мүмкіндік береді. Мысалы, ARCHANGEL жобасында цифрлық қоғамдық құжаттардың шығу тегі, өзгермейтіндігі және тұтастығын криптографиялық түрде қамтамасыз ету мүмкіндігі қарастырылған (Collomosse et al. 2018). Алайда мұндай технологиялар архивтік сипаттау, құндылық сараптамасы және құқықтық бағалау функцияларын алмастырмайды.

Цифрлық технологиялар архив жұмысына жаңа мүмкіндіктер береді, бірақ архившінің кәсіби рөлін толық алмастырмайды. Технология құжатты жылдам өңдеуі, сақтауы, іздеуі немесе талдауы мүмкін. Алайда құжаттың архивтік мәнін, қорлық контекстін, құқықтық режимін, тарихи маңызын және дәлелдік құндылығын архивші анықтайды. Сондықтан цифрлық ортада архившінің жауапкершілігі әлсіремейді; керісінше, қате сипаттау, контекст жоғалту, деректер сапасының төмендеуі, құқықтық режимнің бұзылуы және құжаттың дәлелдік күшінің әлсіреуі сияқты тәуекелдерді кәсіби басқару арқылы күшейе түседі.

Болашақ архившінің кәсіби моделі және құзыреттері

Цифрлық құжат пен технологиялық ортаның күрделенуі архившінің кәсіби бейнесін қайта қарауды талап етеді. Цифрлық технологиялар архив жұмысын жеделдетіп, үлкен көлемдегі деректерді өңдеуге, құжаттарды іздеуге, сипаттауға, сақтауға және қорғауға мүмкіндік береді. Алайда құжаттың архивтік мағынасы, контексті, дәлелдік күші және қоғамдық сенімі технологияның өзінен емес, кәсіби бағалаудан қалыптасады. Цифрлық құжаттың түпнұсқалығы мен сенімділігі оның шығу тегі, жасалу жағдайы, контексті және дәлелдік күшінің сақталуымен тығыз байланысты (Duranti et al. 2022).

Осы тұрғыдан болашақ архившіні техникалық оператор немесе дайын жүйенің пассивті пайдаланушысы ретінде қарастыру жеткіліксіз. Ол цифрлық ортада құжаттың шығу тегін, контекстін, құқықтық режимін, ғылыми маңызын және дәлелдік құндылығын сақтайтын кәсіби сарапшы әрі басқарушы тұлға ретінде қалыптасуы тиіс.

Болашақ архившінің кәсіби моделі – цифрлық ортада архившінің негізгі кәсіби рөлін, жауапкершілік аймақтарын және кәсіби әрекет бағыттарын көрсететін тұжырымдамалық құрылым. Бұл модельдің мәні архившіні технологияны жай қолданушы ретінде емес, цифрлық құжаттың дәлелділігін,

контекстін, құқықтық режимін, ғылыми маңызын және қоғамдық сенімін қамтамасыз ететін кәсіби маман ретінде сипаттауда.

Осы модельдің мазмұны келесі негізгі бағыттар арқылы ашылады. Олар 1-кестеде берілді.

1-кесте. *Болашақ архившінің кәсіби моделі*

Модель бағыты	Мазмұны
Архивтік дәлелділік	құжаттың түпнұсқалығы, шығу тегі, контексті, тұтастығы және дәлелдік күші
Цифрлық басқару	метадеректер, репозиторийлер, архивтік ақпараттық жүйелер, форматтар, цифрлық сақтау
Деректер мен алгоритмдерді түсіну	үлкен деректер, машиналық оқыту, автоматтандырылған сұрыптау, жіктеу және семантикалық іздеу нәтижелерін кәсіби бағалау
Құқықтық және этикалық жауапкершілік	дербес деректер, авторлық құқық, қолжетімділік шектеулері, құпиялық, ақпараттық қауіпсіздік және ЖИ нәтижелеріне жауапкершілік
Ғылыми маңыздылық	зерттеу деректері, ғылыми цифрлық жобалар, цифрлық гуманитаристика, дереккөздік сын және деректерді қайта ғылыми пайдалану
Коммуникация және басқару	IT-мамандармен, заңгерлермен, зерттеушілермен, мемлекеттік органдармен, бизнеспен және пайдаланушылармен кәсіби өзара әрекет

1-кестеде берілген модель болашақ архившінің кәсіби қызметі бір ғана технологиялық дағдымен шектелмейтінін көрсетеді. Архивші цифрлық ортада құжаттың дәлелдік күшін, метадеректер мен контекстің сақталуын, ақпараттық жүйелердің архивтік талаптарға сәйкестігін, құқықтық және этикалық шектеулерді, ғылыми қолдану мүмкіндігін және кәсіби коммуникацияны біртұтас жауапкершілік кеңістігі ретінде басқаруы тиіс.

Архивтік дәлелділік бұл модельдің негізгі бағыты болып қала береді. Цифрлық файлдың сақталуы оның толыққанды архивтік құжат ретінде танылуына жеткіліксіз. Құжаттың түпнұсқалығы, шығу тегі, тұтастығы, қорлық контексті және дәлелдік күші сақталуы қажет. Егер цифрлық объектінің қандай жүйеде жасалғаны, кім өзгерткені, қандай құжаттармен байланысы бар екені немесе қандай құқықтық режимде пайдаланылатыны белгісіз болса, оның архивтік дәлелдік құндылығы әлсірейді. Архивтік сипаттаудың қазіргі контекстік тәсілдері де құжатты жеке объект ретінде емес, оның шығу тегі, қызметтік ортасы және өзге құжаттармен байланысы арқылы қарастырудың маңызын көрсетеді (ICA 2023).

Цифрлық басқару архившінің кәсіби кеңістігін кеңейтеді. Болашақ архивші метадеректермен, репозиторийлермен, архивтік ақпараттық жүйелермен, форматтармен, қолжетімділік режимдерімен және ұзақ мерзімді цифрлық сақтаумен жұмыс істей алуы тиіс. Бұл жерде негізгі мәселе жүйені жай пайдалану

емес, оның архивтік логикаға сай жұмыс істеуін қамтамасыз ету болып табылады. Құжатты үздіксіз басқару моделі цифрлық құжаттың өмірлік циклін тұтас қарастыру қажеттігін көрсетеді (Upward 1996).

Деректер мен алгоритмдерді түсіну де болашақ архившінің кәсіби моделінің маңызды бағытына жатады. Үлкен деректер, машиналық оқыту, автоматтандырылған сұрыптау, жіктеу және семантикалық іздеу архив жұмысына жаңа мүмкіндіктер береді. Бірақ алгоритм нәтижесі дайын кәсіби шешім ретінде қабылданбауы тиіс. Мұндай нәтижелер деректер сапасы, контекст, құқықтық режим, алгоритмдік шектеулер және дәлелдік құндылық тұрғысынан кәсіби бағалауды қажет етеді. Жасанды интеллект жүйелерін қолдану тәуекелді басқару, сенімділік, түсіндірілетіндік және адами бақылау қағидаларымен байланыстырылуы тиіс (NIST 2023).

Құқықтық және этикалық жауапкершілік цифрлық архив жағдайында ерекше маңызға ие. Дербес деректер, авторлық құқық, қолжетімділік шектеулері, құпиялық, ақпараттық қауіпсіздік және жасанды интеллект нәтижелерін пайдалану архившінің кәсіби шешіміне тікелей әсер етеді. Ақпараттық қауіпсіздік талаптары цифрлық ортада деректердің құпиялылығы, тұтастығы және қолжетімділігі қағидаларымен байланысты (ISO/IEC 2022).

Ғылыми маңыздылық болашақ архивші моделінің тағы бір бағыты ретінде көрінеді. Архивтік деректер цифрлық гуманитаристикада, тарихи талдауда, желілік талдауда, тақырыптық модельдеуде және пәнаралық зерттеулерде кеңінен қолданыла бастады. Мұндай жағдайда архивші дереккөзді сақтаушы ғана емес, ғылыми деректердің сапасына, шығу тегіне, сипатталуына және қайта пайдаланылуына жауап беретін маман ретінде қарастырылады.

Коммуникация және басқару бағыты архившінің кәсіби рөлін пәнаралық ортада кеңейтеді. Болашақ архивші IT-мамандармен, заңгерлермен, зерттеушілермен, басқарушылармен, қаржыгерлермен, бизнес өкілдерімен және пайдаланушылармен жұмыс істейді. Оның міндеті — архивтік логиканы басқа кәсіби тілдерге түсіндіру: метадеректің маңызын, файлды контекстен бөлуге болмайтынын, жасанды интеллект нәтижесінің тексерілуі қажет екенін және цифрлық сақтау ұзақ мерзімді ресурс пен басқаруды талап ететінін негіздеу.

Осы модель нақты кәсіби құзыреттер арқылы іске асады. Сондықтан болашақ архившінің құзыреттерін жекелеген техникалық дағдылар ретінде емес, цифрлық архивтік ортада кәсіби шешім қабылдауға мүмкіндік беретін әрекеттер жүйесі ретінде қарастыру қажет. Бұл байланыс 2-кестеде көрсетілді.

***2-кесте**. Болашақ архившінің кәсіби құзыреттер картасы*

Құзырет аймағы	Болашақ архивші нені істей алуы тиіс?
Архивтік негіз және дәлелділік	құжаттың құндылығын, шығу тегін, контекстін, қорлық байланысын және дәлелдік күшін бағалау
Цифрлық құжат және метадеректер	құжаттың контекстін, сипатталуын, сақталуын және қолжетімділігін басқару

Ақпараттық жүйелер және бағдарламалық қамтамасыз ету	архивтік ақпараттық жүйе, репозиторий, электрондық архив немесе лицензиялық платформаның архивтік талаптарға сәйкестігін анықтау
Цифрлау және сапаны бақылау	сканерлеу, мәтінді тану, формат, файл сапасы және цифрлық көшірменің түпнұсқамен байланысын айқындау
ЖИ, үлкен деректер және машиналық оқыту	автоматты сипаттау, сұрыптау, жіктеу және іздеу нәтижелерін кәсіби тұрғыдан тексеру
Сенім, қауіпсіздік және блокчейн	файл тұтастығы, электрондық қолтаңба, әрекеттер журналы, криптография және өзгерістер тарихын түсіну
Құқық және этика	дербес деректер, авторлық құқық, қолжетімділік, құпиялық және ЖИ қолдану салдарын бағалау
Стандарттар және нормативтік құжаттар	метадеректер, цифрлық сақтау, электрондық құжат, ЖИ, қолжетімділік және ақпараттық қауіпсіздік бойынша регламенттерді түсіну және әзірлеуге қатысу
Ғылыми-зерттеу және академиялық коммуникация	архивтік деректерді ғылыми жобаларда қолдану, дереккөздік сын жүргізу, мақала, баяндама және ғылыми жоба түрінде нәтижені ұсыну
Жобалық және қаржылық ойлау	цифрлық жобаның ресурсын, лицензиясын, сақтау құнын, сүйемелдеу шығынын және тұрақтылығын бағалау
Коммуникация, оқыту және фасилитация	IT-мамандармен, заңгерлермен, зерттеушілермен, басқарушылармен, пайдаланушылармен жұмыс істеу, әріптестерді оқыту және пәнаралық келісімді ұйымдастыру

2-кесте кәсіби модельдің нақты әрекет деңгейінде қалай көрінетінін көрсетеді. Мұнда құзырет технологияны қолдану қабілетімен ғана шектелмейді, ол технология нәтижесін архивтік логика, құқықтық режим, деректер сапасы, ғылыми маңыз және қоғамдық сенім тұрғысынан бағалай алу қабілетін білдіреді.

Болашақ архившінің құзыреттері жекелеген техникалық дағдылардың қарапайым жиынтығы емес. Бұл – архивтанулық негізді, цифрлық технологияларды түсінуді, деректермен жұмыс істеуді, құқықтық және этикалық жауапкершілікті, ғылыми ойлауды, басқарушылық қабілетті, қаржылық түсінікті және кәсіби коммуникацияны біріктіретін тұтас дайындық. Осыдан келесі мәселе туындайды: мұндай құзыреттер өздігінен қалыптаспайды, олар білім беру, қайта даярлау, кәсіби сертификаттау және институционалдық әріптестік арқылы жүйелі түрде дамытылуы тиіс.

Білім беру, қайта даярлау және институционалдық әріптестік

Болашақ архившінің кәсіби құзыреттерін қалыптастыру білім беру мен қайта даярлау жүйесін кәсіби мазмұнды жаңартуды талап етеді. Алдыңғы бөлімде көрсетілген модель мен құзыреттер картасы архившінің цифрлық ортадағы жауапкершілік аймақтарын айқындайды, ендігі мәселе — осы құзыреттердің

қандай білім беру тетіктері, кәсіби тәжірибе және институционалдық байланыстар арқылы орнығатынын анықтау.

Архившіні даярлау оқу бағдарламасына жекелеген цифрлық пәндерді қосудан ғана тұрмауы тиіс. Цифрлық архивтік ортада маман құжаттың архивтік табиғатын, метадеректердің маңызын, ақпараттық жүйелердің шектеулерін, жасанды интеллект нәтижелерінің ықтималды сипатын, құқықтық режимдерді және ұзақ мерзімді сақтау тәуекелдерін кәсіби тұрғыдан түсінуі қажет. Сондықтан білім беру мазмұны классикалық архивтанулық негізді сақтай отырып, цифрлық ортада кәсіби шешім қабылдау қабілетін қалыптастыруға бағытталуы керек. Архивтік білім беру туралы зерттеулер де цифрлық дәуірде теориялық біліммен қатар нақты кәсіби дағдылар мен зерттеу қабілеттерін дамыту қажеттігін көрсетеді (Eastwood 2006).

Оқу процесі нақты архивтік кейстерге, практикалық тапсырмаларға, пәнаралық жобаларға және кәсіби шешім қабылдау жағдайларына сүйенгенде ғана нәтижелі болады. Цифрлық құжатпен жұмыс істеу тек техникалық әрекеттер жиынтығы емес; ол құжаттың шығу тегін, контекстін, құқықтық мәртебесін, метадеректерін және болашақта қайта пайдаланылу мүмкіндігін бірге бағалауды талап етеді. Осы тұрғыдан цифрлық архивші мен цифрлық кураторды даярлау мәселесі халықаралық тәжірибеде архив, кітапхана, музей және деректер инфрақұрылымын байланыстыратын пәнаралық дайындық ретінде қарастырылады (Yakel et al. 2011; Madrid 2013).

Қызметтегі архившілерді қайта даярлау ерекше мәнге ие. Цифрлық трансформация тек болашақ мамандарға ғана емес, қазіргі архив саласында жұмыс істеп жүрген мамандарға да қатысты. Алайда қайта даярлау нақты бір бағдарламаның интерфейсін үйретумен шектелмеуі тиіс. Ол архивтік ойлауды жаңартуға, цифрлық құжатпен жұмыс істеудің кәсіби логикасын түсінуге және технологиялық шешімдердің архивтік салдарын бағалауға бағытталуы қажет. Сондықтан мұндай бағдарламалар профильдік сипатта болып, архивтану, құжаттану, цифрлық сақтау, метадеректер, құқықтық режимдер және архивтік ақпараттық жүйелер бойынша ғылыми-білім беру базасы бар құрылымдар арқылы жүзеге асуы тиіс.

Модульдік және онлайн-оқыту, практикалық мектептер, цифрлық архивтану зертханалары, архивтердегі тағылымдамалар және жобалық семинарлар кәсіби дамудың икемді тетіктері ретінде қарастырылуы мүмкін. Мұндай форматтар технология тез өзгеретін жағдайда мамандардың үздіксіз кәсіби дамуына мүмкіндік береді. Цифрлық кураторлыққа арналған оқу шеңберлері кәсіби даярлауда техникалық дағдыны, басқарушылық қабілетті, құқықтық-этикалық жауапкершілікті және коммуникацияны өзара байланысты қарастыру қажеттігін көрсетеді (Molloy et al. 2014).

Кәсіби сертификаттау осы даярлау жүйесінің маңызды бөлігі болуы тиіс. Семинарға қатысу маманның нақты кәсіби құзыретін әрдайым дәлелдей бермейді. Сертификаттау архившінің цифрлық архивтік ортада жұмыс істеуге қабілетті екенін көрсетуі қажет: ол цифрлық құжатпен, ақпараттық жүйелермен, құқықтық режимдермен, деректермен, технологиялық нәтижелермен және кәсіби жауапкершілікпен жұмыс істей алатынын растауы тиіс. Сондықтан

сертификаттау оқуға қатысуды емес, нақты кәсіби құзыретті бағалау тетігі ретінде қарастырылуы керек.

Мұндай даярлау жүйесі жеке оқу бағдарламаларының жиынтығы ретінде ғана емес, институционалдық әріптестікке негізделген кәсіби орта ретінде қалыптасуы тиіс. Университеттер білім беру, әдіснама және ғылыми негізді қамтамасыз етеді; архив мекемелері нақты қорлармен, электрондық құжаттармен, цифрлық коллекциялармен, пайдаланушы сұраныстарымен және архивтік ақпараттық жүйелермен жұмыс істеу тәжірибесін береді. Мемлекеттік органдар нормативтік база, стратегия және қаржыландыру мәселелерін айқындайды, ал кәсіби қауымдастықтар стандарттар, сараптама және сертификаттау өлшемдерін қалыптастыруға қатысады.

Зерттеу орталықтары мен ғылыми жобалар архив саласын жаңартудың тәжірибелік алаңы бола алады. Олар жаңа әдістемелерді сынақтан өткізуге, цифрлық құралдардың тиімділігін бағалауға, архивтік деректерді ғылыми айналымға енгізуге және дәлелге негізделген шешімдер қалыптастыруға мүмкіндік береді. Деректер архивтері мен цифрлық инфрақұрылымдар туралы зерттеулер де архивтердің тек сақтау орны емес, білімді өндіру, тарату және қайта пайдалану процесін қолдайтын институционалдық инфрақұрылым екенін көрсетеді (Borgman et al. 2019).

IT-сектор мен бизнес те бұл жүйеден тыс қала алмайды, себебі олар архивтік ақпараттық жүйелерді, электрондық архивтерді, лицензиялық платформаларды, жасанды интеллект құралдарын, киберқауіпсіздік шешімдерін, бұлттық сервистерді және деректерді өңдеу технологияларын ұсынады. Дегенмен технологиялық шешімді жеткізуші кәсіби талапты алмастырмауы керек. Қандай жүйе қажет, қандай метадерек сақталуы тиіс, қолжетімділік қалай реттеледі, деректер қалай қорғалады және жасанды интеллект нәтижесі қалай тексеріледі деген сұрақтарға кәсіби жауапты архив саласы беруі тиіс.

Болашақ архившіні даярлау бір реттік білім алу процесі емес, үздіксіз кәсіби даму жүйесі ретінде қарастырылуы қажет. Бұл жүйе білім беру, қайта даярлау, практикалық дайындық, кәсіби сертификаттау, ғылыми зерттеу, нормативтік реттеу, технологиялық әріптестік және қаржылық жоспарлау арасындағы тұрақты байланысқа негізделеді. Мұндай байланыс қалыптаспайынша, технологияларды енгізу архившілердің кәсіби даярлығынан, нормативтік реттеуден және архив жүйесінің институционалдық дайындық деңгейінен озып кету қаупі сақталады.

Болашақ архившінің кәсіби құзыреттері жеке пәндер немесе қысқа курстар арқылы емес, архивтік білім, практика, ғылым, технология және кәсіби қауымдастық өзара байланысқан ортада қалыптасады. Бұл тұжырым архив саласын цифрлық трансформация жағдайында дамыту бағыттарын айқындауға негіз болады.

Архив саласын цифрлық трансформация жағдайында дамыту бағыттары

Архив саласын цифрлық трансформация жағдайында дамыту жекелеген технологияларды енгізумен немесе қысқа мерзімді оқыту шараларымен шектелмейді. Бұл үдеріс архивтік құжаттың дәлелдік құндылығын, цифрлық ақпараттың сенімділігін және кәсіби жауапкершілікті қамтамасыз ететін жүйелі орта қалыптастырумен байланысты.

Архив саласының цифрлық дамуы нақты жағдайды ғылыми-әдістемелік тұрғыдан зерттеуге сүйенуі тиіс. Архив мекемелерінің кадрлық әлеуеті, қолданыстағы ақпараттық жүйелері, цифрлық құжаттармен жұмыс тәжірибесі, нормативтік базасы, техникалық инфрақұрылымы және қаржылық мүмкіндіктері анықталмайынша, цифрлық трансформацияның тұрақты әрі нәтижелі үлгісін қалыптастыру қиын. Мұндай зерттеу цифрлық өзгерістерді кездейсоқ жобалар жиынтығы ретінде емес, архив саласының институционалдық дамуын қамтамасыз ететін басқарушылық процесс ретінде қарастыруға мүмкіндік береді.

Технологиялық шешімдерді архив саласының кәсіби талаптарымен сәйкестендіру де негізгі бағыттардың бірі болып табылады. Ақпараттық жүйе, электрондық архив, репозиторий, жасанды интеллект құралы немесе деректерді сақтау платформасы архивтік логиканы өздігінен қалыптастырмайды. Сондықтан технологияны таңдау мен енгізу үдерісінде метадеректердің толықтығы, контекстің сақталуы, қолжетімділік режимі, ақпараттық қауіпсіздік, ұзақ мерзімді сақтау және кәсіби бақылау талаптары алдын ала айқындалуы қажет.

Архив саласының цифрлық дамуы кадрлық даярлау мен кәсіби құзыретті растау жүйесімен тікелей байланысты. Болашақ архившіні даярлау және қызметтегі мамандарды қайта даярлау цифрлық ортада кәсіби шешім қабылдай алатын, технологиялық нәтижені архивтік логика тұрғысынан бағалайтын және құқықтық-этикалық жауапкершілікті ескеретін маман қалыптастыруға бағытталуы тиіс. Бұл тұрғыдан кәсіби сертификаттау оқуға қатысу фактісін емес, нақты құзыреттілікті растау құралы ретінде қарастырылуы қажет.

Нормативтік және институционалдық негізді жаңарту цифрлық трансформацияның орнықтылығын қамтамасыз етеді. Электрондық құжат, метадеректер, цифрлық сақтау, жасанды интеллектті қолдану, қолжетімділік, ақпараттық қауіпсіздік, салалық біліктілік шеңбері және кәсіптік стандарттар өзара байланысқан жүйе ретінде қарастырылғанда ғана кәсіби талаптар мен технологиялық шешімдер арасында үйлесім қалыптасады. Бұл бағыт архив саласындағы білім беру, қайта даярлау, сертификаттау және кәсіби бағалау өлшемдерін бір жүйеге келтіруге мүмкіндік береді.

Ғылыми және пилоттық жобалар цифрлық өзгерістерді кәсіби практикаға енгізудің тәжірибелік негізін қалыптастырады. Жаңа әдістер мен цифрлық құралдар архивтік деректер негізінде сыналып, олардың тиімділігі, шектеулері, құқықтық және этикалық салдары бағаланғаннан кейін ғана кең кәсіби қолданысқа енгізілуі тиіс. Мұндай тәсіл архив саласының дәлелге негізделген, ғылыми тұрғыдан негізделген және институционалдық жағынан тұрақты дамуына ықпал етеді.

Цифрлық жүйелерді ұзақ мерзімді сүйемелдеу архив саласын дамытудың дербес бағыты ретінде қарастырылуы қажет. Цифрлық платформа іске қосылғаннан кейін жұмыс аяқталмайды; керісінше, деректер сапасын бақылау, форматтарды жаңарту, қауіпсіздік талаптарын сақтау, техникалық қолдау, пайдаланушыларды оқыту және регламенттерді қайта қарау тұрақты кәсіби процесс ретінде жалғасады. Сондықтан цифрлық трансформация енгізу кезеңімен ғана емес, оның ұзақ мерзімді басқарылуымен бағалануы тиіс.

Архив саласының цифрлық дамуы тұрақты институционалдық әріптестік жағдайында ғана нәтижелі болады. Университеттер, архив мекемелері,

мемлекеттік органдар, кәсіби қауымдастықтар, зерттеу орталықтары, жұмыс берушілер және IT-сектор ортақ кәсіби кеңістікте әрекет еткен жағдайда технология архивтік логикаға қызмет ететін құралға айналады. Демек, архив саласын цифрлық трансформация жағдайында дамытудың басты бағыты — технологиялық шешімдерді архивтік дәлелділікке, кәсіби жауапкершілікке және қоғамдық сенімге бағындыратын жүйе қалыптастыру.

Қорытынды

Жүргізілген талдау цифрлық трансформация архив саласындағы өзгерістерді тек технологиялық жаңару деңгейінде қарастыруға болмайтынын көрсетеді. Цифрлық ортада құжаттың табиғаты, оның дәлелдік құндылығы, сақталу шарттары, қолжетімділік режимі және қоғамдық сеніммен байланысы күрделене түседі. Осыған байланысты архившінің кәсіби рөлі де кеңейеді: ол құжатты сақтаушы ғана емес, цифрлық архивтік ортаның сенімділігін, құқықтық реттелуін, ғылыми қолданылуын және ұзақ мерзімді тұрақтылығын қамтамасыз ететін кәсіби субъект ретінде қалыптасады.

Мақалада ұсынылған болашақ архившінің кәсіби моделі архивтік дәлелділік, цифрлық басқару, деректер мен алгоритмдерді түсіну, құқықтық-этикалық жауапкершілік, ғылыми маңыздылық, коммуникация және басқару бағыттарының өзара байланысына негізделеді. Бұл модель архившінің қызметін жеке техникалық дағдылармен шектемей, цифрлық құжаттың өмірлік циклін, метадеректерге, контекстке, ақпараттық жүйелерге, технологиялық нәтижелерге және қоғамдық сенімге жауапты кәсіби жүйе ретінде қарастыруға мүмкіндік береді.

Құзыреттер картасы болашақ архившінің нақты кәсіби әрекеттерін айқындайды. Мұнда технологияны меңгеру жеткіліксіз, архивші технологиялық нәтижені архивтік логика, құқықтық режим, деректер сапасы, дәлелдік құндылық және ғылыми маңыз тұрғысынан бағалай алуы тиіс. Жасанды интеллект, үлкен деректер, машиналық оқыту, блокчейн, киберқауіпсіздік және цифрлық сақтау құралдары архив жұмысына жаңа мүмкіндік бергенімен, кәсіби сараптама мен жауапкершілікті алмастырмайды.

Осыған байланысты болашақ архившіні даярлау жүйесі де кәсіби мазмұнда жаңартылуы қажет. Білім беру, қайта даярлау, кәсіби сертификаттау, ғылыми-практикалық жобалар, нормативтік реттеу және институционалдық әріптестік өзара байланысқан орта ретінде дамығанда ғана архив саласының цифрлық трансформациясы тұрақты әрі нәтижелі сипат алады. Цифрлық жүйелерді енгізу архивтік логикадан, кәсіби құзыреттен және нормативтік дайындықтан озып кетпеуі тиіс.

Болашақ архивші технологиямен бәсекелесетін маман емес, технологияны архивтік мақсатқа бағындыратын кәсіби тұлға ретінде қалыптасуы тиіс. Оның негізгі міндеті цифрлық құжаттар мен архивтік ақпараттың дәлелділігін, контекстін, құқықтық негізін, ғылыми құндылығын және қоғам алдындағы сенімділігін қамтамасыз ету. Осы тұрғыдан болашақ архивші өткеннің құжаттық ізін ғана емес, келер ұрпақ үшін тарихи жадының дәлелдік негізін сақтайтын маман ретінде айқындалады.

Әдебиеттер

Borgman et al. 2019 – Borgman C.L., Scharnhorst A., Golshan M.S. Digital Data Archives as Knowledge Infrastructures: Mediating Data Sharing and Reuse // *Journal of the Association for Information Science and Technology*. 2019. Vol. 70, No. 8. P. 888–904. DOI: 10.1002/asi.24172. URL: https://doi.org/10.1002/asi.24172 (accessed 15.05.2026).

Collomosse et al. 2018 – Collomosse J., Bui T., Brown A., Sheridan J., Green A., Bell M., Fawcett J., Higgins J., Thereaux O. ARCHANGEL: Trusted Archives of Digital Public Documents // *Proceedings of the ACM Symposium on Document Engineering*. 2018. URL: https://doi.org/10.1145/3209280.3229120 (accessed 15.05.2026).

Duranti et al. 2022 – Duranti L., Rogers C., Thibodeau K. Authenticity // *Archives and Records*. 2022. Vol. 43, No. 2. P. 188–203. https://doi.org/10.1080/23257962.2022.2054406 (accessed 15.05.2026).

Eastwood 2006 – Eastwood T. Building archival knowledge and skills in the digital age // *Archival Science*. 2006. Vol. 6. P. 163–170. https://doi.org/10.1007/s10502-006-9026-2 (accessed 15.05.2026).

ICA 2000 – International Council on Archives. ISAD(G): General International Standard Archival Description. 2nd ed. Ottawa: ICA, 2000. URL: https://www.ica.org/resource/isadg-general-international-standard-archival-description-second-edition/ (accessed 15.05.2026).

ICA 2023 – International Council on Archives. Records in Contexts – Conceptual Model. Paris: ICA Expert Group on Archival Description, 2023. URL: https://www.ica.org/ica-network/expert-groups/egad/records-in-contexts-conceptual-model/ (accessed 15.05.2026).

ISO 2016 – International Organization for Standardization. ISO 15489-1:2016. *Information and documentation – Records management – Part 1: Concepts and principles*. Geneva: ISO, 2016. URL: https://www.iso.org/standard/62542.html (accessed 15.05.2026).

ISO 2025 – International Organization for Standardization. ISO 14721:2025. *Space Data System Practices – Reference model for an open archival information system (OAIS)*. Geneva: ISO, 2025. URL: https://www.iso.org/standard/87471.html (accessed 15.05.2026).

ISO/IEC 2022 – International Organization for Standardization. ISO/IEC 27001:2022. *Information security, cybersecurity and privacy protection – Information security management systems — Requirements*. Geneva: ISO, 2022. https://www.iso.org/standard/27001 (accessed 15.05.2026).

Madrid 2013 – Madrid M.M. A study of digital curator competences: A survey of experts // *The International Information & Library Review*. 2013. Vol. 45, No. 3–4. P. 149–156. https://doi.org/10.1080/10572317.2013.10766382 (accessed 15.05.2026).

Molloy et al. 2014 – Molloy L., Gow A., Konstantelos L. The DigCurV Curriculum Framework for Digital Curation in the Cultural Heritage Sector // *International Journal of Digital Curation*. 2014. Vol. 9, No. 1. P. 231–241. https://doi.org/10.2218/ijdc.v9i1.314 (accessed 15.05.2026).

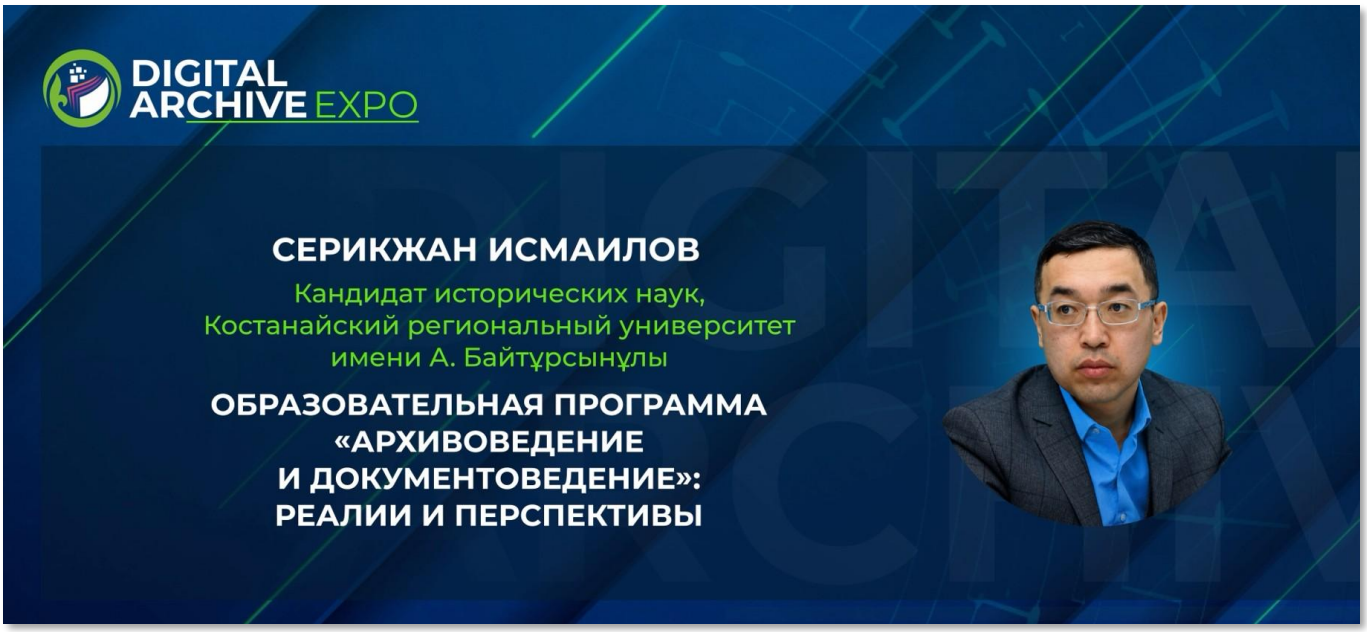
NIST 2023 – National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. Gaithersburg: NIST, 2023. https://doi.org/10.6028/NIST.AI.100-1 (accessed 15.05.2026).

PREMIS Editorial Committee 2015 – PREMIS Editorial Committee. PREMIS Data Dictionary for Preservation Metadata. Version 3.0. Washington, DC: Library of Congress, 2015. URL: https://www.loc.gov/standards/premis/v3/

Rosenthal et al. 2005 – Rosenthal D.S.H., Robertson T.S., Lipkis T., Reich V., Morabito S. Requirements for Digital Preservation Systems: A Bottom-Up Approach // *D-Lib Magazine*. 2005. Vol. 11, No. 11. http://www.dlib.org/dlib/november05/rosenthal/11rosenthal.html (accessed 15.05.2026).

Upward 1996 – Upward F. Structuring the Records Continuum — Part One: Postcustodial Principles and Properties // *Archives & Manuscripts*. 1996. Vol. 24, No. 2. P. 268–285. https://publications.archivists.org.au/index.php/asa/article/view/8583 (accessed 15.05.2026).

Yakel et al. 2011 – Yakel E., Conway P., Hedstrom M., Wallace D. Digital Curation for Digital Natives // *Journal of Education for Library and Information Science*. 2011. Vol. 52, No. 1. P. 23–31. https://www.jstor.org/stable/25764651 (accessed 15.05.2026).



ОБРАЗОВАТЕЛЬНАЯ ПРОГРАММА «АРХИВОВЕДЕНИЕ, ДОКУМЕНТОВЕДЕНИЕ И МУЗЕЙНОЕ ДЕЛО»: РЕАЛИИ И ПЕРСПЕКТИВЫ

Серикжан Сагындыкович Исмаилов
Кандидат исторических наук, ассоциированный профессор,
заведующий кафедрой социально-гуманитарных дисциплин,
Костанайский региональный университет им. А. Байтурсынулы, Казахстан
ismailov_7777@mail.ru

Аннотация. В публикации рассматриваются ключевые аспекты, современные реалии и перспективы подготовки специалистов по образовательной программе «Архивоведение, документоведение и музейное дело» в Костанайском региональном университете им. А. Байтурсынова. Особое внимание уделено междисциплинарному характеру обучения, сочетающему теоретическую базу с развитием цифровых компетенций в условиях автоматизации, внедрения электронных архивов и виртуальных экспозиций. Освещается проблема дефицита кадров, обладающих одновременно гуманитарными знаниями и ИТ-навыками. Обосновывается значимость тесной интеграции с профильными базами практик (государственными архивами, музеями и библиотеками) как фундаментального фактора адаптации выпускников к требованиям современного рынка труда, сохранения историко-культурного наследия и формирования цифровой инфраструктуры общества.

Ключевые слова: Костанайский региональный университет им. Ахмет Байтұрсынұлы; архивоведение; документоведение; музейное дело; цифровизация; электронные архивы; базы практик; высшее образование; культурное наследие

«АРХИВТАНУ, ҚҰЖАТТАНУ ЖӘНЕ МУЗЕЙ ІСІ» БІЛІМ БЕРУ БАҒДАРЛАМАСЫ: БҮГІНГІ ЖАЙ-КҮЙІ МЕН БОЛАШАҒЫ

Серікжан Сағындықұлы Исмаилов
Тарих ғылымдарының кандидаты, қауымдастырылған профессор,
әлеуметтік-гуманитарлық пәндер кафедрасының меңгерушісі,
А. Байтұрсынұлы ат. Қостанай өңірлік университеті, Қазақстан
ismailov_7777@mail.ru

Аңдатпа. Бұл жарияланымда Ахмет Байтұрсынұлы атындағы Қостанай өңірлік университетіндегі «Архивтану, құжаттану және музей ісі» білім беру бағдарламасы бойынша мамандар даярлаудың негізгі аспектілері, қазіргі жай-күйі мен даму перспективалары қарастырылады. Оқыту үдерісінің пәнаралық сипатына ерекше назар аударылып, онда теориялық білім беру цифрлық құзыреттерді дамытумен ұштастырылатыны көрсетіледі. Бұл бағыт қазіргі таңда автоматтандырудың кеңінен енгізілуі, электрондық архивтердің дамуы және виртуалды экспозициялардың қалыптасуы жағдайында ерекше өзектілікке ие. Сонымен қатар, гуманитарлық біліммен қатар ақпараттық технологиялар саласындағы дағдыларды қатар меңгерген мамандар тапшылығы мәселесі талданады. Бітірушілердің қазіргі еңбек нарығының талаптарына бейімделуін қамтамасыз етуде, тарихи-мәдени мұраны сақтауда және қоғамның цифрлық инфрақұрылымын дамытуда бейіндік практика базаларымен (мемлекеттік архивтермен, музейлермен және кітапханалармен) тығыз ықпалдастықтың маңызы негізделеді.

Түйін сөздер: Ахмет Байтұрсынұлы атындағы Қостанай өңірлік университеті; архивтану; құжаттану; музей ісі; цифрландыру; электрондық архивтер; практика базалары; жоғары білім; мәдени мұра

«EDUCATIONAL PROGRAMME «ARCHIVAL STUDIES AND RECORDS MANAGEMENT»: REALITIES AND PROSPECTS»

Serikzhan S. Ismailov
Candidate of Historical Sciences, Associate Professor,
Head of the Department of Social Sciences and Humanities,
A. Baitursynuly Kostanay Regional University, Қазақстан
ismailov_7777@mail.ru

Abstract. The publication examines the key aspects, current realities, and future development prospects of specialist training within the "Archival Studies, Records Management, and Museology" educational program at Akhmet Baitursynuly Kostanay Regional University. Special attention is paid to the interdisciplinary nature of education, which combines a theoretical foundation with the development of digital competencies amidst automation, the deployment of electronic archives, and the creation of virtual exhibitions. The study highlights the issue of a shortage of personnel possessing both humanities knowledge and IT skills. It substantiates the importance of close integration with specialized internship placements (state archives, museums, and libraries) as a fundamental factor for graduates' adaptation to the demands of the modern labor market, the preservation of historical and cultural heritage, and the formation of society's digital infrastructure.

Keywords: Kostanay Regional University after A. Baitursynov; archival studies; records management; museology; digitalization; electronic archives; internship placements; higher education; cultural heritage

Образовательная программа «Архивоведение, документоведение и музейное дело» направлена на подготовку специалистов в области управления документами, сохранения историко-культурного наследия и организации музейной деятельности. Программа охватывает теоретические и практические аспекты архивного дела, документоведения, источниковедения, а также музееведения, включая учет, хранение, реставрацию и экспонирование фондов. Особое внимание уделяется современным информационным технологиям, цифровизации архивов и созданию электронных баз данных. Выпускники программы обладают компетенциями в сфере организации документооборота, экспертизы ценности документов, научного описания архивных и музейных коллекций, а также популяризации культурного наследия.

Образовательная программа «Архивоведение, документоведение и музейное дело» ориентирована на подготовку специалистов в сфере сохранения и управления документальным и культурным наследием.

Программа сочетает междисциплинарный подход, объединяя архивоведение, документоведение и музееведение.

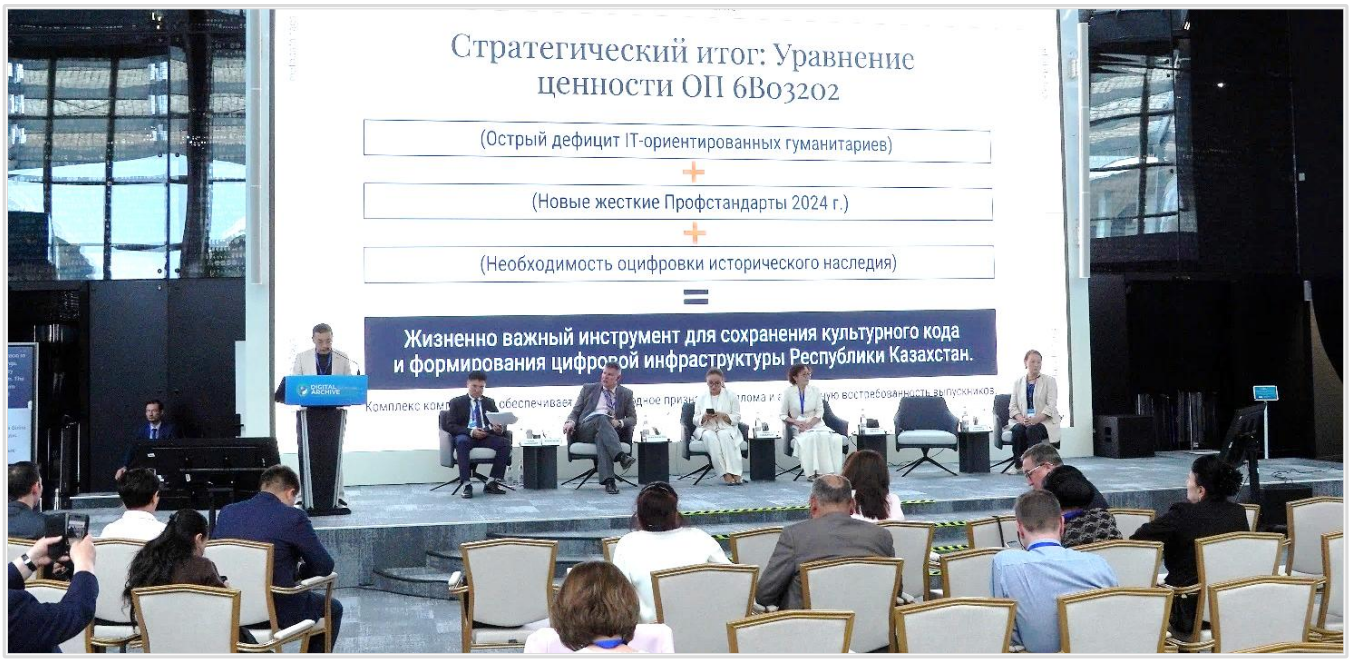
В учебный процесс включены как теоретические основы (источниковедение, история архивного дела), так и практические навыки (учет, хранение и описание фондов). Учитывая современные тенденции по внедрению цифровых технологий и искусственного интеллекта, особое внимание уделяется цифровизации архивов и внедрению современных информационных технологий.

Основные разделы программы включают сдедующие аспекты: изучение истоков и развития архивного и музейного дела, ключевых понятий и теорий; основы архивного управления, включая организацию архивов, методы хранения и обработки документов, а также правила доступа и использования архивных материалов; изучение типов документов, их классификация, создание и использование деловых документов, а также методы работы с ними; основы работы музеев, включая экспозиционное дело, организацию выставок, сохранение музееведческих коллекций и работа с посетителями; понимание значимости культурного наследия, его охраны и сохранения, а также работа с культурными объектами; использование современных технологий в архивном деле и музейном управлении, включая цифровизацию документов и управление электронными архивами; основы правовых норм и этических принципов в области архивного дела и музейного управления; овладение навыками проведения научных исследований, написания курсовых и дипломных работ, участие в проектах по сохранению документов и музейных коллекций; прохождение практики в архивах, музеях или других организациях для получения практического опыта работы в данной сфере.

Для снижения избыточной теоретической нагрузки в образовательную программу введены дисциплины практического цикла: Методика историко-краеведческой работы; Архивоведение и электронный документооборот;

Музеография; Практикум по информационно-аналитической работе и документоведение; Вспомогательные исторические дисциплины; Информационные технологии в деятельности архивиста и музееведа; Научно-фондовая работа в архиве и музее; Информационные технологии в документационном обеспечении управления; Комплектование, учет и хранение музейных фондов; Электронные архивы: организация и управление; Историческая информатика.

В процессе разработки образовательной программы были задействованы руководители и работники баз и филиалов кафедры: КГУ «Костанайский областной историко- краеведческий музей» Управления культуры Акимата Костанайской области, КГУ «Государственный архив Костанайской области» Управления информатизации, оказания государственных услуг и архивов Акимата Костанайской области, Костанайский областной мемориальный музей Ибрая Алтынсарина, КГУ «Қоғамдық келісім» Управления общественного развития Акимата Костанайской области, КГУ «Центр исследования, реставрации и охраны историко-культурного наследия» Управления культуры Акимата Костанайской области, КГКП «Рудненский социально-гуманитарный колледж им. И.Алтынсарина».



Реалии и перспективы

Современное состояние подготовки специалистов характеризуется возрастанием роли цифровых технологий в архивном, документном и музейном деле. Наблюдается переход от традиционных форм хранения к электронным архивам и цифровым музейным коллекциям. Сохраняется проблема недостаточного финансирования учреждений культуры и архивов, что влияет на материально-техническую базу. Отмечается дефицит специалистов, обладающих одновременно гуманитарными знаниями и цифровыми компетенциями. Это связано также с низкой оплатой труда. Мы провели устный

опрос среди работников, связанных с архивным обязанностями. Например, зарплата специалиста в областном архиве составляет примерно 200-203 тыс. тенге, архивиста в бюджетных учреждениях – 150-130 тыс. тенге, архивируса – 120 тыс. тенге. Стоит острый вопрос по подготовке специалистов данного направления в регионах. Многие работники архивных учреждений не имеют специального образования.

Усиливается потребность в стандартизации процессов управления документами и цифрового хранения данных. Перспективы развития связаны с внедрением технологий цифровизации, автоматизации и использования искусственного интеллекта в обработке и анализе документов. Актуальным направлением становится развитие электронных музеев и виртуальных выставок.

Ожидается расширение профессиональных компетенций выпускников за счет междисциплинарной подготовки и международного сотрудничества. В долгосрочной перспективе программа будет играть важную роль в сохранении культурного наследия и формировании цифровой инфраструктуры общества.

Но внедрение цифровых методов обработки архивного, музейного и других материаловтребует у квалифицированного работника особых знаний компьютерной графики. Поэтому преподавателями были изучены интересы и запросы обучающихся для профессиональной сертификации, разрабатываются новые дисциплины с учетом новых условий работы и предложений работодателей. Для усиления подготовки были добавлены следующие дисциплины: Аудиовизуальные архивы; Архивы и музеи мира и Казахстана; Историческая информатика; Музейная антропология; Электронные архивы: организация и управление; Информационные технологии в деятельности архивиста и музееведа; Информационные технологии в документационном обеспечении управления; Управление информацией и документацией электронного правительства; Практикум по работе с архивными и музейными документами; Научно-инновационные технологии и технические средства в музейной работе; Архивы и электронный документооборот; Практикум по информационно-аналитической работе и документоведение.

Следует отметить, что с внедрением искусственного интеллекта преподавание вышеназванных дисциплин претерпело большие изменения. Основное преимущество использования искусственного интеллекта в соответствующей области заключается в его возможностях для обработки большого объема информации и передачи знаний. Компьютеры могут обучаться на большом количестве данных и хранить информацию без ограничений памяти. Это позволяет им оперативно анализировать информацию и предоставлять точные ответы на вопросы студентов. Кроме того, искусственный интеллект способен предоставить персонализированное образование, учитывая уровень подготовки и особенности каждого студента.

Связь с базами практик

Образовательная программа тесно интегрирована с профильными базами практик, включая государственные архивы, музеи, библиотеки и организации с развитым документооборотом. Практическая подготовка является ключевым элементом формирования профессиональных компетенций студентов.

Заключены договоры о сотрудничестве с учреждениями культуры и управления, обеспечивающими доступ к реальным фондам и документам. В процессе практики студенты осваивают навыки учета, описания, систематизации и хранения архивных и музейных материалов.

Образовательная программа «Архивоведение, документоведение и музейное дело» реализуется в тесном взаимодействии с ведущими профильными учреждениями региона, выступающими базами практики. Ключевыми партнерами являются Государственный архив Костанайской области, Костанайский областной историко-краеведческий музей и Мемориальный музей им. И. Алтынсарина.

Сотрудничество с данными учреждениями обеспечивает студентам доступ к подлинным архивным и музейным фондам, позволяет на практике осваивать методы учета, хранения, описания и цифровизации документов и экспонатов. Практика в профильных организациях способствует формированию профессиональных навыков, развитию исследовательских компетенций и адаптации к условиям будущей профессиональной деятельности.

Значительное внимание уделяется работе с электронными архивами и цифровыми информационными системами. Базы практик способствуют адаптации студентов к профессиональной среде и требованиям рынка труда. Взаимодействие с работодателями повышает уровень трудоустройства выпускников. Практико-ориентированный подход обеспечивает преемственность между теоретическим обучением и профессиональной деятельностью.

При разработке образовательной программы участвуют представители заинтересованных лиц (работодатели, профессорско-преподавательский состав, обучающиеся) в составе коллегиальных органов управления образовательной программой, а также их репрезентативность при принятии решений по вопросам управления образовательной программой. Руководители филиалов кафедры социально-гуманитарных дисциплин принимают активное участие в разработке и утверждении образовательной программы . На постоянной основе проводятся встречи «Обсуждение образовательных программ «История» и «Архивоведение, документоведение и музейное дело» с работодателями и стейкхолдерами».

В филиалах кафедры реализуется научно-исследовательская деятельность студентов кафедры. 28 февраля 2024 года в здании Государственного архива Костанайской области состоялось открытие «Ғылыми залы» кафедры социально-гуманитарных дисциплин КРУ им. А. Байтурсынова и презентация научных трудов профессорско-преподавательского состава кафедры. Открытие в областном архиве аудитории кафедры отвечает требованиям рынка труда и запросам работодателей, которые на сегодняшний день хотят получить дипломированного и компетентного специалиста в своей сфере, а также отвечает требованиям подготовки специалиста гуманитарного направления данной образовательной программы.

За годы обучения студенты осваивают теоретические знания во время практико-ориентируемых занятий в филиалах кафедры в Государственном архиве Костанайской области (дисциплины: Архивоведение, Археография,

Теоретическое источниковедение), Историко-краеведческом музее Костанайской области (дисциплины: Методика историко-краеведческой работы, Музееведение, Музейная педагогика, Музеография).

С учетом специфики профилирующих дисциплин теоретические навыки подкрепляются практическими занятиями непосредственно в филиалах или базах практики. На специальности «Архивоведение, документоведение и музейное дело» особое внимание уделяется практико-ориентированным занятиям, проведение которых организовано преимущественно в филиалах кафедры и базах практик.

Многие дисциплины учебного плана образовательной программы «Архивоведение, документоведение и музейное дело» имеют практико-ориентированную направленность: проведение исследовательской работы, создание исторических сайтов, работа с архивными материалами и музейными экспонатами, проведение исторических игр «Квест», «Исторический лабиринт», проведение выставок и экспозиций, работа в качестве гидов.

Стабильность набора студентов по образовательной программе «6B03202-Архивоведение, документоведение и музейное дело» в дальнейшем может обеспечиваться несколькими условиями. Вот некоторые из них:

1. Актуальность программы: образовательная программа должна соответствовать современным требованиям рынка труда и учитывать тенденции в области архивирования, документооборота и музейного дела. Включение современных технологий и методов в учебные планы может повысить интерес к программе.

2. Партнерство с учреждениями: сотрудничество с архивами, музеями и другими организациями, которые могут предложить стажировки и практику, поможет студентам получить практический опыт и сделает программу более привлекательной.

3. Привлекательные условия обучения: обеспечение качественного образовательного процесса, наличие квалифицированного преподавательского состава, а также современного оборудования и материалов для обучения.

4. Информационные кампании: эффективное продвижение программы, включая участие в образовательных выставках, активное присутствие в социальных сетях и привлечение внимания к успешным выпускникам.

5. Обратная связь от студентов: регулярный анализ мнений студентов о программе и адаптация курса на основе их отзывов помогут улучшить качество образования и удовлетворить потребности обучающихся.

6. Разработка новых модулей и курсов: введение новых дисциплин или модулей, которые будут сочетать теорию с практикой, поможет привлечь студентов с разнообразными интересами.

Эти условия могут способствовать привлечению студентов в области архивоведения, документоведения и музейного дела.

Данная образовательная программа была сформирована на основании анализа рынка труда и запросов потенциальных потребителей, а также рекомендаций попечительских советов, преподавателей-практиков, которые работают на кафедрах в качестве совместителей. Для этого были проведены следующие мероприятия: в рамках кафедральных недель – встречи со

специалистами, выпускниками прошлых лет, анализ анкет руководителей производственных практик от предприятий, студентов, обсуждение выводов итоговых конференций по производственной практике, участие в ярмарках вакансий, обзор газет, интернет-ресурсов по имеющимся вакансиям.

Регулярно проводятся анкетирование работодателей, встречи и беседы с ними, где высказываются замечания и пожелания по улучшению дальнейшей работы.

Образовательная программа «6В03202 – Архивоведение, документоведение и музейное дело» является инновационной. Инновационная образовательная программа определяет: комплексность мероприятий по созданию и введению в образовательную практику новых и качественно усовершенствованных образовательных программ, отвечающих требованиям работодателей; применение новых, в том числе информационных, образовательных технологий; внедрение прогрессивных форм организации образовательного процесса и активных методов обучения, а также учебно-методических материалов, соответствующих современному мировому уровню.

Также изучен кадровый потенциал учреждений Костанайской области, где специалистами многих учреждений архивного, музейного и документационного обеспечения в большинстве являются выпускники образовательной программы «История» и «Филология».

Уникальность образовательной программы «Архивоведение, документоведение и музейное дело» определяется отсутствием подобных в Реестре образовательных программ, а также тем, что в ней оптимально соединены три основные равнодействующие специализации. В Реестре образовательных программ отдельно определены «Архивоведение, документоведение и документационное обеспечение», «Музейное дело и охрана памятников культуры». Поэтому соединение трех специализаций в одну предопределило важность и значимость специалистов по сохранению культурного наследия в архивах и музеях в виде письменных источников.

