



Технологическая карта процесса оцифровки

1. Общие сведения

Оцифровка - технологический процесс перевода печатного документа в электронный вид.

Цель: организация эффективного процесса оцифровки документов, обеспечивающего высокое качество электронных копий и оперативность обработки.

Задача: стандартизация последовательности действий и временных рамок для минимизации затрат ресурсов и повышения производительности труда сотрудников.

2. Область применения

Технологическая карта предназначена для обеспечения единого подхода к выполнению операций по оцифровке печатных документов в отделе информационных технологий ГБУК «СКУНБ им. Лермонтова».

Технологическая карта определяет организацию труда и методы работы специалистов отдела информационных технологий при приеме/возврате печатных документов на оцифровку (с оцифровки), при сканировании печатных документов, при корректировке сканов печатного документа в графическом редакторе, при загрузке обработанных сканов печатного документа в информационную систему цифрового контента.

3. Техничко-экономические показатели

Продолжительность рабочего дня (рабочей смены) 8 часов, из них технологические перерывы – 1 час.

Непосредственно заняты в процессе оцифровки 4 специалиста. Все работы по оцифровке проводятся в строгом соответствии с техническими требованиями к качеству оцифровки, определенными подразделением-заказчиком в техническом задании на оцифровку, инструкциями по работе

в специализированных программах, должностными инструкциями работников отдела и инструкциями по охране труда.

4. Аппаратно-программные комплексы, используемые при оцифровке

Сканирование печатных документов проводится на следующем оборудовании:

- сканирующий комплекс ЭларПланСкан А2-ц с ПО ВСS-2;
- широкоформатный сканер ПланСкан А2.Х-VII

При корректировке сканов печатных документов используется:

- процессор: 11thGenIntel(R) Core(TM) i5-11400@2.60GHz; оперативная память: 16,0 ГБ (доступно: 15,6 ГБ); тип системы: 64-разрядная операционная система, процессор x64;

- программное обеспечение: операционная система; драйвера для сканеров; программа для оптического распознавания символов, имеющая набор инструментов для создания и редактирования документов формата pdf.

Контроль качества сканов печатных документов проводится на следующем оборудовании:

- процессор: Intel(R) Pentium(R) CPU G340@2.80GHz; оперативная память (ОЗУ): 4,00 ГБ (3,41 ГБ доступно); тип системы: 32-разрядная операционная система, процессор x64; монитор ≥ 19 дюймов;

- устройство для чтения карт памяти;

- источники бесперебойного питания.

- программное обеспечение: операционная система; драйвера для сканеров; программа для оптического распознавания символов, имеющая набор инструментов для создания и редактирования документов формата pdf; программа записи cd/dvd.

5. Оцифровка включает в себя следующие технологические операции:

№	Наименование операции	Содержание работ	Время протекания операции
Этап 1. Прием и подготовка документа			
1.	Отбор документов в отделе хранения основного фонда (ОХОФ)	Изучение фонда, отбор документов, подлежащих оцифровке	1 мин.- 3 мин.
2.	Формирование электронной заявки, отправка заявки по электронной почте	Создание заявки: подготовка сведений с указанием наименования документа, года издания, шифра хранения; отправка заявки по электронной почте в ОХОФ	1 мин.- 6 мин.

3.	Выдача документа из ОХОФ	Поиск документа в ОХОФ; ручное обеспыливание документа с шести сторон; транспортировка документа на кафедру выдачи; подъем на 4-й ярус; отметка в учетных документах	8 мин.30 сек.- 20 мин.30 сек.
4.	Доставка документа сотрудником отдела информационных технологий (ОИТ) на оцифровку	Транспортировка документа сотрудником ОИТ из ОХОФ на 1этаж в ОИТ	5 мин.-7 мин.
5.	Подготовка документа к сканированию: проверка нумерации, расшивка	Проверка правильности нумерации страниц; распаковка и расшивка документа	5 мин.-7 мин.
Этап 2. Сканирование документа			
6.	Запуск оборудования: включение сканера, программы сканирования	Включение ПК; запуск сканера и программы сканирования VCS-2.	1 мин. 15 сек.- 3 мин.
7.	Установка формата и настроек в соответствии с техническим заданием (ТЗ)	Ожидание полной загрузки программы сканирования VCS-2; установка формата с учетом ТЗ в программе сканирования VCS-2	42 сек.- 1 мин. 20 сек.
8.	Сканирование (газета формата А3, объемом 2 стр.) с контролем качества	Размещение документа на сканировочном столе; расправление страниц; укладка прижимного стекла на документ; сканирование; подъем стекла после сканирования 1-й страницы документа; переворот на 2-ю страницу документа; укладка прижимного стекла на документ; сканирование; подъем стекла после сканирования; перемещение документа; осуществление первичного контроля сканов в прикладной программе VCS-2	4 мин.- 5 мин.40 сек.
Этап 3. Редактирование цифровых копий			
9.	Пакетная обработка сканов с использованием программного обеспечения на основе искусственного интеллекта	Загрузка цифровой копии в прикладную программу; исправление перекосов и искажений строк; исправление трапециевидных искажений; корректировка яркости и контрастности; подбор единого размера для всех страниц документа; обрезка страницы документа; распознавание и сохранение в форматах jpg и pdf; контроль качества цифровой копии.	1мин.- 3мин. 50 сек.

