

Современные технологии оцифровки и практика реализованных проектов

Корпорация ЭЛАР

Григорий Козлов,
ведущий эксперт по оцифровке и сканерам ЭЛАР



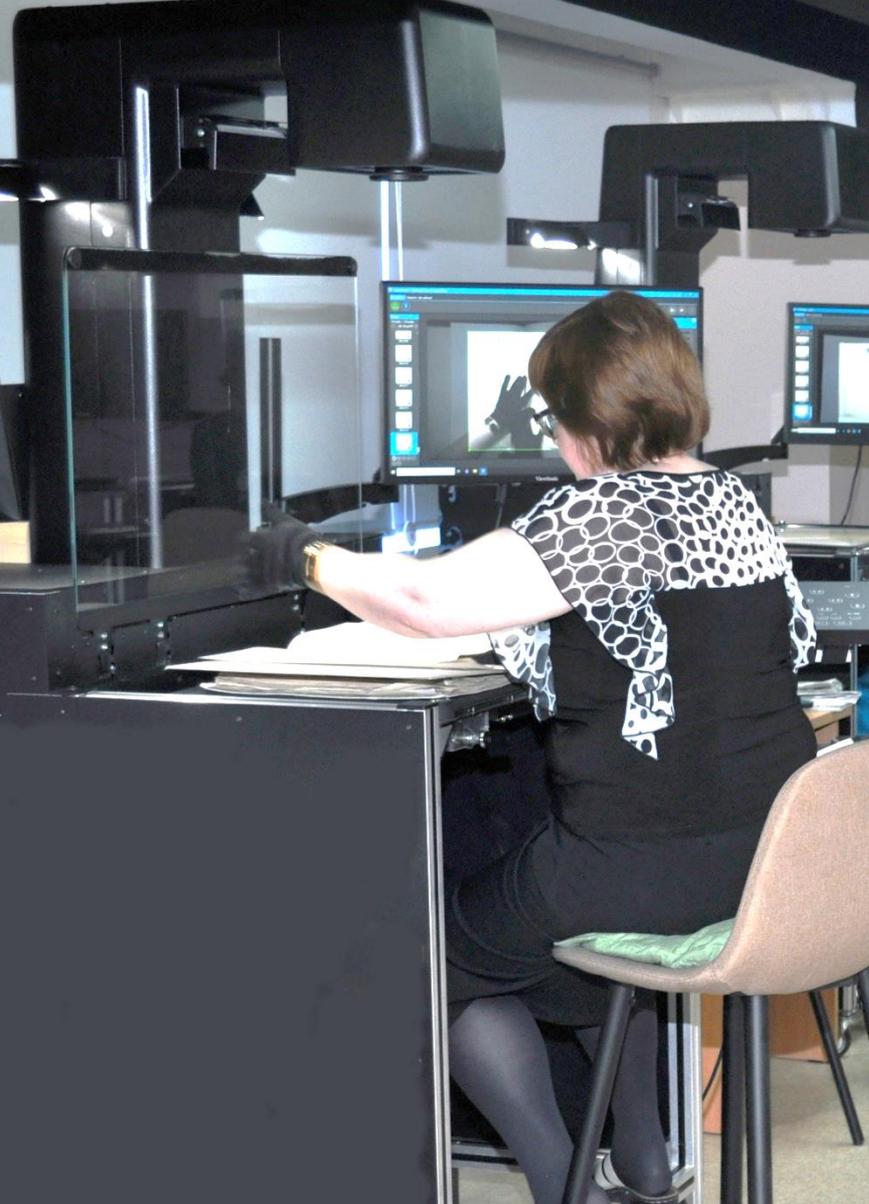


Важные вопросы цифровизации научно-технических библиотек

- ◆ **Большой научный потенциал еще в бумажном виде**
- ◆ **Обеспечение сохранности и долговременного хранения цифровой информации**
- ◆ **Эффективное использование знаний**

Перечень поручений Президента РФ Пр-616 п.27

«Создание цифровых центров на базе научно-технических библиотек и оцифровка фондов в научных и образовательных организациях»



