

## Применение искусственного интеллекта в практике научно-технических библиотек: исследование потенциала, опыт использования, оценка перспектив

М. В. Гончаров<sup>1</sup>, К. Е. Соколинский<sup>2</sup>, Я. Л. Шрайберг<sup>3</sup>

<sup>1,3</sup>ГПНТБ России, Москва, Российская Федерация

<sup>2</sup>Международная ассоциация электронных библиотек и новых информационных  
технологий (Ассоциация ЭБНИТ), Санкт-Петербург, Российская Федерация

<sup>1</sup>[goncharov@gpntb.ru](mailto:goncharov@gpntb.ru)

<sup>2</sup>[info@irbis-service.com](mailto:info@irbis-service.com)

<sup>3</sup>[shra@gpntb.ru](mailto:shra@gpntb.ru), <https://orcid.org/0000-0001-6110-3271>

**Аннотация.** Статья представляет результаты комплексного исследования применения больших языковых моделей (LLM) и технологий искусственного интеллекта в отечественной библиотечно-информационной деятельности. Проведён сравнительный анализ эффективности российских (GigaChat, YaGPT) и зарубежных (ChatGPT, Claude, LLaMA, Mistral, DeepSeek) языковых моделей для решения трёх ключевых библиотечных задач: семантической деконструкции пользовательских запросов, распознавания рукописных каталожных карточек и автоматической коррекции ошибок в текстовых данных.

Методология исследования включала систематическое тестирование через программные интерфейсы API на репрезентативных выборках запросов, образов рукописных карточек и полнотекстовых документов. Выявлены принципиально значимые ограничения существующих решений: нестабильность и деградация моделей, избыточная цензура с высоким процентом ложных срабатываний, генерация несуществующих данных (гальлюцинации), непредсказуемость структурированного вывода, культурно-лингвистические барьеры.

На основе полученных результатов разработана специализированная система ИРБИС AI с архитектурой Mixture of Experts, обеспечивающая стабильную и быструю обработку библиографических данных. Представлена практическая реализация в модуле J-ИРБИС 2.0 с тремя режимами ИИ-поддержки: справочное обслуживание, семантическая обработка запросов и интеллектуальный подбор литературы. Обозначены перспективы развития: создание библиотечного портала с полностью ИИ-управляемым интерфейсом, автоматиза-

ция каталогизации через мультимодальные модели, комплексная обработка неструктурированных данных и преобразование имидж-каталогов.

**Ключевые слова:** искусственный интеллект, большие языковые модели, семантическая деконструкция, каталогизация, имидж-каталоги, справочное обслуживание

**Для цитирования:** Гончаров М. В., Соколинский К. Е., Шрайберг Я. Л. Применение искусственного интеллекта в практике научно-технических библиотек: исследование потенциала, опыт использования, оценка перспектив // Научные и технические библиотеки. 2025. № 12. С. 144–164. <https://doi.org/10.33186/1027-3689-2025-12-144-164>

UDC 004.8:02 + 004.93

<https://doi.org/10.33186/1027-3689-2025-12-144-164>

## **Artificial intelligence in the practice of sci-tech libraries: The study of potential, experience and prospects evaluation**

**Mikhail V. Goncharov<sup>1</sup>, Kirill E. Sokolinsky<sup>2</sup> and Yakov L. Shrayberg<sup>3</sup>**

<sup>1, 3</sup>*Russian National Public Library for Science and Technology,  
Moscow, Russian Federation*

<sup>2</sup>*International Association of Users and Developers of Electronic Libraries and New  
Information Technologies (ELNIT Association), St. Petersburg, Russian Federation*

<sup>1</sup>*goncharov@gpntb.ru*

<sup>2</sup>*info@irbis-service.com*

<sup>3</sup>*shra@gpntb.ru, https://orcid.org/0000-0001-6110-3271*

**Abstract.** The paper comprises the findings of the integrated study of using large language models (LLMs) artificial intelligence (AI) technologies in the national library information practice. The comparative analysis is accomplished for Russian (GigaChat, YaGPT) and foreign (ChatGPT, Claude, LLaMA, Mistral, DeepSeek) LLM applications for three key library tasks, i. e. semantic deconstruction of user queries, recognition of handwritten catalog cards, and automated correction of text errors.

The system testing on representative query selection, handwritten cards and fulltext document images via APIs was accomplished. The critical limitations of existing solutions are revealed: instability and degradation of models, excessive censorship with high percentage of false triggering, nonexistent data generation (hallucinations), and unpredictability of structured inference, or cultural and linguistic barriers.

Based on the study findings, the specialized IRBIS AI system based on the Mixture of Experts architecture was designed to ensure prompt and stable processing of bibliographic data. The authors discuss the practical implementation in J-IRBIS 2.0 module with three modes of AI-support: reference services, query semantic processing, and intellectual literature selection. The prospects for building the library portal with AI-controlled interface, cataloguing automation through multimodal models, integrated processing of unstructured data, and image catalogs transformation, are outlined.

