

Историческая информатика

Правильная ссылка на статью:

Мащенко Н.Е., Гайдарь Е.В. Технологии искусственного интеллекта при формировании архивной среды: проблемы и перспективы // Историческая информатика. 2025. № 1. DOI: 10.7256/2585-7797.2025.1.73393 EDN: QEIGBR URL: https://nbpublish.com/library_read_article.php?id=73393

Технологии искусственного интеллекта при формировании архивной среды: проблемы и перспективы

Мащенко Наталья Евгеньевна

ORCID: 0000-0003-0126-545X

кандидат экономических наук

доцент по научной специальности "Документалистика, документоведение и архивоведение"; доцент; кафедра информационных систем управления; Донецкий государственный университет

283015, Россия, Донецкая Народная респ., г. Донецк, ул. Челюскинцев, д. 198А

✉ maschenko_n@mail.ru



Гайдарь Елена Валентиновна

ORCID: 0009-0008-3353-8831

кандидат экономических наук

доцент; кафедра информационных систем управления; Донецкий государственный университет

283015, Россия, Донецкая Народная респ., г. Донецк, ул. Челюскинцев, д. 198А

✉ e.gaydar.dongu@mail.ru



[Статья из рубрики "Искусственный интеллект и наука о данных"](#)

DOI:

10.7256/2585-7797.2025.1.73393

EDN:

QEIGBR

Дата направления статьи в редакцию:

17-02-2025

Аннотация: В статье рассматриваются перспективы использования технологий искусственного интеллекта (ИИ) для создания и развития цифровой архивной среды, а также их влияние на оптимизацию, автоматизацию процессов работы с архивными данными. Основной целью работы является анализ современных цифровых решений, направленных на улучшение процессов хранения, поиска и обработки архивных

документов (в том числе рукописных, поврежденных, многоязычных). В работе исследуются ключевые технологии, применяемые в цифровых архивах, включая интеллектуальное сканирование, обработку естественного языка (NLP), компьютерное зрение, машинное обучение и методы интеллектуального поиска. Особое внимание уделяется проблемам утраты архивных материалов, необходимости их восстановления, обеспечения безопасности и доступности данных, что особенно актуально в условиях нестабильной политической ситуации и ограниченных ресурсов для новых территорий. Исследование основано на системном анализе современных информационных технологий и их применении в архивном деле. В работе используются методы сравнительного анализа, классификации и прогнозирования, что позволяет определить ключевые направления внедрения ИИ в архивную сферу. Новизна работы заключается в комплексном подходе к анализу применения ИИ в архивной сфере, выявлении проблемных аспектов цифровизации архивов и предложении по автоматизации процессов хранения, обработки и поиска архивных данных. Сделан вывод о том, что технологии искусственного интеллекта способны значительно повысить эффективность работы архивов, обеспечивая ускоренную обработку документов, интеллектуальную классификацию, защиту данных и удобный доступ к информации. Кроме того, подчеркивается необходимость разработки новых алгоритмов на основе машинного обучения, которые позволят улучшить распознавание рукописных текстов, обработку поврежденных документов и многоязычных архивных материалов. Внедрение таких технологий становится важной частью стратегии цифровой трансформации архивного дела и играет ключевую роль в сохранении исторического наследия.

Ключевые слова:

архивы, цифровая архивная среда, цифровая трансформация, искусственный интеллект, машинное обучение, компьютерное зрение, обработка естественного языка, безопасность данных, интеллектуальное сканирование, интеллектуальное прогнозирование

Введение. Архивные системы сталкиваются с растущими объемами данных, разнообразием форматов и требованиями к надежности и доступности информации. В этих условиях внедрение искусственного интеллекта становится стратегически важным шагом для создания эффективной архивной среды, способной справляться с вызовами цифровой эпохи.

Архивная среда представляет собой сложную систему, включающую в себя архивные учреждения и архивные подразделения, а также совокупность архивных материалов, которые образуют архивное пространство и обеспечивают сохранение, обработку и использование архивных данных посредством различных методов и технологий под влиянием определенных факторов [\[1\]](#).

