Уважаемые коллеги!

Иркутская областная государственная универсальная научная библиотека им. И. И. Молчанова-Сибирского с 2007 г. занимается формированием и поддержкой электронной краеведческой библиотеки Иркутской области «Хроники Приангарья» (https://i.irklib.ru/hronp/). Это один из самых крупных и известных региональных проектов в России такого направления (более 220 тыс. документов). Значимость «Хроник Приангарья» во многом связана с тем, что мы отдали приоритет в оцифровке местным газетам. И в этом же заключается большая ценность проекта.

Большой вклад в его реализацию вносят муниципальные библиотеки. Почти 25 тыс. документов уже отсканировано и передано в «Хроники Приангарья» нашими коллегами из городов и поселков Иркутской области (Саянска, Шелехова, Черемхово, Усть-Кута, Кутулика, Заларей, Новонукутска, Михайловки, Баяндая, Нижнеилимска, Слюдянки, Зимы и Нижнеудинска). Мы очень благодарны всем участникам проекта, но еще больше признательны им пользователи электронной библиотеки, среди которых есть жители и уроженцы этих населенных пунктов.

Областная библиотека не просто принимает сделанное партнерами. Мы обрабатываем полученные файлы, стандартизируем их, описываем каждый документ (с обязательным указанием библиотеки-источника), а, главное, распознаем изображения страниц и обеспечиваем возможность текстового поиска для всех отсканированных документов.

В 2019 г. ИОГУНБ инициировала масштабный «Большой проект», который должен продемонстрировать важность и нужность библиотек для властей и населения. Одним из подпроектов его является «Электронная память Приангарья». В нем мы в числе других ставим перед собой задачу предоставить людям доступ к максимально возможному числу газет и журналов, издающихся и издававшихся на территории нашей области.

Коллеги! Без вашей помощи мы не сможем собрать этот ценнейший архив во всей полноте или будем делать это еще очень и очень долго. Подключайтесь к проекту – сканируйте ваши местные газеты и отправляйте их нам.

Мы принимаем также летописи поселений, неопубликованные документы, фотографии (атрибутированные) и краеведческие материалы, собранные библиотеками и местными энтузиастами городов и сел Иркутской области. Так же принимаем библиографические пособия (кроме малых форм), подготовленные сотрудниками библиотек.

Все библиотеки, сканирующие документы, считаются участниками «Большого проекта» и по его итогам самые активные будут отмечены!

**Требования к газетным материалам для включения в электронную краеведческую библиотеку Иркутской области (ЭКБ) «Хроники Приангарья»**

1. **Подготовка газеты к оцифровке**

Для включения в ЭКБ принимаются только сканированные материалы. Цифровые изображения страниц, снятые на фотокамеры – только в исключительных случаях[[1]](#footnote-1).

В целях улучшения качества (читаемости, достоверности оптического распознавания) желательно сканирование отдельными номерами (листами), то есть, если газета переплетена – подшивку на время сканирования лучше «расшить». Это позволит избежать потери части текста, ушедшего «в переплет» и неизбежных геометрических искажений, затрудняющих корректное распознавание текста на изображении. Если это по каким-то причинам невозможно (например, подшивка получена на время из архива) нужно максимально постараться избежать искажения результирующего изображения.

1. **Сканирование**

Страницы сканируются по одной, последовательно (в порядке их нумерации или в том порядке, в котором они расположены в номере, если в номере используется раздельная пагинация). Ориентация изображений должна быть портретной (высота больше ширины).

В случае, если блок текста и/или изображение размещены на развороте и приходятся на «сгиб» разделяющий страницы, допустимо сканирование разворота целиком. В этом случае ориентация изображения, естественно, будет альбомной (ширина больше высоты).

При необходимости производится обрезка изображения по границе листа и выравнивание «перекоса».

