

УДК 025.3

DOI 10.26118/2782-4586.2025.58.63.037

Бойкова Анна Викторовна

Тверской государственный технический университет

Рыльцин Игорь Александрович

Военная академия воздушно-космической обороны

имени Маршала Советского Союза Г.К. Жукова

Подход к технико-экономическому обоснованию автоматизации процесса каталогизации данных, размещенных в электронных библиотеках и информационных системах высшего военного учебного заведения

Аннотация. Статья посвящена технико-экономическому обоснованию автоматизации процесса каталогизации данных в электронных библиотеках высших военных учебных заведений на основе технологий искусственного интеллекта. Цель исследования – оценка эффективности внедрения специализированного ИИ-сервиса «Agregator», предназначенного для автоматического сбора, индексирования и семантического анализа разнородных документов. В работе представлена структура капитальных и эксплуатационных затрат на разработку и внедрение системы, а также расчет ожидаемого экономического эффекта. Показано, что автоматизация позволяет сократить трудозатраты на обработку документов на 70–80%, обеспечивает повышение качества метаданных и оперативности их учета. Расчетный срок окупаемости проекта составляет 4,1 года за счет экономии фонда оплаты труда. Делается вывод о технической реализуемости и экономической целесообразности внедрения предложенного решения в контуре электронной информационно-образовательной среды вуза в соответствии со стратегией цифровой трансформации образования.

Ключевые слова: информация, данные, знания, технико-экономическое обоснование, цифровизация, цифровая зрелость, искусственный интеллект, многомодальная информация, большие языковые модели

Boykova Anna Viktorovna

Tver State Technical University

Ryltsin Igor Alexandrovich

Military Academy of Aerospace Defense named after

Marshal of the Soviet Union G.K. Zhukov

An approach to the feasibility study for automating the process of cataloging data located in electronic libraries and information systems of a higher military educational institution

Abstract. This article presents a feasibility study for automating data cataloging in digital libraries of higher military educational institutions using artificial intelligence technologies. The objective of the study is to evaluate the effectiveness of implementing a specialized AI service, "Agregator," designed for the automatic collection, indexing, and semantic analysis of heterogeneous documents. The paper presents the structure of capital and operating costs for the development and implementation of the system, as well as a calculation of the expected economic impact. It is shown that automation reduces labor costs for document processing by 70–80%, improves the quality of metadata, and improves the efficiency of its accounting. The estimated payback period for the project is 4.1 years due to savings in the payroll fund. A conclusion is drawn regarding the technical feasibility and economic feasibility of implementing the proposed solution within the university's electronic information and educational environment in accordance

with the digital transformation strategy for education.

Keywords: Information, data, knowledge, feasibility study, digitalization, digital maturity, artificial intelligence, multimodal information, large language models

В соответствии с Указом Президента РФ от 07.05.2024 N 309 "О национальных целях развития Российской Федерации на период до 2030 года и на перспективу до 2036 года" поставлена задача к 2030 году достичь «цифровой зрелости» ключевых отраслей экономики и социальной сферы, включая образование, что предусматривает «автоматизацию большей части транзакций в рамках единых отраслевых цифровых платформ и модели управления на основе данных с учетом ускоренного внедрения технологий обработки больших объемов данных, машинного обучения и искусственного интеллекта» [1].

Данная задача ставилась и в предыдущей редакции данного нормативно-правового акта [2]. Для ее достижения Министерством науки и высшего образования была разработана Стратегия цифровой трансформации отрасли науки и высшего образования [3].

В процессе ее разработки «был проведен анализ стратегий цифровой трансформации (цифровизации) в сфере науки и высшего образования стран с наиболее развитыми образовательными системами» [3, 4]. Так, в частности, в ходе анализа было установлено, что многим успешным компаниям, таким как Amazon или Google, удалось добиться организации рационального использования и хранения массивных объемов информации.

И «Google, и Microsoft используют NLP (Natural Language Processing) (обработка естественного языка) – области искусственного интеллекта, посвященной обучению компьютеров понимать нашу письменную речь, чтобы понимать своих пользователей. Исследование Вашингтонского университета рассмотрело Яндекс, четвертую по величине поисковую систему в мире, и отметило, что у нее также есть некоторые передовые приложения NLP и машинного обучения» [5, 6].

