

Акмаев Роман Николаевич,
Магистрант направления «Бизнес-информатика»,
Кубанский Государственный Университет,
Краснодар

**ЦИФРОВИЗАЦИЯ ПРОЦЕССА
ИНФОРМАЦИОННО-ДОКУМЕНТАЦИОННОГО
ОБЕСПЕЧЕНИЯ УПРАВЛЕНИЯ С ПОМОЩЬЮ
СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
DIGITIZATION OF THE PROCESS OF INFORMATION
AND DOCUMENTATION SUPPORT OF MANAGEMENT
USING ARTIFICIAL INTELLIGENCE SYSTEMS**

Аннотация: в статье рассматривается роль цифровизации в контексте развития информационно-документационного обеспечения управления. Рассмотрены варианты применения систем искусственного интеллекта в документообороте организации. Выделены основные аспекты, влияющие на успешное применение систем искусственного интеллекта в организации.

Ключевые слова: цифровизация, искусственный интеллект, цифровизация документооборота, инновации.

Вероятнее всего, каждый из нас регулярно слышит о таком понятии как цифровизация, однако далеко не все смогут дать четкое определение данного понятия. Так что же такое цифровизация и как она меняет наш мир?

Цифровизация – это внедрение и использование цифровых технологий в различных сферы жизни общества [1].

Однако цифровизация не означает просто применение информационных технологий в бизнесе или повседневной жизни. Цифровизация – это своего рода смена парадигмы восприятия мира, которая позволяет нам по-другому взглянуть на наши подходы к ведению бизнеса, бизнес-процессы, выстроенные в компании и бизнес-модели, по которым наши компании развиваются.

Кроме того, необходимо понимать, что цифровизация бизнеса, в особенности нужна тем компаниям, которые хотят укрепить или хотя бы сохранить свои позиции на рынке в ближайшем будущем. Из-за геополитической напряженности, роста инфляции и экономического кризиса, российским компаниям предстоит решать множество проблем в разных направлениях, таких как: проблемы с поставками импортных товаров в страну, постоянно растущими потребностями потребителей и устойчивым развитием. Но также стоит помнить о том, что кризисные ситуации дают огромные возможности для развития, и успеха добиваются те компании, которые вовремя смогли воспользоваться данными возможностями.

Многие люди путают понятие цифровизация и автоматизация. Так некоторые считают, что если в компании используются современные информационные технологии, то эта компания уже можно назвать цифровизированной, однако нужно понимать, что информационные технологии – это лишь инструменты, которые помогают облегчить и ускорить осуществление некоторых процессов в работе компании, но сами эти инструменты не ведут к цифровизации компании.

Основным отличием цифровизации от автоматизации можно назвать то, что цифровизация направлена на изменение самой бизнес-модели и бизнес-процессов компании. То есть при автоматизации происходит усовершенствование производственных и рабочих



процессов в компании, но при этом сам способ ведения дел не изменяется или изменяется слабо. А вот при цифровизации происходит именно качественное изменение продукта или услуги производимых компаний, изменяются взаимоотношения с контрагентами, клиентами и обществом. Изменяется даже позиционирование самой компании на рынке [7].

Если же говорить о цифровизации в контексте развития информационно-документационного обеспечения управления (ДООУ) в современной организации, то многие согласятся с тем, сейчас уже довольно-таки трудно найти большую и успешную компанию, использующую в своей работе только бумажные документы. Так как цифровизация документооборота существенно ускоряет бизнес-процессы, выстроенные в компании, и кроме того, повышает их эффективность и делает их более удобными и понятными для работников компании. Так же цифровизация документооборота существенно ускорит процесс принятия и исполнения управленческих решений, и однозначно повысит их эффективность.

После того, как в конце 2022 года вышла в свет языковая модель ChatGPT, искусственный интеллект (ИИ) пережил новый взлет своей популярности. ИИ является одной из сквозных технологий, заложенных в национальную программу «Цифровая экономика». Так сквозные технологии – это перспективные технологии, способные радикально изменить ситуацию на нынешних рынках. Представители научного сообщества, по-разному дают определение понятию ИИ. Однако в самом широком значении под ИИ понимают способность компьютера решать такие же интеллектуальные задачи, которые решает человек [2].