**Keywords:** artificial intelligence, large language models, LLM, semantic deconstruction, cataloguing, image catalogs, reference service

**Cite:** Goncharov M. V., Sokolinsky K. E., Shrayberg Y. L. Artificial intelligence in the practice of sci-tech libraries: The study of potential, experience and prospects evaluation // Scientific and technical libraries. 2025. No. 12, pp. 144–164. <https://doi.org/10.33186/1027-3689-2025-12-144-164>

## 1. Введение

Современная библиотечно-информационная деятельность переживает период кардинальных трансформаций, обусловленных стремительным развитием технологий искусственного интеллекта. Революционные изменения в области больших языковых моделей (Large Language Models, LLM) и сопутствующих технологий открывают принципиально новые возможности для автоматизации традиционно трудо затратных библиотечных процессов, требующих высокой квалификации специалистов.

За последние годы произошёл качественный скачок в развитии систем искусственного интеллекта: от экспериментальных разработок к практически применимым решениям, способным обрабатывать сложные неструктурированные данные с точностью, сопоставимой с человеческой деятельностью. Особенно значимые изменения затронули области семантического анализа текста, оптического распознавания текстов и автоматической коррекции ошибок – ключевые направления, имеющие первостепенное значение для библиотечной практики.

В российском библиотечном сообществе наблюдается устойчивый интерес к потенциалу технологий искусственного интеллекта, однако их практическое внедрение сдерживается отсутствием систематических исследований эффективности различных подходов применительно к специфическим задачам библиотечно-информационной деятельности. Большинство существующих работ ограничивается анализом пользовательского опыта применения онлайновых чатов (причём далеко не лучших моделей) и генерацией изображений. Это не позволяет делать адекватные заключения о потенциале искусственного интеллекта в библиотеках.

Актуальность исследования определяется необходимостью преодоления разрыва между стремительно развивающимися возможностями искусственного интеллекта и их практическим применением в отечественной библиотечной практике. Особую значимость приобретает разработка методологических подходов к оценке эффективности языковых моделей для решения узкоспециализированных библиотечных задач, а также создание программного инструментария, адаптированного к специфике российского библиотечно-информационного пространства.

## **2. Цель и задачи исследования**

Целью настоящего исследования является комплексная аналитическая оценка наиболее перспективных направлений применения больших языковых моделей и родственных технологий искусственного интеллекта в библиотечной практике и задачах информационного обслуживания.

Для достижения поставленной цели определены следующие исследовательские задачи:

1. Провести сравнительный анализ эффективности современных языковых моделей в контексте решения специфических библиотечных задач.
2. Выявить и систематизировать основные ограничения существующих решений на основе искусственного интеллекта применительно к библиотечно-информационной деятельности.
3. Разработать и апробировать методологию оценки качества работы языковых моделей для задач семантической деконструкции библиографических данных.
4. Определить оптимальные параметры настройки и способы интеграции технологий искусственного интеллекта в существующие библиотечные системы.

Исследование фокусируется на трёх приоритетных направлениях использования искусственного интеллекта, имеющих наибольшее значение для библиотечно-информационной практики: семантическом анализе текста, распознавании текста (преобразовании графических образов в текстовый формат) и автоматической коррекции ошибок в текстовых данных.

Необходимо отметить, что упорядочивание неструктурированных данных составляет сущность большинства библиотечных процессов, включая библиографирование, каталогизацию, комплектование и информационный поиск. С технической точки зрения задачи извлечения метаданных из пользовательских запросов и выделения элементов из библиографических описаний представляют собой близкие по своей природе процессы, что позволяет рассматривать их в рамках единого методологического подхода.

### **3. Методология исследования**

#### ***3.1. Обзор исследуемых языковых моделей и обоснование их выбора***

Исследование проводилось на основе анализа текстовых языковых моделей общего назначения, отобранных по критериям доступности, производительности и потенциальной применимости в российском библиотечном контексте.

В первую очередь анализировались модели, условно классифицируемые как российские разработки (GigaChat, YaGPT), хотя более кор-

ректно их следует характеризовать как продукты российских технологических компаний, базирующиеся на зарубежных архитектурах. Так, GigaChat предположительно основан на архитектуре GPT-3 [5], а YaGPT 5 предположительно использует QWEN 2.5 [2] в качестве базовой модели. Российские модели сегодня безальтернативны для официального прямого (без организаций-посредников) использования в государственных организациях.

Во вторую очередь исследовались дистиллированные и квантованные языковые модели с открытым исходным кодом (LLaMA, DeepSeek R1, Mistral), а также проприетарные западные решения, демонстрирующие наилучшие показатели по стандартным метрикам (ChatGPT, Grok, Claude). Тестирование проводилось с использованием моделей различных версий, отличающихся размерностью и уровнем квантования, с целью определения оптимального соотношения качества результатов и вычислительных затрат. Основная часть исследований проводилась в 2024–2025 гг. с актуальными на этот момент исследованиями моделями.

