

УДК 025:004+930.25:004

**И. В. Лобузин,**  
науч. сотр. НБУВ

### **ЦИФРОВОЙ ПРОЕКТ НАУЧНОЙ БИБЛИОТЕКИ: ТЕХНОЛОГИЧЕСКИЕ РЕШЕНИЯ И ОРГАНИЗАЦИЯ ДОСТУПА К ИНФОРМАЦИИ**

Предложены пути организации технологического цикла создания цифровых ресурсов научной библиотеки, их долгосрочного хранения и предоставления к ним доступа. Рассмотрены особенности формирования цифрового фонда научной библиотеки: научное описание ресурсов и интеллектуальный доступ к информации.

*Ключевые слова:* оцифровка историко-культурного наследия, цифровая библиотека, цифровые коллекции, научная библиотека.

Процессы развития информационно-коммуникационных технологий привели сегодня к значительным изменениям в природе информации и форме ее представления. Эти процессы также затрагивают вопросы сохранения мирового культурного наследия как основы всей человеческой цивилизации в целом. Современной формой обеспечения сохранности документных информационных ресурсов и реализации равных прав доступа к историко-культурному наследию является цифровая библиотека.

Анализ опыта международных и национальных цифровых инициатив дает возможность сделать следующие выводы:

1) ключевыми организациями в создании цифровых ресурсов национального уровня являются национальные библиотеки, поскольку в этих учреждениях накоплен значительный опыт по описанию и управлению документной информацией различного формата;

2) при создании цифровой библиотеки важным является определение технических и технологических стандартов на основе принятой на международном уровне системы требований, что в дальнейшем обеспечит успешную интеграцию цифрового ресурса в глобальные цифровые проекты;

3) уже на первых этапах создания масштабного цифрового ресурса необходимо определить информационную инфраструктуру проекта:

архитектуру информационного комплекса, его основные подсистемы, аппаратное, программное и кадровое обеспечение проекта;

4) в соответствии с инициативой цифрового кураторства, основными технологическими звеньями управления жизненным циклом цифрового библиотечного ресурса являются: подсистема создания цифровых копий, подсистема описания и учета цифровых ресурсов, подсистема обработки и управления цифровыми ресурсами; подсистема администрирования и хранения цифрового фонда; подсистема поиска и публикации цифровых ресурсов;

5) при формировании цифровой библиотеки необходимо:

✓ создать условия для надежного хранения больших объемов цифровой информации;

✓ определить основы долгосрочного сохранения и надежного доступа к созданному цифровому ресурсу;

✓ внедрить систему организации знаний для обеспечения совершенного доступа к оцифрованным документам;

✓ создать предпосылки для интеграции традиционных библиотечных ресурсов и информации цифровой библиотеки [1, 3, 5].

Таким образом, цифровой проект научной библиотеки предполагает создание сложного информационного комплекса, который будет организовывать взаимодействие технических средств, программного и информационного обеспечения, многих профильных специалистов. Объем работы по подготовке материалов для поддержания цифрового проекта очень велик, предстоит значительное изменение традиционных способов и самого содержания работы самой библиотеки и библиотечных специалистов. Технологическая цепочка процессов включает подготовку цифровой версии традиционного документа как физического объекта, его исследование и описание.

Кроме организации технологического цикла создания цифрового фонда необходимо также учитывать такие сопутствующие факторы:

◇ динамичное развитие информационно-коммуникационных технологий приводит к необходимости постоянно приспосабливать уже имеющиеся коллекции к новым программным разработкам, обеспечивать надежность носителей информации, сохранность записанных данных и гарантию их долгосрочного использования;

◇ обеспечение интеллектуального доступа к цифровым библиотекам стоит намного дороже самого процесса создания цифровых копий документов;

◇ информационные потребности современных пользователей требуют включения в цифровые коллекции архивных и современных информационных источников, что приводит к необходимости решения многочисленных задач, связанных с соблюдением авторских прав.

В большинстве случаев цифровой проект научной библиотеки предполагает обеспечение сохранности и организацию широкого доступа к уникальным документным ресурсам. Национальная библиотека Украины имени В. И. Вернадского (НБУВ) активно участвует в межнациональных и международных процессах создания историко-культурного и научного цифрового наследия. В НБУВ в 2010 г. создана группа по оцифровке документов на традиционных носителях, перед которой были поставлены две задачи: создание страхового фонда документов из фондов НБУВ и создание общедоступной цифровой библиотеки. Каждый из этих цифровых ресурсов имеет свои технологические особенности формирования [2].