Образовательная программа «Архивоведение, документоведение и музейное дело» была разработана на основании следующих документов:

- ГОСО высшего образования, утвержден приказом Министра образования и науки Республики Казахстан от 31 октября 2018 года № 604 (с изменениями и дополнениями от 05.05.2020 г.);

- Национальная рамка квалификаций, утвержденная протоколом от 16 марта 2016 года Республиканской трехсторонней комиссией по социальному партнерству и регулированию социальных и трудовых отношений;

- Отраслевая рамка квалификаций в сфере архивного дела и управления документацией. Утверждена протоколом от «25» августа 2016 года № 3 Отраслевой комиссией по социальному партнерству и регулированию социальных и трудовых отношений.

В связи с введением в 2024 году Профессиональных стандартов, в образовательную программу «Архивоведение, документоведение и музейное дело» были включены следующие Профстандарты:

- Библиотекарь-архивист от 17.04.2024 г.

- Обеспечение сохранности Национального архивного фонда (приказ МКС РК от 25 ноября 2022 года № 338, приказ от 28 сентября 2016 года № 265) 25.11.2023 г.
- Музеевед от 17.04.2024 г.
- Хранитель фонда от 17.04.2024 г.
- Методист экскурсионной деятельности музея от 17.04.2024 г.
- Архивист (в музейном деле) от 17.04.2024 г.

Кроме работы в образовательных учреждениях, выпускники данной ОП могут работать в исторических музеях в качестве научных сотрудников.

Выпускающей кафедрой является кафедра социально-гуманитарных дисциплин, которая ежегодно проводит мониторинг карьеры бакалавров, создан и постоянно обновляется банк данных о выпускниках. Для обеспечения связей с работодателями ежегодно проводится активная работа: направляются запросы-согласования на тематику выпускных квалификационных работ; организуются и проводятся ярмарки вакансий; проводится анкетирование работодателей о качестве знаний, навыков, умений выпускников; ведущие специалисты-практики привлекаются в качестве рецензентов и оппонентов выпускных работ; работодатели получают возможность присутствия на открытых защитах выпускных работ. Данная работа способствует потенциальным работодателям обеспечить свои вакансии за счет привлечения лучших выпускников.

Комплекс многопрофильных компетенций бакалавров социальных знаний по образовательной программе «6B03202 – Архивоведение, документоведение и музейное дело» открывает широкие перспективы их дальнейшего трудоустройства и профессионального роста; обеспечение признания обладателя данного диплома в международном образовательном пространстве.

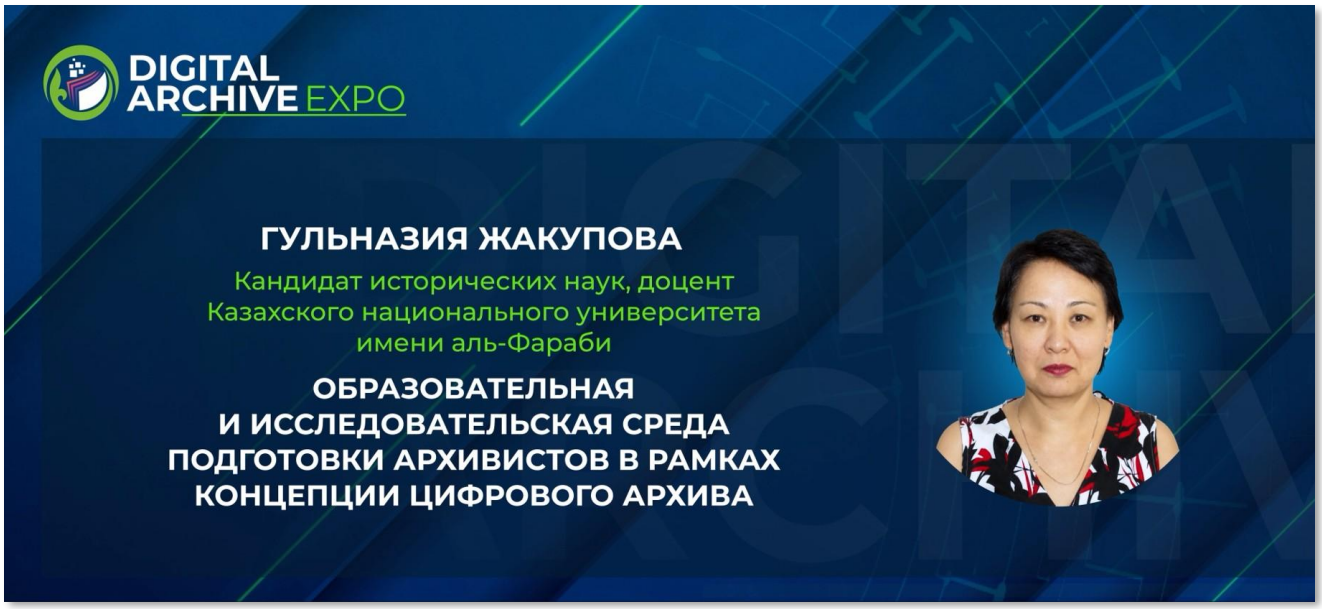
Литература

В Казахстане закроют свыше 700 образовательных программ 01.04.2026. https://24.kz/ru/news/obrazovanie-i-nauka/761372-v-kazakhstane-zakroyut-svyshe-700-obrazovatelnykh-programm

«Музейдегі түн» https://www.facebook.com/groups/261135097922941/posts/1494343721268733/?__cft__[0]=AZaRJmYUHByK9m5fp1DnLEOZKXm5XPH9crIGR518jVaNP0m3KVF89NXFEuHRGYTSUx5p4FtQOvcV8ibMKfxCpkrRjJhXc0MXx--GJQYo2pZgY3rqgMUo0YSCEymEQxgyxhfg8t5Ug7AE8LYw8OvJavCRmIUI8h2jYoDIdlUH-GRnfJBd4tkmZBsSmz8rIVxJ9Z0aB9701BPEWoAOHLTeaYEJ&__tn__=%2CO%2CP-R

Об утверждении государственных общеобязательных стандартов высшего и послевузовского образования (ГОСО РК) https://adilet.zan.kz/rus/docs/V2500036016#z7

30 октября 2025 года в Костанайском областном историко-краеведческом музее состоялся круглый стол, посвященный 190-летию со дня рождения великого казахского ученого Чокана Валиханова. https://www.facebook.com/groups/261135097922941/posts/1761374447898991/?__cft__[0]=AZb_9vi2Po2ICQw_vVXjH0EIcts8G0AjShrWLE6me0gWb2thuGCXDnRpIZjeBStdIHdPwhziksClrJtMl8Puw5VHLTgnxZFmP-8CSHBuxjjmE4AAEKOhhfNshl0wGsQkznr9ZT4AsgKZ7ZIlSwmumzn16GAWTgrrCyz0abQLUZ10WQkTF6sBVHmIUHXpcDtcasTid0Q3CTeZaUZEt7j2QBvy&__tn__=%2CO%2CP-R



ОБРАЗОВАТЕЛЬНАЯ И ИССЛЕДОВАТЕЛЬСКАЯ СРЕДА ПОДГОТОВКИ АРХИВИСТОВ В РАМКАХ КОНЦЕПЦИИ ЦИФРОВОГО АРХИВА

Гульназия Толгаевна Жакупова
Кандидат исторических наук, доцент кафедры всемирной истории, источниковедения и историографии, Казахский национальный университет им. аль-Фараби, Казахстан
gulnaztolgaevna777@gmail.com; https://orcid.org/0000-0003-1519-9043

Аннотация. На современном этапе происходит трансформация традиционного труда в цифровой формат в рамках Государственного проекта Digital Qazaqstan. Президент К. Токаев на Национальном курултае заявил о необходимости создания национального цифрового архива, который станет основой для популяризации культуры и наследия страны в глобальном цифровом пространстве. Концепция цифровых архивов оказывает значительное влияние на осуществление архивной деятельности и архивного образования. В XXI веке модель профессиональной компетентности архивиста, безусловно, будет меняться и развиваться. Архивисты должны обладать различными компетенциями для повышения эффективности своей работы. Цель статьи – определить особенности образовательной и исследовательской среды для подготовки кадров на трех уровнях: бакалавр, магистратура, Phd. Задачи исследования: анализ текущей среды подготовки кадров; рекомендации по улучшению академической и исследовательской среды и формирование информационной культуры. В исследовании использованы методы анализа литературы, анализа образовательной среды подготовки архивистов. На основе результатов исследования было установлено, что выпускники данных уровней должны обладать разными компетенциями и их академическая и исследовательская среда должна иметь связь с архивной отраслью, программами сертификации, стратегиями научного развития.

Ключевые слова: архивоведение; документоведение; цифровой архив; архивное образование; компетенции; академическая и исследовательская среда; бакалавр; магистратура; докторантура; коллаборация; информационная культура, цель и результаты дисциплин

САНДЫҚ АРХИВ ТҰЖЫРЫМДАМАСЫ АЯСЫНДА АРХИВШІЛЕРДІ ДАЯРЛАУҒА АРНАЛҒАН БІЛІМ БЕРУ ЖӘНЕ ЗЕРТТЕУ ОРТАСЫ

Гүлназия Толғақызы Жақұпова
Тарих ғылымдарының кандидаты, әлемдік тарих,
деректану және тарихнама кафедрасының доценті,
әл-Фараби ат. Қазақ ұлттық университеті, Қазақстан
gulnaztolgaevna777@gmail.com; https://orcid.org/0000-0003-1519-9043

Аннотация. Қазіргі уақытта дәстүрлі еңбек «Цифрлық Қазақстан» мемлекеттік жобасы аясында сандық форматқа түрленуде. Ұлттық құрылтайда Президент Қасым-Жомарт Тоқаев елдің мәдениеті мен мұрасын жаһандық сандық кеңістікте танымал етуге негіз болатын ұлттық сандық архив құру қажеттілігін мәлімдеді. Сандық архив тұжырымдамасы архив қызметі мен архивтік білім беруге айтарлықтай әсер етеді. XXI ғасырда архившілердің кәсіби құзыреттілігінің моделі сөзсіз өзгеріп, дамиды. Архившілер жұмыс тиімділігін арттыру үшін әртүрлі құзыреттіліктерге ие болуы керек. Мақаланың мақсаты – үш деңгейде: бакалавр, магистр және PhD деңгейінде кадрларды даярлауға арналған білім беру және зерттеу ортасының сипаттамаларын анықтау. Зерттеудің мақсаттары - қазіргі оқу ортасын талдау; академиялық және зерттеу ортасын жақсарту бойынша ұсыныстар беру; және ақпараттық мәдениетті дамыту. Зерттеуде архившілерді даярлау үшін әдебиеттерге шолу және білім беру ортасын талдау қолданылды. Нәтижелерге сүйене отырып, осы деңгейдегі түлектердің әртүрлі құзыреттіліктерге ие болуы керек екендігі және олардың академиялық және зерттеу ортасы архив саласымен, сертификаттау бағдарламаларымен және ғылыми даму стратегияларымен байланысты болуы керек екендігі анықталды.

Түйін сөздер: архив ісі; құжаттама; цифрлық архив; архивтік білім беру; құзыреттіліктер; академиялық және зерттеу ортасы; бакалавриат; магистратура; докторантура; ынтымақтастық; ақпараттық мәдениет; пәндердің мақсаттары мен нәтижелері

EDUCATIONAL AND RESEARCH ENVIRONMENT FOR ARCHIVIST TRAINING WITHIN THE CONCEPT OF A DIGITAL ARCHIVE

Gulnaziya T. Zhakupova
Candidate of Historical Sciences, Associate Professor,
Department of World History, Source Studies, and Historiography,
Al-Farabi Kazakh National University, Kazakhstan
gulnaztolgaevna777@gmail.com; https://orcid.org/0000-0003-1519-9043

Abstract. Currently, traditional labor is being transformed into a digital format within the framework of the Digital Kazakhstan state project. At the National Kurultai, President Kassym-Jomart Tokayev declared the need to create a national digital archive, which will serve as the basis for popularizing the country's culture and heritage in the global digital space. The concept of digital archives has a significant impact on archival activities and archival education. In the 21st century, the model of archivist professional competence will undoubtedly change and develop. Archivists must possess various competencies to improve their work effectiveness. The purpose of this article is to identify the characteristics of the educational and research environment for training personnel at three levels: bachelor's, master's, and PhD. The objectives of the study are to analyze the current training environment,

provide recommendations for improving the academic and research environment, and develop an information culture. The study utilized literature review and an analysis of the educational environment for training archivists. Based on the results, it was determined that graduates at these levels should possess various competencies, and their academic and research environment should be linked to the archival industry, certification programs, and scientific development strategies.

Keywords: archival science; records management; digital archive; archival education; competencies; academic and research environment; bachelor's; master's; doctoral; collaboration; information culture; purpose and results discipline

Введение

Важность подготовки современных архивистов актуализируется в контексте Государственного проекта Digital Qazaqstan и взаимодействия с ЮНЕСКО: «Мы уделяем большое внимание внесению объектов материального и нематериального наследия Казахстана в международные списки. Правительству следует активизировать работу по включению нашего документального наследия в реестр «Память мира», который наглядно демонстрирует интеллектуальный вклад наций в сокровищницу знаний человечества (Информационный бюллетень о выступлении Президента на Национальном Курултае 2026).

Казахский национальный университет им. аль-Фараби проводит подготовку кадров в области архивоведения и документоведения на трех уровнях согласно Болонскому процессу: бакалавриат, магистратура и докторантура (Phd) (2017 г. – докторантура; 2019 г. – магистратура «7М03207 – IT Архивоведение и документоведение»; 2021 г. – бакалавриат «Цифровое архивоведение и документоведение»).

Современная архивная практика и образовательная среда подготовки архивистов будет продолжать трансформироваться под возрастающим давлением технологий. В данной статье рассмотрим некоторые аспекты образовательной среды подготовки современных архивистов при новой концептуальной структуре архивов, которая интегрирует состояния физических и цифровых архивных документов в сложном цифровом мире, чтобы удовлетворять информационные потребности общества

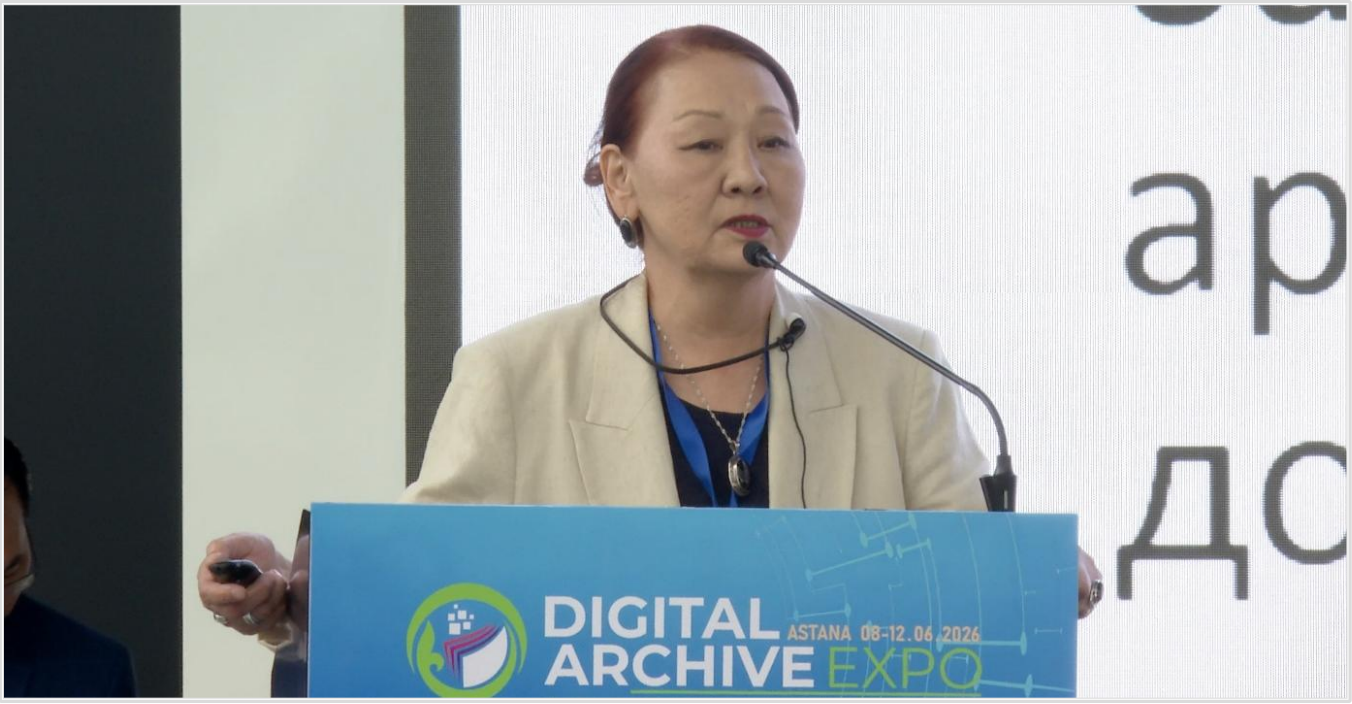
Методология и материалы исследования

В исследовании использован системный подход, который способствует выявлению особенностей академической и исследовательской среды подготовки кадров на трех уровнях как системы. Сравнительный метод позволяет определить особенности формирования компетенций в трехуровневой системе образования. Материалами исследования являются договор с РГГУ, учебные планы образовательных программ и каталоги, план развития по образовательным программам университета по направлению подготовки «Цифровое архивоведение и документоведение и IT-Архивоведение и документоведение», отраслевые рамки квалификации Республики Казахстан.

Обсуждение

Анализ исследований проводится в двух направлениях: компетенции современного архивиста и создание образовательной и исследовательской среды и связи с профессиональной сферой. М. Шленова подчеркивает

необходимость «отхода от традиционных подходов в обучении и внедрении компетентности ориентированных, междисциплинарных и технологических интегрированных моделей. Важно обеспечить практико-ориентированность обучения, что реализуется через стажировку, проектирование долговечности и двойного образования, которое позволяет обучающимся продвигать теоретическую и практическую подготовку профессионального образования» (Shlenova 2025: 173).



Необходимые современные компетенции и подходы к подготовке кадров в области архивоведения и документоведения отмечает К. Коринн, Л. Дюранти: «Важным является предоставить архивистам и специалистам по работе с документами необходимые теоретические и методологические знания, а также процедурные и стратегические навыки для разработки, внедрения и мониторинга аспектов системы цифрового документооборота и сохранения» (Rogers et al. 2011: 1). М.П. Ристийоно, Т.Д. Рахарджа, М.В. Дарвина отмечают, что, профессия – это работа, требующая компетентного образования. Профессионализм человека демонстрируется его результатами работы, исходя из имеющихся у него компетенций. Компетентность связана с профессионализмом. Уровень компетентности, обеспечивающий добросовестность, измеряется результатами работы человека (Ristiyono et al. 2022: 120). Систему компетенций предоставляют «Ключевые профессиональные организации, такие как Ассоциация канадских архивистов (ACA) и Ассоциация архивов и записей (ARA)» (Professional Competence of an Archivist 2017: 149).

Результаты

«Предстоит создать национальный цифровой архив, который будет доступен отечественным и зарубежным разработчикам нейросетей», подчеркнул Президент К. Токаев (Информационный бюллетень о выступлении Президента

на Национальном Курултае 2026). Безусловно, роль электронной информации в архивах оказывает значительное влияние на многие секторы, особенно на подготовку кадров и создание образовательной, исследовательской среды для формирования современных компетенций.

Анализ текущей академической среды по реализации образовательных программ на трех уровнях (единственный университет Республики Казахстан, который готовит кадры на трех уровнях):

- на уровне магистратуры – реализует двухдипломную программу с вузом-партнером – Российским государственным гуманитарным университетом (РГГУ);
- на уровне докторантуры – с РГГУ и зарубежными учеными проводится руководство и оппонирование докторских диссертаций; совместно с зарубежными со руководителями публикуются статьи в журналах с высоким импакт-фактором;
- на уровне бакалавриата – по образовательной программе «Цифровое архивоведение и документоведение» привлекаются преподаватели факультета информационных технологий университета;
- совместно с государственными архивными учреждениями проводится ряд мероприятий: научно-практические конференции, Международная летняя школу;
- работодатели архивной сферы привлекаются для разработки образовательных программ и международной аккредитации, ведения занятий и практик для обучающихся;
- имеется на базе работодателей совместная кафедра с Архивом Президента;
- на трех уровнях образовательные программы прошли международную аккредитацию.

При анализе учебных планов подготовки кадров необходимо отметить, что они имеют модули по формированию традиционных и цифровых компетенций выпускников образовательных программ (Учебные планы ОП 2025).

Каталог элективных дисциплин. Описание дисциплины (Интегрированный учебный план 2025) (*таблица 1*).

***Таблица 1**. Формирование цифровых компетенций по учебным планам образовательных программ*

Уровень подготовки	**Название дисциплин**	**Результаты / цель дисциплины**
бакалавриат	Основы алгоритмизации и программирования	приобретение студентами знаний и навыков в области разработки прикладных программ, обработки информации, системного применения средств информационной технологии для решения прикладных задач
	Информационная безопасность и защита информации	понимать – состав защищаемой информации, структуру угроз информации видов и методов деятельности по обеспечению информационной безопасности, основы правового и организационного обеспечения

		информационной безопасности, классификации информационной безопасности, анализ методов и средств защиты информации
	Основы искусственного интеллекта	изучение основных направлений исследований в области искусственного интеллекта и решение задач с интеллектуальным содержанием
	Программирование на языке Python	сформировать способность применять подходы объектно-ориентированного и функционального программирования, позволяющие быстро перейти к решению задач в соответствующих предметных областях
	Методы и технологии разработки, автоматизированных НСА	сформировать у студентов практические навыки и умения в области разработки автоматизированного НСА
	Веб-архивирование	способность сформировать навыки и знания в области поиска, сбора, обработки, хранения, распространения и использования информации
	Технотронные архивы	сформировать у студентов теоретические знания и практические навыки в области создания, сохранения и использования документов технотронных архивов
	Веб-программирование	подготовка студентов к проектно-технологической деятельности в области создания компонентов программных комплексов и баз данных, автоматизации технологических процессов с использованием современных инструментальных средств и интернет-технологий
	Базы данных	проектирование, создание и ведение баз данных
	Создание и эксплуатация архивов оцифрованных документов	сформировать у студентов системное представление о современных технологиях оцифровки документов, обучить методам работы со сканерами и другими техническими средствами в архиве
	Автоматизированные архивные технологии	сформировать базовые знания о роли автоматизированных технологий для архивного дела
	Введение блокчейн-технологий	формирование способности обосновывать, проектировать и применять технологию блокчейна в практической работе
	Облачные технологии	использование облачных вычислений в формировании новой ИТ- инфраструктуры, освоение технологий создания облачного сервиса, работа с существующими облачными сервисами
	Разработка приложений баз данных	формирование умения применять методологию проектирования и разработки приложений баз данных к крупномасштабным проектам

магистратура	Цифровые технологии обеспечения доступа к архивным документам	комплексное представление о современных формах и путях обеспечения доступа к архивной информации, введении архивных документов в научный оборот, информационной деятельности архивов
	Информационные системы в управлении	формирование теоретических знаний и навыков их применения в решении практических задач использования информационных систем управления
	Облачные сервисы хранения и использования документов	дать представление об «облачных сервисах» для архивной работы, знания в области электронного государственного управления и их применении в решении практических задач управления документными коммуникациями ЭП
	Системы искусственного интеллекта в управлении документами	сформировать теоретические знания и практические навыки использования систем искусственного интеллекта в управлении документами
магистратура двудипломная ОП с вузом-партнером (Интегрированный учебный план)	Цифровое государственное управление	решения задач по управлению документами и государственными данными, в том числе обеспечения их качества, сбора, хранения, обработки, предоставления, распространения и уничтожения посредством инфраструктуры управления данными в государственных органах
	Управление метаданными электронных документов	изучение методов формирования, хранения и использования метаданных для обеспечения аутентичности и сохранности документов
	Роботизированные технологи в управлении документами	способен совершенствовать систему ДОУ организации и архива на основе современных информационных технологий
докторантура	Управление метаданными электронных документов	сформировать способность систематизировать и интерпретировать стандарты метаданных и применять их в информационном обеспечении управления
	Информационные системы в архивоведении и документоведении	формирование способности работать в современных информационных системах, применяемых в архивном деле и документационном обеспечении управлении
	Инновационные технологии в архивном деле	сформировать способность применять современные инновационные технологии для решения проблем автоматизации основных процессов деятельности (комплектование, учет, хранение, использование) архивных учреждений

Обучающиеся КазНУ им. аль-Фараби разрабатывают ряд тем.

Бакалавриат – «Создание приложения по классификации документов»;

Магистратура – «Интеллектуальная платформа генерации архивных документов с искусственным интеллектом»;

Докторантура – «Археографическая деятельность архивов на современном этапе в электронной среде» – разработан веб-сайт «E-Doc Complex» и т.д.

Исследовательская среда в образовательной сфере – это создание оригинальных научных разработок по фундаментальным и прикладным изысканиям транс- и кроссдисциплинарного характера на основе формирования исследовательского климата взаимодействия между основными участниками научного процесса. Ее основной принцип – сотрудничества (коллаборации) – объединение ресурсов для достижения результатов и дальнейшего развития. Коллаборация благодаря объединению человеческих и материальных, информационных и финансовых, внутренних и внешних ресурсов усиливает взаимодействие и кумулятивный эффект развития организации. Формируются сети взаимодействия научного, образовательного, академического, инновационного направления, отражающие диффузию академической культуры, трансфер знаний, перетоки знаний и инноваций, создание новых инновационных товаров и продуктов. Стратегии сотрудничества и коллаборации укрепляют внешние и внутренние связи, академическое лидерство организации.

Текущая образовательная и исследовательская среда по формированию современных компетенций требует совершенствования, т.к. не все темы диссертационного исследования магистрантов и докторантов связаны с профессиональной сферой (коллаборация научных исследований). По личной инициативе ряд докторантов создает программные продукты для архивной практики, которые не находят применения.

Заключение

Необходимо усовершенствовать академическую и исследовательскую среду для подготовки кадров, адаптированных к современным потребностям, в сфере информационных технологий. Необходима ориентация учебного содержания на освоение теории и практики информационно-аналитической, информационно-технологической и социально-коммуникационной деятельности современного специалиста, способного применять технологии в сфере документоведения и архивоведения. Интеграции цифровых ресурсов в учебные программы трансформация и изменение традиционных методик преподавания в ответ на цифровизацию.

Ключевым является стратегическое развитие образовательных программ и продвижение цифровой компетентности как основного компонента профессиональной подготовки. Включать в учебный план комбинированные курсы. Освещать теоретические концепции с помощью практических примеров из реальной жизни, взятых из тематических исследований, основанных на конкретных административных и технологических контекстах. Предоставлять обучающимся содержание и структуру для курсов по управлению и сохранению цифровых документов и ссылки на дополнительные исследования и ресурсы.

Формировать способности создавать цифровые/электронные архивы, управлять динамическими и статическими архивами в электронной форме, предоставлять электронные архивные услуги населению в цифровой форме, сохранять архивы в цифровой форме в электронном виде, поддерживать цифровые архивы с использованием цифровых/электронных средств. Основные технические и цифровые компетенции: сохранение и архивирование цифровых данных: управление изначально цифровыми и оцифрованными записями, включая понимание форматов файлов, технологий хранения и стратегий миграции; метаданные и стандарты: внедрение стандартов описания архивов (например, EAD, DACS) как для физических, так и для цифровых объектов; по информационным системам: владение системами управления базами данных, системами управления контентом (CMS) и программным обеспечением для цифровых хранилищ.

На трех уровнях подготовки – бакалавриат, магистратура и докторантура – необходимо создавать исследовательскую среду по заказу работодателей с дальнейшем внедрением в сферу. Для этого архивные учреждения предоставляют вузу научные планы и проекты. По этим планам выпускники разрабатывает конкретные темы исследования, необходимые для профессиональной сферы (особенно на уровне магистратуры и докторантуры).

Пересмотреть план развития образовательных программ на трех уровнях подготовки и создания технологической академической среды совместно с университетом и архивной сферой. Провести профессиональную аккредитацию образовательных программ, которая усилит взаимную ответственность за подготовку современных кадров.

Разработка сертифицированных образовательных программ на всех уровнях подготовки (на уровне бакалавриата – сертифицированный экзамен по важным учебным дисциплинам (создать сертифицированные центры совместно с вузами и архивными учреждениями); на уровне магистратуры и докторантуры – сертификат об апробации или внедрении в профессиональную сферу.

Литература

Professional Competence of an Archivist 2017 – Professional Competence of an Archivist: Changes That Can Be Expected. 2017, 27 (2). P. 149-153. https://doi.org/10.33700/2670-451X.27.2.149-153(2017)

Ristiyono et al. 2022 – Ristiyono M.P., Raharja T.D., Darwina M.W. The competencies of archivist on the digitalization into one data. NTERNATIONAL *JOURNAL FOR INNOVATIVE RESEARCH IN MULTIDISCIPLINARY FIELD*. Vol. 8. Issue 3. P. 119-125. DOIs:10.2015/IJIRMF/202203019

Rogers et al. 2011 – Rogers C., Duranti L. Educating Records Professionals on Topics in Digital Preservation. University of British Columbia. P. 1-8. http://www.ica-sae.org/Education%20Modules.pdf

Shlenova 2025 – Shlenova M. Modern trends in the professional training of future specialists in library, information, and archival studies in higher technical education institutions. Modern Information Technologies and Innovation Methodologies of Education in Professional Training Methodology Theory Experience Problems. DOI:10.31652/2412-1142-2025-75-173-182 Kharkiv National Aerospace University. P. 173-182.

Интегрированный учебный план 2025 – Интегрированный учебный план. Договор о реализации образовательной программы магистартуры в сетевой (совместной) форме. Ноябрь 2025. РГГУ (46.04.02 Документоведение и архивоведение с направленностью «Теория и практика работы с электоронными документами и управлении и архивах»)и КазНУ им. Аль-Фараби (7М03208 – Документоведение, архивоведение, документационное обеспечение управления). https://univer.kaznu.kz/ (дата обращения: 23.03.2026).

Информационный бюллетень о выступлении Президента на Национальном Курултае 2026 – Информационный бюллетень о выступлении Президента Касым-Жомарта Токаева на Национальном Курултае. 20 января 2026 года. https://www.inform.kz/ru/natsionalniy-tsifrovoy-arhiv-sozdadut-v-kazahstane-22e55f (дата обращения: 05.02.2026).

Учебные планы ОП 2025 – Учебные планы образовательных программ (для набора 2025). Каталог элективных дисциплин. Описание дисциплины. https://univer.kaznu.kz/ (дата обращения: 23.03.2026).

ФОТОГАЛЕРЕЯ

PHOTO GALLERY

Встреча руководства Архива Президента Республики Казахстан с представителями Президентского архива Республики Корея: заместителем директора отдела административного планирования Президентского архива Кореи Ким Ёнхи (вторая слева) и архивистом Отдела архивной службы Президентского архива Республики Корея Ко Имджон (третья слева)
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

Встреча руководства Архива Президента Республики Казахстан с делегацией Национального Управления архивов и документации Султаната Оман
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

Встреча директора Архива Президента Республики Казахстан А. Мустафиной (справа)
с вице-президентом Национального архива Японии И. Фуруя
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

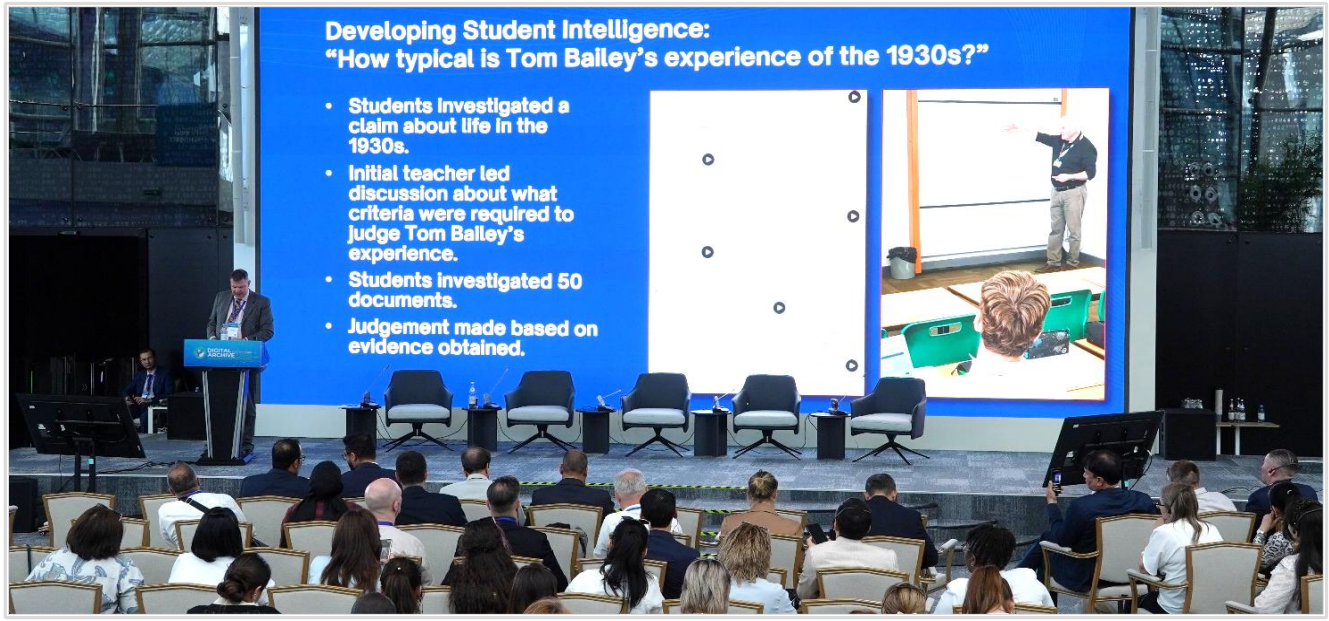
Во время работы одной из сессий Конгресса
Слева направо: директор Департамента качества Национального Управления архивов
и документации Султаната Оман Азан бин Хамис бин Мохаммед аль-Субхи,
председатель Национального Управления архивов и документации Султаната Оман
Хамад бин Мохаммед аль-Дувайни
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

Выступление с приветственным словом директора Архива Президента Республики Казахстана А. Мустафиной на семинаре ЮНЕСКО «Искусственный интеллект: последствия для международного реестра «Память мира» г. Астана, Международный центр искусственного интеллекта ALEM.AI 10 июня 2026 г.

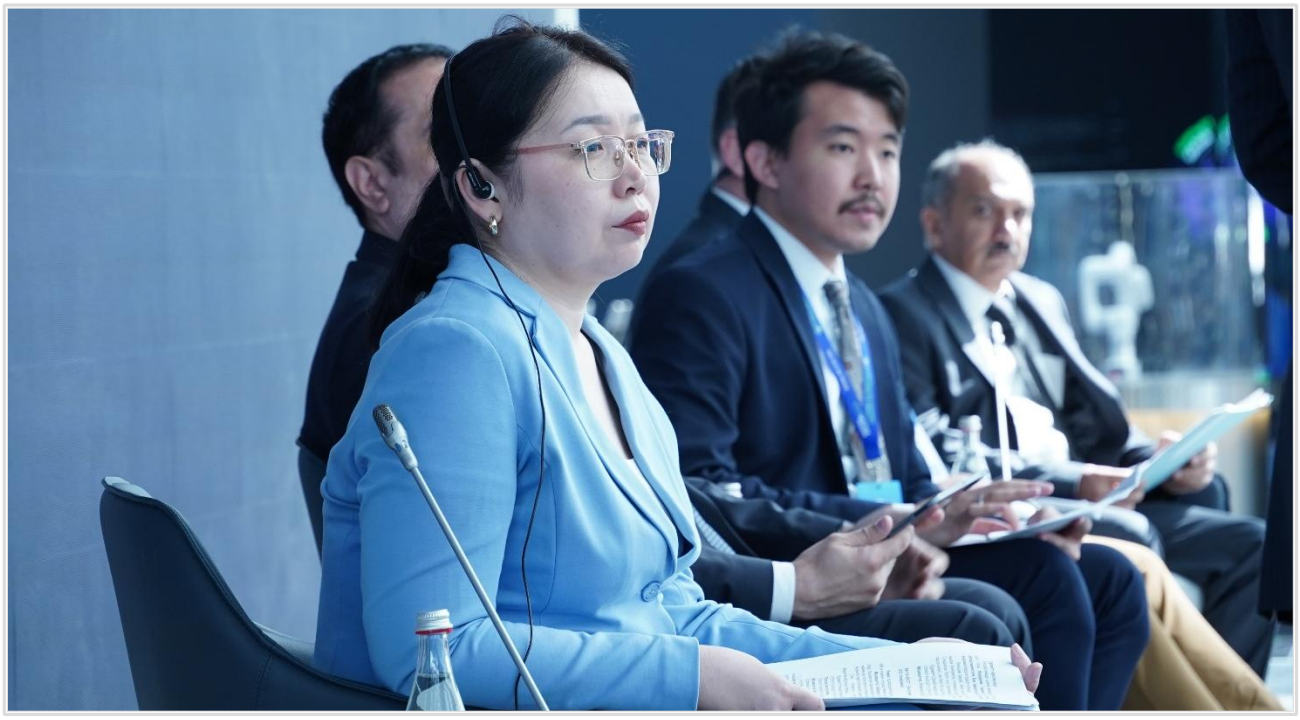
Слушатели семинара ЮНЕСКО «Искусственный интеллект: последствия для международного реестра «Память мира» г. Астана, Международный центр искусственного интеллекта ALEM.AI 10 июня 2026 г.

Неформальное общение участников Конгресса
с руководством Архива Президента Республики Казахстан
Слева направо: 1) документалист Африканского союза САФГРАД Дж. Ризгаллах
(Буркина-Фасо); 2) научный сотрудник Университета Сорбонна Абу-Даби Л. Шонхе (ОАЭ);
3) советник директора Архива Президента Республики Казахстан Р. Карибжанова;
4) директор Архива Президента Республики Казахстан А. Мустафина
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

Памятная фотография участников Конгресса
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

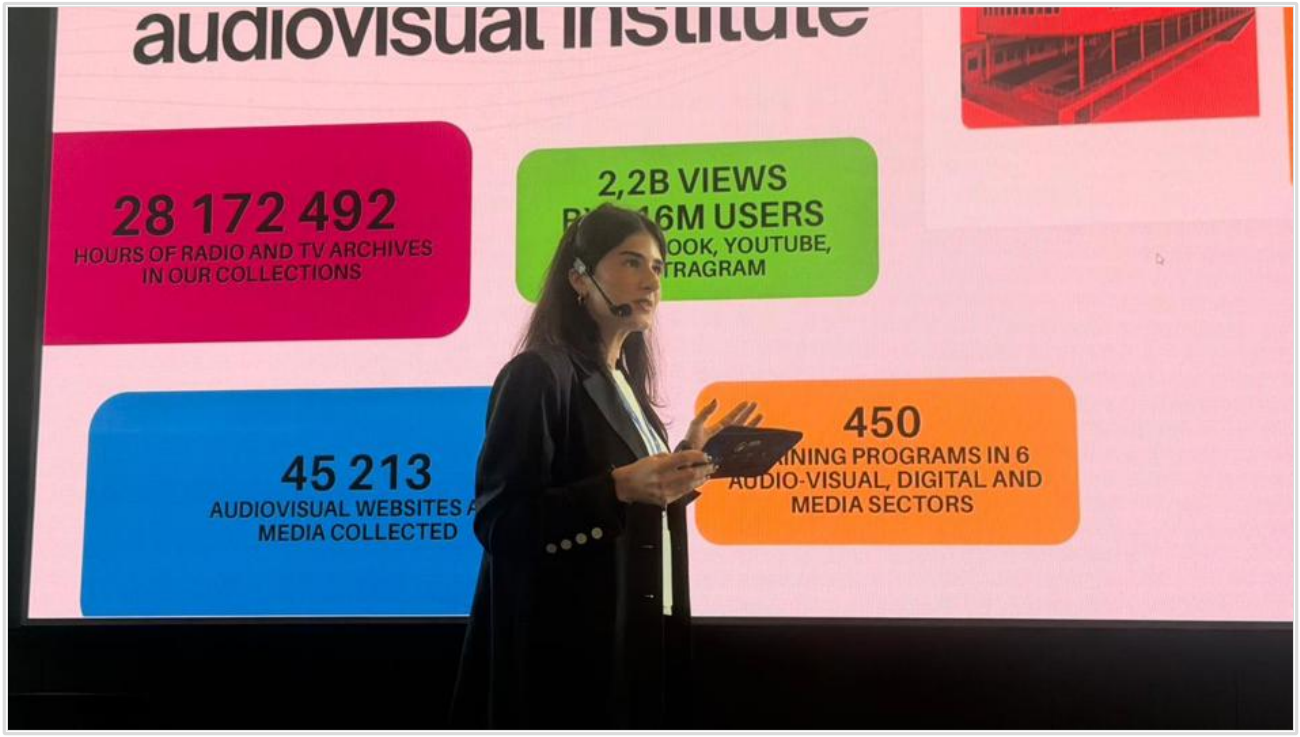


Авторский семинар руководителя отдела образования и информационно-просветительской работы в Национальном архиве Великобритании Э. Пейна «Улучшенная история: использование ИИ, чтобы помочь учащимся развить более эффективное мышление при работе с документами»
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.



На заседании 3-й панельной сессии «Инновационные экосистемы архивов: инфраструктура, сотрудничество и устойчивость»
Слева направо: руководитель Службы цифровых технологий и инноваций Архива Президента Республики Казахстан К. Кабарова; помощник специалиста Управления государственных архивов при Президенте Турецкой Республики О. Куртулуш; директор Архива научно-технической документации Азербайджанской Республики Ф. Аббасов
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

Выступление руководителя информационной службы Агентства «Узархив» (Узбекистан) А. Алиева на сессии MURA talks «Цифровые архивы: новые возможности для государства, науки, бизнеса и креативной индустрии»
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.



Выступление координатора регионального проекта по сохранению аудиовизуального и кинематографического наследия в Центральной Азии Э. Грасии (Франция) на сессии MURA talks
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

Коллективное фото участников сессии MURA talks
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

Модераторы, спикеры и эксперты 4-й панельной сессии
«Архивное будущее: новые компетенции и лидерство»
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

Модераторы 5-й панельной сессии «Архивы для общества»
Слева направо: заведующий сектором перспективных исследований Центра научных и прикладных исследований Архива Президента Республики Казахстан Ш. Тлепина, архивист аудиовизуальных архивов Национального архива Сингапура (NAS) Джун Пок
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

Спикеры и модераторы 5-й панельной сессии «Архивы для общества»
г. Астана, Международный центр искусственного интеллекта ALEM.AI
10 июня 2026 г.

Встреча директора Архива Президента Республики Казахстан А. Мустафиной (справа) с генеральным директором Департамента Национальной библиотеки Иордании Фирассом Аталлахом Мохаммадом Дарабой
г. Астана, Международный центр искусственного интеллекта ALEM.AI
11 июня 2026 г.

На торжественной церемонии закрытия III Конгресса архивистов «Digital Archive. EXPO»
г. Астана, Международный центр искусственного интеллекта ALEM.AI
11 июня 2026 г.

Подведение итогов Конгресса. Выступление заместителя директора Архива Президента Республики Казахстан К. Алимгазинова г. Астана, Международный центр искусственного интеллекта ALEM.AI 11 июня 2026 г.

Выступление с заключительным словом директора Архива Президента Республики Казахстан А. Мустафиной г. Астана, Международный центр искусственного интеллекта ALEM.AI 11 июня 2026 г.

Коллектив Архива Президента Республики Казахстан
г. Астана, Международный центр искусственного интеллекта ALEM.AI
11 июня 2026 г.

До встречи на IV Конгрессе архивистов
г. Астана, Международный центр искусственного интеллекта ALEM.AI
11 июня 2026 г.

«ЦИФРЛЫҚ АРХИВТЕР: МЕМЛЕКЕТ, ҒЫЛЫМ, БИЗНЕС ЖӘНЕ КРЕАТИВТІ ИНДУСТРИЯ ҮШІН ЖАҢА МҮМКІНДІКТЕР»
MURA-talks

MURA-talks
«ЦИФРОВЫЕ АРХИВЫ: НОВЫЕ ВОЗМОЖНОСТИ ДЛЯ ГОСУДАРСТВА, НАУКИ, БИЗНЕСА И КРЕАТИВНОЙ ИНДУСТРИИ»

«DIGITAL ARCHIVES: NEW OPPORTUNITIES FOR THE STATE, SCIENCE, BUSINESS AND THE CREATIVE INDUSTRY»
MURA-talks

THE FUTURE OF HUMAN MEMORY:
HOW TO PRESERVE DIGITAL HERITAGE IN 100 YEARS

Fackson Banda
Chief of the Documentary Heritage Unit, UNESCO
f.banda@unesco.org

АДАМЗАТ ЖАДЫНЫҢ БОЛАШАҒЫ:
ЦИФРЛЫҚ МҰРАНЫ 100 ЖЫЛ БОЙЫ ҚАЛАЙ САҚТАУҒА БОЛАДЫ

Факсон Банда
ЮНЕСКО-ның Құжаттық мұра бөлімінің басшысы
f.banda@unesco.org

БУДУЩЕЕ ЧЕЛОВЕЧЕСКОЙ ПАМЯТИ:
КАК СОХРАНИТЬ ЦИФРОВОЕ НАСЛЕДИЕ НА 100 ЛЕТ

Факсон Банда
Руководитель подразделения документального наследия ЮНЕСКО
f.banda@unesco.org

Imagine for a second that Toy Story 2 almost disappeared forever.

Not because of a fire or a flood – but because someone typed a wrong command and started deleting files.

Pixar's entire movie – years of work – gone in seconds.

Only one person had a backup at home. That's what saved it.

Now multiply that problem by billions.

Our generation is creating more digital memory than all of humanity before us.

But without deliberate action, most of it will vanish within 100 years – not from malice, but from neglect, obsolete formats, or broken policies.

So today, I want to give you two keys to unlock the next century of human memory: *Policy Innovativeness and Institutional Adaptiveness*.

Policy innovativeness is about moving from «preservation as storage» to «preservation as a system»

We often think digital preservation is a technical problem.

Better servers. Newer tapes. More cloud storage.

But technology alone fails – because technology changes every 5-10 years.

Policy is what connects one generation of tech to the next.

Here's what *policy innovativeness* means in practice:

1. *Policies that support R&D, not just maintenance*

Most governments fund archives to keep things, not to research how to keep them. UNESCO's Memory of the World Programme is encouraging its *Knowledge Centres* to actively research AI for preservation – like using machine learning to automatically identify at-risk file formats or to restore damaged digital audio.

We need national policies that say: *5% of any preservation budget goes to R&D into tomorrow's problems*.

2. *Policies that mandate "preservation-ready" creation*

The Handbook by the Digital Preservation Coalition (DPC), produced with support from UNESCO, warns: if you don't capture context (units of measurement, software dependencies, relationships between files), the data becomes meaningless gibberish in 50 years.

A policy-innovative government would require that *publicly funded digital research and born-digital records* include a preservation metadata package from day one – not as an afterthought.

3. *Policies that enable legal exceptions for preservation*

Today, copyright often blocks archives from migrating old software or emulating obsolete systems. Policy innovativeness means updating laws so that *non-commercial preservation actions* are legal – even if they require breaking DRM (Digital Rights Management) or copying software.

Without these policies, we are not preserving memory.

We are just hoarding bits that future generations cannot read.

Now, turning to *Institutional Adaptiveness*, this is *the human heart of digital preservation*.

Policies are dead words unless institutions learn to adapt. The DPC Handbook makes this painfully clear: "Organisational change is a major risk to be managed."

Why?

Because digital preservation takes decades, but institutions restructure every 3–5 years.

Departments merge. Leaders leave. Funding shifts. And digital memory gets orphaned.

So, what does an adaptive institution look like?

It distributes responsibility instead of centralizing it

One person or one department owning preservation = a single point of failure.

Adaptive institutions embed preservation skills into multiple roles – IT, records management, legal, even communications.

When someone leaves, the memory doesn't leave with them.

1. *It treats collaboration as a survival strategy*

No single archive can afford to preserve everything alone.

UNESCO's model encourages *MoW Knowledge Centres* to act as hubs:

• Sharing low-cost digitization methods (especially for small island states or LDCs)

• Pooling software emulation tools

• Creating shared training programmes

The most adaptive institutions don't compete – they connect.

2. *It embraces AI without losing human judgment*

AI can help:

• Automatically check file integrity (fixity checks)

• Identify obsolete formats at scale

• Even generate metadata from unlabeled collections

But AI cannot decide what is worth preserving for 100 years.

That is a human, ethical, community-based decision. Adaptive institutions use AI as a **tool**, not as a replacement for curatorial responsibility.

4. *It plans for organisational change*

One practical example: include in every funding agreement a "digital heritage succession clause" – if your institution closes or changes mission, your digital collections must be transferred to a certified trusted repository.

That is adaptiveness written into governance.

To bring these two ideas together, let me give you a concrete scenario.

In 2070, a historian wants to study climate change policy from the 2020s. She finds a government report from 2026 – but the file won't open. The format is obsolete. The software is gone. The metadata is empty. She has nothing but a filename.

That future is preventable – if we act now.

• *Policy innovativeness* would have required that report to be saved with open formats, preservation metadata, and legal permission to migrate it every decade.

• *Institutional adaptiveness* would have ensured that when the ministry closed in 2040, the report moved automatically to a national digital archive, with staff trained to update it.

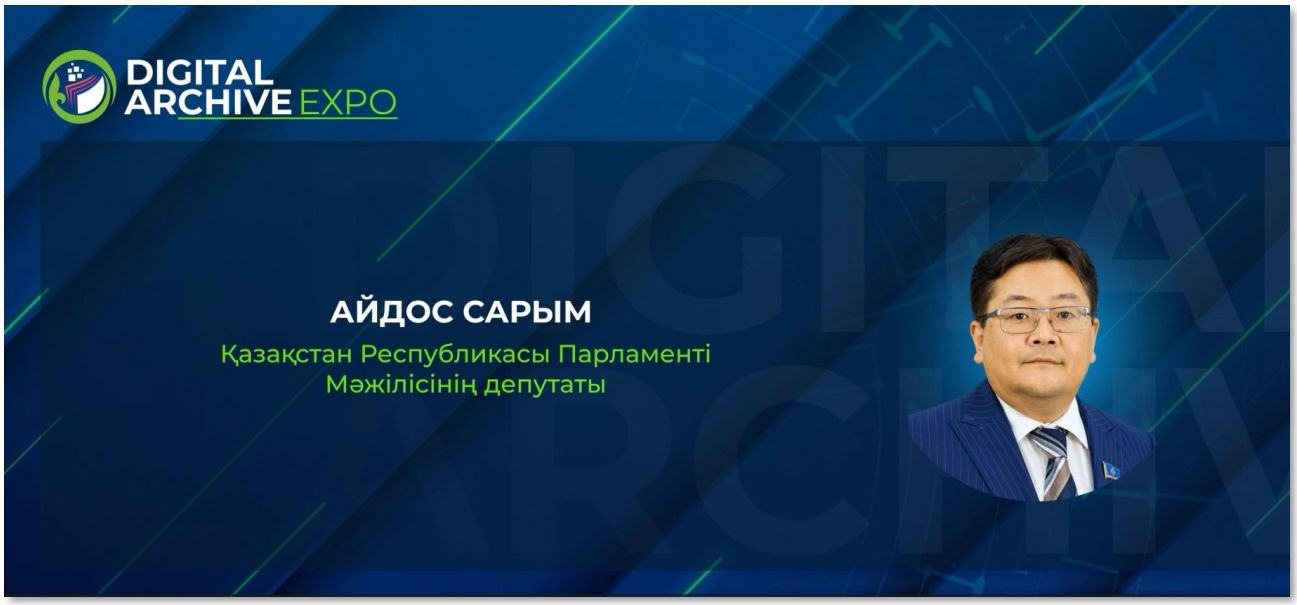
We are the first generation to face the fragility of digital memory.

And we may be the last generation that can still do something about it at low cost.

The future of human memory is not about bits or clouds.

It is about choices – choices to write smarter policies and to build institutions that learn, share, and last.

Let's not leave our grandchildren with empty filenames. Let's give them a living, readable, trustworthy memory of our time.



АРХИВ ЖӘНЕ ЦИФРЛЫҚ ЗАМАН: ТАРИХИ ЖАДЫНЫ САҚТАУДЫҢ ЖАҢА ФОРМАТЫ

Айдос Сарым
Қазақстан Республикасы Парламенті Мәжілісінің депутаты
Sarym@parlam.kz

АРХИВ И ЦИФРОВАЯ ЭПОХА: НОВЫЙ ФОРМАТ СОХРАНЕНИЯ ИСТОРИЧЕСКОЙ ПАМЯТИ

Айдос Сарым
Депутат Мажилиса Парламента Республики Казахстан
Sarym@parlam.kz

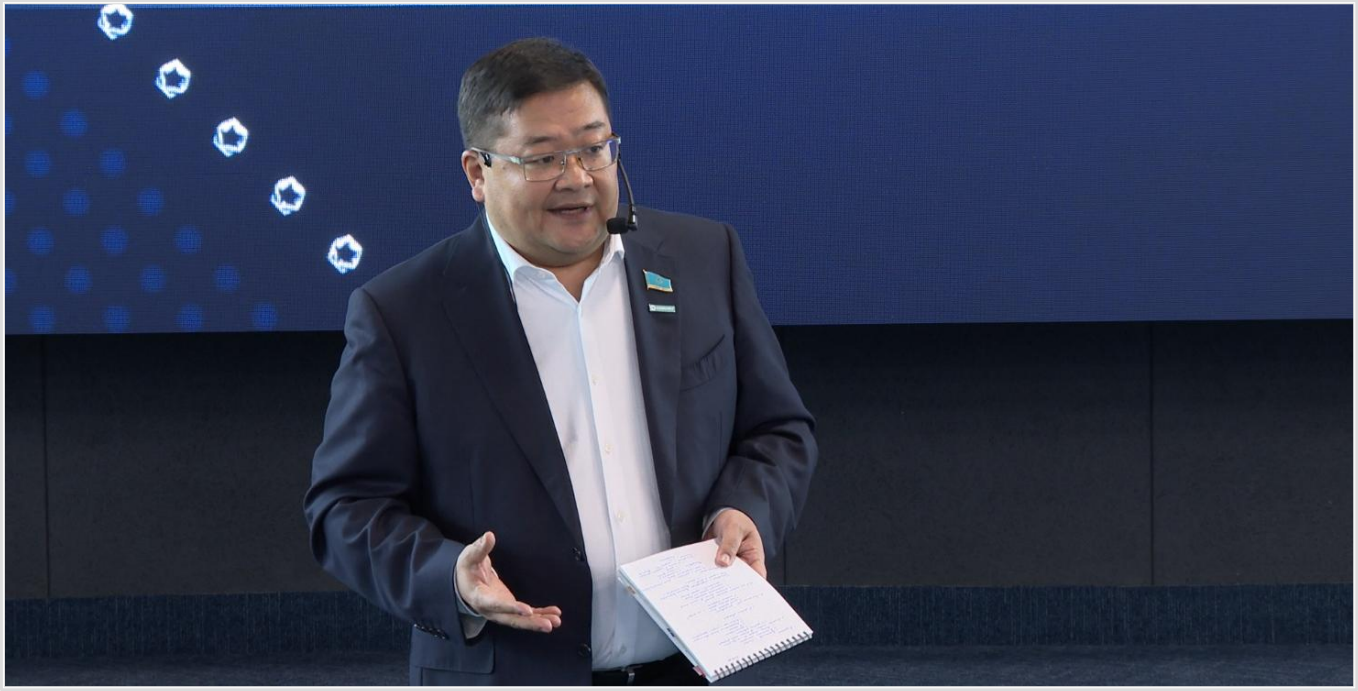
ARCHIVES AND THE DIGITAL ERA: A NEW FORMAT FOR PRESERVING HISTORICAL MEMORY

Aidos Sarym
Member of the Mazhilis of the Parliament of the Republic of Kazakhstan
Sarym@parlam.kz

Қайырлы күн, құрметті достар мен әріптестер!