Оцифровка - ключевой этап

В России проводится
регулярная оцифровка
библиотечных фондов

*Сформулированы
методические
рекомендации*

Оцифровка бумажных фондов

Соблюдение требований для подготовки
качественного ресурса

Мастер-копия не ниже 600 dpi

Формат: TIFF

Пользовательская копия не ниже 300 dpi

Формат: JPEG, PDF

Просмотровая копия - 150-300 dpi

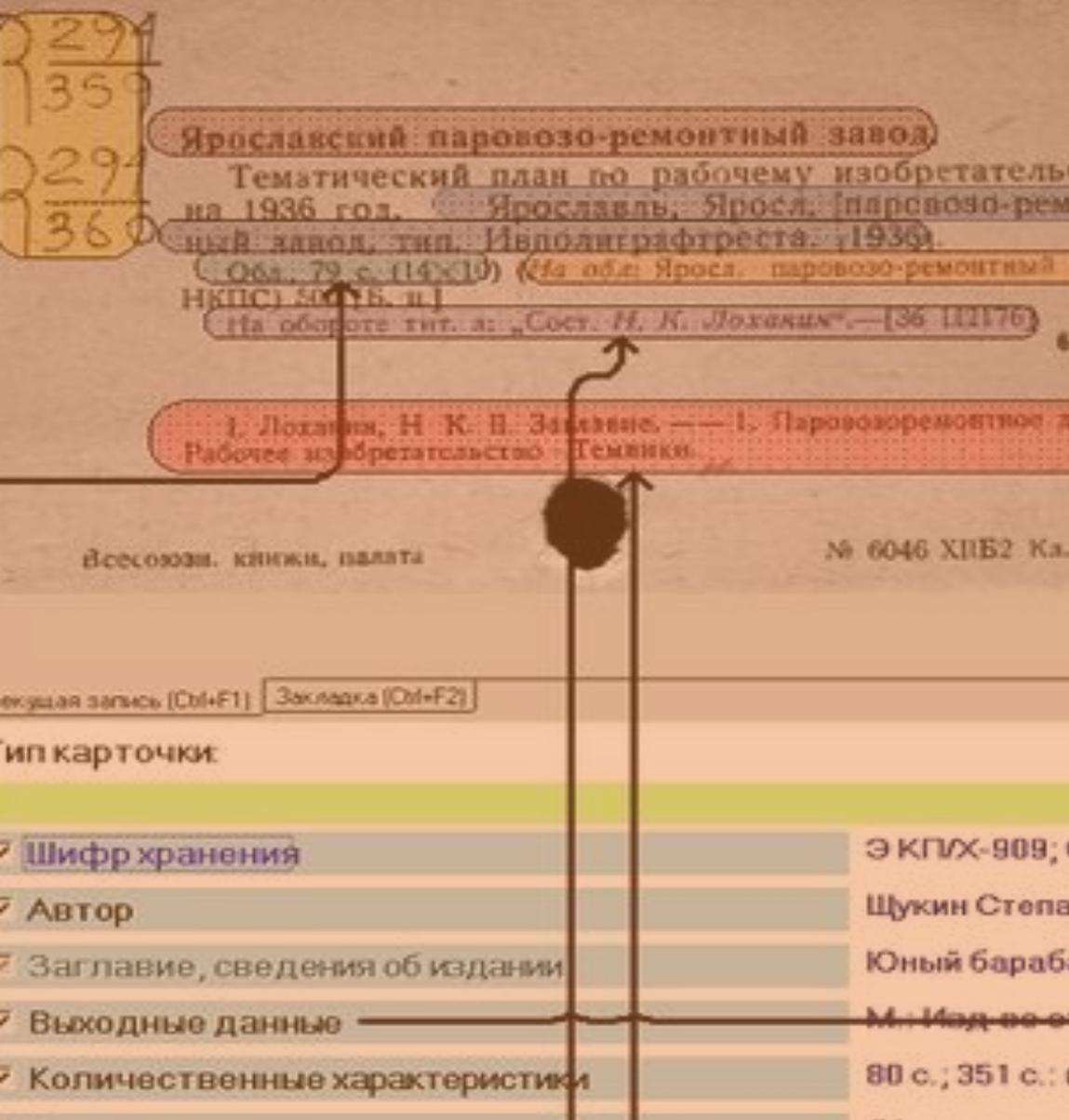
Формат: PDF





Электронный каталог

Применяется
промышленный способ
индексирования
библиотечных каталогов
с использованием
конвейерной технологии



ДЕДУПЛИКАЦИЯ

технология на основе искусственного интеллекта, позволяющая автоматически нормировать и исключать дублирование библиографических записей

Применение дедупликации в рамках проекта НЭБ обеспечило формирование сводного и достоверного источника информации и оптимизацию поисковых возможностей портала

Для оцифровки библиотечных фондов используются профессиональные сканеры



Сканеры
формата А2-А3



Сканеры с глубокой
колыбелью



Сканеры
формата А1-А0



Сканеры
формата А0+

Сканирующее оборудование ЭЛАР признано профессионалами архивной и библиотечной отрасли в России и за рубежом

Сканеры ЭЛАР в реестре РЭП Минпромторга

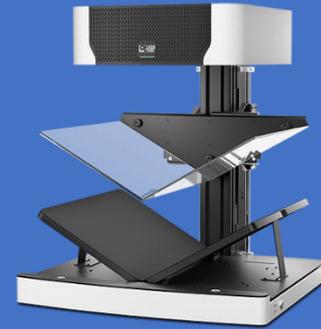
Основные линейки сканирующего оборудования ЭЛАР

ЭларСкан (А3-А0)



Сканирующая система на основе мощной промышленной матрицы

ЭЛАР ПланСкан (А3-А1)



Сканирующая система с линейным сенсором высокого разрешения

Сканирующие системы собственной разработки

Планетарные сканеры ЭЛАР

Включены в реестр РЭП Минпромторга



Универсальное семейство ЭларСкан

оцифровка архивных фондов, крупноформатных материалов и толстых подшивок

Форматы A0, A1, A2 и A3

Матричный сенсор



Интерактивный зум

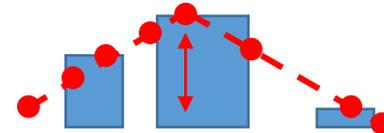


Новое поколение сканеров ПланСкан

высококачественная оцифровка библиотечных фондов

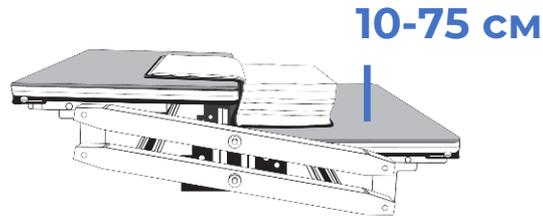
Форматы A2 и A1

Линейный сенсор

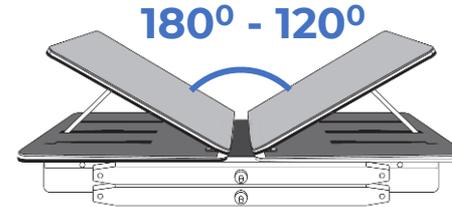


Следящий автофокус

Безопасная оцифровка



Размещение толстых книг для удобного сканирования



Сканирование плотно сшитых и ветхих материалов под углом

Специальные книжные колыбели с подстройкой под особенности оригинала

Режим Антиблик



позволяет получать идеальные копии глянцевых, ламинированных и других бликующих оригиналов.