В случае, если часть листа утрачена, изображение страницы должно соответствовать ее исходному размеру, то есть:

Фрагмент оригинальной страницы

Неверно сканированное изображение

Изображение страницы, отсканированное правильно

*Пример отсканированной и собранной страницы газеты на сканере формата А4*

Оригинальная страница

Неверно сканированное изображение

Изображения страницы отсканированное правильно

1. Верх

2. Низ



*Пример отсканированной и собранной страницы газеты на сканере формата А4*

В случае, если формат страницы превышает формат сканера (например, у вас сканер формата А4, а газетный лист формата А3), допустимо сканирование и передача страниц двумя частями («верх» и «низ»). Мы автоматически объединим изображения. В этом случае ориентация изображений должна быть альбомной, а файлов изображений должно быть ровно в два раза больше, чем страниц и пронумерованы они должны быть в правильной последовательности (т. е. «1» – верх первой страницы, «2» – низ первой страницы, «3» – верх второй страницы, «4» – низ второй страницы и т. д.). Даже если у вас нет половины листа, обязательно должен присутствовать файл с изображением его замещающим.

1. **Технические параметры изображений**

Документы сканируются постранично. Разрешение – 300 dpi. Желательно в формате \*.TIFF, но допустим и \*.JPG. Цвет несущественен, но лучше «Оттенки серого» (grayscale). Обязательно сканировать в цвете страницы с цветной печатью (например с заголовками).

1. **Упорядочение наименований файлов**

Если это отдельные сканы страниц – они должны быть разложены по отдельным папкам и называться так, чтобы при алфавитной сортировке от А до Я первая страница была первой, а последняя – последней. Обычно штатные программы сканера первое изображение не нумеруют, а последующие начинают нумеровать с единицы, то есть получается, например: Изображение.jpg, Изображение001.jpg. В таком случае нужно, чтобы первый файл все-таки назывался Изображение000.jpg, для чего его можно переименовать «вручную».

1. **Названия папок**

Они должны быть однотипны и формализованы следующим образом (на примере газеты «Ленинский путь» за 15 июля 1970 г.):

*lenput\_1970\_001\_\_07\_15*

Т. е. сначала наименование газеты, одно нижнее подчеркивание, год (4 цифры), одно нижнее подчеркивание, номер (ТРИ цифры), ДВА нижних подчеркивания, МЕСЯЦ (ДВЕ цифры), одно нижнее подчеркивание, ДЕНЬ (2 цифры).

Это нужно для быстрой автоматической регистрации в базе данных, а также для сортировки по дате в результатах поиска и поиска по дате. Если номер выпуска сдвоенный, например 9/10, то пишем через дефис, но оба номера с ведущими нулями (если это необходимо):

*lenput\_1980\_009-010\_\_\*\*\_\*\**

Если номер повторяется (бывает такое, ошибочно в типографии поставили дублет), то просто в скобках добавляем двойку. Т.е. есть № 8, скажем, за 10 марта и опять № 8, но за 17 марта. Папки (или файлы) будут соответственно называться:

*leput\_1980\_008\_\_03\_10*

и

*lenput\_1980\_008(2)\_\_03\_17*

Передавать лучше логическими блоками, допустим, сразу за один год.

Способ передачи – загрузка в любой облачный сервис (Яндекс.Диск, облако Майл.ру, Google Drive и т. п.)

**ИОГУНБ осуществляет следующую обработку присланных материалов:**

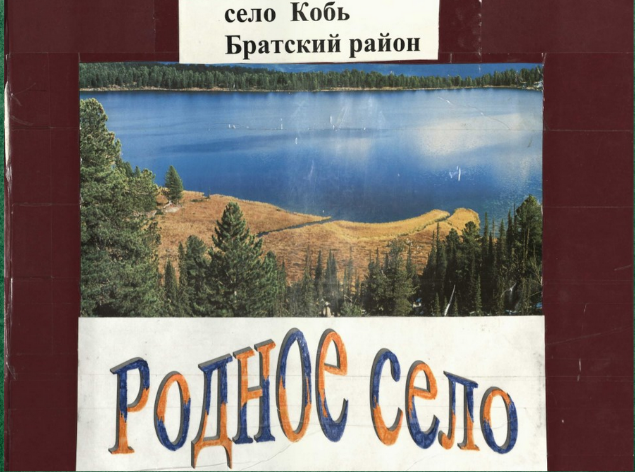
1. Собирает страницы номера в один файл.
2. Осуществляет оптическое распознавание текста (без вычитки).
3. Создает библиографические описания на каждый обработанный номер в базе данных.
4. Осуществляет привязку файла цифровой копии к библиографической записи.
5. Включает распознанные документы в полнотекстовую часть ЭКБ.
6. Если это нужно, передает итоговые pdf создателю сканов.