### *3.2. Критерии оценки эффективности*

Эмпирическая база исследования сформирована на основе результатов систематического тестирования нейросетевых моделей, а также данных, полученных в ходе опытной эксплуатации модулей на основе LLM в составе системы Я-ИРБИС 2.0. Сравнительная оценка проводилась по следующим прикладным критериям.

Для задач семантической деконструкции:

точность следования инструкциям (измеряемая как процент корректно выполненных операций);

скорость обработки запросов (время от подачи запроса до получения результата);

воспроизводимость результатов (степень идентичности ответов при повторных запросах);

способность к исправлению ошибок (эффективность коррекции опечаток и грамматических неточностей).

Для задач распознавания текста:

качество распознавания (процент корректно распознанных символов и слов);

экономическая эффективность (стоимость обработки страницы текста).

Скорость обработки запросов имеет значение при оказании информационных услуг для читателей в реальном времени, но играет небольшую роль при распознавании образов.

Экономическая эффективность приобретает особую значимость при работе с крупными массивами графических и полнотекстовых данных, такими как имидж-каталоги, содержащие сотни тысяч карточек, где совокупные затраты на распознавание могут достигать нескольких сотен тысяч рублей. При этом стоимость семантической де-конструкции записи или запроса редко превышает 10 копеек.

### *3.3. Методы и технологии тестирования*

Исследование моделей осуществлялось через программные интерфейсы (API), совместимые со стандартом OpenAI. Данный подход обеспечивает существенно более высокую точность, воспроизводимость и контролируемость результатов по сравнению с диалоговыми интерфейсами, традиционно используемыми конечными пользователями.

Использование API позволяет точно настраивать параметры генерации ответов. Для повышения качества следования инструкциям применялся системный промпт, придающий инструкциям более высокий приоритет с точки зрения языковой модели. В рамках исследования были разработаны три различных провайдера, поддерживающих основные вариации API.

Тестирование проводилось с единым набором параметров, оптимизированных для достижения максимальной точности и воспроизводимости:

temperature = 0,1 (минимизация вариативности ответов);

top-k sampling = 1 (обеспечение детерминированности результатов);

repetition penalty = 1,2 (снижение вероятности повторений в генерируемом тексте).

Для отдельных моделей дополнительно настраивались параметры top-p и min-p в соответствии с их спецификациями.

### ***3.4. Формирование тестовых выборок***

Для оценки способности к семантической деконструкции была сформирована выборка из 50 наиболее сложных для обработки запросов, зафиксированных в системах СПбГУ, СПбГУТ и НМИЦ онкологии им. Н. Н. Петрова.

При тестировании потенциала распознавания рукописного текста из каталога ВМедА им. С. М. Кирова были отобраны 20 образов рукописных каталожных карточек различной степени качества.

В целях оценки возможностей парсинга объёмных полнотекстовых документов использовался корпус из 10 электронных иллюстрированных журналов и статей в формате PDF.

Промпты подвергались итеративной оптимизации для наиболее перспективных моделей с целью достижения максимальной эффективности на тестовом датасете. Для автоматизации процесса сравнения результатов была разработана программа LLM Arena, позволяющая получать и сопоставлять результаты выполнения запросов различными моделями.

## **4. Результаты исследования**

### ***4.1. Семантическая деконструкция и нормализация текстовых данных***

Исследование велось в трёх взаимосвязанных направлениях: семантическая деконструкция, нормализация текстовых данных, а также анализ мультимодальной (в данном случае визуально-языковой, позволяющей трансформировать графические образы в текст) работы для извлечения метаданных из изображений.

Задача семантической деконструкции решалась на примере свободно сформулированных пользовательских запросов. Языковым моделям предлагалось выполнить комплекс операций: извлечь фамилию автора, выделить инициалы, определить заглавие произведения, год издания, устраниТЬ опечатки, нормализовать падежные формы, предложить стандартизованный вариант и определить наиболее вероятные инициалы автора на основе контекста.

Все указанные задачи ставились в рамках единого промпта, поскольку требование к скорости обработки в реальном времени не поз-

воляет использовать итеративный подход. Целевым показателем было установлено время обработки не более 10 секунд, определённое как максимально приемлемое для современного пользователя библиотечных систем.

#### **4.2. Анализ эффективности текстовых языковых моделей**

Семантическая деконструкция запросов была успешно выполнена для 86% тестовой выборки. Однако анализ оставшихся 14% случаев выявил ряд существенных проблем. В рамках данной статьи невозможно подробно изложить все детали, поэтому ограничимся здесь лишь перечислением проблем и выразительными примерами их проявления:

**Нестабильность и деградация моделей.** Некогда лидирующая по метрикам отечественная модель GigaChat после 6 декабря 2024 г. продемонстрировала признаки деградации качества и нестабильность результатов. В частности, система начала ошибочно идентифицировать словосочетание «искусственный интеллект» как имя индивидуального автора, включать в фамилию автора заглавие публикации.

