

Применение технологий искусственного интеллекта в библиотечной каталогизации: опыт Национальной библиотеки Республики Саха (Якутия)

А. Г. Бурнашева¹, Я. С. Чабыева², А. Р. Баторов³, С. В. Максимова⁴

^{1, 2, 4}Национальная библиотека Республики Саха (Якутия),
Якутск, Российская Федерация

³Арктический государственный институт культуры и искусств,
Якутск, Российская Федерация

¹ag.burnasheva@nlrs.ru, <http://orcid.org/0009-0001-9276-0247>

²ys.chabieva@nlrs.ru, <https://orcid.org/0000-0002-6782-064X>

³batorov@mail.ru, <http://orcid.org/0000-0003-2061-597X>

⁴maksimova@nlrs.ru, <http://orcid.org/0000-0002-7972-2078>

Аннотация. В статье рассмотрен опыт внедрения технологий искусственного интеллекта (ИИ) в процессы каталогизации Национальной библиотеки Республики Саха (Якутия). Основное внимание уделено двум пилотным проектам: Yakutia.Online – системе интеллектуального веб-архивирования региональных интернет-ресурсов – и проекту по ИИ-каталогизации одночастных изданий. Описаны архитектура систем, этапы автоматизации, механизмы семантической обработки и проверка библиографических записей. Отмечена важность участия человека в финальной верификации данных для обеспечения высокого качества описания. Показано, как применение ИИ позволило ускорить процесс обработки на 40–80%, повысить точность метаданных и интегрировать полученные записи в действующие каталоги библиотеки. Приведено сравнение с ведущими международными практиками (Annif, Library of Congress, Национальные библиотеки Германии и Франции), что подтверждает актуальность и конкурентоспособность опыта библиотеки. Сделан вывод о перспективах масштабирования ИИ-решений в библиотечной сфере и создании Центра компетенций в Арктическом государственном институте культуры и искусств в рамках программы технологического лидерства вузов «Приоритет 2030».

Ключевые слова: искусственный интеллект, библиотека, автоматическая каталогизация, веб-архивирование, цифровое наследие, Национальная библиотека Республики Саха (Якутия)

Для цитирования: Бурнашева А. Г., Чабыева Я. С., Баторов А. Р., Максимова С. В. Применение технологий искусственного интеллекта в библиотечной каталогизации: опыт Национальной библиотеки Республики Саха (Якутия) // Научные и технические библиотеки. 2026. № 4. С. 88–110. <https://doi.org/10.33186/1027-3689-2026-4-88-110>

UDC 025.355:004.8(571.56) + 004.8:02
<https://doi.org/10.33186/1027-3689-2026-4-88-110>

Application of artificial intelligence technologies in cataloging: The experience of the National Library of the Republic of Sakha (Yakutia)

**Anna G. Burnasheva, Yana S. Chabieva, Afanasy R. Batorov
and Sargylana V. Maksimova**

^{1, 2, 4}*National Library of the Republic of Sakha (Yakutia), Yakutsk, Russian Federation*

³*Arctic State Institute of Culture and Arts, Yakutsk, Russian Federation*

¹*ag.burnasheva@nlrs.ru, <http://orcid.org/0009-0001-9276-0247>*

²*ys.chabieva@nlrs.ru, <https://orcid.org/0000-0002-6782-064X>*

³*batorov@mail.ru, <http://orcid.org/0000-0003-2061-597X>*

⁴*maksimova@nlrs.ru, <http://orcid.org/0000-0002-7972-2078>*

Abstract. The experience of implementing artificial intelligence (AI) technologies into cataloging processes at the National Library of the Republic of Sakha (Yakutia) is discussed. The authors focus on two pilot projects: Yakutia.Online – the system of intellectual web archiving of regional Internet resources, – and the project of AI cataloging of one-part editions. The architecture of systems, stages of automation, mechanisms of semantic processing and bibliographic records verification are described. The authors emphasize the value of human contribution to the final verification of data and high quality of bibliographic records. AI application enabled to speed up processing by 40–80%, to improve metadata accuracy and to integrate records into the Library's catalogs. The systems are compared to the renowned international practices (Annif, Library of Congress, National Libraries of Germany and France). The study findings demonstrate the relevancy and competitiveness of the

experience of the National Library of the Republic of Sakha (Yakutia). The authors conclude on the prospects for AI decisions scaling up in library sector. They also inform about the establishment of the Competence Center at the Arctic State Institute of Culture and Arts within the framework of "Priority 2030" academic technological leadership program.

Keywords: artificial intelligence, library, automatic cataloging, web archiving, digital heritage, National Library of the Republic of Sakha (Yakutia)

Cite: Burnasheva A. G., Chabyeva Ya. S., Batorov A. R., Maksimova S. V. Application of artificial intelligence technologies in cataloging: The experience of the National Library of the Republic of Sakha (Yakutia) // Scientific and technical libraries. 2026. No. 4, pp. 88–110. <https://doi.org/10.33186/1027-3689-2026-4-88-110>

Введение

Развитие технологий искусственного интеллекта открывает новые возможности для библиотечного дела. Искусственный интеллект (далее – ИИ) позволяет автоматизировать рутинные процессы, повысить точность метаданных и ускорить ввод новых ресурсов в информационный оборот. Разработки в этой сфере активно ведутся по всему миру: ведущие библиотеки, научные учреждения и технологические компании внедряют ИИ-решения для обработки больших массивов данных, интеллектуального поиска, распознавания текста и изображений, автоматической классификации и тематического анализа документов. Благодаря этим достижениям библиотеки получают инструменты для более эффективного управления информацией и создания персонализированных сервисов для пользователей [1–4].

В 2023 г. Национальная библиотека Республики Саха (Якутия) (НБ РС (Я)) запустила амбициозную программу по внедрению технологий ИИ в свою работу [5]. Для этого библиотека выступила с инициативой о сотрудничестве с одной из ведущих технологических компаний России – ПАО Сбербанк, предложив совместную реализацию проектов по интеграции ИИ в библиотечную среду. В результате партнерства НБ РС (Я) стала первой библиотекой в стране, где на практике применяется API

нейросетевой модели GigaChat для создания специализированных информационных сервисов и цифровых решений [6].

