

ОБ ОСОБЕННОСТЯХ СУЩЕСТВУЮЩИХ ПОДХОДОВ К СОЗДАНИЮ ЦИФРОВОГО ОБРАЗА СТАРОПЕЧАТНОЙ КНИГИ И ЕГО ПРЕДСТАВЛЕНИЮ В ЭЛЕКТРОННЫХ БИБЛИОТЕКАХ МИРА

В. В. Свёкла

Центральная научная библиотека НАН Беларуси, Минск

Проведено сравнение подходов к созданию цифровых копий старопечатных книг и их представлению в Интернете, используемых различными электронными библиотеками мира. Особое внимание уделено таким аспектам создания цифровых копий изданий данного типа, как выбор цветового режима оцифровки; выбор между созданием постраничных цифровых образов и образов, выполненных в разворот; требования к полноте цифровой копии старопечатной книги; выбор подхода к созданию цифрового образа отдельной книжной страницы.

Введение

Участники процессов оцифровки библиотечных фондов и создатели электронных библиотек из разных стран мира нередко придерживаются дифференцированного подхода к оцифровке различных категорий документов. Наряду с общими требованиями, предъявляемыми к созданию цифровых копий всех без исключения документов, существует ряд специфических требований к оцифровке определенных групп документов, выделенных по тому или иному признаку (виду документа, времени его создания, историко-культурной или информационной ценности и т. д.).

Определенное представление о дифференцированном подходе к оцифровке различных категорий документов дают встречающиеся в печати и Интернете документы методического и нормативного характера. Ни один из известных документов такого рода не использует определение «старопечатные книги» для отдельной категории документов, требующей особых подходов к оцифровке. Однако сформулированные в них принципы категориального деления указывают на то, что процесс оцифровки старопечатных изданий обладает своей спецификой, отличающей его, к примеру, от процесса оцифровки современных изданий или иных видов документов.

1. О специфике технических требований к оцифровке старопечатных книг

Показателен пример категориального деления библиотечных фондов в отношении подходов к их оцифровке, сформулированный в документе под названием «Единые технические требования к оцифровке изданий, включаемых в НЭБ»¹ (далее ЕТТ) [1]. ЕТТ предлагают условно разделить все документы, подлежащие оцифровке, на три категории: раритеты, художественно-культурные ценности и социально значимые издания, информационно значимые издания.

Согласно предложенной классификации к первой из названных категорий относятся «книжные памятники (до 1830 г.), особо ценные и редкие книги, рукописи»

¹ Данный документ был разработан в качестве нормативно-методической базы для оцифровки документов в рамках создания Национальной электронной библиотеки России.

[1, с. 2]². Таким образом, ЕТТ причисляют старопечатные книги к категории «раритеты», не оставляя оснований для отнесения их к категории «художественно-культурные ценности и социально значимые издания» (сюда входят «издания, представляющие особую художественную и культурную ценность, содержащие авторские иллюстрации и рисунки, а также издания, имеющие большую социальную значимость») и тем более к категории «информационно значимые издания» (ее составляет «тиражная литература, которая представляет ценность только как информация, а не оригинальное издание» [1]).

ЕТТ – один из ряда нормативно-методических документов, основывающих категориальное деление библиотечных фондов в отношении подходов к их оцифровке на противопоставлении документов двух типов:

- имеющих исключительно информационную значимость;
- обладающих культурной, исторической, духовной ценностью наряду с информационной значимостью.

Старопечатные книги, без сомнения, относятся ко второй из названных групп документов. Подходы, используемые при их оцифровке, во многом схожи с подходами, применяемыми при оцифровке иных документов данной группы, но существенно отличаются от подходов к оцифровке документов, обладающих сугубо информационной значимостью. Вызвано это тем, что цифровая копия старопечатного издания должна не только в полном объеме воспроизводить содержащуюся в книге информацию, но и являться адекватной оригиналу реконструкцией предмета искусства, историко-культурного артефакта. Только цифровой образ старопечатной книги, соответствующий данным требованиям, способен удовлетворить потребности всех пользователей.

При оцифровке книг часто возникает дилемма: какой цифровой образ издания предпочтительнее для пользователя – позволяющий с максимальным удобством воспринимать текстовую и иную символьную информацию, содержащуюся в книге, или же дающий возможность составить наиболее органичное представление об издании в целом? В зависимости от предпочтений пользователей их самих можно разделить на две категории. Типичный представитель первой из них (условно назовем его «пользователь-читатель») ориентирован в основном на восприятие текстовой составляющей книги. Второй тип пользователя (его уместно наделить характеристикой «пользователь-зритель») в меньшей степени проявляет интерес к тексту книги, зато более восприимчив к ее зрительному образу. Любой подход к созданию цифрового образа старопечатной книги должен учитывать тот факт, что пожелания к нему названных типов пользователей могут существенно различаться.

