



Боброва Елена Викторовна

Российский университет транспорта (МИИТ),
Всероссийский научно-исследовательский институт
документоведения и архивного дела
г. Москва, Российская Федерация
SPIN-код: 9882-3748, AuthorID: 638778
oumnique@gmail.com

Обзорная статья

ОБ ОСНОВНЫХ НАПРАВЛЕНИЯХ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В АРХИВНОМ ДЕЛЕ

Рассматриваются результаты работы сотрудников ВНИИДАД по подготовке научного доклада «Применение искусственного интеллекта в управлении документами и в архивах: практика, тенденции и перспективы». Раскрываются актуальность и степень изученности темы. Сформулированы наиболее актуальные проблемы применения искусственного интеллекта в сфере архивного дела. Рассмотрены организационные формы внедрения интеллектуальных информационных систем и аспекты внедрения, на которые нужно обратить внимание.

Ключевые слова: архивное дело, данные, документ, информационные системы, искусственный интеллект, управление документами, электронный документ.

Для цитирования: Боброва Е.В. Об основных направлениях применения искусственного интеллекта в архивном деле // Вестник ВНИИДАД. 2024. № 1. С. 75–83.

Review article

ON THE MAIN AREAS OF AI APPLICATION IN ARCHIVAL AFFAIRS

The results of VNIIDAD's scientific report "Application of artificial intelligence in records management and archives: practice, trends and prospects" are presented. The relevance and degree of knowledge of the topic is revealed. The most pressing problems are formulated. The organizational forms of intelligent information systems implementation and aspects of implementation requiring special attention are considered.

Keywords: archival affairs, artificial intelligence, data, document, electronic document, information systems, records management.

Поступила в редакцию: 11.01.2024
Поступила после рецензирования: 17.01.2024
Принята к публикации: 18.01.2024

Received: 11.01.2024
Revised: 17.01.2024
Accepted: 18.01.2024

For citation: Bobrova E.V. On the main areas of AI application in archival affairs. Vestnik VNIIDAD = Herald of VNIIDAD. 2024;1:75–83. (In Russian).

Проблема внедрения технологий искусственного интеллекта (ИИ) во все отрасли экономики и социальной сферы России остро встала в последние несколько лет. Необходимость цифровой трансформации традиционных процессов в государственном управлении диктует необходимость широкого внедрения прорывных информационных технологий, в том числе и технологий ИИ.

В течение последних 10 лет страны мира ведут гонку по внедрению интеллектуальных систем, максимально обострившуюся в конце 2022 года с появлением в открытом доступе генеративной нейронной сети ChatGPT 3.5, что повлекло за собой как возникновение аналогичных нейронных сетей от других производителей, так и множество новых программных продуктов, основанных на этой или аналогичных технологиях.

Совершенно очевидно, что мы вступаем в эпоху очередной информационной революции, когда в результате кардинальных изменений в сфере обработки информации происходит преобразование общественных отношений, следствием чего является приобретение человеческим обществом нового качества жизни. Предыдущая информационная революция была связана с широким распространением персональных компьютеров (ПК) и телекоммуникационных сетей, нынешняя – с искусственным интеллектом.

Основатель фирмы Microsoft Билл Гейтс на странице своего блога пишет: «Развитие ИИ столь же фундаментально, как создание микропроцессора, персонального компьютера, Интернета и мобильного телефона. ИИ изменит то, как

люди работают, учатся, путешествуют, получают медицинскую помощь и общаются друг с другом. Вокруг него будут переориентироваться целые отрасли» [1].

Понимая необходимость и неизбежность включения в информационную гонку, связанную с ИИ, высшие органы власти Российской Федерации уделяют этому направлению особое внимание.

