

Протокол № 103

Заседания информационно-библиографической группы «Каталогизация основных фондов»

от 7 октября 2026 г.

Присутствовали:

1. ТОУНБ им. А. М. Горького: Мян Н. А., Просолова Ю. В., Маркина А. А., Савина Н. В., Бакотина О. Ф.
2. ЦДСЧ им. А.С. Пушкина: Сергеева Т. Ю.
3. МАУ «МБС г. Твери»: Меликова Н. Е.
4. ТвГТУ: Виноградов Д. В.
5. ОНМБ: Володина М. А.

Повестка дня:

1. Подведение итогов 1-го этапа эксперимента по тестированию модуля «ИИ-каталогизация» Глобальной платформы ИИ-решений для библиотек (ГПИИРБ)

Выступили:

1. Мян Н. А. (зав. отделом ОКиК ТОУНБ им. А. М. Горького)
2. Просолова Ю. В. (зав. сектором ИБО ТОУНБ им. А. М. Горького)

Мян Н. А. (зав. отделом ОКиК ТОУНБ им. А. М. Горького)

Областная библиотека приступила к 1-му этапу эксперимента внедрение и тестирование модулей «ИИ – каталогизация» и «Веб-архивирование».

Каталогизаторы Тверской областной библиотеки им. А.М. Горького участвуют в экспериментальном проекте по созданию и внедрению Глобальной платформы ИИ решений для библиотек Российской Федерации.

ИИ= – каталогизация в модуле «Электронный образ документа» (ЭОД) должен ускорить процесс каталогизации ресурсов, то есть сделать его еще более прозрачным, эффективным и управляемым. И, конечно, снизить количество ошибок.

За основу взят опыт работы Национальной библиотеки Республики Саха (Якутия). С марта 2025 г. коллеги тестируют систему пока только на одночастных документах. По прогнозам специалистов, ИИ должен облегчить рутинную работу каталогизаторов, но не заменить их. Проект пилотный, многие вопросы остаются открытыми: ИИ не тестировался на ретроспективной конверсии, платформа еще плохо распознает нечеткий текст старых изданий и др., как раз те моменты, где помощь ИИ необходима нам.

Модуль ЭОД автоматически распознает текст и создает библиографическую запись, включает 4 функционала.

Комплектатор формирует маршрут обработки и выстраивает задачу на создание ЭОД, система автоматически передает данные сканировщику.

Сканировщик создает электронный образ документа. Сканируется не весь документ, а только те элементы оформления издания, которые необходимы для генерации библиографической записи. Это, как правило, обложка, титульная страница, оборот титульного листа, оглавление, аннотация.

После этого в Национальной библиотеке Республики Саха (Якутия) пути печатной копии и ЭОД расходятся. В комплектовании книге присваивается инвентарный номер и она передается в отделы-фондодержатели. Такой технологичный вариант действий возможен при инвентарной расстановке фондов. В Тверском регионе преобладает систематическая расстановка фонда, поэтому печатная копия ресурса и его электронный образ идут вместе до конца обработки.

Сформированный сканировщиком электронный образ документа поступает к каталогизатору. Специалист проверяет, при необходимости редактирует и нажимает кнопку «Запись готова». Система сама проставляет классификационные индексы, сканированные с оборота титульного листа и предметные рубрики через авторитетные файлы. Для создания качественной

библиографической записи очень важно профессиональное взаимодействие ИИ и каталогизатора. Например, нас не устраивают классификационные индексы с оборота титульного листа. Или, на мой взгляд, стандартных отсканированных элементов издания недостаточно для адекватной предметизации. Комплекатор в таком случае должен предугадать «мысли» ИИ и в полном объеме обеспечить его информацией.

Затем, если ресурс оцифрован полностью, контент-менеджер размещает его в Электронной библиотеке.

Вот так выглядит путь книги с использованием искусственного интеллекта. И как мы можем убедиться – это уже реалии дня сегодняшнего, а не фантастика.

Просолова Ю. В. (зав. сектором ИБО ТООУНБ им. А. М. Горького)

Для данного эксперимента было отобрано 10 однотомных изданий из разных отраслей знаний. На все книги мы создали библиографические записи в формате RUSMARC. Затем мы отсканировали все элементы изданий, которые необходимы для составления библиографического описания. Обложка, карман (на случай, если ИИ сможет считать и написать инвентарный номер и отдел библиотеки), шмуцтитул, титульная страница, оборот титульного листа с аннотацией и макетом каталожной карточки, титульная страница на параллельном языке (при наличии), концевая титульная страница и страница, на которой наклеен штрих-код и написана цена. Кроме этого, для некоторых изданий дополнительно сканировались вкладки с иллюстрациями, пронумерованные страницы с разными видами иллюстраций, страницы, которые могли помочь ИИ создать рубрики, например, с датами жизни людей, которым посвящено данное издание и т. д., страницы, на которых есть библиографические списки и подстрочные примечания, указатели, содержание.