Этап 4. Загрузка цифровой копии документа в автоматизированную библиотечно-информационную систему			
10.	Передача оцифрованного документа в отдел автоматизации для включения в электронный ресурс	Проведение сверки цифровой копии с оригиналом (проверка пагинации, качества сканов и др.); сохранение на электронном носителе; передача USB флэш-накопителя в отдел автоматизации	15 мин.- 17мин.
11.	Загрузка отделом автоматизации на сервер хранения ресурса оцифрованного документа	Сохранение и индексация цифровой копии документа	3 мин.-5мин.
12.	Загрузка описания документа, установление связи каждого документа и редактирование описания	Добавление подробного описания документа; создание связей между документами; внесение изменений в описание	2 мин.-3 мин.
13.	Размещение цифровой копии в электронном ресурсе «Периодика Ставрополья»	Интеграция цифровой копии в публичный ресурс	3 мин.-5 мин.

Итоговые временные рамки всего процесса: минимальное время: 65 минут 27 секунд; максимальное время: 125 минут 20 секунд.

6. Требования к сканированию документов и обработке их электронных копий

Печатные издания, рукописи, а также листовые материалы должны быть переведены в электронную форму с разрешением не менее 300 dpi в режиме 24-bit RGB (цветной режим) в постраничном представлении. Отдельные экземпляры сканируются с разрешением 600 dpi. Размер цифрового образа в пикселях с учетом разрешения не должен отличаться от физического размера документа более чем на 5%.

Сканирование осуществляется в отношении всех страниц книги. Пустые страницы (вакаты) также подлежат сканированию, независимо от наличия или отсутствия на них номера и другой значащей информации. Листовые изобразительные материалы, имеющие на обороте листа владельческие записи и/или штампы, сканируются с оборотом, включая библиотечные штампы и регистрационные номера.

Цифровой образ должен соответствовать оригиналу произведений на бумажном носителе. Не допускается присутствие графических и иных артефактов сканирования и обработки (клинья, полосы, мусор, волосы и пр.). Мелкоформатные страницы должны быть отсканированы на фоне основной части страниц книжного блока, при условии, что на полученном образе не видна информация с соседних страниц. В противном случае и лицевую, и оборотную стороны мелкоформатных страниц рекомендуется сканировать с использованием чёрной подложки. Повторное их сканирование вне страниц книжного блока не требуется.

Распашные таблицы и страницы, содержащие иллюстрации, планы, схемы, чертежи и т.д., расположенные цельно на двух страницах (развороте), должны быть отсканированы разворотами.

При наличии в оригинале ярко выраженного просвечивания текста (иллюстраций), а также утраты части листа (отверстия, рваные края), приводящих к искажению или утрате информации, сканирование должно проводиться с подкладыванием черного (в необходимых случаях иного однотонного) листа бумаги, обеспечивающего выравнивание цветового фона страницы и чёткое определение характера повреждения оригинала.

Настройки яркости и контрастности производятся для достижения наилучшей резкости и цветности изображения. Графические образы должны быть сфокусированными. Глубина и насыщенность цвета образов должны быть максимально единообразны в пределах одного документа. Рекомендуется при сканировании избегать образования светлой или тёмной полосы (тени от переплёта) у корешка. Это достигается установкой и правильной регулировкой дополнительных световых приборов.

Для полной передачи образа книги сканирование производится постранично, включая крышки переплета (обложку) и каскад страниц с запасом полей 15–30 мм от края образа издания и переплета. Во избежание избыточного фона сканирования и соблюдения рекомендуемых полей обрезки, книгу следует располагать параллельно горизонтальным и вертикальным направляющим сканера. Допускается подкладывание черного листа под сканируемый каскад страниц частями по 10–20 страниц.