**Избыточная цензура.** Встроенная система фильтрации запрещённого контента в GigaChat срабатывала приблизительно на 30% запросов социально-экономической тематики. Например, на запрос «история русской философии Бердяев» система отказывалась предоставлять информацию.

**Галлюцинации.** Тенденция к генерации правдоподобных, но фактически неверных данных наблюдалась даже у моделей-лидеров. Так, модели добавляли несуществующие инициалы даже в случаях, когда генерация инициалов была явно запрещена в инструкциях. Например, это относилось к инициалам, связанным с известными писателями, общественными деятелями. Зачастую модели генерировали инициалы известных персон для их однофамильцев.

**Непредсказуемость результата.** Модели демонстрировали непоследовательность в соблюдении заданной структуры вывода, произвольно изменяя разделители и форматирование. Сами результаты обработки одного и того же промпта с течением времени также могли меняться.

**Произвольная адаптация к структуре.** Использование структурированного вывода в формате JSON парадоксальным образом повышало вероятность галлюцинаций, поскольку модели стремились заполнить все поля схемы даже при отсутствии соответствующих данных.

**Ограничения дистиллированных и квантизированных моделей.** Модели с открытым исходным кодом и высокой степенью квантизации (например, Mistral 12b4k) оказались в принципе неспособны к эффективному решению задач семантической деконструкции.

**Высокие временные затраты.** Корректная обработка сложных запросов требовала значительных временных ресурсов. Например, правильная обработка запроса «ильф и петров золотой теленок» с генерацией корректных инициалов первого автора осуществлялась моделью ChatGPT 5 только в режиме *high* (высоких ресурсозатрат) за 20 секунд.

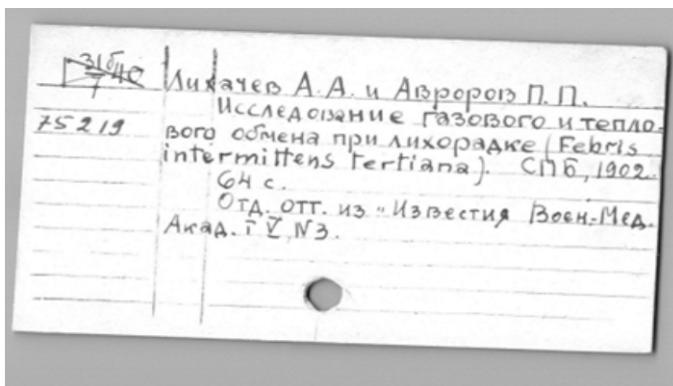
**Культурно-лингвистические барьеры.** Большинство моделей некорректно интерпретировало понятия, специфичные для российской культуры и библиографической практики, такие как «инициалы» и «отчество».

**Влияние опечаток на качество обработки.** Несмотря на то, что устранение опечаток успешно выполняется всеми рассмотренными моделями, их наличие существенно влияет на качество обработки. Это объясняется отсутствием специализированного предобработчика, выполняющего проверку орфографии.

В рамках настоящей статьи невозможно дать пояснения относительно причин указанных проблем (это тема самостоятельных исследований [3]), однако можно обозначить одну: LLM, как правило, не обучались на библиографических данных и не контролировались при решении задач с ними.

#### 4.3. Результаты тестирования мультимодальных моделей

В области распознавания рукописных карточек были получены неожиданно позитивные результаты. Удалось выявить и настроить мультимодальную модель (VLM), способную успешно распознать все карточки из тестовой выборки с последующим корректным выделением основных элементов библиографического описания. Количество ошибок при этом не превышало 0,5%. Эта модель включена в состав системы ИРБИС AI.



**Первый автор:** Лихачев А.А.

**Последующие авторы:** Авроров П.П.

**Заглавие:** Исследование газового и теплового обмена при лихорадке (Febris intermittens tertiana)

**Год:** 1902

**ISBN:** Не указан

**Количество страниц:** 64

**Издательство:** Не указано (упоминается только источник публикации — "Известия Вoen.-Мед. Акад. ")

**Город издания:** Санкт-Петербург (СПб.)

**Рис. 1. Каталожная карточка и результат её распознавания**

Анализ полнотекстовых журналов в формате PDF также продемонстрировал высокую эффективность: успешно извлекался полный набор метаданных, необходимых для библиографического описания. Таким образом, если технологии оптического распознавания текста начала 2000-х гг. были неспособны в принципе интерпретировать ру-

кописный текст, современные системы компьютерного зрения демонстрируют качество, сопоставимое с человеческим восприятием.

## 5. Практическая реализация результатов исследования

### 5.1. Разработка системы ИРБИС AI

Выявленные ограничения существующих LLM обусловили необходимость создания специализированной системы искусственного интеллекта, обеспечивающей требуемый уровень надёжности, предсказуемости и производительности для решения узкоспециализированных библиотечных задач.