Еще в 2016 году в соответствии с Указом Президента Российской Федерации от 1 декабря 2016 года № 642 «О Стратегии научно-технологического развития Российской Федерации» переход к передовым цифровым, интеллектуальным производственным технологиям, роботизированным системам, новым материалам и способам конструирования, создание систем обработки больших объемов данных, машинного обучения и искусственного интеллекта были обозначены в качестве приоритетного направления научно-технологического развития Российской Федерации на ближайшие 10–15 лет.

22 ноября 2023 года Президент России В.В. Путин на конференции по искусственному интеллекту и машинному обучению «Artificial Intelligence Journey 2022» сформулировал магистральную задачу «нового этапа в горизонте текущего десятилетия – обеспечить массовое внедрение искусственного интеллекта. Оно должно охватить все отрасли экономики, социальной сферы и систему госуправления» [2].

29 января 2023 года на сайте Президента Российской Федерации был опубликован Перечень поручений Президента Российской Федерации по итогам конференции «Путешествие в мир искусственного интеллекта», в котором сказано, что

до 15 октября 2023 года Правительство обязано провести мониторинг результатов применения технологий искусственного интеллекта в целях оценки уровня внедрения этих технологий в отраслях экономики и социальной сферы, а также результативности деятельности государственных органов, органов местного самоуправления по обеспечению внедрения таких технологий («индекс интеллектуальной зрелости») и в дальнейшем проводить его ежегодно.

26 сентября 2023 года председатель Правительства Российской Федерации М.В. Мишустин провел стратегическую сессию «Развитие искусственного интеллекта», на которой озвучил следующую информацию: «Объем рынка искусственного интеллекта в прошлом году достиг почти 650 млрд рублей. Рост составил примерно 18%. Показатель внедрения искусственного интеллекта в отраслях экономики достигает порядка 20%. В лидерах – финансовый сектор, где такие технологии применяют 95% предприятий. Всего на развитие искусственного интеллекта из федеральных средств на будущий год предусмотрено 5,2 млрд рублей» [3].

Из его выступления становится ясно, что для обновления Национальной стратегии развития искусственного интеллекта до 2030 года необходимо прежде всего определиться с базовыми сценариями поддержки искусственного интеллекта в России, которые и лягут в основу обновленной стратегии.

Очевидно, что сложившаяся ситуация в полной мере касается и сферы управления документами, и архивного дела. В настоящий момент одной из важнейших задач для органов управления архивным делом становится поиск ответов на следующие вопросы:

- Как обеспечить необходимый «индекс интеллектуальной зрелости» в архивной сфере регионов?

- Что нужно сделать, чтобы внедрить технологии ИИ в работу органа управления архивным делом и подведомственных учреждений?

- Какие существующие информационные платформы можно использовать для решения предыдущих вопросов или нужно разработать их с нуля?

- Какие вызовы несет за собой внедрение технологий ИИ?

В 2023 году Всероссийский научно-исследовательский институт документо-ведения и архивного дела (ВНИИДАД) в рамках выполнения Плана научно-исследовательских и опытно-конструкторских работ по государственному заданию Федерального архивного агентства подготовил научный доклад «Применение искусственного интеллекта в управлении документами и в архивах: практика, тенденции и перспективы» (далее – научный доклад).

Целями научного доклада являлись выявление основных направлений и методов применения искусственного интеллекта в управлении документами и в архивном деле, оценка эффективности и рисков использования технологий, формулирование предложений по перспективным направлениям применения ИИ в поддержке процессов документооборота и в архивной практике.

Для достижения намеченных целей в рамках научного доклада поставлены следующие задачи:

- проведение анализа отечественной и зарубежной литературы по разработке и применению методов и технологий искусственного интеллекта в управлении документами и в архивах;

- проведение анализа опыта применения технологий искусственного интеллекта в сфере делопроизводства и архивного дела;

- подготовка предложений по направлениям внедрения технологий искусственного интеллекта в системы электронного документооборота и системы хранения электронных документов в органах государственной власти;

- подготовка предложений по внедрению перспективных решений в области искусственного интеллекта в архивное дело.