Над созданием цифрового фонда НБУВ работает множество разных специалистов: научные сотрудники, библиотекари, специалисты в области информационно-коммуникационных и цифровых технологий. Для их совместной работы на общий результат необходимо было создать единый технологический комплекс формирования цифровых ресурсов. Основные принципы организации модульной системы формирования цифрового фонда в НБУВ разработаны согласно рекомендованным принципам управления цифровыми ресурсами (digital curation) [8].

*Подсистема создания цифровых копий.* Эта подсистема является аналогом технологии комплектования цифровой библиотеки. Задача отбора документов для оцифровки решается на основе предложений подразделений НБУВ – формируется план оцифровки и тематико-видовой план цифровых коллекций. Цифровые коллекции разделены на три основных типа: 1) документные коллекции (рукописи, архивные документы, редкие книги, старопечатные книги, журналы, газеты, ноты, карты, изобразительные материалы, листовки, календари); 2) тематические коллекции (история, государство и право, краеведение, этнография, языковедение, литературоведение, книговедение, библиотековедение, религия, наука, психология, философия, образование, культура, музыка, искусство); 3) отдельные цифровые проекты (Вернадский В. И., Шевченко Т. Г., нотная коллекция Разумовских, коллекция киноплакатов, «кужбушки» (учебные рисунки) Лаврской иконописной школы,

архитектура Украины на почтовых открытках, еврейский календарь и др.).

Подготовительный этап формирования плана оцифровки, включает распределение запланированного объема документов согласно имеющемуся в наличии оборудованию (характеристики и основные параметры оригиналов документов, пропускная возможность цифрового оборудования). Характеристики документов предварительно оцениваются специалистами группы технологий оцифровывания фондов, для каждого типа документов в базе данных предварительных заказов указывается соответствующее оборудование (сканер или цифровая фотокамера).

*Подсистема описания и учета цифровых ресурсов.* В НБУВ для организации системы учета работ по изготовлению и описанию цифровых ресурсов разработана служебная база данных на платформе АБИС на основе дополнения стандартных записей описаний документов необходимыми технологическими метаданными. База данных содержит записи двух видов: «цифровой ресурс» и «страховой фонд». Запись «цифровой ресурс» включает информацию об этапах и состоянии выполнения работ по созданию цифрового ресурса. Запись «страховой фонд» содержит описания документов, записанных на носитель долгосрочного хранения. Записи базы данных предоставляют возможность получать разноплановую отчетную документацию.

*Подсистема обработки и управления цифровыми ресурсами.* Цифровой проект НБУВ предусматривает создание страхового фонда цифровых копий документов и общедоступной цифровой библиотеки. Форматы изображений для хранения цифровых страховых копий документов НБУВ соответствуют следующим основным техническим требованиям: высокое разрешение мастер-копий (не менее 400 dpi); схемы сжатия данных без потерь; подробное документирование метаданных, соответствие оригиналу. Изображения, предназначенные для страховых копий, не подлежат дополнительной цифровой обработке и записываются на носители длительного хранения в оригинальном виде. Для создания пользовательских копий в НБУВ используется пакетная обработка изображений (уменьшение размера, обрезка, выравнивание и пр.), после чего собирается электронный аналог документа для публикации на сайте. Пользовательская копия сохраняется в двух форматах: pdf и swf (для просмотра с использованием flash-технологии).

*Подсистема администрирования и хранения цифрового фонда.* Система хранения в НБУВ решает три ключевые задачи: оперативный

доступ к информации, резервное копирование и архивное хранение. Оперативный архив реализован на основе RAID-массива класса RAID-5, который удовлетворяет требования скорости и надежности, обеспечивает средства резервного копирования и дублирования информации. RAID-массив используется также для хранения метаданных и учетной базы данных цифрового фонда НБУВ. Для организации области долгосрочного (архивного) хранения в НБУВ используются сменные носители информации: UDO-диски (объемом 30-60 Гб) с предположительным сроком хранения информации около 50 лет. Для хранения UDO-дисков организован их учет и специализированный сейф в помещении, где поддерживается необходимый температурно-влажностный режим.

