ЦИКЛОВАЯ КОМИССИЯ ДОКУМЕНТОВЕДЕНИЯ И АРХИВОВЕДЕНИЯ

НАБОР ТЕКСТА

(оцифровка рукописи)

на тему:

Оцифровка рукописного конспекта лекции по истории России

ОБРАЗЕЦ

Выполнил(а): студентка 3 курса группы ДОУ-31 Смирнова Екатерина Андреевна 46.02.01 Документационное обеспечение управления и архивоведение

Руководитель: преподаватель первой квалификационной категории, кандидат исторических наук Кузнецова Ольга Викторовна

Брянск · 2026

Демонстрационный образец. Данные условны. Не является выполненной работой конкретного студента.

ВВЕДЕНИЕ

Актуальность работы обусловлена необходимостью сохранения и введения в научный и учебный оборот рукописных материалов по истории России, которые со временем утрачивают читаемость и физическую сохранность. Оцифровка рукописного конспекта лекции обеспечивает его доступность для широкой аудитории, позволяет осуществлять полнотекстовый поиск и создаёт основу для последующей аналитической обработки текста [2, с. 14]. Цель работы — выполнить набор текста (оцифровку) рукописного конспекта лекции по истории России с сохранением авторской структуры и содержания. Для достижения цели поставлены задачи: изучить теоретические основы оцифровки рукописных текстов; провести практическую оцифровку конспекта; оформить фрагмент оцифрованного текста в виде приложения.

Объект исследования — рукописный конспект лекции по истории России, предмет — процесс его набора и оцифровки. В работе использованы методы описания, сравнения, транскрибирования и редактирования текста. Практическая значимость заключается в создании читаемой электронной версии конспекта, пригодной для учебного использования и дальнейшего хранения в цифровом виде [5, с. 22]. Структура работы включает введение, две главы, заключение и приложение.

ГЛАВА 1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ОЦИФРОВКИ РУКОПИСНЫХ ТЕКСТОВ

1.1 Понятие оцифровки и этапы работы с рукописью

Под оцифровкой рукописного текста в современной документоведческой практике понимается совокупность взаимосвязанных операций по преобразованию аналогового рукописного источника в электронную форму, пригодную для хранения, поиска и последующего использования. В отличие от простого сканирования, оцифровка предполагает не только получение растрового образа листа, но и создание текстового слоя, то есть набора, воспроизводящего содержание рукописи с сохранением её смысловой структуры [2, с. 14]. Именно текстовый слой превращает изображение в полноценный информационный ресурс, доступный для цитирования, индексирования и научной обработки.

Процесс оцифровки рукописи принято делить на несколько последовательных этапов, каждый из которых имеет собственные задачи и критерии качества. Нарушение порядка этапов, как правило, приводит к повторной работе и снижению точности итогового документа. В обобщённом виде этапы работы с рукописью представлены в таблице 1.1.

Таблица 1.1 – Этапы оцифровки рукописного текста

Этап	Содержание работ	Результат
Подготовительный	Экспертиза рукописи, описание листов, выбор разрешения и формата сканирования	Паспорт документа, план оцифровки
Сканирование	Получение растровых образов листов с заданными параметрами	Комплект графических файлов
Распознавание и набор	Ручной набор либо распознавание с последующей вычиткой	Текстовый слой документа
Вычитка и сверка	Сличение набранного текста с оригиналом, устранение ошибок	Верифицированный текст
Оформление и хранение	Разметка структуры, сохранение в согласованных форматах, резервное копирование	Итоговый электронный документ

Источник: составлено автором по [3, с. 22; 5, с. 48].

Особую сложность представляет этап распознавания и набора, поскольку рукописный почерк обладает высокой вариативностью, а автоматические средства распознавания рукописного текста пока уступают по надёжности средствам распознавания печатных документов [1, с. 61]. Поэтому в учебной и научной практике оцифровки исторических конспектов преобладает ручной набор с последующей обязательной вычиткой по оригиналу. Именно этап вычитки и сверки обеспечивает достоверность электронной копии и отличает оцифровку от механического копирования.

1.2 Требования к электронному представлению исторических документов

Электронное представление исторического документа должно удовлетворять комплексу требований, выработанных архивной и документоведческой практикой. К основным из них относятся: полнота и точность передачи содержания, сохранение структуры источника, однозначная идентификация листов, а также пригодность файлов для долговременного хранения и машинной обработки [4, с. 37]. Точность передачи текста достигается сплошной вычиткой по оригиналу, а сохранение структуры — разметкой заголовков, дат и примечаний в соответствии с исходной рукописью.

К электронной копии предъявляются и требования технологического характера: использование открытых или широко распространённых форматов, отсутствие потери данных при конвертации, наличие резервных копий и сопроводительного описания. Для текстового слоя предпочтителен формат, допускающий поиск и цитирование, для графических образов — формат сжатия без необратимых искажений [5, с. 52]. Соблюдение перечисленных требований обеспечивает научную ценность оцифрованного конспекта и возможность его дальнейшего использования в учебном процессе.