Поиск информации «представляет собой важную функциональность, которую предоставляют почти все информационные системы, и в особенности электронные библиотеки (ЭБ), автоматизирующие работу пользователей с электронными ресурсами, такими как документы, изображения, аудио- и видеофайлы» [7, 8].

В «электронных библиотеках поиск информации считается одной из основных задач, решение которой позволяет пользователям находить требуемые им информационные ресурсы. В них предоставляется возможность выполнять полнотекстовый поиск (поиск терминов) для текстовых электронных документов или поиск в метаданных для других типов ресурсов» [7].

Начиная с 1 сентября 2016 года в рамках реализации инновационного проекта «Электронный ВУЗ», все образовательные организации Министерства Обороны РФ перешли на электронные учебники [9].

В соответствии с Приказом Министра обороны РФ от 30 мая 2022 г. № 308 «Об организации образовательной деятельности в федеральных государственных организациях, осуществляющих образовательную деятельность и находящихся в ведении Министерства обороны Российской Федерации» «при реализации основных профессиональных образовательных программ образовательные организации (институты, филиалы) образовательные организации Министерства обороны могут применять электронное обучение и дистанционные образовательные технологии» [10].

В рамках данной статьи, для решения перечисленных задач предлагается использовать сервис на основе искусственного интеллекта, получивший название «Agregator» (здесь и далее ИИ-сервис «Agregator»).

Он предназначен для автоматизации процесса каталогизации большего объема документов и материалов, размещенных в электронных библиотеках и информационных системах высшего учебного заведения (ВУЗа).

Как правило, все многообразие затрат на разработку и внедрение ИИ-сервис

«Agregator» можно разделить на:

единовременные (капитальные) (Z_{kan}) – осуществляются, как правило, единожды и представляют собой вложения в основные средства, их модернизацию, реконструкцию.

эксплуатационные (операционные) ($Z_{экс}$) – производится в течение всего срока полезного использования оборудования для поддержания его в работоспособном состоянии.

Тогда совокупные затраты на разработку и внедрение ИИ-сервис «Agregator» ($Z_{общ}$) определим на основе соотношения:

$$Z_{общ} = Z_{кан} + Z_{экс}, \quad (1)$$

Капитальные затраты в данном случае будут складываться из затрат на закупку необходимого оборудования (Z_{oc}) и программного обеспечения (ПО) (Z_{PO}), а также на обучение персонала работе с ней ($Z_{об}$):

$$Z_{кан} = Z_{oc} + Z_{PO} + Z_{об}, \quad (2)$$

В отличие от капитальных затрат, эксплуатационные издержки повторяются с определённой периодичностью, поэтому они измеряются в денежных единицах за единицу времени (рублей в год, рублей в месяц).

Пилотное внедрение предполагает установку и апробацию системы ИИ-«Агрегатор» в высшем военном учебном заведении. Для этого потребуется сервер или высокопроизводительный персональный компьютер. Рекомендуемая конфигурация – CPU минимум с 8 ядрами, 32 ГБ ОЗУ, SSD 1-2 ТБ, дополнительно GPU среднего уровня (для ускорения NLP/ASR). Оценочно стоимость такого оборудования составляет 150 тысяч рублей.

Работы по внедрению включают установку ПО на сервер, перенос разработанной системы и подключение к локальной сети. Также необходимо реализовать интеграцию ИИ-сервиса «Agregator» с электронной информационно-образовательной средой (ЭИОС) ВУЗа, например, выполнить настройку единой авторизации (так сотрудники и пользователи библиотеки смогут входить в модуль «Агрегатор» через свои учётные записи ЭИОС), настройку обмена данными (при необходимости, экспорт метаданных из ИИ-сервиса «Агрегатор» в основную базу данных библиотеки или наоборот импорт существующих записей). Эти работы оцениваются в 30 тысяч рублей.

Для успешного пользования предлагаемой системой провести обучающий семинар-тренинг для сотрудников библиотеки учебного заведения, в рамках которого будет продемонстрирован интерфейс, порядок действий при добавлении новых материалов, способы поиска и формирования отчётов. Затраты на подготовку методических материалов и проведение обучения оцениваются в 20 тысяч рублей.

В первые 2-3 месяца эксплуатации ИИ-сервиса «Агрегатор» для оперативного устранения выявленных проблем и ответов на вопросы пользователей необходима поддержка разработчика. Затраты на техническое сопровождение в указанный период оцениваются в 50 тысяч рублей.