Основная часть.

Современный этап развития общества характеризуется эпохой цифровых преобразований, оперированием колоссальными объемами информации, доступом к мировым инновационным процессам и значительным развитием способов использования предприятиями различных информационных технологий. В цифровой среде развитие и эффективное ведение деятельности становится невозможным без применения современных информационных систем и технологий [\[2\]](#).

Формирование цифровой архивной среды предполагает комплексный процесс перехода архивных учреждений и организаций к цифровым технологиям, направленный на обеспечение сохранности, доступности и удобства использования архивных данных. Это важный этап для архивной сферы, который включает в себя как технические, так и организационные изменения, способствующие улучшению работы с документами в цифровом формате.

Цифровизация архивов начинает активное использование инструментов искусственного интеллекта.

Искусственный интеллект позволяет автоматизировать, оптимизировать и улучшить процессы, связанные с созданием, хранением, поиском и анализом архивных данных. Основные задачи, решаемые с помощью ИИ: ускорение обработки больших объемов данных; обеспечение интеллектуального поиска и управления архивами; повышение безопасности и защита информации; автоматизация процессов классификации и индексации документов [\[3,4\]](#).

Эти возможности делают ИИ ключевым инструментом для построения современной архивной среды.

Формирование цифровой архивной среды в ДНР является важной частью стратегии по модернизации государственных и муниципальных институтов, а также обеспечения сохранности, доступности и безопасности архивных данных. Архивы, как хранители исторической, юридической и культурной информации, играют ключевую роль в социально-экономическом развитии региона. В условиях нестабильной политической ситуации и ограниченных ресурсов, переход к цифровым технологиям помогает решать задачи сохранения и упрощения доступа к архивным материалам, повышая эффективность их управления.

Цифровизация архивов в ДНР становится не только способом улучшения архивного дела, но и важной составляющей национальной безопасности и государственной политики по сохранению исторического наследия. В этом контексте создание цифровой архивной среды включает в себя несколько ключевых этапов, от создания цифровых реплик документов до их защиты с помощью современных технологий.

В условиях военных действий на территории Донецкой Народной Республики, как и на других освобожденных территориях, проблема повреждения и утраты архивов и архивных документов остается крайне актуальной. Утрата документов приводит к серьезным трудностям для граждан в процессе обеспечения их гражданских прав и свобод.

Так, в Мариуполе, Авдеевке и ряде других городов, процедура восстановления документов для подтверждения личности связана с длительными и сложными процессами, а восстановление трудового стажа и прав собственности в ряде случаев не представляется возможным. Архивы организаций часто уничтожались в результате пожаров, взрывов вместе с трудовыми книжками, личными делами и другими документами, что лишает граждан возможности подтвердить стаж трудовой деятельности, уровень образования и другие важные аспекты их биографии.

Кроме того, во время эвакуации из населенных пунктов организации оставляли документы и архивы на месте, где они находятся в ненадлежащих условиях, например, в мешках или подвалах, россыпью, требуют дальнейшей сортировки, обработки и передачи

в архивы.

Архивные учреждения Донецка в связи с отсутствием необходимых средств и другими организационными факторами, еще в период прежнего украинского руководства находились в неудовлетворительном состоянии. Здания архивов находятся в состоянии физического износа: протекают крыши, не соблюдаются условия температурно-влажностного режима и нарушены другие параметры, необходимые для сохранности документов. В результате это привело к повреждению и частичной или полной утрате архивных материалов: документы залиты водой, рассохлись и т.д.

Граждане, в свою очередь, зачастую не знают, куда обращаться для поиска информации, необходимой для восстановления документов. Это приводит к серьезным социальным последствиям: люди вынуждены жить без документов, что ограничивает их доступ к социальным пособиям, жилью и другим правам, усугубляя социальное неблагополучие и провоцируя конфликты.