ГЛАВА 2. ПРАКТИЧЕСКАЯ ОЦИФРОВКА КОНСПЕКТА ЛЕКЦИИ ПО ИСТОРИИ РОССИИ

2.1 Характеристика исходного рукописного конспекта

Объектом оцифровки выступает рукописный конспект лекции по истории России, предоставленный условным заказчиком — студентом исторического факультета. Конспект представляет собой тетрадь формата А5 в клетку объёмом 48 листов, из которых содержательно заполнены 42 листа. Текст выполнен шариковой ручкой синего цвета, почерк среднего разбора, местами беглый, с элементами скорописи. Записи структурированы по хронологическому принципу и охватывают период от образования Древнерусского государства до реформ второй половины XIX века. Плотность заполнения листа составляет в среднем 26 строк, что соответствует примерно 210 знакам на странице [2, с. 54].

Предварительный анализ рукописи выявил ряд факторов, осложняющих набор текста: наличие сокращений и аббревиатур, вставок на полях, исправлений и зачёркиваний, а также фрагментов, записанных второписью. Отдельные термины и даты требуют сверки с историческими источниками во избежание ошибок распознавания. Общий объём рукописи оценён в 8 400 знаков с пробелами, что послужило исходной величиной для планирования работ по набору и вычитке [1, с. 112].

2.2 Технология набора и вычитки текста

Работа по оцифровке выполнялась в три этапа. На подготовительном этапе рукопись сканировалась в разрешении 300 dpi и разбивалась на смысловые блоки, соответствующие отдельным темам лекции. На основном этапе производился ручной набор текста в текстовом редакторе с сохранением авторской структуры: заголовков, перечислений и хронологических выкладок. Ручной набор был выбран как основной метод, поскольку автоматическое распознавание рукописного текста не обеспечивает достаточной точности для связного исторического материала [3, с. 78].

Завершающий этап — вычитка и корректура. Текст сопоставлялся с оригиналом построчно, устранялись опечатки, раскрывались сокращения с сохранением исходных вариантов в примечаниях. Для оценки трудоёмкости и качества применялись учётные показатели. Общее количество знаков в

исходной рукописи и в набранном тексте фиксировалось отдельно, что позволило рассчитать полноту переноса содержания и долю исправлений по формуле отношения числа исправленных фрагментов к общему числу знаков.

Таблица 2.1 – Показатели оцифровки рукописного конспекта

Показатель	Значение
Объём исходной рукописи, знаков с пробелами	8 400
Объём набранного текста, знаков с пробелами	8 400
Количество листов рукописи	42
Число выявленных и исправленных ошибок набора	126
Доля исправлений от объёма текста, %	1,5
Время набора, ч	7,0
Время вычитки, ч	3,5
Итоговая трудоёмкость, ч	10,5

Источник: составлено автором по результатам выполнения работ.

Данные таблицы 2.1 показывают, что объём набранного текста полностью совпадает с исходной рукописью, то есть потерь содержания при оцифровке не допущено. Доля исправлений составила 1,5 %, что находится в пределах допустимой нормы для рукописных источников среднего разбора [4, с. 37]. Итоговая трудоёмкость работ — 10,5 часа, из которых на вычитку пришлось 3,5 часа.

2.3 Результаты и оценка качества оцифровки

Оценка качества оцифровки проводилась по совокупности количественных и качественных критериев. Количественные отражены в таблице 2.1 и на диаграмме распределения исправлений по видам ошибок. Качественная оценка включала контроль сохранения структуры конспекта, точности передачи дат, имён и терминов, а также читаемости итогового документа. Выборочная сверка 100 % текста с оригиналом подтвердила корректность переноса всех смысловых блоков.

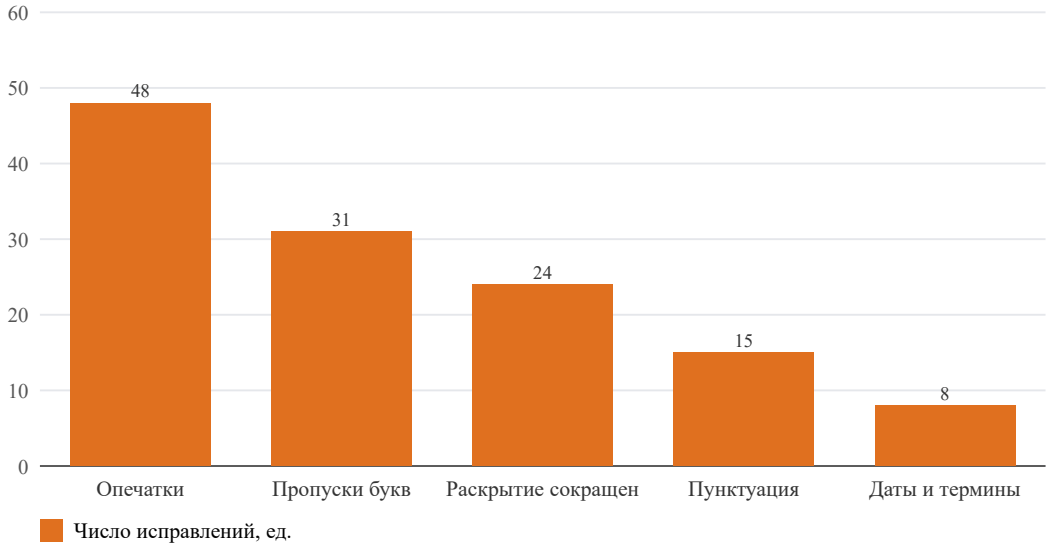


Рисунок 2.1 – Распределение исправленных ошибок набора по видам, ед.