В настоящее время идут работы по запуску двух крупных пилотных проектов, направленных на применение ИИ в каталогизации: Yakutia.Online (в сфере веб-архивирования краеведческих интернет-ресурсов) и эксперимент по каталогизации одночастных изданий. Оба проекта реализованы в тесной связи с общей цифровой стратегией развития библиотеки и стали во многом новаторскими не только для региона, но и для России [7]. Ниже описаны цели, функции и результаты каждого проекта, а также приведены параллели с опытом других библиотек.

Проект Yakutia.Online: ИИ-каталогизация веб-ресурсов

Современный интернет представляет собой высокодинамичную среду, где информация постоянно обновляется, перемещается или исчезает. Поэтому практически одновременно с бурным развитием интернета возникла идея веб-архивирования для сохранения большого количества информации, имеющей историческую, культурную и научную ценность. Именно национальные библиотеки стали играть ключевую роль в веб-архивировании, так как изначально созданы для сбора, хранения и обеспечения доступа ко всем значимым документам, появляющимся в стране. В цифровую эпоху веб-контент – это тоже часть национального наследия.

В Российской национальной библиотеке разработана модель краеведческого веб-архива, отражающего историю и культуру региона [8]. Для НБ РС (Я) задача веб-архивирования тоже актуальна, потому что множество ценных материалов о жизни республики публикуется исключительно в Сети.

Проект Yakutia.Online направлен на создание регионального веб-архива и автоматическое описание сетевых источников, связанных с Якутией [9]. Комплексное веб-архивирование систематически сохраняет ценные интернет-ресурсы, имеющие отношение к истории, культуре и современному развитию республики. Речь идет о локальных новостных сайтах, электронных периодических изданиях, блогах, порталах органов власти и т. п., где публикуются материалы о регионе. Принципиальное требование – использовать только открытые источники и соблюдать права правообладателей: каждый архивируемый материал сопровождается указанием авторов и первоисточника, и не сохраняются ресурсы, требующие специальных разрешений. Такой подход обеспечивает пра-

вомерность создания архива и его последующего использования. Формируется тематический архив веб-страниц (в международном формате WARC) – по сути, цифровой фонд региональной интернет-памяти.

В начале 2024 г. был проведен масштабный мониторинг интернет-источников, публикующих материалы о Республике Саха (Якутия). Критериями отбора сайтов стали: соответствие краеведческой тематике, географическая привязка (региональные ресурсы или федеральные, освещающие регион), культурная и историческая значимость публикуемых материалов, наличие материалов на языках народов республики, востребованность информации у потенциальных пользователей. В итоге сформирован корпус источников: включено 26 республиканских СМИ (новостные сайты ЯСИА, «Якутия24», SakhaLife, SakhaNews, SakhaPress, Ulus.media и др.), 32 федеральных российских ресурса с разделами по Якутии («РИА Новости», EastRussia, «Региональная Россия» и др.), а также шесть зарубежных сайтов, публикующих статьи о Якутии и Арктике. Все выбранные ресурсы открыты и общедоступны, что позволило сразу приступать к их архивации без правовых препятствий.

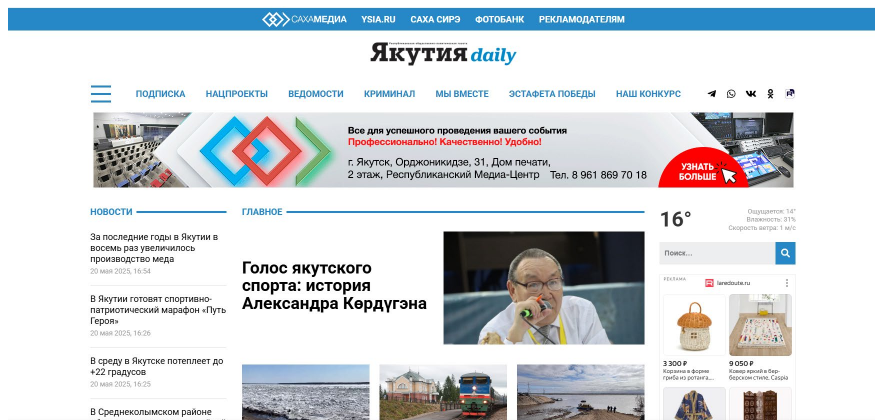


Рис. 1. Веб-страница сайта «Якутия.daily»

Для каждого включенного сайта определены параметры отбора материалов. Они похожи на подход к отбору печатных краеведческих документов, но адаптированы к специфике веба. Учитываются содержательные характеристики (информационная ценность, достоверность, полнота раскрытия темы), язык и страна происхождения, авторитетность

источника; технические параметры (стабильность URL, риск исчезновения сайта, удобство копирования контента); а также объем и значимость информации (насколько полно тема представлена, долговременная ценность). Материалы, попадающие под эти критерии, классифицируются по форматам: текстовые (статьи, документы), мультимедийные (аудио/видео), визуальные (фотографии, инфографика), что затем учитывается при каталогизации (разные отделы занимаются разными типами, см. ниже).

После завершения отбора начинается этап библиографической обработки отобранных ресурсов. В НБ РС (Я) выстроена распределенная схема: Отдел каталогизации составной части ресурсов (ОКСЧР) занимается статьями из электронных периодических изданий – создает библиографические записи на каждую статью и размещает их в сводной базе краеведческих статей «Саха сирэ». Отдел каталогизации ресурсов (ОКР) обрабатывает публикации из сетевых СМИ – формирует монографические записи на ресурсы и включает их в Сводный каталог электронных ресурсов РС (Я). Такой подход позволяет охватить все типы электронных краеведческих материалов: и отдельные статьи, и целые издания, соблюдая при этом единые стандарты описания и координируя работу двух отделов.

Технологии ИИ применяются на этапе отбора и каталогизации ресурсов. ИИ-модуль системы (робот) выполняет ежедневный мониторинг заданных сайтов, выявляет релевантные публикации и формирует заготовки записей в формате RUSMARC. Далее каталогизатор проверяет заявку и подтверждает добавление в электронный каталог. Таким образом, значительная часть рутинной работы по поиску и первичному описанию осуществляется без участия человека.