Ориентация всех страниц должна соответствовать оригиналу. Если в издании есть листовые материалы (иллюстрации, карты, вкладки и т.д.), не скрепленные с переплетом, то они располагаются в соответствии с ориентацией изображения или текста.

Для улучшения внешнего вида графических образов после работ по сканированию при необходимости производится дополнительная обработка.

При проведении дополнительной обработки должны быть выполнены следующие требования:

- графические образы страниц должны быть выровнены по большей стороне для книжной (альбомной) ориентации по центральной вертикали (горизонтали), если изображение геометрически расположено трапецией;

- строки текста должны быть без изгибов, за исключением тех случаев, когда изгиб является особенностью самого издания;

- размеры графических образов страниц одного формата в пикселях должны быть одинаковыми. Размеры образов большеформатных или мелкоформатных страниц (вклеек, карт, вкладышей и пр.) должны сохранить пропорции относительно размеров основного массива страниц;

- окончательный размер полей после выравнивания образов страниц должен составлять 3–10 мм. Не рекомендуется попадание текста с соседней страницы в поле сканирования.

При обработке конволютов образ каждого аллигата (составной части конволюта) должен быть сохранен в отдельный файл. Файл первого аллигата должен начинаться с образа верхней крышки переплета. Последующие (кроме последнего) аллигаты собираются в PDF-файл без крышек переплета. Файл последнего аллигата должен заканчиваться образом нижней крышки переплета. В случае сплетенных томов/частей одного издания запрещается добавлять в файлы второго и последующих томов/частей любые листы (например, титульный лист многотомника, оглавление и т.д.) из первого тома/части.

В результате сканирования документов цифровые образы должны быть сохранены в формате, который бы обеспечивал отсутствие потерь в исходном качестве изображения (например, TIFF без компрессии, TIFF с ZIP-компрессией, JPEG2000 с компрессией без потери качества), в постраничном виде с разрешением 300 или 600 dpi (одинаковым для всего документа).

Комплект полученных при сканировании графических образов страниц после обработки собирается в PDF-файлы в формате PDF/A-1b с компрессией JPEG (уровень качества не ниже 70%).

Текст, расположенный на страницах массива, полученного при сканировании документов, не должен быть распознан. Так называемый подслои с распознанным текстом должен отсутствовать.

Каждая страница полученного PDF-файла должна содержать только один графический объект (файл). Не должно быть скрытого разбиения образа страницы на части.