Бүгінгі таңда елордамыз Астана қаласында ауқымды іс-шара – Архившілердің III Конгресі өтіп жатыр. Оған әлемнің түкпір-түкпірінен келген азаматтар, сондай-ақ ЮНЕСКО мен Халықаралық архивтер кеңесінің өкілдері қатысуда. Сарапшылардың баяндамаларымен таныса келе, жұмыс кестемнің тығыздығына байланысты аталған Конгреске толықтай қатыса алмайтыныма өкініш білдіремін. Дегенмен, болашақта осы жиынның материалдар жинағы жарық көргенде, олармен егжей-тегжейлі танысуды жоспарлап отырмын.

Менің бұл алаңға келуімнің бірнеше маңызды себебі бар. Біріншіден, мен өз еңбек жолымды осы архив саласынан бастағандықтан, өзімді осы саланың түлегі ретінде сезінемін. Кәсіби қызығушылығым мен мүмкіндігіме қарай сіздердің пікірлеріңізді тыңдап, жүздесу мен үшін маңызды. Екіншіден, мамандығым тарихшы болғандықтан, архивтің қазіргі жағдайы мен болашақ даму перспективалары мені бейжай қалдырмайды. Үшіншіден, біз белгілі бір дәрежеде тарихи контентті қалыптастырушы ұйым және мемлекеттік биліктің өкілі ретінде бұл саланың даму бағытын сіздермен бірге талқылауды жөн көрдік. Сіздердің кәсіби пайымдарыңыз бен ұсыныстарыңыз біздің заңнамалық және мемлекеттік жұмысымызға оң ықпал етеді деп есептеймін. Мен сіздерге ақыл айтудан аулақпын, тек өз тәжірибем мен түйген ойларымды бөліскім келеді.



Тарихи сабақтастық пен «Архив революциясы»

Осыдан 35 жыл бұрын, мектеп бітіргеннен кейін, маусым айында мемлекеттік архивке «Қор сақтаушысы» атты ең төменгі сатыдағы қызметке орналастым. Сол кезеңдегі бақылауым бойынша, архив саласына адамдар, негізінен, екі түрлі жолмен келетін. Біріншісі – тарихқа бала кезден ғашық болып, осы саланы серік еткендер, екіншісі – бұл ортаға кездейсоқ келіп, уақытша тұрақтап, кейін ауысып кететіндер.

Сол жылдары архивке алғаш келген қызметкерлерді ең алдымен нөлдік (цокольді) қабатқа түсіріп, XVII ғасырға тиесілі ең алғашқы тарихи құжаттарды көрсететін. Содан кейін жеке архивтердегі ұлы ғалымдарымыз бен жазушы-әдебиетшілеріміздің қолжазбаларымен таныстыратын. Тарихқа ықыласы бар адам үшін бұл өте үлкен әсер қалдыратын сәттер еді. Осындай кезеңдерді бастан өткере келе, біз жай ғана тарихшы емес, кәсіби архившіге айналдық. Құжаттардың артында үлкен идеологиялық мұра мен мемлекеттік биліктің тірегі жатқанын терең түсіндік.

Соңғы 35 жыл ішінде біздің буын жүйелік трансформацияларды ғана емес, Қазақстандағы екі үлкен «Архив революциясын» көрді.

Бірінші толқын: 1990 жылдардың басында Кеңес Одағы (КПСС) заманынан бері құпия сақталып келген, саяси қуғын-сүргін мен ату жазаларына қатысты көптеген құжаттар құпиясыздандырылды. Бұл үдеріс белгілі бір уақыттан кейін баяулап, тоқтап қалды.

Екінші толқын: 2019 жылдан бастап, Мемлекет басшысының бастамасымен Қазақстанда екінші архив революциясы басталды. Бұған дейін ашуға мүмкіндік болмаған Бас прокуратураның, НКВД-ның, КГБ-ның, МВД-ның және сот жүйесінің депортация, саяси қуғын-сүргін мен репрессияларға қатысты құпия істері толықтай ашылды.

Қазіргі таңда 1 миллионнан астам құжат пен 700 мыңға жуық архив ісі құпиялықтан шығарылды. Мұндай ауқымды шара Қазақстан тарихында бұрын-соңды болған емес. Бүгінде Қазақстан Республикасы Президентінің Архиві мен өзге де уәкілетті тұлғалар уақыт өте келе тозып, биологиялық зақымдануға (грибок, бактерия) ұшыраған құжаттарды реставрациялап, ғылыми айналымға енгізу бойынша ауқымды жұмыс жүргізуде. Арнайы комиссияның жұмысы нәтижесінде репрессия тарихына қатысты 30-дан астам том, сондай-ақ ашаршылық пен Алаш қозғалысына арналған көптомдық кітаптар жарық көрді.

Ең бастысы, қазіргі уақытта отандық тарихшылардың басын қосқан «Академиялық 7 томдықты» жазу үстіндеміз. Жақында Тәуелсіздік кезеңін қамтитын 7-томды аяқтауға жақындадық. Бұл — Қазақстан тарихында өз шежіремізді өз бетімізше жазудың алғашқы ірі қадамы. Мемлекет басшысы Алтын Ордаға қатысты халықаралық симпозиумда: *«Осы уақытқа дейін тарихты жеңімпаздар жазып келді, ал жеңілгендердің мұндай мүмкіндігі болмады»*, — деп атап өткен еді. Біз архивтік құжаттарға сүйенеміз, бірақ сол құжаттардың қандай жағдайда және не мақсатпен жазылғанын да сын көзбен ескеруіміз қажет. Ақыр соңында, өз тарихымызды объективті түрде зерделеуге мүмкіндік алдық.

Цифрлық дәуірдегі архив саласы және жаңа сын-қатерлер

Бүгінгі таңда архив саласының алдында «Болашаққа қалай қадам басу керек?» деген стратегиялық мәселе тұр. Мәселе тек құжат сақтайтын қоймалардың толуында немесе қағаздардың тозуында емес. Қазіргі цифрлық дәуірде бұл тақырыпты басты драйвер ретінде алға шығарғандарыңыз өте орынды.

Тарихқа көз жүгіртсек, архив алғашқыда билеушілердің, хандар мен патшалардың өз билігін, жаулап алған жерлерін және шежіресін заңдастыру үшін мемлекеттік институт ретінде құрылған. Сондықтан ертедегі архивтер негізінен соғыс пен билік тарихын қамтиды.

Жаныма жақын сала болғандықтан ба, менің түсінігімде «жұмақтың жұпар иісі» осы архивтің шаңымен үндесетіндей көрінеді.

Уақыт өте келе қоғам дамып, экономикалық, техникалық және архитектуралық құжаттарды сақтау қажеттілігі туындап, өмір күрделене түсті. Бір жағынан, архивтер саяси үдерістердің құралына да айналды. Мұны біз қазіргі Ресей мен Украина арасындағы қалыптасқан жағдайдан көріп отырмыз: архивтерді өзінің алғашқы функциясында, яғни саяси ықпал ету құралы ретінде

пайдалану үрдісі байқалады. Сондықтан тарихтың үлкен қоғамдық-саяси күш екенін ұмытпауымыз керек.

ХХ ғасырда азаматтардың еңбек құқықтары мен әлеуметтік жағдайын растайтын әлеуметтік сервистік функция қалыптасты. Бүгінде біз «креативті индустрия» мен «интертейнмент» (ойын-сауық, танымдық индустрия) кезеңіне жеттік. Бұл ретте архив қызметкерлерінің жалақысы мен штат санының аздығы сияқты өзекті мәселелер бар екенін жақсы білеміз, алайда заман ағымынан қалыс қала алмаймыз.

Кеңес дәуіріндегі танымал туындыларда «телевидение келгенде театр мен кино жойылады» деген болжамдар айтылатын. Бірақ қазіргі технологиялардың ерекшелігі – олар алдыңғы функцияларды жоймайды, қайта жаңа бағыттармен байытады. 1990 жылдары архив саласында негізінен әйел азаматтар қызмет ететін, ал ер азаматтар аз мөлшерде (шамамен 300 адамның ішінде 20–25 адам) тек техникалық, күзет және ауыр құжаттарды тасу жұмыстарымен айналысатын. Бүгінде бұл үрдіс пен білім беру жүйесі түбегейлі өзгерді.

Біз қазір үшінші үлкен революцияға – *«Цифрлық архивтік революцияға»* қадам бастық. Әлемдік тәжірибеге сүйенсек, біз тек мемлекеттік органдар мен тарихи тұлғалардың құжаттарын ғана емес, жеке сектор мен трансұлттық компаниялардың да деректерін жүйелеуіміз қажет.

Себебі болашақта олардың қайсысы тарихи маңызға ие болатынын дөп басып айту қиын. Дамыған елдерде жекеменшік архивтер институты жақсы дамыған.

Біздің елде бұл үдеріс енді ғана басталып келеді. Азаматтар ерте ме, кеш пе, өздерінің туыстарын, зейнетақы немесе еңбекақы мәліметтерін, мүліктік құқықтарын іздеп архивке жүгінеді. Бұрын білім беру, денсаулық сақтау сияқты архив те тек мемлекет монополиясында болса, қазір жекеменшік мектептер мен ауруханалар сияқты архив ісінің де жіктелетін уақыты келді.

Мемлекет стратегиялық маңызы бар құжаттарға өз монополиясын сақтауы тиіс, бірақ өзге салаларды да назардан тыс қалдырмау керек.

Мысалы, АҚШ-тың Конгресс кітапханасы болашақ тарихқа ықпалын ескере отырып, «Твиттер» (Х) желісіндегі азаматтардың жазбаларын жинақтап жатыр. Еуропа елдерінде Youtube платформасындағы ұлттық маңызы бар бейнематериалдар авторлық құқықты сақтай отырып, мемлекеттік архивке жинақталуда. Бізде баспа өнімдерінің (кітап, газет) міндетті данасы Орталық кітапханаға тапсырылатын қалыптасқан формат бар. Ендігі кезекте архив ісінде де осындай тетіктерді енгізу қажет. Бүгін бар сайттар ертең хостинг төлемінің жоқтығынан немесе авторының дүниеден өтуіне, шетелге көшуіне байланысты жойылып кетуі мүмкін. Instagram, Facebook сияқты әлеуметтік желілердегі маңызды мәтіндік және аналитикалық деректерді жинақтау мәселесі де күн тәртібінде тұр. Көптеген телеарналар мен ақпараттық сайттар ашылып, жабылып жатыр, ал бұл – біздің ұлттық тарихымыз бен тағдырымыз.

Мен 1994 жылдан бастап басқарушы органдарда қызмет еттім. Сол кезде ұйымдастыру бөліміндегі 11 қызметкердің 6-уы «Ятрань» баспа машинасында қағаз басатын хатшы-машинисткалар еді. Алғашқы компьютерлер енгізілгенде, бұл лауазымдар қысқарып, олардың орнына жүйелік әкімшілер (сисадминдер) мен компьютерлік сауаттылығы жоғары мамандар қажет болды. Сондықтан біз

цифрландырудан қорықпауымыз керек, алайда бұл үдеріс біздің кәсіби біліктілігімізге қойылатын талаптарды күрделендіре түседі.

Қазіргі таңда Астана мен Алматыда және кейбір аймақтарда заманауи, автоматтандырылған архив қоймалары бой көтеріп келеді. Бұл сала химиктерден, биологтардан бастап, материалтанушыларға дейінгі жаңа жұмыс орындарын тудыруда. Мысалы, Сталиндік репрессия құжаттарын қалпына келтірумен айналысып жатқан мамандар мұны жақсы түсінеді. Сонымен қатар, цифрлық дәуірде құжаттың түпнұсқалығын анықтау, оған өзгерістердің енгізілмегенін тексеру, яғни верификация мәселесін қамтамасыз ету үшін білікті заңгерлер қажет. Контенттің шынайылығын растайтын мамандарға архив есігі ашық болуы тиіс. Болашақта архив – жастар, соның ішінде ер азаматтар көптеп келетін заманауи мекемеге айналуы қажет.

Цифрлық нигилизм және құқықтық реттеу

Бұл ретте біз «цифрлық нигилизмге» бой алдырмауымыз керек. 1990 жылдары кейбір мемлекеттік қызметкерлер арасында «бұл қағаздарды сақтап қажеті не, өртеп жіберу керек» деген теріс көзқарас болған еді. Бүгінде бұл үрдіс «электронды құжат айналымы бар, қағазды үнемдеп, экологияны аялайық, сондықтан мөрсіз, тек цифрлық архивке көшейік» деген сипатта қайталануы мүмкін. Мұнда біздің алдымызда үлкен таңдау тұр: мемлекеттік маңызы бар құжаттарды қағаз жүзінде де, цифрлық форматта да қатар сақтауымыз қажет. Себебі, техникалық іркілістер немесе электр энергиясының өшуі жағдайында цифрлық деректерге қолжетімділік шектеледі. Сұлтанмахмұт Торайғыровтың *«Күн сөнгенше сөнбейміз»* деген сөзі бар, ал электронды құжаттар — тоқ сөнгенше сөнбейтін дүниелер. Біздің XVII ғасырдан бергі құжаттарымыз Алматыдағы және өзге өңірлердегі түрлі қиындықтарға қарамастан, архив саласында жаны ашып қызмет еткен апайларымыз бен қарындастарымыздың арқасында бүгінге аман жетті. Ендігі кезекте жасанды бұрмалауларға («дипфейк», фейк-технологиялар) төтеп бере алатын жаңа кәсіби біліктілікті қалыптастыруымыз қажет.

Қазақстан өз Конституциясына адамның цифрлық құқықтарын енгізген әлемдегі алғашқы мемлекеттердің бірі. Цифрлық даму — біздің мемлекеттік стратегиялық бағдарымыз. Жақында Мемлекет басшысының «Цифрлық Қазақстан» туралы Жарлығы шықты, сонымен қатар Орталық Азияда алғашқы болып Цифрлық кодексті қабылдадық және жасанды интеллект туралы заңнамалық негіздерді әзірлеп жатырмыз. Архив ісіне де көптеген өзгерістер енгізілді, алайда осы цифрлық заманда нақты қандай заңнамалық түзетулер қажет екенін сіздерден күтеміз.

Жаңа ғана Мақпал ханыммен пікір алмасу барысында қазіргі үлкен тілдік модельдерді (LLM) және жасанды интеллектті архив ісіне енгізу мәселесін талқыладық. Әлеуметтік желілер мен сайттардағы ақпараттық базаларды адам құқықтары мен жеке деректердің құпиялылығын сақтай отырып, ғылыми айналымға енгізу кезек күттірмейтін міндет. Бұл — тіл мәселесіне де тікелей қатысты. Қазіргі жастардың әдеби тілден гөрі аралас, трансформацияға түскен тілде сөйлейтінін айтып налимыз, бірақ бұл - бүгінгі күннің шынайылығы. Болашақта бұл нормаға айналуы да мүмкін, бұл да белгілі бір дәрежеде тәуекел

(риск) болып табылады. Осы тұрғыда цифрлық технологиялардың бізге ауадай қажет екені даусыз.

Құрметті әріптестер, біз өзгерістерге бейім болуымыз керек. Тек қатырма қағаздан жасалған қораптардың ішіндегі құжаттармен ғана шектеліп қалмай, ақпараттық кеңістікте, «интертейнмент» саласында белсенді болуымыз қажет. Қоғамды мазалаған кез келген сұраққа, мысалы, пандемия кезіндегі тарихи тәжірибелер сияқты өзекті сауалдарға архив құжаттары арқылы жедел жауап бере алуымыз керек.

Қазіргі таңда еліміздегі кітапханалар жүйесі заман талабына сай трансформацияланып, коворкинг орталықтары мен кофеханалар ашып, жас ұрпақты тартудың жаңа форматтарына көшуде. Архив — мемлекеттік маңызды институт, сондықтан бізге деген сұраныс ешқашан жоғалмайды.

Бірақ біз қоғамға қалай жақындай аламыз, халықтың сұранысын қалай өтейміз деген сұрақ төңірегінде ойлануымыз қажет. Егер барлық деректер цифрландырылып, ғылыми-анықтамалық аппаратпен негізделсе, ақпаратты іздеу мен табу үдерісі едәуір жеңілдейді. Бұл үлкен уақыт пен қомақты қаражатты талап етеді. Біз осындай экожүйені қалыптастыруға, менталитетімізді өзгертуге дайынбыз ба?

Мен сіздерден осы сұрақтарға кәсіби жауап күтемін. Болашақ цифрлық дәуірдегі архив заңнамасының архитектурасын, құрылымы мен даму бағытын бірлесе анықтауға шақырамын.

Назарларыңызға рахмет!

ИСТОРИЯ В ОДИН КЛИК: ПОЧЕМУ АРХИВЫ БУДУЩЕГО СОЗДАЮТСЯ ВАМИ

Анваржон Алиев
PhD по историческим наукам, руководитель информационной службы Агентства «Узархив», Узбекистан
a.aliev@archive.uz

ТАРИХТЫ БІР ҒАНА ШЕРТУ: БОЛАШАҚ АРХИВТЕРІН НЕГЕ СІЗДЕР ҚАЛЫПТАСТЫРАСЫЗДАР

Анваржон Әлиев
Тарих ғылымдары бойынша PhD,
«Өзархив» агенттігінің ақпараттық қызметінің жетекшісі, Өзбекстан
a.aliev@archive.uz

HISTORY IN ONE CLICK: WHY THE ARCHIVES OF THE FUTURE ARE CREATED BY YOU

Anvarjon Aliev
PhD in History, Head of the Information Service of the Uzarchive Agency, Uzbekistan
a.aliev@archive.uz

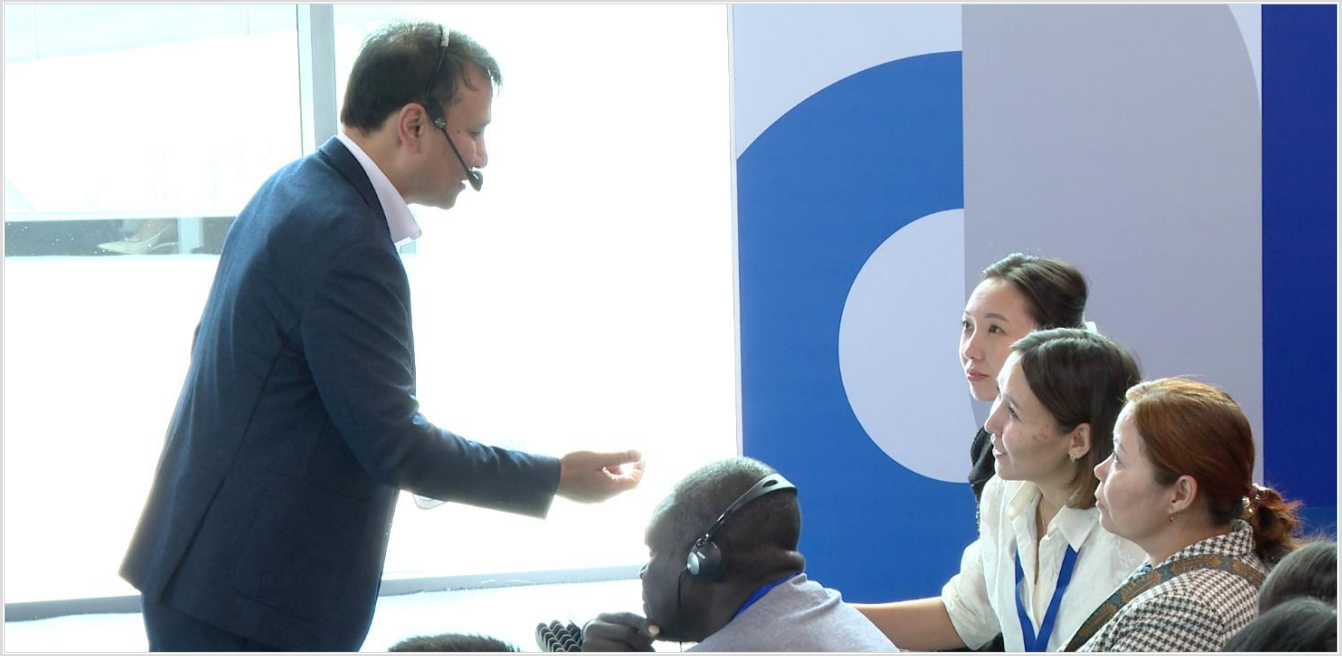
Что вы представляете себе, когда вы слышите слово «архив»? Большинство видит бесконечные ряды серых стеллажей, полумрак, строгого человека в нарукавниках, который призывает к тишине любого, кто посмеет громко

перевернуть страницу. Архив кажется местом, где время замерло, а прошлое заперто под замок. Я здесь сегодня для того, чтобы разрушить этот миф.

Мы живем в XXI веке, и прямо сейчас происходит тихая революция. Архивы сбрасывают с себя вековую пыль. Из закрытых хранилищ они превращаются в открытые цифровые платформы, в экосистемы, которые создаются вами.

Долгое время архив был монополистом на историю. Архивисты решали, что сохранить, а историки – как это трактовать. Обычный человек был просто зрителем. Но история – это не только указы высших должностных лиц и сухие отчеты министерств. История – это ткань, сотканная из миллионов человеческих судеб. Из писем вашей бабушки, из старых фотографий соседа, из воспоминаний репрессированных, чьи голоса пытались стереть. Если мы оставим архивы закрытыми, мы потеряем свою идентичность.

В Агентстве «Узархив» мы задали себе дерзкий вопрос: что если сделать общество не просто потребителем, а *соавтором* истории? Так родилась концепция живого соучастного архива.



Как мы это делаем? Через медиа и технологии, которые понятны каждому. Мы запустили *три ключевых проекта*, и каждый из них меняет правила игры.

Мы вышли на телевидение с проектом *«Архивариус»* на канале «Дунё буйлаб». Мы показали, что за каждым пожелтевшим документом скрывается детективная история, драма или триумф. Знаете, какой был результат? Люди увидели, что архив – это захватывающе. Как бы нескромно ни прозвучало, в 2024 году я держал в руках главную журналистскую премию Узбекистана «Олтин калам». Именно за этот проект. Мы доказали: архивный контент может быть рейтинговым.

Но мы пошли дальше. Письменная история часто официальна. А как же эмоции?

Мы запустили краудсорсинговый проект *«Оғзаки тарих»* («Устная история»). Что это значит? Мы начали записывать интервью у лиц с интересным прошлым, очевидцев нескольких эпох. Записывать воспоминания старожилов, участников исторических событий, ученых. Мы объединили строгие государственные акты с живыми порой дрожащими голосами людей. Это микроистория. Это антропология. Теперь каждый житель Узбекистана может вписать голос своего деда в официальную летопись страны.

Для молодежи, которая привыкла к быстрому контенту, мы создали проект *«Миллий хотира»* («Национальная память») – короткие, динамичные видеоролики на основе редких кино- и фотодокументов. Мы принесли историю туда, где проводит время новое поколение – в их смартфоны.

И, наконец, самый частый барьер: «В архиве сложно что-то найти, там куча законов и регламентов».

Встречайте – наш новый цифровой коллега *«Анваржон архивариус»*. Это нейросеть, искусственный интеллект, обученный на законодательстве Республики Узбекистан и международных стандартах ISO. Вам больше не нужно часами изучать юридические акты по делопроизводству или правилам хранения. Вы просто задаете вопрос Анваржону на человеческом языке, и специалист дает экспертный ответ. Технологии сделали архивы прозрачными и доступными 24/7.

Кто же такой современный архивист?

Это больше не смотритель бумажных древностей. Сегодня это одновременно глубокий историк, подкованный юрист и IT-специалист, управляющий цифровой средой.

Но самое главное – меняется роль самого гражданина. Когда вы делитесь устным воспоминанием, когда вы ищете предков через цифровые базы данных, вы восстанавливаете историческую справедливость. Вы укрепляете фундамент национальной идентичности. Прошлое – это не застывший песок. Это живая глина, из которой мы прямо сейчас лепим наше будущее.

Опыт Узбекистана показывает: когда архивы открывают двери и используют силу медиа и искусственного интеллекта, общество отвечает взаимностью.

Я призываю каждого из вас: загляните сегодня вечером в свои домашние альбомы. Поговорите со своими родителями, пока они рядом. Запишите их голоса.

Зайдите на сайт архива. Потому что история – это не то, что написано в школьных учебниках. История – это то, что мы помним и сохраняем вместе.

FROM INFORMATION OVERLOAD TO TRUSTED KNOWLEDGE: THE POWER OF AUDIOVISUAL ARCHIVES

Elise Gracia
Project Manager, National Audiovisual Institute (INA), Coordinator of the Regional Project for the Preservation of Audiovisual and Cinematographic Heritage in Central Asia, France
egracia@ina.fr

АҚПАРАТТЫҚ АРТЫҚШЫЛЫҚТАН СЕНІМДІ БІЛІМГЕ: АУДИОВИЗУАЛДЫ АРХИВТЕРДІҢ ҚУАТЫ

Элиз Грасиа
Ұлттық аудиовизуалды институттың (INA) жоба жетекшісі, Орталық Азиядағы аудиовизуалдық және кинематографиялық мұраны сақтау жөніндегі өңірлік жобаның үйлестірушісі, Франция
egracia@ina.fr

ОТ ИНФОРМАЦИОННОЙ ПЕРЕГРУЗКИ К ДОСТОВЕРНОМУ ЗНАНИЮ: СИЛА АУДИОВИЗУАЛЬНЫХ АРХИВОВ

Элиз Грасиа
Руководитель проекта Национального аудиовизуального института (INA), координатор регионального проекта по сохранению аудиовизуального и кинематографического наследия в Центральной Азии, Франция
egracia@ina.fr

Introduction

We are living in an age of over-information.

Every second, we produce more images, more videos, more data than ever before. And now, with artificial intelligence, this flow is accelerating.

At first, this might seem like a blessing. More data, more access, more knowledge.

But in reality, something else is happening.

Because here is the paradox:

We have never had so much information… and yet we have never been so uncertain about what is true. So today, the challenge is no longer access.

The real challenge is trust. And this leads us to a fundamental question: How can audiovisual archives become trusted resources to understand, learn, and transmit history in the age of AI?

Framing the problem

In a world shaped by artificial intelligence and the multiplication of information sources, we are no longer lacking content.

We are lacking:

• context,

• meaning,

• and reliable reference points.

So the question is not simply how to preserve archives. But, it is how to activate them.

INA: from preservation to activation

At the French National Audiovisual Institute – the INA – this question is at the heart of what we do.

Today, INA:

• preserves millions of hours of radio, television and web archives,

• captures media flows 24 hours a day,

• and trains thousands of professionals every year.

But the key point is this: What matters is not the quantity of archives.

It is what we do with them.

Because at INA, archives are not just preserved.

They are:

• collected,

• analyzed,

• contextualized,

• and shared.

In other words: they are activated. And this activation happens through different uses – different ways of learning from the same resource. Let me show you how.

Producing knowledge – The INA Lab

Everything starts with knowledge.

And more importantly: with reliable knowledge.

In recent years, research has deeply evolved. We no longer study a single document.

We study patterns, trends, large-scale data. So, we had to rethink what archives could be for research.

That's why INA created the INA Lab. The Lab is not just a place. It is an environment where:

• researchers access large audiovisual datasets,

• they use artificial intelligence and data analysis tools,

• and they explore media over long periods of time.

So instead of asking: "What does this archive show?"

Researchers can now ask: "What do thousands of hours of archives reveal about our society?"

This is a paradigm shift. From individual documents... to massive corpora of data.

Archives become scientific infrastructures.

They allow us to understand not just an event – but entire media ecosystems.

Analyzing the world – data.ina.fr

But research should not be limited to experts. We all need tools to understand the world. Because misinformation does not come from a lack of data.

It comes from a lack of interpretation.

This is the idea behind data.ina.fr. This platform transforms archives into data that anyone can explore.

You can:

• visualize media trends,

• track how a topic evolves,

• analyze how narratives are constructed over time.

For example, you can see how certain keywords rise and fall in the media.

How visibility shifts. How attention is distributed.

Suddenly, the archive is no longer something you watch.

It is something you explore. The user becomes an analyst. The archive becomes a tool for critical thinking. Archives help us decode the present through the past.

Transmitting knowledge – Lumni

But understanding is not enough. We also need to transmit. Because knowledge only exists if it is shared and taught.

This is where Lumni comes in. Lumni is an educational platform designed for students and teachers, where audiovisual archives are transformed into learning resources. Here, archives are:

• selected,
• contextualized,
• adapted to educational needs.

A historical video becomes a lesson. A media archive becomes a teaching tool.

The archive is no longer just a trace of the past. It becomes a way to make history understandable. Archives become living educational material.

Building skills – Training with and through archives

And learning does not stop at school.

At INA, this is where our role goes beyond preservation. Because historically, we are also a training center, with decades of experience teaching how to collect, describe and preserve audiovisual archives. And here, something essential happens.

Because none of what I just described – not the Lab, not data.ina.fr, not Lumni – can exist without well-structured and enriched data. And this requires skills.

At INA, we have long been a training center, teaching professionals how to collect, describe, and preserve archives.

But today, the paradigm has shifted.

We are no longer working only with documents. We are working with data.

And this data is not just something to organize. It is something to question, analyze, and interpret.

So professionals themselves must evolve.

They are no longer only archivists. They are becoming data interpreters. They must understand: how data is produced, how it is enriched, how it shapes what we see.

Because in the age of AI, the ability to question data is a critical skill. And this is exactly what archives now help us learn.

Sharing knowledge – ina.fr

And finally, archives must be shared.

Because an archive that is not accessible cannot create knowledge.

Through platforms like ina.fr, archives are made available to everyone:

• citizens,
• journalists,
• creators,
• researchers.

They circulate. They feed public debate. They become part of our collective culture. Conclusion So across all these examples, something important happens. The archive itself does not change. What changes... is how we use it.

• In the Lab → we produce knowledge

• On data.ina.fr → we analyze
• On Lumni → we transmit
• In training → we develop skills

One resource. Multiple activations. And this leads to a simple but powerful idea:

Archives are not passive storage

They are:

• resources to understand,
• tools to learn,
• and foundations for collective narratives. In the age of artificial intelligence, the challenge is no longer to preserve memory.

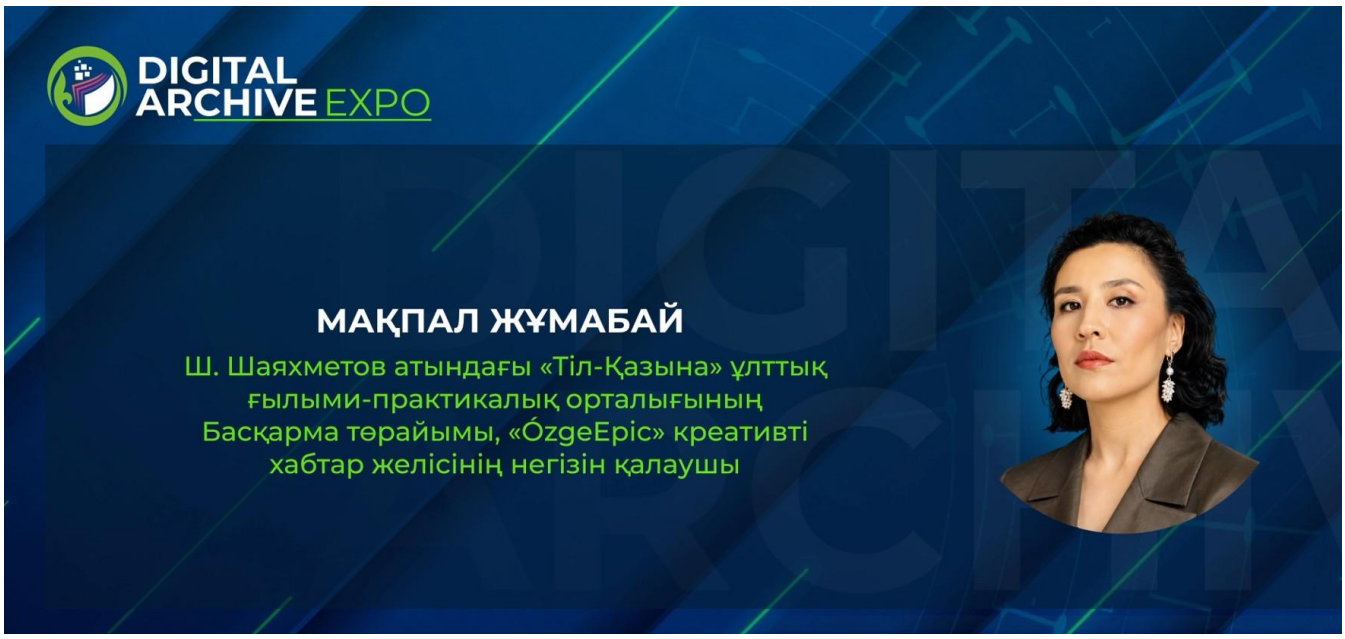
It is to ensure:

• trust,
• understanding,
• and transmission.

Because tomorrow, the risk is not that we will lack information. The risk is that we will lack meaning.

So we do not preserve archives to remember the past.

We activate them to make sense of the present and to build a shared future.



АРХИВ – КРЕАТИВТІҢ ҚАЙНАР КӨЗІ

Мақпал Жұмабай
«Ш. Шаяхметов ат. «Тіл-Қазына» ұлттық ғылыми-практикалық орталығы» КЕАҚ Басқарма төрайымы, ÓzgeEpic креатив хабының негізін қалаушы, Қазақстан
Maqpalmaqam@gmail.com

АРХИВ – ИСТОЧНИК КРЕАТИВА

Макпал Жумабай
Председатель Правления НАО «Национальный научно-практический центр «Тіл-Қазына» им. Ш. Шаяхметова»,
основатель креативного хаба ÓzgeEpic, *Казахстан*
Maqpalmaqam@gmail.com

ARCHIVES AS A SOURCE OF CREATIVITY

Maqpal Zhumabai
Chairperson of the Board, Sh. Shayakhmetov National Scientific and Practical Center «Til-Qazyna» (Non-profit Joint-Stock Company),
Founder of the ÓzgeEpic Creative Hub, Kazakhstan
Maqpalmaqam@gmail.com

Архив адамзат өркениетінің жадын сақтайтын бірегей кеңістік. Онда әр дәуірдің ойлау мәдениеті, шығармашылық ізденісі, ғылыми жаңалығы мен рухани тәжірибесі жинақталған. Адамзаттың ең ұлы әдеби шығармалары, музыкалық мұралары, сәулеттік жобалары, тарихи хаттары мен қолжазбалары дәл осы архивтерде сақтаулы. Сондықтан архивті өткеннің естелігі ретінде емес, жаңа идеялар мен жаңа шығармашылықтың қайнар көзі ретінде қарастыру қажет.

Себебі алдағы уақытта мәдени, ғылыми және креативті жобалардың негізі – адамзаттың жинақталған тәжірибесіне қайта үңілуден туындамақ.

Соңғы онжылдықтарда әлемнің ірі технологиялық компаниялары мен жетекші мәдени институттары дәл осы ақиқатты терең түсіне бастады. Олар архивті жаңа технологиялар мен жаңа интеллектуалдық өнімдерді қалыптастыратын стратегиялық ресурс ретінде қарастырып келеді. Бүгінде тарихи құжаттар, қолжазбалар, аудио және бейнемұралар цифрландырылып, жасанды интеллект жүйелерін оқытуға, фильмдер мен ойындар жасауға, ғылыми зерттеулер жүргізуге және креативті индустрияның жаңа бағыттарын дамытуға пайдаланылып жатыр. Бұл үдеріс архивті өткеннің көлеңкесінде қалған кеңістік тұжырымынан ажыратып, болашақты қалыптастыратын маңызды күшке айналып келе жатқанын көрсетеді.

2004 жылы Google компаниясы жер бетіндегі барлық кітапты сандық форматқа ауыстыруға бағытталған «Google Books» жобасын іске қосты. Бүгін сол жобада 40 миллионнан астам кітап жинақталған. Ортағасырлық монастырь жертөлесінде жатқан пергаменттен бастап, XX ғасырдың сирек кітаптарына дейін, миллиондаған адамның бір шертумен ие бола алатын рухани қазынасына айналды. Электронды білімнің алдындағы есіктер ашылды.

Одан кейін Archive.org жүйесі пайда болды. Интернет архив деп аталатын бұл жобада кітаптардан бөлек, музыка, кино, тарихи бейнежазбалар, радио хабарлар, тіпті 1996 жылдардағы алғашқы веб-сайттар да осы жерде сақталуда. Бұл платформаның «Wayback Machine» деп аталатын мүмкіндігі арқылы сіз интернеттің өз тарихына бойлай аласыз. Яғни, адамзат тек кітапты ғана емес, өзінің цифрлық тынысын да архивтей бастады.

Үшінші үлгі ретінде Еуропа одағының Europeana платформасын атайын. Елу миллионнан астам мәдени мұра объектісін бойына сыйғызған бұл іргелі жобада картиналар, ноталар, карталар, фотосуреттер, қолжазбалар жинақталған. Еуропаның ұлттық мұражайлары, кітапханалары, архивтері бір ортақ платформада бас қосты. Ескі дүние жаңа дүниемен тіл табысты.

Аталған үш жоба бізге архивтің қоғамның интеллектуалдық қозғаушы күші екендігін көрсетіп берді. Кезінде тек санаулы зерттеушілердің ғана иелігінде болған бұл қазына бүгінде жасы да, дәрежесі де бөлек, дүниенің кез келген бұрышындағы ізденуші адамның ортақ мүлкіне айналды.

Қазақстан да бұл жаһандық үдеріске өз үнін қосуда.

Мемлекет басшысы Қасым-Жомарт Тоқаев «Зерделі ұлт» тұжырымдамасын ұсына отырып, қоғам дамуының жаңа бағытын айқындады. Зерделі ұлт деген өз тарихын терең танитын, архивтегі деректі тірі жады ретінде қабылдайтын, технологияны рухани мұрамен ұштастыра алатын қоғам. Сонымен қоса, Президент Ұлттық құрылтайдың IV отырысында Ұлттық цифрлық архив құру мәселесін тікелей күн тәртібіне қойды.

Мен бүгін осы бағытта жүзеге асырылған нақты тәжірибелер туралы айтып өткім келеді. Кейбіріне тікелей қатысып, кейбіріне сырттай тілекші болдым.

Сол жобалардың алғашқысы – қазақ телевизиясы мен радиосының «Алтын қорын» цифрлау жобасы. Бұл жобаның жұмыс тобы құрамында болдым. Сондықтан оның қалай дүзеге асқанын жақсы білемін.

«Қазақстан» телерадиокорпорациясының «Алтын қорында» 1930 жылдардан бері жинақталған бейне және аудиоматериалдар сақталған. Ұзақ уақыт бойы ескі таспалар архив сөрелерінде жатты. Алайда сол таспалардың ішінде тұтас дәуірдің үні бар еді. Тұлғалардың сөздері, халықтың жыры, ұрпаққа деген өсиет бәрі бәрі сол таспаларда сақталды.

Ұлттық телевизия мен радио архиві бірнеше жыл бойы цифрланып, толықтай жаңғыртылды. Цифрлы мұраның едәуір бөлігі altynqor.com платформасы арқылы баршаға қолжетімді болды. Бұл оқиға Қазақстан тарихындағы архив ашықтығын паш еткен елеулі оқиға еді. Нәтижесінде 800-ден астам бөлімге топтастырылған мыңдаған видео көпшілікке қолжетімді болды. «Қарыз бен парыз», «Тамаша», «Халық қазынасы», «Асыл мұра», «Адам іздеп жүрмін» бағдарламалары қайта көпшілік назарына оралды. «Тұлғалар» бөлімінде Мұхтар Әуезов, Ғабит Мүсірепов, Шәкен Айманов сынды ұлы тұлғалардың жанды дауыстары мен нақты бейнелері жарыққа шықты.

«Аудиоқор» бөлімінде 1954 жылдан бері сақталған жазбалар жаңа тыныспен қайта жаңғырды. Жазбалар қатарында Батырлар жырының 50 нұсқасы, 100 күй, 100 ертегі нұсқасы, «Жар-жар», «Беташар», «Жоқтау» секілді дәстүріміздің көне үлгілері болды. Тарихи жад тұңғиықтағы қойнаудан қолданушылар кеңістігіне шықты.

Бүгінде цифрлы платформада 110 мыңнан астам материал бар. Жыл сайын тағы да 1000 жаңа материал қосылуда. Бұл зерттеуші ғалымға да, мектеп мұғаліміне де, мектеп оқушысына да, жай ғана ата-баба мұрасын іздеген адамға да бірдей ашық кеңістік. Алтын қордың цифрланып, көпшілікке ұсынылуы — архив контентін арқау еткен жаңа жобалардың өмірге келуіне түрткі болды. Әлеуметтік медиа архивтік ленталарға толды. Мәселен, Мемлекеттік Кеңесші Ерлан Тынымбайұлының идеясымен жүзеге асқан *«Асыл мұра» аудиожинағы*. Жобаға қазақ фольклорының таңдамалы 1500 аудиожазбасы енді. Атап айтсақ,

400 күй, 300 жыр-терме, 400 халық әні, 100 ертегі, 50 бесік жыры дыбыстық өңдеумен ұсынылды.

«Ozen» музыкалық лейблімен бірге жүзеге асырылған бұл жоба Apple Music, Spotify, Yandex Music, YouTube Music сынды 100-ден астам халықаралық платформаға жүктелді. «Асыл мұра» аудиожинағы Қазақ музыкасының стриминг платформаларда кеңінен таралуына, жас буын арасында көптеп тыңдалуына себепкер болған бірден бір бастама болды.

Маңғыстаудың күйі, Жетісудың жыры бүгін Токиода, Лондонда, Нью-Йоркте еркін тыңдалатын болды. Қазақтың дауысы шекарадан шығып, тек тарихи артефакт ретінде ғана емес, бүгінгі күннің тірі мәдениеті ретінде де тарай бастады.

Үшінші жоба мен тікелей қатысқан жұмыстардың бірі – Қазақстанның ашық кітапханасы

Kitap.kz қазақ тіліндегі толыққанды интеллектуалдық кеңістік. Ол қазақ және әлем әдебиетінің үздік шығармаларын оқу және тыңдау мүмкіндігін беретін ауқымды платформа. Бүгінде қордан 5000-нан астам кітапты оқып және аудио форматта тыңдауға болады. Жоба еліміздегі ұлттық әдеби мұраны тұңғыш рет интернет кеңістігіне шығарып, кітаптарды түпнұсқа кітаптан немесе қолжазба құжаттардан әзірледі. Kitap.kz бастау алған 2010 жылдардың бастапқы кезеңінде, отандық және шетелдік авторлардың еңбектерімен таныстыратын алаң елімізде болмады. Нәтижесінде, Абай Құнанбайұлы, Мұхтар Әуезов, Ілияс Жансүгіров, Бауыржан Момышұлы, Әбіш Кекілбаев т.б. алыптардың еңбектері, мектеп бағдарламасына арналған оқулықтар, балаларға арналған ертегілер мен өлеңдер бұл платформада жинақталды. Одан бөлек, Стивен Ковидің «Жасампаз жандардың 7 дағдысы», Юваль Ноа Хараридің «Sapiens», Клаус Швабтың «Төртінші индустриялық революция» еңбектерінің барлығы қазақ тіліне аударылып, аудионұсқасымен ұсынылды.

Көріп отырғандықтарыңыздай, Академиялық бэкграундым мен Кәсіби жолым цифрландыру мен креатив саласымен тікелей байланысты болды. Қазақтың мәдени мұрасын жаңғырту жолындағы тәжірибем мені бір уақыттарда креатив кеңістіктің негізін қалауға итермеледі. Осылайша, дүниеге OzgeEpic креатив хабы келді. 2019 жылы Алматыда бастау алған бұл жобамыз қазіргі уақытта Қазақстанның 4 аймағына масштабталып, 3000-нан астам жобаны жүзеге асырды.

Креатив кеңістіктің іргелі жобаларының бірі – өңірлік жыршылық-жыраулық дәстүрді насихаттайтын Жауһар жобасы болды. Бұл жоба әр аймақтың мақам-жырларын, күйлерін «тірі» орындаушылардан студияларда таспалап алады. Қазіргі кезеңде Маңғыстау мұрасы аудио және видеоформатта түсіріліп, Youtube сервисінде қолжетімді болып отыр. Алдағы уақытта стриминг платформаларға тегін тарату жоспарланған. Жақын уақытта еліміздің батыс өңірі Ақжайық мектебін таспалаймыз. Ендігі кезекте таспаланған жазбаларды «cultural entertainment» форматымен көпшілікке ұсыну көзделуде. Жобаға ЖИ технологияларын енгізіп, көрермен сұрауына сай аймақтардың дәстүрлі музыкалық өнерін орындап бере алатын және далалық өркениетке дәл уақытта бойлауға мүмкіндік беретін жаңа модельді оқытудамыз. Жоба атауын далалық мәдени кодты насихаттау мақсатында – Dala AI деп атап отырмыз.

Жобаның мәні дәстүрлі мұраны цифрлық инсталляциялар арқылы тірілтіп, музейдің жаңа форматын ұсыну. Мәдениетті адаммен тікелей беттестіру – ендігі кезеңдегі креатив пен архивтің ұштасуының нәтижесі болмақ.

Сөзімді қорытындылай келе, баяндаманың басындағы ойға оралайын.

Бір кездері сарғайған қағаздардың арасында адамдардың қолтаңбасы қалды. Бүгін сол дүниелерді смартфоннан тыңдай алатын заманға жеттік. Бірақ мұның бәрі өздігінен болған жоқ. Әр мұраның артында оны сақтап қалған үлкен ұжымдар бар.

Бүгінгі таңда архив пен креатив бір-бірінен алшақ ұғымдар қатарында қаралмайды. Керісінше, дәл осы екеуінің тоғысында жаңа фильмдер, музыка, цифрлық жобалар, жасанды интеллект модельдері және мәдени өнімдер дүниеге келуде. Өткеннің мұрасы жаңа буынның тілімен сөйлей бастаған сәтте ғана мәдениет өміршеңдігін сақтай алады.

Мұра өздігінен ояна алмайды. Оны ояту керек. Зерделі ұлт үлкен технологиядан ғана емес, өз дауысын жоғалтпай сақтай алған қоғамнан туады.

АРХИВ КАК ЖИВАЯ ПАМЯТЬ: КАК ПАРТНЕРСТВО IT И АРХИВОВ МЕНЯЕТ ОПЫТ ВЗАИМОДЕЙСТВИЯ С ИСТОРИЕЙ

Владислав Шин
Руководитель коммерческого отдела ТОО «Сайлет», Казахстан
v.shin@sailet.pro

АРХИВ – ТІРІ ЖАД: IT САЛАСЫ МЕН АРХИВТЕРДІҢ ӘРІПТЕСТІГІ ТАРИХПЕН ӨЗАРА ӘРЕКЕТТЕСУ ТӘЖІРИБЕСІН ҚАЛАЙ ӨЗГЕРТУДЕ

Владислав Шин
«Сайлет» ЖШС коммерциялық бөлімінің басшысы, Қазақстан
v.shin@sailet.pro

THE ARCHIVE AS A LIVING MEMORY: HOW THE PARTNERSHIP BETWEEN IT AND ARCHIVES IS TRANSFORMING THE WAY WE ENGAGE WITH HISTORY

Vladislav Shin
Head of Commercial Department, Sailet LLP, Kazakhstan
v.shin@sailet.pro

Несколько месяцев назад я открыл систему, которую мы делали для Архива Президента Республики Казахстан. Просто хотел посмотреть, как она работает. Протестировать. Система предложила начать с экспозиции о Герольде Карловиче Бельгере – казахстанском писателе, переводчике, человеке, которого я, честно говоря, раньше не знал.

Первым появилось фото. Черно-белое. Маленький мальчик с родителями. Стоят где-то во дворе. Подпись: «Герольд Карлович Бельгер в раннем детстве вместе с родителями. Снимок выполнен в г. Энгельсе Автономной Советской Социалистической Республики немцев Поволжья в 1938 году незадолго до трагических событий депортации советских немцев».

Я остановился. Потому что я знаю, что такое депортация. Не из книг.

Меня зовут Владислав Шин. Я кореец. Точнее – коре-сарам: потомок советских корейцев, которых Сталин депортировал с Дальнего Востока в Казахстан и Узбекистан в 1937 году.

История моей семьи – это история трех переселений.

Первое – середина XIX века: предки по отцовской линии перешли границу между Кореей и Российской империей, искали землю, осели на Дальнем Востоке.

Второе переселение – 1920 год. По материнской линии. Моему прадеду Хан Ерсену было шесть лет. Его бабушка перенесла его через реку Туманган – тогдашнюю границу между Кореей и советской территорией. Они бежали от японской оккупации.

Третье переселение – 1937 год. Уже не добровольное. Сталин депортировал около ста семидесяти тысяч советских корейцев с Дальнего Востока в Казахстан и Узбекистан. За несколько дней, без предупреждения. Люди брали что успели. Наши предки попали в Узбекистан. Там родились мои бабушки и дедушки. Там родились мои родители. Там родился я.

Но в нашем семейном архиве есть дыры. Некоторые имена, названия деревень – только в памяти бабушки. Когда эта память уйдет – они уйдут вместе с ней.

Вот в чем ирония: я занимаюсь тем, что делает архивы доступными для людей, а собственный семейный архив знаю фрагментами.

Я руковожу коммерческим отделом IT-компании Sailet. Мы разрабатываем сложные цифровые системы. Обычно выступаю перед бизнесом – бюджеты, сроки, технические решения.

Но сегодня я хочу поговорить не про IT. А про то, что происходит с историей, когда ее оцифровывают, и что происходит с ней, когда ее делают живой. Потому что это не одно и то же.

Значительную часть архивных документов сегодня оцифровали. Документы сканируют, загружают в облако, индексируют. И это огромный труд, результат десятилетий работы. Благодаря вам у человечества есть возможность сохранить историю. Вы – те, кто это прошлое хранит.

Но давайте зададим честный вопрос: для кого? Для специалистов это бесценно. Историк, исследователь, архивист – они знают, как искать, умеют читать документ 1923 года с нечитаемым шрифтом и понимают его ценность без подсказок. Но если мы говорим об обычных людях – о молодежи, о детях – насколько удобно им взаимодействовать с делами, файлами, рукописями? С сухой выжимкой без контекста?

Для вас это золото. Для обывателя – это темнота. Это как дать ключ от библиотеки, но в этой библиотеке темно, и книги расставлены в произвольном порядке.

Информация хранится – но не рассказывает историю. Хотя способ передавать историю – он старше любого архива.

До письменности история народа жила в рассказе. Старейшина садился рядом с молодыми и говорил: кем были эти люди, что с ними случилось, почему это важно помнить. Не список имен – история с живыми людьми, с контекстом, с логикой. Слушали. Запоминали. Передавали дальше.

Сегодня у нас терабайты данных и глобальные архивы – а молодежь не знает собственных корней. Не потому, что не хочет. А потому что история хранится как документ, а не рассказывает судьбы.

В прошлом году нам позвонили из Архива Президента Республики Казахстан. Сказали: мы хотим, чтобы история не лежала в файлах, а рассказывала – отвечала на вопросы, показывала лица. Мы начали строить систему, которую назвали *ИИ-архивариус*.

Архив передал нам материалы – фотографии, документы, тексты – по конкретным людям и темам. И мы научили систему работать с этим, как с живой памятью. Открываешь экспозицию о Герольде Бельгере. Система не выдает список файлов. Она начинает рассказывать историю – по главам: детство, юность, становление, творчество. По каждому этапу фотографии, подписи, контекст. Ты идешь по жизни человека, как по маршруту. И в любой момент можешь остановиться и спросить что-то свое: «Кто стоит рядом с Бельгером на этой фотографии?», «А кем была его жена?» и т.д. Система находит ответ в своих материалах, показывает фотографию, которую ты не видел, и рассказывает историю. Это как разговор с памятью.

Эта система уже работает здесь, на Конгрессе – в выставочном зале. Возможно, кто-то из вас уже успел задать ей вопросы.

В Казахстане пока нет сервиса, где архивы и люди по-настоящему встретились. А ведь Казахстан – страна, где история многих народов переплетена

в одну. Казахи с их тысячелетней памятью степи. Корейцы, немцы, русские, уйгуры, поляки – десятки народов, каждый со своей памятью, своими именами, своими фотографиями в старых альбомах. И у каждого – свои пробелы, свои пропавшие страницы.

В казахской культуре есть традиция – Жеті ата: каждый должен знать своих предков до седьмого колена. Это не просто обычай – это способ помнить, кто ты и откуда.

У нас в Центральной Азии до сих пор идут споры, кто является прямым потомком Чингизхана. Я в этой конкуренции, как вы понимаете, не участвую. Но было бы здорово, если бы молодые казахи могли поднять историю своих жузов и родов не только со слов бабушки – а из реальных архивных данных.

Молодежь не ходит в архивы, но она сидит в смартфоне. И у нее есть вопросы: чем был известен мой прапрадед? как жили люди его времени?

Представьте: молодой человек открывает приложение, вводит фамилию деда и видит документ из архива. Фотографию. И может спросить: «А чем занимался мой дед? Как выглядела жизнь тогда?».

Технология уже существует, архивы уже существуют. Не хватает одного – партнерства.

Я возвращаюсь к той фотографии. Маленький Герольд. 1938 год. Энгельс. Его семья пережила депортацию. Он стал писателем. Написал сотни текстов о памяти, о языке, о том, что значит помнить, кто ты. Он умер в 2015-м. Мы создали для него цифровую выставку, и теперь любой человек может открыть экран, и система покажет детскую фотографию Бельгера, расскажет про его родителей, ответит на вопрос: «А кем была его жена?» – найдет ее фотографию и расскажет о ней. Этот человек останется живым в разговоре. Не в папке. В разговоре.

А мой прадед Хан Ерсен, которого бабушка перенесла через реку в 1920 году, он в разговоре остался. Того архива нет.

Когда архива нет, человек постепенно превращается в несколько обрывков памяти: имя, дата, семейный рассказ, который с каждым поколением становится все тише. Мы строим системы, чтобы таких пробелов в памяти становилось меньше.

Вы – архивисты. Вы лучше меня понимаете ценность сохраненного документа. Вы сохраняете не бумагу. Вы сохраняете саму возможность помнить.