Специалисты различных областей стали искать варианты применения ИИ в своей профессиональной деятельности, в частности и в процессе документационного обеспечения управления. Одним из направлений развития ИИ в данной сфере является оптимизация рабочих процессов, сокращение затрат на выполнение делопроизводственных задач и повышение качества работы специалистов отрасли ДООУ. Внедрение ИИ в работу компании помогает не только сэкономить денежные и временные ресурсы, но и делает работу с документами значительно легче и доступнее. ИИ начал заменять старые методы управления документами, когда приходилось вручную, даже с использованием СЭД, создавать, проверять, согласовывать и подписывать огромное количество документов. С помощью ИИ можно оптимизировать документооборот, сократив количество бумажных документов, переведя их в цифровой формат. Это позволит повысить координацию работы, облегчить обмен информацией, улучшить коммуникацию внутри компании и получить доступ к необходимой информации в режиме реального времени.

В настоящее время уже реализуются проекты по внедрению ИИ в корпоративные системы электронного документооборота. Такие ИИ-модули могут распознавать текст документа, определять его вид, преобразуют для последующего атрибутивного либо полнотекстового поиска. Данные из документа вносятся в учетную систему, а специалисту по документообороту нужно только проверить корректность этих данных и при необходимости скорректировать или до заполнить данные.

На основании содержания документов, ИИ определяет состав утверждающих и инициирует задачи по определенным маршрутам. ИИ анализирует эффективность всех бизнес-процессов, выявляет наиболее трудоемкие этапы и перегруженных исполнителей. Кроме этого, генеративные текстовые модели ИИ могут создавать осмысленные проекты документов, соответствующие контексту задачи и заданным параметрам.

Типовой процесс работы с входящим документом в организации можно разделить на 8 основных этапов: регистрация, рассмотрение, исполнение, подготовка проекта ответного документа, согласование, подписание, исходящая регистрация, рассылка. На некоторых этапах можно успешно применять технологии ИИ.



На этапе регистрации документа, системы ИИ можно применять для автоматического заполнения данными регистрационной карточки документа. ИИ хорошо справляется со считыванием данных из формализованных документов, таких как УПД, транспортная накладная, счет-фактура, КС-2, КС-3 и т.д. А вот когда, приходит неформализованный документ, с неструктурированным текстом, произвольного вида, такими как письмо, коммерческое предложение, электронная переписка с контрагентами, то здесь системы ИИ сталкиваются с большими трудностями и не показывают таких же хороших результатов.

На этапе рассмотрения документа, ИИ можно применять для автоматической маршрутизации документа, в частности определения его исполнителя. Перед системой ИИ стоит задача автоматически маршрутизировать входящий документ до структурного подразделения и конкретного сотрудника, который будет отвечать за его обработку и исполнение. В данном случае, входящими данными для ИИ системы будут являться текстовые слои с файлов, приложенных к документу, и дополнительная метаданная (отправитель, вид доставки и т. д.). Применяется для подобных задач многоклассовая (мультилейбелная) классификация. Многоклассовая классификация – это задача машинного обучения, в которой модель обучается различать входные данные на несколько классов [4]. Одним из простых и логичных способов добиться автоматической маршрутизации можно было бы, отбирая документы, которые пришли в ответ на отправленный документ из организации. Так как в большинстве случаев, структурное подразделение в котором создавался изначальный документ, будет заниматься разбором документов от контрагента.

Для обработки документов, поступивших в ответ на исходящий документ из организации, используют Rule-based подход [6]. Для распознавания текста в приложенных файлах можно применять технологию оптического распознавания символов (OCR), или же её более продвинутую версию ICR – технологию интеллектуального распознавания символов [8]. После извлечения текстового слоя из приложенного файла будет происходить поиск пар (исходящий номер документа и дата). Это необходимо для последующего их поиска в базе документов в СЭД организации. Что позволит определить структурное подразделение, в котором создавался исходный документ.

Ещё одна задача, с которой поможет справиться ИИ – это во входящем документе найти и извлечь требуемые реквизиты и используя их создать и заполнить регистрационную карточку документа. Входящими данными здесь является файл документа. С точки зрения ML стоит NER задача с поиском объектов в документе (bounding boxes) и распознавание текста OCR [3]. Необходимо извлечь следующие данные – отправитель, получатель, подписант, исходящий номер, исходящая дата, краткое содержание, после чего можно будет заполнить ими регистрационную карточку документа. Однако стоит отметить, что такие задачи гораздо проще решать с документами, имеющими формализованную структуру, т.к. в них реквизиты можно определить не по контексту, а по структуре документа, то есть реквизит подписант обычно располагают справа снизу, реквизит шапка документа – сверху и т.д.