удаление бликов и засветов осуществляется автоматически прямо в момент сканирования



Адаптивная система освещения и программные алгоритмы позволяют сканировать «сложные» документы

rs101000968652	116	планскан
rs101000907179	278	планскан
rs101001863547	632	планскан
rs101002369903	284	планскан
rs101001263724	112	планскан
rs101002394236	128	планскан
rs101000688073	328	планскан
rs101001249680	488	планскан
rs101001226276-	270	планскан
Пересканы		планскан
		планскан

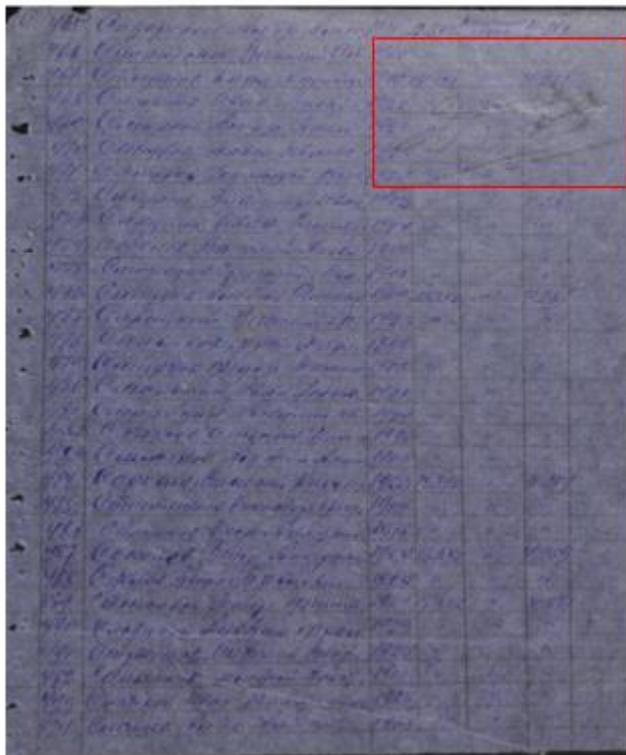
Офисный сканер

rs101000968652	116	планскан
rs101000907179	278	планскан
rs101001863547	632	планскан
rs101002369903	284	планскан
rs101001263724	112	планскан
rs101002394236	128	планскан
rs101000688073	328	планскан
rs101001249680	488	планскан
rs101001226276-	270	планскан
Пересканы	3846	планскан
		планскан

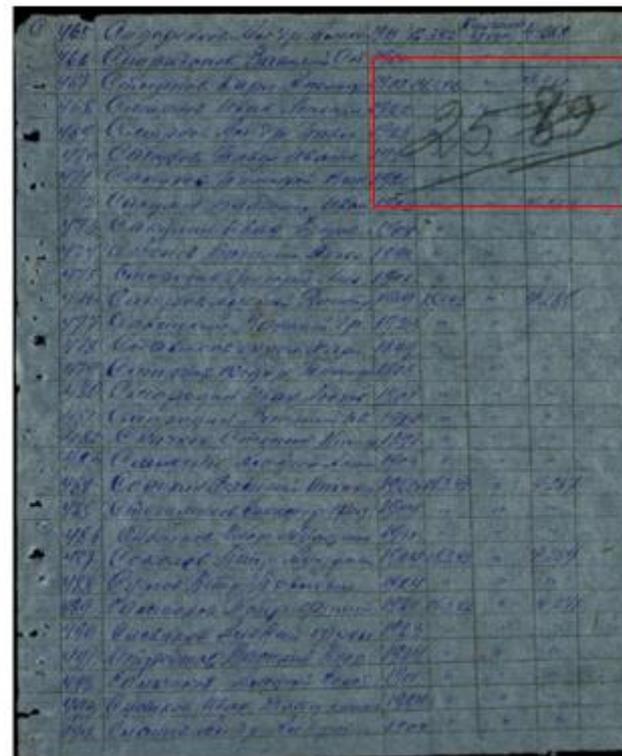
Планетарный сканер

Проявление угасающего текста

Обычное сканирование



Профессиональное сканирование



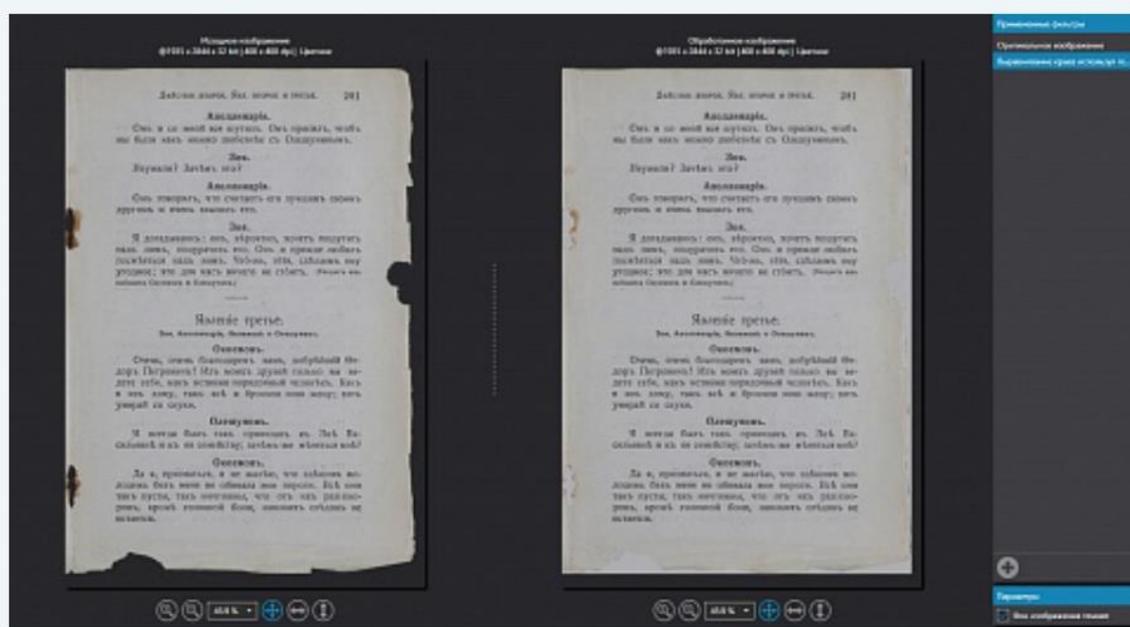
Встроенное профессиональное ПО «СканИмидж» с использованием ИИ на этапе оцифровки

Профессиональная обработка электронных копий может выполняться при сканировании

Обрезка, удаление фона, мусора

Восстановление границ

Коррекция яркости, контраста и др.