Давайте сделаем эту возможность доступной для каждого.

«АРХИВТЕР ҚОҒАМ ҮШІН»
№ 5 сессия

Сессия № 5
«АРХИВЫ ДЛЯ ОБЩЕСТВА»

«ARCHIVES FOR SOCIETY»
Session No. 5

ОТ АРХИВНОГО ФОНДА К ВИРТУАЛЬНОМУ МУЗЕЮ: СОЗДАНИЕ ЦИФРОВОЙ ЭКОСИСТЕМЫ НАСЛЕДИЯ

Джемма Болатовна Аманжолова
Кандидат филологических наук, ассистент-профессор,
Astana IT University, Казахстан
jemma.aman@gmail.com

Аннотация. В статье на основе междисциплинарного подхода анализируются эволюция понятия «виртуальный музей», его основные функции (сохранная, репрезентационная, образовательная, научно-исследовательская). Рассматриваются международные и казахстанские практики, выявляются барьеры и перспективы развития. Обосновывается необходимость создания интегрированной платформы, объединяющей архивную аутентичность и музейную интерпретацию в едином цифровом пространстве.

Ключевые слова: виртуальный музей, цифровая экосистема наследия, архивная атрибуция, функции виртуального музея

АРХИВ ҚОРЫНАН ВИРТУАЛДЫ МУЗЕЙГЕ ДЕЙІН: МҰРАНЫҢ ЦИФРЛЫҚ ЭКОЖҮЙЕСІН ҚҰРУ

Джемма Болатқызы Аманжолова
Филология ғылымдарының кандидаты, ассистент-профессор,
Astana IT University, Қазақстан
jemma.aman@gmail.com

Аңдатпа. Мақалада пәнаралық тәсіл негізінде «виртуалды мұражай» ұғымының эволюциясы, оның негізгі функциялары (сақтау, өкілдік, білім беру, ғылыми-зерттеу) талданады. Халықаралық және қазақстандық тәжірибелер қарастырылады, дамудың кедергілері мен болашағы анықталады. Бірыңғай цифрлық кеңістікте архивтік

түпнұсқалықты және мұражай интерпретациясын біріктіретін интеграцияланған платформа құру қажеттілігі негізделеді.

Түйін сөздер: виртуалды музей, мұраның цифрлық экожүйесі, архивтік атрибуция, виртуалды музейдің функциялары

FROM ARCHIVAL FONDS TO A VIRTUAL MUSEUM: CREATING A DIGITAL HERITAGE ECOSYSTEM

Dzhemma B. Amanzholova
Candidate of Philological Sciences, Assistant Professor,
Astana IT University, Kazakhstan
jemma.aman@gmail.com

Abstract. Based on an interdisciplinary approach, the article analyses the evolution of the concept of a virtual museum, its main functions (preservation, representation, education, and research). International and Kazakhstani practices are considered, and barriers and development perspectives are identified. The need to create an integrated platform that combines archival authenticity and museum interpretation in a single digital space is substantiated.

Key words: virtual museum, digital heritage ecosystem, archival attribution, functions of a virtual museum

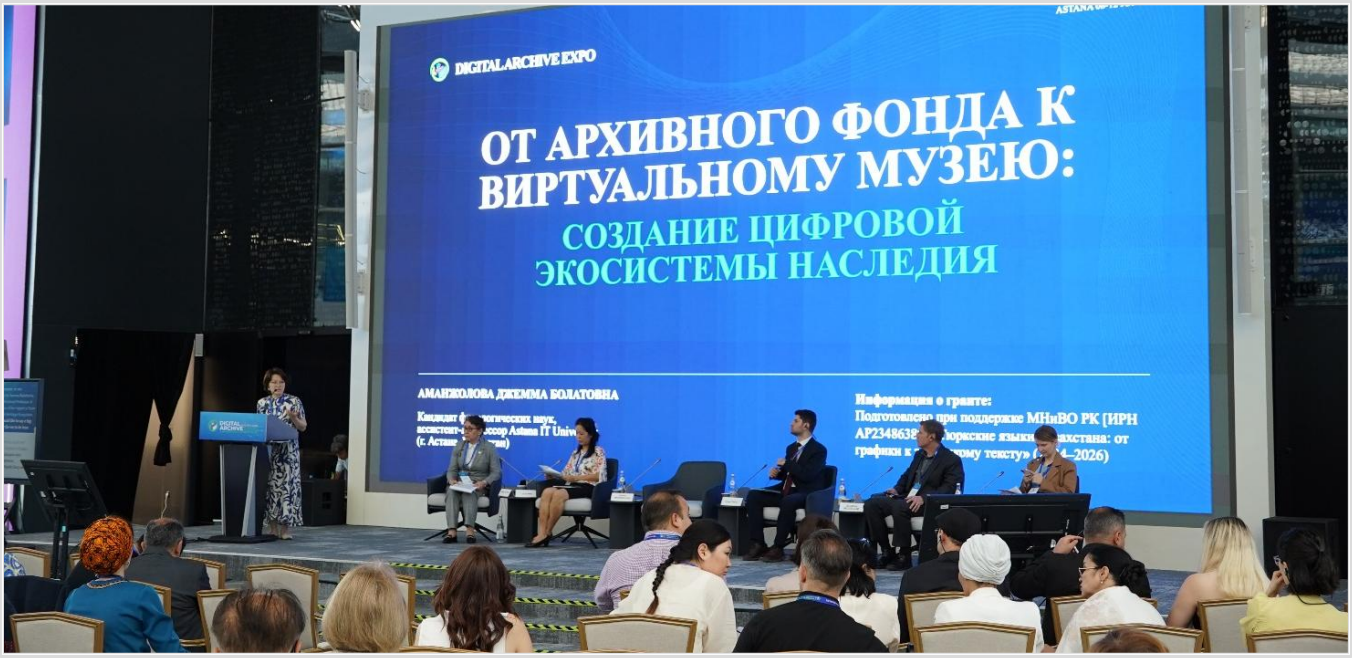
Введение

В современной гуманитарной науке и музейной практике понятие «*виртуальный музей*» утвердилось как обозначение качественно новой культурной формы, порожденной цифровыми технологиями и глобальной сетью Интернет. В отличие от традиционного музея, который понимается как институция, обладающая физическим пространством, материальными экспонатами и системой хранения, виртуальный музей существует исключительно в электронной среде. Согласно определению, предложенному Т.Е. Максимовой, виртуальный музей представляет собой «новую культурную форму интегративного характера, многофункциональный комплекс, характеризующийся прежде всего присутствием в сети Интернет» (Максимова 2014: 165). При этом он не обязательно является цифровой копией реально существующего музея – напротив, многие виртуальные музеи не имеют физического аналога и создаются «с нуля» на основе оцифрованных коллекций, архивных документов или даже полностью сгенерированных цифровых артефактов.

Важной терминологической дискуссией остается вопрос о правомерности использования слова «музей» по отношению к интернет-ресурсам. Часть специалистов настаивает на том, что виртуальным музеем можно называть только ресурсы, созданные по законам музейного проектирования – с продуманным трехмерным пространством, кураторской концепцией и экспозиционной логикой (Лебедев 2010). Другая точка зрения заключается в том, что термин «виртуальный музей» уже приобрел самостоятельное значение, отличное от классического понимания музея, и может применяться к любым интернет-ресурсам, которые выполняют музейные функции – собирают,

сохраняют, интерпретируют и экспонируют культурное наследие в цифровой форме (Максимова 2014: 168).

В настоящем исследовании мы опираемся на расширенное понимание виртуального музея как цифровой платформы, обеспечивающей доступ к оцифрованным объектам культурного наследия (включая архивные документы, рукописи, предметы материальной культуры, аудиовизуальные материалы), структурированным и представленным таким образом, чтобы пользователь мог не только просматривать, но и изучать, интерпретировать и использовать эти объекты в образовательных, научных или просветительских целях.



Среди наиболее известных примеров – виртуальные туры Лувра (открывшего доступ к более чем 480 тысячам произведений), Британского музея (с возможностью виртуального посещения более 60 галерей и 4 миллионов экспонатов), а также виртуальные экспозиции Музеев Ватикана, Метрополитен-музея, Смитсоновского музея естественной истории, интернет-музея «Гулаг Онлайн», а также Музея Мирового океана, который создает уникальные VR-погружения для осмотра затонувших кораблей и мн. др.

Актуальность виртуальных музеев обусловлена несколькими факторами. Во-первых, ускоренная цифровизация всех сфер жизни, усилившаяся после пандемии COVID-19, сделала удаленный доступ к культуре не просто удобством, а необходимостью. Во-вторых, военные конфликты (например, разрушение памятников на Ближнем Востоке) и природные катастрофы (землетрясения, наводнения) демонстрируют уязвимость физического культурного наследия – цифровые копии часто остаются единственным способом сохранить информацию об утраченных объектах. В-третьих, развитие технологий искусственного интеллекта и машинного обучения открывает новые возможности для обработки и анализа больших массивов оцифрованных

данных, что превращает виртуальные музеи из пассивных хранилищ в активные исследовательские платформы.

Для Республики Казахстан тема виртуальных музеев обладает особой значимостью в силу нескольких обстоятельств. Прежде всего Казахстан – полиэтничное государство, на территории которого исторически сложились и продолжают развиваться культуры многих народов. Значительная часть материального и нематериального наследия этих этносов оказалась разрозненной, хранится в зарубежных архивах или в частных коллекциях. Виртуальные музеи могут предоставить механизм для «цифрового воссоединения» этих коллекций, что имеет важное значение для сохранения культурного разнообразия и формирования гражданской идентичности. Кроме того, в Казахстане уже реализуется ряд передовых проектов в этой области, которые заслуживают научного осмысления и обобщения (например, проект «Виртуальный музей немцев Казахстана»).

Целью настоящей работы является теоретико-методологический анализ феномена виртуального музея как элемента цифровой экосистемы наследия, а также выявление ключевых принципов и технологических решений, обеспечивающих переход от традиционного архивного фонда к виртуальному музейному пространству.

Материалы и методы

Для достижения поставленной цели в работе использован междисциплинарный подход, сочетающий методы культурологии, музеологии, источниковедения и цифровых гуманитарных наук (Digital Humanities).

Основу теоретической базы составили работы по вопросам сохранения культурного наследия (Тань Сичжу 2025; Трофимова и Шевелев, 2024); дефинирования и определения функций виртуального музея (Максимова 2014; Мальцева 2022); междисциплинарного изучения виртуальных музеев в контексте цифровой культуры (Дзюба 2019).

Для выяснения реального уровня осведомленности и понимания феномена виртуального музея среди населения Казахстана, а также для верификации теоретических положений, полученных из анализа литературы в рамках настоящего исследования, был проведен онлайн-опрос, в котором приняли участие 500 респондентов из 17 областей и 3 городов республиканского значения (Астана, Алматы, Шымкент). Выборка формировалась методом квот по возрастным группам, полу и типу поселения (город/село). Сбор данных путем анкетирования (структурированная анкета, включавшая как закрытые, так и открытые вопросы) осуществлялся с февраля по апрель 2026 года.

Анализ функционирующих виртуальных музеев (включая казахстанские проекты «Виртуальный музей тюркских народов», «Виртуальный музей немцев Казахстана», платформу E-museum, а также такие международные проекты, как «Гулаг Онлайн», Музей Мирового океана, Palmyra 3D и др.) проводился методом структурно-функционального анализа интерфейсов и контента. Оценивались такие параметры, как тип экспонатов (2D/3D/текст/аудио/видео), навигационная логика, наличие образовательных и исследовательских модулей, соответствие выделенным в литературе функциям виртуального музея. Оценивалась также

глубина метаданных, сохранение архивной атрибуции (фонд, опись, дело, лист) и способ перехода к музейной интерпретации.

Вместе с тем исследование было ограничено рядом факторов. Так, онлайн-формат опроса не охватил население без доступа к Интернету (прежде всего в сельской местности), а квотный метод выборки не смог обеспечить строгой статистической репрезентативности. Кроме того, анализ виртуальных музеев базировался только на публично доступных интерфейсах.

Теоретические основания и реальные проекты

(1) Дискуссии относительно понятия «виртуальный музей»

Феномен виртуального музея стал объектом научного осмысления относительно недавно – с середины 1990-х годов, когда музеи начали создавать свои цифровые прототипы в Интернете. Однако активная дискуссия развернулась в основном в XXI веке. Как показал обзор источников, в зарубежной науке сложилось несколько подходов к определению этого понятия.

Одни исследователи рассматривают виртуальный музей, прежде всего, как цифровую копию реально существующего музея (Самаковская 2012). В этом ключе виртуальный музей понимается как вспомогательный инструмент, расширяющий доступ к физическим коллекциям, но не обладающий самостоятельной ценностью. Такой подход до сих пор доминирует в практиках крупных музеев (например, Лувр[1*]), где виртуальные туры позиционируются как дополнение к основному экспозиционному пространству.

Другие авторы настаивают на самостоятельном статусе виртуального музея, который не обязан иметь физический аналог. Как отмечает Т.Е. Максимова, термин «виртуальный музей» уже приобрел самостоятельное значение, отличное от классического понимания музея, и «может применяться к любым интернет-ресурсам, которые выполняют музейные функции – собирают, сохраняют, интерпретируют и экспонируют культурное наследие в цифровой форме» (Максимова 2014: 168). Эту позицию разделяет и Д.Н. Дзюба, характеризующая виртуальный музей как постмодернистское формирование, посредством которого современный музей беспрепятственно приспосабливает свои традиционные функции к новой культурной среде цифрового века (Дзюба 2019).

Отдельные ученые считают, что виртуальность является имманентным свойством любого музея. Так, Э. Джаккарди говорит: «Все музеи виртуальны, независимо от информационных технологий» (Giaccardi 2004). По ее мнению, традиционный музей также создает «виртуальную» реальность, извлекая артефакты из их исконной среды и помещая в новое контекстуальное пространство. Следовательно, цифровые технологии не создают виртуальность с нуля, а лишь усиливают и предлагают новые формы для той виртуальности, которая уже в нем заложена.

В настоящем исследовании мы опираемся на расширенное понимание виртуального музея, которое позволяет преодолеть противопоставление «архивной аутентичности» и «музейной интерпретации» и рассматривать виртуальный музей как ключевой элемент цифровой экосистемы наследия.

[1*] Louvre Museum. (n.d.). Online tours. Retrieved May 10, 2026, from https://www.louvre.fr/en/online-tours

(2) Функции виртуального музея

Функциональный подход является одним из наиболее разработанных в литературе о виртуальных музеях. Обобщая работы Д.С. Василиной (2016), Д.Н. Дзюбы (2019) и Тань Сичжу (2025), можно выделить следующие ключевые функции, которые сегодня признаются большинством исследователей: сохранная (архивно-консервационная), репрезентационная (иногда называют экспозиционная), образовательная (или педагогическая), коммуникативная и интеграционная, научно-исследовательская и досуговая, а также просветительская функции.

(2.1) Виртуальный музей позволяет оцифровывать и сохранять объекты, находящиеся под угрозой физического разрушения: ветхие рукописи, документы на уязвимых носителях, предметы из деградирующих материалов, а также объекты, расположенные в зонах военных конфликтов или природных катастроф. Так, например, *проект «Пальмира 3D»*[1*], который был запущен в 2016 году, можно считать одним из наиболее показательных и технологически совершенных примеров реализации сохранной (архивно-консервационной) функции. Он создавался как прямой ответ на акт культурного вандализма для спасения наследия, оказавшегося под угрозой полного уничтожения.

(2.2) Виртуальный музей предоставляет возможность показать объекты, которые невозможно или нежелательно экспонировать в реальном пространстве (из-за хрупкости, размера, удаленности или правовых ограничений). Благодаря технологиям 3D-моделирования и виртуальной реальности создаются иммерсивные экспозиции, позволяющие пользователю «войти» в реконструированное пространство. В Казахстане одним из примеров репрезентационной (экспозиционной) функции является *виртуальный музей немцев Казахстана*[2*]. Основной миссией этого проекта является сохранение этнокультурной идентичности казахстанских немцев через оцифровку и систематизацию наследия. Для сбора экспонатов с 2018 по 2020 годы было организовано несколько экспедиций, охвативших 22 населенных пункта, было обследовано 17 государственных и школьных музеев, а также множество частных коллекций. Сейчас на платформе представлено 732 экспоната: от мебели и предметов быта до религиозных книг.

(2.3) Виртуальные музеи активно используются в школах, университетах и системах дополнительного образования, выполняя образовательную (педагогическую) функцию. Они предоставляют доступ к первоисточникам и научным материалам, адаптированным для разных возрастных категорий. Ю.Б. Шатон вводит понятие «культуры университетской памяти» и подчеркивает, что виртуальный музей способствует укреплению связей между студентами, сотрудниками, выпускниками и университетом (Шатон 2023: 104). В качестве примера можно привести виртуальные экспозиции, посвященные жизни и

[1*] Institute of Archaeology of the Russian Academy of Sciences. (n.d.). Palmyra 3D. Retrieved May 10, 2026, from https://rac.archeo.ru/projects/palmyra-3d/

[2*] Public Foundation Kazakhstan Association of Germans "Wiedergeburt". (n.d.). Virtual Museum of the Germans of Kazakhstan: About. Retrieved May 10, 2026, from https://wiedergeburt.asia/about/

наследию философа и мыслителя аль-Фараби, представленные музеем Казахского национального университета им. аль-Фараби[1*].

(2.4) Виртуальный музей выступает площадкой для объединения разрозненных коллекций – феномен, который в зарубежной литературе получил название «virtual reunification» (цифровое воссоединение). Р. Пунзалан (Punzalan 2014: 294-323) показывает, что этот процесс требует тесной кооперации между учреждениями, которые должны придерживаться единых стандартов представления информации на общей цифровой платформе. Так, например, в 2025 году ЮНЕСКО и Интерпол запустили «*Виртуальный музей похищенных культурных ценностей*»[2*]. В его основе лежит база данных Интерпола, насчитывающая более 52 000 украденных объектов. Сотни отобранных артефактов были воссозданы в виде 3D-моделей с помощью генеративного ИИ, поскольку их реальное местонахождение неизвестно и увидеть их вживую невозможно. Это один из самых успешных проектов, реализующих коммуникативную и интеграционную функцию.

(2.5) В последние годы исследователи все чаще акцентируют, что виртуальный музей может служить не только просветительским, но и исследовательским целям. Систематизированные и аннотированные цифровые коллекции становятся базой для текстологических, источниковедческих, лингвистических и культурологических исследований. Как отмечает Тань Сичжу, «виртуальные музеи становятся важным инструментом для получения данных о культурных ценностях со всего мира, [...] помогая ученым изучать малодоступные объекты культурного наследия» (Тань Сичжу 2025: 12). На наш взгляд, проект «*Гулаг Онлайн» (Gulag Online*)[3*] является одним из ярких примеров реализации научно-исследовательской функции в виртуальном музейном пространстве. Проект был создан на стыке археологии, архивных исследований (результаты многолетних экспедиций), картографии, устной истории, а также передовых IT-технологий.

(2.6) Наконец, выделяют также досуговую и просветительскую функцию, которая заключается в привлечении широкой аудитории к культурному наследию через визуально привлекательные, интерактивные и доступные форматы. Исследователи подчеркивают, что виртуальные музеи позволяют показывать объекты культурного наследия, которые просто невозможно поместить в рамки классического музейного пространства (Трофимова и др. 2024). Примечательно, что в Казахстане создана целая виртуальная *платформа E-museum*[1*], которая представляет собой пример реализации такой функции. Данный агрегатор объединяет в едином цифровом пространстве 285 музеев Казахстана, предоставляя пользователям 15 интерактивных 3D-туров, 750 трехмерных моделей уникальных артефактов и доступ к свыше 66 тыс. экспонатов. Благодаря порталу посетители могут в любое время и в любом месте

[1*] KazNU named after al-Farabi. (n.d.). Museum of KazNU named after al-Farabi. Retrieved May 10, 2026, from https://elibrary.kaznu.kz/ru/virtualnyj-tur-muzeya-al-farabi

[2*] UNESCO. (n.d.). Gold Dacian bracelet with scaled dragon. Virtual Museum of Stolen Cultural Objects. Retrieved May 10, 2026, from https://museum.unesco.org/stolen-objects/VM-RO-25-003_Gold_Dacian_bracelet_with_scaled_dragon

[3*] Gulag.cz. (n.d.). *Gulag Online – virtual museum*. Retrieved May 10, 2026, from https://gulag.online/

[1*] *E-museum:* единое цифровое пространство музеев Казахстана. URL: https://e-museum.kz/ru/main/ (дата обращения: 11.05.2026).

совершить виртуальное путешествие по залам даже самых удаленных региональных музеев, детально рассмотреть редкие артефакты и познакомиться с историей страны.

Безусловно, приведенный перечень функций не является исчерпывающим – по мере развития технологий (искусственный интеллект, иммерсивные среды, блокчейн для подтверждения подлинности) виртуальные музеи будут приобретать новые возможности.

(3) Архивный фонд как объект виртуализации

Представленные функции виртуального музея в равной мере относятся и к оцифрованным архивным фондам. Однако архивный документ (рукопись, дело, опись) обладает рядом особенностей, отличающих его от традиционного музейного экспоната. В отличие от предмета материальной культуры, изначально созданного для экспонирования или бытования, архивный документ ценен прежде всего своим контекстом происхождения (принцип *respect des fonds*) и аутентичностью. В этой связи превращение архивного фонда в виртуальный музей требует не просто сканирования и размещения изображений в сети, а сохранения архивной атрибуции (таких как фонд, опись, дело, лист) и одновременно выстраивания музейной интерпретации (т.е. нарратив, тематические маршруты, визуализация связей). Именно переход от линейной описи к интерактивной экспозиции, от учетной единицы к экспонату составляет суть проблемы, решаемой в рамках цифровой экосистемы наследия.

Являясь крупнейшим в мире цифровым собранием свидетельств о Холокосте и других геноцидах, *Visual History Archive* (VHA)[2*] фонда USC Shoah представляет собой пример портала, где атрибуция «архивного» документа сохранена. Каждое из более чем 56 000 видеоинтервью (общей продолжительностью свыше 117 000 часов) сопровождается тремя ключевыми элементами атрибуции (такие как уникальный идентификатор (ID) и индексные метаданные). Примечательно, что на разработку системы индексации фонд потратил годы: к 2001 году, через семь лет после основания, было обработано лишь около 4000 свидетельств из запланированных 52 000. На сегодняшний день Фонд USC Shoah является обладателем 11 патентов на технологии управления цифровыми коллекциями.

Таким образом, несмотря на развитие цифровых проектов, подход к сохранению оригинальной архивной атрибуции в международной системе виртуальных музеев остается фрагментарным. Чаще всего он реализуется в специализированных порталах государственных архивов, в то время как в гибридных виртуальных музеях и университетских проектах он либо упрощен, либо отсутствует.

Результаты опроса: цифровая экосистема глазами пользователей

В онлайн-опросе приняли участие 500 человек. Квотная выборка обеспечила пропорциональное представительство по трем параметрам: возраст (18-30 лет – 32%, 31-50 лет – 42%, старше 50 лет – 26%), пол (48% мужчин, 52% женщин) и тип поселения (город – 68%, село – 32%). Примечательно, что среди городских жителей осведомленность о виртуальных музеях составила 84%,

[2*] USC Shoah Foundation. (n.d.). Visual History Archive. Retrieved May 11, 2026, from https://vha.usc.edu/home

среди сельских – 62%. При этом доля тех, кто никогда не посещал виртуальный музей по причине «медленного Интернета», в селе оказалась втрое выше (23% против 7% в городе).

Как показал анализ ответов, 78% респондентов (390 человек) заявили, что слышали словосочетание «виртуальный музей». Однако при ответе на открытый вопрос «Что это, по-вашему?» лишь 42% (210 человек) дали определение, которое можно считать содержательно верным (например, «сайт музея с 3D-турами», «коллекция оцифрованных экспонатов в Интернете»). Остальные либо затруднились, либо дали упрощенные трактовки («виртуальный музей – это компьютерная игра», «очки виртуальной реальности»). Как видим, ответы респондентов свидетельствуют о необходимости методического сопровождения цифровых платформ в Казахстане.

Результаты анкетирования показали низкую вовлеченность жителей Казахстана в систему виртуальных музеев (архивов, агрегаторов). Только 29% респондентов (145 человек) сообщили, что когда-либо целенаправленно посещали какой-либо виртуальный музей. В открытых рубриках назывались зарубежные проекты (Эрмитаж, Музей искусства «Метрополитен», Британский музей), казахстанские (Музей искусства им. А. Кастеева, портал E-museum). Среди тех, кто не посещал, основными причинами названы: «не знал, что такое существует» (61%), «не интересуюсь музеями вообще» (23%), «медленный Интернет, не загружается» (12%), «нет компьютера/смартфона» (4%).

При этом среди тех, кто пробовал пользоваться казахстанскими виртуальными музеями, основными барьерами названы: недостаток информации о происхождении экспоната (58%), мало контента на казахском языке (52%), непонятная навигация (44%). В качестве желаемых улучшений пользователи предлагали проведение онлайн-акций по типу «Ночь в музее» с элементами геймификации, что могло бы повысить вовлеченность и привлечь молодежную аудиторию.

Ключевым вопросом исследования был «Могут ли оцифрованные архивные документы считаться полноценными музейными экспонатами?». Только 34% респондентов ответили утвердительно, 41% – «скорее да, но нужны пояснения», 25% – «нет, это разные вещи». В открытых комментариях часто встречалось: «архив – это для ученых, музей – для всех». Этот стереотип указывает на институциональный разрыв между архивным и музейным сознанием, который должна преодолеть цифровая экосистема.

На вопрос «Что для вас важнее при просмотре в виртуальном музее?» ответы распределились следующим образом: «красивая 3D-анимация и рассказ о ее истории» – 38%, «возможность увидеть документальные сведения и скачать копию» – 27%, «и то, и другое одинаково важно» – 35%.

При выборе предпочтения функционала «чисто музейные» функции (3D-туры, экскурсии – 58%) уступают исследовательским (скачивание, комментарии – 67%). Это подтверждает, что массовый пользователь, особенно молодой, видит в виртуальном музее не только развлечение, но и рабочий инструмент.

Таким образом, проведенное анкетирование подтверждает актуальность заявленной проблемы и выявляет конкретные запросы, которые должна учесть цифровая экосистема культурного наследия. Во-первых, несмотря на

формальную осведомленность о виртуальных музеях (78%), содержательное понимание остается поверхностным (42%), а реальная вовлеченность – крайне низкой (29%). Это свидетельствует о том, что простого наличия цифровых платформ недостаточно: требуется образовательное сопровождение и удобная навигация. Во-вторых, запрос на исследовательский функционал (скачивание, комментарии, ссылки на фонд/опись) оказался значительным (35-67%), особенно среди молодежи и городских пользователей. Однако казахстанские проекты пока до конца не отвечают этому запросу.

Опыт создания виртуального музея на базе университетских архивов

В рамках заявленной проблематики на базе Казахского национального университета им. аль-Фараби осуществляется попытка создания виртуального музея, посвященного тюркской ойкумене и тюркскому эпосу[1*].

Рис. 1. *Главная страница (эскиз) Виртуального музея*

Данный проект ориентирован на публичную репрезентацию, а также на сохранение и введение в научный оборот уникальных архивных материалов – рукописей, эпических текстов и личных фондов ученых-тюркологов.

Виртуальный музей включает несколько тематических залов, каждый из которых построен по принципу «от архивного документа – к музейному нарративу». Так, например, в Зале «Тюркский эпос» содержатся оцифрованные версии героического эпоса («Алпамыс батыр» и «Қобыланды батыр») с возможностью чтения переводов эпоса. Зал «Персоналии» содержит (помимо биографических материалов) данные из личных архивов ученых-тюркологов,

[1*] Реализуется в рамках научного проекта [ИРН AP23486380] «Тюркские языки Казахстана в условиях духовной модернизации общества: от графики к эпическому тексту» (по Программе грантового финансирования научных и научно-технических проектов на 2024 –2026 годы).

которые в разные годы были переданы в дар КазНУ (личный архив профессора Ш.Ш. Сарыбая, эпистолярный архив профессора А.С. Аманжолова, эпистолярный архив профессора И.Г. Добродомова и др.). Отдельный раздел «Репозиторий» включает структурированное собрание научных трудов по тюркологии, а также оцифрованные редкие рукописи (в том числе на древнетюркском, арабском, персидском и чагатайском языках) с возможностью поиска и скачивания.

Проект реализует основные функции виртуального музея, выделенные в теоретическом обзоре. Сохранная (архивно-консервационная) функция проявляется в оцифровке физически уязвимых рукописей (ветхие бумажные носители, выцветающие чернила) и личных архивов, которые ранее хранились в нерегулируемых условиях. Научно-исследовательская функция обеспечивается благодаря сохранению архивной атрибуции и внедрению полнотекстового поиска по рукописям, что позволяет исследователям цитировать цифровые копии. Образовательная функция реализуется через тематические залы и мультиязычные аннотации (казахский, русский, английский), адаптированные для студентов-тюркологов и широкой аудитории. Коммуникативная и интеграционная функция проявляется в том, что проект собирает воедино разрозненные коллекции: рукописи из лаборатории рукописей КазНУ, научные труды из разных библиотек – это пример «цифрового воссоединения» (virtual reunification) применительно к тюркскому наследию. Предполагается включение отдельных элементов с 3D-презентациями для некоторых экспонатов зала «Тюркский эпос».

В целом исследовательский функционал (наполнение контента, составление комментариев, 3D-моделирование и т.д.) уже частично реализован. Но на текущем этапе проект сталкивается с типичными ограничениями: недостаточный объем уже оцифрованных личных архивов, необходимость постоянного пополнения метаданных. Как показал опыт работы над созданием подобного ресурса, вузовский архив и музей, библиотечный фонд и научные исследования ученых университета могут стать ядром виртуального музея, удовлетворяющего одновременно и научные, и просветительские потребности.

Заключение

Проведенный анализ научной литературы, посвященной виртуальным музеям, позволяет констатировать, что, несмотря на значительное количество работ, посвященных данному феномену, практически отсутствуют работы, в которых систематически рассматривалось бы, как именно архивные документы (рукописи, дела, описи) трансформируются в музейные экспонаты в цифровой среде с сохранением их атрибуции и архивной аутентичности.

Эмпирические данные (результаты опроса, n=500) обосновывают необходимость создания цифровой экосистемы, которая будет направлена на сохранение культурного наследия, а также обеспечит исследовательскую ценность музейного и архивного фондов.

Приоритетными направлениями развития в этой области должны стать качественная интеграция метаданных, языковая доступность (казахский, английский и русский), образовательные модули и надежный исследовательский функционал.

Информация о гранте

Статья подготовлена при поддержке Министерства науки и высшего образования Республики Казахстан в рамках научного проекта [ИРН AP23486380] «Тюркские языки Казахстана в условиях духовной модернизации общества: от графики к эпическому тексту» (по Программе грантового финансирования научных и научно-технических проектов на 2024 –2026 годы).

Литература

Giaccardi 2004 – Giaccardi E. Memory and Territory: New Forms of Virtuality for the Museum. In: Museums and the Web 2004: Proceedings. Toronto: Archives & Museum Informatics, 2004. URL: https://museumsandtheweb.com/mw2004/papers/giaccardi/giaccardi.html (дата обращения: 26.04.2026).

Punzalan 2014 – Punzalan R.L. Understanding Virtual Reunification. The Library Quarterly: Information, Community, Policy. 2014. 84(3), 294–323. DOI: https://doi.org/10.1086/676489

Василина 2016 – Василина Д.С. Виртуальный музей как феномен современной культуры // Международный журнал исследований культуры. 2016. № 3 (24). С. 96–102.

Дзюба 2019 – Дзюба Д.Н. Виртуальный музей в контексте цифровой культуры: автореф. дис. … канд. культурологии: 24.00.01. Саранск, 2019. 24 с.

Лебедев 2010 –Лебедев А.В. Виртуальные музеи и виртуализация музея // Мир музея. 2010. № 10. С. 5-9.

Максимова 2014 – Максимова Т.Е. Виртуальные музеи: анализ термина // Вестник Российского государственного гуманитарного университета. Серия: Философия. Социология. Искусствоведение. 2014. № 14. С. 163-169.

Мальцева 2022 – Мальцева С.С. Виртуальный музей: понятие, инструменты работы, перспективы использования // Документ. Архив. История. Современность: сборник научных трудов. Екатеринбург: Издательство Уральского университета, 2022. Вып. 22. С. 14–23.

Самаковская 2012 – Самаковская О.В. Продвижение этнографических музеев Западной Сибири в электронном информационном пространстве: автореф. дис. … канд. культурологии: 24.00.03. Кемерово, 2012. 24 с.

Тань Сичжу 2025 – Тань Сичжу. Виртуальные музеи как феномен современной культуры – влияние виртуальных музеев на доступность культурного наследия в контексте культурного разнообразия // Культура и цивилизация. 2025. Т. 15. № 3А. С. 7–16.

Трофимова и др. 2024 – Трофимова Т.В., Шевелев В.Д. Виртуальный музей: сохранение и репрезентация историко-культурного // Актуальные проблемы гуманитарных и социально-экономических наук: сборник статей Международной научно-практической конференции. 2024. С. 14–17.

Шатон 2023 – Шатон Ю.Б. Виртуальный музей как инструмент репрезентации культуры университетской памяти // Культура и искусство Урала: история, реалии, перспективы. 2023. С. 101–105.

INVİSİBİLİTY İN DİGİTAL ARCHİVES: THE SELECTİON AND PRESENTATİON OF INFORMATİON

Ersen Özkul
Directorate of State Archives, Türkiye
ersen.ozkul@devletarsivleri.gov.tr

Abstract. Digital archives are positioned as one of the fundamental infrastructures of the information society, promising unlimited and equal access to information. However, this study argues that this promise is significantly constrained in practice and that digital archives produce new forms of invisibility. The frequently experienced problem of "not being able to find the sought information" is not merely a technical deficiency, but a structural issue related to metadata production, classification systems, linguistic differences, and algorithmic ranking processes. In addition, records that exist within archives yet remain inaccessible or marginalized are evaluated within the framework of "latent information" and "archival silence." The study analyzes how invisibility in digital archives is produced at the intersection of the selection and presentation of information, demonstrating that this process extends beyond technical dimensions to include epistemological and political implications. By examining international metadata standards, open access policies, participatory archiving approaches, and AI-based applications, the study discusses current interventions addressing these challenges. Ultimately, it argues that access in digital archives is determined not only by the existence of data, but by how it is defined, organized, and presented, and proposes a framework for more inclusive, transparent, and accountable archival systems.

Keywords: digital archives; invisibility; information access; metadata; algorithmic ranking; archival silence

ЦИФРЛЫҚ АРХИВТЕРДЕГІ КӨРІНБЕЙТІН ДҮНИЕЛЕР: АҚПАРАТТЫ ІРІКТЕУ ЖӘНЕ ҰСЫНУ ТӘСІЛДЕРІ

Эрсен Өзкүл
Түркия Республикасы Мемлекеттік архивтер басқармасы, Түркия
ersen.ozkul@devletarsivleri.gov.tr

Аңдатпа. Цифрлық архивтер ақпараттық қоғамның іргелі инфрақұрылымдарының бірі ретінде қарастырылып, ақпаратқа шектеусіз және тең қолжетімділікті қамтамасыз етеді деп болжанады. Алайда, бұл зерттеу тәжірибе жүзінде аталған тұжырымның айтарлықтай шектеулі екенін және цифрлық архивтердің «көрінбеушіліктің» жаңа формаларын тудыратынын дәлелдейді. Жиі кездесетін «іздеген ақпаратты таба алмау» мәселесі – тек техникалық кемшілік қана емес, ол метадеректерді жасауға, жіктеу жүйелеріне, тілдік айырмашылықтарға және алгоритмдік ранжирлеу процестеріне байланысты туындайтын құрылымдық мәселе. Сонымен қатар, архивтерде бар, бірақ қолжетімсіз немесе маргиналданған жазбалар «жасырын ақпарат» (latent information) және «архивтік үнсіздік» (archival silence) шеңберінде бағаланады. Зерттеу цифрлық архивтердегі көрінбеушіліктің ақпаратты іріктеу мен ұсынудың түйісу нүктесінде қалай қалыптасатынын талдайды және бұл процестің техникалық өлшемдерден тыс, эпистемологиялық және саяси салдары бар екенін көрсетеді. Халықаралық метадеректер стандарттарын, ашық қолжетімділік саясатын, партисипаторлық (қатысушылық) архивтеу тәсілдерін және ЖИ негізіндегі қосымшаларды зерттей отырып, бұл еңбек аталған қиындықтарды шешудің қазіргі жолдарын қарастырады. Қорытындылай келе, цифрлық архивтердегі қолжетімділік тек деректердің бар болуымен емес, олардың қалай анықталуы, ұйымдастырылуы және ұсынылуымен айқындалатыны тұжырымдалады және инклюзивті, ашық әрі есеп беруге тиіс архивтік жүйелерді құрудың тұжырымдамасы ұсынылады.

Түйін сөздер: цифрлық архивтер; көрінбеушілік; ақпаратқа қолжетімділік; метадеректер; алгоритмдік ранжирлеу; архивтік үнсіздік

НЕВИДИМОЕ В ЦИФРОВЫХ АРХИВАХ: КАК ОТБИРАЕТСЯ И ПРЕДСТАВЛЯЕТСЯ ИНФОРМАЦИЯ

Эрсен Озкул
Управление государственных архивов Турецкой Республики, Турция
ersen.ozkul@devletarsivleri.gov.tr

Аннотация. Цифровые архивы позиционируются как одна из фундаментальных инфраструктур информационного общества, предполагающая неограниченный и равный доступ к информации. Однако в данном исследовании утверждается, что на практике этот тезис существенно ограничен и что цифровые архивы создают новые формы невидимости. Часто возникающая проблема «невозможности найти искомую информацию» является не просто техническим недостатком, а структурной проблемой, связанной с созданием метаданных, системами классификации, языковыми различиями и процессами алгоритмического ранжирования. Кроме того, записи, существующие в архивах, но остающиеся недоступными или маргинализированными, рассматриваются в рамках понятий «скрытая информация» (latent information) и «архивное молчание». В исследовании анализируется, как невидимость в цифровых архивах формируется на пересечении процессов отбора и представления информации, демонстрируя, что этот процесс выходит за рамки технических аспектов и включает в себя эпистемологические

и политические последствия. Изучая международные стандарты метаданных, политику открытого доступа, подходы партисипаторного архивирования и приложения на базе ИИ, исследование обсуждает текущие методы решения этих проблем. В конечном итоге делается вывод, что доступ к цифровым архивам определяется не только наличием данных, но и тем, как они определяются, организуются и представляются; также предлагается концепция для создания более инклюзивных, прозрачных и подотчетных архивных систем.

Ключевые слова: цифровые архивы; невидимость; доступ к информации; метаданные; алгоритмическое ранжирование; архивное молчание

Introduction: The Paradox of Invisibility in Digital Abundance

Digital archives are often presented as fundamental infrastructures of the information society, promising unlimited and equal access to information. However, recent user experiences and critical archival studies suggest that this promise must be critically reconsidered. Large-scale digitization projects have expanded access quantitatively by making extensive archival collections available online. Nevertheless, the frequently encountered problem of “not being able to find the sought information” clearly demonstrates that access cannot be reduced to the mere availability of data. In this context, digital archives should not be understood as neutral access tools, but as socio-technical systems that actively structure and shape access (Gilliland 2014).

This contradiction can be described as a condition of “invisibility within 3abundance.” The growth in data volume does not automatically lead to increased accessibility; rather, it produces a visibility regime in which certain types of information are foregrounded while others remain in the background. In this sense, invisibility cannot be explained solely by missing or absent data, but must be understood as the

result of technical and institutional processes that render existing information inaccessible or unnoticed. In archival literature, this condition is often associated with the concept of “archival silence,” emphasizing that archives do not passively reflect social memory but actively construct it (Cook 2001; Schwartz and Cook 2002; Caswell 2016).

The selective nature of archives continues to shape digital environments. Decisions regarding what is considered valuable, how it is described, and in which context it is presented remain central to archival processes. These decisions are shaped not only by technical requirements but also by epistemological and institutional frameworks. As Derrida argues, the act of archiving always involves practices of selection and exclusion, determining not only what is preserved but also what is forgotten (Derrida 1995).

Within this framework, visibility in digital archives is shaped not only by what is selected, but also by how it is presented. Metadata structures, classification systems, and knowledge organization determine how information is represented within systems, while search engines, ranking algorithms, and user interfaces shape how this information is encountered by users. In particular, the ranking of search results and recommendation mechanisms increase the visibility of certain content while pushing others into the background. This demonstrates that digital archives are not merely repositories of information, but systems that actively produce visibility (Noble 2018; Bucher 2018).

This study examines how the selection and presentation of information shape invisibility in digital archives, focusing on the problem of “unfindable information,” the persistence of latent records, and the role of algorithmic processes. It also explores current approaches and international efforts aimed at addressing these challenges.

Selection, Presentation, and Invisibility in Digital Archives

Digital archives are often presented as comprehensive and unlimited information systems; however, the processes of selection and presentation that underpin these systems function as primary mechanisms shaping the visibility of information. Historically, archives have never been neutral structures representing all information equally; rather, they have evolved as systems that prioritize certain types of knowledge while excluding others. Therefore, archival processes should be understood not merely as technical arrangements, but as integral components of knowledge production and representation (Cook 2001; Schwartz and Cook 2002).

Digitization has not eliminated this selective nature but has instead rendered it more complex. While the expansion of data volume appears to increase inclusivity, it actually creates an environment in which visibility is reorganized across multiple levels. In this context, it is not only the inclusion of information in archives that matters, but also how it is defined and in what context it is presented. This demonstrates that the practices of selection and exclusion inherent in archival systems continue to operate in digital environments (Derrida 1995).

The mode of presentation plays a central role in shaping visibility within digital archives. Users access information primarily through search systems and interfaces, which determine not only accessibility but also how content is encountered. This indicates that digital archives function not as passive repositories, but as systems that actively structure visibility (Noble 2018).

In this framework, invisibility in digital archives cannot be explained solely by the absence of information. Rather, it emerges when existing information remains in the background or goes unnoticed due to particular forms of representation and presentation. This highlights that archival processes determine not only what is preserved, but also what becomes visible.

These processes are also shaped by institutional and social contexts. Digitization priorities, resource allocation, and representational choices directly influence which types of information are foregrounded. The limited representation of records related to minority groups and alternative narratives reveals that invisibility also has a significant social dimension (Caswell 2016).

Mechanisms of Invisibility in Digital Archives: Technical and Structural Dimensions

One of the primary mechanisms shaping invisibility in digital archives is metadata production. Metadata constitutes the fundamental structure through which information is described and made accessible, while also playing a critical role in determining visibility. How a record is labeled, which keywords it is associated with, and which classification system it is assigned to directly influence whether it appears in search results. For this reason, incomplete or inaccurate metadata may render records effectively "present but inaccessible" within the system. As Gilliland emphasizes, findability in digital archives often depends more on the quality of metadata than on the content itself (Gilliland 2014: 45).

Another key mechanism of invisibility in digital archives lies in search systems and algorithmic ranking processes. Users' access to information is largely mediated through these systems, which determine not only what can be found but also what becomes visible. Search engines and database systems operate according to specific algorithmic rules and present results based on these criteria. However, these ranking processes are often opaque and may systematically privilege certain types of content. Noble's analysis of the non-neutrality of search algorithms clearly demonstrates this issue (Noble 2018: 5).

The role of algorithmic systems in producing invisibility has become even more pronounced with the increasing use of artificial intelligence. Automated tagging, recommendation systems, and data classification algorithms operate based on patterns within large datasets, making some content visible while relegating others to the background. As Bucher argues, algorithms function as "invisible forces" that shape the informational landscape encountered by users (Bucher 2018: 67). In this sense, algorithms should be understood not merely as technical tools but as structural actors that regulate access to information.

Linguistic and terminological differences also constitute a significant dimension of invisibility. Although digital archives are often multilingual, search systems tend to operate within specific terminological frameworks. This can lead to unsuccessful queries when users employ alternative conceptual vocabularies, thereby rendering existing information inaccessible. Hjørland emphasizes that information retrieval processes are fundamentally shaped by conceptual frameworks and approaches to knowledge organization, making terminological differences a decisive factor in access (Hjørland 2021).

Finally, invisibility in digital archives is closely related to institutional and political choices. Decisions regarding which collections are digitized, which datasets are prioritized, and which materials are highlighted play a central role in shaping archival visibility. These processes, often influenced by limited resources and policy priorities, may result in the systematic exclusion or marginalization of certain types of information. As Caswell notes, archives are not only institutions that document the past, but also those that determine which past is remembered (Caswell 2016: 14).

International Approaches to Reducing Invisibility in Digital Archives

The problem of invisibility in digital archives is increasingly addressed by the international archival community through a range of policies, standards, and practical initiatives. These approaches aim to promote more equitable access to information; however, rather than eliminating invisibility entirely, they primarily function as interventions that mitigate its effects.

Metadata standards constitute one of the most fundamental tools in this field. Standards such as ISAD(G), EAD, and Dublin Core aim to ensure greater consistency, interoperability, and accessibility of archival data. In this context, frameworks developed by International Council on Archives seek to establish a shared language for archival description at an international level. Nevertheless, the assumption of universality underlying these standards has been criticized from the perspective of cultural and linguistic diversity. Gilliland argues that metadata standards are not always flexible enough to accommodate diverse forms of knowledge production, which may result in certain records remaining invisible (Gilliland 2014: 52).

Open access and open data movements represent another important set of initiatives aimed at expanding access to information. These approaches are based on the principle that information should not only be preserved but also made available to the widest possible audience. Policies developed by UNESCO emphasize that universal access to knowledge is a fundamental component of democratic societies (UNESCO 2012). At the same time, open access practices have been criticized for reproducing inequalities, particularly in cases where institutions with limited resources are unable to fully participate.

Participatory archives have emerged as a significant approach in recent years. This model envisions users not only as consumers of information but also as contributors to metadata production and content creation. In doing so, it aims to enhance the visibility of records related to local communities and marginalized groups. Caswell highlights that participatory archival practices have the potential to reduce archival silence and facilitate the inclusion of alternative narratives (Caswell 2016). However, issues such as sustainability and data reliability remain ongoing concerns.

Artificial intelligence and machine learning-based systems also offer new possibilities for addressing invisibility in digital archives. Automated tagging, content analysis, and semantic search technologies enable the identification of previously unrecognized relationships within large datasets. These tools are particularly valuable in addressing limitations related to insufficient metadata. However, they also carry the risk of reproducing bias due to their dependence on training data, potentially reinforcing invisibility (Bucher 2018: 89; Noble 2018: 121). This highlights the need for technological solutions to be complemented by ethical oversight mechanisms.

International digital archival platforms provide additional insights into the dynamics of visibility. Projects such as Europeana and Digital Public Library of America aim to increase access by aggregating collections from multiple institutions within a single framework. While these platforms improve accessibility through data integration and user-friendly interfaces, they also generate new visibility dynamics by determining which content is foregrounded.

Conclusion

The widespread assumption that digitization automatically expands access to information is significantly constrained in practice. Rather than creating an environment of equal access, the increase in data volume produces a new regime of visibility in which certain types of information are foregrounded while others remain in the background. This demonstrates that digital archives should be understood not merely as technical infrastructures, but as epistemological and political systems that shape access to information.

Invisibility in digital archives emerges at the intersection of the selection, description, and presentation of information. In this sense, invisibility should be understood not only as a result of missing data, but also as an outcome of representational and access processes. In particular, the growing influence of algorithmic systems indicates that visibility is increasingly shaped through automated and often opaque processes from the user's perspective (Noble 2018; Bucher 2018).

Access in digital archives is determined not only by the existence of information, but also by how it is defined, organized, and presented to users. Therefore, the development of more inclusive, transparent, and accountable archival systems requires not only technical solutions, but also the adoption of critical, ethical, and participatory approaches. In this context, participatory archival practices offer significant potential for making visible the knowledge of underrepresented groups (Caswell 2016).

Ultimately, digital archives should be reconsidered not only as repositories of the past, but as active systems that determine what becomes visible and shape the future organization of knowledge.

References

Bucher 2018 – Bucher T. If... Then: Algorithmic Power and Politics. Oxford: Oxford University Press, 2018. 224 p.

Caswell 2016 – Caswell M. "The Archive" Is Not an Archives: Acknowledging the Intellectual Contributions of Archival Studies. Reconstruction, 2016. Vol. 16, No. 1. Available at: https://escholarship.org/uc/item/7bn4v1fk (date of access: 23.04.2026).

Caswell 2016a – Caswell M., Migoni A., Geraci N., Cifor M. "To Suddenly Discover Yourself Existing": Uncovering the Impact of Community Archives. The American Archivist, 2016. Vol. 79, No. 1. P. 56–81.

Cook 2001 – Cook T. Archival Science and Postmodernism: New Formulations for Old Concepts. Archival Science, 2001. Vol. 1. No. 1. P. 3–24.

Derrida 1995 – Derrida J. Archive Fever: A Freudian Impression. Chicago: University of Chicago Press, 1995. 112 p.

Gilliland 2014 – Gilliland A. Conceptualizing 21st-Century Archives. Chicago: Society of American Archivists, 2014. 24 p.

Hjørland 2021 – Hjørland B. Information Retrieval and Knowledge Organization: A Perspective from the Philosophy of Science. Information, 2021. Vol. 12, No. 3. 135. DOI: 10.3390/info12030135.

International Council on Archives 2000 – International Council on Archives. *ISAD(G):* General International Standard Archival Description. 2nd ed. Paris: International Council on Archives, 2000.

International Council on Archives 2016 – International Council on Archives. Records in Contexts Conceptual Model. Paris: International Council on Archives, 2016.

Noble 2018 – Noble S. U. Algorithms of Oppression: How Search Engines Reinforce Racism. New York: New York University Press, 2018. 256 p.

Schwartz 2002 – Schwartz J. M., Cook T. Archives, Records, and Power: The Making of Modern Memory. Archival Science, 2002. Vol. 2, No. 1–2. P. 1–19.

UNESCO 2012 – Swan A. Policy Guidelines for the Development and Promotion of Open Access. Paris: UNESCO, 2012. 76 p.

HISTORY REIMAGINED: HOW THE NATIONAL ARCHIVES OF SINGAPORE'S COLLECTION TURNS IDEAS INTO CREATIVE WORKS IN AN ISLAND-WIDE COMPETITION

June Pok
Master of Science (Information Studies), Nanyang Technological University,
Archivist (Audio Visual Archives), National Archives of Singapore
June_pok@nlb.gov.sg

Abstract. Twenty twenty-five was a jubilant year for the National Library and Archives Board (NLB). That year, the statutory board turned 30. To commemorate its journey to bring reading and learning to Singaporeans and to celebrate the Republic of Singapore's 60th year of independence, NLB launched, among its exciting series of events, a branding campaign to position the National Archives of Singapore (NAS) and National Library (NL) as the key memory institution for Singapore's documentary heritage. Several initiatives were rolled out. These included:

1. the launch of "Be Inspired by History", a creative contest which was conducted via social media whereby members of the public (including students, content creators, writers, and researchers) were invited to create original works that capture Singapore's past and future using archival materials from NAS and NL,
2. producing a series of campaign videos and collaterals, and
3. the hosting of three public programmes led by local creators and creatives who have tapped into NAS-NL materials for their own projects.

As an institution of the NLB, the National Archives of Singapore was involved in the campaign in three areas. First, it supplied suitable materials from its collection for the campaign videos. Second, it provided participants with copyright-cleared materials that they could freely use for the contest. Third, it collected the contest video submissions from the top three prize winners.

Through outreach efforts such as these, the NAS not only increases public awareness of its rich archival collection and Singapore's heritage and history, but it also fulfils its mission to actively acquire records that will serve as the social memory of our people.

Keywords: National Archives of Singapore; National Library and Archives Board; National Library; memory institution; Be Inspired by History; archival collection; records, public engagement

ТАРИХТЫ ҚАЙТА ПАЙЫМДАУ: СИНГАПУР ҰЛТТЫҚ АРХИВІ ҚОРЫНЫҢ ЖАЛПЫАРАЛДЫҚ КОНКУРС АЯСЫНДА ИДЕЯЛАРДЫ ШЫҒАРМАШЫЛЫҚ ТУЫНДЫЛАРҒА АЙНАЛДЫРУ ТӘЖІРИБЕСІ

Джун Пок
Жаратылыстану ғылымдарының магистрі (ақпараттық зерттеулер),
Наньян технологиялық университеті; архивші (аудиовизуалды архивтер бөлімі),
Сингапур Ұлттық архивтері
June_pok@nlb.gov.sg

Аңдатпа. 2025 жыл Ұлттық кітапхана және архивтер кеңесі (NLB) үшін мерейлі жыл болды. Дәл сол жылы бұл заңды орган өз қызметінің 30 жылдығын атап өтті. Сингапур халқына оқу мен білімді насихаттаудағы жолын еске алу және Сингапур Республикасының тәуелсіздігіне 60 жыл толуын мерекелеу аясында, NLB бірқатар қызықты іс-шаралармен қатар, Сингапурдың Ұлттық архиві (NAS) мен Ұлттық кітапханасын (NL) елдің деректі мұрасының негізгі жад институты ретінде айқындауға бағытталған брендтік науқанды іске қосты. Осы науқан аясында бірнеше бастама қолға алынды, олардың ішінде:
NAS және NL архивтік материалдарын пайдалана отырып, Сингапурдың өткені мен болашағын бейнелейтін түпнұсқа жұмыстар жасауға шақыратын әлеуметтік желілердегі «Тарихтан шабыт ал» («Be Inspired by History») атты шығармашылық байқау;
науқанға арналған бейнероликтер мен жарнамалық материалдар топтамасын өндіру;
өз жобаларында NAS-NL материалдарын пайдаланған жергілікті авторлар мен креативті тұлғалар жетекшілік ететін үш қоғамдық бағдарламаны өткізу.
NLB институты ретінде Сингапурдың Ұлттық архиві бұл науқанға үш бағытта атсалысты: біріншіден, науқандық бейнероликтер үшін өз қорынан сәйкес материалдарды ұсынды; екіншіден, қатысушыларға байқауда еркін пайдалану үшін авторлық құқықтан тазартылған материалдар берді; үшіншіден, байқаудың жүлдегері атанған үздік үш видеоны қабылдап алды. Осындай ауқымды ақпараттық-түсіндіру жұмыстары арқылы NAS өзінің бай архивтік қоры және Сингапурдың тарихы мен мұрасы туралы жұртшылықтың хабардарлығын арттырып қана қоймай, сонымен бірге халқымыздың әлеуметтік жады ретінде қызмет ететін құжаттарды белсенді түрде жинақтау миссиясын орындайды.

