ФОРМИРОВАНИЕ ЦИФРОВЫХ КОЛЛЕКЦИЙ: НАУЧНО-ПРАКТИЧЕСКОЕ ПОСОБИЕ ПРЕЗИДЕНТСКОЙ БИБЛИОТЕКИ



doi 10.47267/2181-8207/2020/2-002



Елена Жабко, Директор по информационным ресурсам Президентской библиотеки им. Б.Н.Ельцина, доктор пед. наук, Председатель секции РБА «Электронные ресурсы и информационно-библиотечное обслуживание», (Россия)

Юлия Селиванова, Начальник отдела лингвистического и программно-технологического обеспечения Президентской библиотеки им. Б.Н.Ельцина, канд пед. наук, (Россия)

В настоящее время формирование цифровых коллекций становится актуальной научно-практической задачей для российских библиотек. Процесс создания цифровой коллекции – многоэтапный процесс, начинающийся с разработки концепции коллекции и заканчивающийся представлением коллекции на портале библиотеки. Детальному рассмотрению всех этапов формирования цифровой коллекции посвящено научно-практическое пособие Президентской библиотеки «Формирование цифровых коллекций». Пособие подготовлено на основе результатов теоретических исследований и практической деятельности Президентской библиотеки в области формирования цифровых коллекций.

Ключевые слова: цифровых коллекций, электронных документов, информационных объектов, метаданными, международных стандартов, фондодержателей, цифровой формат, редких книг, архивных материалов

В первой главе пособия «Цифровая коллекция: определения, функции, основные виды коллекций» рассматриваются различные трактовки понятия «коллекция». Отмечается, что в электронной среде понятие «коллекция» приобретет особый смысл, связанный с как с определенной формой представления и организации информации для пользователей, так и с особой технологией формирования цифровой коллекции. В профессиональной литературе можно встретить различные варианты термина и его определения – «коллекция электронных документов», «коллекция информационных ресурсов», «коллекция информационных объектов», «электронная или цифровая коллекция» и тд.

Например, В ГОСТ Р 7.0.96–2016 «Электронные библиотеки. Основные виды. Структура. Технология формирования» содержится определение *коллекции электронных документов* – «часть фонда электронной библиотеки, представляющая логическое объединение электронных документов, отобранных по какому-либо критерию принадлежности и предназначенная для реализации задач, которые определены политикой электронной библиотеки». В этом же ГОСТе указывается, что в состав фонда электронной библиотеки, наряду с электронными документами, входят описывающие их метаданные, коллекции с метаданными, гиперссылки на внешние ресурсы.

В «Глоссарии по информационному обществу» (М., 2009)[1] приводится определение *коллекции информационных объектов*, в соответствии с которым в состав коллекции включаются метаданные, описывающие информационные объекты: «Коллекция информационных объектов – организованная совокупность информационных объектов, обладающих общностью каких-либо их характеристик (тематика, авторство, хронология, источники и т. д.), вместе с описывающими их метаданными».

Авторы пособия, основываясь на своей практики формирования коллекций в электронной среде, предлагают использовать термин *«цифровая коллекция»* под которой понимается систематизированная совокупность электронных документов и других цифровых объектов, объединенных общим признаком (тематическим или формальным), отобранных и организованных в целях содействия

[1] Глоссарий по инфрмационному обществу / под общ. ред. Ю.Е. Хохлова. – М. : Институт развития информационного общества, 2009.

раскрытию их содержания, доступу к ним и их использованию, вместе с описывающими их метаданными. Цифровые объекты могут объединяться в коллекцию по общности содержания, по типам документов (текстовые, картографические, нотные и т. д.), по хронологическому признаку, по назначению, видам источников, кругу пользователей и т. д.).

Цифровые коллекции ПБ создаются с учетом международных принципов, разработанных NISO (National Information Standards Organization, Национальная организация международных стандартов (США)) и опубликованных в Framework of Guidance for Building Good Digital Collections») («Руководство по созданию цифровых коллекций»). На основе этих принципов формируют свои коллекции многие библиотеки, архивы и музеи мира.

В издании Президентской библиотеки приводится описание принципа NISO и его интерпретация с учетом практики библиотеки.