Говоря о степени изученности темы, следует отметить, что в настоящий момент сформировалась уже довольно обширная историография работ, посвященных различным аспектам применения ИИ и интеллектуальных информационных систем (ИИС) в архивном деле и в сфере управления документами как в нашей стране, так и за рубежом¹. Однако значительная часть работ носит слишком общий характер либо, напротив, описывает только конкретные аспекты создания или использования ИИ, т.е. носит фрагментарный характер.

Существующие исследования можно условно разделить на следующие тематические направления:

- нормативная правовая база применения ИИ;

- анализ понятийного аппарата и терминов ИИ;

- анализ международных и национальных стандартов, используемых в управлении документами и в архивах и связанных с ИИ;

- описание конкретных проектов, использующих технологии ИИ;

- описание технологий ИИ при работе с документами;

- выявление различных проблем и рисков, возникающих при применении ИИ в работе с документами.

Есть работы, в которых формулируются конкретные цели и подходы к разработке технологий ИИ в архивах.

Проблемы внедрения ИИС активно обсуждаются на семинарах, круглых столах и конференциях в нашей стране. Только за последний год этой теме были посвящены девять мероприятий. Наиболее интересным и информационно насыщенным, на наш взгляд, стал проведенный во ВНИИДАД 10 апреля 2023 года круглый стол «Практические задачи внедрения технологий искусственного интеллекта в деятельность архивов»², на котором было представлено 15 докладов архивистов, специалистов в области технологий применения искусственного интеллекта, историков. В удаленном режиме к прямой трансляции заседания подключились более 200 участников [4], что говорит об огромном интересе архивистов к этой проблеме.

Об этом также свидетельствуют и тот факт, что в Интернете по инициативе архивных учреждений проводятся опросы пользователей. Так, Архивное агентство Красноярского края проводит опрос на тему «Нейросети как инструмент распознавания рукописных текстов архивных документов»³, а Управление по делам архивов Калужской области на портале

² Практические задачи внедрения технологий ИИ в деятельность архивов: круглый стол (10 апреля 2023 г.) // ВНИИДАД. URL: <https://www.youtube.com/watch?v=KNzhpS42vqk> (дата обращения: 12.11.2023).

³ Нейросети как инструмент распознавания рукописных текстов архивных документов [голосование]. URL: <https://24ag.ru/voting/details/1165?locality=1> (дата обращения: 12.11.2023).

¹ В списке источников и литературы к докладу указаны 560 наименований.

«Госуслуги» проводит опрос на тему «Искусственный интеллект и архивы»⁴.

Аналогичное повышенное внимание наблюдается и со стороны пользователей. В 2023 году ВНИИДАД проводил на портале «Архивы России» опрос пользователей архивных сайтов. Результаты опроса показали, что среди тех пользователей архивных сайтов, кто предлагает усовершенствовать поисковый механизм Центрального фондового каталога, 6% советуют использовать для улучшения поиска искусственный интеллект. Очень много пожеланий направлено на совершенствование поисковых средств виртуальных читальных залов с помощью нейросетей, настроенных на автоматическое распознавание оцифрованных образов документов и поиск по распознанным страницам (с подсветкой найденного в документе текста) [5].

25 января 2023 года в сети Интернет стал доступен новый сервис от «Яндекса» «Поиск по архивам»⁵. В 2023 году все социальные сети, в которых общаются пользователи, интересующиеся генеалогией, были наводнены петициями к различным государственным структурам с просьбами рассмотреть вопрос и содействовать появлению на данном ресурсе архивных документов того или иного региона.

Спустя год (по состоянию на 11 января 2024 года) на сайте проекта доступны для поиска 10 281 260 документов из 63 803 отсканированных дел из 13 415 фондов,

находящихся на хранении в 14 архивных учреждениях и одном музее.

Все вышеизложенное свидетельствует о своевременности и необходимости подготовки научного доклада, посвященного вопросам применения ИИ в управлении документами и в архивах.