Далее определим затраты на разработку ИИ-сервиса «Агрегатор». Они будут складываться из оплаты труда программиста ($Z_{п}$), затрат на под конкретную ЭИОС и локальные требования (Z_M), нормативно-справочную документацию (Z_D):

$$Z_{PO} = Z_{п} + Z_D + Z_M, \quad (2)$$

Прежде всего, необходимо определить трудоемкость выполнения работ, которая приведена в таблице 1.

Таблица 1 – Трудоемкость разработки программного обеспечения

| Этап разработки | Трудоёмкость, чел.-мес. |
|---------------------------|-------------------------|
| 1. Проектирование системы | 1,0 |

| Этап разработки | Трудоёмкость, чел.-мес. |
|----------------------------------|-------------------------|
| 2. Реализация (программирование) | 3,0 |
| 3. Тестирование и доработка | 1,5 |
| 4. Внедрение (пилот) | 1,5 |
| Итого | 7,0 |

Таким образом, на разработку ИИ-сервиса «Агрегатор» потребуется 7 человеко-месяцев. Фонд оплаты труда программиста определим на основе соотношения:

$$Z_{\Pi} = (T \cdot \mathcal{C}) \cdot (1 + tax), \quad (3)$$

где

T – трудоемкость разработки программного обеспечения, человеко-месяцев;

\mathcal{C} – часовая ставка разработчика программного обеспечения, рублей/человеко-месяц;

tax – страховые взносы.

Тогда совокупные затраты на разработку ИИ-сервиса «Агрегатор» составят:

$$Z_{\Pi} = (7\mathcal{C} - m * 100000 \text{руб}/\text{ч} - m) \cdot (1 + 0,3) = 910000 \text{руб},$$

Подготовка полного комплекта документации: техническое описание системы, руководство администратора (настройка, развёртывание) и руководство пользователя (для библиотекарей) оценивается оцениваются в 0,5 человека-месяца работы или в 50000 рублей.

Затраты на адаптацию ИИ-сервиса «Агрегатор» под конкретную ЭИОС и локальные требования, а именно доработка интеграции (например, модуль авторизации, соответствие ГОСТ по отображению метаданных), настройка интерфейса под бренд организации оцениваются в 0,8 человека-месяца работы или 80000 рублей.

Тогда затраты на разработку программного обеспечения составят:

$$Z_{PO} = 910000 \text{руб} + 50000 \text{руб} + 80000 \text{руб} = 1040000 \text{руб}$$

В этом случае, капитальные затраты на внедрение ИИ-сервиса «Агрегатор» составят:

$$Z_{kan} = 1040000 \text{руб} + 150000 \text{руб} + 30000 \text{руб} + 20000 \text{руб} + 50000 \text{руб} = 1290000 \text{руб}$$

Автоматизация процесса каталогизации прямо сокращает трудоёмкость работы сотрудников библиотеки образовательной организации по вводу данных. Ранее каждый новый поступающий документ предполагал чтение аннотации или самого текста, извлечение информации (название, авторы, классификационные индексы, ключевые слова) и ручное занесение в электронный каталог. Например, обработка одной диссертации могла занимать 1-2 часа квалифицированного труда, статьи – порядка 30 минут.

ИИ-сервис «Агрегатор» позволяет сократить это время в несколько раз. Практически, роль сотрудника сведётся к контролю и валидации автоматически заполненной карточки: достаточно просмотреть сгенерированные системой данные и при необходимости откорректировать неточности.

Опытная эксплуатация системы показала, что на проверку одной записи уходит порядка 10-15 минут. Иными словами, экономия времени может достигать 70-80% на источник.

Если предположить, что в среднем, в год библиотека ВУЗа обрабатывает около 1000 новых документов, то в пересчете на год экономия рабочего времени персонала составит 1000 часов. Это эквивалентно высвобождению 0,5–1 ставки сотрудника, что может быть направлено на другие задачи (консультирование читателей, пополнение фонда и т.д.) либо отражено как прямая экономия фонда оплаты труда.

Автоматизация процесса каталогизации данных обеспечивает сплошной учет документов без задержек и потерь. ИИ-сервис «Агрегатор» на постоянной основе сканирует указанные директории и оперативно регистрирует появление новых файлов (например, новых выпусков сборников или загруженных электронной системой студенческих работ). Иными словами, электронный каталог всегда соответствует фактическому состоянию

фонда.