*Подсистема поиска и публикации цифровых ресурсов.* Окончательной целью любого проекта по оцифровке историко-культурных и научных библиотечных информационных ресурсов является реализация доступа к созданному информационному массиву. Первостепенное внимание уделялось развитию когнитивных возможностей интерфейса цифровой библиотеки, относящихся к организации интеллектуального поиска информационных ресурсов. К описанию ресурсов привлечены экспертные специалисты из отделов-фондодержателей (библиографы, книговеды, документоведы и другие научные сотрудники), что обеспечило высокий научный и культурно-образовательный потенциал цифрового ресурса. Дополнительными средствами раскрытия содержания цифровых ресурсов, особенно для неструктурированной графической информации оцифрованных страниц, стало введение содержания документов, аннотаций, предметных рубрик, организация тематических коллекций [4].

Для обеспечения интеграции цифровых коллекций в мировые цифровые библиотеки (World Digital Library, Europeana) были разработаны средства экспорта записей в специальный XML-формат, основанный на базовом наборе метаданных Dublin Core.

Предложенный модульный цифровой комплекс был сформирован и апробирован в течение пяти лет в рамках цифрового проекта Национальной библиотеки Украины имени В. И. Вернадского, что позволило оцифровать более 0,5 млн изображений и предоставить пользователям около 10 тыс. электронных версий документов. Разработка отдельных модулей с определенной степенью автономными задачами создала предпосылки для детального исследования особенностей каждого

технологического звена, более рациональной организации и распределения профессиональных обязанностей специалистов библиотеки (рис. 1).

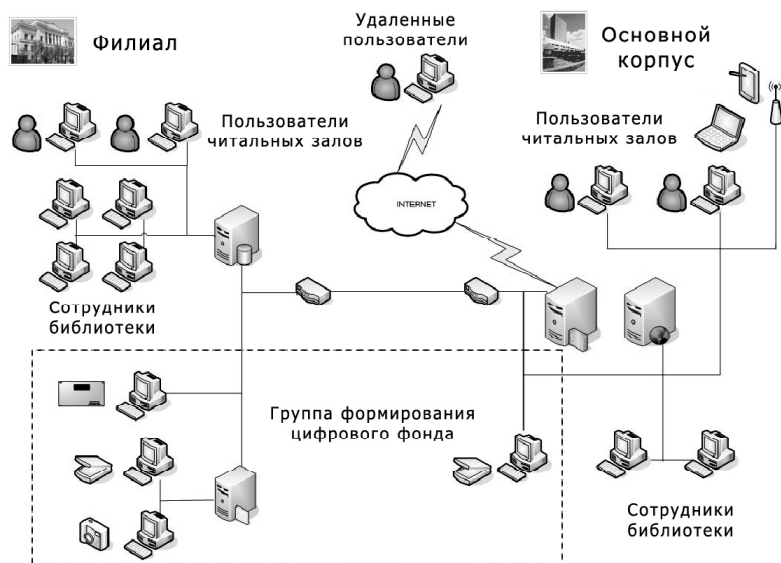


Рис. 1. Организация сетевого взаимодействия пользователей в процессе формирования и использования цифрового фонда НБУВ

Как подчеркивалось ранее, принципиальное отличие цифрового проекта научной библиотеки состоит в наличии научного описания документов и организации интеллектуального доступа к информации цифровой библиотеки. Ключевое звено в этом процессе – метаданные. Метаданные являются необходимым элементом структурированных цифровых коллекций для поддержки удобного поиска, длительного хранения и взаимодействия в интернет-масштабе. Для всех информационных цифровых объектов (физической или интеллектуальной формы) принято определять три основных атрибута: содержание, контекст и структуру. Именно эти атрибуты необходимо отразить через метаданные: содержание отвечает за информационные свойства объекта; контекст указывает, кто, что, почему, где и каким образом связан с созданием объекта и внешними факторами по отношению к информационному объекту; структура содержит формальный набор

атрибутов взаимосвязей между элементами внутри информационного объекта или различных информационных объектов.

От качества системы метаданных, ее продуманности и соответствия международным стандартам зависит успех и интероперабельность цифрового проекта научной библиотеки. Как показывает международный опыт, процесс изучения, отбора и развития системы метаданных цифрового фонда делится на несколько этапов: изучение основных потребностей в метаданных; рассмотрение и анализ существующих стандартов и проектов; подготовка спецификации метаданных и разработка системы метаданных; подготовка нормативно-методической документации; обучение сотрудников и поддержка службы метаданных [6, 7].

В соответствии с этими рекомендациями в НБУВ был осуществлен ряд исследований и разработок в направлении обеспечения цифрового проекта качественной и развитой системой метаданных. Сформированный в библиотеке цифровой ресурс имеет свои специфические особенности:

- ♦ полидокументность (рукописные и архивные документы, старинные и более современные книги, ноты, изобразительные материалы, карты, периодические издания и газеты),
- ♦ полилингвистичность (на разных языках, в том числе древних),
- ♦ политематичность (широкий тематический и информационный спектр ресурсов различных отраслей знания),
- ♦ полиформатность (различные размеры документов, шрифтов, изображений и широкая вариабельность их качества и четкости).