Например, принцип NISO «Формирование и развитие коллекции должно осуществляться в соответствии с четкой стратегией. По данным исследователей, еще в 2008 г., по меньшей мере одна треть от всех запускаемых цифровых проектов терпели неудачу по причине отсутствия проектного управления и последующего контроля, еще одна треть - потому не было точного определения цели и недостаточной информированности участвующих специалистов об объеме и сложности работ. Именно поэтому стратегическое планирование стало обязательным условием создания электронных библиотек, цифровых коллекций и любых массивов цифрового контента. В стратегии определяется цель создания ресурса, ставятся задачи которые предстоит решить, определяется потенциальная аудитория пользователей, включается конкретный план мероприятий. В стратегии могут закрепляться самые разнообразные задачи: создание ресурсов национального уровня, интеграция коллекции в программы дистанционного обучения, подготовка специализированных пользовательско-ориентированных коллекций, привлечение пользователей к формированию метаданных и др. Важным аспектом при разработке стратегии является прогнозирование дальнейшего использования ресурса не только в привычных областях образования и науки, но в области маркетинга, рекламы, проведения выставок, привлечения общественности и сбора финансовых средств в рамках фандрайзинга. Поскольку основным источником цифрового контента при создании коллекций являются результаты оцифровки документов различных фондодержателей, то в стратегии в обязательном порядке определяются критерии отбора ресурсов для перевода в цифровой формат. Данное направление работы сродни традиционному комплектованию библиотек, но вместо оценки изданной литературы для покупки, библиотекари дают оценку тем ресурсам, которые должны стать основной или частью цифровой коллекции. При это важным является осведомленность об аналогичных коллекциях, создаваемых в других организациях во избежание дублирования при сканировании. Поскольку большая часть ресурсов будет состоять не только из редких книг и архивных материалов, но из карт, фотографий, произведений искусства, аудио и видеозаписей, то при отборе данных документов в стратегии фиксируются технические требования к цифровым копиям каждого вида документов. Также в общие стратегии могут включаться положения относительно условий доступа и предоставления контента пользователям, длительности хранения ресурсов и курирования отдельных разделов.

Важным аспектом при разработке стратегии является прогнозирование дальнейшего использования ресурса не только в привычных областях образования и науки, но в области маркетинга, рекламы, проведения выставок, привлечения общественности и сбора финансовых средств в рамках фандрайзинга

Применительно к деятельности Президентской библиотеки данный принцип интерпретирован следующим образом: «Уставной целью Президентской библиотеки является формирование общегосударственного хранилища информационных ресурсов в виде представленных в электронно-цифровой форме произведений и документов по истории, теории и практике российской государственности, по вопросам русского языка как государственного языка Российской Федерации. В соответствии с уставной целью Президентская библиотека формирует базовые коллекции «Государственная власть», «Территория России», «Российский народ», «Русский язык», раскрывающие различные аспекты понятия российской государственности. Кроме того, в Президентской библиотеке формируются тематические коллекции, посвященные отдельным периодам и событиям российской истории, деятельности известных общественных деятелей и политиков.

«Уставной целью Президентской библиотеки является формирование общегосударственного хранилища информационных ресурсов в виде представленных в электронно-цифровой форме произведений и документов по истории, теории и практике российской государственности

При формировании коллекции важно определить круг потенциальных пользователей. В «Руководстве» NISO приводится следующая рекомендация: «*организация должна определить основную целевую аудиторию коллекции, но при этом необходимо учитывать и других потенциальных пользователей». Ос-*

новная пользовательская группа Президентской библиотеки включает исследователей и специалистов в области истории, так как большую часть фонда библиотеки составляют архивные материалы, опубликованные исторические источники, научные труды. Однако, следуя одной из задач библиотеки по привлечению студентов и школьников к работе с документами и историческими источниками, реализуется возможность использования коллекций Президентской библиотеки в просветительской и образовательной деятельности. Поэтому в коллекции включаются записи лекций известных ученых, ведущих преподавателей высших учебных заведений, политиков по отдельным проблемам и вопросам.