На следующем этапе мы загрузили электронные образцы документов (PDF-файлы) через личный кабинет участника глобальной платформы ИИ-решений для библиотек. На основе загруженных образцов модуль автоматически сформировал RUSMARC-записи, которые мы скачали в формате MARCXML.

Далее мы провели сравнительный анализ наших записей и записей, сделанных ИИ.

В 010 поле ИИ нигде не заполнил подполе \$b (в пер., в обл.), \$d (цена), но везде написал ISBN и тираж.

В 100 поле в позиции «набор символов G0» (позиция 26, длина 2) стоит цифра 50. В нашей библиотеке принято решение использовать это значение (50) для заимствованных записей. Код целевого назначения ИИ определял «u» (неизвестно), даже когда на концевой тит. с. было указано, что книга, например, для среднего и старшего школьного возраста. В таком случае код целевого назначения (позиция 17, длина 1) должен быть «d» (для детей среднего возраста). Код целевого назначения (позиция 18, длина 1) должен быть «e» (для юношества).

101, 102 всегда были заполнены верно, даже для переводных изданий.

В 105 поле не распознавал виды иллюстраций (портреты, карты и т. д.), ставил общее обозначение «a» - иллюстрации. Но были случаи, когда проставлял в подполе «коды иллюстраций» (позиция 0, длина 1), (позиция 1, длина 1), (позиция 2, длина 1), (позиция 3, длина 1) «элемент данных не используется», хотя иллюстрации были отсканированы.

В позициях «коды конференции», «индикатор юбилейного издания», «индикатор указателя» при отсутствии этой информации ИИ выбирал значение 0. Если издание публикуется не от имени конференции, не является юбилейным, не имеет указателя, в нашей библиотеке принято решение проставлять значение | (элемент данных не используется).

В позиции «код литературного жанра» ИИ ставил то значение «u» (нелитературный текст), то |. Художественную литературу определял верно. В нашей библиотеке, если издание не является художественной литературой, мы применяем «элемент данных не используется».

В позиции «код биографии» ИИ ставил «u» (не биография), если издание не является биографическим. При отсутствии каких-либо биографических данных, в нашей библиотеке принято решение ставить значение | (элемент данных не используется).

106 поле (обычная печать) ИИ не заполнял, в нашей библиотеке это поле заполняется автоматически.

181, 182, 203 всегда были заполнены верно. В нашей библиотеке при описании книг эти поля

не используются.

В 200 поле под \$e не указывал целевое назначение. В основном не писал вторичные сведения об ответственности, а когда писал, то брал их из макета каталожной карточки в том виде, в котором они были представлены, например, в сокращенном виде.

210 поле заполнял верно.

Информацию об объеме (\$a 215 поле) скорее всего, брал из макета. Помимо этого, в квадратных скобках указывал еще страницы. Возможно, что ИИ считал переплет, форзацные и нахзацные страницы. Размер обычно проставлял верно.

Если в 225 поле надо было написать дополнительные сведения, например, сведения, относящиеся к заглавию, редактора, выпуск, ИИ ошибался. Сведения, относящиеся к заглавию писал с заглавной буквы, редактора не указывал, выпуск указал под \$d (параллельное заглавие серии) вместо \$v.

Библиографию (320 поле) указывал, но не всегда корректно. Например, мог не написать количество источников.

Не заполнял 327 и 464 поля.

330 поле (аннотация) заполнял всегда, если информация уже была в книге.

333 поле, в котором мы указываем возрастные ограничения, никогда не заполняет.

Не заполнил 510 поле (параллельное заглавие), хотя в 200 указал эту информацию под \$d.

Предметные рубрики ИИ не формирует.

Заполнял 675, 676 поля, но мы не используем эти классификации.

686 поле заполнял, брал информацию с оборота титульного листа.

При заполнении 700 поля не берет информацию из авторитетного файла. Заполнено с ошибками. Под \$b указаны не инициалы, а расширение инициалов. \$g нет.

При заполнении 702 полей также не использует авторитетные файлы. Заполняет с ошибками. Под \$b указывает не инициалы, а расширение инициалов. \$g нет.

Скорее всего, берет информацию о вторичной интеллектуальной ответственности с концевой титульной страницы, а в 200 поле под \$g эти сведения не указывает.

Не заполняет 712 поля, \$g (последующие сведения об ответственности) не пишет и в 200 поле, хотя на тит. с. все сведения есть.

Решения:

Провести следующее заседание группы КОРБИС в ноябре 2026 г. в ТОУНБ им. А. М. Горького.

_____/Мян Н.А./