Түйін сөздер: Сингапурдың Ұлттық архиві; Ұлттық кітапхана және архивтер кеңесі; Ұлттық кітапхана; жад институты; «Тарихтан шабыт ал»; архивтіккор; құжаттар; қоғаммен жұмыс

ПЕРЕОСМЫСЛЕНИЕ ИСТОРИИ: КАК КОЛЛЕКЦИЯ НАЦИОНАЛЬНОГО АРХИВА СИНГАПУРА ПРЕВРАЩАЕТ ИДЕИ В ТВОРЧЕСКИЕ РАБОТЫ В РАМКАХ ОБЩЕОСТРОВНОГО КОНКУРСА

Джун Пок
Магистр естественных наук (информационные исследования), Наньянский технологический университет; архивист (отдел аудиовизуальных архивов), Национальные архивы Сингапура
June_pok@nlb.gov.sg

Аннотация. 2025 год стал знаменательным для Совета национальных библиотек и архивов (NLB). В этом году уставный совет отметил свое 30-летие. Чтобы отметить путь по продвижению чтения и знаний среди сингапурцев и отпраздновать 60-летие независимости Республики Сингапур, NLB, наряду с рядом других захватывающих мероприятий, запустил брендинговую кампанию, направленную на позиционирование Национального архива Сингапура (NAS) и Национальной библиотеки (NL) в качестве ключевого института памяти документального наследия Сингапура. Было реализовано несколько инициатив, в том числе: запуск творческого конкурса в социальных сетях «Вдохновляйтесь историей» («Be Inspired by History»), в рамках которого общественность (включая студентов, создателей контента, писателей и исследователей) была приглашена к созданию оригинальных работ, отражающих прошлое и будущее Сингапура с использованием архивных материалов NAS и NL; производство серии рекламных видеороликов и сопутствующих материалов кампании; проведение трех публичных программ под руководством местных авторов и креативных деятелей, использовавших материалы NAS-NL в своих проектах.

Как институт NLB Национальный архив Сингапура участвовал в кампании по трем направлениям. Во-первых, он предоставил подходящие материалы из своего фонда для видеороликов кампании. Во-вторых, он предоставил участникам материалы, очищенные от авторских прав, которые они могли свободно использовать для конкурса. В-третьих, он принял конкурсные видеоработы от трех призеров. Благодаря таким просветительским усилиям NAS не только повышает осведомленность общественности о своей богатой архивной коллекции, а также наследии и истории Сингапура, но и выполняет свою миссию по активному сбору документов, которые будут служить социальной памятью нашего народа.

Ключевые слова: Национальный архив Сингапура; Совет национальных библиотек и архивов; Национальная библиотека; институт памяти; «Вдохновляйтесь историей»; архивный фонд; документы; взаимодействие с общественностью

Introduction

The National Archives and Library Board

The National Library Board (NLB) was established on 1 September 1995 as a statutory board to promote reading and learning, improve information literacy, (The NLB Brand, n.d.) and foster an appreciation of history and heritage among Singaporeans (About NLB, n.d.). Managing a network of 21 public libraries across Singapore, the board also manages the National Library (NL), a reference library responsible for the preservation of the nation's cultural and literary heritage, and the National Archives of Singapore (National Library Board 2024: 2), the official custodian of government records of Singapore.

The National Archives of Singapore

The National Archives of Singapore (NAS) was established in 1968 by the National Archives and Records Centre Act (1967) (About Us: History, n.d.). On 1 August 1993, the NAS merged with the Oral History Department and the National Museum of Singapore to form the National Heritage Board (About NHB, n.d.). However, as part of a restructuring of government ministries and portfolios, the NAS joined NLB on 1 November 2012 (About Us: History, n.d.). In 2025, to reflect that the National Archives of Singapore is part of NLB, the board's name was changed to the National Library and Archives Board (The NLB Brand, n.d.).

In its mission to fulfil the corporate memory of the Singapore Government and the social memory of the people, the NAS has been actively collecting records from both public agencies and the private sector that are deemed to be of national or historical significance. These records come in different media and formats and are preserved securely in its holdings (About Us: Our Roles, n.d.). The rich collection includes audiovisual recordings, oral history interviews, photographs, speeches and press releases, posters, maps, building plans, private records and copies of records from overseas, Straits Settlements Records, Federated Malay States records, and other volumes, title deeds and other registers. With the advent of technology and as organisations undergo digital transformation (About Singapore Digital, n.d.), an increasing number of archival records in the collection is in digital content — both born-digital and digitised.

Outreach

One mandate that is key to the NAS is advocacy (About Us: Mandate, n.d.). Over the years, it has rolled out several outreach initiatives to engage and educate the public about its collections and services, chiefly via exhibitions, publications and public programming (Tan 2019: 18). *The Albatross File: Singapore's Independence*

Declassified exhibition (Discover the Story of Singapore's Road to Independence — Ticketing Now Open for The Albatross File: Singapore's Independence, 2025), which was launched on 18 November 2025, is among the many exhibitions that the NAS has been involved in. Organised to commemorate Singapore's 60th anniversary as an independent nation, this SG60 capstone showcase features, for the first time, declassified extracts of oral histories from Singapore's founding Prime Minister Mr Lee Kuan Yew and former Deputy Prime Minister Dr Goh Keng Swee.

Besides exhibitions, the NAS organises programmes for the public. Some are targeted at young children, educating them about the country's history and heritage through fun, hands-on activities. There are also public talks, engagement sessions, conferences, and movie screenings attended by members of the public such as historians, students, veterans and history enthusiasts, all of which were well received (Public Programmes, n.d.). As technology advanced, so did NAS' operations — this meant digitising its vast collection to make it more accessible on Archives Online, the NAS' search engine. Digitisation also enabled the NAS to engage with its younger audiences and change how outreach was conducted (Jayakumar 2019: 3).

From online children's programmes to monthly series of online talks given by NAS staff or invited guests, the NAS has curated a number of initiatives featuring specific collections. To view NAS' online offerings, visitors can visit Archives Online, a platform where researchers can submit requests for access to and the reproduction of records. The NAS also took to social media as a means of reaching digital natives and millennials. One experiment that proved to be extremely successful was the NAS TikTok Initiative (NASTI). Not only did it increased awareness of the NAS' collection, it put our archivists, specialists, and officers working behind the scenes into the spotlight.

Be Inspired by History

As we have seen, using digital content to connect with one's audience certainly has its advantages. As the NAS continued to grow its digital collection, what better way to let it be discovered than by partnering with NL in a branding campaign — one that is aligned with NLB's Singapore Storytellers vision of nurturing storytellers and inspiring the discovery and creation of Singapore stories. This took the form of "Be Inspired by History", a branding campaign to position NAS and NL as the key memory institutions for Singapore's documentary heritage. Launched on 23 July 2025 as part of NLB's celebratory efforts to commemorate its 30-year journey to bring reading and learning to Singaporeans and to celebrate the Republic of Singapore's 60th year of independence, the campaign rolled out several initiatives.

These included organising programmes led by local creators and creatives who had tapped into NAS-NL materials for their own projects. Held from August to September 2025, these free public programmes were open to members of the public, who could register online. Some of the speakers included Ho Yong Min, the innovative storyteller and heritage enthusiast behind the Instagram and TikTok accounts @Urbanist.Singapore, and Felix Cheong, a local poet and author whose works have been longlisted for the Frank O'Connor Award. (Create your masterpiece with rich resources from NLB's libraries and archives, 2025).

The highlight of the branding campaign, however, was the launch of the "Be Inspired by History", a creative contest conducted via social media whereby members

of the public (including students, content creators, writers, and researchers) were invited to create original works that captured Singapore's past and future using archival materials from NAS and NL (Create your masterpiece with rich resources from NLB's libraries and archives, 2025).

For the contest, materials from the collections of both the NAS and the NL were curated according to the eight selected topics and uploaded onto the contest website, where they were available for free download during the contest period. The topics featured, including their descriptions, are as follows:

1. Food Heritage – Look back at the beginnings of our hawker scene, wet markets and food centres and how Singapore quite literally became a melting pot of culture and cuisine.
2. Attractions and Entertainment – Amusement parks, Botanic Gardens, the Zoo, our tiny island has never been in want of things to see and do.
3. Customs and Cultures – Singapore’s multicultural society means we are privy to the celebrations, rituals, beliefs and festivals of a myriad of cultures – each distinct and unique, but all a part of Singapore.
4. Childhood Days – Fewer screens, more time outside, painting the kampung red with neighbourhood friends – what do you recall, miss, or heard of those days gone by?
5. Vanishing Trades – Basket-weaving, trishaw riding, fortune-telling, these are just a handful of occupations that are fading into time.
6. Transport – It’s hard to imagine that just a few decades ago rickshaws, trishaws, and mosquito buses were commonplace and the airconditioned bus, a much sought after rarity.
7. Urban Spaces – Kampungs, bungalows, shophouses, colonial architecture and landmarks – Singapore’s history is all around if we’d just take a moment to look.
8. Surprise Me – Here are some surprising and unexpected finds from the National Archives and National Library!

Participants were allowed to use the readily available resources for their creations or browse for additional resources on Archives Online and/or National Library Online. To enter the contest, participants were required to create a piece of digital content (e.g. a video, photo carousel, poem, song, etc.) featuring the resource(s), explain how and why the resource(s) inspired their creation in their video, image(s) and/or captions, post their content on Instagram, TikTok, Facebook, or YouTube, and include the hashtag #NLBInspiredByHistory in the post caption while tagging three friends. Open to all residents in Singapore, submissions had to be in English, and participants had to ensure that their posts and profile were made public. Only post published on TikTok, Instagram, Facebook, and YouTube accounts from 23 July to 23 October 2025 were considered.

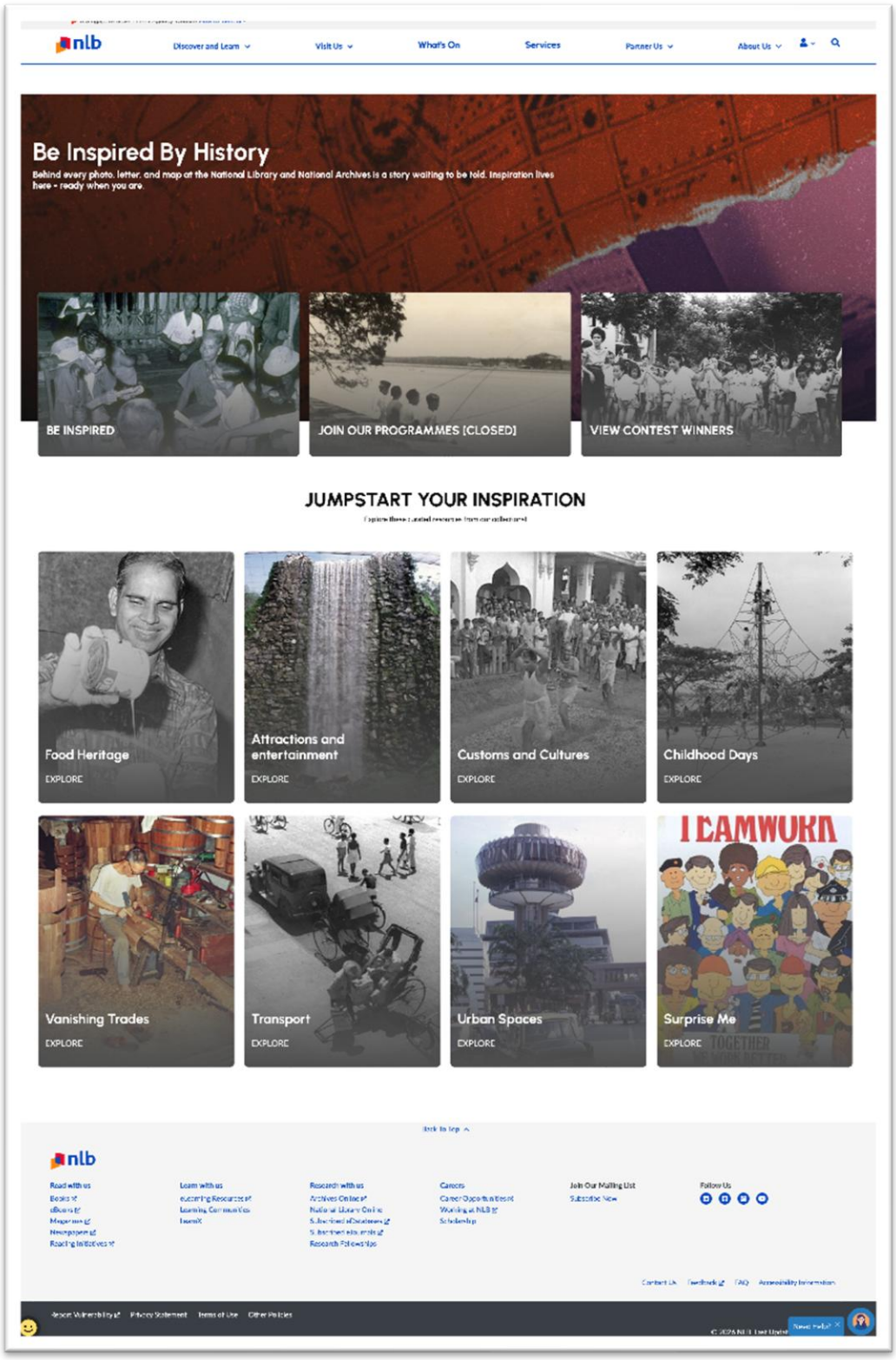


Fig. 1. *'Be Inspired by History' curated page, courtesy of NLB*

To generate publicity for the campaign and contest, creation guides were produced to help contest participants get ideas on how to get started. Publicity materials in the form of electronic banners were also disseminated. Additionally, a

campaign video and profile videos featuring creators from different fields sharing their stories on how they consulted NLB's resources for inspiration were produced.

The response was overwhelming. Despite organising a contest like this for the first time, approximately 100 entries were received by the deadline, out of which 35 were selected as winners, including consolation prize winners. The winning entry videos will be added to and accessioned into the NAS' collection, and made accessible on the Audio Visual Archives database on Archives Online.

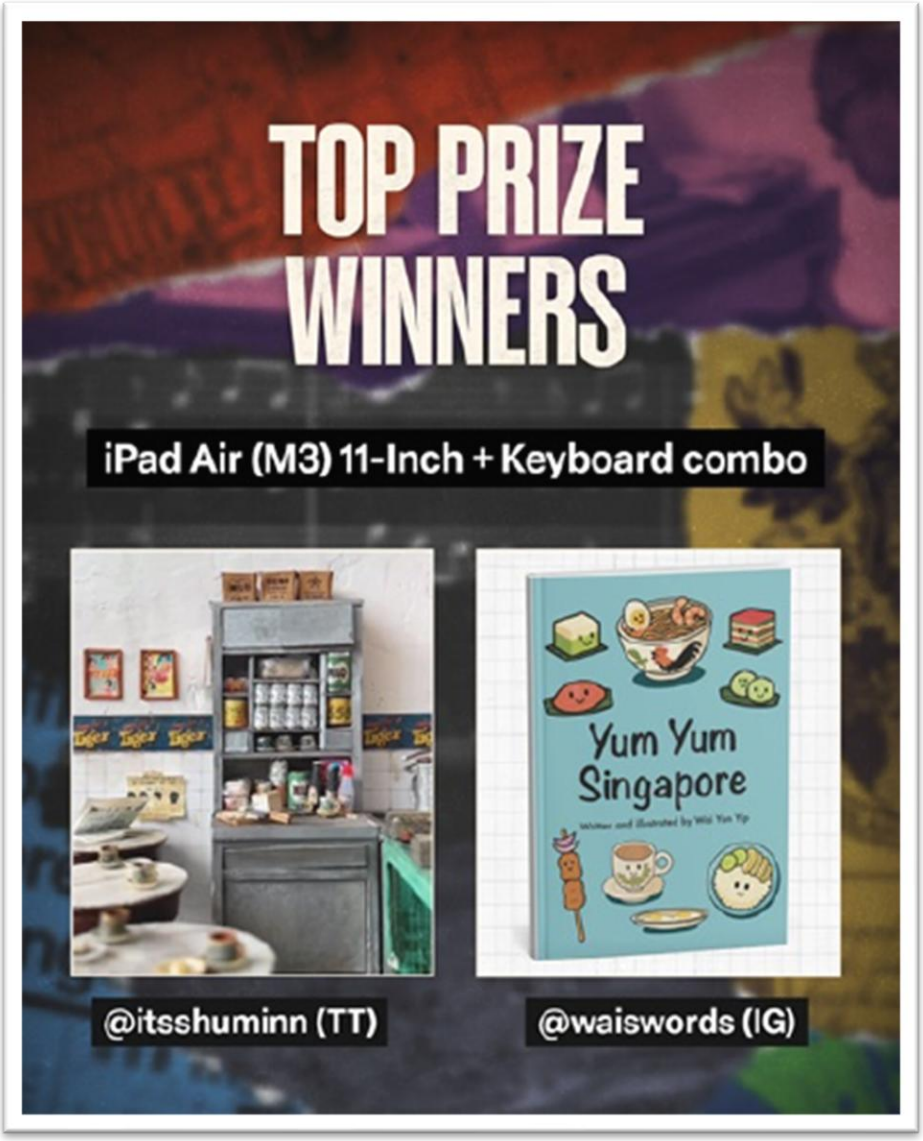


Fig. 2. *'Be Inspired by History' Top Prize Winners, courtesy of NLB*

The Audio Visual Archives department was formed in April 1995 'to appraise and acquire audiovisual records that are of national, historical and socio-cultural significance to Singapore, and documents such records that are transferred to NAS. The department also manages the prioritisation and digitisation of at-risk audiovisual formats, to ensure that the valuable contents in the moving images and

recorded sound heritage remain accessible to future generations' (About Us: Our Roles, n.d.).

Currently, the Audio Visual Archives holds about 390,000 audiovisual and sound recordings (189,000 hours) covering a wide range of genres, from current affairs and documentaries to infotainment and entertainment (About Audiovisual Records, n.d.). Among the over 3,500 shellac and vinyl records preserved at the NAS are the very first commercial sound recordings produced locally in 1903.

Other genres preserved by the Audio Visual Archives include government productions, radio and TV archives, home movies and more. Berita Singapura (or *News Singapore*) is an example of a government production. Produced by the Ministry of Culture in the 1960s and shot on 16-mm and 35-mm reels, these films documented the economic and socio-cultural developments of Singapore in that decade. Produced in Singapore's four official languages (English, Mandarin, Malay and Tamil), as well as in two Chinese dialects (Hokkien and Cantonese), these clips were screened on television and in public spaces such as community centres and rural villages as the government's way of disseminating information to the public (Chen 2019: 106).

As a valuable research and public resource for AV materials in Singapore (About Audiovisual Records, n.d.), the winning videos from the "Be Inspired by History" contest will not only enrich the existing AV materials held by the NAS, but will also be preserved for posterity by the NAS' Sound and Moving Image Laboratory, which is responsible for the conservation, digitisation, and preservation of audiovisual recordings in NAS' holdings (Conservation and Preservation: Sound and Moving Image Laboratory, n.d.).

Conclusion

The National Archives of Singapore has come a long way since its inception in 1968. As the repository for the nation's corporate and social memory, the NAS plays an important role in preserving, conserving and safeguarding our country's archival treasures. However, it cannot tell the nation's story if its records are inaccessible, kept in archival boxes and stored away from the public eye. It is only through outreach and education efforts that the country's heritage, culture and history can be made known. Thanks to technology — and increasingly, artificial intelligence — the NAS has come up with interactive and innovative ways to reach out to its audience. Will traditional means of public engagement go away? Probably not, judging from the number of exhibitions put on by other heritage institutions such as museums and galleries every year. As the National Archives celebrates its 60th anniversary in two years' time, we can certainly look forward to more exciting and engaging initiatives and programmes from the archives as it marks another milestone in its history.

References

Chen 2019 – Chen J. "A Film Clip in 1965." In *50 Records from History: Highlights from the National Archives of Singapore*, ed. by Chee V., Lau J., Zhuang J., Ng S. Singapore: National Archives of Singapore, 2019. P. 106.

Jayakumar 2019 – Jayakumar J. Looking Back, Looking Forward. *BiblioAsia*. 2019. Vol. 15, Issue 1. P. 3.

Tan 2019 – Tan F. Pioneers of the Archives. *BiblioAsia*. 2019. Vol. 15, Issue 1. P. 18.

Electronic Resource

About Audiovisual Records n.d. – About Audiovisual Records. National Archives of Singapore. https://www.nas.gov.sg/archivesonline/audiovisual_records/about-us (accessed 20.05.2026).

About NHB n.d. – History and Milestones. National Heritage Board. https://www.nhb.gov.sg/who-we-are/about-us/history-and-milestones (accessed 18.05.2026).

About NLB n.d. – About NLB. National Library Board. https://www.nlb.gov.sg/main/about-us (date of access: 18.05.2026).

About Singapore Digital n.d. – About Singapore Digital. IMDA. https://www.imda.gov.sg/about-imda/research-and-statistics/sgdigital (accessed 19.05.2026).

About Us: History n.d. – History. National Archives of Singapore. https://corporate.nas.gov.sg/about-us/history (accessed 18.05.2026).

About Us: Mandate n.d. – Mandate. National Archives of Singapore. https://corporate.nas.gov.sg/about-us/mandate/ (accessed 18.05.2026).

About Us: Our Roles n.d. – Our Roles. National Archives of Singapore. https://corporate.nas.gov.sg/about-us/our-roles/ (accessed 19.05.2026).

Conservation and Preservation n.d. – Sound and Moving Image Laboratory. National Archives of Singapore. https://corporate.nas.gov.sg/conservation-and-preservation/sound-and-moving-image-laboratory/ (accessed 22.05.2026).

Create your masterpiece 2025 – Create your masterpiece with rich resources from NLB's libraries and archives. National Library Board. https://www.nlb.gov.sg/main/about-us/press-room-and-publications/media-releases/2025/Be-Inspired-by-History (accessed 22.05.2026).

Discover the Story of Singapore's Road to Independence 2025 – Discover the Story of Singapore's Road to Independence Ticketing Now Open for The Albatross File: Singapore's Independence. National Library Board. https://www.nlb.gov.sg/main/about-us/press-room-and-publications/media-releases/2025/Albatross (accessed 21.05.2026).

National Library Board 2024 – National Library Board. *Discovering Singapore Stories Together: National Library Board Annual Report 2024/2025*. https://www.nlb.gov.sg/main/-/media/NLBMedia/Documents/About-Us/Press-Room-Publication/Annual-Reports/NLB-AR-24-25-Web.pdf (accessed 18.05.2026).

Public Programmes n.d. – Public Programmes. National Archives of Singapore. https://corporate.nas.gov.sg/outreach/public-programmes/ (accessed 21.05.2026).

The NLB Brand n.d. – The NLB Brand. National Library Board. https://www.nlb.gov.sg/main/about-us/The-NLB-Brand (accessed 18.05.2026).

ORAL FAMILY HISTORY: CAPTURING, PRESERVING, AND CREATING SEARCHABLE RECORDS

Jonathan McCollum
Content Manager for the Middle East and North Africa (MENA) Region, FamilySearch International, USA

ОТБАСЫНЫҢ АУЫЗША ТАРИХЫ: ЖАЗБАЛАРДЫ ЖИНАУ, САҚТАУ ЖӘНЕ ІЗДЕУГЕ ҚОЛАЙЛЫ ЖАҒДАЙ ЖАСАУ

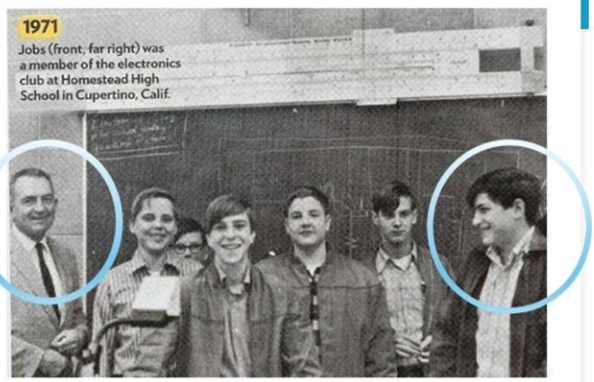
Джонатан МакКоллум
FamilySearch International компаниясының Таяу Шығыс және Солтүстік Африка (MENA) өңірі бойынша контент-менеджері, АҚШ

УСТНАЯ СЕМЕЙНАЯ ИСТОРИЯ: СБОР, СОХРАНЕНИЕ И СОЗДАНИЕ ДОСТУПНЫХ ДЛЯ ПОИСКА ЗАПИСЕЙ

Джонатан МакКоллум
Контент-менеджер по региону Ближнего Востока и Северной Африки (MENA), FamilySearch International, США

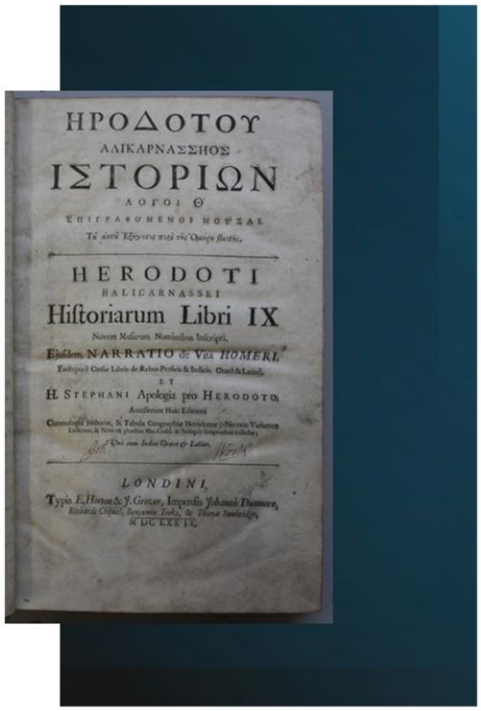
History and Technology: Oil and Water?

- Does the “Apple” fall far from the tree?
- What I love about history:
  - The Humanity
  - The Messiness
  - The Past
- But what does history mean?



Is history stuck in the past?

- The beginnings of “history”
  - From the Greek word *Ἱστορίαι* which is the title Herodotus of Halicarnassus gave his “investigations” or “inquiries” into the causes of the Greco-Persian Wars.
  - Ἱστορίαι vs. تاريخ (тарих)
    - יֶרַח
- Leopold von Ranke and the Professionalization of History
  - “*Wie es eigentlich gewesen ist*”
- But oral sources have always been part of history
  - Herodotus, Jules Michelet, Pre-Columbian America, American Slavery, Genocides, independence movements

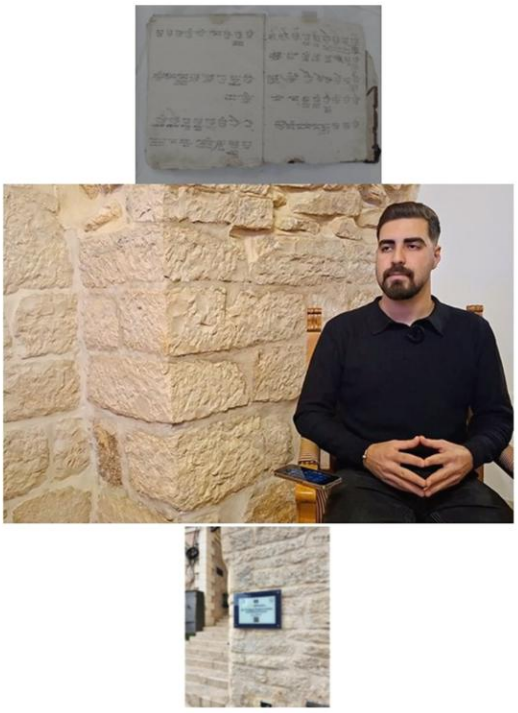


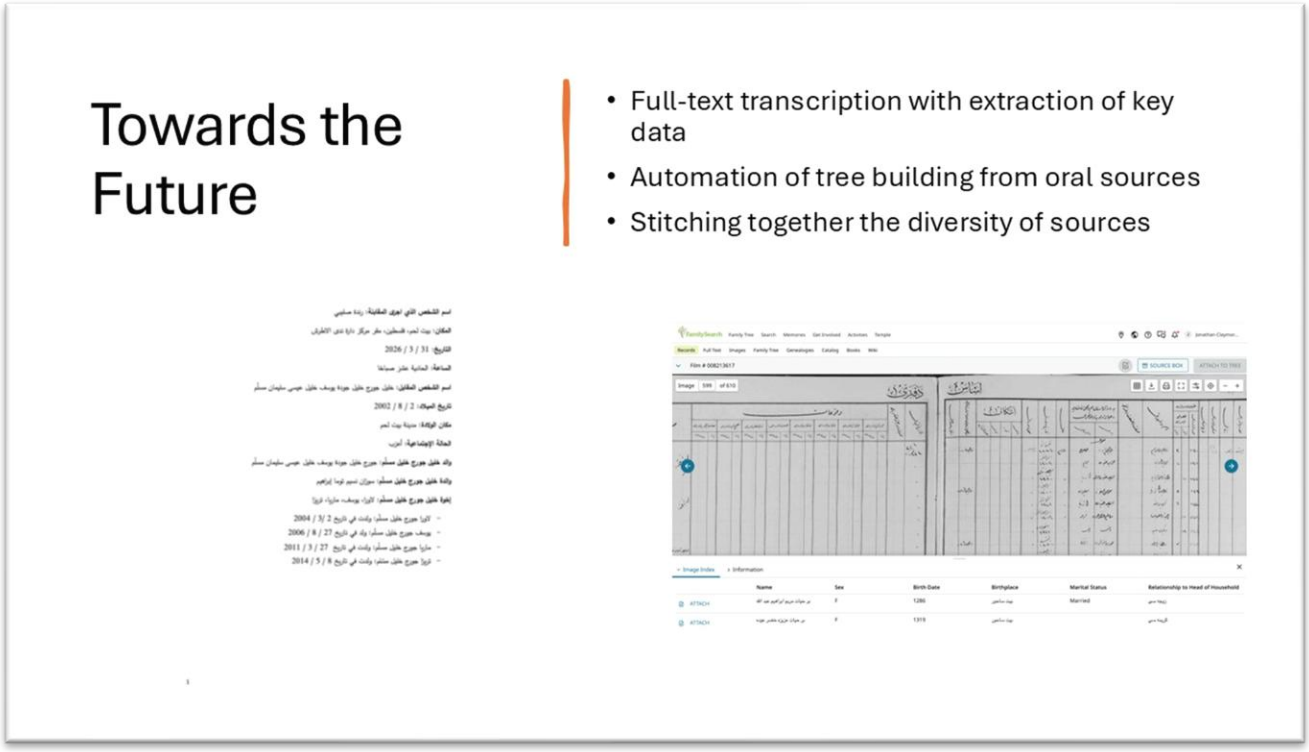
Oral History at FamilySearch

- Over 2.8 million interviews conducted
- 28 countries
- Active in 17 countries

Bethlehem: A Test Case

- Trained professionals
- Diversity of resources
- A tradition of oral genealogy
- A need for capturing the present




Towards the Future
• Full-text transcription with extraction of key data
• Automation of tree building from oral sources
• Stitching together the diversity of sources

FUNCTIONALITIES AND PROSPECTS OF MODERN ARCHIVAL DATABASES BASED ON THE PROJECTS OF THE WITOLD PILECKI INSTITUTE OF SOLIDARITY AND VALOR

Anna Olesińska
BA, Specialist in arhcive, Witold Pilecki Institute of Solidarity and Valor, Poland
a.olesinska@instytutpileckiego.pl
Paweł Lesisz
PhD, Chief specialist in archive, Witold Pilecki Institute of Solidarity and Valor, Poland
p.lesisz@instytutpileckiego.pl

Abstract. The development of digital archival ecosystems has significantly transformed the ways historical sources are collected, processed and made accessible, while also introducing new new technological and methodological challenges. This paper examines the Digital Archive of the Pilecki Institute and the Chronicles of Terror portal as elements of a broader digital archival infrastructure. The analysis is based on a case study of the collection Exile in Central Asia, created from documents obtained from the Hoover Institution. The paper demonstrates how digital archival systems function not only as repositories of digitised materials but also as collaborative infrastructures supporting research, accessibilty and long-term preservation. It also discusses the processes of digitisation, metadata creation and use of OCR technologies, as well as future perspectives, including the application of artificial intelligence tools in archival practice.

Keywords: digital archives, archival infrastructure, digital humanities, metadata, OCR

ЗАМАНАУИ АРХИВТІК ДЕРЕККӨЗДЕРДІҢ ФУНКЦИОНАЛДЫҚ МҮМКІНДІКТЕРІ МЕН ПЕРСПЕКТИВАЛАРЫ: ПИЛЕЦКИЙ ИНСТИТУТЫ ЖОБАЛАРЫНЫҢ МЫСАЛЫНДА

Анна Олесиньска
Бакалавр, архив ісі бойынша маман, Витольд Пилецкий атындағы ынтымақтастық және ерлік институты, Польша
a.olesinska@instytutpileckiego.pl
Павел Лесиш
PhD, архив ісі бойынша бас маман, Витольд Пилецкий атындағы ынтымақтастық және ерлік институты, Польша
p.lesisz@instytutpileckiego.pl

Аңдатпа. Цифрлық архив экосистемаларының дамуы тарихи дереккөздерді жинау, өңдеу және оларға қолжетімділікті қамтамасыз ету тәсілдерін айтарлықтай өзгертті, сонымен қатар жаңа технологиялық және әдіснамалық сын-қатерлерді туындатты. Бұл мақалада Пилецкий институтының Цифрлық архиві және «Террор хроникасы» порталы кең ауқымды цифрлық архив инфрақұрылымының элементтері ретінде қарастырылады. Талдау Гувер институтынан алынған құжаттар негізінде жасалған «Орталық Азиядағы жер аудару» коллекциясының мысалында жүргізіледі. Мақалада цифрлық архив жүйелерінің жай ғана цифрланған материалдар қоймасы емес, сонымен қатар зерттеулерді, қолжетімділікті және ұзақ мерзімді сақтауды қолдайтын ынтымақтастық инфрақұрылымы ретінде қалай қызмет ететіні көрсетіледі. Сондай-ақ цифрлау, метадеректерді жасау және OCR технологияларын қолдану процестері, сондай-ақ архив тәжірибесінде жасанды интеллект құралдарын қолдануды қоса алғандағы болашақ перспективалар талқыланады.

Түйін сөздер: цифрлық архивтер, архив инфрақұрылымы, цифрлық гуманитарлық ғылымдар (digital humanities), метадеректер, OCR

ФУНКЦИОНАЛЬНЫЕ ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ СОВРЕМЕННЫХ АРХИВНЫХ БАЗ ДАННЫХ: НА ПРИМЕРЕ ПРОЕКТОВ ИНСТИТУТА СОЛИДАРНОСТИ И ДОБЛЕСТИ ИМЕНИ ВИТОЛЬДА ПИЛЕЦКОГО

Анна Олесиньска
Бакалавр, специалист по архивному делу,
Институт солидарности и мужества им. Витольда Пилецкого, Польша
a.olesinska@instytutpileckiego.pl
Павел Лесиш
PhD, главный специалист по архивному делу,
Институт солидарности и мужества им. Витольда Пилецкого, Польша
p.lesisz@instytutpileckiego.pl

Аннотация. Развитие цифровых архивных экосистем существенно трансформировало способы сбора, обработки и обеспечения доступа к историческим источникам, одновременно породив новые технологические и методологические вызовы. В данной статье рассматриваются Цифровой архив Института Пилецкого и портал «Хроники террора» как элементы более широкой цифровой архивной инфраструктуры. Анализ базируется на кейсе коллекции «Ссылка в Центральную Азию», созданной на основе документов, полученных из Гуверовского института. В статье демонстрируется, как цифровые архивные системы функционируют не только как хранилища оцифрованных материалов, но и как инфраструктуры сотрудничества, поддерживающие исследования, доступность и долгосрочное хранение. Также обсуждаются процессы оцифровки, создания метаданных и использования технологий OCR и будущие перспективы, включая применение инструментов искусственного интеллекта в архивной практике.

Ключевые слова: цифровые архивы, архивная инфраструктура, цифровые гуманитарные науки (digital humanities), метаданные, OCR

The development of digital archives in recent years has significantly transformed the ways in which historical sources are collected, processed and made accessible (Smołucha 2021; Pawłowski 2023). At the same time, it has introduced new technological and methodological challenges for archival institutions. These include the digitisation of dispersed collections, the creation of precise metadata and search systems, and ensuring the long-term preservation and sustainability of digital resources.

This paper argues that modern archival databases should be understood not only as repositories of digitised materials but also as elements of broader digital ecosystems that enable institutional collaboration, knowledge production and public access to historical sources (Radomski et al. 2013).

Digital Archival Infrastructure of the Pilecki Institute

The Digital Archive of the Pilecki Institute functions as a platform integrating copies of documents obtained from Polish and international institutions, including the Bundesarchiv, United Nations archives and the Hoover Institution (Stępniak 1997). By

bringing together dispersed materials within a single system, the archive represents an emerging archival hub based on cooperation between institutions.

The collections include textual documents, photographs and audiovisual materials. Their integration within one platform allows researchers to access sources remotely, significantly reducing the need for time-consuming archival queries.

The Chronicles of Terror portal complements this infrastructure by publishing testimonies of victims and witnesses of totalitarian regimes, including accounts of civilians from occupied Poland and testimonies of individuals deported to the Soviet Union. Together, these projects create a digital environment that supports both academic research and the broader dissemination of historical knowledge.

Case Study: "Exile in Central Asia" Collection

The analysis presented in this paper is based on a collection created from documents obtained from the Hoover Institution entitled Exile in Central Asia. The materials concern the experiences of Polish deportees sent after 1939 to Kazakhstan and other regions of Central Asia.

A significant part of this collection consists of documents produced as part of post-war documentation initiatives, including survey campaigns aimed at collecting testimonies from former deportees. These materials include personal accounts, questionnaires and reports describing the conditions of deportation, everyday life in exile and the processes of adaptation and survival in extreme circumstances. Such sources provide valuable insight into individual experiences while also enabling broader comparative analysis.

The archival materials preserved in the Hoover Institution constitute an important and, in many cases, unique body of sources documenting the fate of Polish citizens deported to the Soviet Union. Their digitisation and integration into a digital archival

system make them accessible to researchers who would otherwise face significant barriers in conducting archival queries abroad.

The use of this collection as a case study reflects a digital humanities approach (Warwick et al. 2012). It allows for a detailed examination of the processes involved in creating and managing digital archival resources and demonstrates how dispersed archival materials can be integrated into a coherent digital system, enabling new forms of research and interpretation.

Digitisation Process and Metadata Management

The process preceding the publication of materials in the digital archive includes several stages. These involve archival research, the acquisition of digital copies through licensing agreements, OCR processing, and the introduction of metadata into the system.

The dLibra platform, widely used by institutions in Poland, serves as the technological backbone for managing digital collections. It enables the organisation, description and dissemination of archival materials in a structured and accessible way (Matysek et al. Tomaszczyk 2020).

In addition, the implementation of geolocation data allows for the spatial visualisation of historical information. This enhances the usability of the archive and introduces new possibilities for analysing historical sources.

Challenges and Future Perspectives

Modern digital archives face a number of challenges, including the digitisation of large and dispersed collections, the creation of accurate metadata, and ensuring long-term data preservation. Legal issues related to access and copyright also remain significant.

At the same time, new technologies create opportunities for further development. The use of artificial intelligence tools, such as Translab, can support OCR

processes and improve the efficiency of working with large volumes of archival materials. These developments indicate that digital archives are evolving into complex infrastructures that combine technological solutions with institutional collaboration and scholarly practice (Kuklo et al. 2024).

Conclusion

The analysis demonstrates that digital archival systems should be understood as more than repositories of digitised sources. They constitute dynamic and collaborative infrastructures that enable new forms of access, research and interpretation.

The case of the Pilecki Institute's digital projects shows how the integration of dispersed materials, the application of digital tools and cooperation between institutions contribute to the development of sustainable archival ecosystems. In this context, digital archives play a crucial role not only in preserving historical sources but also in shaping the future of historical research.

References

Kuklo et al. 2024 – Kuklo C., Walczak W. (eds.). Człowiek twórcą historii. Warsztat nowoczesnego humanisty historyka na progu XXI wieku., vol. 5–6. Białystok, 2024.

Matysek et al. 2020 – Matysek A., Tomaszczyk J. Cyfrowy warsztat humanisty. Warszawa, 2020.

Pawłowski 2023 – Pawłowski A. Od Gutenberga do Zuckerberga. Wstęp do humanistyki cyfrowej. Kraków, 2023.

Radomski et al. 2013 – Radomski A., Bomba R. (eds.). Zwrot cyfrowy w humanistyce. Lublin, 2013.

Smołucha 2021 – Smołucha D. Humanistyka cyfrowa w badaniach kulturowych. Kraków, 2021.

Stępniak 1997 – Stępniak W. Archiwalia polskie w zbiorach Instytutu Hoovera Uniwersytetu Stanforda. Warszawa, 1997.

Warwick et al. 2012 – Warwick C., Terras M., Nyhan J. (eds.). Digital Humanities in Practice. London, 2012.

РОЛЬ АРХИВОВ В СОХРАНЕНИИ ПРЕЕМСТВЕННОСТИ ГОСУДАРСТВЕННОЙ ВЛАСТИ И ИСТОРИЧЕСКОЙ ИДЕНТИЧНОСТИ

Куляйша Кабдуалиевна Актаева
Кандидат исторических наук, директор Музея Первого Президента, Казахстан
aktayeva.k@gmail.com

Аннотация. В докладе рассматривается историко-мемориальный ресурс архивных учреждений в формировании исторической идентичности современных обществ, переживающих периоды критического переосмысления исторического опыта предшествующих поколений. Категории объективности и беспристрастности, которые традиционно ассоциируются с трактовкой роли архивов в накоплении элементов наследия, испытывают наибольшее негативное давление. Утилитарные и стратегические функции архивов вступают в открытые противоречия с условиями их институционального развития в контексте вызовов и рисков информационной эры. Современные тренды медиапотребления диктуют необходимость оперативного управления скоростью и качеством распространения исторической информации, обладателями которой в современных технологических условиях выступают не только архивы, но и электронные средства информации, облачные системы хранилищ, частные лица – иными словами, понятия подлинности и полноты исторических свидетельств и просто архивных данных весьма условны. В век господства не только информационных, но и политических технологий вопросы преемственности государственной власти и исторической идентичности вступают в критическую стадию.

Ключевые слова: институты памяти; историческое сознание; историческая идентичность; медиапотребление; архивные данные; цифровой архив; информационные войны

МЕМЛЕКЕТТІК БИЛІКТІҢ САБАҚТАСТЫҒЫ МЕН ТАРИХИ БІРЕГЕЙЛІКТІ САҚТАУДАҒЫ АРХИВТЕРДІҢ РӨЛІ

Күләйша Қабдуәлиқызы Ақтаева
Тарих ғылымдарының кандидаты,
Тұңғыш Президент музейінің директоры, Қазақстан
aktayeva.k@gmail.com

Аңдатпа. Баяндамада тарихи тәжірибені сын тұрғысынан қайта ой елегінен өткізіп жатқан қазіргі қоғамдардың тарихи бірегейлігін қалыптастырудағы архив мекемелерінің тарихи-мемориалдық ресурсы қарастырылады. Архивтердің мұра элементтерін жинақтаудағы рөлімен дәстүрлі түрде байланыстырылатын объективтілік және бейтараптылық санаттары қазіргі таңда ең қатты теріс қысымға ұшырауда. Архивтердің утилитарлық және стратегиялық функциялары ақпараттық дәуірдің сын-тегеуріндері мен тәуекелдері жағдайында олардың институционалдық даму шарттарымен ашық қайшылыққа түсуде. Қазіргі медиа-тұтыну трендтері тарихи ақпаратты таратудың жылдамдығы мен сапасын жедел басқару қажеттілігін талап етеді. Қазіргі технологиялық жағдайда мұндай ақпараттың иелері тек архивтер ғана емес, сонымен қатар электронды ақпарат құралдары, бұлтты сақтау жүйелері және жеке тұлғалар болып табылады – басқаша айтқанда, тарихи куәліктердің, тіпті қарапайым архив деректерінің түпнұсқалығы мен толықтығы туралы түсініктер өте шартты сипатқа ие. Тек ақпараттық қана емес, сонымен қатар саяси технологиялар үстемдік құрған ғасырда мемлекеттік биліктің сабақтастығы мен тарихи бірегейлікті сақтау мәселелері сын кезеңге еніп отыр.

Түйін сөздер: жад институттары; тарихи сана; тарихи бірегейлік; медиа-тұтыну; архив деректері; цифрлық архив; ақпараттық соғыстар

THE ROLE OF ARCHIVES IN PRESERVING THE CONTINUITY OF STATE AUTHORITY AND HISTORICAL IDENTITY

Kulaisha K. Aktayeva
Candidate of Historical Sciences,
Director of the Museum of the First President, Kazakhstan
aktayeva.k@gmail.com

Abstract. This paper examines the historical and memorial resources of archival institutions in shaping the historical identity of contemporary societies undergoing periods of critical reevaluation of the historical experience of previous generations. The greatest negative pressure now faces the categories of objectivity and impartiality, which are traditionally associated with the interpretation of archives' role in preserving heritage. The utilitarian and strategic functions of archives are increasingly at odds with the conditions of their institutional development, given the challenges and risks posed by the information age. Contemporary trends in media consumption dictate the need for operational management of the speed and quality of historical information dissemination, the custodians of which in today's technological environment include not only archives but also electronic media, cloud storage systems, and private individuals—in other words, the concepts of authenticity and completeness of historical evidence, and indeed of archival data itself, are highly relative. In an age dominated not only by information technologies but also by political technologies, issues of the continuity of state power and historical identity are entering a critical phase.

Keywords: memory institutions; historical consciousness; historical identity; media consumption; archival data; digital archive; information wars

Введение

III Конгресс архивистов проходит в год, определяющий вектор развития современной казахстанской государственности. В широком информационном поле обсуждаются вопросы, определяющие категории исторической памяти, основанной на историческом знании.

Эволюция идей конституционного строительства, зафиксированная в архивных массивах предшествующих десятилетий казахстанской независимости, актуализирует запрос на исторические контексты, формирующие наши знания и представления о конституции как юридической основе государственности.

Для архивной сферы и более широко – для сферы хранения и презентации историко-культурного наследия – стратегически важным представляется определение принципов развития институтов памяти в рамках общественного договора, каковым является суть государства как института, выполняющего функцию доверительного управления правами и обязанностями граждан.

Основной вызов современности, как показывает практика сегодняшнего дня, заключается в свободе манипулирования историческим знанием, ревизии многих его страниц.

Повестка дня взаимодействия архивов и общества, архивов и государства стремительно меняется. Профессиональное архивирование побеждает в битве за сохранность уже созданных документов, но начинает проигрывать битву за фиксацию самого процесса принятия решений, который все чаще скрывается от бюрократического учета (Бонч-Осмоловская 2024).

Основная часть

Архивы обеспечивают правовую непрерывность государства и формируют коллективную память народа. Без них невозможна легитимность власти и сохранение культурного кода. Динамика политических процессов, фиксируемая в архивных массивах, отражает все многообразие идей и посылов, направленных на совершенствование системы государственного управления.

В архивах общество гарантированно получает ответ на запросы по сохранению преемственности власти в виде оригиналов конституций, законов и международных договоров. Документы доказывают непрерывность государственных институтов при смене правительств. В вопросах исторической идентичности первоисточники служат главным аргументом для противоборствующих сторон. Национальный нарратив, выраженный в образовательных программах и музейных экспозициях, покоится на фундаменте архивных данных, накапливаемых из поколения в поколение и превращающих мифы в доказанную историю. Где в этой цепочке осмысленных, веками отработанных алгоритмов пополнения исторического знания происходит смысловой разрыв, позволяющий подвергать сомнению подлинность исторических сведений?

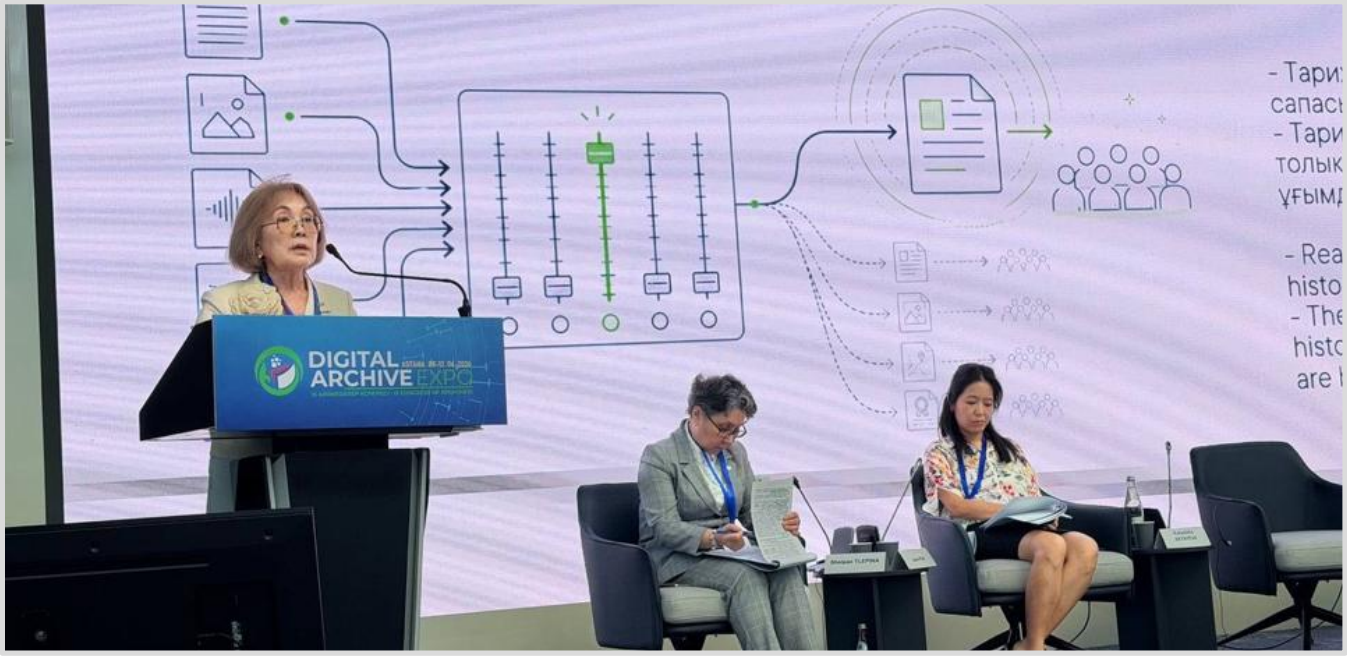
Решая технологические аспекты архивирования истории, современные общества сталкиваются с проблемой того, что архив как документальный

комплекс стал так многолик и многообразен, что трудно сказать, насколько полными и достоверными являются материалы, поступающие на архивное хранение. «Что написано пером, не вырубишь топором» – эта сентенция перестает быть очевидной по известным причинам.

Объективная реальность, написанная «широким взмахом кисти», которая оказывается в наших высокотехнологичных руках, скрывает другую проблему. На ней и хотелось бы остановиться. Динамика современных политических процессов развивается в активном информационном поле, которое неуправляемо в привычных параметрах административного воздействия

Практически каждый день мы слышим о том, как с помощью современных технологий поменяется наше будущее. Но, кажется, не меньшее влияние эти технологии будут иметь на прошлое – вернее на то, каким образом мы будем обращаться из этого будущего к нашему уже совершенно цифровому прошлому.

Относительно краткий опыт человечества в создании и накоплении цифрового архива актуализировал новые запросы и ожидания. От осознания безграничности возможностей цифрового архива к пониманию критически важных функций: 1) оперативного хранения цифрового документа и 2) обеспечения неизменности и аутентичности в процессе решения задачи освоения огромных потоков данных.



Мы преодолели пространства и границы. И это не география. Пространства и границы – в приобщении к культурно-историческому наследию стран и народов, невзирая на географическое, идеологическое, политическое измерение, соотношение глобального и национального, аутентичного и синтезированного. И при этом система сдержек и противовесов в распределении указанных характеристик основывается только на одной категории – коллективной памяти.

Феномен коллективной памяти, как известно, заключается в способности людей хранить ее и передавать последующим поколениям (Рамси 2022). Это

когда-то породило и архивы, в том числе как институты памяти, наряду с музеями и библиотеками. На протяжении нескольких десятилетий с момента технологической революции 70-х годов XX века человечество все более взрослеет в информационной среде, формируемой технологическими ресурсами, порожденными им самим в стремлении усовершенствовать практику хранения и продвижения своей коллективной памяти.

«Мы ежесекундно создаем огромный цифровой архив человечества, но при этом совершенно непонятно, куда это все денется дальше и сложится ли из этого действительно коллективная память – та самая, без которой невозможно будет представить наше будущее» (Бонч-Осмоловская 2024).

И это только один аспект осознания рисков, сопровождающих деятельность институтов памяти. Нормативное и организационное деление функций архивов, музеев и библиотек в части хранения коллективной памяти в современных условиях активного общественного интереса к прошлому с позиций настоящего все больше становится условным. Работники этих сфер не являются более «стражами», ревностно не допускающими к сокровищам своих хранилищ посторонних «непосвященных». Профессиональная компетенция, в отличие от любопытствующего специалиста смежной отрасли или активного дилетанта, должна предполагать компетенцию архивиста, музейного хранителя, библиотекаря – не собирателя, а куратора всех данных, регулирующего как процессы отбора, так и алгоритмы распределения и презентации фондовых собраний.

Риски неизбежно сопровождают современную цифровую эпоху, порождающую архивы общественного, частного, государственного происхождения изначально в цифровом формате «born digital» (рожденные цифровыми). Они никогда не существовали в материальном виде, а сразу уже были созданы «в цифре» – тексты, написанные на компьютере, письма, отправленные по электронной почте, электронные фотографии, аудио и видеозаписи, веб-сайты, датасеты и базы данных.

Общественное и государственное, частное и коллективное – где граница, которую мы не переходим, приобщаясь к новой информации о своем прошлом, настоящем и грядущем? Фейки стали неотъемлемой частью информационной картины современного человека независимо от его возраста, образования, рода занятий, вероисповедания, политической принадлежности и пр.

Кураторство таких данных просто необходимо и формируется автоматически при систематизации цифровых массивов. Критерии такой систематизации, как правило, формируют те, в чьих руках весь массив, и конечная цель такой систематизации может быть обусловлена разными задачами, возможностями распределения и хранения этой информации.

При этом нельзя забывать, что фундаментальный принцип классического архивного дела состоит в том, что архивист не должен упорядочивать коллекции самостоятельно, а должен сохранять ту организацию архива, которая была заложена в него изначально. Нельзя навязывать свое индивидуальное видение в условиях постоянно растущих фондов – миссия архивиста состоит в сохранении и передаче объектов памяти, а не в их интерпретации.

Сохранение же документов born digital требует ровно обратного: без категоризации, разметки документов и создания тематических подборок по множеству разнообразных параметров цифровой архив вообще нельзя считать сохраненным. Такой архив невозможно использовать, потому что документы в нем не будут найдены. Современные цифровые архивисты являются активными акторами процесса создания коллекций, кураторами, а не хранителями цифровых архивов. Соответственно, их роль значительно возрастает – можно сказать, что они обладают определенной властью над конечной формой архива, их субъективное представление и профессиональная оптика приобретают особенное значение. Это ставит дополнительные этические и профессиональные вопросы о работе архивистов-кураторов, возникающие как на глобальном, так и национальном уровнях.

