
БИЗНЕС-ИНФОРМАТИКА

DOI: 10.34020/2073-6495-2020-3-214-230

УДК 004.738.52:339(063)

НЕКОТОРЫЕ ОЦЕНКИ ТРУДОЗАТРАТ НА СОЗДАНИЕ ЭЛЕКТРОННЫХ БИБЛИОТЕК¹

Каленов Н.Е., Кириллов С.А., Соболевская И.Н., Сотников А.Н.

Научно-исследовательский институт системных исследований
Российской академии наук

E-mail: nekalenov@mail.ru, skirillov@jssc.ru,
nik_first@mail.ru, asotnikov@jssc.ru

В работе рассматриваются вопросы создания цифровых копий книг и 3D-моделей музейных предметов как составляющих интегрированных электронных библиотек. На примере электронной библиотеки «Научное наследие России» (ЭБ ННР) выделены и проанализированы технологические операции, выполняемые при формировании этих объектов библиотечными и музейными работниками, редакторами и техническими специалистами. На основании утвержденных норм на библиотечные и музейные работы и 10-летнего опыта наполнения ЭБ ННР рассчитаны временные трудозатраты на выполнение отдельных операций, формирование цифровых объектов и их коллекций. Проведенный анализ позволяет оптимизировать временные затраты на формирование элементов электронных библиотек путем распараллеливания технологических процессов.

Ключевые слова: электронные библиотеки, оцифровка книг, 3D-модели, музейные предметы, сканирование, трудозатраты, технологические процессы.

SOME ESTIMATES OF THE LABOR COSTS FOR CREATING ELECTRONIC LIBRARIES

Kalenov N.E., Kirillov S.A., Sobolevskaya I.N., Sotnikov A.N.

Scientific Research Institute for System Analysis
of the Russian Academy of Sciences

E-mail: nekalenov@mail.ru, skirillov@jssc.ru,
nik_first@mail.ru, asotnikov@jssc.ru

The issues of creation of books digital copies and 3D-models of museum objects as a content of integrated digital libraries are discussed in the paper. The process steps to take a digital content for digital library are presented and analyzed in terms of the Scientific Heritage of Russia digital library (SHR DL). Librarians, museum employees, editors and

¹ Работа выполнена в рамках государственного задания МСЦ РАН – филиале ФГУ ФНЦ НИИСИ РАН.

technical specialists carry out these steps. Time labor costs for performing certain operations such as the formation of digital objects and collections of these objects are calculated based on approved standards for library and museum works and on the basis of 10 years of provisioning experience with SHR DL. The analysis allows us to optimize the time spent on the formation of digital library elements by parallelizing technological processes.

Keywords: digital libraries, books capture, 3D-models, museum objects capture, scanning, labor expenditures, technological process.

ВВЕДЕНИЕ

Создание электронных (цифровых) библиотек является одним из интенсивно развивающихся направлений современной информатики. В мире существует огромное количество коллекций научных информационных ресурсов, называемых электронными библиотеками, по разным тематическим и видовым направлениям, с различными подходами к формированию их фондов (контента). Необходимо отметить, что хотя термин «электронная библиотека» используется в огромном количестве публикаций, единого четкого его определения пока не дано. Не касаясь определений этого термина, обсуждаемых 20 и более лет назад, приведем несколько вариантов, относящихся к более позднему времени.

В монографии 2007 г. [2] электронная библиотека определяется как информационная система, позволяющая надежно сохранять и эффективно использовать разнообразные коллекции электронных документов (текстовых, изобразительных, звуковых, видео и др.), локализованных в самой системе, а также доступных ей через телекоммуникационные сети.

Через 10 лет в обзорной работе белорусских коллег указывается, что «Электронные библиотеки (ЭБ) представляют собой формы сложных распределенных информационных систем, предоставляющих новые возможности работы с неоднородной информацией, и рассматриваются как основа создания глобального распределенного хранилища знания» [3, с. 133]. Правда, далее авторы замечают, что «общепринятого определения электронной библиотеки в настоящее время не существует».

Последнее утверждение соответствует действительности, доказательством чему служит то, что современные справочно-энциклопедические сетевые ресурсы определяют термин «электронные библиотеки» по-разному.

На сайте академических словарей [19] читаем: *«Электронная библиотека – распределенная информационная система, позволяющая надежно сохранять и эффективно использовать разнородные коллекции электронных документов через глобальные сети передачи данных в удобном для конечного пользователя виде».*

Сайт Юридической энциклопедии приводит другое определение [20]: *«ЭЛЕКТРОННАЯ БИБЛИОТЕКА – тематически ориентированная или структурированная иным образом среда доступа к удаленным или локальным электронным ресурсам, способная обслуживать электронными ресурсами локальных или удаленных пользователей (например, через Интернет)».*

Википедия [28] определяет электронную библиотеку как *«упорядоченную. Коллекцию разнородных электронных документов (в том числе книг, журналов), снабженных средствами навигации и поиска».*

Не вдаваясь в обсуждение приведенных определений, заметим, что, как бы ни трактовался термин «электронные библиотеки» (далее ЭБ), для реализации на практике «среды доступа», «информационной системы» или «упорядоченной коллекции» необходима специальная технология. Если проводить сравнение традиционных и электронных библиотек, то общим между ними является предоставление информационных ресурсов пользователям. И те, и другие библиотеки должны создавать справочный аппарат, позволяющий пользователю найти то, что его интересует среди ресурсов, которые предоставляют библиотеки. Для традиционных библиотек это – различного рода каталоги (в том числе электронные) со сложившимися в многолетней библиотечной практике элементами поиска, для электронных библиотек (в широком понимании этого термина) – базы метаданных с поисковым интерфейсом той или иной сложности. Существенным отличием электронных библиотек от традиционных является то, что последние используют готовую первичную информацию – печатные издания, опубликованные другими организациями. Формирование же электронных библиотек зачастую связано с необходимостью целенаправленной оцифровки тех или иных изданий. Это, в первую очередь, характерно для научных ЭБ, формируемых по узкотематическим или другим фиксированным принципам. К ним относятся ЭБ диссертаций, формируемая РГБ [29]; ЭБ Президентской библиотеки имени Б.Н.Ельцина, посвященная вопросам российской государственности [23]; ЭБ «Научное наследие России» [24] и ЭБ «Научное наследие Урала» [25], отдельных научных институтов (см., напр., [29, 23] и др.)

