

# БИБЛИОТЕЧНО-ИНФОРМАЦИОННАЯ ДЕЯТЕЛЬНОСТЬ: ТЕОРИЯ И ПРАКТИКА

УДК 025:004+930.25:004

**И. В. Лобузин**

*Национальная библиотека Украины им. В. И. Вернадского*

## Технология обеспечения жизненного цикла цифровых ресурсов научной библиотеки

Рассмотрены технологические звенья жизненного цикла библиотечных цифровых ресурсов. Обоснованы состав и последовательность технологических процессов оцифровки библиотечных документов на основе опыта Национальной библиотеки Украины им. В. И. Вернадского; представлен её модульный цифровой комплекс, опробованный в течение четырёх лет.

**Ключевые слова:** цифровая библиотека, библиотечный цифровой ресурс, жизненный цикл цифрового ресурса, цифровой комплекс научной библиотеки, Национальная библиотека Украины им. В. И. Вернадского.

---

UDC 025:004+930.25:004

**Ivan Lobuzin**

*V. I. Vernadsky National Library of Ukraine, Kiev, Ukraine*

## Technology of life cycle support of research library's digital resources

Technological elements of library digital resources lifecycle are examined. Content and sequence of technological processes of library documents digitization are discussed as a case study of V. I. Vernadsky National Library of Ukraine; its module digital complex trialed for four years.

**Keywords:** digital library, library digital resources, digital resources lifecycle, digital complex of the research library.

---

Главными направлениями сохранения документального национально-го достояния сегодня являются организация оцифровки фондов библиотек,

создание цифровых коллекций, отражающих как национальную самобытность отдельных регионов, так и всё разнообразие мирового научного и культурного опыта. Однако следует отметить: несмотря на достаточно детально разработанные стратегии оцифровки фондов библиотек, отсутствует единый системный подход к организации самой технологии формирования библиотечных цифровых ресурсов.

Согласно главным принципам управления цифровыми ресурсами (*digital curation*) [6, 7], среди основных модулей (подсистем) информационной системы цифрового фонда, которые будут обеспечивать его стабильное и эффективное функционирование, целесообразно выделить следующие: создание цифровых копий; описание и учёт цифровых ресурсов; обработка цифровых ресурсов и управление ими; хранение цифрового фонда; поиск и публикация цифровых ресурсов.

Рассмотрим детально главные принципы организации модулей системы формирования цифровых ресурсов научной библиотеки на основе опыта оцифровки историко-культурных фондов в Национальной библиотеке Украины им. В. И. Вернадского (НБУВ).

### **Подсистема создания цифровых копий**

Эта подсистема – аналог технологии комплектования цифровой библиотеки, которая решает задачи отбора документов для оцифровки. Одна из основных проблем библиотек на этапе отбора документов для оцифровки связана с авторским правом [1, 3]. Исходя из жёстких ограничений украинского законодательства об авторском праве на создание любых копий произведений (только спустя 70 лет после смерти автора), цифровые коллекции НБУВ формируются на основе документов, перешедших в общественное достояние (уникальных, редких коллекционных книг и архивных документов).

Следует отметить положительный опыт ГПНТБ России по организации оцифровки фонда редких и ценных изданий [2]. В НБУВ также есть некоторый опыт по взаимодействию с издательствами научных журналов с целью получения готовых оцифрованных архивов (*backfiles*) с ретроспекцией номеров, которые не имеют оригинальных электронных аналогов (*born-digital*). Полученные файлы включаются на тех же принципах, что и новые поступления, в электронную библиотеку «Научная периодика Украины», снабжаются необходимыми метаданными для поиска и цитирования публикаций.

В формировании «Цифровой библиотеки историко-культурного наследия» (основной проект оцифровки в НБУВ) принимают участие фондодержатели уникальных документных фондов (Институт рукописи, отделы: библиотечных собраний и исторических коллекций, старопечатных книг и редких изданий, музыкальных фондов, изобразительных искусств, баз данных право-

вой информации). На основе предложений подразделений НБУВ формируются план оцифровки и тематико-видовой план цифровых коллекций.

Цифровые коллекции разделены на три типа: 1) документные (рукописи, архивные документы, редкие книги, старопечатные книги, журналы, газеты, ноты, карты, изобразительные материалы, листовки, календари), 2) тематические (история, государство и право, краеведение, этнография, языкознание, литературоведение, книговедение, библиотековедение, религия, наука, психология, философия, образование, культура, музыка, искусство), 3) отдельные цифровые проекты (Вернадский В. И., нотная коллекция Разумовских, коллекция киноплакатов, кужбушки (учебные рисунки) Лаврской иконописной школы, архитектура Украины на почтовых открытках, еврейский календарь и др.).