Повышение оперативности учёта особенно важно для новых публикаций сотрудников университета – они сразу становятся видимыми и доступными для поиска, что повышает индекс цитирования вуза и прозрачность научной деятельности.

Кроме того, по вине человеческого фактора формуляры по учету библиотечных фондов часто содержат опечатки и ошибки, неточности. ИИ-сервис «Агрегатор» работает по единым алгоритмам, исключающим случайные ошибки ввода. Например, фамилия, имя, отчество автора будут извлечены точно как в оригинале (исключая вариативность написания). Встроенные справочники (например, соответствие русских и английских названий типов документов) нормализуют данные до единого формата. Это обеспечивает высокую степень консистентности базы данных, что впоследствии облегчает поиск и аналитику. Кроме того, централизованная обработка документов гарантирует единообразие: система для каждого документа пытается найти все ключевые показатели (автора, год, тип, и т.д.), что уменьшает процент пропущенной информации.

Благодаря тому, что ИИ-сервис «Агрегатор» не только сохраняет данные, но и анализирует связи между ними (например, авторы – организации – публикации), сотрудники библиотеки могут проводить глубокую аналитику в короткие сроки. Это позволит быстро ответить на вопросы, которые ранее потребовали бы длительной ручной сводки: например, сколько диссертаций защищено по определенной специальности за последние 5 лет.

Интерактивный граф связей позволит научному отделу ВУЗа выявить насколько активно сотрудники ВУЗа взаимодействуют друг с другом и с сотрудниками других ВУЗов (то есть имеют совместные публикации), какие авторы являются центральными фигурами (имеют наибольшее число работ, связанных с разными соавторами). Полученные данные можно использовать при составлении отчетов для Министерства обороны РФ, при планировании поддержки ведущих научных школ, для популяризации достижений университета. Фактически, система превращает библиотеку из просто хранилища книг в активный инструмент аналитики, подкрепляя решения руководства точными данными.

Качественно каталогизированный и быстро пополняемый фонд – это прямая выгода для конечных пользователей: ученых, студентов, сотрудников. Поиск нужной информации становится проще и точнее, т.к. имеются разнообразные точки доступа (по ключевым словам, организациям, полнотекстовый поиск по фрагментам и т.д.). Пользователи смогут обнаруживать работы коллег по интересующей тематике, видеть взаимосвязанные материалы.

Например, при исследовании научной проблемы студент легко найдёт связанные с ней диссертации и статьи, где одним из авторов, например, выступает его научный руководитель, или все работы, выпущенные в определенном институте.

Это стимулирует использование электронного каталога и повышает авторитет библиотеки. Внедрение ИИ-сервис «Агрегатор» также положительно скажется на имидже ВУЗа: демонстрируется внедрение цифровых инноваций в научную инфраструктуру, что может стать дополнительным аргументом при аккредитациях и оценке эффективности деятельности (например, по показателю цифровизации).

Экономический эффект от внедрения ИИ-сервиса «Агрегатор» будет заключаться в экономии рабочего времени сотрудников библиотеки. Как было отмечено выше, можно ожидать высвобождения сотен часов ежегодно. Среди косвенных эффектов отметим: снижение необходимости перепроверок и исправлений (экономия времени заведующих отделом на контроль качества); отсутствие затрат на приобретение коммерческого ПО для решения указанных задач (это обходится ВУзу в сотни тысяч рублей ежегодно); потенциальный рост цитируемости и показателей публикационной активности учёных.

Далее определим срок окупаемости предлагаемого ИИ-сервиса «Агрегатор». Как показали расчеты, совокупные затраты на реализацию системы составляют 1230 тысяч рублей.

Как отмечалось ранее, внедрение программного обеспечения позволит сократить затраты рабочего времени сотрудников ВУЗа. Рассмотрим консервативный сценарий: библиотека ежегодно принимает на учёт 1000 новых документов (статьи преподавателей, выпускные квалификационные работы, диссертации, отчёты и т.п.). Если ранее на обработку одного документа в среднем затрачивался 1 час, то после внедрения ИИ-сервиса «Агрегатор» это время сокращается до 15 минут. Экономия времени составляет 45 минут на один или 750 часов в год на 1000 документов. При среднечасовой ставке сотрудника библиотеки 400 рублей, это позволит сократить фонд заработной платы на 300 тыс. рублей в год. В этом случае срок окупаемость (без дисконтирования) составит:

$$T_{ок} = \frac{1230000 \text{ руб}}{300000 \text{ руб / год}} = 4,1 \text{ года}$$

Таким образом, по итогам технико-экономического анализа можно сделать вывод, что ИИ-сервис «Агрегатор» является технически реализуемым и экономически целесообразным для научной библиотеки вуза. Система достигла высокой степени готовности: разработан и опробован прототип, демонстрирующий заявленные функции автоматической каталогизации и анализа. Инвестиции в доработку и внедрение данной системы оправданы ожидаемыми выгодами – значительным снижением ручного труда, улучшением качества учётных данных и расширением аналитических возможностей.