Создание и поддержка каждой из электронных библиотек требуют значительных трудозатрат. Часть из них, связанная с разработкой онтологии ЭБ (выбор структуры базы данных, определение классов объектов, включаемых в ЭБ, и их связей, профилей метаданных и форматов представления информации), созданием или адаптацией соответствующей программной оболочки, приобретением и установкой оборудования, является однократной. Трудозатраты, связанные с сопровождением ЭБ, требуются постоянно, и они связаны не только и не столько с обеспечением технической поддержки ЭБ, сколько с формированием контента, включающего как собственно цифровые объекты, так и их метаданные, обеспечивающие необходимое качество поиска.

Предметом данной статьи является оценка трудозатрат постоянного характера на поддержку электронных библиотек на примере ЭБ «Научное наследие России» [5, 7–9, 11].

Электронная библиотека «Научное наследие России» (ЭБ ННР) функционирует в промышленном режиме начиная с 2010 г. Основной целью создания Библиотеки является формирование, обеспечение сохранности и предоставление пользователям достоверной цифровой информации о выдающихся ученых, внесших вклад в развитие российской науки, и их научных достижениях. В ЭБ ННР отражаются биографические данные об ученых, наиболее значительные их публикации (библиография и отсканированные полные тексты), архивная и музейная информация, относящаяся к деятельности ученого. Библиотека включает текстовую информацию, оцифрованные печатные издания, архивные документы, фотографии и кинофильмы, 3D-модели музейных предметов.

К настоящему времени в ЭБ ННР представлена информация о более чем 6100 ученых, работавших в России с XVIII по первую четверть XX в., оцифровано и доступно пользователям около 25 000 книг, изданных за этот период.

На начальном этапе создания ЭБ ННР, которая разрабатывалась в рамках целевой программы Президиума РАН, были выработаны основные принципы разработки и функционирования Библиотеки, определены правила формирования контента, технология подготовки данных участниками проекта.

В основу функционирования ЭБ ННР положен принцип распределенной подготовки данных с централизованной редакторской обработкой, загрузкой контента и поддержкой технологии. Информацию для ЭБ готовят по единым правилам более 20 библиотек (в том числе центральные академические библиотеки – БАН, БЕН РАН, ЦНБ УрО РАН, библиотека ИНИОН РАН), институтов и музеев. Координацию работ по формированию и сопровождению фондов ЭБ ННР, включая редактирование и загрузку данных, осуществляют сотрудники МСЦ РАН – филиала ФГУ ФНЦ НИИСИ РАН.

В задачи поставщиков контента входит отбор материалов в соответствии с установленными принципами, формирование метаданных о принятых к включению в ЭБ объектах (персоналии, издания, архивные документы, музейные предметы, фотографии, мультимедийные материалы), оцифровка изданий и обработка информации в соответствии с правилами системы (для ЭБ ННР принято решение, согласно которому отсканированный текст не распознается, за исключением оглавления, по которому обеспечивается навигация внутри издания), передача обработанных материалов в редакторскую группу.

Редакторская группа принимает окончательное решение о включении в ЭБ тех или иных изданий, предложенных поставщиками, осуществляет «выходной контроль» подготовленной информации, включающий проверку правильности метаданных и качества (постранично) каждого оцифрованного издания, подготавливает прошедшие контроль издания для загрузки в программную оболочку демонстрационной части ЭБ, доступную пользователям по адресу <http://e-heritage.ru>.

Объектно-ориентированный подход, выбранный при проектировании ЭБ ННР, использование технологии распределенной подготовки данных, отражение в Библиотеке различных цифровых объектов научного назначения и положительный опыт многолетней эксплуатации позволяют рассматривать ЭБ ННР как возможный прототип Единого цифрового пространства научных знаний (ЕЦПНЗ) [1]. Создание ЕЦПНЗ является одной из важнейших задач в области развития системы информационного обеспечения научных исследований, научно-образовательной и научно-популярной деятельности, сохранности научных знаний.

В этой связи представляется важным, исходя из опыта эксплуатации ЭБ ННР, построить модель оценки трудозатрат при формировании контента ЕЦПНЗ. Модель построим на примерах наиболее трудоемких процессов, которыми являются оцифровка печатных изданий (с формированием электронных книг) и создание 3D-моделей музейных объектов.

1. ПОДГОТОВКА ЦИФРОВЫХ КОПИЙ КНИГ

Включение в контент ЭБ ННР книг предполагает, что на первом этапе в библиотеку вводится информация об их авторах – ученых, сыгравших важную роль в развитии российской науки. В соответствии с профилем метаданных персон, принятым в ЭБ ННР, в библиотеку вводятся биографические данные об ученых, их научные интересы в терминах ГРНТИ [17], библиография их основных работ. Эту работу выполняют информационно-библиотечные работники. Она включает три этапа – поиск источников биографических данных ученого и составление развернутой биографии, подбор библиографии, ввод данных в технологический блок ЭБ ННР. Обозначим средние временные затраты на выполнение каждого этапа соответственно через t_p^1, t_p^2, t_p^3 .

Технологические процессы, выполняемые при подготовке издания к включению в ЭБ ННР, и их последовательность представлены на рис. 1.

В табл. 1 представлены те же процессы с указанием, какие специалисты их выполняют, что является единицей их учета и обозначением времени выполнения каждого процесса.

