

НАУЧНАЯ СТАТЬЯ
УДК 930.25



Искусственный интеллект в документной деятельности

Ольга Владимировна МЕДВЕДЕВА ✉, **Анастасия Сергеевна ПОПОВА**
ФГБОУ ВО «Тамбовский государственный университет им. Г.Р. Державина»
392000, Российская Федерация, г. Тамбов, ул. Интернациональная, 33
✉ movlad2009@yandex.ru

Аннотация. Рассмотрены вопросы использования технологий искусственного интеллекта в документной деятельности. Изучены возможности использования искусственного интеллекта в работе с оперативной документацией на примере «1С: Документооборот 8 КОРП»: с целью распознавания текстов, голосового ввода, подготовки протоколов, резюме и др. Обобщены сведения о возможностях применения искусственного интеллекта в архивном деле, в первую очередь с целью поиска и распознавания архивной информации.

Ключевые слова: искусственный интеллект, документная деятельность, архивное дело, цифровые технологии в работе с документами

Для цитирования: Медведева О.В., Попова А.С. Искусственный интеллект в документной деятельности // Державинский форум. 2025. Т. 9. № 1. С. 106-111.

ORIGINAL ARTICLE
UDC 930.25

Artificial intelligence in documentary production

Olga V. MEDVEDEVA ✉, **Anastasia S. POPOVA**
Derzhavin Tambov State University
33 Internatsionalnaya St., Tambov, 392000, Russian Federation
✉ movlad2009@yandex.ru

Abstract. The study is devoted to the using artificial intelligence technologies issues in document activities. The using artificial intelligence possibilities in working with operational documentation are considered using the example of “1С: Document Management 8 CORP”: for the purpose of recognizing texts, voice input, preparing protocols, summaries, etc. Information on the using artificial intelligence possibilities in archival work is summarized, primarily for the purpose of searching and recognizing archival information.

Keywords: artificial intelligence, document management, archival work, digital technologies in working with documents

For citation: Medvedeva, O.V., & Popova, A.S. (2025). Artificial intelligence in documentary production. *Derzhavinskii forum = Derzhavin Forum*, vol. 9, no. 1, pp. 106-111.

ВВЕДЕНИЕ

Тема искусственного интеллекта на современном этапе становится одной из самых популярных, в том числе в связи с изданием «Национальной стратегии развития искусственного интеллекта на период до 2030 г.». Под искусственным интеллектом понимается комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их¹.

Казалось бы, наиболее масштабное использование искусственного интеллекта началось примерно в последние пять лет, за эти годы все узнали про ChatGPT, голосовые помощники и т. п. Однако на самом деле история искусственного интеллекта имеет длинные корни, уходящие в середину XX века. Именно в 1950-х гг. начали появляться экспертные системы, описывающие алгоритм действий по выбору решения в зависимости от конкретных условий, что уже тогда помогало принимать человеку решения. На смену им пришло машинное обучение, при помощи которого информационные системы самостоятельно формируют правила и находят решение на основе анализа зависимостей. Они используют исходные наборы данных без составленного человеком заранее перечня возможных решений. Учитывая такую значительную предысторию, можно предположить мощное развитие искусственного интеллекта во всех сферах, включая документную деятель-

ность. Однако в сфере управления документацией уже более десяти лет существуют мнения, что электронные секретари вытеснят человека, но этого до сих пор не произошло. Попытаемся выяснить, применимы ли технологии искусственного интеллекта в сфере работы с оперативной и ретроспективной (архивной) документацией [1].

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Основными задачами использования цифровых технологий в работе с документами являются оптимизация документооборота, сокращение затрат и, в конечном итоге, повышение качества работы специалистов сферы управления документацией. Непосредственно системы искусственного интеллекта применимы и полезны в операциях классификации документов, их поиска, получения из текстов необходимых метаданных.

Благодаря искусственному интеллекту процессы распознавания и поиска становятся более совершенными. Если ранее распознавание текстов бумажных документов и перевод их в электронный вид были основаны на разборе семантики, то есть распознавании букв и символов, исходя из их привязки к определенным позициям в бумажной форме документа, то современные системы искусственного интеллекта работают иначе. Новые технологии позволяют выделять отдельные смысловые данные из документа не в соответствии с привязкой к позиции в тексте документа, а в соответствии с их смыслом. Например, в тексте договора могут быть выделены атрибуты и словосочетания, которые характеризуют предмет договора, юридические адреса контрагентов и фамилии ответственных лиц, суммы договора и другие структурированные данные.