Последствия ковида2019 активно обсуждались на II Конгрессе, в очередной раз можно упомянуть его в данном контексте. Ковид стал не только коллективной травмой, но и общечеловеческим коллективным опытом выстраивания коллективной памяти. Он дал толчок еще большей цифровизации всех аспектов общественной жизни. Локдаун стимулировал развитие каналов трансляции частного опыта отдельного человека, небольших сообществ, локальных объединений, которые ранее были бы обречены на забвение. В целом феномен ковида заключается в цивилизационной практике активного конструирования памяти в частных и глобальных измерениях и масштабах.

Обращаясь к казахстанскому опыту последних 5-6 лет, нельзя не упомянуть *кейс Alatau City*. В мае 2026 года Президент Казахстана подписал Конституционный закон «О специальном правовом режиме города Алатау». Закон вводит революционную норму, напрямую касающуюся архивной сферы – любые административные процедуры или акты недействительны, если они не размещены изначально в открытом цифровом формате в городском репозитории. На практике это означает простое правило: если решения нет в открытом цифровом репозитории, юридически его не существует.

При всей необычности проекта сама логика особого режима для Казахстана не нова. Страна уже не раз пыталась создавать внутри старой системы отдельные «островки эффективности» – пространства, где можно быстрее менять правила, тестировать новые подходы и выращивать управленческие практики, которые потом могут повлиять на всю систему. В качестве примеров можно привести НИШ, Назарбаев Университет, Air Astana и МФЦА – каждый из этих проектов по-своему выходил из привычной административной логики. Но все бюрократические процедуры касательно архивирования инициативы соблюдались в традиционном формате: вначале бумажный формат, затем цифровой. К 2050 году, когда предполагается увидеть реальный Alatau City в завершенном виде, казахстанские архивы будут оперировать born digital как секторальным навыком, формирующим историческое знание без привычных бумажных массивов. Это не прогноз, а заложенная в Конституционном законе основа преемственности исторического опыта, преемственности принципов государственного управления.

Не вызывает сомнений, что именно способность к коллективной памяти о прошлом неразрывным образом связана с нашим пониманием будущего и

умением создавать это будущее. Кейс Alatau City, развивая идеи, заложенные в первых амбициозных проектах независимого Казахстана, формирует траекторию того, что нас ждет в нашем цифровом будущем. Но неизбежны сомнения, насколько нам удастся создать цифровую коллективную память.

Важнейшим эффектом перехода современных коммуникаций в цифровой формат стало стирание границ между общественной и частной информацией. Поисковые системы не делают различий между государственным постановлением или частным высказыванием в блоге. На каком месте в иерархии ценностей информации должны быть те или иные источники информации определяет сам потребитель. Современные вызовы для национального нарратива формируются в ходе информационных войн, на фоне которых происходит децентрализация памяти. Во взаимодействии архивов и общества наступила эра
«мини-нарратива», в которой доминирует история, создаваемая человеком или локальным сообществом. Диалектика развития этих взаимодействий открыла, наряду со сложностями, новые возможности, горизонты развития для архивов.

Вместо борьбы с децентрализацией памяти передовые архивные системы используют стратегию интеграции и партнерства, превращая общество из «клиента» в «соавтора». На площадках нынешнего Конгресса прозвучало много вдохновляющих примеров краудсорсинга, архивации веб-пространства, внедрения ИИ для умного поиска. Собственно, сама тематика, заявленная в Конгрессе, говорит о готовности архивных учреждений артикулировать новые идеи, способствующие закреплению их роли в преемственности государственности и исторической идентичности.

Заключение

Уровень развития архивного дела влияет на процесс исторического познания, это, в свою очередь, обеспечивает повышение общественного сознания и способствует самоидентификации современного социума. Государство может потерять богатство, казну, как это часто бывает в истории. Но все это со временем можно восполнить и восстановить. Но невозможно восстановить потерянные архивные материалы.

Основной урок, извлеченный архивами из практики цифровой эпохи в том, чтобы архив перестал быть закрытым хранилищем и стал открытой платформой-верификатором, которая предоставляет гражданам и специалистам разных сфер качественную информацию для создания исторически достоверных нарративов.

Проблемы, актуальные сегодня: нужно ли нам сохранять гигабайты информации, сгенерированные искусственным интеллектом, имеет ли это какое-то отношение к миссии институтов памяти и как отличить документ, представляющий историческую ценность, от документа, созданного моделью GPT следующих поколений?

Литература

Бонч-Осмоловская 2024 – Бонч-Осмоловская А. Проблема цифрового архива: почему информацию надо спасать / А. Бонч-Осмоловская // Системный блокъ. 2024. 4 июня. URL: sysblok.ru (дата обращения: 20.05.2026).

Гинатулина 2022 – Гинатулина О.А. Социальное бытие архива в свете вызовов современного общества / О. А. Гинатулина // Философская мысль. 2022. № 9. С. 18–24. URL: cyberleninka.ru (дата обращения: 25.05.2026).

Паркова – Паркова Н.А. Значение архивов в жизни современного общества. [Электронный ресурс]. Н.А. Паркова // Материалы научных конференций ТГУ им. Г.Р. Державина. URL: tsutmb.ru (дата обращения: 21.05.2026).

Рамси 2022 – Рамси Э.С. Зачем нам помнить? Из истории коллективной памяти / Эбби Смит Рамси // Как читать медиа: [онлайн-проект] / Гете-Институт. 2022. Июнь. URL: goethe.de (дата обращения: 25.05.2026).

АРХИВЫ КАК ОБЩЕСТВЕННЫЙ ИНСТИТУТ: ДОСТУП К ПАМЯТИ, ЗНАНИЯМ И ИДЕНТИЧНОСТИ

Сауле Карибаевна Удербаева
Кандидат исторических наук, ассоциированный профессор,
и.о. профессора кафедры истории Казахстана,
Казахский национальный университет им. аль-Фараби, Казахстан
saule-uderbaeva@mail.ru; https://orcid.org/0000-0002-0322-964X

Аннотация. В статье рассматривается становление и трансформация архивного дела в Казахстане в контексте изменения роли архивов как общественного института. Целью исследования является анализ эволюции функций архивов – от инструментов административного управления к институтам, обеспечивающим доступ к исторической памяти, знаниям и формированию общественной идентичности.

Особое внимание уделено деятельности Оренбургской ученой архивной комиссии как одному из ключевых этапов формирования архивной практики в регионе. Показано, что комиссия не только осуществляла сбор и хранение документов, но и активно вводила их в научный оборот, способствуя развитию исторических исследований и формированию принципов систематизации и описания архивных материалов. Подчеркивается ее роль в расширении источниковой базы за счет привлечения этнографических, статистических и фольклорных материалов.

Анализируется изменение статуса архивов в советский период, характеризующийся усилением государственного контроля над информацией и ограничением доступа к документам. На примере Переписи населения СССР 1937 года показано, каким образом архивы использовались как инструмент политического регулирования исторической памяти. Отмечается, что в этот период архивы в значительной степени утратили функции открытого общественного института.

Рассматриваются процессы реформирования архивной системы в постсоветский период, включая принятие нормативно-правовой базы и постепенную демократизацию доступа к архивным материалам. Особое значение придается современным

тенденциям цифровизации архивов, развитию электронных платформ и расширению доступа к документам для исследователей и широкой общественности.

В заключение делается вывод о том, что архивы в современном Казахстане выступают важным элементом общественной инфраструктуры, способствующим сохранению исторического наследия, развитию научного знания и формированию национальной и культурной идентичности.

Ключевые слова: архивоведение; архивы; Казахстан; архивные институты; историческая память; доступ к информации; советский период; архивная политика; Оренбургская ученая архивная комиссия; цифровизация; культурное наследие

АРХИВТЕР ҚОҒАМДЫҚ ИНСТИТУТ РЕТІНДЕ: ЖАДҚА, БІЛІМГЕ ЖӘНЕ СӘЙКЕСТІККЕ ҚОЛЖЕТІМДІЛІК

Сәуле Кәрібайқызы Үдербаева
Тарих ғылымдарының кандидаты, қауымдастырылған профессор,
Қазақстан тарихы кафедрасының профессор міндетін атқарушы,
әл-Фараби ат. Қазақ ұлттық университеті, Қазақстан
saule-uderbaeva@mail.ru; https://orcid.org/0000-0002-0322-964X

Аңдатпа. Мақалада Қазақстандағы архив ісінің қалыптасуы мен трансформациясы архивтердің қоғамдық институт ретіндегі рөлінің өзгеруі контекстінде қарастырылады. Зерттеудің мақсаты – архив функцияларының эволюциясына талдау жасау: олардың әкімшілік басқару құралынан тарихи жадқа, білімге қолжетімділікті қамтамасыз ететін және қоғамдық бірегейлікті қалыптастыратын институттарға айналу жолын көрсету.

Орынбор ғылыми архив комиссиясының қызметіне өңірдегі архив тәжірибесін қалыптастырудың негізгі кезеңдерінің бірі ретінде ерекше назар аударылады. Комиссия құжаттарды жинау және сақтаумен қатар, оларды ғылыми айналымға белсенді енгізіп, тарихи зерттеулердің дамуына, архив материалдарын жүйелеу мен сипаттау қағидаларының қалыптасуына ықпал еткені көрсетілген. Оның этнографиялық, статистикалық және фольклорлық материалдарды тарту арқылы деректік базаны кеңейтудегі рөлі атап өтіледі.

Кеңес кезеңіндегі архив мәртебесінің өзгеруі, ақпаратқа мемлекеттік бақылаудың күшеюі және құжаттарға қолжетімділіктің шектелуі талданады. КСРО халық санағының 1937 жылғы деректері мысалында архивтердің тарихи жадты саяси реттеу құралы ретінде қалай пайдаланылғаны көрсетілген. Осы кезеңде архивтердің ашық қоғамдық институт функциясын айтарлықтай жоғалтқаны ерекше атап өтіледі.

Посткеңестік кезеңдегі архив жүйесін реформалау процестері, соның ішінде нормативтік-құқықтық базаның қабылдануы және архив материалдарына қолжетімділіктің біртіндеп демократияландыруы қарастырылады. Архивтерді цифрландырудың қазіргі үрдістеріне, электронды платформалардың дамуына және зерттеушілер мен қалың жұртшылық үшін құжаттарға қолжетімділіктің кеңеюіне ерекше мән беріледі.

Қорытындылай келе, қазіргі Қазақстандағы архивтер тарихи мұраны сақтауға, ғылыми білімді дамытуға, ұлттық және мәдени бірегейлікті қалыптастыруға ықпал ететін қоғамдық инфрақұрылымның маңызды элементі ретінде танылады.

Түйін сөздер: архивтану; архивтер; Қазақстан; архив институттары; тарихи жад; ақпаратқа қолжетімділік; кеңес кезеңі; архив саясаты; Орынбор ғылыми архив комиссиясы; цифрландыру; мәдени мұра

ARCHIVES AS A SOCIAL INSTITUTION: ACCESS TO MEMORY, KNOWLEDGE AND IDENTITY

Saule Karibaevna Uderbayeva
Candidate of Historical Sciences, Associate Professor, Acting Professor,
Department of History of Kazakhstan,
Al-Farabi Kazakh National University, Kazakhstan
saule-uderbaeva@mail.ru; https://orcid.org/0000-0002-0322-964X

Abstract. This article examines the formation and transformation of archival practice in Kazakhstan within the context of the evolving role of archives as a public institution. The aim of the study is to analyze the evolution of archival functions – from instruments of administrative governance to institutions that ensure access to historical memory, knowledge, and the formation of public identity.

Particular attention is paid to the activities of the Orenburg Scientific Archival Commission as a key stage in the development of archival practices in the region. It is demonstrated that the Commission not only collected and preserved documents but also actively introduced them into scholarly circulation, thereby contributing to the development of historical research and the establishment of principles for the systematization and description of archival materials. Its role in expanding the source base through the inclusion of ethnographic, statistical, and folkloric materials is also emphasized.

The article analyzes the transformation of the status of archives during the Soviet period, characterized by increased state control over information and restricted access to documents. Using the example of the 1937 Soviet Census, the study illustrates how archives were employed as instruments of political regulation of historical memory. It is argued that during this period archives largely lost their function as open public institutions.

The processes of reforming the archival system in the post-Soviet period are also considered, including the development of a legal framework and the gradual democratization of access to archival materials. Particular importance is attached to contemporary trends in the digitalization of archives, the development of electronic platforms, and the expansion of access to documents for researchers and the broader public.

In conclusion, the article argues that archives in modern Kazakhstan function as an essential component of public infrastructure, contributing to the preservation of historical heritage, the advancement of scholarly knowledge, and the formation of national and cultural identity.

Keywords: archival studies; archives; Kazakhstan; archival institutions; historical memory; public access; Soviet period; archival policy; Orenburg Scientific Archival Commission; digitalization; cultural heritage

Архивы в Казахстане, как и во всем мире, играют ключевую роль в сохранении исторической памяти. В условиях глобализации и трансформации исторических нарративов они выступают важным инструментом обеспечения доступа к знаниям и формирования общественной идентичности.

У истоков архивного дела в Казахстане стояла Оренбургская ученая архивная комиссия, созданная в конце XIX века. Ее деятельность имела принципиальное значение для сохранения значительного массива документов по истории Казахстана нового времени, многие из которых сохраняют свою научную ценность и в настоящее время (Uderbaeva 2016: 77). Уже на данном этапе

исследователи – фактически основоположники архивного дела в регионе – осознавали необходимость системного и бережного хранения архивных материалов.

Деятельность Оренбургской ученой архивной комиссии (далее – ОУАК) выходила за рамки простого комплектования документальных материалов. Комиссия фактически выступала одним из первых научных центров по изучению истории Степного края. Ее члены – историки, краеведы, представители административного аппарата – не только выявляли и сохраняли документы, но и вводили их в научный оборот посредством публикаций, составления описей фондов и подготовки исторических исследований по периоду Российской империи (Удербаева 2010: 113).

ОУАК сыграла значимую роль в формировании принципов архивного дела в регионе. Разрабатываемые ею подходы к систематизации, описанию и учету документов оказали влияние на последующее становление государственных архивов в советский период. Комиссия также выполняла функцию посредника между центром и регионом: благодаря ее деятельности значительное количество документов из местных канцелярий, военных архивов и частных собраний было выявлено, сохранено и частично передано в центральные архивохранилища.

Существенным является вклад ОУАК в сохранение не только официальных, но и этнографических, статистических и фольклорных материалов, что позволило расширить источниковую базу по истории и культуре казахского народа. Это способствовало сохранению уникальных сведений о традициях, быте и социальных отношениях, а также стимулировало интерес к историческому наследию среди местной интеллигенции. В данном контексте деятельность

комиссии может рассматриваться как один из этапов институционализации исторической науки и архивного дела в Казахстане.

По масштабам и результатам научной деятельности ОУАК во многом опережала другие губернские ученые архивные комиссии Российской империи. Комплексный характер ее работы включал архивную, исследовательскую и просветительскую составляющие. Значительный вклад деятелей комиссии в развитие архивного дела нашел отражение, в частности, в издании «Трудов» ОУАК, представляющих собой ценный исторический источник (Удербаева 2013: 200).

Преемственность в развитии научно-архивной деятельности проявилась в создании Общества изучения Киргизского края, в рамках которого формировалась общественная активность представителей казахской демократической интеллигенции. Принципы, заложенные ОУАК, – систематический сбор, научное описание и публикация документов – легли в основу современного понимания архивов как общественного института.

В дореволюционный период наблюдалась тенденция к постепенной трансформации архивов из преимущественно управленческих структур в учреждения общественного характера, ориентированные на научно-исследовательскую и справочно-информационную деятельность. Однако данный процесс был прерван революционными событиями 1917 года (Мазур 2021: 101).

В 1920-е годы советская власть приняла ряд нормативных актов, направленных на ограничение доступа к архивным документам. Было институционализировано понятие «секретного документа», а архивная система стала включать закрытые фонды, доступ к которым регулировался государством (Мазур 2021: 101).

Показательным примером является Перепись населения СССР 1937 года, результаты которой, не соответствовавшие ожиданиям власти, были засекречены, а часть организаторов подверглась репрессиям. Документы переписи были изъяты из научного оборота и помещены в закрытые архивные фонды, что иллюстрирует трансформацию архивов в инструмент государственного контроля над информацией.

Таким образом, в советский период архивы в значительной степени утратили функции открытого общественного института. Доступ к документам стал определяться политическими факторами, что ограничивало возможности объективного исторического исследования.

В более широком контексте это свидетельствует о том, что роль архивов как института доступа к памяти и знаниям является исторически изменчивой и зависит от характера политического режима. Если на рубеже XIX–XX веков формировалась традиция научной публикации источников, то в 1930-е годы произошел переход к политике закрытости архивной информации.

С 1940-х годов начинается процесс централизации архивной системы, в рамках которого ключевое место заняли партийные архивы. В период хрущевской оттепели предпринимались попытки либерализации доступа к архивным материалам, однако системные изменения были реализованы лишь в 1990-е годы в ходе архивной реформы, сопровождавшейся ликвидацией

партийных архивов и началом процессов рассекречивания документов (Мазур 2021: 101).

После обретения независимости в Казахстане начался новый этап развития архивного дела. Как отмечает директор Архива Президента Республики Казахстан А.Х. Мустафина, только в 1998 году был принят первый Закон «О Национальном архивном фонде и архивах». И именно на начало 2000-х пришлось формирование современной нормативно-правовой базы в сфере архивного дела Казахстана (Каштелюк 2025).

Закон 1998 года регулирует общественные отношения в области сбора, комплектования, хранения, государственного учета и использования документов Национального архивного фонда Республики Казахстан и Государственного управления архивным делом в Республике Казахстан. Основной его целью определено создание правовых основ для обеспечения сохранности историко-документального наследия Республики Казахстан в интересах граждан, общества и государства (Закон РК 1998).

Сегодня архивы в Казахстане и во всем мире уже не воспринимаются лишь как «хранилища документов». Они выполняют гораздо более широкую общественную миссию – обеспечивают доступ к коллективной памяти, знаниям и способствуют формированию идентичности. Именно благодаря усилиям первых архивистов, в том числе членов ОУАК, была заложена прочная источниковая база. Продолжая эту традицию, современные казахстанские архивисты активно сохраняют и систематизируют огромный массив ценных документов, который сегодня востребован исследователями, в том числе зарубежными, студентами и широким кругом общества.

Современные архивы становятся открытыми: развивается цифровизация фондов, создаются электронные каталоги, онлайн-доступ к документам.

Из новостей Архива Президента Республики Казахстан широкая общественность узнает, что архив активно расширяет процесс цифровизации. На сегодняшний день специалистами оцифровано более 235 тыс. дел, что составляет около 29 млн листов архивных документов.

Это делает историческое наследие более доступным для широкой аудитории, а не только для узкого круга специалистов. В этом смысле архивы выступают как инструмент демократизации знания.

Кроме того, архивы играют важную роль в формировании национальной и культурной идентичности. Документы, собранные еще в дореволюционный период, позволяют переосмысливать историю, возвращать забытые имена и события, формировать более целостное представление о прошлом страны.

Таким образом, если ОУАК заложила основы системного сохранения исторических источников, то современные архивы развивают эту традицию, превращаясь в активный общественный институт, который не только хранит прошлое, но и влияет на настоящее – через образование, науку и общественное сознание.

В настоящее время архивы Казахстана являются хранилищами памяти о сложных и значимых страницах национальной истории, таких как имперский период, советская эпоха и становление независимости. После обретения

независимости страна столкнулась с необходимостью переосмысления этих исторических этапов.

Особое значение имеют материалы советского периода, включая документы о политических репрессиях и голоде 1930-х годов, затронувших значительную часть населения страны. Эти источники позволяют восстанавливать историческую справедливость, исследовать события, которые долгое время замалчивались или искажались в официальной историографии. Открытие архивных документов важно для того, чтобы новые поколения могли осмыслить трагические страницы прошлого и извлечь из них уроки.

С момента обретения независимости архивисты Казахстана активно занимаются модернизацией системы, включая цифровизацию и создание электронных баз данных. Вместе с тем сохраняются проблемы доступа к отдельным материалам и неравномерного распределения цифровых ресурсов между крупными городами и регионами.

Программы цифровизации архивных документов расширяют возможности работы с историческими источниками для исследователей и граждан из разных регионов. Важным шагом стало развитие электронной платформы «Электронный архив Казахстана», упрощающей доступ к ключевым документам.

Интересно, что по инициативе Президента страны Касым-Жомарта Токаева национальный цифровой архив будет доступен отечественным и зарубежным разработчикам нейросетей (Кущанова 2025).

Доступ к архивным материалам способствует распространению научных знаний и углублению исторических исследований. Современные архивы создают условия для накопления, систематизации и анализа информации, что позволяет историкам, исследователям, включая зарубежных, а также журналистам и широкой общественности изучать прошлое и лучше понимать современные вызовы.

Таким образом, архивы как общественный институт имеют большое значение для сохранения документального наследия, развития научных исследований и формирования общественной, национальной и культурной идентичности.

Литература

Uderbaeva 2014 – Uderbaeva S.K. (2016). «Proceedings» of the Orenburg Scientific Archival Commission as a Historical Source. Oriente Moderno. 96(1): 76-98. Publisher: Brill.

Каштелюк 2025 – Каштелюк Ю. Вклад в сохранение истории// Мысль. 20 августа 2025 г. // https://mysl.kazgazeta.kz/news/16457 (дата обращения: 20.03.2026).

Кущанова К. Национальный цифровой архив создадут в Казахстане// Режим доступа: https://www.inform.kz/ru/natsionalniy-tsifrovoy-arhiv-sozdadut-v-kazahstane-22e55f (дата обращения: 20.03.2026).

Мазур 2021 – Мазур Л.Н. Архив как институт памяти: особенности институциональной трансформации// Архив в социуме – социум в архиве : материалы четвертой Всерос. науч.-практ. конф. /сост., науч. ред. Н. А. Антипин. Челябинск: ОГАЧО, 2021. С. 100-103.

О Национальном архивном фонде и архивах. Закон Республики Казахстан от 22 декабря 1998 года// Режим доступа: № 326-1.https://adilet.zan.kz/rus/docs/Z980000326_ (дата обращения: 20.03.2026).

Около 9 миллионов страниц документов оцифровано в Архиве Президента РК// Режим доступа: https://archive.president.kz/ru/news/prezident-arkhivinde-29-million-a-zhuy-zhat-para-y-tsifrly-format-a-k-shirildi/ (дата обращения: 20.03.2026).

Удербаева 2010 – Удербаева С.К. «Труды» Оренбургской ученой архивной комиссии как источник по истории Казахстана (период деятельности 1887–1917 гг.): Дис.... канд. ист. наук: 07.00.09. Алматы, 2010. 173 с.

Удербаева 2013 – Удербаева С.К. Возникновение и формирование «Трудов» Оренбургской ученой архивной комиссии как исторического источника // Вестник Национальной академии наук Республики Казахстан. № 2. 2013. С. 200-210.

ЦИФРОВЫЕ АРХИВНЫЕ РЕСУРСЫ КАК ИНСТРУМЕНТ ПОПУЛЯРИЗАЦИИ ИСТОРИЧЕСКОГО НАСЛЕДИЯ: ПОРТАЛ «ДОКУМЕНТАЛЬНОЕ НАСЛЕДИЕ ЮЖНОГО УРАЛА XVII–XVIII ВЕКОВ»

Светлана Рифовна Ардашова
Первый заместитель директора,
Объединенный государственный архив Челябинской области, Россия
ardashova@archive74.ru

Аннотация. В статье анализируется создание цифрового информационного ресурса по истории южноуральского региона на основе архивных документов «Документальное наследие Южного Урала». Цифровой ресурс создается специалистами Объединенного государственного архива Челябинской области в рамках одноименного проекта, целью которого является широкая популяризация исторического наследия региона, восполнение исторических пробелов документами из федеральных и региональных архивов, создание условий для свободного изучения архивных документов, инициативное вовлечение исследователей в развитие портала. Рассматриваются подходы к созданию научно-справочного аппарата, новизна представления информации, удобные сервисы поиска информации, обеспечение возможности поиска одновременно по десяти фильтрам, в том числе и полнотекстовый поиск. Выделяются особенности цифрового архивного ресурса: археографически полное и точное описание подлинника архивного документа и его электронной копии, представляющее собой подробную характеристику; обеспечение возможности понимания текста непрофессиональными историками, для чего частью археографического описания документов является транслитерация текста, которая выполнена профессиональными филологами.

В качестве перспектив развития рассматриваются инициативное взаимодействие с пользователями, предоставление им возможности делиться имеющимися копиями

архивных документов или предложить собственную транслитерацию текста, использование нейросетей при транслитерации отдельных текстов архивных документов.

Ключевые слова: электронный информационный ресурс, база данных, транслитерация текста, электронные ресурсы на основе архивных документов, тематические порталы

ЦИФРЛЫҚ АРХИВТІК РЕСУРСТАР – ТАРИХИ МҰРАНЫ ДӘРІПТЕУ ҚҰРАЛЫ: «XVII–XVIII ҒАСЫРЛАРДАҒЫ ОҢТҮСТІК ОРАЛДЫҢ ҚҰЖАТТЫҚ МҰРАСЫ» ПОРТАЛЫ

Светлана Рифовна Ардашова
Челябинск облысының Біріккен мемлекеттік архиві директорының бірінші орынбасары, Ресей
ardashova@archive74.ru

Аңдатпа. Мақалада «Оңтүстік Оралдың деректі мұрасы» архив құжаттары негізінде Оңтүстік Орал өңірінің тарихы бойынша цифрлық ақпараттық ресурсты құру мәселесі талданады. Бұл цифрлық ресурс өңірдің тарихи мұрасын кеңінен насихаттау, федералдық және өңірлік архивтерден алынған құжаттар арқылы тарихи олқылықтардың орнын толтыру, архив құжаттарын еркін зерттеу үшін жағдай жасау және порталды дамытуға зерттеушілерді бастамашыл түрде тартуды мақсат ететін аттас жоба аясында Челябинск облысының Біріккен мемлекеттік архивінің мамандары тарапынан жасалуда. Мақалада ғылыми-анықтамалық аппаратты құрудың тәсілдері, ақпаратты ұсынудың жаңалығы, ыңғайлы іздеу сервистері, соның ішінде он фильтр бойынша бір мезгілде іздеу және толықмәтіндік іздеу мүмкіндіктерін қамтамасыз ету қарастырылады. Цифрлық архив ресурсының ерекшеліктері айқындалады: архив құжатының түпнұсқасы мен оның электронды көшірмесін археографиялық тұрғыдан толық және дәл сипаттау, бұл егжей-тегжейлі сипаттама болып табылады; кәсіби емес тарихшылардың мәтінді түсіну мүмкіндігін қамтамасыз ету, бұл үшін құжаттарды археографиялық сипаттаудың бір бөлігі ретінде кәсіби филологтар орындаған мәтінді транслитерациялау қолданылады.

Даму перспективалары ретінде пайдаланушылармен бастамашыл өзара іс-қимыл жасау, оларға қолда бар архив құжаттарының көшірмелерімен бөлісу немесе мәтінді өз транслитерациясын ұсыну мүмкіндігін беру, сондай-ақ архив құжаттарының жекелеген мәтіндерін транслитерациялауда нейрожелілерді пайдалану мәселелері қарастырылады.

Түйін сөздер: электронды ақпараттық ресурс, деректер базасы, мәтінді транслитерациялау, архив құжаттары негізіндегі электронды ресурстар, тақырыптық порталдар

DIGITAL ARCHIVAL RESOURCES AS A TOOL FOR POPULARIZING HISTORICAL HERITAGE: THE PORTAL «DOCUMENTARY HERITAGE OF THE SOUTHERN URALS OF THE 17TH–18TH CENTURIES»

Svetlana R. Ardashova
First deputy director, United State Archives of the Chelyabinsk Region, Russia
ardashova@archive74.ru

Abstract. This article analyzes the creation of a digital information resource on the history of the South Ural region based on archival documents, "Documentary Heritage of the Southern Urals." The digital resource is being developed by specialists from the United State Archives of the Chelyabinsk Region as part of a project of the same name, which aims to widely popularize the region's historical heritage, fill historical gaps with documents from federal and regional archives, create conditions for the free study of archival documents, and proactively engage researchers in the development of the portal. The article examines approaches to creating a scientific reference tool, innovative information presentation, user-friendly information search services, and the ability to search across ten filters simultaneously, including full-text search. The following key features of a digital archival resource are highlighted: an archaeologically complete and accurate description of the original archival document and its electronic copy, representing a detailed characterization; and the ability to understand the text by non-professional historians, which includes transliteration of the text by professional philologists as part of the archaeographic description of the documents.

Proactive interaction with users, providing them with the opportunity to share existing copies of archival documents or offer their own transliteration of the text, and the use of neural networks in the transliteration of individual texts of archival documents are considered as development prospects.

Keywords: Electronic information resource, database, text transliteration, electronic resources based on archival documents, thematic portals

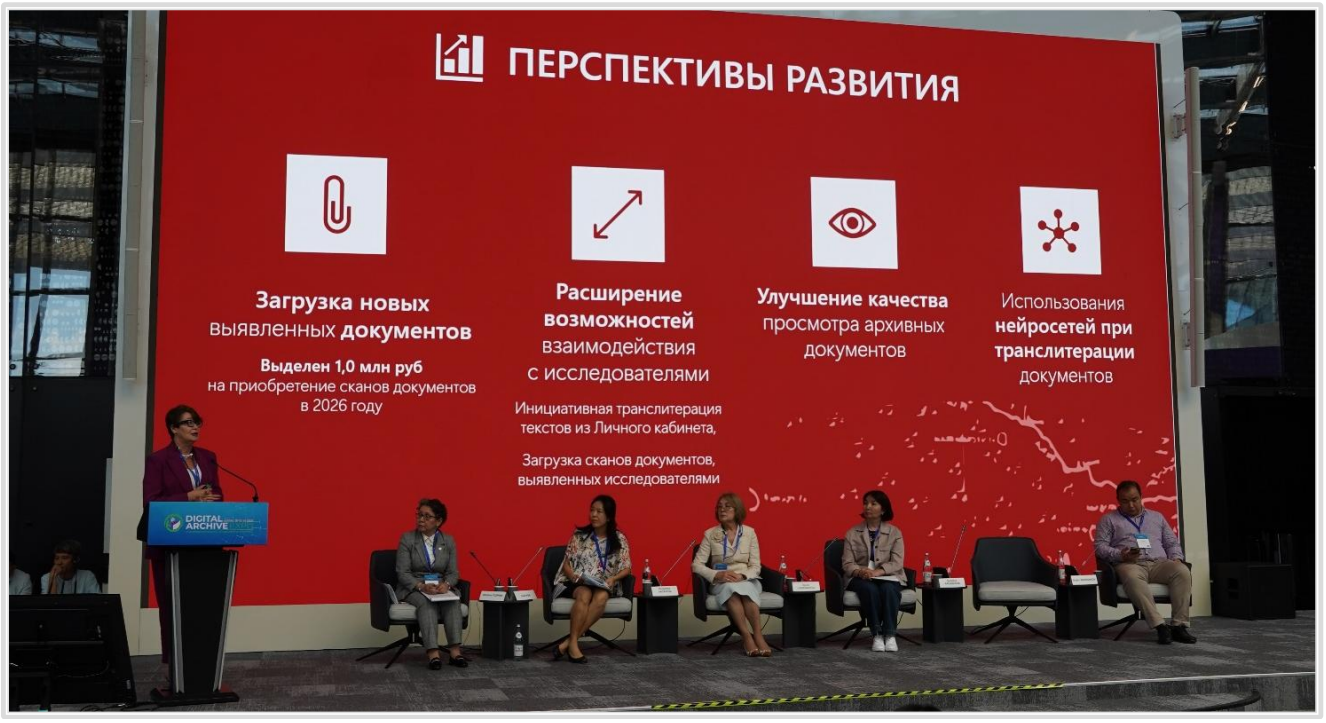
Тематические архивные цифровые ресурсы, на которых размещаются цифровые копии документов, в настоящее время становятся одной из распространенных форм представления и использования электронных архивных документов (Юмашева 2016: 32). Они представляют собой специализированные веб-сайты или информационные системы, где собраны архивные документы по определенным темам или тематическим направлениям, и позволяют пользователям быстро найти нужные документы, используя тематические каталоги и поисковые системы, а также ознакомиться с описаниями и контекстом документов (Афанасьева 2013: 53).

Причины интереса к созданию тематических порталов архивных документов очевидны, т.к. подобный формат позволяет обеспечить как доступность информации, ее свободное использование, так и удобство просмотра и поиска (которые одновременно определяются программными возможностями порталов).

Остановимся более подробно на целях создания, особенностях и перспективах развития портала «Документальное наследие Южного Урала», который был создан специалистами Объединенного государственного архива Челябинской области в рамках проекта «Документальное наследие Южного Урала: возвращая забытое».

Проект был инициирован губернатором Челябинской области Алексеем Текслером в 2023 году и нацелен на изучение интересного, но слабо изученного периода в история Южного Урала и Зауралья – XVIII века. Этот период насыщен необычайно значимыми событиями: освоение края крестьянами, деятельность Оренбургской экспедиции, появление крепостей и заселение территории казаками, крестьянская война под руководством Емельяна Пугачева и жестокое его подавление, создание границы и сопротивление набегам кочевников, зарождение металлургической промышленности, появление городов и развитие

торговли. Однако в Объединенном государственном архиве Челябинской области документальное описание истории этого периода сохранилось фрагментарно, в частности в силу того, что сформированный изначально комплекс документов пострадал в ходе уже упомянутого восстания под предводительством Е.И. Пугачева, и на хранение в результате попала лишь малая часть документов. Сохранились фонды городского и сословного самоуправления Челябинска, Челябинского духовного правления и частично делопроизводство заводов (Антипин 2024: 7).



Тот факт, что документальные комплексы XVIII века сосредоточены в федеральных и региональных архивах Урала и Сибири, а также сложность чтения и понимания скорописных текстов, явились причинами недостаточно внимания к данному периоду со стороны исследователей. Поэтому инициированный губернатором проект «Документальное наследие Южного Урала: возвращая забытое» нацелен на то, чтобы восполнить исторические пробелы документами из федеральных и соседних региональных архивов и тем самым открыть новые страницы ранней истории региона, ответить на дискуссионные вопросы, создать условия для широкой популяризации истории.

Задачи проекта:

- выявить новые архивные документы по истории южноуральского региона XVII–XVIII веков;
- обеспечить публикацию выявленных архивных документов (сборники документов, публикации в СМИ, в интернет-ресурсах);
- создать архивную коллекцию электронных документов, обеспечить к ней свободный доступ;
- способствовать расширению знаний о прошлом Южного Урала, активизации научных исследований, созданию позитивного образа региона.

Результатом работы в рамках первого этапа проекта стало выявление более трех тысяч ранее неиспользовавшихся архивных документов, содержащих более десяти тысяч листов, создание электронного информационного ресурса «Документальное наследие Южного Урала», на котором были размещены цифровые копии 1 439 выявленных документов, а также все иные материалы, относящиеся к реализации проекта (https://southuralhistory.ru/; https://южныйуралистория.рф/), подготовка и издание на основании выявленных и ранее не использованных документов двух сборников.

Проект сразу был запланирован как межархивный. На первом этапе реализации проекта партнерами Объединенного государственного архива Челябинской области выступили семь архивов, в том числе федеральные архивы Российской Федерации: Российский государственный архив древних актов, Российский государственный исторический архив, Российский государственный архив Военно-морского флота, Российский государственный военно-исторический архив, а также региональные: Государственный архив Свердловской области, Объединенный государственный архив Оренбургской области, Государственный архив в городе Шадринске. Также неоценимую помощь и содействие в выявлении документов, их оцифровке оказали Центр микрографии и реставрации архивных документов (г. Екатеринбург), Публичная библиотека Свердловской области им. Белинского.

Завершением первого этапа реализации проекта стала успешная его презентация губернатору Челябинской области, имевшая большой резонанс в региональных средствах массовой информации, интернет-ресурсах, вызвавшая широкий интерес профессионального исторического сообщества и генеалогов-любителей и имевшая результатом решение о продолжении проекта еще на три года, до 2029 года.

К настоящему времени круг партнеров расширился, на портале представлены архивные документы Государственного архива Пермского края, Национального архива Республики Башкортостан.

При создании портала «Документальное наследие Южного Урала» мы, в первую очередь, были нацелены на то, чтобы не просто представить приобретенные в архивах электронные копии архивных документов на веб-ресурсе, а создать археографически полное и точное описание подлинника архивного документа и его электронной копии и найти максимально качественный и комфортный способ визуализации электронной копии в цифровой среде, разработать и реализовать удобные сервисы поиска информации. Исходя из целей были поставлены следующие задачи. Во-первых, предоставить широкому кругу исследователей электронные копии архивных документов по заявленной тематике в хорошем качестве. Во-вторых, разработать удобную и понятную структуру информационного ресурса, которая позволила бы исследователю с любым опытом провести поисковую работу. В-третьих, максимально глубоко индексировать каждый архивный документ, раскрыв как его общее содержание, так и выделив основные тематические, географические, именные ссылки.

Хотя в работе над созданием портала принимало с учетом сжатых сроков реализации (всего полгода) ограниченное число специалистов (двадцать

человек), специализация рабочей группы оказалась очень разной. В нее вошли и профессиональные историки, в том числе пять кандидатов исторических наук, профессиональные филологи, в том числе доктор филологических наук, а также специалисты по созданию научно-справочного аппарата к архивным документам, специалисты сферы информационных технологий.

Остановимся на особенностях данного тематического электронного информационного ресурса.

Для обеспечения возможности многоуровневого и разнопланового поиска по архивной коллекции применены различные принципы систематизации документов. Во-первых, систематизация, в основе которой лежит хронологическая периодизация. Выделены разделы «Сибирская губерния», «Исетская провинция» (Исетская провинция 2024), «Оренбургская губерния», «Академические экспедиции». Во-вторых, в отдельные группы выделены документы, представляющие особый интерес для разных групп исследователей: «Карты и планы», «Генеалогические источники», «Полное собрание законов». В-третьих, исследователям дана возможность сформировать подборки архивных документов по архивам в «Архивные коллекции». Кроме того, в ходе поиска и отбора документов, представленных на портале, все принципы систематизации могут быть использованы в совокупности, исследователь может сформировать подборку документов, задав несколько фильтров одновременно. Например, «Исетская провинция» (Исетская провинция 2025) и «Российский государственный архив древних актов».

Археографическое описание каждого архивного документа представляет собой подробную его характеристику и включает в себя следующие элементы: заголовок документа, дату документа, основные темы, затронутые в документе (отсылка к предметному указателю), географические названия, встречающиеся в документе (отсылка к географическому указателю), персоналии, упомянутые в документе (отсылка к именному указателю), указание на систематический раздел, в который размещен документ. Археографическая легенда включает архивные поисковые данные (наименование архива, шифр документа), указание на подлинность.

Одной из главных особенностей археографического описания документов, размещенных на портале, является наличие транслитерации текста. Транслитерация текстов архивных документов выполнена доктором филологических наук Н.В. Глухих и независимым исследователем А.У. Шапоренко (г. Екатеринбург) по цифровым копиям документов. Археографическая обработка документов, включавшая сверку полученной транслитерации с подлинником документа, редактирование текста в соответствии с методическими рекомендациями, составление примечаний, заголовков документов и легенд к ним, датировку, произведена главными археографами Объединенного государственного архива Челябинской области – кандидатом исторических наук М.А. Базановым и кандидатом исторических наук С.А. Кусковым; археографическая обработка карт и планов – заведующим отделом научно-справочного аппарата государственного архива Е.В. Будановой. Археографическая обработка документов производилась согласно «Методическим рекомендациям по публикации архивных документов в

печатном виде» (Методические рекомендации 2022) с учетом специфических особенностей комплекса документов.

При транслитерации текстов вышедшие из употребления буквы заменялись современными, обозначающими тот же звук: например, буква Ѣ заменялась буквой Е, буква Ѳ – буквой Ф, буква Ѡ – буквой О; буква Ъ в конце слова не ставилась. Сохранялись орфография и грамматика текста; суффиксы, приставки, окончания не заменялись современными. Выносные буквы внесены в состав слова, что не оговаривается в примечаниях. В случаях слитного написания слов при воспроизведении их в сборнике они пишутся отдельно. Орфографические ошибки не исправлялись, сохранялась вариативность написания. Правки, изменение способа воспроизведения, наличие маргиналий оговариваются в примечаниях. Многие из публикуемых документов имеют повреждения, а тексты – утраты, принесенные временем; в случаях, когда таких повреждений много, в примечании в начале документа указано: «Здесь и далее часть документа повреждена». Это означает, что последующее воспроизведение в квадратных скобках или многоточие являются следствием утрат.

В результате такой подробной археографической обработки текстов архивных документов и их индексации поиск в электронном ресурсе возможен одновременно по десяти фильтрам, в том числе и полнотекстовый поиск. Кроме того, в состав научно-справочного аппарата входят список сокращений, словарь терминов и устаревших слов, библиотека по теме проекта. Темы, персоналии и географические названия, упомянутые в документах, собираются в соответствующие предметный, именной и географический указатели. Количество ссылок в указатели из одного архивного документа не регламентировалось: один архивный документ может содержать в себе до ста ссылок на географические названия.

Для привлечения внимания широкого круга исследователей-любителей на портале размещена база данных «Первые жители: первопоселенцы Миасской, Чебаркульской, Челябинской и Еткульской крепостей», в основу которой легла информация из переписных книг населения вышеуказанных крепостей Исетской провинции, составленных в 1740 году. Переписные книги были выявлены в 1960-х годах челябинским краеведом И.В. Дегтяревым в фонде Сената[1], в дальнейшем краеведом была проведена транслитерация документов, которая хранится в его личном фонде[2]. Позднее транслитерация И.В. Дегтярева была переработана и уточнена челябинским краеведом В.В. Поздеевым[3], именно она взята за основу при составлении базы данных.

В переписных книгах учтено 588 семей. Для размещения на портале текст был преобразован в табличную форму, соответствующую формату базы данных, содержащую следующие графы: фамилия, имя, отчество, возраст главы семьи, возраст членов семьи и их отношение к главе семьи (жена, сын, дочь, сноха, внук и прочие члены семьи), место рождения и место жительства до переселения на Южный Урал, в какую крепость записан и кем. Поиск информации возможен по любому из этих критериев. Результатом поиска может быть как в целом информация о семье, так и карточка на отдельного его члена. Всего в базе данных 2 500 записей об отдельных персоналиях. Также информация из базы данных входит в глобальный поиск по порталу.

Можно сказать, что нам удалось создать полифункциональный интегрированный информационно-поисковый ресурс, в котором соединены описания первичной архивной документации и сведения вторичной документации (научно-справочный аппарат), при наличии возможности работать с полнотекстовыми версиями архивных документов (их электронных копий). Информационно-поисковая система снабжена интеллектуальной системой классификаторов, рубрикаторов, указателей, развитыми поисковыми механизмами и исследовательским инструментарием, адаптированным к условиям цифровой среды. При этом, на наш взгляд, наличие на портале полнотекстовых версий документов, снабженных транслитерированным переводом, устраняет неизбежную субъективность научно-справочного аппарата.

Перспективы развития электронного информационного портала связаны, во-первых, с его последовательным наполнением контентом. Причем речь идет как о продолжении выявления архивных документов по истории Южного Урала XVIII–XIX веков в федеральных и региональных архивах, их сканировании, археографической обработке, включая при необходимости транслитерацию текста, так и о размещении на портале разнообразной сопутствующей информации о проекте, публикаций по приобретенным проектам, медиафайлов, созданных выставочных проектов, новостных сюжетов, в том числе телевизионных. Все это в целом должно еще более повысить пользовательский интерес к ресурсу.

Во-вторых, планируется работа в плане улучшения функции просмотра электронных копий документов, обеспечение возможности полностраничного осмотра, выделения отдельных фрагментов.

Также перспективным направлением мы считаем развитие инициативного взаимодействия с пользователями, предоставление им возможности поделиться имеющимися копиями архивных документов или предложить собственную транслитерацию текста.

Наконец, обсуждается возможность использования нейросетей при транслитерации отдельных текстов архивных документов (в первую очередь различных списков и т.п.).

Источники

1. Российский государственный архив древних актов (РГАДА). Ф. 248. Оп. 3. Д. 144.

2. Объединенный государственный архив Челябинской области (ОГАЧО). Ф. Р-481. Оп. 1. Д. 62, 65, 67, 68.

3. ОГАЧО. Ф. Р-620. Оп. 1. Д. 263, 266.

Литература

Афанасьева 2013 – Афанасьева Л.П. Электронный научно-справочный аппарат архивов: новый этап в развитии // Делопроизводство. 2013. № 1-2.

Исетская провинция 2024 – Исетская провинция: повседневная жизнь в документах Миасской крепости: сб. док / сост. и науч. ред. Н.А. Антипин и др. Челябинск, 2024.

Исетская провинция 2025 – Исетская провинция в документах Российского государственного архива древних актов: Южный Урал и Зауралье в XVIII веке: сб. док. / сост. и науч. ред. Н.А. Антипин и др. Челябинск, 2025.

Методические рекомендации 2022 – Методические рекомендации по публикации архивных документов в печатном виде / Всероссийский научно-исследовательский институт документоведения и архивного дела. М., 2022.

Юмашева 2016 – Юмашева Ю.Ю. Тематические архивные базы данных и информационные ресурсы. Вестник Московского университета. Сер. 8. История. 2016. № 4.

КООРДИНАЦИОННАЯ РАБОТА РЕГИОНАЛЬНОГО ОФИСА UNESCO В Г. АЛМАТЫ В СОТРУДНИЧЕСТВЕ СТРАН ЦЕНТРАЛЬНОЙ АЗИИ В ОБЛАСТИ АРХИВНОГО И МУЗЕЙНОГО ДЕЛА

Маджер Масанов
Национальный программный офицер по коммуникации и информации,
Региональный офис ЮНЕСКО в Алматы, Казахстан
m.massanov@unesco.org

Аннотация. В статье рассматривается роль Регионального офиса ЮНЕСКО в Алматы в сохранении и развитии культурного наследия в странах Центральной Азии. Авторы анализируют стратегические направления работы ЮНЕСКО с музеями и архивами, в частности важность реализации программы «Память мира» и международных рекомендаций по охране культурного наследия. На примерах конкретных международных проектов, таких как цифровизация общего наследия Шелкового пути и изучение буддийского наследия в Центральной Азии, показаны передовые практики межрегионального сотрудничества. Подчеркивается потенциал офиса ЮНЕСКО в Алматы в развитии профессиональных навыков архивных специалистов, обмене передовым опытом и популяризации регионального наследия на глобальном уровне. Статья определяет значение культурного наследия в информационном обществе и роль архивов как общественных институтов в обеспечении подотчетности, прозрачности и исторической справедливости.

Ключевые слова: культурное наследие, архивы, музеи, «Память мира», Шелковый путь, цифровизация, международное сотрудничество, Центральная Азия, историческая память

ОРТАЛЫҚ АЗИЯ ЕЛДЕРІНІҢ АРХИВ ЖӘНЕ МУЗЕЙ ІСІ САЛАСЫНДАҒЫ ЫНТЫМАҚТАСТЫҒЫНДАҒЫ ЮНЕСКО-НЫҢ АЛМАТЫ ҚАЛАСЫНДАҒЫ ӨҢІРЛІК КЕҢСЕСІНІҢ ҮЙЛЕСТІРУ ҚЫЗМЕТІ

Маджер Масанов
Коммуникация және ақпарат мәселелері жөніндегі ұлттық бағдарламалық маман, ЮНЕСКО-ның Алматыдағы өңірлік кеңсесі, Қазақстан
m.massanov@unesco.org

Аңдатпа. Бұл мақалада ЮНЕСКО-ның Алматыдағы Өңірлік кеңсесінің Орталық Азия елдеріндегі мәдени мұраны сақтау және дамыту саласындағы рөлі қарастырылады. Авторлар ЮНЕСКО-ның мұражайлар мен архивтермен жұмыс істеудегі стратегиялық бағыттарын, атап айтқанда, «Память мира» («Әлем жады») бағдарламасын және мәдени мұраны қорғау жөніндегі халықаралық ұсынымдарды жүзеге асырудың маңыздылығын талдайды. Мақалада Жібек жолының ортақ мұрасын цифрландыру және Орталық Азиядағы буддалық мұраны зерттеу сияқты нақты халықаралық жобалардың мысалында өңіраралық ынтымақтастықтың озық тәжірибелері көрсетілген. ЮНЕСКО Алматы кеңсесінің архив мамандарының кәсіби дағдыларын дамыту, озық тәжірибе алмасу және өңірлік мұраны жаһандық деңгейде танымал ету бойынша әлеуеті мен ынтымақтастыққа ашықтығы атап өтіледі. Мақала мәдени мұраның ақпараттық қоғамдағы маңызын және архивтердің қоғамдық институт ретіндегі подотчеттылықты, ашықтықты және тарихи әділеттілікті қамтамасыз етудегі рөлін айқындайды.

Түйін сөздер: мәдени мұра, архивтер, мұражайлар, «Әлем жады», Жібек жолы, цифрландыру, халықаралық ынтымақтастық, Орталық Азия, тарихи жады

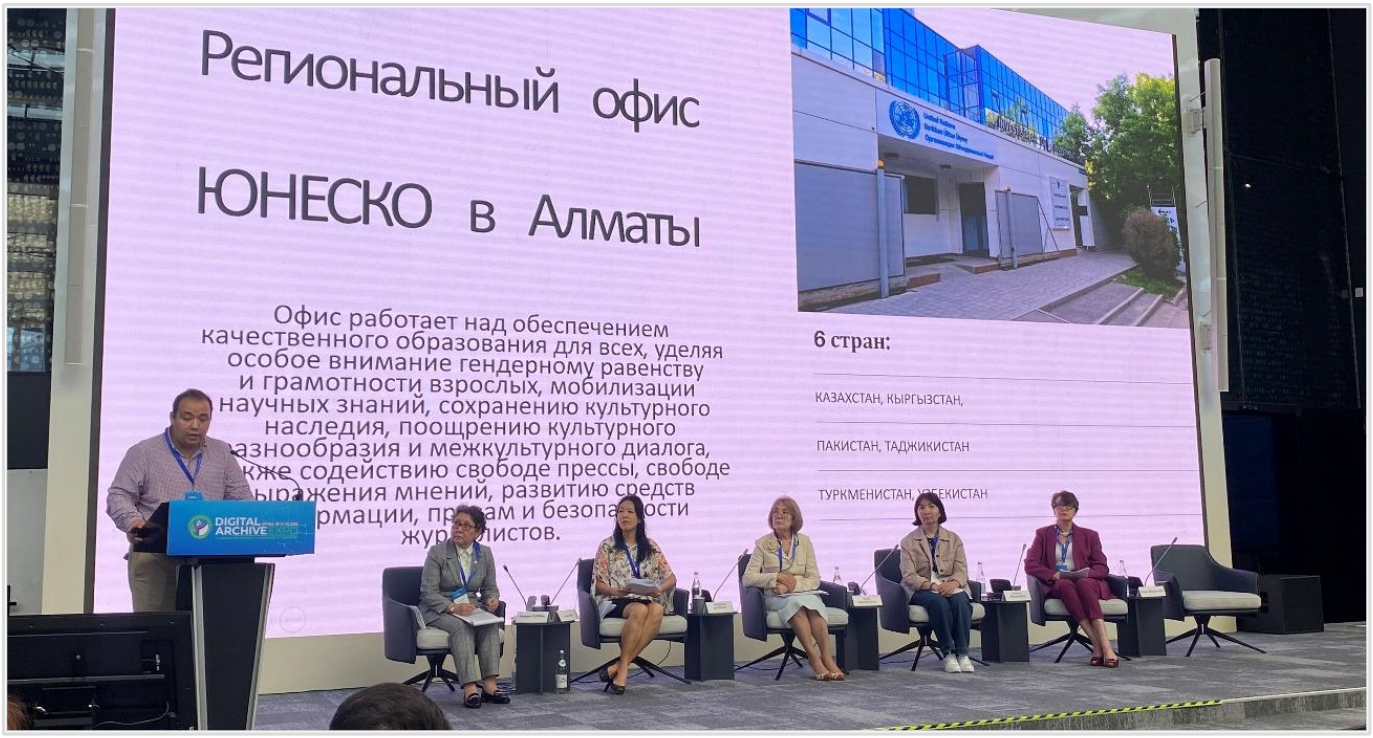
THE COORDINATING ROLE OF THE UNESCO REGIONAL OFFICE IN ALMATY IN FOSTERING COOPERATION AMONG CENTRAL ASIAN COUNTRIES IN THE FIELDS OF ARCHIVES AND MUSEUMS

Madjer Massanov
National Programme Officer for Communication and Information,
UNESCO Regional Office in Almaty, Kazakhstan
m.massanov@unesco.org

Abstract. This article examines the role of the UNESCO Regional Office in Almaty in the preservation and development of cultural heritage in Central Asian countries. The authors analyze UNESCO's strategic approaches to working with museums and archives, with particular emphasis on the implementation of the "Memory of the World" program and international recommendations for the protection of cultural heritage. Through examples of specific international projects, such as the digitalization of the shared heritage of the Silk Road and the study of Buddhist heritage in Central Asia, the paper highlights best practices in interregional cooperation. It underscores the potential of the UNESCO Almaty Office in developing the professional skills of archive specialists, facilitating the exchange of best practices, and promoting regional heritage on a global level. The article defines the significance of cultural heritage in the information society and the role of archives as public institutions in ensuring accountability, transparency, and historical justice.

Keywords: cultural heritage, archives, museums, "Memory of the World", Silk Road, digitalization, international cooperation, Central Asia, historical memory

Региональный Офис ЮНЕСКО в Алматы (далее – ЮНЕСКО Алматы) был создан в 1994 году и в данный момент осуществляет свою деятельность на территории Казахстана, Кыргызстана, Пакистана, Таджикистана, Туркменистана и Узбекистана при поддержке национальных офисов ЮНЕСКО в Исламабаде и Ташкенте. Офис работает над обеспечением качественного образования для всех, уделяя особое внимание гендерному равенству и грамотности взрослых, мобилизации научных знаний, сохранению культурного наследия, поощрению культурного разнообразия и межкультурного диалога, а также содействию свободе прессы, свободе выражения мнений, развитию средств информации, правам и безопасности журналистов.



В том числе ЮНЕСКО работает над развитием навыков среди музеев, библиотек и архивов. «Назначение музеев не ограничивается демонстрацией и хранением экспонатов. Музеи могут играть ведущую роль в укреплении креативной экономики на местном и региональном уровне. Они занимают все более важное место в социальной сфере, выступая в качестве платформ для дебатов и обсуждений, участвуя в решении сложных социальных проблем и привлекая к участию общественность» (Музеи ЮНЕСКО). Ключевое значение для работы ЮНЕСКО с музеями имеет Рекомендация о защите и развитии музеев и коллекций, их разнообразии и роли в обществе (Рекомендация 2015). Разработка данной Рекомендации, официально принятой на 38-й сессии Генеральной конференции 17 ноября 2015 года, была продиктована стремлением дополнить и расширить сферу применения норм и принципов, закрепленных в существующих международных документах, касающихся места музеев, а также связанных с ними функций и обязанностей, в том числе работ Международного совета музеев (ИКОМ).