Таким образом, если в ЭБ ННР вводится книга объемом N страниц, общие временные затраты T_B на ее включение в Библиотеку составят:

$$T_B = \sum_{i=1}^7 t_k^i + N \cdot \sum_{i=1}^3 t_s^i. \quad (1)$$

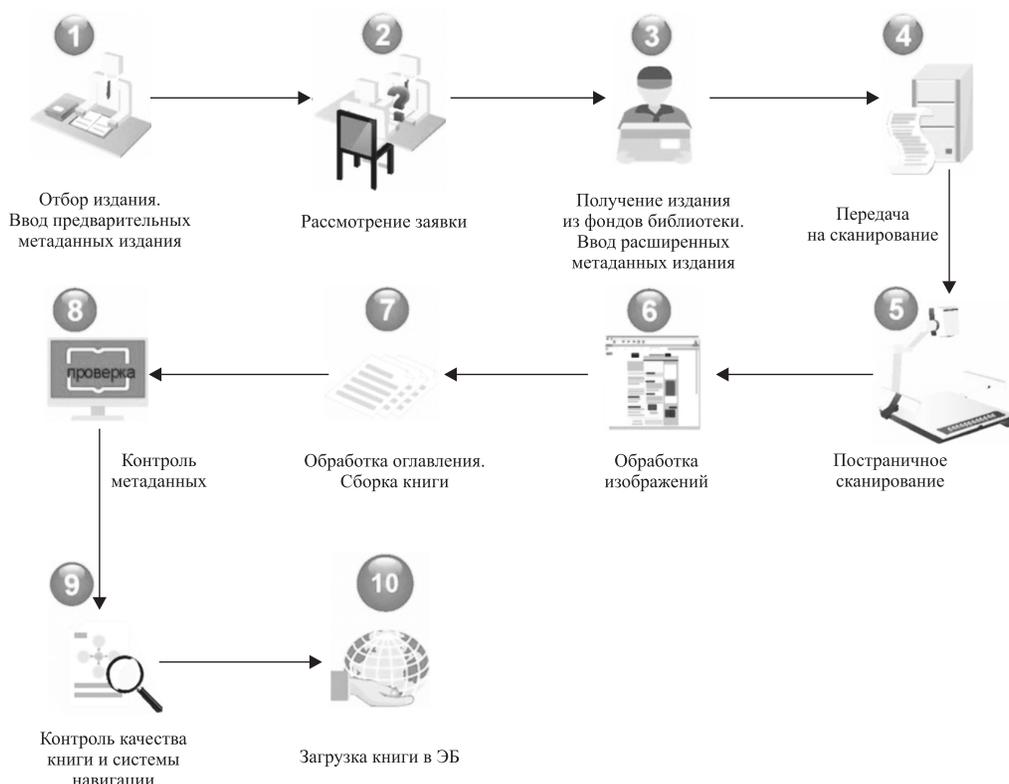


Рис. 1. Последовательность технологических процессов, выполняемых при подготовке издания к включению в ЭБ ННР

Таблица 1

Процессы подготовки издания к включению в ЭБ ННР

Номер этапа	Содержание работы	Кем выполняется	Единица учета	Время
1	Отбор издания, предлагаемого к включению в ЭБ, ввод в технологический блок предварительных метаданных издания, включая связи с персонами	Библиотечный работник организации – участницы проекта	Книга	t_k^1
2	Рассмотрение заявки	Сотрудник редакторской группы	Книга	t_k^2
3	Получение издания из фондов библиотеки, ввод расширенных метаданных издания (включая присвоение книге рубрикационного индекса)	Библиотечный работник организации – участницы проекта	Книга	t_k^3
4	Передача на сканирование, подготовка к сканированию книги	Библиотечный работник организации – участницы проекта	Книга	t_k^4
5	Постраничное сканирование	Оператор-сканировщик	Страница	t_s^1
6	Обработка изображений	Технический специалист	Страница	t_s^2
7	Обработка оглавления, сборка книги	Технический специалист	Книга	t_k^5
8	Контроль метаданных книги	Редактор	Книга	t_k^6
9	Постраничный контроль качества книги и системы навигации	Редактор	Страница	t_s^3
10	Загрузка книги в ЭБ (установление связи макета книги с метаданными)	Технический специалист	Книга	t_k^7

Для оценки реальных временных затрат библиотечных специалистов на выполнение перечисленных работ воспользуемся «типовыми отраслевыми нормами труда на работы, выполняемые в библиотеках», утвержденными министром культуры РФ в конце 2014 г. [18]. Хотя в этом документе в явном виде не присутствуют процессы, перечисленные выше, но, объединив ряд более мелких работ, можно получить оценочные величины искомых трудозатрат.

Формирование сведений об ученом, отражаемых в ЭБ ННР, как указывалось выше, включает три временных интервала.

Для реализации первого этапа (временной интервал t_p^1) необходимо выполнение следующих операций: отбор авторитетных изданий (в терминологии [18] – первичных документов), содержащих сведения об ученом, заказ их в фондах библиотеки, доставку из хранилища, выдачу пользователю (в данном случае работнику, составляющему биографию ученого), составление биографии ученого на основе информации из полученных изданий, возврат изданий в фонды. Временные затраты на отбор изданий можно оценить с помощью нормы «Выполнение тематической справки; поиск и отбор документов». Она приводится в расчете на группу отобранных документов – для группы от 2 до 5 документов (анализ данных ЭБ ННР показывает, что в среднем при составлении биографии ученого используется от 2 до 3 источников) составляет 15 мин.

Технологические библиотечные операции, связанные с выдачей и приемом изданий из фондов, нормируются в расчете на одно издание и составляют в сумме 13 мин. Оценим их, применительно к нашей задаче (выдается в среднем немного больше 2 изданий), в 30 мин.

Для оценки времени, затрачиваемого на составление биографии ученого², воспользуемся нормой «составление реферата: изучение и анализ документа, на который составляется реферат; написание текста», приравнивая условно составление биографии к составлению реферата отобранных изданий). Эта норма приводится в расчете на один авторский лист (40 000 знаков) и составляет 5920 мин. Анализ данных, отраженных в ЭБ ННР, показывает, что объем текста биографии ученого колеблется от 1000 до 31 000 знаков и составляет в среднем около 6000 знаков или 15 % печатного листа. Таким образом, нормативное время на составление биографии ученого и ввод ее в систему составит 888 мин, а суммарное время на выполнение первого этапа формирования данных об ученом $t_p^1 = 15 + 30 + 888 = 933$ мин.