Для описания и атрибуции цифровых объектов в НБУВ была разработана оригинальная система метаданных, основанная на общепринятых международных стандартах (административных, технических, сохранности и описательных). Предложенная система метаданных решает технологическое сопровождение всех технологических звеньев формирования цифрового фонда, обеспечивает доступ к цифровым ресурсам на портале НБУВ и интеграцию метаданных с международными цифровыми проектами.

Административные метаданные обеспечивают управление цифровой коллекцией: права копирования и доступа, местонахождение документа и его копии, принадлежность к коллекции, критерии отбора для оцифровки, назначение копии (долговременное хранение или пользовательский доступ) и др. Технические метаданные характеризуют аппаратные средства и программное обеспечение, форматы данных, характеристики

изображений, ключи шифрования, пароли. Метаданные сохранности документируют физическое состояние ресурсов, принятые меры для сохранения физических и цифровых версий ресурсов, информацию о носителе и дату сохранения цифровых ресурсов. Описательные метаданные отвечают за содержательные характеристики ресурса и контекст его создания, они включают каталожные записи, специализированные тематические инструменты и индексы, гиперссылки и отношения между ресурсами, аннотации и содержание. В НБУВ процесс оцифровки документов сопровождается формированием соответствующих метаданных цифрового ресурса в служебной базе данных «SCAN». Записи базы данных позволяют следить за всеми этапами пополнения электронного фонда: от изготовления изображений до записи на диск UDO страхового фонда, создания и публикации электронной версии издания. Имеется также возможность получать разноплановую отчетную документацию: учет работ (по оборудованию, проектам, заказчикам и типу ресурсов), акты передачи цифровых копий заказчикам, списки оцифрованных документов, описания единиц хранения страхового фонда, индивидуальные отчеты по оцифровке и обработке цифровых копий.

После завершения технологического цикла создания и сохранения цифровой копии документа наступает этап публикации ее для свободного доступа в цифровой библиотеке историко-культурного наследия. Каждая цифровая библиотека ориентирована на свою аудиторию пользователей. Для научной библиотеки основная потребность в метаданных должна быть направлена на удовлетворение потребностей научного поиска информации. Поэтому проект научной библиотеки не может ограничиваться упрощенной схемой метаданных на уровне основных атрибутов (например, пятнадцать основных элементов схемы метаданных Dublin Core). Цифровые объекты научной библиотеки должны быть обеспечены расширенным спектром элементов описания содержания и контекста создания информационных объектов, критически важной информации для научных исследований. Поэтому на этапе публикации цифрового ресурса наибольшего внимания требуют описательные метаданные, которые в большинстве случаев являются единственными точками доступа к контенту, поскольку большинство цифровых копий не имеют распознанного текстового слоя и хранятся в формате изображений. Описательная часть метаданных в НБУВ включает описание библиотечных документов и документов архивных фондов. В структуре описания

документов (где это целесообразно) учтена связь между библиографическим и архивным описанием на уровне основных элементов: автор, название (заголовок), место издания (написания), дата публикации (создания), страна (локализация), язык документа, размер, формат, раздел знаний, предметная рубрика. Организация такой связи обеспечивает единый поисковый аппарат на уровне основных точек доступа для информационных ресурсов цифрового фонда. Среди главных элементов, определяющих уровень и качество описательных метаданных, следует выделить следующие: разработка типологии цифровых коллекций, тематическое упорядочение ресурсов, обеспечение поиска через авторитетные файлы (имен лиц и названий коллективов), хронологические и географические характеристики документов, язык текста, аннотации, исторические справки и содержание документов.

Для успешного представления проекта пользователям необходимо наладить тесное взаимодействие между различными специалистами библиотеки: дизайнерами и разработчиками информационной системы, специалистами по метаданным, экспертами. Благодаря совместной платформе реализации библиотечно-библиографической деятельности и цифровой библиотеки НБУВ имеет возможность привлекать экспертных специалистов из отделов-фондодержателей (библиографов, книговедов и архивистов, научных сотрудников) для получения полноценных профессиональных описаний документов электронных коллекций, что значительно повышает научный и культурно-образовательный потенциал цифрового ресурса.