Расчёты показывают, что проект способен окупиться в среднесрочной перспективе за счёт одной только оптимизации рабочих процессов. Дополнительные же эффекты (оперативность, полнота, новые сервисы для пользователей) усиливают ценность проекта, выводя библиотечно-информационную деятельность на новый уровень.

Рекомендуется внедрить ИИ-сервис «Агрегатор» в контуре ЭИОС образовательной организации Министерства обороны РФ. Предлагаемое решение соответствует целям и задача стратегии цифровизации образования и науки, демонстрируя, как современные технологии искусственного интеллекта могут эффективно применяться в управлении данными.

Список источников

1. Указ Президента РФ от 07.05.2024 N 309 "О национальных целях развития Российской Федерации на период до 2030 года и на перспективу до 2036 года" [Электронный ресурс]. – Режим доступа – https://www.consultant.ru/document/cons_doc_LAW_475991
2. Указ Президента Российской Федерации от 21.07.2020 N 474 "О национальных целях развития Российской Федерации на период до 2030 года" [Электронный ресурс]. – Режим доступа – https://www.consultant.ru/document/cons_doc_LAW_380244/95306fd63c6d78809f4ae93a22b776b264fc0d7
3. Стратегия цифровой трансформации отрасли науки и высшего образования (утв. Минобрнауки России 14 июля 2021 года) [Электронный ресурс]. – Режим доступа – <https://minobrnauki.gov.ru/upload/iblock/e16/dvbedzmr0og5dm57dtm0wyllrbuwtujw.pdf>
4. Тузовский А.Ф. Формирование семантических метаданных для объектов системы управления знаниями // Изв. Томск. политехн. ун-та. – 2007. – Т. 310, № 3. – С. 108-112
5. Лукша Б.Н., Лаптёнок Н.В. Искусственный интеллект в поисковых системах: обзор современного состояния технологий // Информационные системы и технологии: сб. ст. 58-й науч. конф. (Минск, 18-22 апр. 2022). – Минск: БГУИР, 2022. – С. 16-18
6. Калёнова О.В., Перл И.А. Сравнение подходов к обработке гетерогенных данных // Альманах науч. раб. молодых ученых Университета ИТМО. – 2019. – Т. 1. – С. 95-102
7. Ле Хоай, Тузовский А. Ф. Поиск в семантических электронных библиотеках // Доклады ТУСУР. – 2013. – №1 (27). – URL: <https://cyberleninka.ru/article/n/poisk-v-semaniticheskikh-elektronnyh-bibliotekah> (дата обращения: 10.09.2025)
8. Нгуен Б.Н. Модели и методы поиска информационных ресурсов с использованием

семантических технологий: дис. канд. техн. наук. – Томск, 2012. – 198 с

9. Шойгу: все вузы Минобороны РФ перешли на электронные учебники [Электронный ресурс]. – Режим доступа – <https://tass.ru/armiya-i-opk/3586712>

10. Приказ Министра обороны РФ от 30 мая 2022 г. № 308 “Об организации образовательной деятельности в федеральных государственных организациях, осуществляющих образовательную деятельность и находящихся в ведении Министерства обороны Российской Федерации” [Электронный ресурс]. – Режим доступа – <https://www.garant.ru/products/ipo/prime/doc/404833517/#0>

Сведения об авторе

Бойкова Анна Викторовна, профессор, Тверской государственный технический университет, Тверь, Россия.

Рыльцин Игорь Александрович, доцент, Военная академия воздушно-космической обороны имени Маршала Советского Союза Г.К. Жукова, Тверь, Россия

Information about the author

Boykova Anna Viktorovna, professor, Tver State Technical University, Tver, Russia.

Ryltsin Igor Alexandrovich, senior teacher, Military Academy of Aerospace Defense named after Marshal of the Soviet Union G.K. Zhukov, Tver, Russia