В последнее время, системы ИИ стали применять для автореферирования (суммаризация от англ. «Summary») документов. Целью этой задачи, является – сформированное краткое содержание документа, на основе представленного файла. Обычно для решения этой задачи применяют методы экстрактивной или абстрактивной суммаризации [5]. Метрикой, для оценки таких ML-моделей служит «ROUGE-score». Она показывает, как сильно пересекаются два разных текста [10]. Инструменты для решения такой задачи – это Bert-like модели на базе библиотеки transformers [9]. Автореферирование позволяет сформировать более полное и правильное описание документа по сравнению с пользовательским описанием, что в свою очередь существенно сократит, времени на ознакомление с документами.



В заключение хочется выделить основные аспекты успешного внедрения систем ИИ в процесс ИДО управления в организации.

Во-первых, для качественно использования ML-моделей, их надо обучить и натренировать под особенности ИДО в конкретной организации. А для этого необходим оптимальный объем и качество обучающей выборки. Поэтому в организациях с небольшим объемом документооборота, будет трудно сформировать достаточную обучающую выборку, чтобы ML-модели показывали хороший результат.

Во-вторых, в самой организации должны быть четко прописаны регламенты обработки документов (текущие и будущие).

В-третьих, для используемой ИИ-системы нужно установить порог «уверенности», достигая который система выполняет заданные действия. Например, если система на 85% уверена, что данный документ необходимо отправить на установление резолюции данному руководителю, то она отправляет его, а если менее 85%, то отправляет документ на рассмотрение сотруднику службы ДОУ или делопроизводителю.

И в-четвертых, лучше использовать комбинацию алгоритмической обработки документов, заложенной в самой СЭД и модулю с применением ИИ для разных видов задач. Например, для выстраивания цепочек согласования документа, гораздо проще и оптимальнее будет использовать инструменты, встроенные в систему электронного документооборота, основанные на простых алгоритмических правилах. Так как применение нейронных сетей или ML-моделей потребует привлечения больших ресурсов для обучения, а их результаты все равно могут уступать, простой алгоритмически выстроенной цепочки.

Список литературы:

1. Балашов, А.М. Использование ит-технологий в различных сферах деятельности и формирование новой информационно-цифровой реальности // Теоретическая экономика. 2022. №9 (93). URL: <https://cyberleninka.ru/article/n/ispolzovanie-it-tehnologiy-v-razlichnyh-sferah-deyatelnosti-i-formirovanie-novoy-informatsionno-tsifrovoy-realnosti> (дата обращения: 15.01.2025).

2. Мирошниченко, М. А. Цифровизация и цифровые решения в управлении организацией: Учебник / М. А. Мирошниченко. – Краснодар: Кубанский государственный университет, 2023. – 222 с. – ISBN 978-5-8209-2335-7.

3. СИСТЕМНЫЙ БЛОКЪ: Named Entity Recognition (NER) [Сайт] // – URL: <https://sysblok.ru/glossary/named-entity-recognition-ner/> (дата обращения: 22.01.2025).

4. Хабр: Многоклассовая и многозадачная классификация [Сайт] // – URL: <https://habr.com/ru/companies/otus/articles/766034/> (дата обращения: 21.01.2025).

5. Хабр: Суммаризация текста: подходы, алгоритмы, рекомендации и перспективы [Сайт] // – URL: <https://habr.com/ru/articles/514540/> (дата обращения: 23.01.2025).

6. Хабр: Rule-based neural networks [Сайт] // – URL: <https://habr.com/ru/articles/423393/> (дата обращения: 20.01.2025).

7. Центр2М: Что такое цифровизация и какие сферы жизни она заденет [Сайт] // – URL: <https://center2m.ru/digitalization-technologies?ysclid=liq1obz3em498589452> (дата обращения: 15.01.2025).

8. AWS: Что такое оптическое распознавание символов? [Сайт] // – URL: <https://aws.amazon.com/ru/what-is/ocr/> (дата обращения: 22.01.2025).

9. Medium: A Comparative Study of Different Types of Transformers Using Hugging Face [Сайт] // – URL: <https://medium.com/@rakeshrajpuhith/a-comparative-study-of-different-types-of-transformers-using-hugging-face-5eed98c64ddb> (дата обращения: 23.01.2025).

10. Medium: Understanding BLEU and ROUGE score for NLP evaluation [Сайт] // – URL: <https://medium.com/@sthanikamsanthosh1994/understanding-bleu-and-rouge-score-for-nlp-evaluation-1ab334ecadcb> (дата обращения: 23.01.2025).