Временные затраты на выполнение второго этапа (формирование библиографического списка публикаций ученого) можно оценивать на основании нормы по составлению библиографического указателя, которая приводится в расчете на один авторский лист и составляет 13 500 мин. Объем библиографии ученого существенно зависит от периода его научной деятельности – среднее количество публикаций одного ученого выросло за последнее столетие в несколько раз. Анализ данных, введенных в ЭБ ННР, показывает, что библиографический список одного ученого в среднем составляет 2200 знаков или 5,5 % авторского листа. В соответствии с нормами на его составление требуется 742 мин.

Ввод структурированных данных об ученом можно интерпретировать как приведенную в [18] операцию «набор на клавиатуре сведений о читателе: фамилия, имя, отчество, характеризующие его признаки (образование, специальность, иная информация». На эту операцию, согласно нормам, отводится 6 мин.

Таким образом, общие временные затраты на формирование в ЭБ информации об одном ученом ($T_p = t_p^1 + t_p^2 + t_p^3$) составляют 1681 мин или (округленно) 28 ч работы библиотечного специалиста.

При оценке трудозатрат библиотечных работников t_k^1 , t_k^3 и t_k^4 воспользуемся наряду с уже рассмотренными нормами на отбор литературы, нормы на «формирование библиографической записи на документы на русском языке (описательная каталогизация)» (18 мин на документ), «индексирование документов (содержательная каталогизация)» (18 мин) и «ввод в компьютер основных сведений о документе (автор, заглавие) в специализированной программе» (5 мин), «подготовка документов на микрофильмирование и сканирование» (5 мин), «передача документов на микрофильмирование и сканирование» (16 мин). В результате получим следующие данные: $t_k^1 + t_k^3 + t_k^4 = 75$ мин.

² Биографии ученых составляются библиотечными специалистами на основании анализа энциклопедических данных и публикаций, посвященных данному ученому.

Рассмотрим процессы (обозначенные в качестве этапов в табл. 1), выполняемые сотрудниками редакторской группы, операторами-сканировщиками и техническими специалистами. За основу возьмем опыт работы по наполнению ЭБ ННР и нормы на сканирование документов бесконтактным способом (именно такая технология используется в ЭБ ННР), представленные в [12].

На этапе рассмотрения заявки (этап 2 в табл. 1) сотрудник редакторской группы принимает решение о целесообразности ввода в ЭБ книг, предложенных участниками проекта. Для этого выполняется проверка книг по следующим параметрам: соответствие издания профилю библиотеки, соблюдению законов об авторском праве, выявление дублирования, правил оформления библиографической записи. Затем электронной книге присваивается уникальный номер (ID). Если зарегистрировать предложенную книгу не представляется возможным, выставляется отметка с указанием причины такого решения. Норма на одного сотрудника составляет 30 книг в смену. Исходя из этого, получаем

$$t_k^2 = 16 \text{ мин.}$$

В соответствии с [12] норма на одного оператора при постраничном сканировании (этап 5) составляет 800 страниц в смену. Это означает, что

$$t_s^1 = 0,6 \text{ мин.}$$

Основная задача 6-го этапа (обработка изображений) – проверить и отредактировать графические образы страниц электронных книг. Она включает в себя три технологических процесса: автоматическая обработка сканов с помощью специальной программы, проверка автоматической обработки и ручная коррекция.

В результате выполнения первого процесса происходит исправление типичных дефектов сканирования страниц до приемлемого уровня.

Второй процесс включает в себя:

- проверку последовательного отображения страниц (нумерация страниц должна быть последовательной, поиск пропущенных при сканировании страниц);
- проверку качества сканирования (степень читаемости текста, не менее 98 % информации, представленной на странице, должно быть читаемо);
- проверку качества автоматической обработки отсканированных страниц (корректная обрезка страниц, геометрическая коррекция текста, изгибов текста и иных искажений);
- простейшую правку отсканированных страниц (обрезка, удаление посторонних элементов).

В рамках третьего процесса производится ручная обработка страниц в одном из графических редакторов. Этот этап предусмотрен для наиболее сложных изданий, содержащих многочисленные формулы, таблицы, иллюстрации и т.п. Он включает в себя:

- геометрическую коррекцию текста, изгибов текста и иных искажений;
- удаление посторонних элементов на страницах электронных книг (пальцы оператора, полосы, тени и другие посторонние элементы);
- цветовую коррекцию.

Норма на одного оператора при выполнении этого этапа составляет 800 страниц в смену. Таким образом

$$t_5^2 = 0,6 \text{ мин.}$$

Основными задачами 7-го этапа являются: формирование оглавления книги (распознавание и редактирование текста или его ручной ввод); верстка электронной книги в специальной программе на основе подготовленных качественных графических образов страниц и сформированного оглавления; создание максимально точной навигационной системы электронного издания. В процессе создания навигационной системы технический специалист должен обеспечить:

- правильность набора текста, заголовков, примечаний и других частей навигационной системы;
- правильность оформления электронных ссылок и точность навигационной системы;
- комплектность электронной книги: порядковую нумерацию страниц, порядок разделов.

Норма выработки на одного специалиста составляет 5 электронных книг в смену.

$$t_k^5 = 96 \text{ мин.}$$

Этап 8 (контроль метаданных книги) включает:

- проверку соответствия фамилий авторов, заголовка, перевода названия на русский язык, выходных данных тем, которые имеются на титульном листе книги;
- проверку оформления записей – орфографии, пунктуации, принятых сокращений слов в библиографических данных;
- проверку соответствия сведений, внесенных в поля «вид издания», «язык», «страницы», оригиналу. Поле «страницы» выверяется строго по электронной версии книги и включает в себя общее количество файлов в электронной версии, подготовленной к загрузке на сайт, проверку наличия индекса ГРНТИ, его соответствия тематике публикации;
- проверку оформления библиографического описания (согласно ГОСТ 7.1–2003).