Ведущий ученый в сфере архивного дела Ю.Ю. Юмашева, изучив зарубежный опыт, приводит примеры разных разрабо-

¹ О развитии искусственного интеллекта в Российской Федерации: Указ Президента РФ от 10 октября 2019 г. № 490 (с изм. и доп.). URL: <https://base.garant.ru/72838946/?ysclid=m6gg5rfpha763489452> (дата обращения: 01.10.2024).

ток: «Итака», которая восстанавливает древнегреческие тексты, платформа OCR-распознавания, которая позволяет пользователям конвертировать изображения документов, написанных на иероглифическом японском языке, в печатный текст, и др. [2].

В качестве примера отечественных разработок рассмотрим возможности достаточно популярной системы, которая используется в службах документационного обеспечения, – «1С: Документооборот 8 КОРП». Разработчики заявляют о внедрении следующих цифровых технологий, построенных на работе искусственного интеллекта: распознавание аудио- и видеофайлов мероприятий с помощью сервиса «Таймлист», голосовой ввод данных для работы с процессами и документооборотом, автосоздание записей в календаре по анализу текста документа, распознавание документов, а также чат-бот «Ася»².

Сервис «Таймлист» предназначен для расшифровки аудио- и видеозаписей. Известно, что на смену стенографии, которая использовалась в процессе протоколирования заседаний и совещаний коллегиальных органов, давно пришли аудио- и видеопотоколирование в цифровых форматах, а на основе полученных записей человеком готовятся текстовые протоколы. Расшифрованный с помощью искусственного интеллекта текст может сразу поступать в работу, что позволяет сократить трудозатраты сотрудников. «Таймлист» помогает сделать из аудиозаписи и стенограмму, и краткое резюме (автопротокол). Данный сервис предлагает удобный доступ к результатам расшифровки и автопротоколам как для всех участников процесса, так и для пользователей, не участвовавших в мероприятии, что позволяет быстро оце-

нить содержание и принять необходимые решения.

Следующей функцией от разработчиков является голосовой ввод данных, который позволяет не только писать письма, но и подбирать для них адресатов, заполнять заголовки, описания для задач, вводить отчеты и резолюции. Эти возможности очень важны, в частности, для лиц с полной или частичной потерей зрения.

В записях календаря «1С: Документооборот 8 КОРП», созданных на основании письма, дата, время, место и наименование автоматически заполняются из текста данного письма. Программа распознает множество различных комбинаций текстов. В том случае, если данные не распознаны или отсутствуют в тексте письма, реквизиты новой записи календаря остаются незаполненными.

В «1С: Документообороте 8 КОРП» предусмотрены два режима распознавания документов: с использованием внешней программы CuneiForm и с использованием сервиса распознавания документов. Разработчики включили возможность распознавания различных форматов файлов: BMP, TIF, TIFF, JPG, JPEG, PNG, GIF, PDF, что сокращает издержки на ввод данных в программу с графических изображений документов.

Кроме всех вышеперечисленных функций, разработанных на основе технологий искусственного интеллекта, в программе присутствует чат-бот «Ася». Он может отвечать на вопросы, заданные на естественном языке, тем самым закрывая различные профессиональные задачи: кадровые (оформление отпусков, командировок, ознакомление с правилами для новых сотрудников), производственные (поиск государственных стандартов, строительных норм и правил и другой информации), административные (помощь сотрудникам в освоении и настройке компьютерной и офисной техники).

² Отраслевые и специализированные решения // 1С: Предприятие: официальный сайт. URL: <https://solutions.1c.ru/catalog/doc8-corp/digital> (дата обращения: 01.10.2024).

По мнению специалистов, с помощью искусственного интеллекта можно также решать такие задачи, как сравнительный анализ версий документов, формирование аннотаций, составление финансовой отчетности и категорирование финансовых документов [3]. Технологии искусственного интеллекта полезны и для сотрудников подразделений по управлению персоналом, так как, например, проверка резюме на соответствие требованиям вакансии может производиться без участия человека, что значительно экономит время.