Система «Веб-архивирование» включает два модуля с веб-интерфейсами: «Настройки робота» и «Заявки на веб-архивирование». Модуль «Настройки робота» позволяет легко управлять тематическими критериями отбора контента для архивирования. При необходимости специалисты могут добавлять новые тематические критерии – гибкая настройка дает возможность адаптировать поиск под актуальные информационные запросы. Робот работает непрерывно: ежедневно сканирует страницы выбранных сайтов в поисках новых публикаций, сопоставляет их с критериями отбора и передает подходящие материалы в обработку.

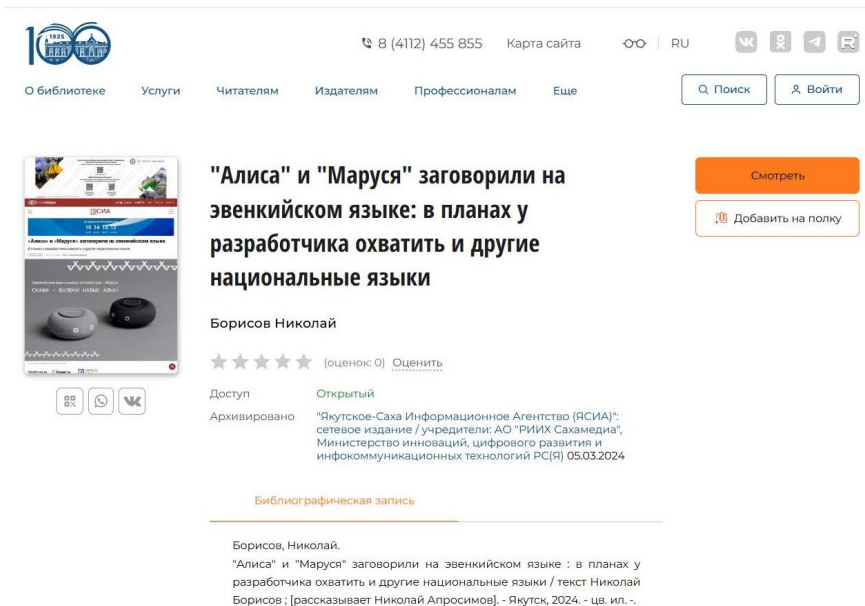


Рис. 2. Архивированная публикация на сайте НБ РС (Я)

Интерфейс «Заявки на веб-архивирование» реализует модель Human-in-the-Loop: подтверждение созданной роботом заявки и проверка автоматически сгенерированной RUSMARC-записи производится человеком. Когда робот обнаруживает новый ресурс, соответствующий критериям, ИИ-каталогизатор автоматически создает заявку и черновик библиографической записи (в формате RUSMARC) на этот ресурс, перед этим делая автоматическую проверку на дублетность, исключая повторное архивирование. Далее каталогизатор просматривает список заявок через веб-интерфейс: для каждой заявки он открывает содержимое ресурса, верифицирует его соответствие критериям (убеждается, что материал действительно краеведческий, качественный, полная версия и т. п.), после чего либо подтверждает заявку, либо отклоняет. После подтверждения заявки на создание библиографической записи система автоматически переносит заготовку записи в базу (Сводный каталог электронных ресурсов) и открывает ее для редактирования в интерфейсе АИБС OPAC-Global. Отклоненные заявки помечаются и не поступают в архив.

В завершение процесса происходит редактирование библиографической записи и архивация ресурса. Каталогизатор проверяет автоматически сгенерированную запись в формате RUSMARC. Как правило, ИИ вносит основную информацию (заглавие, автор/источник, дата, ссылка, аннотация). Задача специалиста – дополнить и откорректировать запись: например, добавить соавторов, переводчиков, если имеются. Важный момент – использование авторитетных данных: если встречается новое имя, отсутствующее в Национальном авторитетном файле Республики Саха (Якутия), каталогизатор заводит соответствующую авторитетную запись.

После финализации библиографической записи робот осуществляет веб-архивирование страницы путем ее сохранения на сервере библиотеки и создания электронного объекта. Для этого используется уникальный идентификатор записи (OPAC ID), который увязывает запись и копию страницы. Сохранение веб-страницы производится в формате WARC, принятом международным стандартом для долгосрочного хранения веб-контента. Перед сохранением контент проверяется на целостность и отсутствие вредоносного кода, после чего помещается на защищенные серверы библиотеки.

Доступ пользователей к архивированным ресурсам организован через единый поисковый сервис «ИЛИМ» и официальный сайт библиотеки: веб-архивы представлены в виде структурированных электронных объектов с полным библиографическим описанием. Интеграция с поисковой системой «ИЛИМ» обеспечивает пользователям быстрый доступ к архивированным ресурсам. «ИЛИМ» (поисковая строка) предоставляет пользователям единую точку доступа ко всем оцифрованным материалам и веб-архиву, удобный поиск по каталогам, тематическим коллекциям и возможность работы с архивами через портал библиотеки. В результате создана единая экосистема, в которую включаются библиографические записи, краеведческие веб-ресурсы и цифровые копии из документного фонда библиотеки.