Как уже отмечалось, подходы ряда электронных библиотек мира к созданию цифровых копий старопечатных книжных изданий во многом специфичны и неприменимы к оцифровке некоторых иных видов документов. В то же время особенности подходов к созданию такого рода копий у различных субъектов процессов оцифровки могут существенно различаться между собой. В этом убеждает не только анализ существующих нормативно-методических документов, касающихся оцифровки библиотечных фондов, но и непосредственное ознакомление с цифровыми образами старопечатных книг, представленных в разных электронных библиотеках мира.

В данном докладе выявление сходства и специфики существующих подходов к созданию цифрового образа старопечатного издания осуществляется последним из обо-

² Характерно, что дата, используемая в «Единых технических требованиях» как хронологическая граница для документов категории «Раритеты», совпадает с датой, которая, согласно доминирующей в отечественном книговедении точке зрения, завершает период создания старопечатной книги.

значенных способов. С этой целью автором рассмотрены десять электронных библиотек: Мировая цифровая библиотека (<https://www.wdl.org>); электронная библиотека Российской государственной библиотеки (РГБ) (<http://elibrary.rsl.ru>); открытая электронная библиотека Государственной публичной исторической библиотеки России (ГПИБ России) (<http://elib.shpl.ru>); электронные коллекции Библиотеки им. Врублевских Литовской академии наук (<http://elibrary.mab.lt>); электронная библиотека Национальной библиотеки Армении (<http://www.nla.am/arm/?q=en/node/3>); цифровая национальная библиотека Polona (Польша) (<https://polona.pl>); цифровая библиотека Gallica Национальной библиотеки Франции (<http://gallica.bnf.fr>); цифровая библиотека Hispánica Национальной библиотеки Испании (<http://www.bne.es/en/Catalogos/BibliotecaDigitalHispanica/Inicio/index.html>); цифровые коллекции Национальной и университетской библиотеки (Загреб, Хорватия) (<http://digitalna.nsk.hr>); электронные коллекции Национальной парламентской библиотеки Японии (http://dl.ndl.go.jp/?_lang=en).

При выборе для рассмотрения именно этих библиотек принимались во внимание их статус и значимость, принадлежность к различным странам и регионам мира, но в первую очередь – наличие в их составе цифровых копий старопечатных книг в объеме, достаточном для выявления определенных закономерностей в подходах, используемых при создании цифровых образов такого рода изданий.

Поскольку спектр требований, лежащих в основе того или иного подхода к созданию цифрового образа старопечатной книги, весьма широк, остановимся лишь на некоторых частных, однако важных аспектах существующих подходов. В их числе: выбор цветового режима оцифровки старопечатной книги, выбор между созданием постраничных цифровых образов и выполненных в разворот, требования к полноте цифрового образа старопечатного издания, выбор подхода к созданию цифрового образа отдельной страницы старопечатной книги.

2. Выбор цветового режима оцифровки старопечатной книги

Как известно, при оцифровке документов путем сканирования возможно использование трех цветовых режимов: RGB 24 bit (полноцветный режим), Grayscale 8 Bit (256 оттенков серого цвета), B&W (черно-белое изображение).

В большинстве рассмотренных библиотек представлены полноцветные цифровые образы старопечатных книг. Одни библиотеки (Hispánica, национальные библиотеки Хорватии, Армении) абсолютно последовательны в использовании данного режима оцифровки, в других же (Национальной парламентской библиотеке Японии, Polona) наряду с доминирующими полноцветными образами встречаются образы, выполненные в иных цветовых режимах. Не поддерживают тенденцию лишь электронные библиотеки РГБ и ГПИБ России. Преобладающий в данных библиотеках цветовой режим для создания цифрового образа старопечатной книги – 256 оттенков серого цвета. Копии старопечатных изданий, выполненные в полноцветном режиме, встречаются гораздо реже. РГБ нередко прибегает и к созданию черно-белых цифровых образов, при этом закономерности выбора того или иного цветового режима порой трудно установить.