Важная перспектива направления в развитии искусственного интеллекта связана с экспертизой рисков при заключении договоров, так как при использовании данных технологий будет возможно выявление отложенных или неоплаченных платежей, штрафных санкций и сроков, отведенных на предоставление документации, формирование вероятных замечаний. Одной из возможностей является также проверка договоров на соответствие законодательству [4].

Однако работа с оперативной документацией – не единственная возможность применения искусственного интеллекта. Новые перспективы открываются в работе с архивными документами. Подтверждением этому является то, что еще в 2018 г. в Международном совете архивов была создана профильная группа – Центр исследований искусственного интеллекта в архивах. Довольно активны в этом отношении национальные и профильные общественные объединения архивистов и специалисты крупнейших архивов мира, в частности, Великобритании, США, Нидерландов, Израиля и др. Новые программные продукты разрабатываются и применяются в целях автоматического извлечения и индексации контента, поиска в архивных информационных системах, создания связанных данных, в том числе с ресурсами музеев и библиотек. Венгерский разработчик программного обеспечения для

архивов и библиотек П. Кирали отмечает, что искусственный интеллект в архивной сфере может применяться с целью анализа текстов документов для архивной обработки, определения возможности доступа к документам, идентификации личных данных, оценки, регистрации, описания документов и др. Особое значение имеет облегчение процессов архивирования и научно-технической обработки интернет-ресурсов, в том числе из социальных сетей, электронной почты и т. п. [2, с. 225]. Способность обрабатывать и автоматически классифицировать большие объемы данных специалисты считают одной из наиболее перспективных возможностей для использования искусственного интеллекта в работе с архивами материалов, изначально созданных в цифровых форматах.

В нашей стране также имеется определенный опыт в данной сфере. Так, в Государственном архиве Российской Федерации круглосуточно работает цифровой ассистент НИКА, который выполняет функцию справочного бюро. Там же функционирует одноименная поисковая система НИКА, наполненная архивными данными из путеводителей архива и 134 фондов. НИКА используется в тех случаях, когда возникает потребность в сложном анализе информации и выявлении связей между архивными данными, и находится на стадии обучения³.

По данным Главархива Москвы, с 2020 г. функционирует онлайн-сервис «Моя семья», который содержит оцифрованные метрические книги, исповедные ведомости храмов, ревизские сказки податного населения Москвы и Московской губернии. На его материалах разработчики обучали нейросеть, которую стали ис-

³ В архивной отрасли задействован искусственный интеллект // АО «НЛЭ». URL: <https://www.nle.ru/media-center/corporate-media/magazine/May-2024/v-arkhivnoy-otrasli-zadeystvovan-iskusstvennyy-intellekt/> (дата обращения: 08.10.2024).

пользовать в проекте «Поиск по архивам»⁴. С помощью нейросети он распознает текст рукописного документа и делает его доступным для поиска неподготовленными пользователями. «Поиск по архивам» постоянно дополняется новыми видами документов, например, справочниками из фондов Национальной электронной библиотеки и Российской государственной библиотеки, газетами «Советский спорт», «Вечерняя Москва», «Сенатские ведомости». В базе сервиса представлено уже более 12 миллионов страниц исторических документов из архивов Москвы, Московской, Оренбургской, Вологодской, Иркутской, Астраханской и других областей, а также нескольких муниципальных архивов.

Однако при всех достоинствах новых технологий возникает опасность возникновения и тиражирования различных ошибок, а также доступа к конфиденциальным данным. Специалисты считают, что внедрение технологий искусственного интеллекта должно сопровождаться четкой структурой управления, соблюдением принципов этики и конфиденциальности. Крайне необходим междисциплинарный подход – с опорой на опыт архи-

вистов, социологов, историков и антропологов [5, с. 55].

ВЫВОДЫ

Таким образом, возможности использования искусственного интеллекта в документной деятельности пока еще не получили широкого распространения и должной разработки, несмотря на то, что данные вопросы стали активно подниматься уже на протяжении последнего десятилетия, что свидетельствует о спорности вопроса и сложности замены человека машиной. Однако выявлены те направления и операции, которые могут значительно ускорить и облегчить участие искусственного интеллекта. Отмечаются возможности более совершенного распознавания текстов, голосового ввода данных, отчасти – создания документов. Особенно большие перспективы открываются в отношении работы с большими массивами интернет-ресурсов, расшифровки рукописных архивных документов и поиска информации в массивах оцифрованных документов. Важно то, что в имеющиеся проекты вовлекается не только коммерческая сфера, но и государственные архивы, что способствует росту открытости и большей популяризации архивной информации.