000 000000nam0 220000003: 450
001 NLR\$ BVL_0002480118
005 20250520170709 0
021 ##\$aRU
100 ##\$a20250326:20220422##\$u0tsey50#####a
101 0#Sarus
102 ##\$aRUSb411
105 ##\$y##\$##000y
135 ##\$dref#...umun
139 ##\$scat#
184 #0\$u#f\$bwX#e##
184 #0\$u#b
200 ##\$aГригорий Савинов, "Результаты исследований Вилейской экспедиции являются уникальными"
210 ##\$aГригорий Савинов, "Результаты исследований Вилейской экспедиции являются уникальными"
216 ##\$aБайтук\$Q2022
225 1#S\$Экология
225 1#S\$Экология
300 ##\$aЗагл. с титул. экрана
300 ##\$aДата публикации: 22.04.2022
330 ##\$aДиректор НИИ прикладной экологии Севера СВФУ, доктор биологических наук Григорий Савинов рассказал об итогах Комплексных научных исследований пойменных экосистем реки Виллой и ее притоков с 2020 по 2021 гг.
333 ##\$a16+
337 ##\$aURL... https://yakutia-daily.ru/grigorij-savinov-rezultaty-issledovaniy-vilyujskoj-ekspedicii-v-yvuyuzure-unikalnyy (дата обращения: 26.03.2025)
371 ##\$aПри использовании любых материалов активная гиперссылка обязательна
606 1#S2#tr sakda_#b\$3NLR\$AF_0000056041S\$Экологические системы пойменныхУВиллой, река (Сибирь, Восточная)
607 #S2#tr_#b1S3RUNLR_AUTH661924108Виллой, река (Сибирь, Восточная)\$X\$Экология
675 ##\$a502.5:1(282.2:36.66)\$v4\$2rus
686 ##\$b26.222:\$fПоч.Як.5:\$lVLBC/AS2:nbkk
800 #1S\$UNL\$#P\$AF_000002438\$савинов\$BГ. Н.\$gГригорий Николаевич\$ст-р биол. наук\$1957
800 #1S\$UNL\$#P\$AF_000002438\$савинов\$BГ. Н.\$gГригорий Николаевич\$ст-р биол. наук\$1957
801 #1S\$UNL\$#P\$Cav\$S20250326\$2rus
856 40S#tr#p://yakutia-daily.ru/grigorij-savinov-rezultaty-issledovaniy-vilyujskoj-ekspedicii-v-yvuyuzure-unikalnyy#s\$т-р биол. наук\$1957
899 ##\$aННР Сава\$В\$б\$архив
909 ##\$aЭлектронный ресурс. Однотомник
915 ##\$aЭлектронные издания

Закрыть

Рис. 4. ИИ-запись веб-архивированного ресурса в формате RUSMARC (модель НІТЛ)

Внедрение ИИ-технологий в процесс веб-архивирования продемонстрировало явные преимущества. Существенно ускорилась обработка ресурсов: по внутренним оценкам, время составления одной записи сократилось примерно на 40% относительно прежнего уровня. Повысилась точность формирования метаданных – доля ошибок и пропусков сведений снизилась, поскольку алгоритм строго следует заданным шаблонам и проверяет дублирования. Оптимизировались трудозатраты: вместо того чтобы вручную мониторить десятки сайтов, сотрудники теперь занимаются только проверкой и уточнением предложенных системой записей.

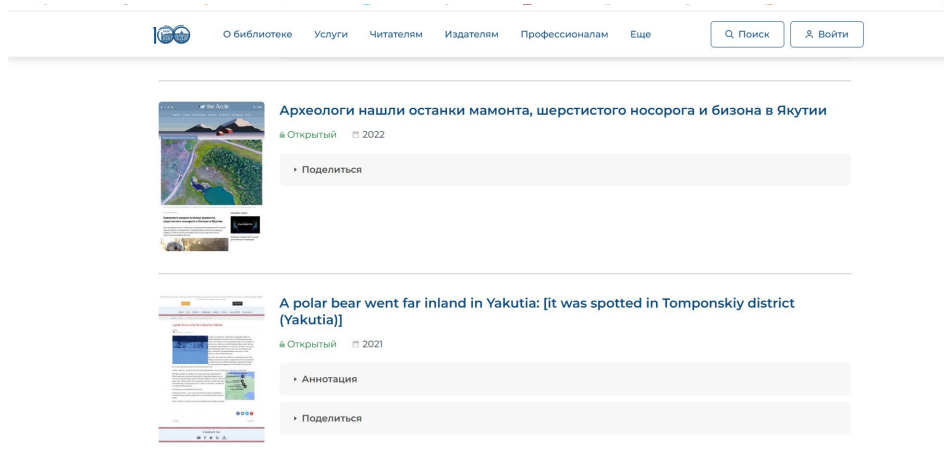


Рис. 5. Представление веб-архива на сайте библиотеки

Проект ИИ-каталогизации одночастных изданий

В марте 2025 г. Национальная библиотека РС (Я) запустила в тестовом режиме систему интеллектуальной каталогизации одночастных изданий (оцифрованных книг и брошюр) с помощью технологий ИИ. Цель нового проекта – автоматизировать трудоемкие операции, повысить точность и скорость описания документов, сохраняя при этом высокое качество библиографических записей.

Данный проект реализует ИИ-каталогизацию одночастных изданий – цифровых копий книг и брошюр с интеграцией в библиотечные технологические процессы.

Комплексный процесс включает несколько ключевых этапов:
автоматизированное распознавание текста,
семантический анализ содержания,
автоматическая генерация библиографических записей,
верификация и доработка.

В Национальной библиотеке РС (Я) внедрена система оцифровки всех поступающих документов, что позволяет организовать удаленную работу каталогизаторов и существенно повысить эффективность технологического процесса. Оцифровка интегрирована в рабочий цикл библиотеки через автоматизированную информационную систему комплектования и каталогизации фондов – АИС «Электронный образ документа» (ЭОД).

Система ЭОД предусматривает два основных сценария работы:

1. Использование уже существующих полных электронных копий изданий, поступающих по местному обязательному экземпляру.

2. Создание электронных образов документов непосредственно для целей каталогизации.

Оцифровка проводится по единым техническим стандартам:

Исходный формат: TIFF.

Конечный формат: PDF с текстовым слоем (OCR).

Минимальное разрешение: 300 dpi.

Обязательные элементы сканирования: обложка, титульный лист с оборотом, оглавление, предисловие.

Процесс работы в ЭОД начинается с создания рабочей записи в системе:

Комплектатор инициирует рабочую запись на каждое наименование, присваивая каждому экземпляру уникальный инвентарный номер.

Корректность заполнения базовых и служебных метаданных (автор, заглавие, выходные сведения) является обязательным условием для прохождения документа по заданному маршруту обработки в АИС ЭОД.

Особенности работы в системе:

1. Все этапы оцифровки и каталогизации проходят через автоматизированные контрольные точки и документируются в системе.

2. Реализована модель распределенной командной работы: каждый участник процесса видит только задачи, соответствующие его роли.

3. Обеспечена полная прослеживаемость всех изменений, внесенных в записи.

4. Каталогизаторы могут работать дистанционно, без потери качества и скорости обработки информации.