Следует отметить, что почти все рассмотренные библиотеки придерживаются принципа цветового единообразия при создании цифрового образа старопечатного издания (все без исключения страницы и другие элементы отдельно взятого издания оцифровываются в одном и том же цветовом режиме независимо от того, какой режим используется). Исключением является ГПИБ России. Многие представленные здесь цифровые копии старопечатных книг выполнены следующим образом: основная масса страниц оцифрована с использованием режима «256 оттенков серого цвета», но некото-

рые книжные элементы (чаще всего переплет, корешок переплета, страницы, созданные с использованием цветной печати, содержащие цветные иллюстрации, нередко – титульный лист) воспроизведены в полноцветном режиме.

Аргументы в пользу того или иного цветового режима очевидны. Выполнение оцифровки в любом режиме, отличном от полноцветного, уменьшает затраты времени на создание копии книги, включение ее в состав электронной библиотеки и работу с ней пользователя. Однако очевидны и недостатки копий старопечатных изданий, выполненных в таких цветовых режимах. Эти копии способны обеспечить адекватную и полную передачу текстовой составляющей книги, но не позволяют составить представление о старопечатной книге как о предмете искусства, историко-культурном артефакте. Как следствие, цифровой образ старопечатной книги, созданный в ином цветовом режиме, нежели полноцветный, неприемлем для некоторых категорий пользователей (искусствоведов, книгovedов и т. д.). В отдельных случаях такие копии не удовлетворяют даже пользователей, обращающихся исключительно к тексту издания (в частности, текстологов).

3. Выбор между созданием постраничных цифровых образов и выполненных в разворот

По мнению автора, проблема выбора между данными способами представления цифрового образа сводится к выбору между предпочтениями двух названных выше типов пользователей электронных библиотек – «пользователя-читателя» и «пользователя-зрителя». Очевидно, что книга, представленная в электронной библиотеке в виде комплекса цифровых образов отдельных страниц, более приемлема для восприятия ее текстово-информационной составляющей и в большей степени способна удовлетворить «пользователя-читателя». В то же время цифровой образ книги, представленной в разворот, позволяет составить более целостное представление об издании как о предмете искусства, об историко-культурном артефакте, что соответствует запросам «пользователя-зрителя».

В случае с созданием цифрового образа книги, имеющей сугубо информационную ценность, выбор естественным образом делается в пользу первого из названных способов и на практике сводится к созданию цифровых копий отдельных книжных страниц. Ситуация с выбором подхода к созданию цифрового образа старопечатной книги, в равной степени обладающей свойствами как источника информации, так и предмета искусства, не столь однозначна. Однако, являясь в первую очередь информационными ресурсами, электронные библиотеки ориентируются прежде всего на пользователя – потребителя информации. Их создатели чаще склоняются к тому, что старопечатная книга должна быть представлена в виде совокупности цифровых образов отдельных страниц, а не книжных разворотов³. Одни из электронных библиотек последовательны в реализации такого подхода (Gallica, национальные библиотеки Хорватии, Армении), в других (РГБ, Polona, Hispánica) при явном преобладании постраничных цифровых образов встречаются и цифровые копии книжных разворотов.

³ В отличие, например, от создателей некоторых мультимедийных ресурсов, посвященных книге, виртуальных книжных выставок и т. д., ориентирующихся в первую очередь на интересы «пользователя-зрителя» и поэтому нередко прибегающих к использованию образов книжных разворотов.

4. Требования к полноте цифрового образа старопечатного издания

В контексте создания цифровых копий книг слова «полнотекстовый» и «полный» не всегда синонимичны. Полнотекстовая копия документа предполагает наличие в ней в абсолютном объеме текстовой информации, присутствующей в оригинале. Однако в состав книжного издания входит ряд элементов, не содержащих такого рода информации. Цифровой образ старопечатной книги, из которого исключены такие элементы, не всегда способен удовлетворить ожидания пользователей электронных библиотек.

Большинство электронных библиотек, рассматриваемых в данном докладе, стремится к созданию абсолютно полного цифрового образа старопечатной книги. Однако даже в цифровых образах старопечатных книг, включенных в состав таких библиотек, порой отсутствуют те или иные элементы, присущие оригиналу издания. В качестве наиболее часто отсутствующего элемента следует назвать корешок книжного переплета. Из рассматриваемых библиотек лишь Национальная библиотека Испании (в меньшей степени – Мировая цифровая библиотека и Национальная библиотека Франции) включает изображение корешка в цифровой образ старопечатной книги.

Как эпизодическое явление можно отметить наличие в рассматриваемых библиотеках цифровых копий старопечатных книг, не содержащих изображений переплета, авантитула, шмуцтитула, чистых страниц и т. д. Порой отсутствие в цифровом образе старопечатной книги того или иного элемента издания является не исключением, а правилом. Так, в электронных коллекциях Библиотеки им. Врублевских Литовской академии наук большинство цифровых копий не содержит изображения переплета, а в РГБ и вовсе стало тенденцией то, что просмотрные копии старопечатных изданий включают лишь элементы книги, являющиеся носителями символической информации.