В разделе 1 научного доклада дан подробный обзор нормативной правовой и методологической базы, на основании которой действуют государственные органы и организации в России и в других странах в вопросах применения искусственного интеллекта. Обозначены стратегические подходы, проблемы создания и внедрения технологий ИИ. Поставлена задача разработки специальных требований к системам искусственного интеллекта, предназначенным для решения различных задач в сфере управления документами и архивного дела.

В разделе 2 рассмотрены вопросы, связанные с применением искусственного интеллекта для решения задач автоматизированной обработки реквизитов документов и распознавания информации, необходимых при классификации документной информации, а также вопросы использования ИИ при проведении экспертизы ценности электронных документов.

В разделе 3 представлены практика и перспективы применения искусственного интеллекта в сферах комплектования, учета, формирования справочно-поисковых средств, обеспечения сохранности и использования архивных документов.

В докладе рассмотрены:

- вопросы поддержки процесса передачи электронных документов и их метаданных между информационными системами в рамках комплектования архивов организаций и государственных и муниципальных архивов;

⁴ Искусственный интеллект и архивы [опрос]. URL: <https://pos.gosuslugi.ru/lkp/polls/384002/> (дата обращения: 12.11.2023).

⁵ Поиск по архивам // Проект компании «Яндекс». URL: <https://yandex.ru/archive> (дата обращения: 12.11.2023).

- вопросы использования искусственного интеллекта в сфере учета;
- вопросы использования семантического анализа для формирования справочно-поисковых средств;
- возможности внедрения интеллектуальных технологических систем в архивохранилища;
- практика использования искусственного интеллекта для выявления и установления степени биологического заражения архивных документов;
- практика восстановления текстов и изображений в документах, физическое состояние которых препятствует извлечению информации;
- вопросы интеллектуального распознавания документов (как текстовых, так и аудиовизуальных);
- перспективы разработки и внедрения интеллектуальных электронных помощников для исполнения запросов;
- использование систем интеллектуального поиска в архивных информационных ресурсах;
- возможности применения искусственного интеллекта при создании электронного фонда пользования архивных документов;
- вопросы использования искусственного интеллекта при разработке пользовательских сервисов на сайтах и в информационных системах читальных залов архивов.

В заключении научного доклада представлены выводы по итогам исследования и предложены пути дальнейшего изучения темы.

Выявлены следующие наиболее актуальные проблемы:

1. Нерешенность вопроса ответственности в случае применения систем искусственного интеллекта, ответственности за

причинение вреда с использованием систем искусственного интеллекта.

2. Необходимость повышения уровня доверия к технологиям искусственного интеллекта.

3. Обеспечение охраны прав на результаты интеллектуальной деятельности, а также достоверности генерируемого контента.

4. Отсутствие большого количества доступных размеченных и структурированных данных для обучения алгоритмов искусственного интеллекта и информационно-коммуникационной инфраструктуры для обеспечения доступа к наборам таких данных.

5. Отсутствие полноценной системы стандартизации в области технологий искусственного интеллекта.

6. Неунифицированность этических принципов разработки и применения технологий искусственного интеллекта.

Обозначены организационные формы внедрения интеллектуальных информационных систем:

1. Архивное учреждение может по своей инициативе пытаться собственными силами, без централизованного финансирования, но с привлечением в каких-то вопросах некоммерческих партнеров создать свой собственный интеллектуальный продукт.

2. Архивное учреждение может купить интеллектуальную технологию общего назначения, приспособив ее решения для собственных задач.

3. Архивное учреждение может вступить в некоммерческое партнерство с организацией – разработчиком интеллектуальной технологии или интеллектуального продукта.

4. Коммерческая фирма может разработать и предложить архивным учреждени-

ям различные интеллектуальные решения, предназначенные для автоматизации именно архивных процессов или же создания специализированных архивных сервисов для пользователей.