Эскалация военно-политической ситуации в регионе показывает на высокую вероятность повторения подобных сценариев в будущем. Таким образом, для решения указанных проблем необходимо разработать и внедрить эффективные алгоритмы и инструменты на основе технологий искусственного интеллекта, что позволит существенно ускорить и упростить процессы восстановления архивных данных.

Для предотвращения описанных выше негативных сценариев в будущем необходима масштабная работа по их оцифровке, которая предполагает внедрение современных архивных информационных систем, использование специализированных сканирующих устройств, а также применение инструментов искусственного интеллекта для автоматизации процессов обработки, классификации, поиска, применения аналитики и обеспечения безопасности.

В условиях цифровой трансформации возможность применения решений на основе искусственного интеллекта является одной из основных целей и в качестве приоритетного направления научно-технологического развития Российской Федерации на ближайшие 10–15 лет согласно Указу Президента Российской Федерации от 1 декабря 2016 года № 642 «О Стратегии научно-технологического развития Российской Федерации» [\[5\]](#).

Согласно Presidential Actions. Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, October 30, 2023 «Модель ИИ» означает компонент информационной системы, который реализует технологию ИИ и использует вычислительные, статистические методы или методы машинного обучения для получения результатов из заданного набора входных данных.

Выделим основные инструменты ИИ, которые могут применяться в подобных ситуациях.

1. Одним из наиболее перспективных направлений является **интеллектуальное сканирование и распознавание архивных документов**. В этом случае ИИ значительно экономит время сотрудников и нивелирует роль так называемого человеческого фактора – снижает количество непредвиденных ошибок на этапе обработки данных [\[6\]](#).

Интеллектуальное сканирование – это процесс, при котором документы преобразуются в цифровой формат с использованием технологий анализа и обработки изображений. Распознавание архивных документов включает использование алгоритмов искусственного интеллекта (ИИ) и машинного обучения для извлечения текстовой и

структурной информации из цифровых изображений.

Эти технологии объединяют такие инструменты, как:

оптическое распознавание символов (OCR): наибольшее применение получили движки PyLaia, Kraken, Calamari, Tesseract. Исходные коды перечисленных движков и документация к ним размещены на сайте GitHub в открытом доступе [\[7,8\]](#);

обработка естественного языка (NLP): анализ текста для выделения ключевой информации (Amazon Textract и Google Cloud Vision): облачные решения для анализа и обработки документов с использованием ИИ; ИИ-модели на базе OpenAI и GPT: используются для обработки и анализа текста, включая выделение смысла и создание аннотаций;

машинное зрение: распознавание форм, печатей и других визуальных элементов.

Технологии интеллектуального распознаванияприобретают особую инновационность в задачах обработки рукописных текстов (HTR), особенно когда речь идет о сложных или нестандартных почерках, характерных для старинных документов, что требует дополнительной обработки и обучения алгоритма для высокой точности. В 2022 г. M. Terras [\[9\]](#) представила платформу Transkribus.

Особое внимание необходимо уделить обработке поврежденных материалов, где процесс осложняется наличием пятен, разрывов, выцветшего или обгоревшего текста. Для таких случаев требуются специальные алгоритмы, способные адаптироваться к частичной утрате информации.

Многоязычность (работа с документами, представленными на разных языках, например, на украинском, что актуально для новых территорий) – отдельная проблема для интеллектуального распознавания, предполагает использование адаптированных языковых моделей, лексических и грамматических знаков.

Наконец, большой объем данных в архивах, часто включающий миллионы документов, вызывает необходимость использования высокопроизводительных вычислительных ресурсов для эффективной обработки, хранения и анализа массивов информации.

2 . Интеллектуальная классификация архивов – это процесс автоматического разделения архивных материалов на категории с использованием передовых технологий ИИ: машинного обучения (ML), обработки естественного языка (NLP) и компьютерного зрения. Этот процесс отличается от традиционной классификации тем, что ИИ анализирует не только текстовые и визуальные элементы документов, но и их контекст, значение и взаимосвязи. Такой подход значительно улучшает организацию и доступность архивных данных.