Выравнивание страниц и текста

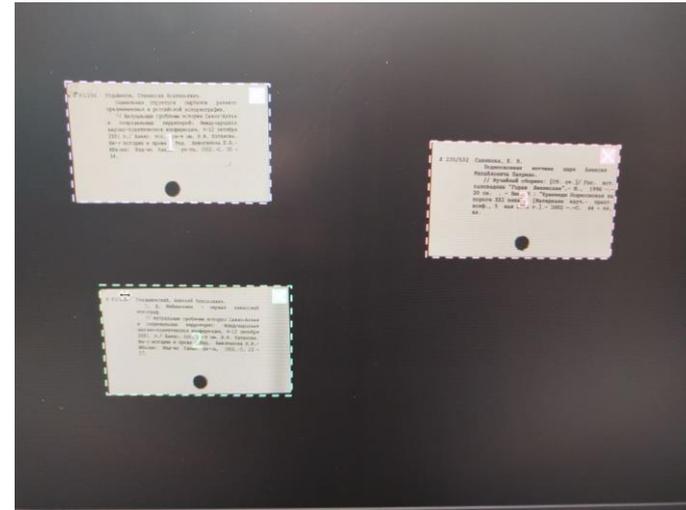
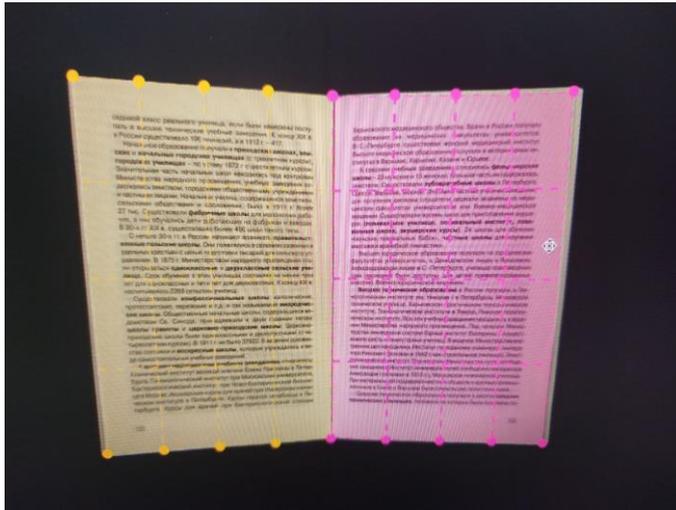
Распознавание текста

Распознавание пустых страниц, штрих-кода и др.

Более 50 интеллектуальных алгоритмов для обработки и коррекции

Развитие ИИ в ПО «СканИмидж»: ИИ-ассистенты для упрощения оцифровки

Интеллектуальный режим с автоматической настройкой алгоритмов обработки под особенности оригинала



Применение ИИ в несколько раз снижает трудозатраты оператора при сканировании

Роботизированные архивы ЭЛАРобот НСМ

В реестре РЭП Минпромторга

50+ лет срок хранения

Защита от любых угроз

Онлайн / офлайн хранение

Низкая стоимость хранения

для ЦОД



← 1036 ТБ

для библиотеки / ВУЗа

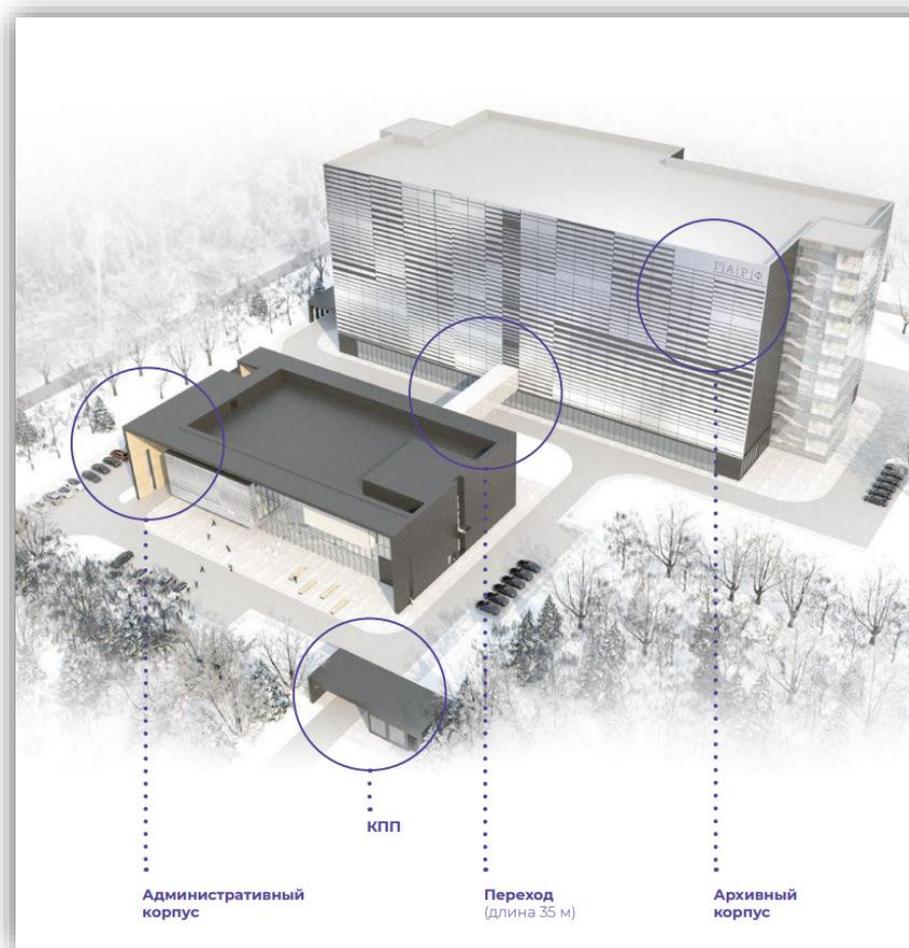


← до 56 ТБ

Оптика – стандарт долговременного хранения

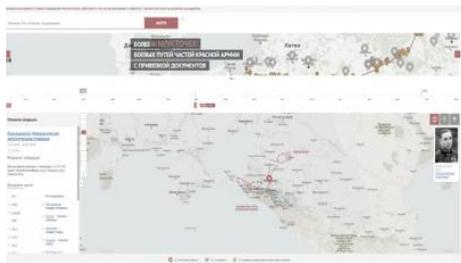
Единственная технология, гарантирующая
долговременное хранение любых данных