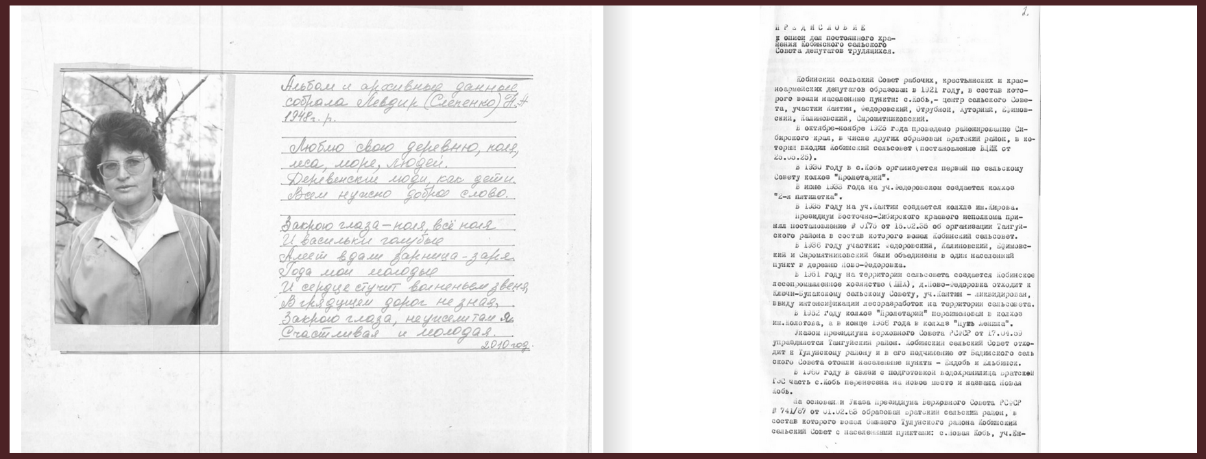
**Требования к «Летописям поселений» для включения в электронную краеведческую библиотеку Иркутской области (ЭКБ) «Хроники Приангарья»**