Алгоритмы *машинного обучения* играют ключевую роль в классификации архивных данных. Модели обучаются на примерах уже отсортированных данных, что позволяет системе научиться автоматически классифицировать новые документы. Для этого применяются следующие подходы:

обучение с учителем: модели обучаются на заранее размеченных данных, где каждый документ уже имеет соответствующую метку (например, «финансовый отчет», «научная статья», «личное письмо»);

обучение без учителя: используется для кластеризации материалов, когда заранее

неизвестно, к какой категории относится документ; система автоматически группирует материалы по схожим признакам;

полуобучаемые алгоритмы: подход, когда система сначала работает с размеченными данными, а затем использует малую информацию о новых данных для их классификации.

Например, E. Shang и др. [\[10\]](#) представили эффективный метод классификации архивов с использованием XGBoost и Spark computing.

Технологии *обработки естественного языка* помогают ИИ «понимать» текстовый контент документов и классифицировать их по смыслу, а не только по ключевым словам. Например, ИИ может различать юридические документы, научные статьи и письма на основе их содержания. Основные методы NLP, применяемые для классификации, включают:

тематическое моделирование: разделение документов на категории, связанные с их темой (например, политика, экономика, искусство) [\[11\]](#);

анализ тональности: оценка эмоциональной окраски документа для классификации документов по типу (например, официальный, личный, деловой);

выделение ключевых слов и фраз: это помогает создать индексы и теги, которые облегчают поиск и сортировку документов.

В архивах, содержащих визуальные материалы (фотографии, карты, чертежи и т.д.), ИИ применяется для классификации изображений. С помощью *компьютерного зрения* система способна распознавать объекты на изображениях и классифицировать их по категориям. Например, фотографии могут быть автоматически разделены на группы (например, «портреты», «пейзажи», «события»). Это возможно благодаря использованию алгоритмов глубокого обучения (Deep Learning): нейронные сети (например, сверточные нейронные сети, CNN) могут распознавать объекты и сцены на изображениях, а рекуррентные нейронные сети (RNN) могут работать с последовательными текстовыми данными, такими как документы с большим объемом текста или исторические рукописи.

K. Carter [\[12\]](#) использовали двухэтапный алгоритм контролируемого машинного обучения, который использует такие инструменты, как Google AutoML Vision для обработки оглавления страниц. A. Männistö [\[13\]](#) представили платформу автоматического извлечения содержимого изображений (AICE).

При интеллектуальной классификации используется также распознавание текста на изображениях: для классификации документов, содержащих текст (например, письма, книги), используется технология OCR, которая помогает извлечь текст из изображений и провести классификацию по его содержанию; в архивах, содержащих аудио- и видеоматериалы, используется распознавание речи для преобразования аудио в текст, а также анализ видеопотока для классификации сцен, объектов или событий, происходящих в видео.

Процесс интеллектуальной классификации архивных материалов включает несколько этапов.

На этапе предобработки данных система очищает и подготавливает архивные данные для анализа. В случае текстовых данных это может быть удаление шума (ненужных символов, пробелов), исправление ошибок и форматирование текста, для изображений –

улучшение качества изображения, выравнивание или восстановление поврежденных материалов; для мультимедиа – очистка звука, синхронизация данных.

На этапе извлечения признаков система анализирует содержание документов, извлекая из них значимые признаки для классификации. (для текстов – это извлечение ключевых слов, тем, фраз, адресов, имен и др.; для изображений – это распознавание объектов, фотографий, лиц, текста и т.д., для мультимедиа – извлечение речи и преобразование ее в текст, анализ контекста).

На этапе обучения модели система обучается на примерах, учится распознавать структуру данных, создается интеллектуальная модель классификации архивных материалов на основе алгоритмов машинного обучения. Обучение включает построение классификационных моделей, таких как нейронные сети, деревья решений, кластерные алгоритмы; тестирование модели на новых данных для проверки ее эффективности и точности; постепенное дообучение по мере поступления новых данных.