Среди последних примеров того, как ЮНЕСКО Алматы координировала сотрудничество между музеями Центральной Азии, можно выделить виртуальную выставку «Общее наследие Шелкового пути» (Виртуальная выставка), которая была выполнена в 2024 году рамках проекта «Коридоры наследия Шелкового пути в Афганистане, Центральной Азии и Иране – международное измерение Европейского года культурного наследия» (Коридоры наследия Шелкового пути 2023) при финансовой поддержке Европейского Союза. Данная виртуальная выставка была разработана в сотрудничестве с национальными музеями Ирана, Казахстана, Кыргызстана, Таджикистана, Туркменистана и Узбекистана и доступна на платформе Google Arts and Culture. Она включает в себя подробные электронные 3D-копии шестидесяти выдающихся музейных экспонатов, демонстрирующих культурное наследие Шелкового пути. Кроме того, в качестве образовательного и информационного материала для виртуальной выставки были подготовлены две серии видеороликов для YouTube-аккаунта ЮНЕСКО Алматы. В первой представлен общий обзор всех музейных объектов, включенных в виртуальную выставку. Вторая содержит короткое видео о каждом из выбранных музейных объектов и их выдающейся художественной ценности. Эти короткие видеоролики доступны в двух специальных плейлистах на английском[1*] и на русском[2*] языках.

В данный момент ЮНЕСКО Алматы реализуется другой проект под названием «Содействие сохранению буддийского наследия и его документирование», который был поддержан правительством Японии и рассчитан до марта 2028 года. Проект реализуется в Казахстане и Кыргызстане в сотрудничестве с местными научно-исследовательскими организациями и архивами и направлен на проведение новых полевых и историографических исследований с целью повышения осведомленности и понимания этих объектов, их распределения на территории обеих стран, а также их исторической значимости.

ЮНЕСКО является единственной международной и межправительственной организацией, работающей на местах, которая обладает необходимым опытом для привлечения различных партнеров, заинтересованных сторон и учреждений к решению вопросов, связанных с материальным и документальным наследием. По этой причине, в рамках проекта также будут разработаны меры по обеспечению их сохранности и возможного включения в Список всемирного наследия и Международный реестр «Память мира».

Что касается архивов, то они являются местами открытий, способными пробудить воображение и дать нам возможность путешествовать во времени, а с появлением цифровых версий физических материалов – даже в пространстве. Они являются исходным материалом, из которого (пере)создаются и передаются нерассказанные истории, знания и новые идеи. Таким образом, архивы обладают потенциалом раскрывать, зачастую неожиданным образом, кто мы такие и как мы поступали, а также помогать нам формировать лучшее и более осознанное

[1*] URL: https://www.youtube.com/playlist?list=PLfHGi_qV1sN7AIyU4cqPLEEYUVqH612Gu
[2*] URL: https://www.youtube.com/playlist?list=PLfHGi_qV1sN4m7YuTQ1W0P7Nir2rnp2a-

настоящее и будущее. Являясь свидетельством наших действий как наций, обществ, организаций и отдельных людей, архивы играют основополагающую роль в обеспечении подотчетности, прозрачности, справедливости и доступа к информации.

В 1992 году ЮНЕСКО запустила программу «Память мира» (Memory of the World)[1*] с целью противостоять коллективной амнезии, призывая к сохранению ценных архивных фондов и библиотечных коллекций по всему миру и обеспечивая их широкое распространение. Идея программы заключается в том, что документальное наследие мира принадлежит всем, должно быть полностью сохранено и защищено для всех и, с должным учетом культурных традиций и практических аспектов, должно быть постоянно доступно для всех без каких-либо препятствий.

Таким образом, цели программы заключаются в следующем:

- содействовать сохранению документального наследия мира;
- способствовать всеобщему доступу к документальному наследию;
- повышать осведомленность общественности во всем мире о значении документального наследия.

Таким образом, важно отметить, что ЮНЕСКО Алматы имеет прошлый и действующий опыт активной работы с музеями и архивами стран Центральной Азии. В отдельных случаях наш офис может координировать международное сотрудничество в широко известных областях знания, предоставляя площадку для обмена опытом, развития навыков специалистов и межстранового сотрудничества в области культуры, как это было в случае наследия Шелкового пути. В других мы можем привлекать внимание к относительно малоизвестным фактам, связанным с наследием региона, как это осуществляется в данный момент в рамках проекта по буддийскому наследию. Любой из данных видов сотрудничества возможен и в рамках работы с архивами. ЮНЕСКО Алматы открыта для обсуждения идей потенциального сотрудничества и поиска финансирования для поддержки архивных специалистов региона включая развитие их профессиональных навыков, обмена передовым опытом и популяризации наследия центральноазиатского региона на местном и глобальном уровнях.

Литература

Виртуальная выставка – Виртуальная выставка «Общее наследие Шелкового пути. https://artsandculture.google.com/pocketgallery/nAWxtrFiZ4dQ1Q (дата обращения: 12.05.2026).

Коридоры наследия Шелкового пути 2023 – Коридоры наследия Шелкового пути в Центральной Азии. 26 июля 2023 г. https://www.unesco.org/ru/articles/koridory-naslediya-shelkovogo-puti-v-centralnoy-azii (дата обращения: 12.05.2026).

Музеи ЮНЕСКО. https://www.unesco.org/ru/museums (дата обращения: 12.05.2026).

Рекомендация 2015 – Рекомендация об охране и популяризации музеев и коллекций, их разнообразия и их роли в обществе. Принята генеральной конференцией на ее 38-й сессии Париж, 17 ноября 2015 г. https://unesdoc.unesco.org/ark:/48223/pf0000246331 (дата обращения: 12.05.2026).

[1*] URL: https://www.unesco.org/en/memory-world/

СИРЕК КІТАПТАР МЕН ҚОЛЖАЗБАЛАРДЫ ЦИФРЛАНДЫРУ: ҰЛТТЫҚ МҰРАНЫ МӘДЕНИЕТАРАЛЫҚ ИНТЕГРАЦИЯЛАУ

Шынарай Дүбірқызы Бүркітбаева
PhD, доцент м.а., әл-Фараби ат. Қазақ Ұлттық университеті жанындағы Түркітану және Алтайтану ғылыми-зерттеу институтының Қолжазбалар зертханасының меңгерушісі, Қазақстан
burkitbaeva.1975@mail.ru; https://orcid.org/0000-0001-8235-4213

Аңдатпа. Зерттеу жұмыста ұлттық тарихи-мәдени мұраны сақтау, қорғау және оны қоғам игілігіне тиімді пайдалану мәселелері жаһандану жағдайындағы рухани жаңғыру үдерісімен сабақтастыра қарастырылады. Еңбекте мәдени мұраны қастерлеу дәстүрінің түркілерде көне дәуірлерден бастау алатыны, соның ішінде Геродот еңбектеріндегі сақ тайпаларының ата-баба мұрасына деген құрметі арқылы дәлелденетіні көрсетіледі. Сонымен қатар түркі өркениетіндегі тарихи жадты сақтаудың ерекше үлгісі ретінде Орхон-Енисей жазба ескерткіштері талданып, олардың тек тарихи дерек қана емес, ұлттық бірегейлік пен идеологиялық тұтастықты қалыптастырушы рухани код ретіндегі маңызы айқындалады. Жұмыстың негізгі бағыты – сирек кітаптар мен қолжазбаларды цифрландырудың мәдениетаралық интеграциядағы рөлін анықтау. Цифрлық технологиялардың дамуы нәтижесінде жазба мұралардың қолжетімділігі артып, олардың халықаралық ғылыми айналымға ену мүмкіндігі кеңейгені талданады. Сонымен қатар Қазақстандағы қолжазба қорларының қазіргі жай-күйі, сақтау мәселелері, маман тапшылығы және материалдық-техникалық базаның жеткіліксіздігі сияқты өзекті проблемалар сараланады. Зерттеу қорытындысында сирек қолжазбалар мен кітаптарды цифрландыру ұлттық мұраны сақтаудың, жүйелеудің және оны әлемдік мәдени кеңістікке енгізудің тиімді құралы екені негізделеді. Бұл үдеріс мәдениетаралық диалогты дамытуға, ғылыми ынтымақтастықты кеңейтуге және ұлттық бірегейлікті сақтай отырып, оны жаһандық деңгейде танытуға мүмкіндік беретіні тұжырымдалады.

Түйін сөздер: ұлттық мұра, сирек кітаптар, қолжазба, архив, цифрлық технология, мәдениетаралық интеграция

ЦИФРОВИЗАЦИЯ РЕДКИХ КНИГ И РУКОПИСЕЙ: МЕЖКУЛЬТУРНАЯ ИНТЕГРАЦИЯ НАЦИОНАЛЬНОГО НАСЛЕДИЯ

Шынарай Дубиркызы Буркитбаева
PhD, и.о. доцента, заведующий Лабораторией рукописей
Научно-исследовательского института тюркологии и алтаистики
при Казахском национальном университете им. аль-Фараби, Казахстан
burkitbaeva.1975@mail.ru; https://orcid.org/0000-0001-8235-4213

Аннотация. В исследовании рассматриваются вопросы сохранения, охраны и эффективного использования национального историко-культурного наследия в контексте процессов духовной модернизации в условиях глобализации. Показано, что традиция бережного отношения к культурному наследию у тюркских народов восходит к глубокой древности, что подтверждается, в частности, свидетельствами Геродота о почитании предками саков наследия своих предков. Особое внимание уделено Орхоно-Енисейским письменным памятникам как уникальному примеру сохранения исторической памяти в тюркской цивилизации. Подчеркивается, что они представляют собой не только ценный исторический источник, но и духовный код, формирующий национальную идентичность и идеологическое единство.

Основное направление исследования связано с определением роли цифровизации редких книг и рукописей в развитии межкультурной интеграции. Анализируется влияние современных цифровых технологий на повышение доступности письменного наследия и расширение возможностей его включения в международный научный оборот. Кроме того, рассматриваются современное состояние рукописных фондов Казахстана, проблемы их сохранности, дефицит квалифицированных специалистов, а также недостаточность материально-технической базы.

В заключение обосновывается, что цифровизация редких рукописей и книг является одним из наиболее эффективных инструментов сохранения, систематизации и интеграции национального наследия в мировое культурное пространство. Делается вывод, что данный процесс способствует развитию межкультурного диалога, расширению международного научного сотрудничества и популяризации национального наследия на глобальном уровне при сохранении национальной идентичности.

Ключевые слова: национальное наследие, редкие книги, рукописи, архив, цифровые технологии, межкультурная интеграция

DIGITIZATION OF RARE BOOKS AND MANUSCRIPTS: CROSS-CULTURAL INTEGRATION OF NATIONAL HERITAGE

Shynarai Dubirkyzy Burkitbayeva
PhD, Acting Associate Professor, Head of the Manuscripts Laboratory,
Research Institute of Turkology and Altaic Studies,
al-Farabi Kazakh National University, Kazakhstan
burkitbaeva.1975@mail.ru; https://orcid.org/0000-0001-8235-4213

Abstract. This study examines the issues of preserving, protecting, and effectively utilizing the national historical and cultural heritage in the context of spiritual modernization under the conditions of globalization. It demonstrates that the tradition of preserving cultural

heritage among the Turkic peoples dates back to ancient times, as evidenced, in particular, by the accounts of Herodotus describing the reverence of the Saka tribes for the legacy of their ancestors. Special attention is given to the Orkhon–Yenisey inscriptions as a unique example of preserving historical memory within Turkic civilization. These monuments are viewed not only as valuable historical sources but also as a spiritual code that shapes national identity and ideological unity.

The main focus of the study is to determine the role of the digitization of rare books and manuscripts in promoting intercultural integration. The paper analyzes how modern digital technologies enhance the accessibility of written heritage and expand its integration into the international scholarly community. It also examines the current state of Kazakhstan's manuscript collections, highlighting challenges related to their preservation, the shortage of qualified specialists, and the inadequacy of material and technical resources.

The study concludes that the digitization of rare books and manuscripts is one of the most effective means of preserving, systematizing, and integrating national heritage into the global cultural space. This process contributes to the development of intercultural dialogue, the expansion of international scholarly cooperation, and the promotion of national heritage at the global level while preserving national identity.

Keywords: national heritage, rare books, manuscripts, archive, digital technology, intercultural integration

Бүгінгі таңда қазақ қоғамының мәдениет саласындағы негізгі стратегиялық міндеттерінің бірі – тарихи-мәдени мұраны тиімді сақтау, жүйелі түрде қорғау және оны қоғам игілігіне белсенді түрде пайдалану арқылы рухани жаңғыру үдерісін қамтамасыз ету болып табылады. Бұл мәселе әлемдік тарихи тәжірибе тұрғысынан қарастырғанда, өзінің бастауын көне дәуірлерден алады және қай кезеңде, қай елде болмасын өзектілігін жоймаған әмбебап құбылыс ретінде сипатталады. Тарихи-мәдени мұраны сақтау мәселесінің даму тарихына назар аударсақ, ескерткіштерді қорғау мен қадірлеудің алғашқы белгілері көне дәуірлерден бастау алатыны байқалады. Мұндай деректердің бірі көне грек тарихшысы Геродоттың еңбектерінде көрініс табады. Онда б.з.б. I мыңжылдықта өмір сүрген сақ тайпаларының қайтыс болған ата-бабаларын, олардың тұтынған заттарын және жерленген орындарын ерекше құрметтегені туралы мәліметтер келтірілген (Кадримбетова 2020 1). Бұл деректер сақтардың тек өткенді еске сақтау ниетін ғана емес, сонымен қатар ата-бабаларға қатысты заттарды киелі нысан деңгейіне көтеру арқылы олардың рухын қастерлеуге ұмтылғанын көрсетеді. Сонымен бірге мұндай әрекеттердің тәрбиелік және идеологиялық мәні де болғаны аңғарылады. Яғни сақ қоғамында қасиетті деп танылған заттар адамгершілік құндылықтардың көрінісі ғана емес, сонымен қатар ұжымдық сананы қалыптастырып, патриоттық сезімді нығайтатын маңызды идеологиялық құрал қызметін атқарған.

Көне сақ дәстүрінің заңды жалғасы ретінде Ұлы даланы мекендеген түркі халықтары да тарихи жадты сақтаудың өзіндік бірегей үлгісін қалыптастырды. Олар өздерінің саяси, әлеуметтік және рухани өмірін бейнелейтін маңызды деректерді табиғи ортада сақталатын материалдарға түсіру арқылы мәңгілік мұра қалдыруға ұмтылды. Соның айқын көрінісі – Орхон-Енисей жазба ескерткіштері. Түркілер Орхон мен Енисей өзендері бойындағы тастарға қашап жазу арқылы өз дәуірінің шежіресін, мемлекет тарихын, билеушілердің қызметін және халықтың

дүниетанымын кейінгі ұрпаққа жеткізуді мақсат етті. Бұл жазбалар тек ақпарат жеткізуші құрал ғана емес, сонымен қатар тарихи сананы қалыптастырушы, ұлт бірлігін нығайтушы идеологиялық мәтіндер ретінде де қызмет атқарды. Күлтегін, Білге қаған, Тоныкөк құрметіне тұрғызылған жазбаларда елдің тәуелсіздігі, бірлігі, ата-баба аманаты, мемлекетті сақтау идеялары кеңінен көрініс тапқан. Орхон-Енисей ескерткіштерінің ерекшелігі – олардың табиғи кеңістікте, ашық аспан астында орналасуы. Бұл жағдай оларды тек жазба дерек қана емес, сонымен бірге кең мағынадағы мәдени-тарихи ландшафттың бір бөлігіне айналдырады. Түркілер үшін тасқа қашалған мәтін – уақыт пен кеңістік шегінен тыс өмір сүретін, ұрпақтан-ұрпаққа жалғасатын мәңгілік үндеу іспетті болды. Сондықтан да бұл ескерткіштерді шартты түрде «ашық аспан астындағы мәңгілік архив» деп сипаттауға болады. Аталған жазбаларда көрініс тапқан идеялар тек тарихи фактілерді баяндаумен шектелмейді. Олар арқылы түркі қоғамының құндылықтар жүйесі, саяси идеологиясы, моральдық-этикалық ұстанымдары айқындалады. Мысалы, ел бірлігі мен тәуелсіздігін сақтау, билік пен халық арасындағы үйлесімділік, сыртқы қауіп-қатерлерге қарсы тұру мәселелері мәтіндердің негізгі өзегін құрайды. Бұл тұрғыдан алғанда, Орхон-Енисей жазбалары тек тарихи дерек көзі емес, сонымен қатар ұлттың рухани-идеологиялық кодын сақтаушы маңызды мұра болып табылады.

Еуропа елдерінің тарихында мәдени ескерткіштерді сақтау дәстүрі ежелгі дәуірлерден бастау алады. Ежелгі Римде мәдени мұраны қорғауға бағытталған яғни, ежелгі құрылыстарды бұзғаны үшін қатаң жаза қарастыратын арнайы жарлық шығарылып, нақты құқықтық шаралар қабылданған (Геродот 1972). Бұдан біз антикалық қоғамның өзінде тарихи нысандарды сақтау қажеттілігінің мемлекеттік деңгейде көтерілген мәселе болғанын көре аламыз. Жалпы, көне дәуірдің өзінде адамдар құнды заттарды арнайы сақтауға мән берген. Ежелгі Греция мен Римде түрлі қазыналар мен бағалы бұйымдарды сақтайтын арнайы қоймалар – алғашқы «қазына сақтау орындары» пайда болған. Дегенмен, бұл кезеңде мұндай заттар толыққанды тарихи-мәдени мұра ретінде емес, көбіне материалдық құндылық немесе байлық ретінде қабылданған. Өткен дәуір ескерткіштеріне мәдени мұра ретінде қарау үрдісі тек орта ғасырларда қалыптаса бастады. Дәл осы кезеңде көне жәдігерлер ғылыми және шығармашылық қауым өкілдерінің назарын аударып, оларды зерттеу мен жүйелеуге деген қызығушылық артты. Ал жаңа заманға қарай Еуропада мәдени ескерткіштер орналасқан аумақтарды мемлекет тарапынан қорғау жүйесі қалыптасты. Осы кезеңде ескерткіштер мемлекеттік меншік ретінде қарастырылып, оларды сақтау мен қорғау заңнамалық негізде жүзеге асырыла бастады. Еуропа мемлекеттері ішінде тарихи-мәдени мұра ескерткіштерін қорғауды құқықтық тұрғыдан алғаш жүйелеген ел – Швеция болып саналады. 1660 жылы Швецияның Мемлекеттік Кеңесі «Мәдениет ескерткіштерін қорғау туралы» арнайы заң қабылдағаны белгілі (Оттенсон 1985). Аталған заңда мемлекет қорғауына алынатын нысандардың кең ауқымы белгіленді. Олардың қатарына бекіністер, тас қамалдар, қорымдар мен қорғандар, көне тұрғын үйлер, шағын бекіністердің қалдықтары, жер иеліктерінің шекара белгілері, сондай-ақ тарихи шайқас өткен орындар енгізілді. Сонымен бірге шіркеулердегі көркемсуреттер, руникалық жазбалар мен монастырь кітаптары, заңдар мен діни-әдеби мәтіндер

қамтылған қолжазбалар да ерекше қорғауға алынды. Бұған қоса эпикалық поэмалар мен халық арасындағы аңыз-әңгімелерді жинақтап, хатқа түсіру де маңызды міндет ретінде қарастырылды. Кейінірек, 1666 жылы Швеция аумағындағы барлық археологиялық ескерткіштер ресми түрде корольдің жеке меншігі деп жарияланды. Бұл шешім тарихи мұраны сақтау ісінің мемлекеттік бақылауға толықтай өткенін және оның құқықтық негізде жүйелене бастағанын айқын көрсетеді (Жолдасбеков 2012). Кейін Дания, Франция, Белгия, Пруссия мен Саксония, Норвегия, Ұлыбритания, Шотландия, Ирландия елдерінде архив ісі қарқынды қолға алынады. Қазақстанда тек тәуелсіздік алған жылдардан бастап қолға алынғаны белгілі.

Қазіргі жаһандану жағдайында еліміздің мәдени мұрасын сақтау, зерттеу және оны әлемдік ғылыми-мәдени айналымға енгізу өзекті мәселелердің бірі болып табылады. Әсіресе сирек кітаптар мен қолжазбалар халықтың тарихи жадын, тілдік дамуын, рухани құндылықтарын, ғылыми және әдеби дәстүрлерін бейнелейтін аса құнды дереккөздер қатарына жатады. Алайда уақыт өте келе қағаздың тозуы, табиғи факторлардың әсері, сақтау жағдайының жеткіліксіздігі сияқты себептер бұл мұралардың жойылу қаупін арттырады. Осы тұрғыдан алғанда сирек кітаптар мен қолжазбаларды цифрландыру мәдени мұраны сақтаудың, жүйелеудің және оны халықаралық ғылыми қауымдастыққа қолжетімді етудің тиімді құралдарының бірі болып табылады.

Ел президенті Қ.К. Тоқаев Түркістан қаласында өткен Ұлттық Құрылтайда «Мен халықтың жады мен ғылым-білімін жинақтайтын «Ұлттық цифрлық мұра» қорын құру идеясын қолдаймын. Цифрландыру және жасанды интеллект жылы аясындағы мұндай бастама тарихи мұрамызды сақтап, оны терең зерделеуге жол ашады. Мен былтырғы Құрылтайда ғылыми еңбектерді, архив материалдарын, мұражай қорларын және өнер туындыларын жүйелеу қажеттігін ерекше атап өттім. «Ұлттық цифрлық мұра» бұл шаруаны заман талабына сай жүргізуге мүмкіндік береді» деген болатын. Қазіргі таңда шын мәнінде ұлттық мұраларымызды цифрландыру кезек күттірмейтін мәселе. Еліміздің ірі ірі кітапханаларында шағатай тілінде жазылған мол қолжазбалар қоры бар. Егер оны цифрландырсақ баршаға қолжетімді болар еді. Қолжетімділіктің нәтижесінде ол қолжазбалар шаң басып жатпай, ғылыми айналысқа түсіп, ел мүддесіне қызмет етер еді.

Цифрлық технологиялардың дамуы мәдени мұраларды сақтаудың дәстүрлі тәсілдерін жаңа деңгейге көтерді. Бұрын сирек қорларда сақталған қолжазбалар мен көне басылымдарға тек белгілі бір ғылыми орталықтар мен зерттеушілер ғана қол жеткізе алса, цифрландыру нәтижесінде оларды әлемнің кез келген нүктесінен қарастыру мүмкіндігі пайда болды. Бұл жағдай мәдени және ғылыми ақпарат алмасуды жеделдетіп, әртүрлі елдердің зерттеушілері арасында ынтымақтастықтың дамуына ықпал етеді. Сонымен қатар цифрлық қорлар мәдени мұраны болашақ ұрпаққа сақтап жеткізудің сенімді тәсілі ретінде де маңызды рөл атқарады.

Сирек кітаптар мен қолжазбаларды цифрландырудың мәдениетаралық интеграция тұрғысынан маңызы ерекше. Біріншіден, цифрлық форматқа көшірілген тарихи деректер түрлі халықтардың ортақ мәдени байланыстарын зерттеуге мүмкіндік береді. Мысалы, түркі халықтарының ортақ әдеби-мәдени

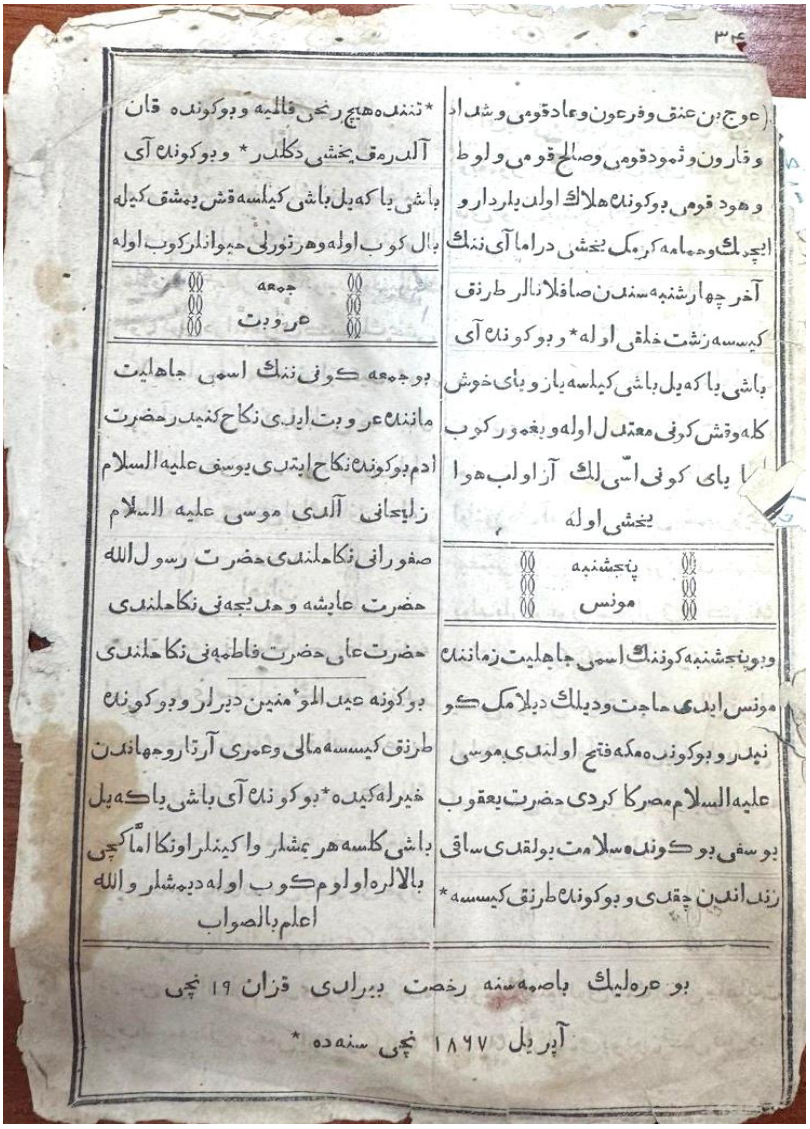
мұрасы болып саналатын ортағасырлық қолжазбалар, діни және тарихи трактаттар, лексикографиялық еңбектер немесе эпикалық шығармалар цифрлық қорларда жинақталып, түрлі елдердің ғалымдарына қолжетімді болады. Бұл жағдай түркі дүниесінің мәдени байланыстарын, тілдік ұқсастықтарын және тарихи даму кезеңдерін кешенді түрде зерттеуге мүмкіндік береді.

Екіншіден, цифрландыру мәдени мұраларды халықаралық ғылыми айналымға енгізуге жол ашады. Цифрлық кітапханалар мен ашық электрондық платформалар арқылы ұлттық мәдени мұра тек белгілі бір елдің шеңберінде ғана емес, әлемдік мәдени кеңістіктің бір бөлігіне айналады. Нәтижесінде ұлттық әдебиет, тарих және тіл мәселелері бойынша жаңа ғылыми зерттеулердің пайда болуына, сондай-ақ мәдениеттер арасындағы өзара түсіністік пен диалогтың дамуына жағдай жасалады.

Үшіншіден, сирек қолжазбалар мен көне кітаптарды цифрландыру білім беру саласында да маңызды рөл атқарады. Университеттер мен ғылыми орталықтар цифрлық қорларды оқу-зерттеу процесінде пайдалану арқылы студенттер мен жас ғалымдардың тарихи дереккөздермен жұмыс істеу мүмкіндігін кеңейтеді. Бұл, өз кезегінде, гуманитарлық ғылымдардың, әсіресе тарих, филология, лингвистика және мәдениеттану салаларының дамуына оң ықпал етеді. Сонымен қатар цифрландыру барысында ғылыми каталогтар мен дерекқорлардың құрылуы да үлкен маңызға ие. Әрбір цифрланған қолжазба немесе сирек кітап библиографиялық сипаттамамен, ғылыми түсіндірмелермен, мәтіндік және палеографиялық ерекшеліктерімен бірге ұсынылады. Мұндай жүйелі ақпараттық база зерттеушілерге дереккөздерді салыстырмалы-тарихи, лингвистикалық немесе мәтінтанулық тұрғыдан жан-жақты талдауға мүмкіндік береді.

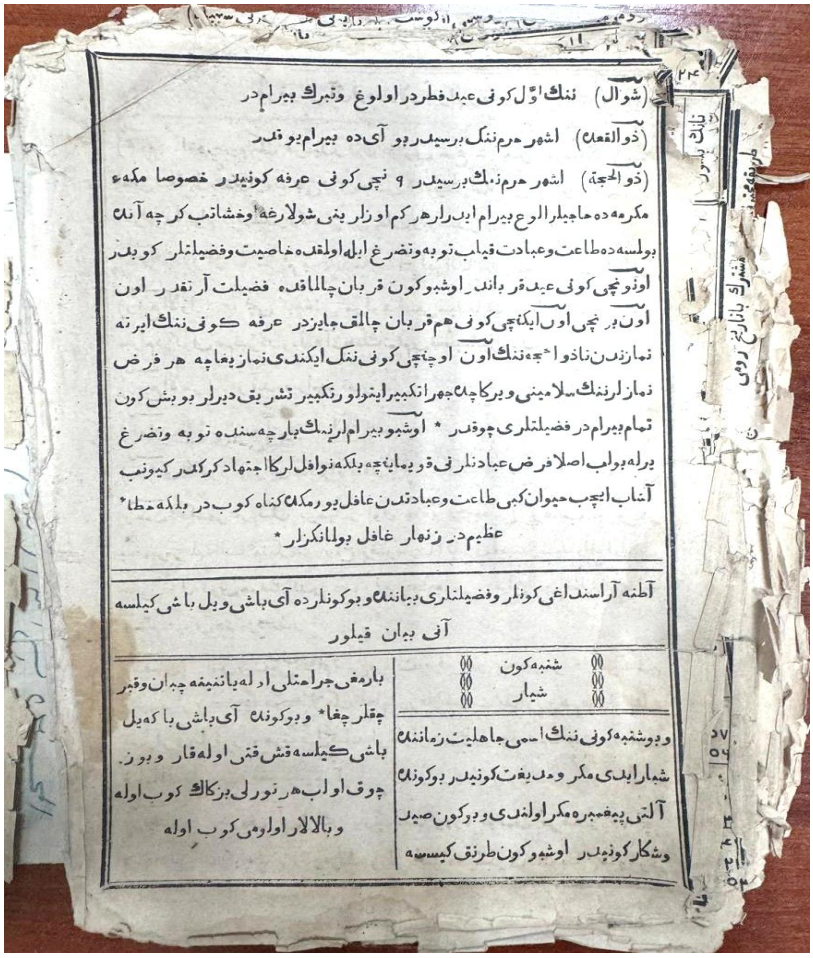
Сирек кітаптар мен қолжазбаларды цифрландыру – мәдени-тарихи құндылығы жоғары жазба мұраларды жоғары сапалы цифрлық форматқа көшіру, оларды электрондық дерекқорлар мен цифрлық кітапханаларда сақтау және көпшілікке ұсыну үдерісі. Бұл процесс тек техникалық әрекет қана емес, сонымен бірге ғылыми, мәдени және ақпараттық интеграцияны қамтамасыз ететін күрделі кешенді жұмыс болып саналады. Цифрландыру барысында мәтіндердің жоғары сапалы сканерленуі, метадеректердің жасалуы, ғылыми сипаттамалардың берілуі, мәтіндерді тану технологиялары арқылы іздеуге қолайлы форматқа келтіру сияқты кезеңдер жүзеге асырылады.

Бізде мәдени мұраны жоғалту қаупі де бар. Сирек қолжазбалар, аудиожазбалар және далалық материалдар жойылу алдында тұр десек те болады. Өйіткені кітапханалардың көпшілігіндегі қолжазбалар ережеге сай, тиісті жағдайда сақталып отырған жоқ. Мысалы, Біздің Әл-Фараби кітапханасындағы қолжазбаларды сканерлеу, цифрландыру ісі енді қолға алына бастады. Қолжазбаларды оқитын мамандардың тапшылығы да бұл мәселені күрделендіріп отыр. Мәселен, төмендегі суреттердегі сирек кітаптарды дер кезінде заманауи технологияны пайдаланып цифрландырып, сақтап қалмасақ, уақыт өте келе қағаз бұзылып жойылып кетеді (1-3 суреттер).



1-сурет. *«Гарелік» кітабы.*
әл-Фараби атындағы ҚазҰУ кітапханасы

Авторы көрсетілмеген, 1867 жылы Қазан қаласы Қазан университеті баспасында басылған кітап 34 беттен тұрады (1-сурет). Кітапта қырық жылға жыл басы, күн мен түннің ұзақтығы, күннің шығу және бату сағаты туралы мәліметтер берілген. Апта күндерінің атаулары туралы мәлімет берілген. Сонымен қоса, кітапта әр айда болатын мейрам күндері жазылған. Бұл бөлім «Хер қайсы айларда олан байрамлар билинур» деп аталады. Кітаптың шетіне айлардың арабша атауы, жанына сол айда болатын мейрамдар жазылған. Кітаптың бір бетінде жыл мезгілдері туралы жазылған. Жаз, қыс, күз, көктем мезгілдері шаршының әр қабырғасына жазылып, астына айлары, олардың неше күннен тұратыны, орыс тіліндегі атауы араб әріптерімен және араб тілінде жұлдыз жорамал атауларының атаулары берілген. Ортасында уақыт өлшемдері, олардың қанша уақыттан тұратыны туралы ақпарат берілген (1-сурет).



2-сурет. *«Гарелік» кітабының мүжілген парақтарының бірінің көрінісі. әл-Фараби атындағы ҚазҰУ кітапханасы*

Сондай-ақ, көптеген материалдардың жергілікті архивтерде сақталып, халықаралық ғылыми қауымдастық үшін қолжетімсіз күйде қалып отырғаны да шындық. Жазба ескерткіштерді, аудиоархивтерді, этнографиялық деректерді және қазіргі зерттеулерді біріктіретін бірыңғай цифрлық платформаның құрылуы аса қажет.

Түркі тілдерінде сөйлейтін өскелең ұрпақ дәстүрлі мәдени мұрамен байланысын жоғалтуда. Білім беру мекемелері заманауи цифрлық ресурстарға мұқтаж. Цифрлық кітапхана мәдениетаралық диалогты дамытуға және бірегейлікті нығайтуға ықпал етері сөзсіз.

Қазіргі таңда материалдар әртүрлі елдердің архивтерінде, көбіне қағаз форматында, каталогсыз және цифрлық қолжетімсіз күйде сақталуда. Еліміздің кітапханаларындағы қолжазбалар мен сирек кітаптардың ашық қолжетімді бірыңғай цифрлық платформасын құру кезек күттірмес мәселе. Ел басшысының «2026 жылды Цифрландыру және Жасанда интеллект жылы» деп жариялағаны белгілі. Осы үндеу аясында кітапханаларда жатқан қолжазбаларды цифрландыруға кіріскен жөн.

Ұлттық мұраны мәдениетаралық интеграциялау ең алдымен мәдениеттер арасындағы диалогты қамтамасыз етуге бағытталады. Әрбір халықтың мәдени мұрасы – тек сол ұлттың ішкі игілігі ғана емес, бүкіл адамзат өркениетінің

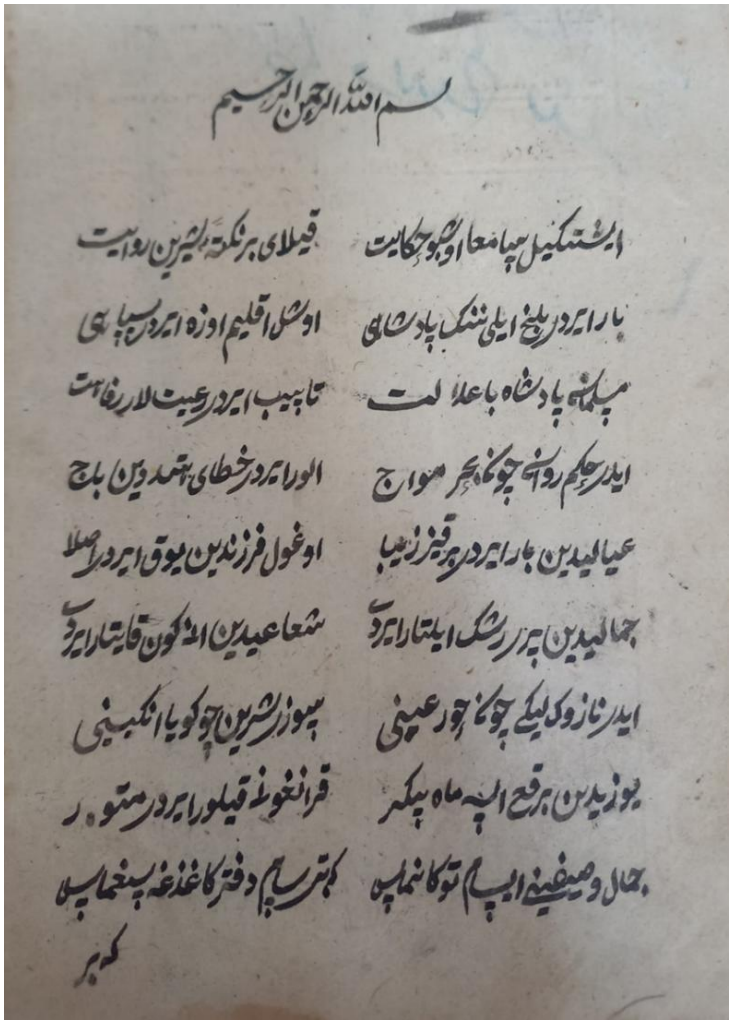
құрамдас бөлігі. Сондықтан ұлттық әдебиет, фольклор, тілдік деректер, тарихи жазбалар мен өнер туындылары басқа мәдениеттермен салыстырмалы түрде зерттеліп, ортақ және ерекшелік белгілері анықталады. Бұл үдеріс мәдениеттер арасындағы өзара түсіністікті тереңдетіп, өркениеттердің үйлесімді дамуына ықпал етеді. Мәдениетаралық интеграцияда тілдің рөлі ерекше. Тіл – мәдени ақпараттың негізгі тасымалдаушысы ретінде ұлттық дүниетанымды, тарихи тәжірибені және әлеуметтік құндылықтарды жеткізеді. Осыған байланысты аударма, лексикография, көптілді білім беру жүйесі ұлттық мұраны басқа тілдік кеңістікке енгізудің негізгі құралдары болып табылады. Әсіресе, тарихи мәтіндер мен көркем әдебиеттің сапалы аудармасы ұлттық мәдениеттің әлемдік деңгейде танылуына мүмкіндік береді. Ұлттық мұраны мәдениетаралық интеграциялаудың маңызды бағыттарының бірі – цифрлық технологияларды пайдалану. Цифрландыру нәтижесінде тарихи құжаттар, сирек кітаптар, қолжазбалар, архив материалдары электрондық форматқа көшіріліп, халықаралық ғылыми қауымдастыққа қолжетімді болады. Цифрлық кітапханалар мен онлайн платформалар арқылы ұлттық мұра географиялық шектеулерден тыс таралып, әртүрлі елдердің зерттеушілері үшін ашық ресурсқа айналады. Бұл өз кезегінде мәдени және ғылыми ақпарат алмасуды жеделдетіп, бірлескен зерттеулердің дамуына ықпал етеді. Сонымен қатар мәдениетаралық интеграция білім беру және ғылыми зерттеу салаларында да көрініс табады. Университеттер мен ғылыми орталықтарда ұлттық мәдениетке қатысты пәндердің енгізілуі, халықаралық ғылыми жобалардың жүзеге асырылуы, академиялық ұтқырлық бағдарламалары мәдениеттер арасындағы байланысты нығайтады. Мұндай интеграциялық үдерістер нәтижесінде ұлттық мұра тек зерттеу нысаны ретінде ғана емес, сонымен бірге мәдени дипломатияның құралы ретінде де қызмет атқарады.

Ұлттық мұраны мәдениетаралық интеграциялау барысында бірқатар мәселелер де туындайды. Олардың қатарында мәдени кодтардың толық аударылмауы, ұлттық ерекшеліктердің бұрмалануы, мәдени ассимиляция қаупі сияқты факторлар бар. Сондықтан бұл үдеріс ғылыми негізде, мәдени сезімталдық пен жауапкершілік принциптерін сақтай отырып жүзеге асырылуы тиіс. Әрбір мәдени элемент өзінің тарихи-мәдени контексімен бірге ұсынылуы қажет. Жалпы, ұлттық мұраны мәдениетаралық интеграциялау – жаһандану жағдайында ұлттық бірегейлікті сақтай отырып, оны әлемдік мәдени кеңістікке енгізудің тиімді жолы. Бұл үдеріс халықтар арасындағы рухани байланысты нығайтып, мәдени әртүрлілікті сақтауға және дамытуға мүмкіндік береді. Ұлттық мұраны халықаралық деңгейде таныту арқылы әрбір халық өз мәдениетін жаһандық өркениеттің ажырамас бөлігі ретінде көрсете алады.

Көптеген қолжазбалар араб, көне түркі, шағатай жазуымен жазылған. Оларды оқу үшін арнайы палеографиялық және филологиялық дайындық қажет. Бұл зерттеушілер шеңберін тарылтады. Қолжазбаларды цифрландыру ісінің дұрыс жүргізілуі өте маңызды. Өйіткені, қолжазбаларға бірінші транслитерация жасалу керек, содан кейін аудармасы жасалады. Әрине, білікті мамансыз транслитерация мен аударма жасау мүмкін емес. Сонымен қатар, мәтіндердің физикалық жағдайы да өте маңызды фактор. Көптеген қолжазбалар уақыт өте келе бүлінген, толық емес немесе оқылуы қиын күйге түскен. Бұл олардың

ғылыми айналымға толыққанды енгізілуіне кедергі келтіреді. Терминологиялық және аударма мәселелері де күрделі үдеріс. Қолжазбалардағы ұғымдарды қазіргі тілге дәл аудару әрдайым мүмкін емес. Кейбір мәдени-тарихи реалийлердің басқа тілде баламасы жоқ, бұл мәдениетаралық түсінуде қиындық туғызады. Келесі, қолжетімділік мәселесі. Қолжазбалардың көп бөлігі шетел архивтерінде, сирек қорларда сақталған. Оларға қол жеткізу құқықтық, қаржылық және институционалдық шектеулерге байланысты қиын болуы мүмкін. Сондай-ақ, интерпретация субъективтілігі де маңызды іс. Әртүрлі зерттеушілер бір мәтінді әрқалай түсіндіруі мүмкін, бұл ұлттық мұраны бұрмалап көрсету қаупін тудырады. Төмендегі мәтінде «шырын» сөзі бір жолында *ширин,* келесі жолда *шрин,* басқа бір жолында *шрин* деп берілген (3-сурет).

Мұны контестке қарай дұрыс оқып, аударуымыз керек.



3-сурет. *«Ибраһим ибн әдһам» қиссасының бірінші беті. әл-Фараби атындағы ҚазҰУ кітапханасы*

Төменде қиссаның транскрипциясы мен аудармасын қоса бердік.

1/a. Ĭşitkĭl sâmiʿā üşbū ḥikāyet	1/b. qīlāy bir nükte-i şĭrĭn rivāyet
1/a. Есітіңіз, тыңдаушы бұл хиқая	1/b. Айтайын керемет мағыналы оқиға

2/a. Bār irdi Bālx īlīniñ pādşāhı 2/a. Бар еді Балқы елінің падишахы	2/b. ūşal īqlīm öze irdi sipāhı 2/b. Бүкіл елі өзіне еді әскер
3/a. Müsülmān pādşāh bā âdālet 3/a. Мұсылман падишах	3/b. tāpīb irdi râiyyetlēr refāhet 3/b Халқы бақытты, бақуатты еді
4/a. İdi ḥükm revān çün baḥr-i mevvāc 4/a. Берген әмірі аққан теңіздей тез жететін	4/b. ālur irdi Xatāy Hind*dīn bāc 4/b. Алатын еді қытай мен үндіден салық
5/a. Ayālīdīn bār irdi bir qīz zībā 5/a. Әйелінен бар еді бір сұлу қыз	5/b. oġul ferzendin yoq irdi āslā 5/b. Ұл перзенті жоқ еді мүлде
6/a. Camālıdīn peri reşk iltēr irdi 6/a. Сұлулығын перінің өзі қызғанар еді	6/b. şuāʿıdīn anı kün qāytār irdi 6/b. (жүзінің) жарықтығынан күннің өзі ұялып кері қайтатын
7/a. İdi nāzūkligi çün ḥur âynı 7/a. Нәзіктігі хор қызындай еді	7/b. sözi şirin çvkvyā* engübini 7/b. Сөзі шырын тәтті балдай еді
8/a. Yüzidin burqâ ālsa māh peyker* 8/a. Жүзінің сұлулығынан ай көрінбей Қалатын	8/b. qarāṅġūnı qılur irdi munevver* 8/b. Қараңғыны жарық қылатын
9/a. Camāl u şıfını āysām tükenmes 9/a. Сұлулығын айтсам таусылмас	9/b. ki bitīsēm defter kāġaẓ*ġa sıġmās ki bir 9/b. Жазсам дәптер қағазға сыймас бір

(Древнетюркский словарь 1969; Eren 2020).

Сирек кітаптар мен қолжазбаларды цифрландыру – жоғары дәлдік пен ғылыми негізді талап ететін күрделі көпсатылы процесс. Бұл жұмысты жүйелі түрде ұйымдастыру үшін оны бірнеше негізгі кезеңдерге бөліп қарастыру қажет. Ең бірінші, цифрландырылатын материалдарды ғылыми тұрғыдан саралап, іріктеу керек. Бұл дайындау кезеңінде құжаттардың тарихи, мәдени және ғылыми құндылығы анықталады, физикалық жағдайы (тозу деңгейі, зақымдануы) тексеріледі, басымдық берілетін нысандар (сирек, бірегей, сұранысқа ие) таңдалады. Сонымен қатар реставрация қажет ететін қолжазбалар алдын ала қалпына келтіру жұмыстарынан өткізіледі. Екіншіден техникалық жағынан толық қамтамасыз етілуі керек. Цифрландыру сапасы қолданылатын жабдыққа тікелей байланысты болады. Сондықтан да, негізгі құралдар яғни, планетарлық сканерлер (қолжазбаларды зақымдамай түсіру үшін), жоғары ажыратымдылықтағы камералар, жарықтандыру жүйесі (ультракүлгін әсерсіз) дұрыс болу керек. Бұл кезеңде сақтау стандарттары (температура, ылғалдылық) да қатаң сақталуы тиіс. Негізгі процесс сканерлеу және цифрлық көшірме жасау кезеңінде жүзеге асады. Әрбір бет жоғары сапада (әдетте 300–600 dpi) сканерленеді, түпнұсқаның түсі, фактурасы дәл берілуі қамтамасыз етіледі, файл форматтары таңдалады (TIFF – архив үшін, PDF/JPEG – қолданыс үшін). Қолжазбаларда беттердің реттілігі мен толықтығы ерекше бақылауда болуы керек. Жұмыс барысының келесі кезеңінде мәтінді тану және өңдеу (OCR) ісі тұрады. Егер мәтін баспа түрінде болса, OCR (Optical Character Recognition) технологиясы арқылы мәтін цифрлық мәтінге айналдырылады және қателерді

түзету үшін редакциялау жүргізіледі. Ал көне жазбалар (араб графикасы, шағатай тілі және т.б.) үшін қолмен транскрипция жасап, қосымша ғылыми түсіндірме жазу қажет. Келесі кезекте метадеректер жасалады. Әрбір цифрлық нысан ғылыми сипаттамамен қамтамасыз етіледі. Мәселен, автор, атауы, жазылған уақыты, тілі, жазу түрі, тақырыптық сипаттамасы, қысқаша аннотациясы беріледі. Бұл деректер іздеу мен ғылыми айналымға енгізу үшін өте маңызды. Цифрланған материалдар электрондық кітапханалар, институционалдық репозиторийлер, бұлттық сақтау жүйелері сияқты арнайы жүйелерге енгізіледі. Сонымен қатар, резервтік көшірмелер жасалады, деректердің ұзақ мерзімді сақталуы қамтамасыз етіледі. Цифрландырудың негізгі мақсаты – қолжетімділікті қамтамасыз ету. Қолжетімділік дегеніміз цифрланған материалдарды ғылыми порталдарда жариялау, халықаралық дерекқорларға енгізу. Бұл кезең мәдениетаралық интеграцияға тікелей ықпал етеді. Әрине, цифрланған материалдардың авторлық құқық, мәдени мұраны қорғау нормалары, сирек қорларға қатысты шектеулер, т.б. құқықтық және этикалық мәселелері міндетті түрде ескеріледі.

Сирек кітаптар мен қолжазбаларды цифрландыру тек техникалық процесс емес, ол ғылыми сипаттау, мәдени мұраны сақтау және оны жаһандық ғылыми кеңістікке енгізу міндеттерін қамтитын кешенді жүйе. Бұл жұмысты тиімді жүргізу үшін кітапханалар, архивтер, ғылыми орталықтар және IT-мамандар арасындағы пәнаралық ынтымақтастық аса маңызды.

Қорытындылай келгенде, сирек кітаптар мен қолжазбаларды цифрландыру – ұлттық мәдени мұраны сақтаудың заманауи әрі тиімді тәсілі ғана емес, сонымен қатар мәдениетаралық интеграцияны дамытудың маңызды құралы екендігін баса айта аламыз. Цифрлық технологиялар арқылы тарихи жазба мұралар ғылыми айналымға кеңінен енгізіліп, түрлі елдердің зерттеушілері арасында мәдени және ғылыми байланыстар нығая түсері анық. Сондықтан мәдени мұраларды цифрландыру жұмыстары ұлттық кітапханалар, архивтер, ғылыми орталықтар және халықаралық ұйымдардың бірлескен жобалары арқылы жүйелі түрде жүзеге асырылуы қажет. Бұл үдеріс ұлттық құндылықтарды сақтаумен қатар, оларды әлемдік мәдени кеңістіктің ажырамас бөлігі ретінде танытуға мүмкіндік береді.

Әдебиеттер

Eren 2020 – Eren H. Türk Dilinin Etimolojik Sözlüğü (ETDES) / Hazırlayan: Şükrü Haluk Akalın. Ankara: Türk Dil Kurumu Yayınları, 2020.Геродот 1972 – Геродот. История. М., 1972. 600 с.

Древнетюркский словарь 1969 – Древнетюркский словарь [Текст]/Акад. наук СССР; Ин-т языкознания; [ред.: В. М. Наделяев [и др.]] Ленинград: Наука. Ленингр. отд-ние, 1969. 676 с.

Жолдасбеков 2012 – Жолдасбеков М. Асыл арналар. I том. Зерттеулер. Мақалалар. Астана: Күлтегін, 2012.

Кадримбетова 2020 – Кадримбетова Н.Н. Тарихи-мәдени мұра ескерткіштерінің сақталу тарихы. Шәкәрім Университетінің Хабаршысы. Техникалық ғылымдар сериясы. 2020. (3(91). 353-358 б.

Оттенсон 1985 – Оттенсон А. Памятники культуры во все времена были объектом охраны государства. // Советское право. 1985. № 3. С. 192-198.

АРХИВЫ И КРЕАТИВНЫЕ ИНДУСТРИИ: СТРАТЕГИЯ ОТКРЫТОСТИ И СОЦИАЛЬНОЙ ПАМЯТИ

Шолпан Валерьевна Тлепина
Доктор юридических наук, профессор, заведующий сектором перспективных исследований Центра научных и перспективных исследований, Архив Президента Республики Казахстан, Казахстан
sh.tlepina@aprk.gov.kz; https://orcid.org/0000-0002-6586-3515

Аннотация. Современные архивы представляют собой практическую, научно-исследовательскую систему комплектования, хранения, обработки и предоставления доступа к документам и информации, имеющим историко-культурную, общественную, юридическую, иную ценность. Перед архивами стоит задача цифровой трансформации в новую форму ее организации как цифрового архива. Цифровая трансформация обусловлена и обеспечивается инновациями индустрий информатики (информационных технологий), телекоммуникаций (телекоммуникационных технологий) и других базирующихся на них креативных индустрий.

Цифровая трансформация открывает перед архивами новые горизонты, позволяя им стать более доступными, востребованными и значимыми в информационном обществе. Оцифровка документов, внедрение современных информационных систем и использование инноваций (искусственный интеллект, блокчейн и др.) становятся неотъемлемой частью работы. Это позволяет не только надежно сохранить уникальные материалы, но и сделать их доступными для широкой аудитории, обеспечивая выполнение задач с новым цифровым качеством. Кроме того, архивы становятся сырьевой базой для креативной экономики.

Динамическая экосистема современных архивов сохраняет данные, генерирует ценность, обеспечивает доверие и строит будущее. Потенциал открытых архивов для общества будет расти пропорционально количеству оцифрованных документов: чем больше документов будет в цифровом виде, тем выше будет количество использований документов, архивных проектов, увеличится количество вовлеченных граждан в науку,

архивную практику – сбор и описание архивных документов, креативное использование архивных документов. Доклад направлен на рассмотрение вариантов коллаборации архивов и общества – киноиндустрия, цифровые медиа-хабы, архивирование «эфемерной» цифровой реальности, «цифровых волонтеров» и др.

Ключевые слова: Казахстан, Международный совет архивов, Архив Президента Республики Казахстан, «Открытому обществу – открытые архивы», цифровизация, искусственный интеллект, III Конгресс архивистов, «Архивы для общества», креативная индустрия, цифровое наследие, «цифровые волонтеры», Web-контент

АРХИВТЕР ЖӘНЕ КРЕАТИВТІ ИНДУСТРИЯЛАР: АШЫҚТЫҚ ЖӘНЕ ӘЛЕУМЕТТІК ЖАДЫ СТРАТЕГИЯСЫ

Шолпан Валерийқызы Тлепина

Заң ғылымдарының докторы, профессор, Ғылыми және перспективалық зерттеулер орталығының перспективалық зерттеулер секторының меңгерушісі, Қазақстан Республикасы Президентінің Архиві, Қазақстан

sh.tlepina@aprk.gov.kz; https://orcid.org/0000-0002-6586-3515

Аңдатпа. Қазіргі заманғы архивтер – тарихи-мәдени, қоғамдық, құқықтық және өзге де құндылығы бар құжаттар мен ақпаратты жинақтау, сақтау, өңдеу және оларға қолжетімділікті қамтамасыз етудің практикалық және ғылыми-зерттеу жүйесі. Архивтердің алдында оларды цифрлық архив ретінде ұйымдастырудың жаңа формасына көшу, яғни цифрлық трансформация міндеті тұр. Цифрлық трансформация информатика (ақпараттық технологиялар), телекоммуникация (телекоммуникациялық технологиялар) индустриялары және оларға негізделген креативті индустриялардың инновациялары арқылы жүзеге асады және қамтамасыз етіледі.