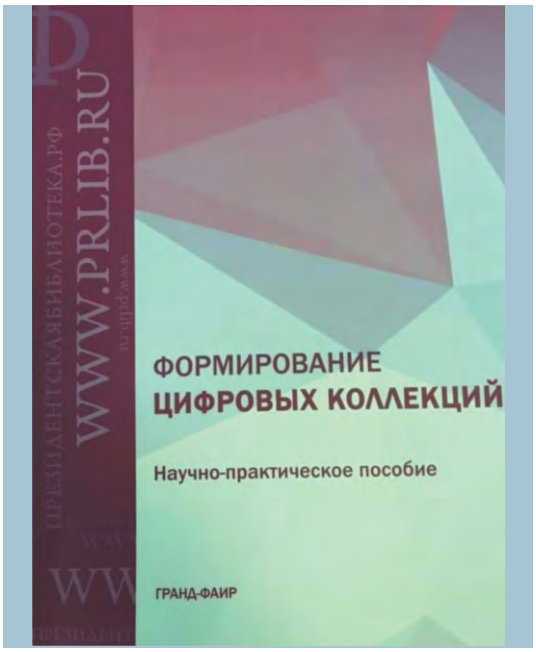
В Президентской библиотеке все производственные этапы формирования коллекций от планирования до размещения на портале, а также последующее пополнение коллекций интегрированы в существующие технологические бизнес-процессы. Над созданием и дальнейшим развитием коллекции работают различные специалисты библиотеки: комплектаторы проводят отбор документов, технические специалисты осуществляют сканирование, каталогизаторы формируют метаданные, куратор коллекции разрабатывает структуру коллекции и распределяет документы по разделам, веб-мастер обеспечивает дизайнерское оформление коллекции, технические специалисты отвечают за публикацию коллекции на портале и обеспечение сохранности.

В «Руководстве» NISO регламентируются принципы, обеспечивающие качественную подготовку цифровых коллекций: создание описания коллекции, организация доступа пользователей, обеспечение устойчивого жизненного цикла функционирования коллекции во времени и пространстве.

Коллекции должны быть описаны таким образом, чтобы пользователь мог получить информацию о характеристиках коллекции – содержании, формате объектов, ограничениях доступа, владельце коллекции, а также другую информацию, существенную для определения целостности коллекции.

Необходимо обеспечить широкий доступ пользователей к коллекции и исключить необоснованные ограничения на использование материалов коллекции.

На интернет-портале Президентской библиотеки (www.prlib.ru) для поиска и просмотра доступны более 300 коллекций, которые содержат аннотацию в целом к коллекции, аннотации к разделам. В случае, если материалы коллекции подпадают под действие авторского права или имеют иные ограничения для открытого просмотра, предусмотрена возможность работы с цифровыми коллекциями в центрах удаленного доступа Президентской библиотеки. В настоящее время открыто свыше 600 центров удаленного доступа в различных регионах Российской Федерации и странах Европы, Азии, Америки.



Кроме принципов, зафиксированных в Руководстве авторами пособия предложены два принципа, которые также могут использоваться при формировании коллекций:

Принцип планового развития предусматривает перспективное планирование создания новых коллекций и пополнение уже существующих. В Президентской библиотеке перспективный план развития составляется для базовых коллекций с учетом значимых дат и событий в экономическом, социальном, политическом, культурном и духовном развитии страны. В планах отражаются юбилейные даты жизни и деятельности представителей науки, культуры, государственных деятелей. Кроме того, составляется план-календарь событий и дат, а также план мероприятий, проводимых в библиотеке.

Принцип интеграции заключается в том, что для поддержания интереса различных групп пользователей к коллекциям необходимо обеспечить включение в коллекцию документов различного типа и вида.

Цифровая коллекция может включать электронные документы различного вида: текстовые, изобразительные и аудиовизу-

В «Руководстве» NISO регламентируются принципы, обеспечивающие качественную подготовку цифровых коллекций: создание описания коллекции, организация доступа пользователей, обеспечение устойчивого жизненного цикла функционирования коллекции во времени и пространстве

В Президентской библиотеке перспективный план развития составляется для базовых коллекций с учетом значимых дат и событий в экономическом, социальном, политическом, культурном и духовном развитии страны

альные документы (включая кино-, видео-, фоно-, фотопродукцию и ее комбинации), мультимедийные документы. Наиболее полное, глубокое раскрытие темы коллекции предполагает включение в цифровую коллекцию, помимо библиотечных материалов, элек-тронных копий документов и образов пред-метов, хранящихся в архивах и музеях. Вклю-чение в одну коллекцию материалов трех ви-дов хранения позволяет раскрыть наиболее полно документальную историю общества, науки, истории государства, жизнедеятель-ности персоны, создать более широкий кон-текст восприятия темы.