Результатом работы системы являются:

Полноценные машиночитаемые библиографические записи в формате RUSMARC с учетом инвентарных номеров и шифров хранения.

Готовые печатные карточки для традиционного каталога.

Связь библиографических записей с электронными образами документов и интеграция с Электронной библиотекой НБ РС (Я).

Синхронизация с системой Обязательного электронного экземпляра (ОЭК).

Автоматическая статистическая отчетность, включающая данные о количестве созданных записей, оцифрованных страниц, поступивших и обработанных экземпляров.

Применение АИС ЭОД значительно снижает трудозатраты сотрудников, устраняет рутинные операции и минимизирует риски потери информации. Правильно выстроенный цифровой процесс обеспечивает своевременную регистрацию, каталогизацию и доступность новых поступлений, делая работу библиотеки прозрачной, управляемой и соответствующей современным требованиям.

Для обработки поступающих в библиотеку документов развернута отдельная технологическая цепочка, встроенная в общую инфраструктуру процесса каталогизации. По завершении сканирования документа полученный PDF-файл прикрепляется к рабочей записи АИС ЭОД. Для одночастных изданий система автоматически извлекает из прикрепленного файла данные, необходимые для каталогизации, с использованием технологий ИИ.

[Область](#)

ЭЗД

Здесь отображаются записи, для которых готов ЭЗД (саны) и есть данные для заполнения краткой БЗ (включая данные для 021, 899 поля). Желтым фоном выделены записи, для которых еще не назначен каталогизатор, а также назначенные Вами записи, у которых отсутствует БЗ.

Введите ID записи/заглавие/номер/номер записи КСУ (партии) и нажмите Enter

Каталогизатор

Дата создания записи комп-ом

Сортировка

Вид документа

Год поступления

ID	№ записи КСУ (партии)	Инд. номера	Экз.	Цена экз. (₽)	Заглавие	Автор	Кол-во отсказ. страниц	Статус	Назначенные работники	Дата создания записи комп-ом	Действия
55185	2025/114	Др4481/38	1	0.44	Серен В. К., Дурья Семена - стихи [для дошкольного и младшего школьного возраста] - Якутск: Якутское книжное издательство, 1977		0	ЭЗД готов	Комплекатор: Абрамов Айтал Юрьевич	20.05.2025 17.13	
54596	2025/376	1663/32	1	189.00	Толстикова А. Я., За Царя и Отечество. Российские генерал-фельдмаршалы [для среднего школьного возраста]. - Москва: Белый город, 2008		12	ЭЗД готов	Сканировщик: Архипов Дмитрий Юрьевич Комплекатор: Афанасьева Светлана Анатольевна	07.05.2025 15.38	
55180	2025/114	Др4481/35	1	25.00	Сивцева И. К., Аан дойду уонна Саха сириг бэргэри күүстэхтэрэ. 1870-1920 - Дьокуускай: Бичик, 2001		0	ЭЗД готов	Комплекатор: Абрамов Айтал Юрьевич	20.05.2025 16.49	
54565	2025/376	1663/31	1	207.02	Салпаев К., Веники - Москва: Мир книг,		28	ЭЗД готов	Сканировщик: Архипов Дмитрий Юрьевич	07.05.2025	

Рис. 6. Главная страница АИС ЭЗД

После извлечения необходимых для каталогизации данных система переходит непосредственно к созданию черновика RUSMARC-записи. Используя заложенные шаблоны и извлеченные данные, ИИ формирует запись в формате RUSMARC по стандартам библиографического описания, совместимую с библиотечной АИБС OPAC-Global. Генерируются все основные поля: заголовок, область заглавия и сведений об ответственности (куда подставляются авторы, заглавие, переводчики и т. д.), выходные данные (место, издательство, дата), физическая характеристика (число страниц – эту информацию можно извлечь из библиографической записи), серия (если выявлена), примечания. Система автоматически проставляет идентификационные номера и контрольные признаки. Особое внимание уделено полям доступа: подключается Национальный авторитетный файл РС (Я) для проверки имен (автор, редактор, герой произведения и пр.). Если имя найдено в авторитетном файле, в запись помещается ссылка на него.

Затем система автоматически извлекает индексы УДК и ББК. В завершение к записи добавляется аннотация – либо извлеченная из документа в исходном виде, либо, при необходимости, сгенерированная системой. По сути, на этом этапе получается почти полноценная библиографическая запись, которую остается проверить.

ID: 54813. История противотуберкулезной службы Верхоянского района - Якутск: Көмүөл, 2024

The screenshot shows a user interface for a library catalog record. At the top, there is a green button labeled 'Сохранить' (Save) and a blue button labeled 'Создать БЗ с помощью ИИ' (Create BZ with IIR). Below this, the record details are displayed:

- Status: ЭОД готов (EOB ready), БЗ готова (BZ ready)
- Date of record creation: 15.05.2025 09:26
- Download PDF: Сканать (Scan)
- Cataloger: Софроньева Ольга Степановна (sofroneeva_os)
- Short BZ ready: 021, 899 полей (не считая место xр.) with radio buttons for Да (Yes) and Нет (No), where 'Да' is selected.
- Additional cataloging (из Сандаль-Биче) with radio buttons for Не требуется (Not required), Требуется (Required), Исполнено (Completed), and Проверено (Checked).
- BZ ready: with radio buttons for Да (Yes) and Нет (No), where 'Да' is selected.

Рис. 7. Кнопка «Создать БЗ с помощью ИИ»

Как и в предыдущем проекте, последний этап – обязательная проверка и редактирование записи каталогизатором. Специалисты просмат-

ривают каждую автоматически созданную запись, чтобы убедиться в ее корректности. Особое внимание обращается на проблемные места: правильно ли определены все авторы (например, соавторы или составители могли быть пропущены), нет ли ошибок в датах (ИИ мог неверно интерпретировать типографские особенности, например римские цифры), корректно ли выбраны предметные рубрики и индексы классификации. Поскольку значительная часть фонда носит краеведческий характер, специалисты проверяют, учел ли ИИ все специфические детали регионального содержания. Иногда автоматическая обработка может не уловить тонких нюансов – к примеру, книга может относиться сразу к двум областям (история и этнография), и тогда каталогизатор добавляет вторую рубрику вручную. В процессе проверки запись дополняется, если чего-то не хватило, и устраняются ошибки. Лишь после такого ревью запись считается завершенной и поступает в электронный каталог, а цифровая копия книги становится доступна пользователям.