Наиболее сложной задачей оказалась организация взаимодействия и координирования подразделений-фондодержателей. В НБУВ с технологической точки зрения эта задача была решена на основе создания базы данных предварительных заказов на оцифровку документов, в которую специалисты соответствующих подразделений вносят описания документов; определяется их принадлежность к цифровым коллекциям, режим доступа к документам (страховая копия, локальный, онлайнный).

Следующий подготовительный этап формирования плана оцифровки предусматривает распределение запланированного объёма документов согласно имеющемуся оборудованию для оцифровки. Для этого необходимо определить характеристики и основные параметры оригиналов документов, цифровые копии которых необходимо изготовить [8]. К основным параметрам относятся: тип бумаги и способы изготовления, локальные недостатки, требующие предварительной реставрации или ретуширования, тональность, резкость (чёткость), габаритные размеры.

Также необходимо учитывать пропускную способность цифрового оборудования, которая в НБУВ составляет в среднем 100 тыс. страниц цифровых копий в год. Все эти характеристики документов предварительно оцениваются специалистами группы технологий оцифровывания фондов, для каждого типа документов в БД предварительных заказов указывается соответствующее оборудование (сканер или цифровая фотокамера).

### **Подсистема описания и учёта цифровых ресурсов**

Учёт и инвентаризация цифровых массивов, в том числе в сфере образования, науки и культуры, – предмет ряда национальных и международных проектов. Важным компонентом стратегий сохранения цифрового наследия являются метаданные (административные, технические, правовые, сохранности). В НБУВ для организации системы учёта работ по изготовлению и

описанию цифровых ресурсов разработана служебная БД на платформе Системы автоматизации библиотек (САБ) ИРБИС – на основе дополнения стандартных записей описаний документов необходимыми технологическими метаданными. БД содержит записи двух видов: «Цифровой ресурс» и «Страховой фонд» [5].

Запись «Цифровой ресурс» включает следующую информацию об этапах и состоянии выполнения работ по оцифровке фондов: даты работы, номер диска страхового фонда, подразделение, предназначение копии, наименование проекта или коллекции, тип документа, объём документа в страницах; описание документа (с необходимыми идентификаторами); выполнение заказа (оборудование, проведённые работы, дата выполнения, исполнитель); объём работ (название папки/файла, количество файлов, объём в байтах, формат файлов, разрешение, цветность, глубина цвета).

Запись «Страховой фонд» содержит описания документов на носителях долгосрочного хранения: номер диска страхового фонда, даты записи, подразделение, коллекция или проект, описание документа (номер заказа, идентификаторы, краткое описание, название папки/файла, количество файлов, объём в байтах, формат файлов, разрешение, цветность, глубина цвета). При формировании описания диска страхового фонда используется уже введённая для каждого документа информация. По этим данным можно определить необходимый для записи объём информации и распечатать акт учёта поступлений в страховой фонд.

Записи БД предоставляют также возможность получать разноплановую отчётную документацию: учёт работ (по оборудованию, проектам, заказчикам, типу ресурсов), акты передачи цифровых копий заказчикам, списки оцифрованных документов, описания единиц хранения страхового фонда, индивидуальные отчёты по оцифровке и обработке цифровых копий.

### **Подсистема обработки и управления цифровыми ресурсами**

В процессе подготовки цифровых копий уникальных и ценных печатных и рукописных объектов в соответствии с современными требованиями обеспечения их сохранности в будущем необходимо соблюдать технологическую политику – залог сохранения электронных копий в течение длительного времени: создание базовой (депозитарной) цифровой копии объекта; формирование на основе депозитарной копии необходимых для пользователя представлений цифрового объекта; возможность возврата с целью модификации и развития цифровой копии объекта в процессе его эксплуатации.

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. Важный этап создания эталонной цифровой копии документов – оценка полноты и качества полученных цифровых изображений (в первую очередь проверяется соответствие оригиналу, чёткость полученных изображений и цветовое соответствие). 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 применяется пакетная обработка изображений (уменьшение размера, обрезка, выравнивание и др.), 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). Flash-формат предусматривает возможность создания оригинального дизайна электронной публикации, подбор параметров качества изображения в зависимости от типа документа (размер шрифта, качество печати и др.), защиту изображений от копирования, масштабирование изображений, ввод оглавления и краткого описания документа, поиск по тексту, возможность загрузить документ в pdf-формате.