Первая глава пособия завершается характеристикой различных видов цифровых коллекций персоноведческих, регионоведческих, страноведческих, тематических и тд.

Вторая глава «Метаданные управлении цифровыми объектами и коллекциями» пособия Президентской библиотеки посвящена характеристике метаданных, используемых при описании цифровых объектов и коллекций. Термин «метаданные» широко используется специалистами в области создания и эксплуатации информационных систем для обозначения структурированной информации о характеристиках цифровых объектов с целью их идентификации, поиска, обработки и хранения.

В ГОСТ Р 7.0.96-2016 «Электронные библиотеки. Основные виды. Структура. Технология формирования» метаданные определяются как «структурированные данные, описывающие контекст, содержание и структуру электронного документа на протяжении его жизненного цикла». Пособие включает характеристику метаданных, используемых для цифровых объектов и коллекций: описательных, административных и структурных метаданных.

Описательные метаданные

представляют информацию, необходимую для идентификации, поиска и обнаружения информационного объекта и связанных с ним объектов. Элементы описательных метаданных, как правило, содержат следующую информацию об объекте: заглавие, автор, предмет, примечания, аннотацию, точки доступа.

Административные метаданные создаются преимущественно для управления цифровыми коллекциями и отдельными цифровыми объектами и включают технические метаданные (тип формата данных, объем в байтах, сжатие, разрешение и пиксельный размер изображения, цветность, глубина цвета, дата и время создания объекта и тд.), метаданные сохранности (информация о физическом состоянии цифрового объекта, об изменениях объекта на протяжении времени и истории его использования), метаданные прав (сведения о правообладателе, о лицензии, об условиях доступа, правах и способах репродуцирования и т. д.).

Структурные метаданные характеризуют внутреннюю структуру объекта и ее компоненты – связи между частями цифрового файла, порядок следования страниц, глав и т. д.. Основное назначение структурных метаданных – обеспечение навигации и представление корректной последовательности отображения цифровых объектов на экране для пользователя. Цифровой объект нуждается в более развернутом наборе метаданных, чем документы в традиционной форме. Для традиционных изданий достаточно библиографической записи, в которой представлена содержательная и физическая характеристика издания (количество страниц, размер и тд). Для цифрового объекта таких данных недостаточно, требуется детальное описание структуры цифрового издания, без подобной структуры, цифровой объект будет восприниматься как набор разрозненных изображений.

В ходе исследования LIFE, проведенного Британской библиотекой и UCL (University College London, Университетский колледж в Лондоне) под эгидой JISC (Joint Information Systems Committee, Объединенный *комитета* информационных систем) было установлено, что до 50% всех затрат, связанных с депонированием и поддержанием жизненного цикла электронного документа, связано с созданием метаданных (включая технические, описательные, административные метаданные и метаданные сохранности).[2]

В международном культурном сообществе для записи метаданных цифровых объектов

Термин «метаданные» широко используется специалистами в области создания и эксплуатации информационных систем для обозначения структурированной информации о характеристиках цифровых объектов с целью их идентификации, поиска, обработки и хранения

Процесс формирования метаданных цифровых объектов начинается с момента создания цифровой копии. На стадии создания цифрового объекта автоматически генерируются данные о типе файла, разрешении, дате создания и тд

[1] *LIFE (Life Cycle Information for E-Literature)* – проект, целью которого было изучение жизненного цикла коллекции и проблем сохранения цифровых материалов. В рамках проекта был исследован жизненный цикл цифровых коллекций UCL и Британской библиотеки. По результатам исследования были выработаны рекомендации по расчёту стоимости каждого этапа жизненного цикла с учетом долгосрочного хранения цифрового контента. Информацию о проекте LIFE см. http://www.life.ac.uk/

широко используется стандарт METS (Metadata Encoding and Transmission Standard, Стандарт кодирования и передачи метаданных). METS позволяет объединить в одном XML-документе всю информацию об объекте – из каких файлов он состоит объект, где они размещаются, как связаны между собой, какие метаданные относятся к объекту и его компонентам, каковы условия использования объекта, какие действия должна выполнять система при открытии объекта.