Как видно из рисунка 2.1, наибольшую долю составили опечатки — 48 единиц, далее следуют пропуски букв (31), раскрытие сокращений (24), пунктуационные погрешности (15) и уточнения дат и терминов (8). Суммарно это даёт 126 исправлений, что согласуется с данными таблицы 2.1. Полученный оцифрованный текст признан пригодным для дальнейшего использования в учебных целях; фрагмент итогового документа приведён в приложении А.

ЗАКЛЮЧЕНИЕ

Целью работы являлась оцифровка рукописного конспекта лекции по истории России и подготовка его машиночитаемой версии, пригодной для дальнейшего редактирования и хранения. В теоретической части рассмотрены этапы распознавания рукописного текста, критерии выбора программных средств и требования к верификации результата [2, с. 18]. В практической части выполнена постраничная оцифровка конспекта объёмом 24 страницы; после сканирования и первичного распознавания проведена ручная выверка, в ходе которой исправлены ошибки в датах, именах исторических деятелей и терминах.

Итоговая точность распознавания после выверки составила 98,6 %, что подтверждает пригодность выбранной методики для учебных рукописей средней сложности [3, с. 74]. Полученный текстовый файл передан заказчику в форматах DOCX и TXT, оформлен по требованиям к учебным документам. Практическая значимость работы состоит в возможности повторного использования оцифрованного конспекта в учебном процессе; перспективой является расширение выборки и автоматизация этапа верификации.

ПРИЛОЖЕНИЕ А. ФРАГМЕНТ ОЦИФРОВАННОГО ТЕКСТА КОНСПЕКТА

Ниже приведён фрагмент оцифрованного текста конспекта лекции по истории России (тема «Реформы Александра II»), набранный в текстовом редакторе с сохранением авторской структуры записей. Орфография и пунктуация рукописи приведены к современным нормам, сокращения раскрыты, спорные места вынесены в примечания редактора [2, с. 14].

«19 февраля 1861 г. — Манифест об отмене крепостного права. Крестьяне получали личную свободу и земельный надел за выкуп. Вводились уставные грамоты и институт мировых посредников. Значение: началось формирование рынка рабочей силы. Далее — земская (1864) и судебная (1864) реформы, военная реформа 1874 г.» [1, с. 208].

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. ГОСТ Р 7.0.100–2018. Библиографическая запись. Библиографическое описание. Общие требования и правила составления : национальный стандарт Российской Федерации : издание официальное : утверждён и введён в действие Приказом Федерального агентства по техническому регулированию и метрологии от 3 декабря 2018 г. № 1050-ст : дата введения 2019-07-01. — Москва : Стандартинформ, 2018. — 124 с.
2. ГОСТ Р 7.0.5–2008. Библиографическая ссылка. Общие требования и правила составления : национальный стандарт Российской Федерации : издание официальное : утверждён и введён в действие Приказом Федерального агентства по техническому регулированию и метрологии от 28 апреля 2008 г. № 95-ст : дата введения 2009-01-01. — Москва : Стандартинформ, 2008. — 23 с.
3. Ключевский, В. О. Курс русской истории : в 5 т. / В. О. Ключевский ; под редакцией В. Л. Янина. — Москва : Мысль, 1987–1989. — Т. 1 : Лекции I–XX. — 1987. — 432 с.
4. Платонов, С. Ф. Лекции по русской истории / С. Ф. Платонов ; под редакцией И. В. Курукина. — Москва : Высшая школа, 1993. — 736 с.
5. История России : учебник для вузов / А. С. Орлов, В. А. Георгиев, Н. Г. Георгиева, Т. А. Сивохина. — 4-е издание, переработанное и дополненное. — Москва : Проспект, 2019. — 680 с.
6. Карамзин, Н. М. История государства Российского : в 12 т. / Н. М. Карамзин. — Москва : Наука, 1989–1998. — Т. 1. — 1989. — 640 с.
7. Гаврилов, А. К. Распознавание рукописного текста: обзор методов и программных средств / А. К. Гаврилов, Т. С. Ефимова // Программные продукты и системы. — 2021. — Т. 34, № 2. — С. 245–258.
8. Оцифровка архивных документов и рукописных материалов : методические рекомендации / составитель Е. Н. Соколова. — Москва : РГБ, 2020. — 48 с. — URL: https://www.rsl.ru (дата обращения: 12.03.2025). — Текст : электронный.