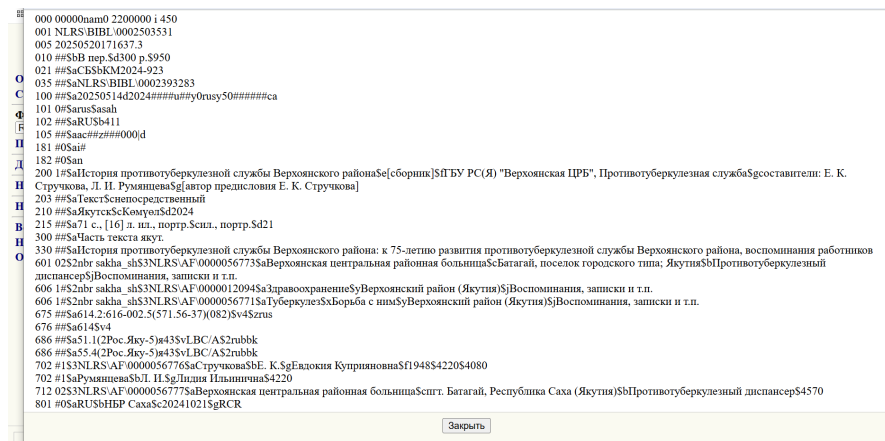


Рис. 8. ИИ-запись одночастного ресурса в формате RUSMARC (модель HITL)

В рамках пилотного тестирования ИИ-каталогизации одночастных изданий были достигнуты многообещающие результаты. Время обработки одной книги сократилось радикально: если раньше на создание записи могло уходить до 40 минут специалиста, то теперь автоматизированная система выдает заготовку библиографической записи за считанные минуты, а каталогизатору требуется лишь немного времени на проверку

и уточнение. Точность библиографического описания повысилась: по оценкам, доля ошибок и упущений в готовых записях снизилась примерно на 40%. Это означает, что записи сейчас более полные и соответствуют стандартам, стало меньше пустых полей или ошибок в именах.

С архитектурной точки зрения процесс генерации RUSMARC-записи обеспечивается отдельным роботом, представляющим собой микросервис, написанный на языке Python. Данный микросервис получает на вход идентификатор рабочей записи в АИС ЭОД, скачивает связанный с этой записью PDF-файл, извлекает необходимые для каталогизации данные, строит и возвращает в своем ответе черновик RUSMARC-записи.

Сравнение с мировыми практиками ИИ-каталогизации

Для оценки проектов Якутии важно сопоставить их с ведущими международными и российскими инициативами. Ниже приведена сравнительная таблица ряда проектов.

Проекты библиотек по ИИ-каталогизации

Проект (организация, страна)	Краткое описание	Результаты/статус
Yakutia.Online (Национальная библиотека РС (Я), Россия)	Региональный веб-архив + ИИ-каталогизация веб-контента	~40% ускорение описания веб-ресурсов; интеграция с каталогом и поиском; пилот стартовал в 2023 г., теперь рабочая служба библиотеки.
ИИ-каталогизация одночастных изданий (Национальная библиотека РС (Я), Россия)	ИИ-каталогизация на основе PDF-файла	Время на запись сократилось с ~40 минут до нескольких минут; ошибок в метаданных ~ на 40% меньше; пилот (2025) успешно интегрирован в АИБС, обеспечивая ~90% соответствия стандартам после контроля.
Annif (Национальная библиотека Финляндии)	Open-source инструмент, объединяющий различные алгоритмы и ML-модели в одном решении для автоматической предметной индексации	После успешного пилотирования внедрен в рабочий процесс для присвоения предметных рубрик новым поступлениям. Сокращает нагрузку на индексаторов и повышает единообразие предметизации.

Проект (организация, страна)	Краткое описание	Результаты/статус
Exploring Computational Description (Library of Congress, США)	Эксперимент по автоматическому описанию электронных книг (использование различных ML-моделей для заполнения полей MARC-записи)	Проведен в 2022–2024 гг.; разработаны прототипы ассистентов каталогизатора. Выводы: ИИ отлично выделяет базовые поля (авторы, заглавия), но тематические рубрики и аннотации все еще требуют участия человека. Библиотекари положительно восприняли AI-помощника, готовы использовать при ясных преимуществах.
KI (Национальная библиотека Германии)	Эксперименты с использованием технологий ИИ для семантической привязки работ к ключевым словам из Gemeinsame Normdatei (GND)	Проводился с 26 апреля 2021 г. по 31 марта 2025 г. Основной результат проекта – разработка и внедрение системы Erschließungsmaschine, которая с апреля 2022 г. используется для автоматической классификации и индексирования публикаций.
FINLAM (Национальная библиотека Франции)	Целью проекта является разработка мультимодальных моделей (текст – изображение) для извлечения информации из оцифрованных или изначально цифровых документов.	Проводится с 2023 по 2027 г. в процессе.

Как видим, проекты НБ РС (Я) находятся в русле современных тенденций, которые демонстрируют переход библиотек от экспериментального тестирования ИИ-решений к их постепенной интеграции в рабочие процессы.

Проект Yakutia.Online стал первым примером регионального веб-архива с автоматической каталогизацией страниц в библиотечном каталоге, тогда как даже национальные библиотеки (например, Библиотека Конгресса), хотя и обладают гигантскими веб-архивами, обычно хранят их обособленно, без столь детальной каталогизации каждой страницы (поиск там чаще осуществляется по полнотекстовому индексу).

Финский Annif, немецкий KI и эксперименты LOC подтверждают востребованность ИИ в предметной обработке и описании, но сосредотачиваются на отдельных участках работы (только предметные рубрики или только книги), тогда как в НБ РС (Я) ИИ внедряется комплексно на разных типах ресурсов.

Вызовы и риски внедрения ИИ-решений в библиотечной практике

Несмотря на очевидные преимущества внедрения технологий ИИ в процессы каталогизации, необходимо учитывать ряд рисков и ограничений, влияющих на долгосрочную устойчивость решений.