В пособии подробно описана структура стандарта. METS.

Третья глава «Технология формирования и ведения коллекций» посвящена описанию жизненного цикла цифровой коллекции. Жизненный цикл коллекции включает следующие этапы: планирование, отбор ресурсов, создание (приобретение) цифрового контента, загрузка в информационную систему, описание коллекций и ресурсов, входящих в них, публикации для общественного доступа.

Создание концепции коллекции предполагает определение предмета/темы коллекции, ее цели и задач, содержательного замысла и полноты раскрытия предмета, принципов и критериев отбора документов, предварительной структуры коллекции

На этапе планирования происходит разработка концепции коллекции. Создание концепции коллекции предполагает определение предмета/темы коллекции, ее цели и задач, содержательного замысла и полноты раскрытия предмета, принципов и критериев отбора документов, предварительной структуры коллекции, стратегии формирования метаданных, перспектив дальнейшего развития коллекции.

В качестве предмета коллекции может выступатьперсона, организация/учреждение, тема, событие, регион, страна, вид / тип документов и др. Содержательный замысел коллекции, т.е. формирование основной идеи представления предмета коллекции и степени его содержательного раскрытия, определяет тематические границы создаваемой коллекции – круг тем (проблем, вопросов), которые предполагается включить в коллекцию, а также необходимую полноту их отражения в коллекции.

Общей целью создания любой коллекции является предоставление пользователю отобранных и систематизированных информационных ресурсов по определенному предмету, теме или проблеме

Общей целью создания любой коллекции является предоставление пользователю отобранных и систематизированных информационных ресурсов по определенному предмету, теме или проблеме. Наиболее общими задачами цифровых коллекций являются: многоаспектное раскрытие предмета коллекции; аккумуляция разнородных информационных ресурсов по теме коллекции; предоставление пользователю наиболее значимых и систематизированных информационных ресурсов по теме коллекции, обеспечение эффективного информационного поиска и др. Конкретные цели и задачи каждой коллекции определяются в зависимости от предмета коллекции, целевого назначения, характера собираемого материала. Для большинства цифровых коллекций актуальной также является такая цель, как сохранение оригинальных документов и других объектов культурного наследия.

В качестве основных критериев отбора документов используются: содержание документов (соответствие содержания документов тематике, закрепленной концепцией формируемой коллекции); целевое назначение документов (категория пользователей, для которой предназначен документ); способ хранения (документы / предметы библиотечного, архивного и музейного хранения); природа основной информации (текстовые, изобразительные, аудиовизуальные, картографические, мультимедийные документы и т. д.); время создания (учитывается с точки зрения необходимости соблюдения действующего законодательства в области авторского и смежных прав); сохранность и комплектность документов (обеспечение целостности данных) и тд.

Структура коллекции определяется, прежде всего, особенностями ее предмета/темы и состава включенных в нее документов.

В рамках концепции также разрабатывается модель ее визуализации, т. е. представления объектов коллекции для пользователя; определяются пути рекламы и продвижения коллекции.

В пособии рассматриваются особенности формирования концепции различных коллекций: персоноведческих, тематических, типо-видовых, регионоведческих.. В приложении 2 к пособию приведена Модель концепции регионоведческой коллекции Президентской библиотеки.

Большое внимание уделено вопросам описания, индексирования, дисплейному представлению цифровых коллекций, организации поиска в электронном каталоге.

В российском библиотечном сообществе необходимость формирования библиографической записи на коллекцию становится актуальной в связи с реализацией ряда проектов, связанных с обменом и передачей коллекций цифровых ресурсов. Президентская библиотека совместно с Центром ЛИБНЕТ в 2016–2017 гг. проводили работы по разработке методики описания цифровых коллекций, организации поиска и дисплейному

представлению коллекций в Сводном каталоге электронных ресурсов (СКЭР

В настоящее время стандартов, четко регламентирующих описание коллекций, нет – ни на национальном, ни на международном уровне.

