

УДК 002.2:091(=161.1):025.13:004  
ББК 76.19,2+78.347.8

## ОСОБЕННОСТИ СОЗДАНИЯ ЦИФРОВЫХ КОПИЙ СОБРАНИЙ РУКОПИСНЫХ КНИГ

© А. В. Шабанов, 2013

*Государственная публичная научно-техническая библиотека  
Сибирского отделения Российской академии наук  
630200, г. Новосибирск, ул. Восход, 15*

Рассмотрена комплексная задача создания цифровых копий для собраний русских рукописных книг.

*Ключевые слова:* русские редкие книги и рукописи, цифровые библиотеки, обработка изображений.

The complex task of creating digital copies of Russian book manuscripts collections is considered.

*Key words:* Russian rare books and manuscripts, digital libraries, image processing.

**З**адача создания высококачественной цифровой копии редкой книги достаточно сложна. Необходимо рассматривать аппаратные, программные и концептуальные вопросы. Следует также отметить, что в этой области за последние три года произошли огромные изменения и многие задачи, которые еще сравнительно недавно были трудноразрешимыми, стали вполне обычными, доступными довольно широкому кругу специалистов.

В нашей работе 2008 г. [1] были подробно рассмотрены основные факторы, влияющие на конечный результат оцифровки на примере русских старопечатных и рукописных книг:

- функции системы,
- состав коллекции,
- аппаратура,
- прикладное программное обеспечение,
- форматы данных,
- система резервирования первичного архива,
- связь с научным описанием источника,
- учет возможности развития как по функциям, так и по объемам.

В связи с отмеченным выше бурным развитием этой области роль перечисленных факторов несколько изменилась, характер этих перемен и будет описан в настоящей статье.

К функциям системы добавилась возможность создания 3D образа книги. Хотя пока еще трудно считать эти работы крайне необходимыми применительно ко всем собраниям и коллекциям редких книг, одно из направлений – музейная деятельность, видимо, может значительно выиграть от использования 3D образов уникальных книжных памятников.

По учету состава коллекции радикальные изменения следующие. Во-первых, стало возможным создание цифровых копий непосредственно в месте «бытования» первоисточника и, во-вторых, оцифрованы огромные массивы научно-справочной литературы археографии, истории и смежных дисциплин, что значительно увеличило полноту фондов – пусть эти копии и не бумажные, а цифровые.

Относительно аппаратуры ситуация радикально улучшилась по части предложений установок, которые специально разработаны именно для оцифровки редких книг, и возможностей цифровых фотоаппаратов [2]. И если, к сожалению, стоимость специализированных комплексов пока велика и они малодоступны для большинства российских хранилищ редких книг, то цифровые фотоаппараты, хоть и в ограниченной степени, но становятся незаменимыми при частичной оцифровке коллекций первоисточников. На конец 2011 г. диапазон общей стоимости аппаратуры, способной качественно решать поставленную задачу (в зависимости от цели и коллекции), – от 30 тыс. до 10 млн руб.

Основные задачи при «доводке» первичного изображения до конечного, адекватного поставленным целям, сохранились: цветокоррекция, повышение резкости и устранение геометрических искажений листов.

Для задачи цветокоррекции на первый план выходит тщательная подготовка помещения, где ведется оцифровка – создание режима «темной комнаты». Также радикально улучшились характеристики цифровых фотоаппаратов.

Повысить резкость можно разными способами. Из чисто программных:

Суммарные данные по кириллическому фонду  
ОРКиР ГПНТБ СО РАН

Собрание	Рукописи	Печатные издания
М. Н. Тихомирова	649	93
Алтайское	50	65
Дальневосточное	18	29
Забайкальское	137	130
Кемеровское	56	68
Красноярское	56	22
Томское	107	118
Тувинское	7	8
Текущие поступления	160	650
<i>Всего</i>	1 240	1 183

- использовать Unsharp Mask с заранее подобранными параметрами и тем или иным способом ослабить контурную резкость;
- изображение перевести в цветное пространство Lab, выбрать канал Lightness и для него применять фильтры, повышающие резкость.

Следует отметить широкие возможности пакетных комплексов программ по обработке изображений, например, предлагаемых фирмой FastStone.

В связи с тем, что кардинального улучшения в реставрационных работах для собраний русских рукописных книг пока не произошло, при оцифровке, как правило, сохраняются геометрические искажения листов. Преобладающими форматами данных для оцифровки редких книг и рукописей являются tiff, jpeg, raw.