Росархив выбрал оптику
в качестве основной
технологии для ЦХОД ГА
РФ (г. Обнинск),
созданного
для приема на
постоянное хранение
электронных архивных
документов органов
власти





Национальные проекты – научный подход ЭЛАР



Министерство обороны РФ

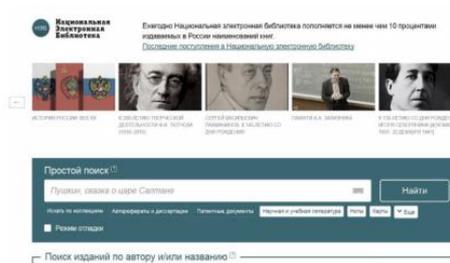
Крупнейший открытый онлайн-ресурс о гeroях Второй Мировой войны

- > 160 000 000 оцифрованных документов,
- > 100 000 исторических карт
- > 12 000 000 координат событий



Министерство обороны РФ Федеральное архивное агентство Российское историческое общество

Банк оцифрованных документов и данных о гeroях Первой мировой войны



Министерство культуры РФ Российская государственная библиотека

Единая технологическая платформа и доступ к объединенным ресурсам более 6 000 публичных библиотек России



Создан один из крупнейших геосервисов с слоями из исторических карт, графикой боевых действий и привязками архивных документов

166,8 млн

оцифрованных страниц архивных документов

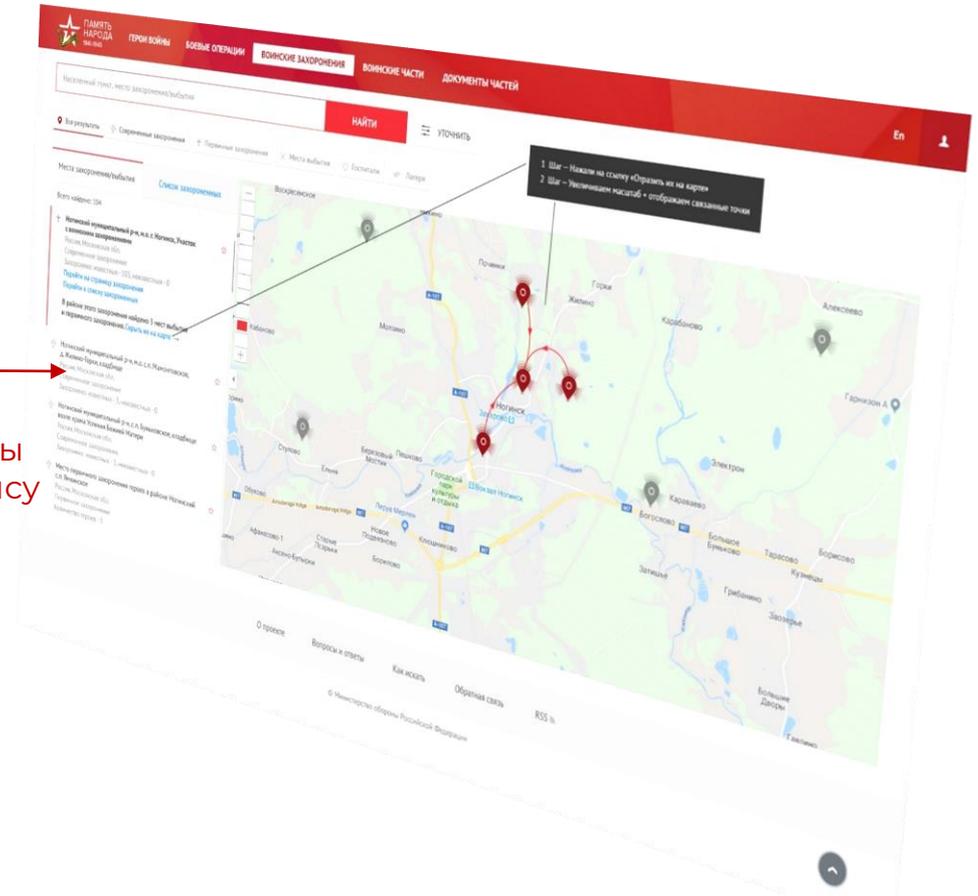
200 млн

записей на участников ВОВ

100 тыс.