После изготовления пользовательских копий и записи дисков страхового фонда в технологической БД ставятся отметки о выполненных работах.

### **Подсистема хранения цифрового фонда**

С целью обеспечения сохранности цифровых данных больших объёмов в последние годы широкое распространение получила концепция управления жизненным циклом информации, в основе которой лежит принцип разделения общего массива данных на классы в зависимости от содержания, частоты обращений и срока хранения. В соответствии с этим подходом выделяются три ключевые задачи хранения цифровых данных: оперативный доступ к информации, резервное копирование и архивное хранение [4]. Система хранения и резервирования цифровых ресурсов в НБУВ организована согласно этим принципам.

Оперативный архив реализован на основе дискового массива *RAID-5*, который отвечает требованиям скорости и надёжности, обеспечивает средства резервного копирования и дублирования информации. *RAID-массив* используется также для хранения метаданных и учётной БД цифрового фонда НБУВ.

Для организации долгосрочного (архивного) хранения в НБУВ используются сменные носители информации: *UDO-диски* (объёмом 30–60 Гб),

которые поддерживают сверхплотный оптический формат (*Ultra Density Optical, UDO*). Технологии *UDO* популярны благодаря возможности записать на один носитель большой объём информации и предположительному сроку хранения информации – около 50 лет. Для хранения *UDO*-дисков установлен специальный сейф в хранилище рукописных документов, где поддерживается необходимый температурно-влажностный режим.

### **Подсистема поиска и публикации цифровых ресурсов**

Окончательная цель любого проекта по оцифровке историко-культурных и научных библиотечных информационных ресурсов – реализация доступа к созданному информационному массиву. Любая цифровая библиотека, предназначенная для определённого круга пользователей, должна обеспечивать поддержку конкретных видов деятельности.

Основную аудиторию пользователей «Цифровой библиотеки историко-культурного наследия НБУВ» составляют студенты, исследователи и научные сотрудники (историки, этнографы, краеведы, филологи). Поэтому первостепенное внимание уделялось развитию когнитивных возможностей интерфейса цифровой библиотеки, относящихся к организации интеллектуального поиска информационных ресурсов.

Благодаря совместной платформе реализации библиотечно-библиографической деятельности и цифрового проекта НБУВ на базе САБ ИРБИС реализована возможность привлечения экспертных специалистов из отделов-фондодержателей (библиографов, книговедов, документоведов и других научных сотрудников) для получения полноценных профессиональных описаний документов электронных коллекций, что значительно повысило научный и культурно-образовательный потенциал цифрового ресурса. Дополнительными средствами раскрытия содержания цифровых ресурсов, особенно для неструктурированной графической информации оцифрованных страниц, стали введение содержания документов, аннотаций, предметных рубрик, организация тематических коллекций, что значительно расширило поисковый потенциал созданного информационного массива.

В дальнейшем предполагается интегрировать в поисковый интерфейс цифровой библиотеки авторитетные файлы и электронный путеводитель по историко-культурным фондам НБУВ. Кроме создания цифровой библиотеки универсального содержания, на основе оцифрованных документов обеспечена возможность формирования отдельных цифровых проектов различного содержания (например, мемориальная коллекция «В. И. Вернадский», созданная к 150-летию со дня рождения выдающегося мыслителя).

Для интеграции цифровых коллекций в мировые цифровые библиоте-

ки (*World Digital Library, Europeana*) разработаны средства экспорта записей в специальный XML-формат, основанный на базовом наборе метаданных *Dublin Core*. Записи и цифровые копии документов прошли успешную апробацию в Библиотеке Конгресса США, что дало возможность представить уникальные документы из фондов НБУВ в рамках международного цифрового проекта.

## **Выводы**

Предложенный модульный цифровой комплекс сформирован и опробован в течение четырёх лет в рамках цифрового проекта НБУВ. Как показал накопленный опыт, каждая из подсистем (модулей) – создание, описание и учёт, обработка и управление, хранение, поиск и публикация цифровых ресурсов – соответствует определённому этапу жизненного цикла цифрового фонда научной библиотеки. Разработка отдельных модулей с автономными задачами создала предпосылки для детального исследования особенностей каждого технологического звена, более рациональной организации и распределения профессиональных обязанностей специалистов библиотеки.