Таким образом, большинство электронных библиотек осознает необходимость создания безоговорочно полного цифрового образа старопечатной книги. В то же время существуют библиотеки, не придающие надлежащего значения оцифровке элементов книги, не несущих текстовой информации. Представленные в таких электронных библиотеках копии старопечатных книг, без сомнения, можно назвать полнотекстовыми, но полными во всех смыслах слова они не являются.

5. Выбор подхода к созданию цифрового образа отдельной страницы старопечатной книги

Существуют два преобладающих подхода к созданию цифрового образа отдельно взятой книжной страницы. При использовании первого из них границы изображения устанавливаются с некоторым отступом от границ оцифровываемого документа. В этом случае образ книжной страницы включает фрагмент соседней страницы, во многих случаях содержит изображение фрагментов книжного обреза и переплета. При втором подходе границы изображения совпадают с краями книжной страницы. Таким образом, фрагменты соседней страницы, обреза книги и переплета из изображения исключаются.

При оцифровке книг, обладающих только информационной значимостью, предпочтение почти всегда отдается последнему из данных подходов. При создании же цифровых образов страниц изданий, представляющих художественную и культурно-историческую ценность (в том числе старопечатных книг) единства в использовании названных подходов не обнаруживается. Например, Мировая цифровая библиотека, Gallica, Hispánica, электронная библиотека Национальной парламентской библиотеки Японии последовательно придерживаются первого из них, а электронные библиотеки

ГПИБ России, Национальной библиотеки Армении являются столь же убежденными приверженцами второго.

Проблему выбора подхода к созданию цифрового образа отдельной страницы вполне уместно рассматривать в контексте выбора между предпочтениями «пользователя-читателя» и «пользователя-зрителя». Однако следует обратить внимание на иной аспект данной проблемы, касающийся разницы в восприятии пользователем тексто-информационной составляющей каждого из двух названных типов цифровых образов.

Нередко в старопечатных изданиях текстовая или иная символьная информация располагается на книжной странице в непосредственной близости от ее края. В том случае если границы цифрового образа такой страницы совпадают с границами самой страницы, при использовании копии может сложиться впечатление, что в ней отсутствует часть информации, имеющейся в оригинале (даже если в действительности это не так). Копии же страниц, выполненные с отступом от границ издания, не вызывают никаких подозрений в том, что часть информации могла быть потеряна в результате неаккуратности создателя цифрового образа. Вышесказанное является весомым аргументом в пользу цифровых образов страниц старопечатных изданий, выполненных с отступом от края книги.

Заключение

Проведенные автором исследования подтверждают, что подходы к созданию цифрового образа старопечатной книги, используемые участниками процессов оцифровки библиотечных фондов и создателями электронных библиотек из разных стран мира, заметно отличаются. Тем не менее в отношении рассмотренных в данном докладе частных аспектов существующих подходов можно сделать вывод, что в мировом масштабе с той или иной степенью отчетливости сформировалась тенденция к созданию: полноцветных цифровых образов старопечатных книг, пограничных образов старопечатных книг, абсолютно полных цифровых копий старопечатных книг, цифровых образов отдельной страницы с отступом от границ издания.

Целенаправленная деятельность по созданию цифровых копий документов, хранящихся в отделе редких книг и рукописей ЦНБ НАН Беларуси, началась в 2013 г. В числе документов, оцифровка которых осуществляется либо уже завершена, присутствуют и старопечатные книги: кириллические издания XVI–XVIII вв., книги гражданского шрифта, западноевропейские инкунабулы и т. д. С 2015 г. действует электронный репозиторий ЦНБ НАН Беларуси (<http://library.basnet.by>), в котором представлены и цифровые копии ряда старопечатных изданий.

Приступая к оцифровке старопечатных книг, ЦНБ НАН Беларуси имела возможность опираться на богатый опыт других библиотек, занимающихся оцифровкой собственных фондов, а также различных электронных библиотек мира. Проанализировав данный опыт, ЦНБ НАН Беларуси пришла к выводу о необходимости соответствия вышеизложенным принципам создания цифрового образа старопечатной книги.

Список литературы

1. Проект для общественного обсуждения «Единые технические требования к оцифровке изданий, включаемых в НЭБ» [Электронный ресурс] / Российская государственная библиотека, 2013. – Режим доступа: http://www.rsl.ru/datadocs/doc_7935ra.pdf. – Дата доступа: 02.05.2018.