Норма выработки на одного специалиста составляет 10 электронных книг в смену, откуда следует:

$$t_k^6 = 48 \text{ мин.}$$

На этапе 9 (постраничный контроль качества электронной книги) осуществляется проверка выпускающим редактором верстки электронной книги на рабочем сервере. Работа редактора предусматривает анализ графических образов страниц и проверку системы навигации. Она включает:

- проверку последовательного отображения страниц;
- проверку качества сканирования (степень читаемости текста, не менее 99 % информации, представленной на странице, должно быть читаемо);
- проверку качества обработки отсканированных страниц (корректная обрезка страниц, геометрическая коррекция текста, отсутствие изгибов текста и иных искажений, отсутствие «посторонних элементов» – полос, теней, следов пальцев операторов и т.п.);

- проверку ссылок на предмет их открытия;
- проверку ссылок на соответствие главам и содержанию книги.

При выявлении тех или иных недостатков соответствующая информация передается оператору 6-го этапа. Норма на эти работы составляет 1200 страниц в смену, исходя из этого, получаем

$$t_s^3 = 0,4 \text{ мин.}$$

На заключительном этапе выпускающий редактор осуществляет публикацию книги и метаданных на портал электронной библиотеки и проверяет доступность загруженной информации. Норма выработки на одного специалиста составляет 50 электронных книг в смену,

$$t_k = 9,6 \text{ мин.}$$

Подставив полученные значения в формулу (1), получим, что средние временные затраты на оцифровку и включение в ЭБ одной книги объемом N страниц составят (мин)

$$T_B = 244,6 + 1,6 \cdot N. \quad (2)$$

Библиотечные работники из этого времени тратят

$$T_L = 75,$$

редакторы

$$T_R = 64 + 0,4 \cdot N,$$

технические специалисты

$$T_c = 105,6 + 0,6 \cdot N,$$

операторы-сканировщики

$$T_0 = 0,6 \cdot N.$$

Для подготовки и ввода в ЭБ первой книги ученого, ранее не отраженного в ЭБ, объемом 200 страниц потребуется около 38 ч, в том числе ~29,5 ч работы библиотечных специалистов, ~2,5 ч работы редактора, ~2 ч работы оператора-сканировщика, ~4 ч работы технического специалиста. При вводе других книг этого автора потребности в работе библиотечных специалистов сократятся до полутора часов, а общее время подготовки книги составит около 10 ч.

2. ПОДГОТОВКА ЦИФРОВЫХ 3D-МОДЕЛЕЙ МУЗЕЙНЫХ ОБЪЕКТОВ

Наряду с цифровыми публикациями ЭБ ННР содержит мультимедийный контент и, в частности, 3D-модели музейных объектов. Эти объекты могут быть привязаны к определенной персоне (или нескольким персонам), а могут быть объединены в самостоятельную коллекцию, посвященную, в том числе некоторому научному направлению или событию. Ниже рассмотрены трудозатраты на создание цифровой 3D-модели одного объекта, а также на формирование цифровых коллекций, включающих несколько таких моделей.

Полное, качественное и подробное цифровое 3D-моделирование объекта достаточно сложная задача, требующая использования мощных вычислительных ресурсов и специального оборудования. Построение так называемой полноценной цифровой 3D-модели музейного объекта предполагает наложение на нее текстуры и цвета. Кроме того, даже если такая полноценная модель построена, возникает задача погружения объекта в электронную библиотеку и обеспечение его визуализации, позволяющей просматривать эту модель через наиболее распространенные интернет-браузеры.

Для визуализации трехмерного объекта используются различные методы [6]. Эти методы могут быть основаны на SfM технологии [13, 14, 16], на программных и технологических решениях, используемых, в частности, при лазерном и оптическом 3D-сканировании, или методе фотограмметрии.

Оптические 3D-сканеры используются для быстрой оцифровки различных мелких и средних предметов, так как одновременно могут оцифровать множество точек или все «поле зрения» сканера. Для оцифровки геометрии и захвата текстуры при низком освещении используются специальные лампы белого или синего цветов в качестве источников света. Одним из недостатков оптических 3D-сканеров является невозможность построения объектов со светоотражающей или светопоглощающей структурой.

Лазерные 3D-сканеры обеспечивают большую, по сравнению с оптическими 3D-сканерами, точность и детализацию при оцифровке объектов. Однако для точной пространственной привязки лазера к сканируемому объекту необходимо применять специальные маркеры, которые крепятся в непосредственной близости от объекта или непосредственно на объекте сканирования, что далеко не всегда приемлемо для музейных объектов. Кроме того, основным недостатком сканера является неспособность передавать цвет текстуры объекта. Тем не менее задачу по построению максимально детальной поверхности сканируемого объекта (без необходимости визуализации текстуры и цвета) лазерные 3D-сканеры решают вполне удовлетворительно.

В отличие от технологий 3D-сканирования метод фотограмметрии [4, 10] позволяет построить высококачественную 3D-модель с передачей текстуры и цвета объекта. Однако в вычислительном плане построение 3D-моделей из набора изображений методом фотограмметрии – достаточно трудоемкая задача. Например, обработка 124 фотографий на одном из узлов кластера МВС-10П [26], установленного в МСЦ РАН, заняла 41 ч расчетов [15].

Для формирования цифровых 3D-моделей в ЭБ ННР была разработана технология интерактивной анимации [15]. Эта технология не предполагает построения полноценной 3D-модели, а основана на программной смене (прокручивании) фиксированного набора видов объекта (кадров) с помощью специализированных интерактивных программ отображения, имитирующих смену точки взгляда на исходный объект. Для создания такого интерактивного мультфильма необходим набор заранее снятых сцен, которые будут использоваться как отдельные кадры экспозиции.

Прежде чем приступать к формированию цифровых 3D-моделей музейных объектов с целью погружения их в электронную библиотеку, необ-

ходимо провести определенную подготовительную работу, выполняемую сотрудниками музея, которому принадлежит моделируемый объект. Эта работа включает подбор объекта для оцифровки с составлением соответствующей документации; осмотр объекта на предмет сохранности с составлением протокола осмотра или записью в книгу учета музейных предметов; выдачу объекта на оцифровку.