Классификация и категоризация – после завершения обучения система применяет полученные знания для анализа новых архивных материалов. На основе полученных знаний система классифицирует новые данные по заданным категориям. Это может быть сделано по меткам, таким как «финансовые документы», «исторические записи», «научные статьи» и т.д. Система может интегрироваться с архивными системами, автоматически добавляя категории и помогая сотрудникам быстрее ориентироваться в данных.

Автоматическая индексация – завершающий этап классификации, на котором классифицированным материалам присваиваются метаданные для упрощения поиска и управления. Все классифицированные документы или материалы автоматически индексируются и получают метаданные, что упрощает их дальнейший поиск. Классифицированные и индексируемые материалы становятся доступными для поиска с использованием простых запросов или сложных фильтров.

3. Интеллектуальный поиск архивных документов – это процесс извлечения информации из архивных данных с использованием технологий ИИ, NLP, ML и других передовых подходов. Он предполагает не только поиск по ключевым словам, но и более глубокое понимание контекста, тематики, связи между документами и их содержимым. Пользователь может задавать вопросы в свободной форме, а ИИ подбирает релевантные документы с учетом контекста. Такой подход значительно повышает эффективность работы с большими объемами данных и улучшает доступность архивов для пользователей.

Технологии NLP позволяют анализировать и понимать текстовый контент архивных документов. NLP помогает выделять ключевые слова, темы и сущности; анализировать синтаксис и грамматику текста; извлекать смысл и контекст из сложных фраз; применять синонимы и контекстуальные термины для улучшения результатов поиска.

A. Alothman и Abdul Sait ^[14] внедрили алгоритм ранжирования для эффективного поиска документов. M. Modiba ^[15] рассматривает несколько преимуществ эффективного поиска информации с особым акцентом на управление записями.

Алгоритмы ML обучаются на исторических данных и могут улучшать свои результаты со временем, адаптируясь к запросам пользователей.

Например, семантический поиск позволяет учитывать не только точные совпадения слов,

но и их значение в контексте. Например, система понимает, что «финансовый отчет» и «бухгалтерский баланс» – это разные фразы, но обе относятся к одному направлению.

Семантические модели включают:

модели на основе векторного представления слов (Word2Vec, GloVe): эти модели представляют слова как векторы, что позволяет выявлять семантические связи между ними;

трансформеры (BERT, GPT): современные модели трансформеров позволяют анализировать текст с учетом контекста, что значительно повышает качество поиска.

В архивах может использоваться метод распознавания сущностей, который применяется для извлечения специализированных терминов, имен, дат и другие важных аспектов, которые могут быть использованы для улучшения поиска. Например, ИИ может распознавать имена авторов, организации, географические объекты и события, что помогает более точно классифицировать документы.

Используя методы идентификации и категоризации по типу документа, ИИ может автоматически распознавать тип документа (например, отчет, письмо, контракт, научная статья) и применять соответствующие фильтры и категории для упрощения поиска.

ИИ-системы также могут извлекать и анализировать метаданные (автор, дата, ключевые слова, тема), что позволяет ускорить поиск и повысить точность результатов.

Архивы часто содержат не только текстовые документы, но и изображения, аудио, видео и другие типы данных. Интеллектуальный поиск может интегрировать данные из разных источников и анализировать мультимедийные материалы:

анализ изображений и видео: с помощью компьютерного зрения можно распознавать текст на изображениях (OCR), а также анализировать содержание изображений и видеоконтента для поиска;

распознавание речи: видеоматериалы и аудиофайлы могут быть обработаны с помощью технологии распознавания речи, что позволяет извлекать текстовую информацию из аудиовизуальных данных.

Y. Yang [\[16\]](#) представил усовершенствованную модель преобразования текста в видео, разработанную специально для аудиовизуальных архивов, демонстрирующую развитие управления мультимедийной информацией.

4. Интеллектуальная аналитика и прогнозирование. Использование машинного обучения позволяет выявлять закономерности в архивных данных, прогнозировать потребности пользователей или оптимизировать процессы хранения.