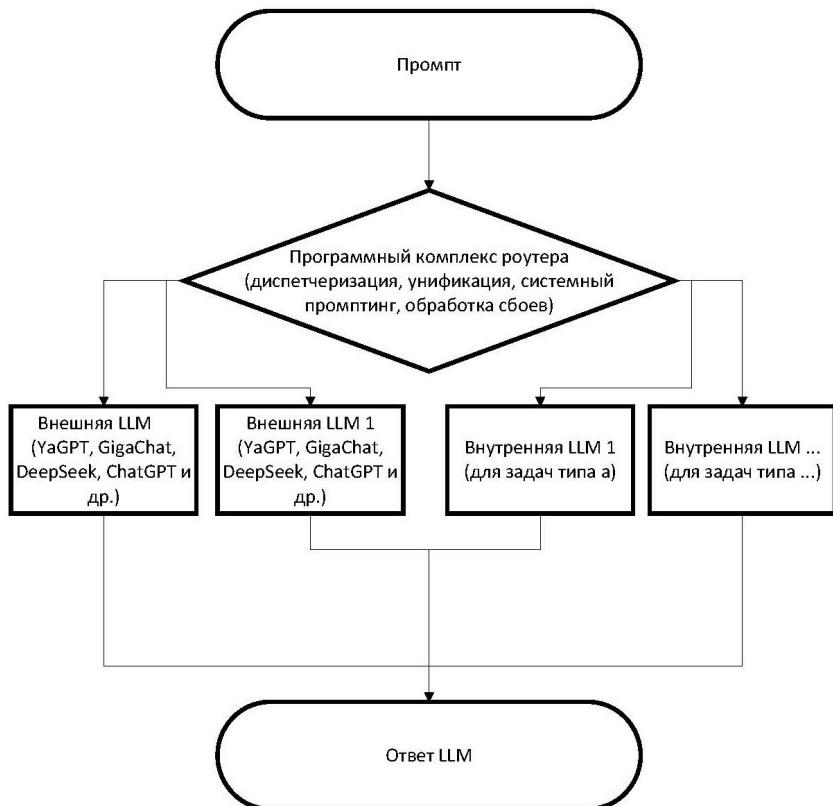


Рис. 2. Упрощённая схема ИРБИС AI

На вычислительных мощностях Объединения «ИРБИС-сервис» и Ассоциации ЭБНИТ развернута система ИРБИС AI, построенная в соответствии с принципами архитектуры Mixture of Experts. Основу архитектуры составляют маршрутизатор запросов и набор локально функционирующих языковых моделей, специально отобранных и адаптированных для библиотечных задач. Такая архитектура является стандартом лидирующих LLM [1], в ИРБИС AI она реализуется на метауровне (в системе используются как плотные, так и разреженные модели). Особенности реализации роутера могут меняться для наилучшего соответствия новым эксплуатационным задачам.

При необходимости система может дополняться облачными моделями, включая модели, заблокированные для прямого использования российскими пользователями.

Скорость работы системы на задачах деконструкции превосходит GPT-5 в режиме low (минимального использования ресурсов). При эквивалентном качестве она примерно на 70% выше. Типовой запрос на семантическую деконструкцию требует примерно 1,5 секунды, что является едва ли не лучшим показателем для LLM.

Система обеспечивает решение задач семантической деконструкции с приемлемым качеством, минимальным уровнем галлюцинаций и полной предсказуемостью результатов, что критически важно для стабильной работы библиотечных систем.

В отличие от облачных LLM, ИРБИС AI может функционировать автономно, без подключения к интернету, что обеспечивает информационную безопасность для пользователей, работающих в закрытом контуре. Кроме того, ИРБИС AI может быть специально обучена на данных организаций-пользователей САБ ИРБИС. Такое обучение позволяет системе обеспечить рецепцию библиотечных традиций, заполнить пробелы в знаниях и усвоить специальную терминологию. Например, в настоящее время выполнено её обучение на предметных рубриках и ключевых словах библиотеки Санкт-Петербургского горного университета.

Для системы разработан веб-интерфейс с поддержкой потокового вывода (вывод текста по мере его генерации), который планируется использовать для упрощения работы с документацией САБ ИРБИС 64 и тестирования LLM.

## 5.2. Интеграция искусственного интеллекта в J-ИРБИС 2.0

Для модуля J-ИРБИС 2.0 разработаны и внедрены три режима работы с использованием искусственного интеллекта, обеспечивающие поддержку пользователя на всех этапах информационного поиска и реализацию семантического поиска. Режимы могут работать как с ИРБИС AI, так и с облачными LLM.