Общие итоги использования ресурсов общедоступной «Цифровой библиотеки историко-культурного наследия» свидетельствуют о востребованности проделанной работы, несмотря на пока еще небольшой объем представленной пользователям информации (рис. 2). Так, с 2011 г. ресурсами «Цифровой библиотеки» воспользовались около 120 тыс. пользователей Украины, России, США, Польши, Беларуси, Германии, Канады, Израиля.

Как показал опыт цифрового проекта НБУВ, популярность цифровых ресурсов напрямую зависит от качества описательных метаданных, их реализации в интерфейсе пользователя, рекламы и популяризации новых поступлений в цифровую библиотеку. Информация о новых цифровых ресурсах автоматически обновляется на портале НБУВ в информационном блоке «Новые ресурсы и поступления», информационный



Рис. 2. Статистика использования ресурсов «Цифровой библиотеки историко-культурного наследия» на портале НБУВ (2011–2016 гг.)

баннер с обложками новых оцифрованных документов сопровождает поиск в электронном каталоге НБУВ. В ближайшем будущем запланированы работы по реорганизации и оформлению цифровых коллекций, их визуальному сопровождению. Особенно важным будет включение сформированного цифрового массива в новый проект НБУВ фундаментальной электронной библиотеки «Украиника», где можно будет раскрыть дополнительные информационные грани оцифрованных документов.

#### Список использованных источников

1. Лобузин І. В. Міжнародні та національні цифрові бібліотечні проекти: технологічні вимоги і стандарти [Електронний ресурс] / І. В. Лобузин // Бібліотека. Наука. Комунікація : матеріали міжнар. наук. конф. (Київ, 6–8.10.2015 р.) / НБУВ. – 2015. – Режим доступу: <http://conference.nbu.gov.ua/report/view/id/651>. – Назва з екрана.
2. Лобузин І. В. Технология обеспечения жизненного цикла цифровых ресурсов научной библиотеки / И. В. Лобузин // Науч. и техн. б-ки. – 2015. – № 9. – С. 3–10.
3. Лобузин І. В. Цифровая библиотека: проблемы интеграции в библиотечно-информационное пространство / И. В. Лобузин // Электронная библиотека : сб. науч. тр. – Санкт-Петербург : Президентская библиотека, 2013. – Вып. 4 : Научные и организационно-технологические основы интеграции цифровых информационных ресурсов. – С. 308–319.
4. Лобузін К. Технології організації знаннєвих ресурсів у бібліотечно-інформаційній діяльності : монографія / Катерина Лобузін ; відп. ред. О. С. Онищенко ; НАН України, Нац. б-ка України ім. В. І. Вернадського. – Київ, 2012. – 252 с.
5. Согласование национальных подходов к сохранению цифрового наследия / ред. Н. Макговерн ; ред. сер. К. Скиннер ; пер. с англ.: И. Н. Андреева, Н. Б. Богданова, Е. А. Губина, Д. Е. Осадчук. – Москва : МЦБС, 2013. – 359 с.

6. Chen Y.-N. A metadata lifecycle model for digital libraries: methodology and application for an evidence-based approach to library research [Electronic resource] / Ya-Ning Chen, Shu-Jiun Chen, Simon C. Lin // World Library and Information Congress : 69th IFLA General Conference and Council, 1–9 August 2003, Berlin. – Mode of access: [http://archive.ifla.org/IV/ifla69/papers/141e-Chen\\_Chen\\_Lin.pdf/](http://archive.ifla.org/IV/ifla69/papers/141e-Chen_Chen_Lin.pdf/). – Title from the screen.

7. Gilliland-Swetland A. J. Setting the stage: defining metadata / A. J. Gilliland-Swetland // Introduction to Metadata: Pathways to Digital Information. – Los Angeles, Calif. : Getty Information Institute, 1998. – P. 1–8.

8. Higgins S. Digital curation: the emergence of a new discipline / Sarah Higgins // The international journal of digital curation. – 2011. – Vol. 6, № 2. – P. 78–88.

Статья поступила 15.07.2016

#### Ivan Lobuzin

#### DIGITAL PROJECT OF RESEARCH LIBRARY: TECHNOLOGICAL SOLUTIONS AND ORGANIZATION OF ACCESS TO INFORMATION

The ways of organization of technological cycle of creating academic and research library digital resources, its long-term storage, and providing facilities of access are proposed. The features of academic and research library digital collection formation (resources scientific description and intellectual access to information) are considered.

*Keywords:* digitization of historical and cultural heritage, digital library, digital collections, academic and research library.