Используя технологии интеллектуального анализа данных, архивные системы могут автоматически проводить *анализ структуры и содержания архивов*, классифицировать документы, выделять ключевые темы и определять взаимосвязи между различными материалами. Это упрощает доступ к информации и повышает точность поиска.

С помощью алгоритмов машинного обучения архивы могут *прогнозировать состояние документов, подверженных риску повреждения*, например, из-за несоблюдения условий хранения. Это позволяет своевременно принимать меры для их сохранения.

Системы ИИ способны *анализировать запросы пользователей*, выявляя наиболее

востребованные документы или темы. Эти данные используются для оптимизации работы архивов, например, для приоритетности оцифровки определенных материалов.

Системы ИИ могут вести *мониторинг и управление архивными фондами*. На основе данных об условиях хранения, таких как температура, влажность или уровень освещенности, интеллектуальные системы могут предлагать рекомендации по улучшению условий и предотвращению разрушения документов.

Используя исторические данные, архивы могут *прогнозировать тенденции в запросах пользователей*, то есть какие темы или материалы станут востребованными в будущем, например, в связи с социальными или политическими событиями.

Аналитические системы на базе ИИ могут *предлагать решения* по оптимальному распределению ресурсов, таких как оцифровка, реставрация или организация новых выставок.

ИИ способен *обеспечивать безопасность данных*, выявлять подозрительные действия (например, несанкционированный доступ) и предотвращать утечку данных. Он также может обеспечивать автоматическое шифрование и резервное копирование.

5. Основные направления применения искусственного интеллекта для обеспечения безопасности данных в архивах:

Системы на базе ИИ способны анализировать огромные объемы данных в режиме реального времени, выявляя потенциальные угрозы, такие как попытки несанкционированного доступа, аномальные действия пользователей или вредоносное программное обеспечение.

ИИ может оптимизировать и адаптировать методы шифрования, обеспечивая защиту данных как при их хранении, так и при передаче. Технологии машинного обучения позволяют создавать динамически изменяющиеся шифровальные алгоритмы, затрудняющие их взлом.

Искусственный интеллект способен анализировать поведение пользователей и создавать индивидуальные модели доступа. Например, при обнаружении подозрительных действий система автоматически ограничивает или блокирует доступ.

Алгоритмы ИИ используются также для прогнозирования и предотвращения кибератак, включая фишинг, SQL-инъекции и DDoS-атаки. Системы машинного обучения анализируют предыдущие инциденты, выявляют слабые места и предлагают меры для их устранения.

ИИ способен отслеживать перемещение данных внутри архивных систем, выявляя несанкционированные копии, передачи или модификации документов.

Системы на базе ИИ могут автоматически анализировать журналы событий и аудита, выявляя подозрительные действия и обеспечивая прозрачность всех операций.

Синергия блокчейн-технологий и ИИ позволяет фиксировать все изменения в архивных данных и предотвращать их фальсификацию. Кроме того, ИИ способствует оптимизации процедур верификации, что упрощает управление цифровыми записями.

Заключение. Искусственный интеллект открывает новые горизонты для архивной среды, обеспечивая скорость, удобство и безопасность работы с информацией. Его внедрение позволяет не только улучшить текущие процессы, но и адаптироваться к вызовам

будущего, создавая более интеллектуальные и эффективные архивные системы. Архивы, интегрирующие ИИ, становятся не просто хранилищами данных, но динамичными центрами знаний, готовыми к использованию в любой момент.

Интеллектуальное сканирование и распознавание архивных документов – это не просто тренд, а необходимость в условиях цифровой трансформации. Эти технологии открывают новые возможности для автоматизации и упрощения работы с архивами, обеспечивая сохранность, доступность и удобство использования данных. Инвестирование в развитие и внедрение таких решений станет ключевым фактором успеха для архивных учреждений в XXI веке.