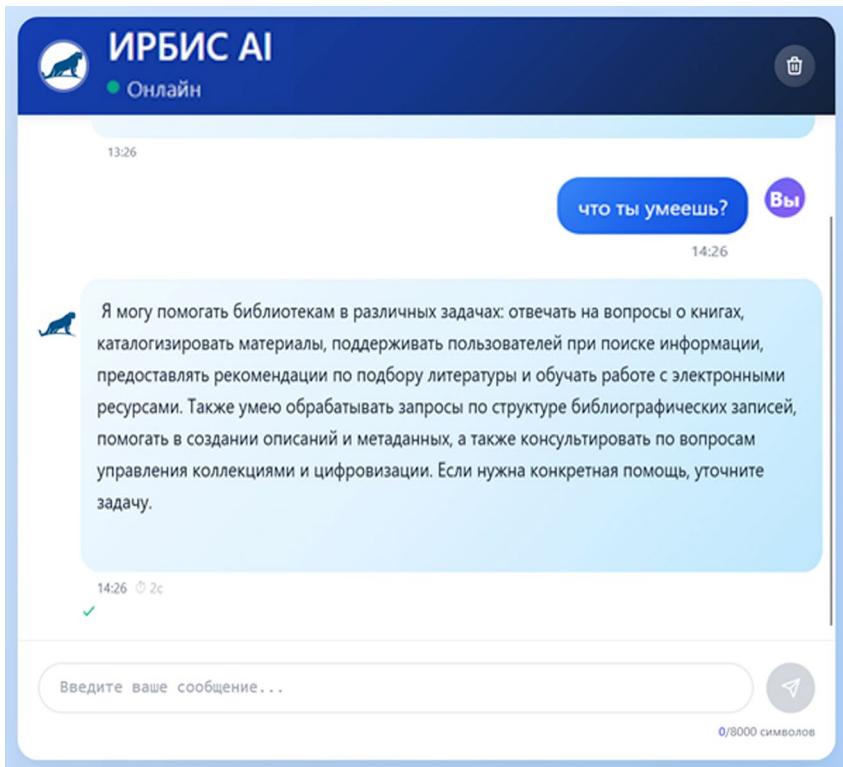


Рис. 3. Веб-интерфейс ИРБИС AI

**Первый режим** – справочное обслуживание. Система предоставляет развёрнутые справки по запросам, сформулированным в строке поиска, используя возможности коммерческих языковых моделей в режиме максимальной точности. Это может быть полезно, если читатель хочет быстро сориентироваться в лексике интересующей его темы.



Рис. 4. Семантическая обработка запросов на естественном языке

**Второй режим** – семантическая обработка запросов на естественном языке запросов. Данный функционал активируется пользователем, если он не удовлетворён результатами первичного поиска. Режим позволяет трансформировать неструктурированные и содержащие ошибки запросы к запросам атрибутивного поиска (автор, заглавие, год издания). Качество работы достигается комбинацией использования языковых моделей и детерминированных алгоритмов J-ИРБИС 2.0, выполняющих анализ баз каталога и оценивающих запрос. Особенно полезной показала себя на практике функция автоматического устранения искусственным интеллектом ошибок и опечаток. При использовании стандартного режима свободного (полнотекстового) поиска слова с опечатками исключаются из запроса и оставшиеся термины приводят к значительному информационному шуму. Устранение опечаток позволяет решить проблему нерелевантных результатов.

**Третий режим** – интеллектуальный подбор литературы в соответствии с семантикой информационного запроса. Из множества найденных записей искусственный интеллект отбирает наиболее релевантные. Это реализуется через многоэтапный информационный обмен с І-ІРБІС 2.0. Отбор выполняется с учётом не только лексического, но и семантического соответствия запросу. Например, литература по производственному браку не будет отобрана, если контекст запроса явно относится к семейной жизни.

## **6. Перспективные направления развития**

### ***6.1. Концепция библиотечного портала нового поколения***

В разрабатываемом модуле І-ІРБІС 3.0 предполагается реализация инновационной концепции библиотечного портала с полностью управляемым искусственным интеллектом поисковым интерфейсом. Данный подход представляет собой качественный скачок в развитии библиотечных систем, связанный с переходом от вспомогательной роли ИИ к его центральной позиции в процессе информационного поиска.

### ***6.2. Проект «ІРБІС ІІ Каталогизатор»***

Революционные изменения в библиотечной практике могут быть достигнуты через использование мультимодальных моделей для обработки печатных документов. В настоящий момент на этапе разработки находится проект «ІРБІС ІІ Каталогизатор». Данное программное решение предполагает фотографирование титульного листа книги мобильным устройством с последующим автоматическим созданием библиографической записи. Сгенерированная в результате распознавания титульного листа запись может быть автоматически обогащена данными из записей ІРБІС-корпорации, ІС ЭКБСОН (минимум 200 млн записей), каталога РГБ (через Z39-50-провайдер) либо просто заменена одним из эталонных вариантов [4].