⁴ Как искусственный интеллект помогает читать исторические рукописи из фондов Главархива // Главархив Москвы. URL: <https://cgamos.ru/news/e262722/> (дата обращения: 01.10.2024).

Список источников

1. Ахметкужин Р.А. История развития и перспективы применения искусственного интеллекта в правотворческой деятельности // E-Scio. 2023. № 2 (77). С. 532-539. <https://elibrary.ru/mdpilq>
2. Юмашева Ю.Ю. Историческая наука, архивы, библиотеки, музеи и искусственный интеллект: музеи и искусственный интеллект: год спустя // Документ. Архив. История. Современность. 2022. № 2. С. 217-241. <https://elibrary.ru/bppuwv>
3. Гордиенко В.В., Лисицин А.Л., Чистилина Е.В. Использование готовых решений на основе искусственного интеллекта в защите документооборота организаций // Auditorium. 2023. № 2 (38). С. 47-57. <https://elibrary.ru/npkxwj>

4. Катасев А.С., Катасева Д.В., Кирпичников А.П. Нейросетевая технология классификации электронных почтовых сообщений // Вестник технологического университета. 2015. Т. 18. № 5. С. 180-183. <https://elibrary.ru/toubhd>
5. Джайан Л., Капуто А. Раскрытие цифровых архивов: междисциплинарный подход к проблеме искусственного интеллекта и изначально цифровых данных // Новости Международной федерации библиотечных ассоциаций и учреждений. 2023. № 1 (122). С. 42-57.

References

1. Akhmetkuzhin R.A. (2023). The history of the artificial intelligence development and prospects in law-making. *E-Scio*, no. 2 (77), pp. 532-539. (In Russ.) <https://elibrary.ru/mdpilq>
2. Yumasheva Yu.Yu. (2022). Historical science, archives, libraries, museums and AI: a year later. *Dokument. Arkhiv. Istoriya. Sovremennost'*, no. 2, pp. 217-241. (In Russ.) <https://elibrary.ru/bppuwy>
3. Gordienko V.V., Lisitsin A.L., Chistilina E.V. (2023). Finished document flow of organizations with the help of the use of ready solutions based on artificial intelligence. *Auditorium*, no. 2 (38), pp. 47-57. (In Russ.) <https://elibrary.ru/npkxwj>
4. Katasev A.S., Kataseva D.V., Kirpichnikov A.P. (2015). Neural network technology for classifying electronic mail messages. *Vestnik tekhnologicheskogo universiteta = Herald of Technological University*, vol. 18, no. 5, pp. 180-183. (In Russ.) <https://elibrary.ru/toubhd>
5. Dzhaian L., Kaputo A. (2023). The disclosure of digital archives: An interdisciplinary approach to addressing the challenge of artificial intelligence and initial digital data. *Novosti Mezhdunarodnoi federatsii biblioteknykh assotsiatsii i uchrezhdenii = Newsletter of the International Federation of Library Associations and Institutions*, no. 1 (122), pp. 42-57. (In Russ.)

Информация об авторах

Медведева Ольга Владимировна, кандидат исторических наук, доцент, заведующий кафедрой библиотечно-информационных ресурсов, Тамбовский государственный университет им. Г.Р. Державина, г. Тамбов, Российская Федерация, movlad2009@yandex.ru

Попова Анастасия Сергеевна, магистрант по направлению подготовки «Документоведение и архивоведение», Тамбовский государственный университет им. Г.Р. Державина, г. Тамбов, Российская Федерация, stasya.kobzar@mail.ru

Information about the authors

Olga V. Medvedeva, Cand. Sci. (History), Associate Professor, Head of Library and Information Resources Department, Derzhavin Tambov State University, Tambov, Russian Federation, movlad2009@yandex.ru

Anastasia S. Popova, Master's Degree Student in "Documentation and archival science" Program, Derzhavin Tambov State University, Tambov, Russian Federation, stasya.kobzar@mail.ru

Статья поступила в редакцию / The article was submitted 16.11.2024

Одобрена после рецензирования / Approved after reviewing 15.01.2025

Принята к публикации / Accepted for publication 03.03.2025