исторических карт и схем боевых действий

Данные привязаны к геосервису



ИИ для использования знаний

создание цифровых сервисов

Генеративный ИИ

большие
возможности поиска

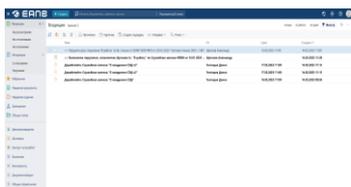
Машинное обучение

аналитика
заданных данных

Научный поиск

термины и
области знаний

ЕАПО



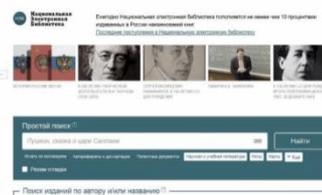
Патентный поиск

Память Народа



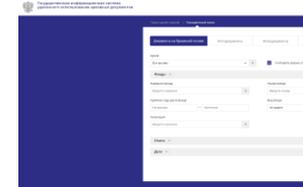
Геосервис

НЭБ



Дедупликация

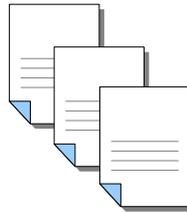
ГИС УИАД



Планы развития

Структурированные запросы

Поиск по базе данных после распознавания

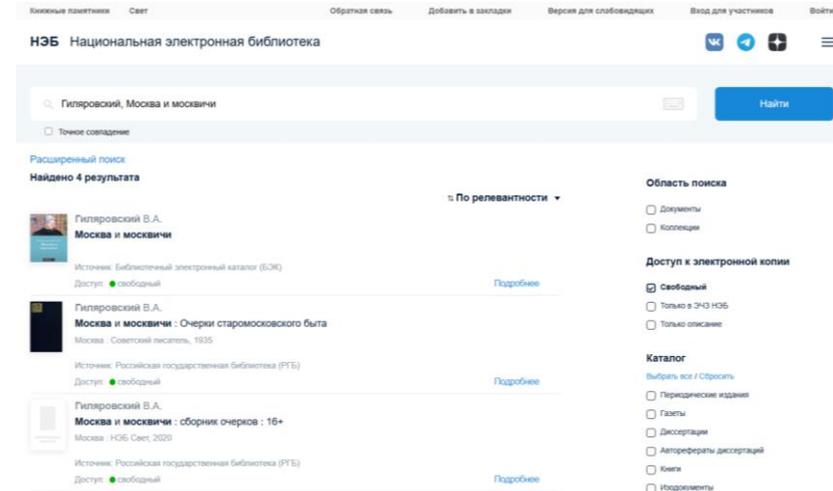


Структурированные запросы



Название, автор, ключевые слова...

(справочники, реквизиты, полнотекстовый поиск)



Классические заметки Счет Обратная связь Добавить в закладки Версия для слабовидящих Вход для участников Войти

НЭБ Национальная электронная библиотека

Гиларовский, Москва и москвичи

Точное совпадение

Найти

Расширенный поиск

Найдено 4 результата

По релевантности

Область поиска

- Документы
- Коллекции

Доступ к электронной копии

- Свободный
- Только в ЭЧЗ НЭБ
- Только описание

Каталог

Выбрать все / Сбросить

- Периодические издания
- Газеты
- Диссертации
- Авторефераты диссертаций
- Книжки
- Издания

Гиларовский В.А.
Москва и москвичи
Источник: Библиотечный электронный каталог (БЭК)
Доступ: свободный

Гиларовский В.А.
Москва и москвичи : Очерки старомосковского быта
Москва : Советский писатель, 1935
Источник: Российская государственная библиотека (РГБ)
Доступ: свободный

Гиларовский В.А.
Москва и москвичи : сборник очерков : 16+
Москва : НЭБ Санкт, 2020
Источник: Российская государственная библиотека (РГБ)
Доступ: свободный

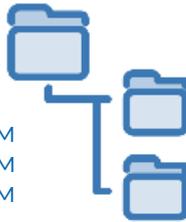
Удобный и быстрый поиск по электронной базе данных на примере НЭБ

Аналитические запросы

Поиск по нераспознанным изображениям



Поиск по файлам и файловым хранилищам



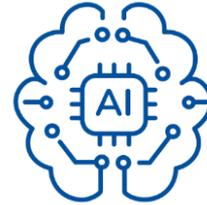
Аналитические запросы



Найди все изделия, которые... и создай таблицу...

Опиши кратко цели данного исследования...

(работа в чате ИИ по областям знаний)



Итог

Решение на базе ЭЛАР Контекст + ЭЛАРобот НСМ — это единая, безопасная, масштабируемая и суверенная экосистема для управления документами на всех этапах их жизненного цикла: от оперативного использования до бессрочного архивного хранения. Оно обеспечивает **максимальную сохранность, доступность и юридическую значимость** информации в условиях цифровой трансформации и геополитической нестабильности.



Редизайнирование изображений веб-разработка Генерация изображений Генерация видео Артефакты

Чем я могу помочь вам сегодня?

+ Мессенджер Поиск

Обработка архива научно-технической документации

Анализ документов. Саммаризация.
Список источников



Национальная электронная библиотека

2013 год

РАЗРАБОТКА МАКЕТА

2014 год

ОСНОВНАЯ РАЗРАБОТКА

Портал и интеллектуальная
система поиска

2015 год

РАСШИРЕНИЕ

В НЭБ вошли
67 новых участника

2017 год

РОСТ

4,1 млн. фондов
и 39,1 млн. записей

2016 год

ПЛАНОМЕРНОЕ РАЗВИТИЕ

Налажена типовая
интеграция с АБИС

2018 год

ПОПУЛЯРИЗАЦИЯ НЭБ

Портал НЭБ

1С-Битрикс
отечественное ПО

Разработка ЭЛАР
с 2014-2017 гг.

Отечественная
интеллектуальная
поисковая система

Разработка ЭЛАР
контракт 2017 г.

Технологическая
платформа

Разработка ЭЛАР
с 2015-2017 гг.

Функционал системы
полностью перенесен
на новую платформу,
код программы
переписан под новую
систему на языке JAVA
Разработка ЭЛАР
2017 г.

Издания

НЭБ – это классическая библиотека



Интернет-портал

Обеспечивает свободный доступ к культурному наследию



Электронные читальные залы

Через 4 589 читальных залов обеспечивается доступ к материалам ограниченным авторским правом, в том числе к диссертациям, научной и образовательной литературе

НЭБ идейно абсолютно такая же библиотека, как и другие мировые аналоги: Gallica, EUROPEANA, Библиотека Конгресса и остальные

Сегодня

За годы развития НЭБ
стала одной
из лучших электронных
библиотек в мире



Умный поиск

Поисковые алгоритмы позволяют пользователю максимально быстро и удобно находить необходимую информацию



Технологическая база

Типовая платформа для подключения библиотек, интеграция с АБИС, доступ к закрытым изданиям из ЭЧЗ, загрузка библиотечной информации и копий



Дополнительные сервисы

Чтение и работа с электронными копиями, система личных кабинетов, сервисы поиска книг в ближайшей библиотеке, мобильные приложения

Удобный поиск и чтение произведений

Промышленная система поиска

Высокая релевантность

Дедупликация изданий

Нечеткий поиск, подсказки и синонимы

Расширенный и многоэтапный поиск

В НЭБ сделано все, чтобы читатели за несколько секунд находили нужное произведение

Gallica

Онлайн-библиотека «Галлика» 4,3 млн. документов

Насколько удобно искать на ведущем французском ресурсе?