Особую актуальность приобретают вопросы информационной безопасности: цифровые системы библиотеки работают с массивами данных, включая персональные сведения пользователей и авторитетные файлы. Нарушения целостности или утечки информации могут поставить под угрозу доверие к библиотечным сервисам.

Также важное значение имеют правовые аспекты, связанные с соблюдением авторского права и лицензионных ограничений. При создании цифровых копий и архивировании интернет-ресурсов необходимо строго следовать правовым нормам, указывая авторов, источники и исключая объекты, требующие специальных разрешений.

Следующим фактором является долгосрочное хранение цифровых данных. Даже при использовании международных стандартов (например, формата WARC для веб-архивирования) есть риск устаревания технологий и необходимости миграции данных в новые форматы. Это требует разработки стратегии цифровой преемственности и устойчивого хранения.

Наконец, нельзя игнорировать этические вызовы: автоматизированные системы не должны подменять профессиональное суждение библиотекаря. Оптимальной является модель, при которой специалист выполняет финальную верификацию записей, обеспечивая сохранение качества и предотвращение возможных искажений [10].

Таким образом, успешное применение ИИ в библиотечном деле требует не только технологической эффективности, но и всестороннего анализа рисков, что позволит выработать устойчивые и воспроизводимые практики на долгосрочную перспективу.

Перспективы дальнейших исследований (вместо заключения)

Представленный опыт Национальной библиотеки РС (Я) отражает первый в России региональный кейс комплексного внедрения ИИ в процессы каталогизации. Вместе с тем авторы осознают ограничения проведенного исследования [11].

Во-первых, описан опыт только одной библиотеки, что не позволяет делать обобщения на национальном уровне. Данный проект рассматривается как пилотный, его результаты планируется масштабировать в сотрудничестве с другими библиотеками России.

Во-вторых, в статье основной акцент сделан на технологическом описании и первых результатах (сокращение времени обработки на 40–80%, снижение ошибок примерно на 40%). В дальнейшем планируется проведение более детального анализа масштабов применения ИИ, расчетов экономической эффективности и оптимизации трудозатрат.

Третьим направлением станет разработка более строгих критериев оценки качества автоматизации. Помимо уже приведенных количественных показателей, будет систематизирована статистика ошибок и успешности верификации, что позволит выработать комплексные методики контроля.

Кроме того, важной задачей остается изучение пользовательского опыта. Планируется организация анкетирования и интервьюирования как сотрудников библиотеки, так и читателей, чтобы оценить влияние ИИ-каталогизации на удобство поиска и качество обслуживания.

Таким образом, представленный проект является лишь первым шагом в систематическом исследовании внедрения технологий ИИ в библиотеках России. Его продолжение позволит учесть обозначенные вызовы и развить научно-методическую основу для комплексного применения ИИ в отрасли. Общий вывод заключается в том, что оптимальной моделью работы признано взаимодействие «человек + ИИ». Во всех успешных примерах ИИ выступает как инструмент ускорения процессов и расширения охвата, а окончательное качество обеспечивается профессиональной экспертизой специалистов.

Перспективным направлением является масштабирование библиотечных ИИ-технологий. Разработки НБ РС (Я) получили поддержку в рамках государственной программы стратегического академического лидерства «Приоритет 2030», реализуемой на базе Арктического государственного института культуры и искусств (г. Якутск). В ближайших планах –

создание Центра компетенций по искусственному интеллекту в библиотечном деле. Центр станет площадкой для научных исследований, внедрения и масштабирования передовых ИИ-решений, а также подготовки нового поколения специалистов в сфере цифровой гуманитаристики.

Список источников

1. **Brador I.** Could Artificial Intelligence Help Catalog Thousands of Digital Library Books? An Interview with Abigail Potter and Caroline Saccucci. The Signal: Library of Congress Blog, 19.11.2024. URL: <https://blogs.loc.gov/thesignal/2024/11/could-artificial-intelligence-help-catalog-thousands-of-digital-library-books-an-interview-with-abigail-potter-and-caroline-saccucci/> (accessed: 10.05.2025).
2. **Suominen O.** Annif: DIY automated subject indexing using multiple algorithms. LIBER Quarterly. Vol. 29, No. 1, 2019. DOI 10.18352/lq.10285 (National Library of Finland open source project Annif) (accessed: 10.05.2025).
3. **Искусственный интеллект в библиотеке.** ЛаЛалась (блог ЭБС «Лань»), [без даты]. URL: <https://lala.lanbook.com/iskusstvennyj-intellekt-v-biblioteke> (дата обращения: 10.05.2025).
4. **«КорФор-2024»:** круглый стол «ИИ в науке, образовании и работе с информацией: от ожиданий к реальности» // Университетская книга (онлайн), 2024. URL: <https://www.unkniga.ru/ai/ii-v-nauke-obrazovanii-i-bibliotekakh.html> (дата обращения: 10.05.2025).
5. **Об ИИ-продуктах** НБ РС (Я) : [видеозапись] / Национальная библиотека РС (Я). Якутск: НБ РС (Я), 2025. 1 видеофайл (MP4; 2 мин 20 с). URL: <https://l.nlrs.ru/open/124841> (дата обращения: 16.04.2025).
6. **Национальная библиотека Якутии** первой в России внедрила в свои сервисы GigaChat: за первые шесть месяцев работы AI-помощник обработал более 33 тысяч запросов пользователей. ЯСИА, 13.12.2024. URL: <https://ysia.ru/natsionalnaya-biblioteka-yakutii-pervoj-v-rossii-vnedrila-v-svoi-servisy-gigachat> (дата обращения: 18.03.2025).
7. **Максимова С. В.** Новая стратегия развития цифровой интегрированной системы Национальной библиотеки Республики Саха (Якутия) // Научные и технические библиотеки. 2024. № 3. С. 127–141.
8. **Балацкая Н. М., Мартиросова М. Б.** Краеведческий веб-архив в структуре информационных ресурсов библиотеки: модель и возможности реализации. Санкт-Петербург, 2023. 207 с.
9. **Ершова В. В.** Краеведческие ресурсы интернета как объект веб-архивирования: опыт реализации проекта Yakutia.Online в Национальной библиотеке Республики Саха (Якутия) // Краеведческие ресурсы сети как объект веб-архивирования и исследования : сб. статей. Санкт-Петербург, 2024. С. 98–114.