Библиография

1. Машенко Н. Е. Формирование архивной среды как элемента социокультурного пространства // Донецкие чтения 2023: Образование, наука, инновации, культура и вызовы современности: Материалы VIII Международной научной конференции. Донецк, 2023. С. 91-93.
2. Гайдарь Е. В. Современные информационные системы и технологии в условиях цифровой трансформации бизнеса // Экономика: Сб. науч. работ ГОУ ВПО «ДОНАУИГС». 2022. Вып. 25. С. 47-57.
3. Белов И. И. Роль технологий искусственного интеллекта в цифровой трансформации делопроизводства и архивного дела // Научный вестник Крыма, 2022. № 4 (39). С. 1-6.
4. Лобачев С. Л. Место искусственного интеллекта в подготовке специалистов по документоведению и архивоведению // Вестник Юридического института МИИТ. 2021. № 2 (34). С. 135-142.
5. Ильина К. Б. Искусственный интеллект в архивах: опыт применения в Российской Федерации, проблемы и перспективы // Архивы и электронные документы: вызовы времени : Доклады и сообщения Международной научно-практической конференции, Самарканд, 20 сентября 2023 года. Москва: ВНИИДАД, 2024. С. 144-152.
6. Шалков Д. Ю. Искусственный интеллект в документоведении: эргономика профессиональной деятельности // Управление информацией и документацией в цифровой среде: сб. науч. ст. по материалам III Всероссийской научно-практической конференции (Донецк, 21-22 ноября 2024 г.) / ред. коллегия: Пономаренко Н. Ш. (председатель) [и др.]; ФГБОУ ВО «Донецкий государственный университет» [и др.]. Донецк: ДонГУ, 2024. С. 124-132.
7. Киселев И. Н. О применении искусственного интеллекта в распознавании текстов // Вестник ВНИИДАД. 2024. № 1. С. 84-95.
8. Давлетов А. Р. Современные методы машинного обучения и технология OCR для автоматизации обработки документов // Вестник науки. 2023. Т. 5, № 10 (67). С. 676-698. – DOI: <https://doi.org/10.24412/2712-8849-2023-1067-676-698>.
9. Terras M. Inviting AI into the archives: The reception of handwritten recognition technology into historical manuscript transcription // *Archives, Access and Artificial Intelligence*. 2022. December. Pp. 179-204. – DOI: <https://doi.org/10.1515/9783839455845-008>.
10. Shang E., Liu X., Wang H., Rong Y., & Liu Y. Research on the application of artificial intelligence and distributed parallel computing in archives classification // *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. 2020. February. Pp. 1267-1271. – DOI: <https://doi.org/10.1109/IAEAC47372.2019.8997992>
11. Haffenden C., Fano E., Malmsten M., & Börjeson L. Making and using AI in the library: Creating a BERT model at the National Library of Sweden // *College & Research Libraries*. 2023. № 84(1). – DOI: <https://doi.org/10.5860/crl.84.1.30>
12. Carter K., Gondek A., Underwood W., Randby T., & Marciano R. Using AI and ML to

- optimize information discovery in under-utilized, Holocaust-related records // *AI & Society*. 2022. № 37. May. Pp. 837-858. – DOI: <https://doi.org/10.1007/s00146-021-01368-w>
13. Männistö A., Seker M., Iosifidis A., & Raitoharju J. Automatic image content extraction: Operationalizing machine learning in humanistic photographic studies of large visual archives // *arXiv*. 2022. April. – DOI: <https://doi.org/10.48550/arXiv.2204.02149>
14. Alothman A., & Sait A. Managing and retrieving bilingual documents using artificial intelligence-based ontological framework // *Computational Intelligence and Neuroscience*. 2022. August. Pp. 1-15. – DOI: <https://doi.org/10.1155/2022/4636931>
15. Modiba M. User perception on the utilization of artificial intelligence for the management of records at the Council for Scientific and Industrial Research // *Collection and Curation*. 2023. № 42(3). Pp. 81-87. – DOI: <https://doi.org/10.1108/CC-11-2021-0033>
16. Yang Y. Write what you want: Applying text-to-video retrieval to audiovisual archives // *arXiv*. 2023. October. – DOI: <https://doi.org/10.48550/arXiv.2310.05825>

Результаты процедуры рецензирования статьи

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов издательства можно ознакомиться [здесь](#).