Поиск одной строкой по отрывку из названия: Un billet... (Лотерейный билет № 9672, Жюль Верн)

Искомое по первому запросу не найдено

The screenshot shows the Gallica search interface. At the top, the search bar contains the query "Un billet" and the results are sorted by "Relevance". The search results are displayed in a list format, showing three items:

- 1** **Billet (d'admission de Mademoiselle Forgeot comme élève) (manuscrit autographe)**
Detailed information
- 2** **Billet [d']1 milliard de marks : [photographie de presse] / [Agence Rol] 1923**
Detailed information
- 3** **Collection de manuscrits d'Auguste Rondel. Série MRo : musique, music-hall, danse, mime, cirque, marionnette, fêtes et carnivals. Section MRo. Sous-section Cirque. Artistes de cirque. F-H. Francesco Hidalgo : billet**
Detailed information

The left sidebar contains filters for "Consultation site" and "Type of document". The "Consultation site" filter shows options like Gallica (135 265), e-rara.ch (332), and others. The "Type of document" filter shows options like Books (119 787) and Press and magazines (14 715).

Аналогичный поиск в НЭБ

Поиск одной строкой: Un billet...

Уже в первой выдаче искомое произведение найдено

НЭБ Национальная электронная библиотека



Un billet



Найти

Точное совпадение

[Расширенный поиск](#)

Найдено 2 результата

По релевантности ▾



Verne J.

Un billet de loterie : (le numero 9672) : suivi de Frritt - Flacc

J. Hetzel et Cie, 1886

Источник: Российская государственная библиотека (РГБ)

Доступ: ● свободный

[Подробнее](#)



Verne J.

Un billet de loterie (le numéro 9672). : Suivi de Frritt-Flacc

Paris : Hetzel, 1886

Источник: Российская государственная библиотека (РГБ)

Доступ: ● свободный

[Подробнее](#)

Область поиска

- Документы
- Коллекции

Доступ к электронной копии

- Свободный
- Только в ЭЧЗ НЭБ
- Только описание

Каталог

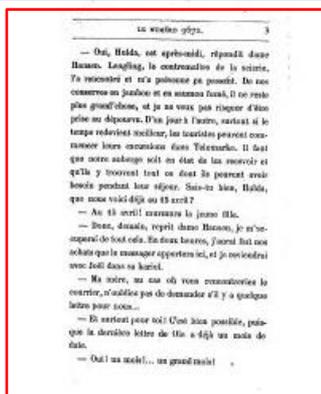
[Выбрать все](#) / [Сбросить](#)

- Периодические издания
- Газеты

Работа с изданиями

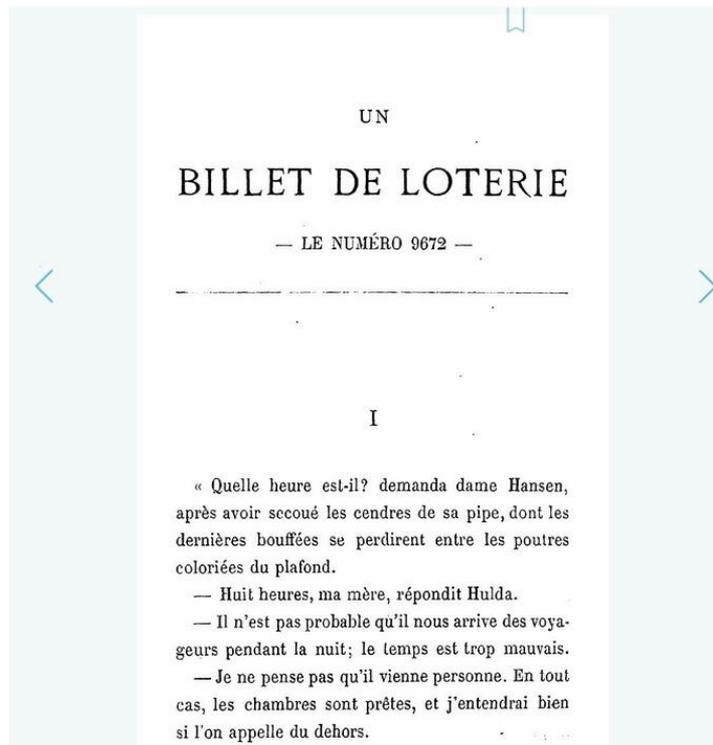
EUROPEANA

(читать с портала затруднительно)



НЭБ

(просмотровщик адаптирован для чтения онлайн)



НЭБ – виртуальная библиотека

Портал предоставляет читателям возможность не выходя из дома знакомиться с научно-образовательной литературой

Задача НЭБ давать пользователям произведения в их оригинальном формате, приобщать к чтению, давать возможности читать в режиме нон-стоп: от произведения к произведению