Гораздо более доступным для любого пользователя стало создание системы резервирования первичных архивов изображений. На конец 2011 г. она состоит из двух частей: копии в локальной сети и дублирование в удаленных глобальных системах поддержки сохранности информации.

Работы по научным описаниям рукописных книг значительно продвинулись вперед благодаря ряду проектов, поддержанных Российской академией наук (РАН), Российским фондом фундаментальных исследований (РФФИ), Российским гуманитарным научным фондом (РГНФ).

Подробное обсуждение этого вопроса не входит в задачи настоящего сообщения. Можно только повторить, что высококачественная оцифровка позволяет существенно повысить уровень и возможности научного описания книжного памятника.

Перейдем к изложению возникающих задач на примере фонда отдела редких книг и рукописей (ОРКиР) ГПНТБ СО РАН, основного в нашей работе. Главные из отработанных концептуальных, аппаратных и программных решений применимы, естественно, для широкого круга сибирских хранилищ кириллических рукописных и старопечатных книг.

Базовая информация об объеме и составе «сибирского фонда» – 18 каталогов серии «Рукописи, старопечатные и редкие книги Сибири Дальнего Востока», выходящей с 1989 г. под общей редакцией академика А. П. Деревянко. Электронные версии этих каталогов включены в состав цифровой библиотеки. В качестве основной для оцифровки выбрана рукописная часть фонда (табл. 1).

Среди сложных задач – обработка первичных изображений с целью получения цифровой копии, адекватно передающей первоисточник [3].

Возможность генерации электронной копии высокого качества определяют: физическая сохранность листов книги, состояние и вид переплета, особенно, величина внутренних полей, ограничения оборудования по степени разрешения (dpi).

По этим факторам были оценены основные рукописные собрания. После чего осуществлена «массовая» оцифровка первых 200 единиц собрания М. Н. Тихомирова и многих рукописных книг из территориальных коллекций, а также фондов сибирских хранилищ. Из первых 200 единиц собрания М. Н. Тихомирова качественные электронные копии созданы для 163 источников.

За один год удастся добавлять в цифровую библиотеку от 100 до 150 единиц – влияет весьма большая разнородность собраний с точки зрения трудоемкости получения первичного архива, последующей обработки изображений и добавления в электронную копию научного описания.

Суммарные данные по цифровой библиотеке «Книжные памятники Сибири» приведены в табл. 2.

Общий объем оцифрованных источников – более 367 Гб, или около 150 тыс. страниц. Значительная часть электронных копий доступна без ограничений в Интернете ([www.spsl.nsc.ru/rbook](http://www.spsl.nsc.ru/rbook)). Полная версия цифровой библиотеки предоставляется в ГПНТБ СО РАН в рамках единой системы доступа к информации зарегистрированных пользователей<sup>1</sup>.

Для проведения оцифровки редких книг и рукописей в ГПНТБ СО РАН используются две установки (типа ПланСкан «Репро» фирмы ЭЛАР и на основе цифрового фотоаппарата Canon 500D) и два источника рассеянного света FalconEyes Softbox SBQ-7575.

Следующий важный этап – выбор системы управления базами данных (СУБД), системы опрелеления полномочий пользователей по доступу

<sup>1</sup> Для получения копий на DVD/CD следует обратиться по e-mail: [rbook@spsl.nsc.ru](mailto:rbook@spsl.nsc.ru).

Суммарные данные по цифровой библиотеке «Книжные памятники Сибири» (по состоянию на май 2012 г.)

Состав по собраниям, коллекциям, месту хранения	Оцифровано, ед.
Рукописи и книги кириллической традиции	259
В том числе:	
Собрание М. Н. Тихомирова (ГПНТБ СО РАН)	188
Территориальные коллекции ГПНТБ СО РАН, 1965–2010 гг. (Алтайское, Красноярское, Дальневосточное, Томское, Забайкальское собрание, текущие поступления)	26
Книги общинных и частных старообрядческих собраний Сибири и Дальнего Востока, «живые» библиотеки, 2003–2010 гг. (Красноярский край, Алтайский край, Кемеровская обл., Приморский край, Тюменская обл.)	15
Сибирские хранилища, за пределами Новосибирска, 2000–2010 гг. (Тобольск, Улан-Удэ, Тюмень, Хабаровск, Бийск)	30
Медицинские сочинения Центра восточных рукописей и ксилографов ИМБТ СО РАН	64
Книжные памятники XIX в., ГПНТБ СО РАН (книги и периодические издания пушкинской эпохи, прижизненные издания А. С. Пушкина, прижизненные издания Ф. М. Достоевского)	487
<i>Всего</i>	810

к полнотекстовой информации и пользовательского программного обеспечения для генерации электронной копии. Здесь вполне достаточными оказались возможности ранее созданных в ГПНТБ СО РАН систем и, соответственно, новые разработки не потребовались.