5. Интеллектуальная информационная система, ориентированная именно на архивные процессы, может разрабатываться специально при централизованном (желательно федеральном) финансировании как типовое решение, реализующее национальный проект.

Безусловно, для внедрения ИИ нужно финансирование, зачастую значительное. Сейчас в России складывается благоприятная ситуация для поиска финансовых и иных ресурсов. Очевидно, что достаточно подготовленные, прочно обоснованные и эффективные проекты будут способствовать повышению цифровой зрелости регионов и получать материальную поддержку.

При этом нужно учитывать еще несколько важных моментов:

1. Внедрение технологий ИИ невозможно там, где еще нет значительной информатизации процессов, потому что нет достаточного количества данных, необходимых для обработки и обучения.

2. Особое внимание нужно уделять процессу подготовки качественных дата-сетов, и готовить их должны не программисты и математики, а архивисты и документоведы. Следовательно, они должны освоить этот новый вид деятельности.

3. При принятии решения о внедрении технологий ИИ важно учесть методологические и экономические аспекты: определить приоритетную задачу, для которой применение ИИ даст наибольший эффект; определить, какой метод даст для достижения этой задачи наиболее значимый результат; оценить затраты на реализацию

выбранного метода для данной задачи (на аппаратные, программные, информационные ресурсы, связь, подготовку дата-сетов и т.д.) и прогнозируемую экономическую эффективность (с учетом необходимости дальнейшей поддержки разработанной интеллектуальной информационной системы). При этом желательно отдавать предпочтение решениям, носящим комплексный характер, а не отдельным сервисам, которые либо в принципе не смогут быть потом интегрированы в информационную систему архива, либо потребуют на это дополнительных затрат.

4. Внедрение интеллектуальных информационных систем неизбежно повлечет за собой автоматизацию рутинных операций, что вызовет проблему вытеснения специалистов низшего и среднего звена. Следовательно, необходимо заранее учитывать необходимость переквалификации таких сотрудников, организации для них переобучения, и эти вопросы нужно тщательно продумывать.

5. Риски, связанные с внедрением технологий искусственного интеллекта, будут влиять не только на архивистов, но и на фондообразователей. Защита персональных данных, обеспечение информационной безопасности, риски принятия на архивное хранение «фейков», проблема объяснимости ИИ при принятии им решений и доверия к объяснениям – это проблемы, требующие специальной проработки и дальнейших исследований.

В приложении А к научному докладу приведен перевод части публикации Б.М. Шабу, Ж. Тьеш, Ж. Кнафу, А. Година «Алгоритмические методы исследования автоматизации анализа структурированных и неструктурированных цифровых данных» [6], касающейся уровней критериев оценки и связанных метрик, кото-

рые можно использовать при проведении экспертизы ценности электронных документов с помощью искусственного интеллекта.

В приложении Б продемонстрированы результаты тестирования базовых моделей платформы Transkribus для распознавания русскоязычных документов.

В приложении В приведен указатель проектов и платформ, упомянутых в научном докладе. В нем перечислены проекты, у которых в публикациях были упомянуты официальные названия.

В приложении Г сформулированы предложения по перспективным направлениям применения искусственного интеллекта в

поддержке процессов документооборота и в архивной практике.

Исследование ВНИИДАД, представленное в научном докладе «Применение искусственного интеллекта в управлении документами и в архивах: практика, тенденции и перспективы», является частью работы по формированию научной и информационной основы для разработки нормативных и методических документов, регулирующих применение искусственного интеллекта в архивной сфере.

Авторы доклада надеются, что исследования и выводы доклада могут быть практически полезны при создании типовых программных решений отраслевого назначения.