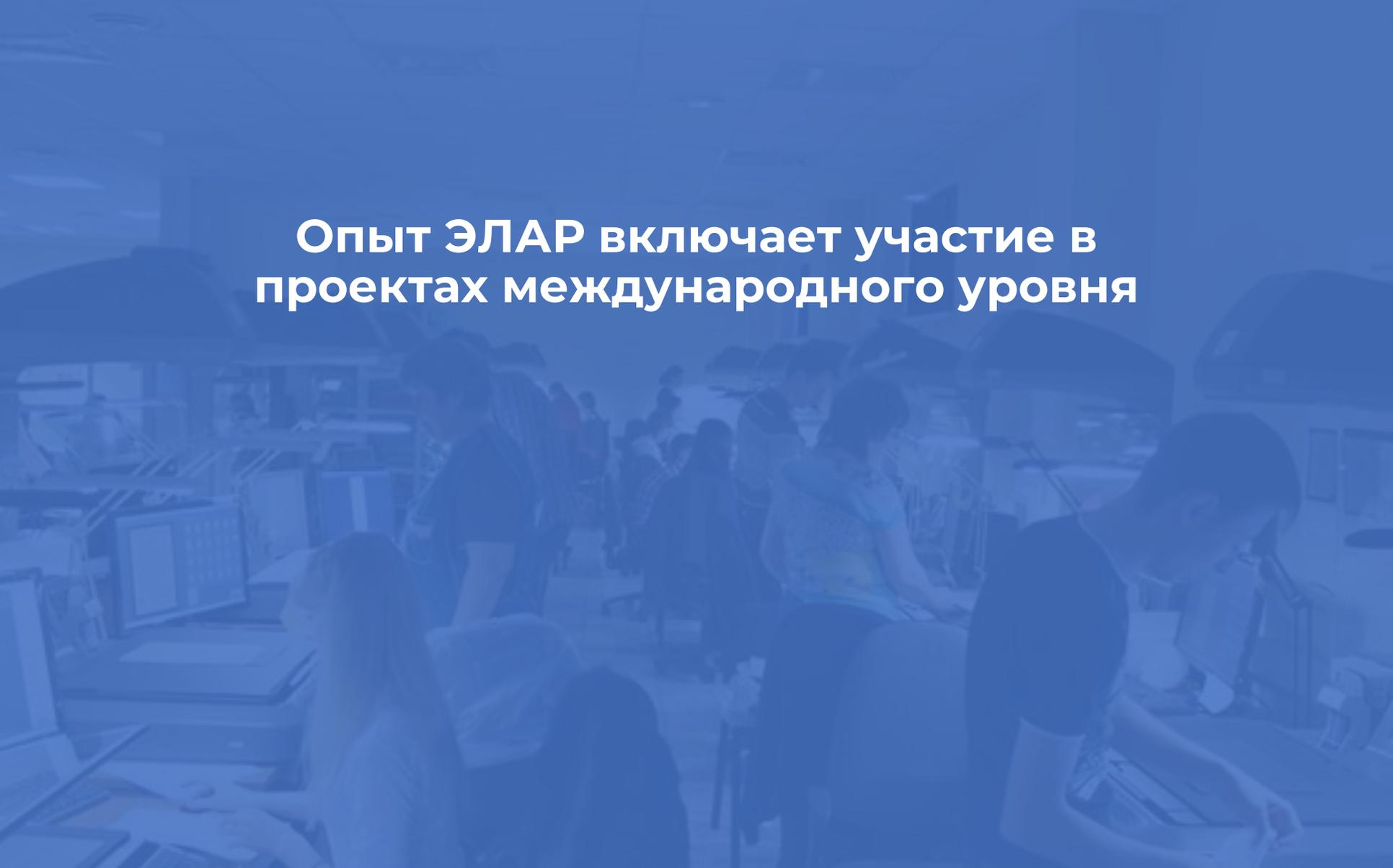
Лучшие коллекции и подборки



Широкие поисковые возможности для разных категорий пользователей



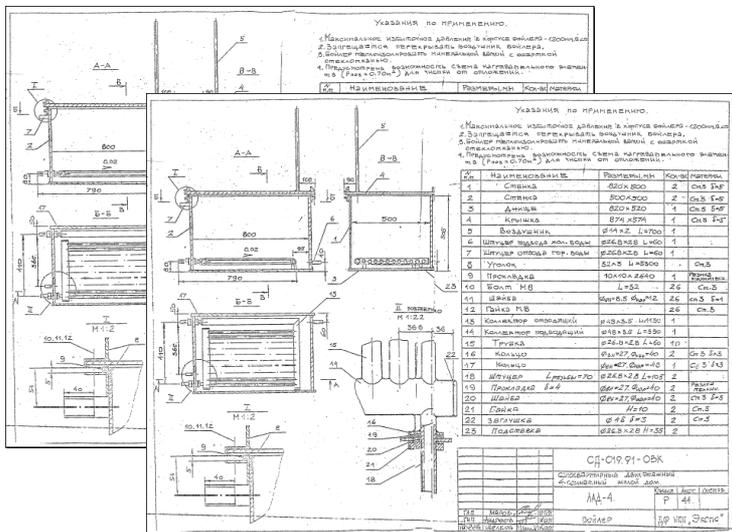
Геосервисы для поиска печатных изданий и библиотек



Опыт ЭЛАР включает участие в проектах международного уровня

Проект для МАГАТЭ

- Международный проект пакетной обработки документов на микрофильмах.
- Общий объём - 20 млн страниц формата А4 - 400 000 документов
- Разработка и внедрение централизованной электронной библиотеки



И другие проекты международного уровня:

- | **ELIBRON:** Оснащение и сопровождение центра массового ввода книг в рамках совместного проекта создания крупнейшей Интернет–библиотеки (плановый объем — 1 млн. изданий)
- | **Международные гуманитарные проекты (МВД/МИД/МинЮст РФ/ Швеция, ФСБ РФ/ НКК)**
Оцифровка и каталогизация документов спецархивов 30–40х гг.
- | **Оцифровка для Германского Красного Креста о военнопленных немцах.** Создание электронной базы данных по таким полям как: номер фонда, описи, дела, ФИО, год и место рождения, номер, дата и место отправления и назначения эшелона, место смерти/захоронения.
- | **Международный библиотечный проект "Встреча на границах"** Более 2 500 материалов и около 70 000 изображений из коллекций редких книг, рукописей, карт и звукозаписей.
- | **Создание Электронного архива Коминтерна**

Корпорация ЭЛАР

www.elar.ru
office@elar.ru



+7 (495) 274-31-31