Нормативное время T_0 , требуемое на выполнение этого вида работ можно оценить, пользуясь «Типовыми отраслевыми нормами труда на работы, выполняемые в зоопарках, фильмофондах, музеях и других организациях музейного типа», утвержденными приказом Министерства культуры РФ [27]. Оно составляет в среднем 130 мин на один объект.

После выполнения этих подготовительных работ, начинается основной цикл работ по созданию цифровой 3D-модели выбранного музейного объекта. Технологические процессы, выполняемые при подготовке 3D-модели музейного объекта к включению в ЭБ ННР, и их последовательность представлены на рис. 2.

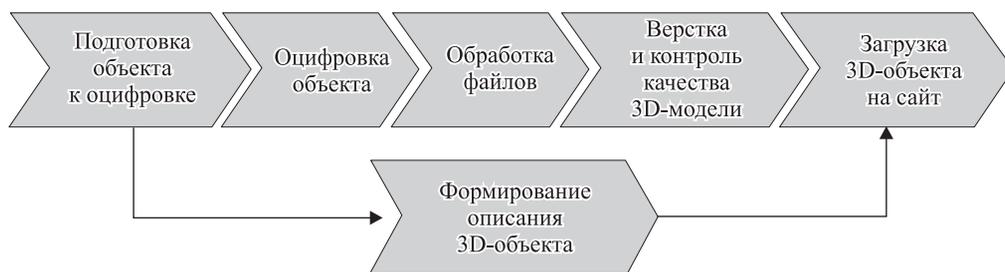


Рис. 2. Технологические процессы, выполняемые при подготовке 3D-модели музейного объекта к включению в ЭБ ННР

Этот цикл работ включает следующие основные этапы:

1. Подготовка к оцифровке – установка объекта на место съемки, наладка освещения и т.п.

2. Оцифровка объекта. Для выполнения этой работы применяется специальный комплекс на основе поворотной платформы Rekam T-50 и цифрового фотоаппарата Canon EOS600D. Конечным результатом выполнения этого этапа является массив данных, содержащий файлы с фотографиями объекта, снятыми со 120 ракурсов.

3. Обработка полученного на первом этапе массива данных. На этом этапе с каждой фотографии убирается фон, на котором снято изображение. Это делается с помощью программного модуля, специально разработанного для данного этапа.

4. Верстка и контроль качества изображения цифрового ресурса. Конечным результатом этого этапа являются цифровые 3D-образы музейных предметов.

5. Описание музейного предмета, цифровая 3D-модель которого включается в ЭБ. Эта работа выполняется сотрудниками музея.

6. Загрузка сформированной модели в ЭБ ННР.

Пусть $T_1, T_2, T_3, T_4, T_5, T_6$ – временные интервалы, необходимые для обработки одного музейного объекта на 1–6 этапах соответственно (табл. 2).

Таблица 2

Процессы подготовки 3D-модели музейного объекта к включению в ЭБ ННР

Номер этапа	Содержание работы	Исполнитель	Единица учета	Время
1	Подготовка объекта к оцифровке	Сотрудник музея	Музейный объект	T_1
2	Оцифровка отобранного объекта	Технический специалист МСЦ РАН	Папка, содержащая 120 файлов формата jpg для каждого сфотографированного объекта	T_2
3	Обработка полученных в результате оцифровки файлов	Технический специалист МСЦ РАН	Сформированные файлы	T_3
4	Верстка и контроль качества изображения цифрового ресурса	Технический специалист МСЦ РАН	Цифровой 3D-объект	T_4
5	Формирование описания цифрового 3D-объекта	Сотрудник музея	Цифровой 3D-объект	T_5
6	Загрузка 3D-объекта в ЭБ ННР	Технический специалист МСЦ РАН	Цифровой 3D-объект	T_6

Если в ЭБ ННР вводятся M цифровых музейных 3D-объектов, то время $T_{\text{ср}}$ на включение этого объема цифровых ресурсов в ЭБ ННР составляет:

$$T_{\text{ср}} = M \cdot \sum_{i=0}^6 T_i.$$

После того, как оцифровано несколько объектов, они могут быть объединены в одну или несколько коллекций. Пусть T_k – время, необходимое на формирование и описание коллекции. Тогда общее время T , требуемое на формирование цифровой коллекции музейных 3D-объектов, составляет:

$$T = T_{\text{ср}} + T_k.$$

Приведенные ниже числовые значения средних временных затрат на формирование цифровых 3D-моделей музейных предметов основаны на опыте формирования контента ЭБ ННР. В процессе пополнения библиотеки было подготовлено более 100 3D-моделей музейных предметов, объединенных в несколько коллекций. Среди них цифровая 3D-коллекция макетов плодов И.В. Мичурина, хранящаяся в Государственном биологическом музее имени К.А. Тимирязева (ГБМТ), цифровая 3D-коллекция антропологических реконструкций М.М. Герасимова, хранящаяся в ГБМТ и Государственном Дарвиновском музее. Средние временные значения $T_1, T_2, T_3, T_4, T_5, T_6$, приведенные ниже, основаны на опыте формирования, в том числе этих коллекций.

Для реализации первого этапа (подготовка объекта к оцифровке, временной интервал T_1) необходимо в среднем 45 мин.

Для реализации второго этапа (оцифровка отобранного контента, временной интервал T_2) необходимо в среднем 20 мин на один объект.

Для реализации третьего этапа (обработка полученных в результате оцифровки файлов, временной интервал T_3) необходимо в среднем 290 мин на один объект.

Для реализации четвертого этапа (верстка и контроль качества изображения цифрового ресурса, временной интервал T_4) необходимо в среднем 25 мин на один объект.

Для реализации пятого этапа (описание цифрового 3D-объекта, временной интервал T_5) необходимо в среднем 15 мин на один объект.

Для реализации шестого этапа (загрузка 3D-объекта в ЭБ ННР, временной интервал T_6) необходимо в среднем 35 мин на один объект.

Таким образом, общие временные затраты на представление в ЭБ ННР одной цифровой 3D-модели музейного объекта составляют:

$$T = 45 + 20 + 290 + 25 + 15 + 35 + 130 = 560 \text{ мин.}$$

Для формирования и описания одной коллекции, насчитывающей не менее 40 цифровых 3D-моделей музейных объектов (время T_k), необходимо в среднем 180 мин.