Окончательно созданный цифровой комплекс и последовательность модулей с определённым кругом технологических процессов обеспечили:

решение двух главных задач цифрового проекта научной библиотеки: 1) создание цифровой библиотеки открытого доступа, 2) организация страхового фонда цифровых копий документов;

логическую последовательность выполнения операций технологического цикла;

необходимую технологическую инфраструктуру для выполнения всех задач, которые предусматривает цифровой проект научной библиотеки;

рациональное использование метаданных цифровых объектов на всех технологических этапах: создание – сохранение – учёт – публикация цифровых ресурсов;

обоснованное распределение функциональных обязанностей и ответственности сотрудников, принимающих участие в формировании цифрового ресурса;

развитую систему поиска информации в цифровой библиотеке, ориентированную на проведение научных исследований;

совместимость с общепринятыми стандартами и возможность интеграции созданных цифровых ресурсов в международные цифровые проекты.

## СПИСОК ИСТОЧНИКОВ:

1. **Антопольский А. Б.** Правовые и технологические проблемы создания и функционирования электронных библиотек / А. Б. Антопольский, Е. А. Данилина, Т. С. Маркова. – Москва : ПАТЕНТ, 2008. – 207 с.

**Antopolskiy A. B.** Pravovoye i tehnologicheskije problemy sozdaniya i funktsionirovaniya elektronnyh bibliotek / A. B. Antopolskiy, E. A. Danilina, T. S. Markova. – Moskva : PATENT, 2008. – 207 s.

2. **Евстигнеева Г. А.** Идеология оцифровки библиотечных фондов на примере ГПНТБ России / Г. А. Евстигнеева // Науч. и техн. б-ки. – 2014. – № 3. – С. 48–54.

**Evstigneeva G. A.** Ideologiya otsifrovki bibliotечnyh fondov na primere GPNTB Rossii / G. A. Evstigneeva // Nauch. i tehn. b-ki. – 2014. – № 3. – С. 48–54.

3. **Земсков А. И.** Возможности оцифровки фондов библиотек, обусловленные особенностями российского законодательства по авторскому праву (на примере фондов ГПНТБ России) / А. И. Земсков, Т. В. Яковлева // Там же. – 2013. – № 5. – С. 39–57.

**Zemskov A. I.** *Vozmozhnosti otsifrovki fondov bibliotek, obuslovlennye osobennostyami rossiyskogo zakonodatelstva po avtorskomu pravu (na primere fondov GPNTB Rossii)* / A. I. Zemskov, T. V. Yakovleva // *Tam zhe*. – 2013. – № 5. – С. 39–57.

4. **Корепанов И.** Как сохранить архив на десятилетия? [Электронный ресурс] / И. Корепанов // Журнал сетевых решений (LAN). – 2008. – № 3. – Режим доступа: <http://www.osp.ru/lan/2008/03/4899898/>.

**Korepanov I.** *Kak sohranit arhiv na desyatiletiya?* [Elektronnyy resurs] / I. Korepanov // *Zhurnal setevykh resheniy (LAN)*. – 2008. – № 3.

5. **Лобузин І. В.** Створення бази даних цифрового фонду бібліотеки: основні проблеми та технологічні рішення / І. В. Лобузин // Науково-технічна інформація (НТИ). – 2013. – № 1. – С. 67–62.

**Lobuzin I. V.** *Stvorenniya bazi danih tsifrovogo fondu biblioteki: osnovni problemi ta tehnologichni rishennya* / I. V. Lobuzin // *Naukovo-tehnichna informatsiya (NTI)*. – 2013. – № 1. – С. 67–62.

6. **Higgins S.** Digital Curation: The Emergence of a New Discipline // The International Journal of Digital Curation. – 2011. – Vol. 6, № 2. – P. 78–88.

7. **Smith A.** Strategies for Building Digitized Collections. / Digital Library Federation ; Council on Library and Information Resources. – Washington (DC), 2011. – 35 p.

8. **Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Image** [Электронный ресурс] / S. T. Puglia, J. Reed, E. Rhodes; U.S. National Archives and Records Administration (NARA). – 2004. – 87 p. – Режим доступа: <http://www.archives.gov/preservation/technical/guidelines.html>.

---

**Ivan Lobuzin**, researcher, Information and Communication department,  
V. I. Vernadsky National Library of Ukraine  
loboozin@nbuv.gov.ru  
3, 40-letiya Oktyabrya prospect, Kiev, Ukraine 03039