В Президентской библиотеке описание как на уровне отдельных документов, так и на уровне цифровой коллекции осуществляется в Российском коммуникативном формате представления библиографических записей в машиночитаемой форме (RUSMARC). Запись RUSMARC на коллекцию создается в соответствии с общими правилами формата. Информация о принадлежности документа к той или иной коллекции вносится в библиографическую запись каждого документа, отобранного для коллекции. Наличие библиографической записи на коллекцию и связи, установленные между описаниями документов и коллекций, обеспечивают возможность организации поиска как на верхнем уровне, по метаданным коллекции в целом, так и поиска документов в пределах отдельной коллекции. Пособие включает подробные рекомендации по формированию библиографических записей на коллекцию, в приложении 8 содержится пример описания на уровне коллекции.

Цифровые коллекции могут рассматриваться как самостоятельный информационный ресурс, включающий отобранную, систематизированную информацию по определенной теме, персоне, географическому объекту, по жанровой или видовой принадлежности и т. д. Это дает основания говорить о коллекции как о новом объекте информационного поиска, появление которого требует разработки поискового интерфейса и это становится новой задачей как для библиотечных специалистов, так и для разработчиков автоматизированных библиотечных систем.

В главе 4 представлены обзоры базовых цифровых коллекций Президентской библиотеки: «Территория России», «Государственная власть», «Российский народ», «Русский язык», а также персоноведческих, регионоведческих, страноведческих, событийных, типовидовых и посвященных различного рода организациям.

Завершает пособие раздел «Приложения», в который включены примеры различной документации, связанной с техническими требованиями на сканирование документов, шаблоны списков на оцифровку, акты-приема и передачи итд.

В ходе работы над пособием авторы использовали большое количество публикаций отечественных и зарубежных авторов, опирались на стандарты и образцы из мировой практики, а также свой многолетний опыт работы с цифровыми коллекциями, что проиллюстрировано значительным количеством примеров и дополнительных материалов в приложениях. Пособие, подготовленное специалистами Президентской библиотеки можно рассматривать как практическое руководство по созданию цифровых коллекций. Издание Президентской библиотеки может быть интересно специалистам различных типов библиотек занимающихся вопросами оцифровки, а также студентам, аспирантам библиотечных факультетов.

Информация о принадлежности документа к той или иной коллекции вносится в библиографическую запись каждого документа, отобранного для коллекции

Цифровые коллекции могут рассматриваться как самостоятельный информационный ресурс, включающий отобранную, систематизированную информацию по определенной теме, персоне, географическому объекту, по жанровой или видовой принадлежности и т. д

Елена Жабко – Б.Ельцин номидаги Россия Президент кутубхонасининг Ахборот ресурслари бўйича директори, педагогика фанлари доктори.

Мақолада муаллиф АҚШда фаолият юритувчи (НИСО) Халқаро стандартлар миллий ташкилоти томонидан ишлаб чиқилган халқаро тамойиллар ёрдамида яратилган ва яхши рақамли тўпламларни яратиш учун ниҳоятда керакли манба сифатида чоп этилган «Рақамли тўплам яратиш учун қўлланма»нинг афзалликлари ҳақида маълумотлар беради. Бу тамойиллар асосида дунёнинг кўплаб кутубхоналари, архив ва музейлари ўз рақамли тўпламларини ташкил этадилар.

Шунингдек, мақолада Россия Федерациясининг тарихи, назарияси ва амалиёти бўйича рақамли шаклда тақдим этилган ишлар ва ҳужжатлар шаклида ахборот ресурсларининг миллий омборини яратиш, Низом мақсадига мувофиқ Президент кутубхонаси «Давлат ҳокимияти», «Россия ҳудуди», «Рус халқи», «Рус тили» каби асосий тўпламларни ташкил этиб, рус давлатчилиги тушунчасининг турли қирраларини очиб берадиган Президент кутубхонасининг фаолияти ҳақида ҳам сўз боради. Бундан ташқари, муаллиф, Президент кутубхонасидаги Россия тарихининг муайян даврлари ва воқеаларига, машҳур жамоат арбоблари ва сиёсатчиларнинг фаолиятига бағишланган тематик тўпламлар ҳақида маълумот беради.