10. **Lowagie H.** From Bias to Transparency : Ethical Imperatives in AI-Based Library Cataloging / Hannes Lowagie // IFLA. 23.09.2023. URL: <https://repository.ifla.org/handle/20.500.14598/2841> (accessed: 29.09.2025).
11. **Искусственный интеллект в деятельности Национальной библиотеки Республики Саха (Якутия) / С. С. Климов, А. Г. Бурнашева, И. К. Федоров и др. // Библиотековедение. 2025. Т. 74, № 4. С. 341–358. DOI 10.25281/0869-608X-2025-74-4-341-358.**

References

1. **Brador I.** Could Artificial Intelligence Help Catalog Thousands of Digital Library Books? An Interview with Abigail Potter and Caroline Saccucci. The Signal: Library of Congress Blog, 19.11.2024. URL: <https://blogs.loc.gov/thesignal/2024/11/could-artificial-intelligence-help-catalog-thousands-of-digital-library-books-an-interview-with-abigail-potter-and-caroline-saccucci/> (accessed: 10.05.2025).
2. **Suominen O.** Annif: DIY automated subject indexing using multiple algorithms. LIBER Quarterly. Vol. 29, No. 1, 2019. DOI 10.18352/lq.10285 (National Library of Finland open source project Annif) (accessed: 10.05.2025).
3. **Iskusstvenny`i` intellekt v biblioteke. LaLaLan` (blog E`BS «Lan`»), [bez daty`]. URL: <https://lala.lanbook.com/iskusstvennyj-intellekt-v-biblioteke> (дата обращения: 10.05.2025).**
4. **«KorFor-2024»: krugly`i` stol «II v nauke, obrazovanii i rabote s informatsiei`: ot ozhidani` k real`nosti` // Universitetskaia kniga (onlai`n), 2024. URL: <https://www.unkniga.ru/ai/ii-v-nauke-obrazovanii-i-bibliotekakh.html> (data obrashcheniia: 10.05.2025).**
5. **Ob II-produktakh NB RS (Ia) : [videozapis`] / Natsional`naia biblioteka RS (Ia). Iakutsk: NB RS (Ia), 2025. 1 videofai`l (MP4; 2 min 20 s). URL: <https://l.nlr.ru/open/124841> (data obrashcheniia: 16.04.2025).**
6. **Natsional`naia biblioteka Iakutii pervoi` v Rossii vnedrila v svoi servisy` GigaChat: za pervy`e shest` mesiatcev raboty` AI-pomoshchnik obrabotal bolee 33 ty`siach zaprosov pol`zovatelei`. IASIA, 13.12.2024. URL: <https://ysia.ru/natsionalnaya-biblioteka-yakutii-pervoj-v-rossii-vnedrila-v-svoi-servisy-gigachat> (data obrashcheniia: 18.03.2025).**
7. **Maksimova S. V.** Novaia strategiia razvitiia tcifrovoi` integrirovannoi` sistemy` Natsional`noi` biblioteki Respubliki Saha (Iakutiia) // Nauchny`e i tekhnicheskie biblioteki. 2024. № 3. S. 127–141.
8. **Balatckaia N. M., Martirosova M. B.** Kraevedcheskii` veb-arhiv v strukture informatcionny`kh resursov biblioteki: model` i vozmozhnosti realizatsii. Sankt-Peterburg, 2023. 207 s.
9. **Ershova V. V.** Kraevedcheskie resursy` interneta kak ob`ekt veb-arhivirovaniia: opyt realizatsii proekta Yakutia.Online v Natsional`noi` biblioteke Respubliki Saha (Iakutiia) // Kraevedcheskie resursy` seti kak ob`ekt veb-arhivirovaniia i issledovaniia : sb. statei. Sankt-Peterburg, 2024. S. 98–114.
10. **Lowagie H.** From Bias to Transparency : Ethical Imperatives in AI-Based Library Cataloging / Hannes Lowagie // IFLA. 23.09.2023. URL: <https://repository.ifla.org/handle/20.500.14598/2841> (accessed: 29.09.2025).

11. *Iskusstvennyĭ intellekt v deiatel'nosti Nacional'noi biblioteki Respubliki Saha (Iakutiia) / S. S. Klimov, A. G. Burnasheva, I. K. Fedorov i dr. // Bibliotekovedenie. 2025. T. 74, № 4. S. 341–358. DOI 10.25281/0869-608X-2025-74-4-341-358.*

Информация об авторах / Authors

Бурнашева Анна Григорьевна – заведующая Центром цифровизации языкового и культурного наследия Национальной библиотеки Республики Саха (Якутия), Якутск, Российская Федерация
ag.burnasheva@nlrs.ru

Чабыева Яна Семеновна – заведующая Центром интегрированных информационных ресурсов Национальной библиотеки Республики Саха (Якутия), Якутск, Российская Федерация
ys.chabieva@nlrs.ru

Баторов Афанасий Раднаевич – канд. техн. наук, доцент, заведующий кафедрой информатики Арктического государственного института культуры и искусств, Якутск, Российская Федерация
batorov@mail.ru

Максимова Саргылана Васильевна – канд. ист. наук, доцент, директор Национальной библиотеки Республики Саха (Якутия), Якутск, Российская Федерация
maksimova@nlrs.ru

Anna G. Burnasheva – Head, Center for Language and Cultural Heritage Digitization, National Library of the Republic of Sakha (Yakutia), Yakutsk, Russian Federation
ag.burnasheva@nlrs.ru

Yana S. Chabieva – Head, Center for Integrated Information Resources, National Library of the Republic of Sakha (Yakutia), Yakutsk, Russian Federation
ys.chabieva@nlrs.ru

Afanasy R. Batorov – Cand. Sc. (Engineering), Associate Professor, Head, Chair for Information Science, Arctic State Institute of Culture and Arts, Yakutsk, Russian Federation
batorov@mail.ru

Sargylana V. Maksimova – Cand. Sc. (History), Associate Professor, Director, National Library of the Republic of Sakha (Yakutia), Yakutsk, Russian Federation
maksimova@nlrs.ru