Структура научного описания рукописного книжного памятника кириллической традиции:

<b>tbl_rbooks Книги</b>	
ID int NOT NULL AUTO_INCREMENT	Ключевое поле
TITLE varchar (255)	Название источника
SELF_TITLE varchar (255)	Самоназвание источника
SOBRANIE varchar (255)	Собрание/коллекция
HREF_A varchar (255)	Ссылка на профессиональную версию источника
HREF_C varchar (255)	Ссылка на интернет-версию источника
DATE varchar (255)	Время создания
DESCR varchar (255)	Ссылка на научное описание
SCI_INF varchar (255)	Ссылка на научно-справочную информацию
PAGECOUNT varchar (255)	Количество страниц
SIZE_A varchar (255)	Размер профессиональной версии источника

SIZE_C varchar (255)	Размер интернет-версии источника
FORMAT varchar (255)	Формат
FOLIAC varchar (255)	Фолиация
SIGNATUR varchar (255)	Сигнатуры
BUMAGA varchar (255)	Бумага
POCHERK varchar (255)	Почерк
OBL_PISMA varchar (255)	Область письма
FILIGRANI varchar (255)	Филиграни
ZAPISI varchar (255)	Записи
COLONTITUL varchar (255)	Колонтитул
OFORMLENIE varchar (255)	Оформление
PEREPLET varchar (255)	Переплет
SHIFR varchar (255)	Шифр
SOSTAV varchar (255)	Состав
FIZ_SOST varchar (255)	Физическое состояние
PRIOBR varchar (255)	Место и время приобретения
BIBLIOGRAF varchar (255)	Библиография
FONET varchar (255)	Фонетическая редакция текста
NOTATION varchar (255)	Нотация
ROSPEV varchar (255)	Роспев
MUZ_OSOB varchar (255)	Музыкальные особенности

tbl_sobranie Собрание	
ID int NOT NULL AUTO_INCREMENT	Ключевое поле
SOBRNAME varchar (100)	Название собрания/ коллекции

Таким образом, комплексная разработка задачи создания электронных копий русских рукописных книг позволила создать цифровую библиотеку «Книжные памятники Сибири» ([www.spsl.nsc.ru/rbook](http://www.spsl.nsc.ru/rbook)), которая активно используется как специалистами, так и в учебном процессе [4, 5].

Следующим этапом развития цифровой библиотеки «Книжные памятники Сибири» является создание цифровых копий для всех рукописных книг фонда ОРКиР ГПНТБ СО РАН, продолжение работ с другими сибирскими хранилищами.

### Литература

1. Факторы, влияющие на выбор технологии оцифровки русских старопечатных и рукописных книг // Библиосфера. – 2008. – № 4. – С. 46–48.
2. Шабанов А. В. Сравнение установок для оцифровки русских старопечатных и рукописных книг и методы обработки изображений // Библиосфера. – 2010. – № 2. – С. 30–32.
3. Цифровая версия Острожской Библии: к проблеме максимального функционального приближения электронного образа к «живому» книжному памятнику / Е. И. Дергачева-Скоп [и др.] // Остромирово Евангелие и современные исследования. – СПб., 2010. – С. 84–93.
4. Создание цифровой библиотеки древнерусских рукописных и старопечатных книг современных сибирских хранилищ как единого информационно-коммуникационного пространства и использование ее в учебном процессе гуманитарного факультета НГУ / Е. И. Дергачева-Скоп [и др.] // Регионы России для устойчивого развития: образование и культура народов Российской Федерации : материалы Междунар. науч.-практ. конф. (Новосибирск, 25–27 марта 2010 г.). – Новосибирск, 2010. – С. 297–306.
5. Бородихин А. Ю., Елепов Б. С., Шабанов А. В. Результаты выполнения интеграционного проекта «Принципы и технология электронного представления книжного памятника» // Информационные технологии и электронные ресурсы в библиотеках : материалы межрегион. науч.-практ. конф. (26–30 сент. 2011 г., Респ. Бурятия, г. Улан-Удэ). – Новосибирск, 2011.

Материал поступил в редакцию 14.02.2013 г.

Сведения об авторе: Шабанов Андрей Васильевич – кандидат технических наук, старший научный сотрудник отдела редких книг и рукописей, тел.: (383) 266-10-91, e-mail: shabanov@spsl.nsc.ru