7. Контроль качества

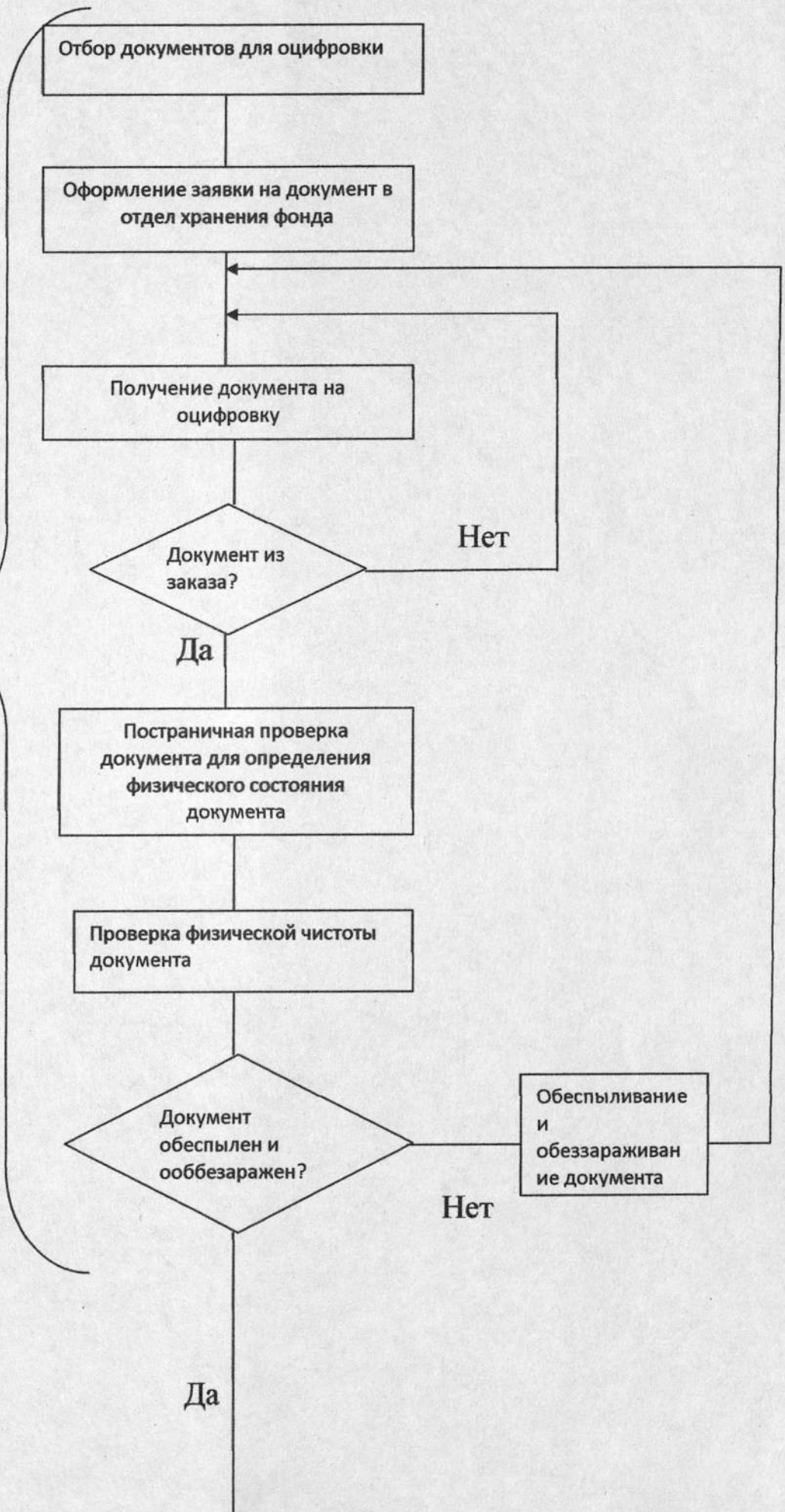
Контроль качества выполняется специалистом отдела информационных технологий и включает проверку: полноты страниц, четкости и соответствия техническому заданию, сборки файлов в единый архив.

8. Заключение

Данная технологическая карта регламентирует последовательность операций, используемое оборудование, контроль качества; способствует снижению рисков ошибок и обеспечивает сохранность культурных ценностей.

АЛГОРИТМ ОЦИФРОВКИ ДОКУМЕНТА

Этап 1
Прием и подготовка
документа



Этап 2
Сканирование
документа

Постановка задачи на сканирование
(ознакомление с техническим заданием на
оцифровку)

Запуск оборудования: включение
сканера, программы сканирования .
Установка формата и настроек в
соответствии с техническим заданием

Сканирование документа

Первичная постобработка цифровой копии
документа(конвертирование обрезка,
выравнивание яркости, удаление шумов) на
ПО из комплекта сканера

Контроль качества сканирования
документа

Качество
соответствует
требованиям ТЗ?

Нет

Да

Постановка задачи на последующее
редактирование цифровой копии
документа(ознакомление с ТЗ на оцифровку)

Пакетная обработка сканов с использованием
программного обеспечения на основе
искусственного интеллекта(в графическом
редакторе):устранение геометрических
искажений, удаление артефактов и т.д.

Контроль качества вторичной графической
обработки цифровой копии документа

Качество вторичной графической
обработки цифровой копии
соответствует ТЗ?

Нет

Этап 3
Редактирование
Цифровых копий
документов

Этап 4
Загрузка
цифровой копии
документа
в АБИС

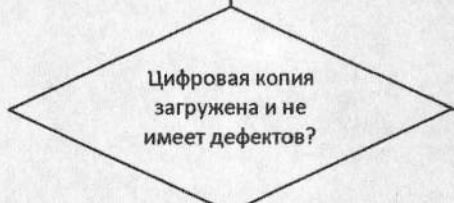
Передача оцифрованного документа в
отдел автоматизации для включения в
электронный ресурс

Загрузка отредактированных цифровых
копий документа на сервер хранения
ресурса оцифрованного документа

Загрузка описания документа,
установление связи каждого документа и
редактирование описания

Размещение цифровой копии в
электронном ресурсе(АБИС)

Проверка загруженной цифровой копии
документа в АБИС



Нет

Этап 5
Возврат
документа
в отдел хранения
фонда

Постраничная проверка
документа для определения
физического состояния документа
после сканирования для возврата
документа в отдел хранения

Возврат документа в отдел
хранения фонда