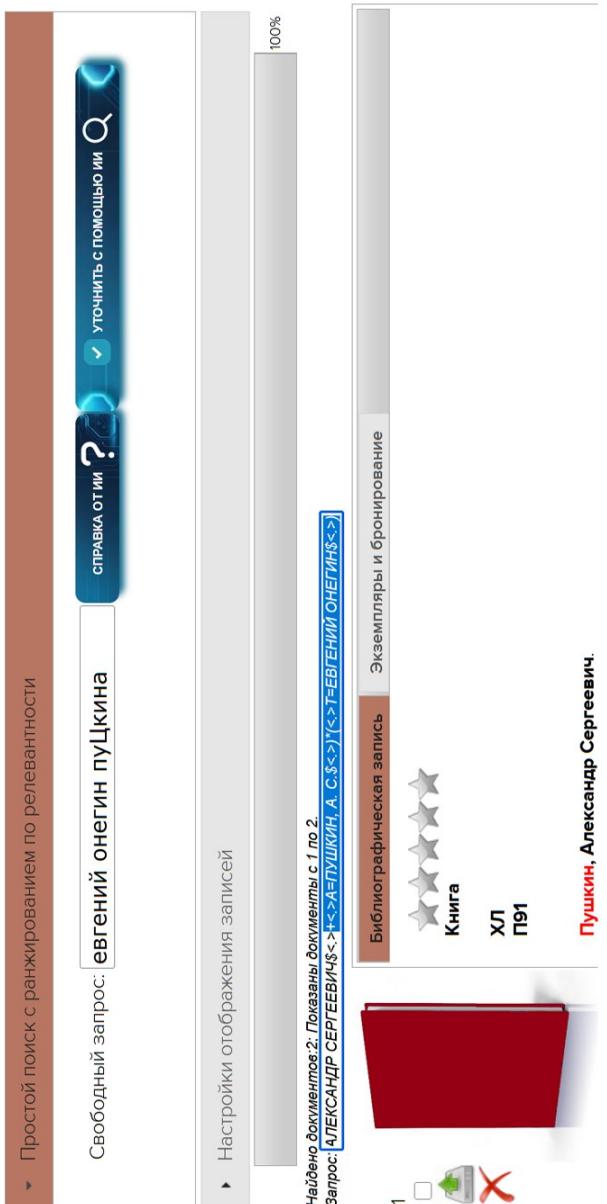


Рис. 5. Результаты исправления ошибки и трансляции на язык запросов ИРБИС  
(можно видеть в строке «Запрос»)

### **6.3. Комплексная автоматизация библиотечных процессов**

Синергия технологий семантической деконструкции, распознавания образов и систем получения готовых библиографических записей формирует основу для решения следующих задач:

1. Автоматическое создание рефератов для электронных публикаций и их фрагментов.
2. Автоматическое формирование оглавлений, списков авторов и других элементов библиографической записи на основе полнотекстового анализа.
3. Автоматизированный перевод фрагментов публикаций.
4. Генерация персонализированных рекомендаций для дальнейшего чтения.
5. Комплексная автоматизация формирования метаданных на основе визуальной и текстовой информации.
6. Автоматическое преобразование имидж-каталогов в структурированные библиографические записи.
7. Обработка неструктурированных данных для импорта.
8. Анализ и сопоставление неформализованных списков литературы с каталогом.
9. Реализация интеллектуального автодополнения при поиске.

## **7. Заключение**

Проведённое исследование демонстрирует, что ограничения и недостатки систем искусственного интеллекта преодолеваются с беспрецедентной скоростью. За последние три года произошли качественные изменения: от изолированных моделей, ограниченных обучающими данными (2022), к системам с доступом к актуальной информации и возможностью управления внешними системами; от неспособности решать логические задачи (2023) к успешному прохождению экзаменов докторского уровня; от кажущегося исчерпания потенциала из-за вычислительных ограничений (2024) к эффективным архитектурам *Mixture of Experts*, радикально снижающим требования к ресурсам.

В условиях стремительной эволюции технологий искусственного интеллекта поиск их недостатков теряет практический смысл. Искусственный интеллект должен стать неотъемлемым компонентом современных информационно-библиотечных систем. Библиотеки, не спо-

собные интегрировать технологию в свою деятельность, рискуют утратить конкурентоспособность в сфере информационных услуг.

Конкурентным преимуществом библиотек в эпоху искусственного интеллекта становятся эксклюзивный верифицированный контент и специализированные сервисы, основанные на глубоком понимании информационных потребностей пользователей. Именно синтез традиционных библиотечных компетенций и инновационных технологий определит будущее библиотек в новой реальности.

## Список источников

1. **Daruvuri V.** Understanding Mixture of Experts (MoE): A Deep Dive into Scalable AI Architecture // International Journal of Scientific Research in Computer Science, Engineering and Information Technology. University of Cincinnati, USA, 2025.  
URL: <https://ijsrcseit.com/index.php/home/article/view/CSEIT251112164/CSEIT251112164> (accessed: 27.10.2025).
2. **RuGPT3** demo // russiannlp.github.io. URL: <https://russiannlp.github.io/rugpt-demo/> (accessed: 28.09.2025).
3. **Sahoo P., Meharia P., Ghosh A., Saha S.** Comprehensive Survey of Hallucination in Large Language, Image, Video and Audio Foundation Models / P. Sahoo, P. Meharia, A. Ghosh, S. Saha, V. Jain, A. Chadha // arXiv.org. 2024. URL: <https://arxiv.org/abs/2405.09589> (accessed: 03.11.2025).
4. **Соколинский К. Е.** Новая технология создания сводных каталогов и корпоративных электронных библиотек в J-ИРБИС 2.0 / К. Е. Соколинский // Научные и технические библиотеки. 2015. № 11. С. 83–100.
5. **Тестируем YandexGPT-5-Pro.** Когда хотелось быть ChatGPT, но в душе всё ещё Алиса // Хабр : блог компании BotHub. 2025. URL: <https://habr.com/ru/companies/bothub/articles/893128/> (дата обращения: 28.09.2025).
6. **Шрайберг Я. Л., Волкова К. Ю.** Вопросы авторского права в отношении произведений, созданных при помощи генеративного искусственного интеллекта // Научные и технические библиотеки. 2025. № 2. С. 115–130.