Цифрлық трансформация архивтерге жаңа көкжиектер ашып, олардың ақпараттық қоғамда қолжетімді, сұранысқа ие және маңызды болуына мүмкіндік береді. Құжаттарды цифрландыру, қазіргі заманғы ақпараттық жүйелерді енгізу және инновацияларды (жасанды интеллект, блокчейн және т.б.) пайдалану жұмыстың ажырамас бөлігіне айналуда. Бұл бірегей материалдарды сенімді сақтап қана қоймай, оларды қалың көпшілікке қолжетімді етуге және міндеттерді жаңа цифрлық сапада орындауға мүмкіндік береді. Сонымен қатар, архивтер креативті экономика үшін шикізат базасына айналуда.

Қазіргі архивтердің динамикалық экожүйесі деректерді сақтайды, құндылық жаратады, сенімділікті қамтамасыз етеді және болашақты құрады. Ашық архивтердің қоғам үшін әлеуеті цифрланған құжаттар санына пропорционалды түрде өседі: құжаттар цифрлық түрде неғұрлым көп болса, оларды пайдалану көрсеткіші мен архивтік жобалар саны соғұрлым артады, сондай-ақ ғылым мен архив практикасына – архив құжаттарын жинау және сипаттау, архив құжаттарын креативті пайдалану ісіне тартылған азаматтар саны көбейеді. Баяндама архивтер мен қоғамның өзара ынтымақтастық нұсқаларын – киноиндустрия, цифрлық медиа-хабтар, «эфимерлік» цифрлық болмысты архивтеу, «цифрлық волонтерлер» және т.б. қарастыруға бағытталған.

Түйін сөздер: Қазақстан, Халықаралық архивтер кеңесі, Қазақстан Республикасы Президентінің Архиві, «Ашық қоғамға – ашық архивтер», цифрландыру, жасанды интеллект, III Архившілер конгресі, «Қоғам үшін архивтер», креативті индустрия, цифрлық мұра, «цифрлық волонтерлер», Web-контент

ARCHIVES AND CREATIVE INDUSTRIES: STRATEGY OF OPENNESS AND SOCIAL MEMORY

Sholpan V. Tlepina
Doctor of Law, Professor, Head of the Advanced Research Sector,
Archive of the President of the Republic of Kazakhstan, Kazakhstan
sh.tlepina@aprk.gov.kz; https://orcid.org/0000-0002-6586-3515

Abstract. Modern archives are a practical, scientific and research system for acquiring, storing, processing and providing access to documents and information of historical, cultural, social, legal, and other value. Archives face the task of digital transformation into a new form of its organization as a digital archive. Digital transformation is conditioned and provided by innovations in the industries of computer science (information technology), telecommunications (telecommunication technologies), etc., and creative industries based on them.

Digital transformation opens up new horizons for archives, allowing them to become more accessible, in demand and significant in the information society. Digitization of documents, the introduction of modern information systems and the use of innovations (artificial intelligence, blockchain, etc.) are becoming an integral part of the work. This allows not only to reliably preserve unique materials, but also to make them accessible to a wide audience, ensuring that tasks are performed with new digital quality. In addition, archives are becoming a raw material base for the creative economy.

The dynamic ecosystem of modern archives preserves data, generates value, provides trust, and builds the future. The potential of open archives for society will grow in proportion to the number of digitized documents: the more documents there are in digital form, the higher the number of uses of documents and archival projects will be, the number of citizens involved in science and archival practice will increase – the collection and description of archival documents, the creative use of archival documents. The report is aimed at considering options for collaboration between archives and society – the film industry, digital media hubs, archiving of "ephemeral" digital reality, "digital volunteers", etc.

Keywords: Kazakhstan, International Council of Archives, Archive of the President of the Republic of Kazakhstan, "Open Society – open archives", digitalization, artificial intelligence, III Congress of Archivists, "Archives for Society", creative industry, digital heritage, "digital volunteers", Web content

Введение: новая парадигма архивного дела в цифровую эпоху

Уважаемые модераторы, коллеги и участники Конгресса!

Тема нашей сессии – «Архивы для общества» – неразрывно связана с обеспечением доверия в цифровом мире. Современный архив – это катализатор коллаборации, где историческая правда встречается с современными технологиями в архивоведении, традиционном творчестве, в отдельных направлениях креативной индустрии.

Особое внимание хотела бы уделить роли архивов в развитии креативных индустрий и формировании качественного информационного пространства.

Наша сессия как свидетельство открытости и инклюзивности кураторов выставок, библиотекарей, музейных профессионалов и архивистов, которые неустанно работают над сохранением и обеспечением доступа к архивам.

Мы находимся в исторической точке трансформации, когда архивная отрасль Казахстана переходит от консервативной функции «статичного хранения» к активной роли «генератора национальных смыслов». Тема нашей сессии – «Архивы для общества» – требует от нас признания архивов, активно формирующих контуры будущего.

Как отмечалось на Конгрессе Международного совета архивов (ICA) (О членстве Республики Казахстан 1996) в Абу-Даби (2023), (Конгресс ICA 2023) архивы сегодня – это не молчаливая или пассивная память, а «голос знаний», который в эпоху глобального информационного шума должен звучать максимально отчетливо, становясь фундаментом для развития креативной экономики и укрепления гражданской идентичности. Этот голос должен быть интегрирован во все сферы общественной жизни, от школьных учебников до сложнейших алгоритмов искусственного интеллекта.

Наша стратегия открытости – это не просто механическое упрощение доступа к делам в читальных залах, а создание новой архитектуры взаимодействия, где архивный документ становится ликвидным активом для кино, медиа, IT-стартапов, образовательных платформ и др. Мы должны построить бесшовный интеллектуальный мост между академическим хранилищем и динамичными потребностями современного творческого сообщества. В этом контексте лозунг «Открытому обществу – открытые архивы» стал со дня I Конгресса главной дорожной картой на ближайшее десятилетие, определяющей успех цифровой трансформации Архива Президента Республики Казахстан. Глава государства поддержал создание фонда «Национальное цифровое наследие» («Национальное цифровое наследие» 2026). Такая инициатива в рамках Года цифровизации и искусственного интеллекта (2026) (Указ Президента РК 2026) открывает новые возможности для сохранения

и глубокого изучения исторического наследия (К.К. Токаев на V заседании Национального курултая, Кызылорда) (Выступление Главы государства 2026). Мы стремимся к тому, чтобы каждый архивный фонд стал открытой лабораторией для инноваций, где исторический факт обретает новую жизнь в современных форматах.

1. Архивы как стратегическая сырьевая база для креативной экономики

Международный тренд «Познавая прошлое, создаем будущее» провозглашает архивы первичной и незаменимой базой для индустрии интеллектуальной собственности. В Казахстане креативный сектор – это не только искусство, но и мощный драйвер ВВП, который остро нуждается в достоверном, глубоком и визуально насыщенном историческом «сырье».

Глава государства отмечал, что Казахстан вышел на передовые позиции в области креативной индустрии (Токаев 2022). Президент высоко оценил успехи, которые молодые специалисты демонстрируют в этой сфере. Он подчеркнул, что поддержка креативного сектора сегодня является одной из важнейших задач государства.

1.1. Киноиндустрия: от исторической декорации к документальной метавселенной

Мировая практика (успешные кейсы студий Netflix, HBO (Home Box Office), Би-би-си (British Broadcasting Corporation) наглядно доказывает, что коммерческий успех и международное признание исторических проектов напрямую зависят от глубины и качества архивных изысканий. Современный зритель, обладая доступом к огромным массивам информации, стал крайне требователен к деталям, и любая историческая неточность или анахронизм мгновенно подрывают авторитет кинопроизведения.

Сотрудничество с «Казахфильмом» и частными продакшн-центрами. При создании масштабных исторических полотен, таких как «Қазақ хандығы. Алтын тақ» или сериал «Қасым хан», творческие группы часто сталкиваются с острым дефицитом верифицированных подробностей материальной культуры, дипломатического протокола и нюансов повседневности прошлых веков.

Мы предлагаем официально внедрить в Казахстане систему «Архивного сопровождения кинопроизводства», которая станет стандартом качества для отрасли. В рамках этой модели эксперты, сотрудники соответствующих Архивов, будут выступать не просто как технические сотрудники, а как полноценные научные консультанты и креативные продюсеры по исторической части, помогая воссоздавать дух времени на основе подлинных чертежей, уникальных фотографий и ранее не публиковавшихся архивных записей. Это позволит поднять планку отечественного исторического кино на мировой уровень.

Документалистика и цифровые медиа-хабы. Проект М. Бекбаевой «Тайны. Судьбы. Имена» демонстрируют широкой аудитории, как один случайно найденный в фондах документ может стать триггером для создания захватывающего фильма-расследования. Мы должны инициировать создание специализированных цифровых хабов для документалистов, где кино- и фотодокументы будут доступны для предварительного просмотра и лицензирования в режиме «одного окна».

Это превратит архив из закрытой бюрократической инстанции в гибкого и современного партнера творческого процесса, что, в свою очередь, обеспечит взрывной рост цитируемости наших национальных фондов в глобальном медиапространстве. Такие хабы станут платформой для обмена идеями между историками и визуализаторами.

1.2. Работа с лидерами мнений: кто формирует память поколения Z?

В современных реалиях социальные сети (TikTok, Instagram, YouTube, др.) фактически подменили собой традиционные учебники для молодежи, став основными источниками получения исторических знаний. Если государственные архивы не предложат лидерам мнений качественный, адаптированный и верифицированный материал, этот вакуум неизбежно заполнится псевдоисторическими фейками, манипуляциями и опасными спекуляциями.

Кейс: Яков Федоров. Известный блогер-исследователь активно путешествует по самым отдаленным регионам Казахстана, работая с локальной историей, старыми картами и уникальными архитектурными планами. Системное сотрудничество архивов с такими авторами позволяет эффективно доносить ценность документального наследия до сотен тысяч зрителей, которые по собственной воле никогда бы не переступили порог архивного учреждения. Мы рассматриваем возможность предоставления топ-блогерам эксклюзивного доступа к наиболее интересным фондам личного происхождения и картам, чтобы их контент становился максимально глубоким, авторитетным и просветительским. Это сотрудничество выгодно обеим сторонам: блогер получает уникальный контент, а архив – колоссальный охват аудитории.

Кейс: Диас Камериданов – автор проекта «Неизвестный Казахстан». Его контент, построенный на наглядном сравнении архивных фотографий городов (Алматы, Семей, Астана) с современными видами, вызывает невероятный эмоциональный отклик у подписчиков всех возрастов. Архивы должны занять проактивную позицию и сами инициировать подобные проекты, формируя готовые «медиа-пакеты», включающие редкие фото и краткие, но яркие исторические справки для инфлюенсеров. Такая стратегия позволит создать вокруг архива «доверительную экосистему знаний», где популярный формат изложения всегда опирается на строгий, документально подтвержденный научный факт. Мы должны стать главным поставщиком «интеллектуальной моды» на знание своей истории.

Кейс: Образовательные IT-проекты и просветительство. Интеграция архивных массивов в современные мобильные приложения и образовательные платформы через популярных просветителей – это самый короткий путь к формированию «архивного интеллекта» у подрастающего поколения. Мы должны не просто предоставлять данные, а приглашать авторов к совместному проектированию цифровых продуктов, делая архивы технологичными и востребованными площадками для самоидентификации молодежи. Это позволит превратить архивную информацию в активную часть образовательного процесса, доступную в один клик.

2. Технологическая трансформация: искусственный интеллект и большие данные

Конгресс в Абу-Даби 2023 года провозгласил фундаментальный тезис: «В конечном итоге вся документация и человеческая память станут цифровыми» (Конгресс ICA 2023). Казахстан должен занять лидирующие позиции в этом процессе в Центральноазиатском регионе, внедряя передовые стандарты обработки данных.

2.1. Искусственный интеллект как главный ассистент современного архивиста

Внедрение технологий искусственного интеллекта в архивную практику – это не угроза вытеснения человека, а единственный способ справиться с лавинообразным ростом информации и многократно повысить эффективность нашей работы. ИИ позволяет нам автоматизировать те процессы, которые ранее требовали десятилетий кропотливого ручного труда целых коллективов.

Распознавание рукописей (HTR) и лингвистические модели. Мировой опыт (проект Transkribus) показывает, что современные алгоритмы машинного обучения способны распознавать сложнейшие тексты – от средневековой арабской вязи до скорописи XIX века – в тысячи раз быстрее самого квалифицированного палеографа. Нам необходимо форсировать внедрение подобных инструментов для массовой индексации национальных фондов Казахстана, превращая «мертвые» массивы сканов в живые «умные данные», доступные для мгновенного полнотекстового поиска. Это откроет перед историками и лингвистами возможности, о которых раньше можно было только мечтать.

Нейросетевая реставрация и колоризация. Использование ИИ для восстановления поврежденных фотопленок и колоризации черно-белой

хроники делает историю визуально понятной и эмоционально близкой для «цифровых аборигенов». Это мощный инструмент, который позволяет архиву эффективно конкурировать за внимание аудитории в агрессивной цифровой среде, превращая сухой документ в яркое визуальное событие. Ожившие образы прошлого помогают восстановить связь поколений на интуитивном уровне.

2.2. Архивирование «эфемерной» цифровой реальности (Web Archiving)

На прошедшем Конгрессе в г. Барселоне (2025) одной из ключевых тем заявлено сохранение «цифрового наследия» (Digital Heritage), которое исчезает быстрее, чем бумага. На сессии, которую в г. Барселоне возглавила директор Архива Президента Республики Казахстан А.К. Мустафина, участники представили результаты исследований о влиянии искусственного интеллекта на архивную профессию и сохранение цифрового наследия (АП РК на Конгрессе ICA Barcelona 2025). Мы сталкиваемся с беспрецедентным вызовом: как сохранить культуру, которая рождается и существует только в битах и байтах.

Современная общественная и культурная жизнь Казахстана сегодня фиксируется преимущественно в социальных сетях, мессенджерах и на динамичных веб-сайтах. Если мы не разработаем государственные стандарты захвата, каталогизации и долгосрочного хранения Web-контента прямо сейчас, будущие исследователи столкнутся с катастрофической «цифровой амнезией» в истории независимого Казахстана.

Нам жизненно необходимы инновационные партнерства с национальным IT-сектором для создания сверхзащищенных кибер-архивов, способных гарантировать сохранность данных на протяжении столетий, независимо от смены форматов и носителей. Это вопрос обеспечения информационного суверенитета страны.

3. Социальная миссия: стратегия «открытого архива» и гражданская наука

Архивы должны окончательно выйти из тени закрытых ведомств и стать полноценным пространством для «Citizen Science» (гражданской науки), где каждый гражданин осознает себя активным участником сохранения коллективной памяти.

3.1. Краудсорсинг и институт «Цифровых волонтеров»

Мы предлагаем инициировать запуск масштабной национальной платформы волонтерства «Память нации», объединяющей усилия государства и общества. На этой платформе студенты, краеведы и просто неравнодушные граждане смогут дистанционно участвовать в транскрибации и индексации оцифрованных архивных документов, помогая делать их доступными для поиска.

Такой опыт уже успешно реализован в крупнейших архивах мира (США, Франция, др.) и показал свою невероятную эффективность как в плане обработки данных, так и в плане патриотического воспитания. Вовлечение молодежи в этот процесс создает прямую эмоциональную связь между личной историей семьи и большой историей государства, превращая архив в общенародное достояние.

3.2. Архивы как антидот против «инфодемии» и исторических манипуляций

В современном мире, где фейковые новости и дезинформация распространяются со скоростью света, архив остается единственным «золотым эталоном» объективной истины. Наша стратегия открытости включает в себя агрессивную просветительскую политику и активное присутствие в медиаполе.

Мы должны не просто пассивно ждать исследователя в читальном зале, а сами выходить на передовую информационной борьбы, публикуя документальные опровержения популярных мифов и спекуляций. Архив должен стать главным и бесспорным верификатором информации для журналистов, учителей и государственных органов, обеспечивая фундаментальную информационную безопасность нашего общества.

4. Будущее профессии: компетенции, экология и устойчивость

Согласно видению ICA, архивист будущего – это сложный сплав медиаменеджера, дата-аналитика, юриста и хранителя смыслов (New Professionals Programme). Нам необходимо радикально пересмотреть национальные программы профессиональной подготовки кадров, чтобы соответствовать этим вызовам.

Инновационные партнерства и ГЧП. Мы должны активно развивать механизмы государственно-частного партнерства в архивной сфере. Привлечение частных инвестиций и IT-компетенций для оцифровки фондов и создания коммерчески успешных креативных продуктов – это единственный путь к технологическому лидерству и финансовой устойчивости отрасли. Мы открыты к сотрудничеству с бизнесом, понимая, что культурное наследие может быть мощным экономическим активом.

Экология и «зеленые» стандарты хранения. В соответствии с глобальной повесткой устойчивого развития (Sustainable Development Goals – SDG), архивы обязаны минимизировать свой экологический след. Проектирование новых архивных зданий и энергоемких дата-центров должно осуществляться с применением «зеленых» технологий и принципов энергоэффективности. Сохранение памяти о прошлом не должно ставить под угрозу экологическое будущее нашей планеты и Казахстан должен быть готов представить свои передовые решения.

Заключение: переход к модели «архивы 3.0»

Реализация масштабной стратегии «Архивы и креативные индустрии» – это наш решающий вклад в прогрессивное развитие современного Казахстана. Мы окончательно уходим от устаревшей модели «архива-склада» и строим современную модель «архива-хаба», где рождаются прорывные идеи, снимаются глубокие фильмы и воспитывается ответственное мыслящее поколение.

Коллаборация с творческим сообществом, системное партнерство с блогерами и внедрение технологий искусственного интеллекта – это не просто дань глобальной моде, а вопрос выживания и процветания нашей национальной идентичности в эпоху тотальной цифровизации. Наш авторитет на международной арене и успех на будущих форумах ICA будут напрямую зависеть от того, насколько открытыми смелыми и технологичными мы станем уже сегодня.

Литература

New Professionals Programme – Программа новых профессионалов. URL: https://www.ica.org/professional-programme/new-professionals-programme/ (дата обращения: 06.06.2026).

АП РК на Конгрессе ICA Barcelona 2025 – Архив Президента представляет опыт Казахстана на Конгрессе ICA Barcelona 2025 URL: https://www.inform.kz/ru/arhiv-prezidenta-rk-predstavlyaet-opit-kazahstana-na-kongresse-ica-barcelona-2025-a98bdf (дата обращения: 05.06.2026).

Выступление Главы государства 2026 – Выступление Главы государства Касым-Жомарта Токаева на V заседании Национального курултая. URL: https://www.akorda.kz/ru/vystuplenie-glavy-gosudarstva-kasym-zhomarta-tokaeva-na-v-zasedanii-nacionalnogo-kurultaya-200332?ysclid=mqhyx3akvi332806272 (дата обращения 01.06.2026).

Конгресс ICA 2023 – Конгресс ICA в Абу-Даби 2023. URL: https://chatru.com/index.php?topic=242602.0&ysclid=mqhyb7v52p599375258 (дата обращения: 21.05.2026).

«Национальное цифровое наследие» 2026 – Президент поддержал идею создания фонда «Национальное цифровое наследие» URL: https://abaidan.kz/ru/categories/kr-prezidenty/prezident-podderzhal-ideyu-sozdaniya-fonda-nacionalnoe-cifrovoe-nasledie-2026-01-21?view=yes&ysclid=mqhy8bt2q4979548482 (дата обращения: 01.06.2026).

О членстве Республики Казахстан 1996 – О членстве Республики Казахстан в Международном Совете архивов. Постановлением Правительства РК от 20 августа 1996 г. URL: https://adilet.zan.kz/rus/docs/P940000728_ (дата обращения: 21.05.2026).

Токаев 2022 –Токаев: Казахстан вышел на передовые позиции в области креативной индустрии URL: https://informburo.kz/novosti/tokaev-kazahstan-vyshel-na-peredovye-pozicii-v-oblasti-kreativnoj-industrii?ysclid=mq51ysoxr9187857645 (дата обращения: 08.06.2026)

Указ Президента РК 2026 – Об объявлении Года цифровизации и искусственного интеллекта. Указ Президента Республики Казахстан. URL: https://www.akorda.kz/ru/ob-obyavlenii-goda-cifrovizacii-i-iskusstvennogo-intellekta-601222?ysclid=mqj2w5tl4t86399713 (дата обращения: 18.06.2026).

КОНГРЕСТІҢ САЛТАНАТТЫ ЖАБЫЛУЫ

ТОРЖЕСТВЕННОЕ ЗАКРЫТИЕ КОНГРЕССА

OFFICIAL CLOSING CEREMONY OF THE CONGRESS

КОНГРЕСТІҢ ҚОРЫТЫНДЫЛАРЫ ТУРАЛЫ

Қ.Ш. Әлімғазинов
Қазақстан Республикасы Президенті Архиві директорының орынбасары, тарих ғылымдарының докторы

Құрметті қатысушылар, форумымыздың қонақтары, әріптестер!

Конгресс жұмысының сессиялар бөлінісіндегі қысқаша қорытындыларын экранға шығаруға рұқсат етіңіздер. Конгресс барысында 1 Пленарлық сессия; мазмұнды 5 панельдік сессия; TED Talks форматындағы «MURA» сарапшылар клубының отырысы; ЮНЕСКО-ның халықаралық семинары (120 адам қатысты); Эндрю Пейннің семинары (61 адам қатысты); IX Жазғы мектеп (оның шеңберінде 100 адам білім алды); үш воркшоп және 42 экспоненттің жетістіктері ұсынылған «DA-EXPO» *көрмесі* өткізілді. Панельдік сессияларға 15 модератор, 65 спикер және 4 халықаралық сарапшы қатысты.

Конгресс күндері архив қауымдастығының парадигманы өзгертуге, яғни цифрлық технологиялардың енжар тұтынушысы болудан жасанды интеллекттің (AI) бейімделген технологияларын жасаушы әрі оны сенімді пайдаланушы белсенді субъектілерге көшуге деген ұмтылысын айқын көрсетті. Бұл ретте Конгресс архивтердің тарих қоймасы ретіндегі өз мандатын мүлтіксіз орындау жөніндегі өзгермейтін мәнін тағы бір мәрте дәлелдеді. Өз бастауларыңды тану ұлттық дамудың көкжиегін кеңейтеді. Қазақстан Республикасының жаңа Конституциясында (2026 жылғы 15 наурыз) осы қағидат терең сезіне отырып, мәдени мұраны сақтау, білім, ғылым мен инновацияларды дамыту, сондай-ақ азаматтардың цифрлық құқықтарын қорғау стратегиялық басымдық ретінде нақты белгіленген.

Панельдік пікірталастардың нәтижелерін талдай келе, біз Конгресс шеңберінен шығып, кең ауқым алған келесідей негізгі тұжырымдарды айқындай аламыз:

№ 1 сессия. «Архивтер және AI: құралдар, этика және болашақтың стратегиялары»

Сессия архив саласының сарапшыларын, халықаралық ұйымдардың өкілдерін және әлемнің 8 елінен – Қазақстан, Әзербайжан, Ирландия, Біріккен Араб Әмірліктері, Ресей, Сингапур, АҚШ пен Түркиядан келген цифрлық технологиялар мамандарының басын қосты.

Сессия аясында архивтік массивтерді зияткерлік жіктеу мен ағынды өңдеудің негізгі құралы ретінде жасанды интеллектті енгізудің ұлттық стратегияларын әзірлеу бағыты айқындалды. Қатысушылар цифрландырудың этикалық аспектілерін егжей-тегжейлі талдап, тарихи шындықты сақтау және алгоритмдік бұрмалауларды болдырмау үшін AI технологияларын енгізу деректерді тексерудің (верификациялаудың) қатаң хаттамаларын жасаумен қатар жүруі тиіс екенін атап өтті.

Тарихи зердені сақтау технологиялары туралы түсініктер үнемі эволюцияны бастан кешіріп келеді: дәстүрлі қоғам жағдайында қағаз тасымалдаушы ең озық технологиялық шешім болды. Бұған ең алғашқылардың бірі болып Ш.Ш. Уәлиханов назар аударған еді. Қазақ халқының бірегей перзенті, ерекше ғалым, зерттеуші, саяхатшы әрі жаңалық ашушының ғылыми мұрасы ұлттық археографиялық дәстүрдің негізін қалады.

Қазіргі таңда архивтерді интеллектуалдандырудың осы ұлттық тарихи ретроспекциясы архивтік ой-пікір дамуының кезекті кезеңін бастан өткеруде. Бұл жаңа кезең жасанды интеллекттің технологиялық құралдарына негізделген.

Архив ісіне ЖИ-ді енгізу дәстүрлі тәсілдерді қайта қарауды және мұқият құқықтық реттеуді талап ететін маңызды этикалық және құқықтық міндеттермен қатар жүріп отыр.

№ 2 сессия. «Digital Archive: заманауи дискурстар»

Екінші панельдік сессияның жұмысына 8 мемлекеттен (Армения, Тәжікстан, Польша, Ұлыбритания, Ресей, АҚШ, Түркия және Қазақстан) келген 9 спикерден тұратын халықаралық құрам қатысты. Сессия қатысушылары архивтерді цифрлық трансформациялау кешенді мемлекеттік тәсілді талап етеді деген ортақ тұжырымға келді. Армения тәжірибесі архив саласын цифрландырудың ұлттық стратегияларын әзірлеудің тиімділігін көрсетсе, Тәжікстан тәжірибесі архив саласын жасанды интеллектті дамытудың мемлекеттік бағдарламаларына интеграциялаудың маңыздылығын дәлелдейді.

Қазіргі заманғы архивтер цифрлық экожүйелерге айналуда. Тәжірибе көрсеткендей, жасанды интеллектті кезең-кезеңімен енгізу, архивтік процестерді автоматтандыру толықтай мүмкін.

Құжаттық, аудиовизуалды және қолжазба мұраларының сақталуын қамтамасыз ету цифрлық форматқа көшірудің (оцифровка) және ұзақ мерзімді цифрлық сақтаудың заманауи технологияларын енгізуді талап етеді.

Архивтерді цифрландыру саласындағы халықаралық ынтымақтастық ортақ тарихи зердені сақтауға және мәдени мұраға қолжетімділікті кеңейтуге ықпал етеді. Бұны Конгресс қатысушыларының түркі әлемінің цифрлық ресурстарын

қалыптастыру жөніндегі жобалары мен Қазақстанның ұлттық мұраны халықаралық ғылыми кеңістікке интеграциялау жөніндегі бастамалары айқын көрсетті.

№ 3 сессия. «Архивтердің инновациялық экожүйелері: инфрақұрылым, ынтымақтастық және тұрақтылық»

Тақырыптық сессия әлемнің 30-дан астам елінен келген архив мекемелерінің 150-ден астам басшылары мен мамандарының, IT-индустрия, ғылыми қоғамдастық және халықаралық ұйымдар өкілдерінің басын қосты.

Талқылау үш негізгі бағытқа шоғырланды:

1. Архивтерді цифрлық трансформациялаудың ұлттық стратегиялары және архивтік деректерді басқаруда экожүйелік тәсілді дамыту;

2. Архивтердің киберқауіпсіздігі, құжаттық мұраны климаттық, биологиялық, техногендік және цифрлық қатерлерден қорғау, сондай-ақ электрондық құжаттардың ұзақ мерзімді сақталуын қаммасыз ету;

3. Жасанды интеллект негізінде архивтік мұраны басқарудың инновациялық технологиялары.

Генеративті жасанды интеллекттің кеңінен таралуы жағдайында цифрлық егемендік, архивтік инфрақұрылымның тұрақтылығы және архивтік құжаттарға деген сенімді сақтау мәселелеріне ерекше назар аударылды.

Пікірталас архивтердің «цифрлық мемлекеттің» жұмыс істеуін қамтамасыз ететін ұлттық деректер инфрақұрылымының іргелі бөлігі ретіндегі рөлін қайта пайымдауға бағытталды. Сарапшылар цифрлық егемендікті қорғау, киберқатерлерге қарсы тұру және физикалық әрі бұлттық сақтау инфрақұрылымы үшін тәуекелдерді азайту мақсатында мемлекеттік архивтер, IT-компаниялар және халықаралық құрылымдар арасында тұрақты серіктестік құру қажеттілігін негіздеді.

№ 4 сессия. «Архивтік болашақ: жаңа құзыреттер мен көшбасшылық»

Сессия қатысушылары – білім беру саласындағы мемлекеттік органдардың (ҚР ҒЖБМ, ҚР жоғары оқу орындары, Жоғары білімді дамыту ұлттық орталығы), «Атамекен» ҰКП өкілдері, Қазақстан мен шетелдік архив мекемелерінің басшылары мен мамандары. Сессияға 5 спикер мен 4 сарапшы қатысты.

Дәстүрлі архивтану ғылымы терең трансформацияны бастан кешуде. Кадрларды даярлау ісі AI құралдарын енгізе отырып, цифрлық экономиканың сұраныстарына сәйкес жедел қайта құрылуы тиіс. Басты тұжырым: қазіргі заманғы маманды даярлау – ортақ жауапкершілік. Жоғары оқу орны нарықтан оқшауланып білім бере алмайды, ал нарық процеске қатыспай, тек дайын «*көктен түскендей дайын мамандарды*» күтіп отыра алмайды.

Бұл мамандық IT, құқық, тарих, деректерді басқару сияқты салалардың тоғысында жаңа білімді үнемі меңгеріп отыруды талап етеді. Пәнаралық тәсіл архившінің табысты кәсіби мансабының жай ғана қалаулы емес, міндетті шартына айналуда.

№ 5 сессия. «Архивтер қоғам үшін»

Қорытынды сессия цифрлық қорлардың ашықтығы азаматтық ғылымның, генеалогиялық зерттеулердің дамуына және тарихи әділдікті қалпына келтіруге қуатты серпін беретінін көрсетті. Талқылау қорытындысы бойынша GLAM халықаралық стандартын енгізу, архивтік волонтерлікті дамыту, сондай-ақ

құжаттық мұра негізінде сапалы ғылыми-танымал контент жасау үшін шығармашылық (креативті) индустриялар мен медиахолдингтермен ауқымды коллаборация құру жөніндегі бастамалар тұжырымдалды.

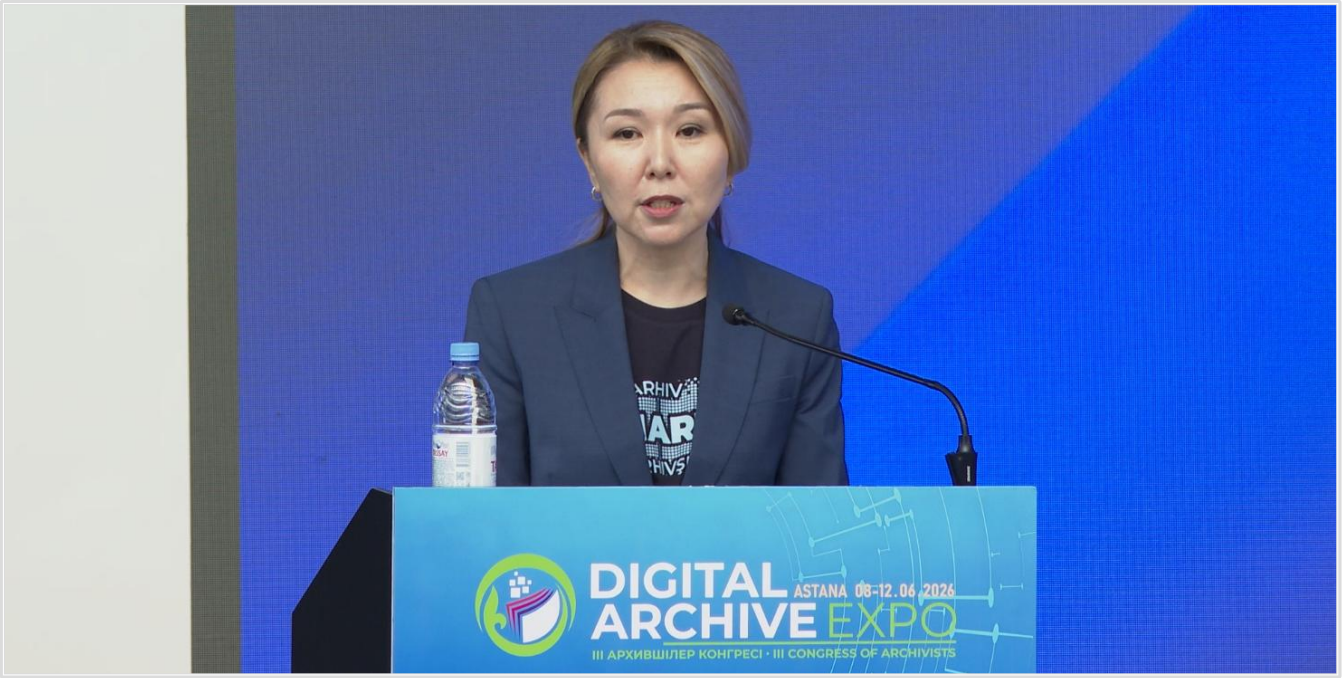
Шығармашылық индустрияларды дамыту қазақстандық контентті цифрлық өнер және мәдени сторителлинг арқылы ілгерілетуді көздейді. Бұған еліміздің мәдени, тарихи және ғылыми құндылықтарын цифрландыруға, сақтауға және жүйелеуге, оны болашақ ұрпаққа жеткізуге бағытталған ауқымды мемлекеттік бастама – «Қазақстанның ұлттық цифрлық мұрасы» қоры ықпал етеді.

Құрметті әріптестер!

Ұсынылған баяндамалар мен пікірталастар архивтерді цифрлық трансформациялау тәсілдерінің алуан түрлілігін және кәсіби міндеттерді шешудегі инновациялық технологиялардың жоғары әлеуетін айқын көрсетті.

Конгрестің барлық жұмысы ортақ мақсатқа – архивтік процестерді жаңғырту, құжаттық мұраның қолжетімділігін арттыру және архивтік қызметтерді сапалы жетілдіру үшін жасанды интеллект мүмкіндіктерін кешенді түсінуді қалыптастыруға бағытталды.

Құрметті әріптестер! Назарларыңызға Конгресс *резолюциясының* жобасын ұсынамын. Ерекше атап өтетін жайт: өз ұсыныстарыңызды, өзгертулер мен толықтыруларыңызды ағымдағы жылдың 1 қарашасына дейін Конгресс мекенжайына жолдауларыңызды сұраймын. Резолюция мәтіні Конгресс материалдарының жинағында жарияланады. Ол ағымдағы жылдың соңына дейін Конгресс сайтында барлық қатысушыларға қолжетімді болады.



ЗАКЛЮЧИТЕЛЬНОЕ СЛОВО

А.Х. Мустафина
Директор Архива Президента Республики Казахстан,
кандидат исторических наук

Құрметті Конгресс қатысушылары!

Осы күндері айтылған барлық тың ойлар, тұжырымдар мен ұсынылған тәжірибелер алдағы жұмысымыздың берік негізіне айналады. Конгресс пікірталастарының ресми түрде жабылуы – осы жасампаз идеяларды іс жүзінде жүзеге асырудың бастауы ғана. Ұйымдастыру комитеті әрдайым байланыста! Біз өзара іс-қимыл мен серіктестікке барынша ашықпыз. Серіктестігімізді одан әрі жалғастыру және ортақ міндеттерді орындау үшін өз ұсыныстарыңыз бен жаңа жобаларыңызды бізге жолдай аласыздар!

Конгресс жұмысын қорытындыламас бұрын, қатысушылардың, баяндамашылардың, халықаралық сарапшылардың, модераторлар мен спикерлердің сертификаттары жақын күндері тіркелу кезінде көрсетілген электрондық пошта мекенжайларыңызға жіберілетінін еске саламын.

Уважаемые участники, гости нашего форума, коллеги!

Прошла пятая панельная сессия Конгресса. Прошли яркие динамичные и живые моменты работы Конгресса. Мой коллега только что подвел итоги Конгресса. Хочу отметить, что они превзошли ожидания. И это благодаря Вам, нашим коллегам и друзьям!

Мы были рады в эти дни принять в г. Астане делегации ООН, ЮНЕСКО, Международного совета архивов и наших коллег из стран СНГ, Европы, Америки, Африки, Ближнего Востока, Восточной и Юго-восточной Азии.

Для нас важно, что интеллектуализация деятельности Архива Президента

была признана Международным советом архивов. Был подписан Меморандум о сотрудничестве. Меморандум, как документ о договоренностях, открывает не только для Казахстана, но и для государств Центральной Азии разнообразные возможности для реализации многостороннего сотрудничества.

Расширена география партнерских отношений с архивными службами ряда государств – участников Конгресса: Великобритании, Узбекистана, Японии, Кореи и др. Верю, наши начинания превратились в функциональное пространство взаимного обмена знаниями, опытом, местом культурного обогащения.

В организации Конгресса нам оказали поддержку Министерство искусственного интеллекта и цифрового развития, Министерство иностранных дел, Министерство культуры и информации РК, акимат г. Астаны, резиденты IT-парка Astana Hub, ведущие ученые и исследователи в области истории, архивоведения и документоведения.

Благодарю наших казахстанских коллег из республиканских и региональных архивов за совместную продуктивную работу, профессиональный диалог и готовность к открытому обмену знаниями и опытом. Все эти дни они вместе с нами принимали активное участие в обсуждениях на площадках Конгресса.

Уверена, что идеи, инициативы и договоренности, достигнутые в эти дни, найдут свое практическое продолжение и будут способствовать дальнейшему укреплению международного профессионального сотрудничества.

Благодарю команду Архива Президента.

Подводя итоги нынешнего Конгресса, мы уже закладываем прочный фундамент и берем четкий ориентир на нашу следующую масштабную встречу – на IV Международный конгресс архивистов, который намечен на 2028 год!

До встречи на следующем IV Конгрессе архивистов 2028 года!

«DIGITAL ARCHIVE. EXPO»
III ХАЛЫҚАРАЛЫҚ АРХИВШІЛЕР КОНГРЕСІНІҢ РЕЗОЛЮЦИЯСЫ
(2026 жылғы 9-11 маусым, Астана қ., Қазақстан)

«Digital Archive Expo» III Архившілер конгресіне қатысушылар осы Резолюцияны қабылдайды және 2028 жылы өтетін кезекті IV Конгрестің күн тәртібін әзірлеу кезінде назарға алуды ұсынады.

Конгресс жұмысы қатысушылардың Қазақстанда архившілер конгрестерін өткізу бойынша Қазақстан Республикасы Президенті Архивінің тәжірибесін қолдауға бейілділігін көрсетті.

Цифрлық трансформация жағдайында архивтер өз функцияларын барынша толық іске асыру және қоғамдық маңызы бар міндеттерді сапалы орындау үшін цифрлық технологияларға бейімделуде. Архивтер құжаттар қоймасынан деректердің түпнұсқалығын, сенімділігін және ұзақ мерзімді сақталуын қамтамасыз ететін деректерді басқарудың белсенді орталықтарына біртіндеп айналып келеді. Архив қызметінің жаңа көкжиектері қоғамның нақты сұраныстарын қамтамасыз ету үшін жағдай жасайды.

Конгресс қатысушылары архивші, ақпарат менеджері мамандығының және Халықаралық архивтер кеңесінің мұраттарына адалдығын растайды, бұл ретте деректі мұраны, адамзаттың тарихы мен жадын сақтауды; ашық қоғамдарда олардың қолжетімділігін үнемі жақсартуды; білім беру, ғылым үшін пайдалануды, үкімет пен қоғам деңгейіндегі есептілік пен ашықтықты қамтамасыз етуді негізге алады.

Конгресс қатысушылары жасанды интеллект архив саласы эволюциясының негізгі факторына айналғанын және көптеген елдердің деректі мұраны басқару саласындағы мемлекеттік саясатының басымдықтары қатарына енгенін мәлімдейді. Бұл ретте жасанды интеллект архив мамандығын алмастырмайды, керісінше архивтердің когнитивтік және басқарушылық мүмкіндіктерін күшейте отырып, оны трансформациялайды.

Архивтерді цифрлық трансформациялау бытыраңқы ақпараттық жүйелерді интеграциялауды, метадеректерді үйлестіруді, сақтау процестерін стандарттауды және архивтік басқарудың бірыңғай орнықты архитектурасын құруды талап етеді.

Конгресс қатысушылары қоғам үшін ашық архивтердің әлеуеті цифрландырылған құжаттардың санына пропорционалды түрде өсетініне келіседі. Құжаттарды, архивтік жобаларды пайдалану жиілігі артады, ғылымға, архивтік практикаға – архивтік құжаттарды жинау мен сипаттауға, оларды шығармашылықпен пайдалануға тартылған азаматтардың саны көбейеді.

Сонымен қатар, поликризис дәуірінде бүкіләлемдік деректі мұраның кепілді сақталуы тәуекелдерді проактивті модельдеуді және трансшекаралық қатерлерге – соның ішінде климаттық аномалияларға, техногендік апаттарға, биологиялық сын-қатерлерге және критикалық ақпараттық инфрақұрылымға жасалатын үйлестірілген кибершабуылдарға қарсы тұруды талап етеді.

РЕЗОЛЮЦИЯ
III МЕЖДУНАРОДНОГО КОНГРЕССА АРХИВИСТОВ
«DIGITAL ARCHIVE. EXPO»
(9-11 июня 2026 года, г. Астана, Казахстан)

Участники III Конгресса архивистов «Digital Archive. Expo» принимают настоящую Резолюцию и рекомендуют принять во внимание при разработке повестки очередного IV Конгресса в 2028 году.

Работа Конгресса продемонстрировала приверженность участников в поддержке практики Архива Президента Республики Казахстан по проведению Конгресса архивистов в Казахстане.

В условиях цифровой трансформации архивы адаптируются к цифровым технологиям для максимально полной реализации своих функций и качественного выполнения общественно значимых задач. Архивы постепенно трансформируются из хранилищ документов в активные центры управления данными, обеспечивающие их подлинность, достоверность и долгосрочную сохранность. Новые горизонты деятельности архивов создают условия для обеспечения реальных запросов общества.

Участники Конгресса подтверждают свою приверженность ценностям архивной профессии и принципам Международного совета архивов, признавая сохранение документального наследия и исторической памяти важнейшей общественной миссией. Они выступают за расширение доступа к архивным ресурсам, их использование в науке и образовании, а также за укрепление прозрачности и подотчетности государственных институтов и общества.

Участники Конгресса констатируют: искусственный интеллект становится ключевым фактором эволюции архивной отрасли, включен в число приоритетов государственной политики многих стран в сфере управления документальным наследием. При этом искусственный интеллект не заменяет архивную профессию, а трансформирует ее, усиливая когнитивные и управленческие возможности архивов.

Цифровая трансформация архивов требует интеграции разрозненных информационных систем, унификации метаданных, стандартизации процессов хранения и создания единой устойчивой архитектуры архивного управления.

Участники Конгресса согласны, что потенциал открытых архивов для общества будет расти пропорционально количеству оцифрованных документов. Увеличится количество использований документов, архивных проектов, вовлеченных граждан в науку, архивную практику: сбор и описание архивных документов, их креативное использование.

Вместе с тем, в эпоху поликризиса гарантированная сохранность всемирного документального наследия требует проактивного моделирования рисков и противодействия трансграничным угрозам, включая климатические аномалии, техногенные катастрофы, биологические вызовы и скоординированные кибератаки на критическую информационную инфраструктуру.

RESOLUTION
OF THE III INTERNATIONAL CONGRESS OF ARCHIVISTS «DIGITAL ARCHIVE. EXPO»
(June 9-11, 2026, Astana, Kazakhstan)

The participants of the III Congress of Archivists "Digital Archive Expo" hereby adopt this Resolution and recommend taking its provisions into account when developing the agenda for the upcoming IV Congress in 2028.

The work of the Congress demonstrated the participants' commitment to supporting the practice of the Archive of the President of the Republic of Kazakhstan in hosting congresses of archivists in Kazakhoкestan.

In the context of digital transformation, archives are adapting to digital technologies to fully realize their functions and effectively fulfill socially significant tasks. Archives are gradually transforming from document repositories into active data management centers that ensure data authenticity, reliability, and long-term preservation. New horizons of archival activity create the necessary conditions to meet the real demands of society.

The participants of the Congress reaffirm their commitment to the ideals of the profession of archivist and information manager, as well as to the principles of the International Council on Archives, viewing the preservation of documentary heritage, memory, and human history as paramount; continuously improving accessibility to them in open societies; and utilizing them for education, science, and ensuring accountability and transparency in governments and societies.

The participants of the Congress state that artificial intelligence is becoming a key factor in the evolution of the archival industry and is included among the priorities of state policy in many countries regarding documentary heritage management. At the same time, artificial intelligence does not replace the archival profession but rather transforms it, enhancing the cognitive and managerial capabilities of archives.

The digital transformation of archives requires the integration of disparate information systems, the unification of metadata, the standardization of preservation processes, and the creation of a single, sustainable architecture for archival management.

The participants of the Congress agree that the potential of open archives for society will grow proportionally to the volume of digitized documents. The utilization of documents and archival projects will increase, and more citizens will be engaged in science and archival practice – including the collection, description, and creative use of archival documents.

At the same time, in the era of polycrisis, ensuring the guaranteed preservation of the world's documentary heritage requires proactive risk modeling and countering transboundary threats – including climate anomalies, man-made disasters, biological challenges, and coordinated cyberattacks on critical information infrastructure.

МАЗМҰНЫ

КОНГРЕСТІҢ САЛТАНАТТЫ АШЫЛУЫ

Қазақстан Республикасының Президенті Қ.К. Тоқаевтың Конгресс қатысушылары мен қонақтарына үндеуі 4

Қазақстан Республикасы Президенті Архивінің директоры Ә. Мұстафинаның құттықтауы 6

Welcome speech of Director of the UN Library and Archives Geneva Ch. Warakaulle 8

Welcome speech of F. Banda, Chief of Documentary Heritage Unit, Communication and Information Sector, UNESCO 10

Welcome speech of President of the International Council on Archives J. Kirps 12

Приветственное слово депутата Сената Парламента Республики Казахстан Р. Рустемова 14

Welcome speech of Senior Vice-President National Archives of Japan F. Ichiro 16

Приветственное слово начальника Главного архивного управления при Правительстве Республики Таджикистан Л.Ш. Шарифзоды 18

«DIGITAL ARCHIVE: ПАРАДИГМАНЫҢ АУЫСУЫ» ПЛЕНАРЛЫҚ СЕССИЯ

A. Mustafina 21
Archives in the AI Era

Ch. Warakaulle 27
Digitizing the Multilateral Process: Challenges and Opportunities for International Organization

F. Ichiro 36
Challenges in Digital Archives Collaboration and Learning Content Development by the National Archives of Japan

M. Bragg 44
Family History in a Digital Age: Enduring Purpose and Global Impact

Цюй Кай 48
Инновационные практики представления архивной информации: взгляд из Китая

J. Sheridan 57
The Economics of Digital Archiving in the AI Age

P. Zawilski 70
Quo vadis, digitization?

I. Ko 78
Current Status and Challenges of Collecting Presidential Private Records at the Presidential Archives

«АРХИВТЕР ЖӘНЕ AI:
ҚҰРАЛДАРЫ, ЭТИКАСЫ МЕН БОЛАШАҚ СТРАТЕГИЯЛАРЫ»
№ 1 СЕССИЯ

М.В. Ларин 88
О возможностях применения технологий ИИ в архивном деле
S. Park 95
What is Artificial Intelligence? Why Does It Matter for the Future of Documents and Archives?
S. Filios 105
Artificial Intelligence for Accessible Archives: Making Records Searchable for All
А.М. Байкенова 111
Цифровизация национальной памяти: мультискриптовый OCR и диаризация для архивного наследия
B. Altındiş 116
From Traditional Archives to Data Centers: The Contribution of Artificial Intelligence to the Processing of Archival Records and Research Processes
N. Soh 130
Using AI in Review of Text-Based Archival Records
Г. Мамадзаде 138
Развитие архивного дела с использованием технологий искусственного интеллекта
L. Shonhe 146
Large Language Models are not a One-Size-Fit All Solution for Archives and Records Management

«DIGITAL ARCHIVE: ЗАМАНАУИ ДИСКУРСТАР»
№ 2 СЕССИЯ

A. Kılınç, H.A. Tokur 167
The Encyclopedia of the Turkic World Project and the Construction of Digital Cultural Memory
Н.Г. Суровцева 181
Влияние цифровой трансформации на комплектование архивов
S. Sahakyan 187
Building National Archival Strategies in the Age of AI: The Case of Armenia
М.Н. Назарализода, М.С. Курбанова 198
Внедрение цифровых технологий и искусственного интеллекта в архивное дело в Республике Таджикистан
P. Popławski 206
From Human Intelligence to Process Automation and Artificial Intelligence: Transition at the Archive of the Institute of National Remembrance
П. Рыженков 215
8K-оцифровка аудиовизуальных архивов как стратегия сохранения и монетизации *наследия*
K. Morris 234
An AI-Based Solution for Converting Paper Documents into Intelligent, Searchable Digital Artifacts
К.Ш. Алимгазинов 238
От бумажной к цифровой архитектонике документа

ФОТОГАЛЕРЕЯ 246

«ЖАСАНДЫ ИНТЕЛЛЕКТТІҢ «ӘЛЕМ ЖАДЫ» ХАЛЫҚАРАЛЫҚ ТІЗІМІНЕ ӘСЕРІ» UNESCO СЕМИНАРЫ

А.Х. Мустафина 260
Приветственное слово
Zh. Shaimenova 262
Welcome Addresses
Н.Ж. Шаймерденова 264
Историко-лингвистическое исследование письменных источников на основе архивных картотек в цифровую эпоху
S. Park 270
Artificial Intelligence Implications for UNESCO Memory of the World Register – Valorisation of AI-generated Works as Documentary Heritage
Lai Tee Phang 276
Preservation of AI-generated Works as Documentary Heritage

«АРХИВТЕРДІҢ ИННОВАЦИЯЛЫҚ ЭКОЖҮЙЕЛЕРІ: ИНФРАҚҰРЫЛЫМ, ЫНТЫМАҚТАСТЫҚ ЖӘНЕ ТҰРАҚТЫЛЫҚ» № 3 СЕССИЯ

Firass Atallah Mohammad Darabah 284
Protecting the National Memory in the Age of AI: The Jordanian National Library project for 2026
У.М. Юсупов 296
Цифровая экосистема архивов Узбекистана: инфраструктура, ИИ и государственные сервисы
Ф.Р. Аббасов 304
Роль архивов в цифровом управлении и применение ИКТ
M. Tokdemir, S. Koyun, Ö. Kurtuluş 312
Natıonal Archıves Management and Integratıon: The Bays Project
T. Kux 321
Digital Long-Term Records Preservation and Migration Strategies
I. Ferrari 326
An Integrated System for the Permanent Protection of Archival Heritage
R. Łoboda 335
Securing digital archives against data loss in the disinformation era
Ж. Жумагалиев 347
Современные проблемы архивов государственных структур и бизнеса и пути их решения с использованием технологий Canon
K. Kabarova 353
Intelligent Archival Infrastructure Management System

« ЖЕТІЛДІРІЛГЕН ТАРИХ: ОҚУШЫЛАРДЫҢ ТАРИХИ ОЙЛАУЫН ЖЕТІЛДІРУ ҮШІН ЖИ ПАЙДАЛАНУ»
СЕМИНАРЫ

A. Payne 359
Better History – Using AI to help students develop more effective thinking about documents

«АРХИВ БОЛАШАҒЫ: ЖАҢА ҚҰЗЫРЕТТЕР МЕН КӨШБАСШЫЛЫҚ»
№ 4 СЕССИЯ

Д.Р. Щеглова 381
AI-Sana: трамплин новых компетенций
A. Payne 388
The UK National Experience in Archivist Training
К.Т. Туенбаева 402
Болашақ архившілер: білім, құзыреттер және технологиялар
С.С. Исмаилов 415
Образовательная программа «Архивоведение, документоведение и музейное дело: реалии и перспективы
Г.Т. Жакупова 424
Образовательная и исследовательская среда подготовки архивистов в рамках концепции цифрового архива
435

ФОТОГАЛЕРЕЯ

«**ЦИФРЛЫҚ АРХИВТЕР: МЕМЛЕКЕТ, ҒЫЛЫМ, БИЗНЕС ЖӘНЕ КРЕАТИВТІ ИНДУСТРИЯ ҮШІН ЖАҢА МҮМКІНДІКТЕР»**
MURA-TALKS

F. Banda 447
The Future of Human Memory: How to Preserve Digital Heritage in 100 Years
А. Сарым 451
Архив және цифрлық заман: тарихи жадыны сақтаудың жаңа форматы
А. Алиев 457
История в один клик: почему архивы будущего создаются вами
E. Gracia 460
From Information Overload to Trusted Knowledge: The Power of Audiovisual Archives
М. Жұмабай 465
Архив – креативтің қайнар көзі
В. Шин 470
Архив как живая память: как партнерство IT и архивов меняет опыт взаимодействия с историей

«АРХИВТЕР ҚОҒАМ ҮШІН»
№ 5 СЕССИЯ

Д.Б. Аманжолова 475
От архивного фонда к виртуальному музею: создание цифровой экосистемы наследия
E. Özkul 487
Invisibility in Digital Archives: The Selection and Presentation of Information
J. Pok 495
History Reimagined: How the National Archives of Singapore's Collection Turns. Ideas into Creative Works in an Island-Wide Competition
J. McCollum 505
Oral Family History: Capturing, Preserving, and Creating Searchable Records
A. Olesińska, P. Lesisz 509
Functionalities and Prospects of Modern Archival Databases based on the projects of the Witold Pilecki Institute of Solidarity and Valor
К.К. Актаева 515
Роль архивов в сохранении преемственности государственной власти и исторической идентичности
С.К. Удербаева 523
Архивы как общественный институт: доступ к памяти, знаниям и идентичности
С.Р. Ардашова 531
Цифровые архивные ресурсы как инструмент популяризации исторического наследия: портал «Документальное наследие Южного Урала XVII–XVIII веков»
М. Масанов 540
Координационная работа Регионального офиса UNESCO в г. Алматы в сотрудничестве стран Центральной Азии в области архивного и музейного дела
Ш.Д. Бүркітбаева 545
Сирек кітаптар мен қолжазбаларды цифрландыру: ұлттық мұраны мәдениетаралық интеграциялау
Ш.В. Тлепина 557
Архивы и креативные индустрии: стратегия открытости и социальной памяти

КОНГРЕСТІҢ САЛТАНАТТЫ ЖАБЫЛУЫ

Қ.Ш. Әлімғазинов 568
Конгрестің қорытындылары туралы
А.Х. Мустафина 572
Заключительное слово
«Digital Archive. EXPO» III Халықаралық архившілер конгресінің резолюциясы 574

Сборник материалов
III Конгресса архивистов Казахстана
«Digital Archive EXPO» 8-12 июня 2026 г.
Компьютерная верстка: Г.Т. Сандыбаева
Шрифт: Montserrat. Кегль: 12 пт.
Электронное издание

ISBN 978-601-7149-30-7