Список источников

1. Gates W.H. The Age of AI has begun [Эра искусственного интеллекта началась] // *GatesNotes – the blog of Bill Gates*, March 21, 2023. URL: <https://www.gatesnotes.com/The-Age-of-AI-Has-Begun> (дата обращения: 12.11.2023).
2. Конференция по искусственному интеллекту (24 ноября 2022 г., Москва) // Президент России – официальный сайт URL: <http://kremlin.ru/events/president/news/69927> (дата обращения: 12.11.2023).
3. Михаил Мишустин провел стратегическую сессию «Развитие искусственного интеллекта» (26 сентября 2023 г., Координационный центр Правительства России) // Правительство России – официальный сайт. URL: <http://government.ru/news/49604/> (дата обращения: 12.11.2023)¹.
4. Киселев И.Н. Круглый стол «Практические задачи внедрения технологий искусственного интеллекта в деятельность архивов» // *Отечественные архивы*. 2023. № 3. С. 122–125.
5. Боброва Е.В. Архивные информационные ресурсы: результаты опроса пользователей // *Отечественные архивы*. 2022. № 4. С. 35–45.
6. Shabou B.M., Tièche J. Algorithmic methods to explore the automation of the appraisal of structured and unstructured digital data [Алгоритмические методы исследования автоматизации анализа структурированных и неструктурированных цифровых данных] // *Records Management Journal*, 2020. URL: <https://www.emerald.com/insight/content/doi/10.1108/RMJ-09-2019-0049/full/pdf> (дата обращения: 30.09.2023).

References

1. Gates W.H. The Age of AI has begun. *GatesNotes – the blog of Bill Gates*, March 21, 2023. URL: <https://www.gatesnotes.com/The-Age-of-AI-Has-Begun> (accessed: 12.11.2023).
2. Konferenciya po iskusstvennomu intellektu (24 noyabrya 2022 g., Moskva) [Conference on artificial intelligence (November 24, 2022, Moscow)]. *President of Russia – official website*. URL: <http://kremlin.ru/events/president/news/69927> (accessed: 12.11.2023). (In Russian).

3. Mixail Mishustin provyol strategicheskuyu sessiyu “Razvitie iskusstvennogo intellekta” (26 sentyabrya 2023 g., Koordinacionny`j centr Pravitel`stva Rossii) [Mikhail Mishustin held a strategic session “Development of artificial intelligence” (September 26, 2023, Coordination Center of the Government of Russia)]. *Russian Government – official website*. URL: <http://government.ru/news/49604/> (accessed: 12.11.2023). (In Russian).
4. Kiselev I.N. Krugly`j stol “Prakticheskie zadachi vnedreniya texnologij iskusstvennogo intellekta v deyatel`nost` arxivov” [Round table “Practical problems of artificial intelligence technologies implementation in archives activities”]. *Otechestvenny`e arxivny` = Domestic Archives*. 2023;3:122–125. (In Russian).
5. Bobrova E.V. Arxivny`e informacionny`e resursy` : rezul`taty` oprosa pol`zovatelej [Archival information resources: user survey results]. *Otechestvenny`e arxivny` = Domestic Archives*. 2022;4:35–45. (In Russian).
6. Shabou B.M., Tièche J. Algorithmic methods to explore the automation of the appraisal of structured and unstructured digital data. *Records Management Journal*, 2020. URL: <https://www.emerald.com/insight/content/doi/10.1108/RMJ-09-2019-0049/full/pdf> (accessed: 30.09.2023).

ИНФОРМАЦИЯ ОБ АВТОРАХ

Боброва Елена Викторовна, старший преподаватель Российского университета транспорта (МИИТ); старший научный сотрудник отдела архивоведения Всероссийского научно-исследовательского института документоведения и архивного дела (ВНИИДАД), Москва, Российская Федерация.

INFORMATION ABOUT THE AUTHORS

Elena V. Bobrova, senior lecturer of the Russian University of Transport (MIIT), senior researcher of Archival Science Department of the All-Russian Scientific and Research Institute for Records and Archives Management (VNIIDAD), Moscow, Russian Federation.