Последние десятилетия отмечены бурным ростом информационно-коммуникационных технологий, которые кардинальным образом меняют нашу повседневную жизнь. Конечно, такие перемены нередко имеют как положительные, так и отрицательные стороны: чего стоят только споры вокруг использования искусственного интеллекта, в том числе применимо к научным исследованиям. В то же время возможности искусственного интеллекта весьма многогранны, в связи с чем представляется важным обратиться к изучению его использования в различных сферах исторической информатики.

Указанные обстоятельства определяют актуальность представленной на рецензирование статьи, предметом которой являются технологии искусственного интеллекта при формировании архивной среды. Автор ставит своими задачами рассмотреть формирование цифровой архивной среды на примере ДНР, проанализировать такие направления, как интеллектуальное сканирование и распознавание архивных документов, интеллектуальная аналитика и прогнозирование, показать направления применения искусственного интеллекта для обеспечения безопасности данных в архивах.

Работа основана на принципах анализа и синтеза, достоверности, объективности, методологической базой исследования выступает системный подход, в основе которого находится рассмотрение объекта как целостного комплекса взаимосвязанных элементов. Научная новизна статьи заключается в самой постановке темы: автор стремится охарактеризовать проблемы и перспективы технологии искусственного интеллекта при формировании архивной среды.

Рассматривая библиографический список статьи как позитивный момент следует отметить его масштабность и разносторонность: всего список литературы включает в себя 16 различных источников и исследований. Несомненным достоинством рецензируемой статьи является привлечение зарубежной англоязычной литературы, что определяется самой постановкой темы. Из используемых автором трудов укажем на работы И.И. Белова, С.Л. Лобачева, М. Терраса и других, в центре внимания которых находятся различные аспекты изучения применения искусственного интеллекта в архивах. Заметим, что библиография статьи обладает важностью как с научной, так и с просветительской точки зрения: после прочтения текста статьи читатели могут

обратиться к другим материалам по ее теме. В целом, на наш взгляд, комплексное использование различных источников и исследований способствовало решению стоящих перед автором задач.

Стиль написания статьи можно отнести к научному, вместе с тем доступному для понимания не только специалистам, но и широкой читательской аудитории, всем, кто интересуется как искусственным интеллектом, в целом, так и его возможностями в архивном деле, в частности. Аппеляция к оппонентам представлена на уровне собранной информации, полученной автором в ходе работы над темой статьи.

Структура работы отличается определенной логичностью и последовательностью, в ней можно выделить введение, основную часть, заключение. В начале автор определяет актуальность темы, показывает, что в современных условиях «внедрение искусственного интеллекта становится стратегически важным шагом для создания эффективной архивной среды, способной справляться с вызовами цифровой эпохи». В работе отмечаются такие перспективные направления, как интеллектуальное сканирование и распознавание архивных документов, что «значительно экономит время сотрудников и нивелирует роль так называемого человеческого фактора – снижает количество непредвиденных ошибок на этапе обработки данных», а также «когда речь идет о сложных или нестандартных почерках, характерных для старинных документов». Автор обращает внимание на то, что «синергия блокчейн-технологий и ИИ позволяет фиксировать все изменения в архивных данных и предотвращать их фальсификацию». В конечном итоге, как справедливо отмечает автор рецензируемой статьи, «архивы, интегрирующие ИИ, становятся не просто хранилищами данных, но динамичными центрами знаний, готовыми к использованию в любой момент».

Главным выводом статьи является то, что «искусственный интеллект открывает новые горизонты для архивной среды, обеспечивая скорость, удобство и безопасность работы с информацией».

Представленная на рецензирование статья посвящена актуальной теме, вызовет читательский интерес, обобщает отечественный и зарубежный опыт, а ее материалы могут быть использованы как в учебных курсах, так и в рамках работы архивных учреждений.

В целом, на наш взгляд, статья может быть рекомендована для публикации в журнале «Историческая информатика».