## References

1. **Daruvuri V.** Understanding Mixture of Experts (MoE): A Deep Dive into Scalable AI Architecture // International Journal of Scientific Research in Computer Science, Engineering and Information Technology. University of Cincinnati, USA, 2025.  
URL: <https://ijsrcseit.com/index.php/home/article/view/CSEIT251112164/CSEIT251112164> (accessed: 27.10.2025).
2. **RuGPT3** demo // russiannlp.github.io. URL: <https://russiannlp.github.io/rugpt-demo/> (accessed: 28.09.2025).
3. **Sahoo P., Meharia P., Ghosh A., Saha S.** Comprehensive Survey of Hallucination in Large Language, Image, Video and Audio Foundation Models / P. Sahoo, P. Meharia, A. Ghosh, S. Saha, V. Jain, A. Chadha // arXiv.org. 2024. URL: <https://arxiv.org/abs/2405.09589> (accessed: 03.11.2025).
4. **Sokolinskii' K. E.** Novaia tekhnologiiia sozdaniia svodny'kh katalogov i korporativny'kh e'lektronny'kh bibliotek v J-IRBIS 2.0 / K. E. Sokolinskii' // Nauchny'e i tekhnicheskie biblioteki. 2015. № 11. S. 83–100.
5. **Testiruem** YandexGPT-5-Pro. Kogda hotelos' by't` ChatGPT, no v dushe vsyo eshchyo Alice // Habr : blog kompanii BotHub. 2025. URL: <https://habr.com/ru/companies/bothub/articles/893128/> (data obrashcheniia: 28.09.2025).
6. **Shrai'berg Ia. L., Volkova K. Iu.** Voprosy' avtorskogo prava v otnoshenii proizvedenii', sozdannyy'kh pri pomoshchi generativnogo iskusstvennogo intellekta // Nauchny'e i tekhnicheskie biblioteki. 2025. № 2. S. 115–130.

## Информация об авторах / Authors

**Гончаров Михаил Владимирович –**  
канд. техн. наук, доцент, ведущий  
научный сотрудник, руководитель  
группы перспективных  
исследований и аналитического  
прогнозирования ГПНТБ России,  
Москва, Российская Федерация  
goncharov@gpntb.ru

**Mikhail V. Goncharov –** Cand. Sc.  
(Engineering), Associate Professor,  
Leading Researcher, Head,  
Perspective Studies and Analytical  
Forecasts Group, Russian National  
Public Library for Science and  
Technology, Moscow,  
Russian Federation  
goncharov@gpntb.ru

**Соколинский Кирилл Евгеньевич** –  
канд. техн. наук, программист  
Международной ассоциации  
электронных библиотек и новых  
информационных технологий  
(Ассоциация ЭБНИТ),  
директор Объединения  
«ИРБИС-Сервис», Санкт-Петербург,  
Российская Федерация  
[info@irbis-service.com](mailto:info@irbis-service.com)

**Шрайберг Яков Леонидович** –  
доктор техн. наук, проф.,  
член-корреспондент Российской  
академии образования, научный  
руководитель ГПНТБ России,  
главный редактор журнала  
«Научные и технические  
библиотеки», заведующий  
кафедрой электронных библиотек  
и научометрических исследований  
Московского государственного  
лингвистического университета,  
заслуженный работник культуры РФ,  
заслуженный деятель науки РФ,  
Москва, Российская Федерация  
[shra@gpntb.ru](mailto:shra@gpntb.ru)

**Kirill E. Sokolinsky** – Cand. Sc.  
(Engineering), Programmer,  
International Association of Users  
and Developers of Electronic  
Libraries and New Information  
Technologies (ELNIT Association);  
Director IRBIS-Service Inc.,  
St. Petersburg, Russian Federation  
[info@irbis-service.com](mailto:info@irbis-service.com)

**Yakov L. Shrayberg** – Dr. Sc.  
(Engineering), Professor,  
Corresponding Member of Russian  
Academy of Education; Director of  
Research, Russian National Public  
Library for Science and Technology,  
Editor-in-Chief, “Scientific and  
Technical Libraries” Journal; Head,  
Department for Electronic Libraries  
and Scientometric Studies, Moscow  
State Linguistic University, Moscow,  
Russian Federation  
[shra@gpntb.ru](mailto:shra@gpntb.ru)