При формировании цифровой 3D-коллекции антропологических реконструкций М.М. Герасимова было создано и загружено на сайт <http://acadlib.ru/>, интегрированный с ЭБ ННР, 50 работ М.М. Герасимова. Общее время, затраченное на создание этой коллекции, составило:

$$T_{\text{Гер}} = 415 \cdot 50 + 180 = 28\,180 \text{ мин. Это около 470 ч.}$$

ЗАКЛЮЧЕНИЕ

Используя полученные результаты, можно решить задачу оптимизации временных затрат на создание цифровых копий печатных материалов и музейных предметов путем «распараллеливания» технологических процессов, выполняемых библиотечными или музейными специалистами (подготовка метаданных объектов) и техническими специалистами (оцифровка материалов и контроль ее качества).

Полученные оценки в дальнейшем могут быть распространены на создание цифровых копий других видов объектов и использоваться для планирования работ по формированию единого цифрового пространства научных знаний.

Литература

1. Антопольский А.Б., Каленов Н.Е., Серебряков В.А., Сотников А.Н. О едином цифровом пространстве научных знаний // Вестник Российской академии. 2019. Т. 89. № 7. С. 728–735.
2. Антопольский А.Б., Майстрович Т.В. Электронные библиотеки: принципы создания. М.: Либерия-Бибинформ, 2007. 288 с.
3. Богданова И.Ф., Богданова Н.Ф. Электронные библиотеки: история и современность // Информационное общество: образование, наука, культура и технологии будущего. 2017. Вып. 1. С. 133–153.
4. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация (Practical optimization). М.: Мир, 1985. 509 с.

5. *Забровская И.Е., Кириллов С.А., Кондратьева Е.А., Пругло О.А., Сотников А.Н.* Вопросы формирования фондов электронной библиотеки «Научное наследие России» // Информационное обеспечение науки: новые технологии / Сб. науч. тр. М.: БЕН РАН, 2017. С. 184–191.
6. *Каленов Н.Е., Кириллов С.А., Соболевская И.Н., Сотников А.Н.* Визуализация цифровых 3d-объектов при формировании виртуальных выставок // Электронные библиотеки. 2020. Т. 23. № 4. С. 418–432.
7. *Каленов Н.Е., Савин Г.И., Серебряков В.А., Сотников А.Н.* Принципы построения и формирования электронной библиотеки «Научное наследие России» // Программные продукты и системы. 2012. Т. 4. № 100. С. 30–40.
8. *Каленов Н.Е.* Управление технологией наполнения электронной библиотеки «Научное наследие России» // Электронные библиотеки: перспективные методы и технологии, электронные коллекции / Тр. XVI Всерос. науч. конф. RCDL-2014. М., 2014. С. 357–361.
9. *Кириллов С.А., Малинин А.И., Кириллов С.А.* Методика обработки отсканированных изображений в проекте электронной библиотеки «Научное наследие России» // Информационное обеспечение науки: новые технологии / Сб. науч. тр. М.: Научный мир, 2009. С. 99–107.
10. *Лобанов А.Н.* Фотограмметрия. М.: Недра, 1984. 552 с.
11. *Сотников А.Н., Кириллов С.А.* Технология подготовки электронных изданий для электронной библиотеки «Научное наследие России» // Информационное обеспечение науки: новые технологии / Сб. науч. тр. М.: БЕН РАН, 2015. С. 178–190.
12. *Юмашева Ю.Ю.* Методические рекомендации по электронному копированию архивных документов и управлению полученным информационным массивом. М.: ВНИИДАД, 2012. 125 с.
13. *Garstki K.* Virtual representation: the production of 3D digital artifacts. *J. Archaeol. Method Theory*. 2017. Vol. 24. P. 726–750.
14. *Scopigno R.* Digital fabrication techniques for cultural heritage: a survey. *Comput. Graph. Forum*. 2017. Vol. 36. P. 6–21.
15. *Sobolevskaya I.N., Sotnikov A.N.* Principles of 3D Web-collections Visualization // Proceedings of the 3rd International Conference on Computer-Human Interaction Research and Applications. 2019. P. 145–151.
16. *Wróżyński R., Pyszny K., Sojka M., Przybyła C., Murat-Błażejewska S.* Ground volume assessment using ‘Structure from Motion’ photogrammetry with a smartphone and a compact camera // *Open Geosciences*. 2017. Vol. 9. P. 281–294.
17. Государственный рубрикатор научно-технической информации. URL: <http://grnti.ru/> (дата обращения: 04.06.2020).
18. Приказ Министерства культуры РФ от 30 декабря 2014 г. № 2477 «Об утверждении типовых отраслевых норм труда на работы, выполняемые в библиотеках». URL: <https://www.garant.ru/products/ipo/prime/doc/70921222/> (дата обращения 04.06.2020).
19. Сайт академических словарей. URL: https://dic.academic.ru/dic.nsf/fin_enc/31885 (дата обращения: 04.06.2020).
20. Юридическая энциклопедия. URL: https://yuridicheskaya_encyclopediya.academic.ru/12949 (дата обращения: 04.06.2020).
21. Электронная библиотека Института философии РАН. URL: <https://iphlib.ru/library> (дата обращения: 04.06.2020).
22. Электронная библиотека Института этнологии и антропологии РАН им. Н.Н. Миклухо-Маклая. URL: <http://library.iea.ras.ru/index.html> (дата обращения: 04.06.2020).
23. <http://diss.rsl.ru> (дата обращения: 04.06.2020).
24. <http://e-heritage.ru> (дата обращения: 04.06.2020).

25. <http://i.uran.ru/nasledie/> (дата обращения: 04.06.2020).
26. <http://new.jscc.ru/resources/hpc/#item1587> (дата обращения: 04.06.2020).
27. <https://minjust.consultant.ru/documents/14547> (дата обращения: 04.06.2020).
28. https://ru.wikipedia.org/wiki/Электронная_библиотека (дата обращения: 04.06.2020).
29. www.prlib.ru/collections (дата обращения: 04.06.2020).