1. **Подготовка летописи к оцифровке**

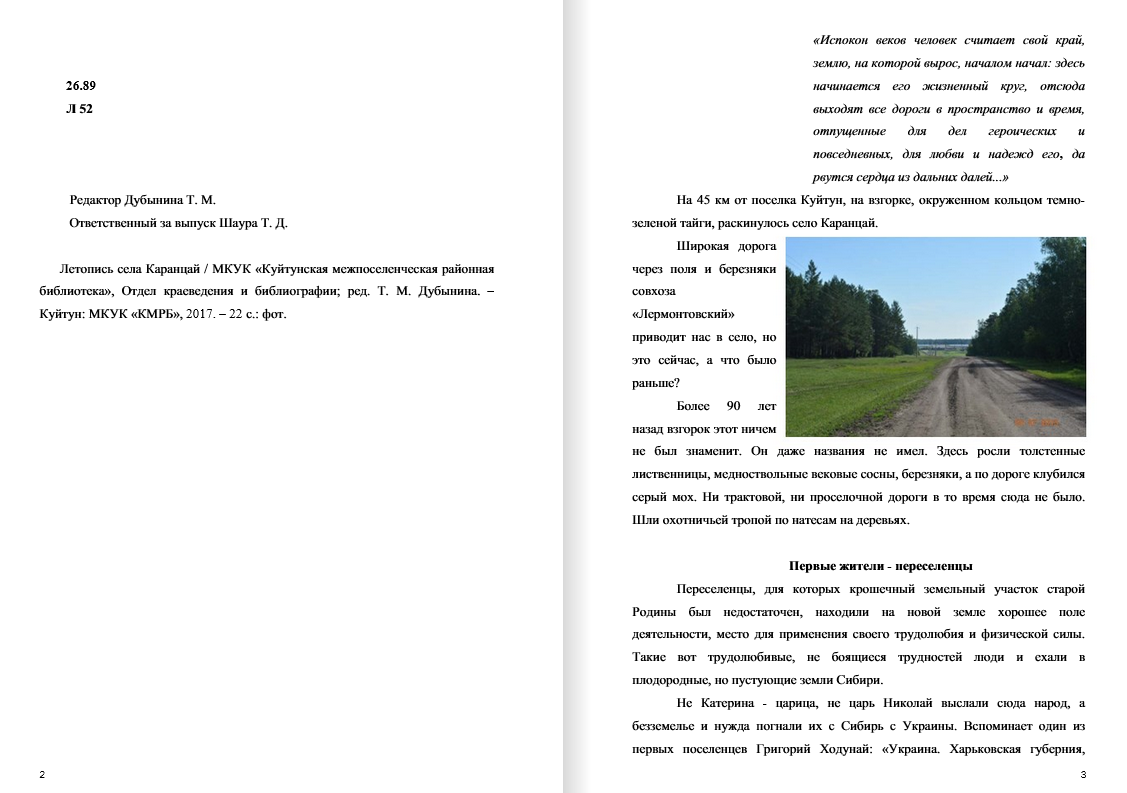
Летописи поселений могут быть:

а) сканированными (если это собранные в советское время материалы или написанные от руки тексты с фотографиями и рисунками, но только не тематические папки с вырезанными газетными статьями – они не являются летописями!);





б) документом в электронном виде (форматы «.doc» и PDF).



Для включения в ЭКБ принимаются сканированные материалы и документы в электронном виде (форматы «.doc» и PDF). Цифровые изображения страниц, снятые на фотокамеры, могут быть приняты только в исключительных случаях. Съемка производилась с использованием штатива, перпендикулярно плоскости листа, при равномерном и достаточном освещении и т. д. Каждый случай рассматривается индивидуально.

В целях улучшения качества (читаемости, достоверности оптического распознавания) желательно сканирование отдельными номерами (листами), то есть, если газета переплетена – подшивку на время сканирования лучше «расшить». Это позволит избежать потери части текста, ушедшего «в переплет» и неизбежных геометрических искажений, затрудняющих корректное распознавание текста на изображении. Если это по каким-то причинам невозможно (например, подшивка получена на время из архива) нужно максимально постараться избежать искажения результирующего изображения.

**2. Технические параметры изображений**

Документы сканируются постранично. Разрешение – 300 dpi. Желательно в формате \*.TIFF, но допустим и \*.JPG. Сканы в цвете.

**3. Упорядочение наименований файлов**

Если это отдельные сканы страниц – они должны быть разложены по отдельным папкам и называться так, чтобы при алфавитной сортировке от А до Я первая страница была первой, а последняя – последней. Обычно штатные программы сканера первое изображение не нумеруют, а последующие начинают нумеровать с единицы. То есть получается, например: Изображение.jpg, Изображение001.jpg. В таком случае нужно чтобы первый файл все-таки назывался Изображение000.jpg для чего его можно переименовать «вручную».

В таком же порядке можно сканировать летописи истории библиотеки, фотоальбомы предприятий и различных учреждений, неопубликованные документы, предварительно проконсультировавшись с нами по почте или телефону.



Отправляйте свои материалы и консультируйтесь по адресу: [**kraeved@irklib.ru**](mailto:kraeved@irklib.ru)

Либо через соцсети, вк: <https://vk.com/id410987300> или fb: <https://www.facebook.com/profile.php?id=100007832253111>

1. Съемка производилась с использованием штатива, перпендикулярно плоскости листа, при равномерном и достаточном освещении и т. д. Каждый случай рассматривается индивидуально. [↑](#footnote-ref-1)