Bibliography

1. *Antopol'skij A.B., Kalenov N.E., Serebrjakov V.A., Sotnikov A.N.* O edinom cifrovom prostranstve nauchnyh znaniy // Vestnik Rossijskoj akademii. 2019. T. 89. № 7. P. 728–735.
2. *Antopol'skij A.B., Majstrovich T.V.* Jelektronnye biblioteki: principy sozdaniya. M.: Liberija-Bibinform, 2007. 288 p.
3. *Bogdanova I.F., Bogdanova N.F.* Jelektronnye biblioteki: istorija i sovremennost' // Informacionnoe obshhestvo: obrazovanie, nauka, kul'tura i tehnologii budushhego. 2017. Vyp. 1. P. 133–153.
4. *Gill F., Mjurrej U., Rajt M.* Prakticheskaja optimizacija (Practical optimization). M.: Mir, 1985. 509 p.
5. *Zabrovskaja I.E., Kirillov S.A., Kondrat'eva E.A., Pruglo O.A., Sotnikov A.N.* Voprosy formirovanija fondov jelektronnoj biblioteki «Nauchnoe nasledie Rossii» // Informacionnoe obespechenie nauki: novye tehnologii / Sb. nauch. tr. M.: BEN RAN, 2017. P. 184–191.
6. *Kalenov N.E., Kirillov S.A., Sobolevskaja I.N., Sotnikov A.N.* Vizualizacija cifrovych 3d-ob#ektov pri formirovanii virtual'nyh vystavok // Jelektronnye biblioteki. 2020. T. 23. № 4. P. 418–432.
7. *Kalenov N.E., Savin G.I., Serebrjakov V.A., Sotnikov A.N.* Principy postroenija i formirovanija jelektronnoj biblioteki «Nauchnoe nasledie Rossii» // Programmnye produkty i sistemy. 2012. T. 4. № 100. P. 30–40.
8. *Kalenov N.E.* Upravlenie tehnologiej napolnenija jelektronnoj biblioteki «Nauchnoe nasledie Rossii» // Jelektronnye biblioteki: perspektivnye metody i tehnologii, jelektronnye kollekcii / Tr. XVI Vseros. nauch. konf. RCDL-2014. M., 2014. P. 357–361.
9. *Kirillov S.A., Malinin A.I., Kirillov S.A.* Metodika obrabotki otskanirovannyh izobrazhenij v proekte jelektronnoj biblioteki «Nauchnoe nasledie Rossii» // Informacionnoe obespechenie nauki: novye tehnologii / Sb. nauch. tr. M.: Nauchnyj mir, 2009. P. 99–107.
10. *Lobanov A.N.* Fotogrammetrija. M.: Nedra, 1984. 552 p.
11. *Sotnikov A.N., Kirillov S.A.* Tehnologija podgotovki jelektronnyh izdaniy dlja jelektronnoj biblioteki «Nauchnoe nasledie Rossii» // Informacionnoe obespechenie nauki: novye tehnologii / Sb. nauch. tr. M.: BEN RAN, 2015. P. 178–190.
12. *Jumasheva Ju.Ju.* Metodicheskie rekomendacii po jelektronnomu kopirovaniju arhivnyh dokumentov i upravleniju poluchennym informacionnym massivom. M.: VNIIDAD, 2012. 125 p.
13. *Garstki K.* Virtual representation: the production of 3D digital artifacts. J. Archaeol. Method Theory. 2017. Vol. 24. P. 726–750.
14. *Scopigno R.* Digital fabrication techniques for cultural heritage: a survey. Comput. Graph. Forum. 2017. Vol. 36. P. 6–21.
15. *Sobolevskaja I.N., Sotnikov A.N.* Principles of 3D Web-collections Visualization // Proceedings of the 3rd International Conference on Computer-Human Interaction Research and Applications. 2019. P. 145–151.
16. *Wróżyński R., Pyszny K., Sojka M., Przybyła C., Murat-Błażejewska S.* Ground volume assessment using 'Structure from Motion' photogrammetry with a smartphone and a compact camera // Open Geosciences. 2017. Vol. 9. P. 281–294.
17. Gosudarstvennyj rubrikator nauchno-tehnicheskoy informacii. URL: <http://grnti.ru/> (data obrashhenija: 04.06.2020).

18. Prikaz Ministerstva kul'tury RF ot 30 dekabrja 2014 g. № 2477 «Ob utverzhdenii tipovyh otraslevykh norm truda na raboty, vypolnjaemye v bibliotekah». URL: <https://www.garant.ru/products/ipo/prime/doc/70921222/> (data obrashhenija: 04.06.2020).
19. Sajt akademicheskikh slovarej. URL: https://dic.academic.ru/dic.nsf/fin_enc/31885 (data obrashhenija: 04.06.2020).
20. Juridicheskaja jenciklopedija. URL: https://yuridicheskaya_encyclopediya.academic.ru/12949 (data obrashhenija: 04.06.2020).
21. Jelektronnaja biblioteka Instituta filosofii RAN. URL: <https://iphlib.ru/library> (data obrashhenija: 04.06.2020).
22. Jelektronnaja biblioteka Instituta jetnologii i antropologii RAN im. N.N. Mikluho-Maklaja. URL: <http://library.iea.ras.ru/index.html> (data obrashhenija: 04.06.2020).
23. <http://diss.rsl.ru> (data obrashhenija: 04.06.2020).
24. <http://e-heritage.ru> (data obrashhenija: 04.06.2020).
25. <http://i.uran.ru/nasledie/> (data obrashhenija: 04.06.2020).
26. <http://new.jssc.ru/resources/hpc/#item1587> (data obrashhenija: 04.06.2020).
27. <https://minjust.consultant.ru/documents/14547> (data obrashhenija: 04.06.2020).
28. https://ru.wikipedia.org/wiki/Jelektronnaja_biblioteka (data obrashhenija: 04.06.2020).
29. www.prlib.ru/collections (data obrashhenija: 04.06.2020).