

ДОСТУП К НАУЧНЫМ ПУБЛИКАЦИЯМ: ВЫБОР ПЛАТФОРМЫ

ACCESS TO SCIENTIFIC PUBLICATIONS: SELECTING THE SOFTWARE

Ковязина Елена Васильевна, научный сотрудник, кандидат технических наук, Институт вычислительного моделирования Сибирского отделения Российской академии наук – обособленное подразделение Федерального исследовательского центра «Красноярский научный центр Сибирского отделения Российской академии наук», e-mail: elena@icm.krasn.ru

Kovyazina Elena, researcher, PhD, Institute of Computational Modelling of the Siberian Branch of the Russian Academy of Sciences, e-mail: elena@icm.krasn.ru

Аннотация. В последнее десятилетие в библиотеках научных и образовательных организаций стало нормой формирование собственных полнотекстовых информационных ресурсов, многообразных по назначению, содержанию и параметрам эксплуатации. В докладе представлен сравнительный анализ двух таких ресурсов – электронно-библиотечной системы (ЭБС) и институционального репозитория научных публикаций. Результаты анализа наглядно иллюстрируют принципиальные различия между такими ресурсами, определяющие особенности их формирования и выбор программной платформы, максимально соответствующей предъявляемым к таким ресурсам требованиям.

Abstract. Over the past decade, the normal thing for libraries of scientific and educational organizations have become to create their own full-text information resources, diverse in purpose, content and operation parameters. A comparative analysis of two such resources – the Electronic Library System (EBS) and Institutional Repository are presented in this report. The results of the analysis clearly illustrate the fundamental differences between such resources, determining the specifics of their creation and the choice of a software platform.

Ключевые слова. Электронно-библиотечная система (ЭБС), институциональный репозиторий, архив открытого доступа.

Keywords. Electronic Library System (EBS), institutional repository, open access archive.

Введение. Последнее десятилетие в библиотеках научных и научно-образовательных организациях характеризуется постоянным ростом количества электронных документов, компонуемых в информационные ресурсы собственной генерации. По форме представления таких документов в информационном ресурсе их можно условно разбить на два типа:

- документы, представленные в информационном ресурсе в виде ссылки на файл в локальном хранилище или Интернет-ссылки;
- документы, интегрированные в информационный ресурс, представляющий в данном случае полнотекстовую базу данных.

Исторически в библиотеке преобладали документы первого типа, точно представленные в преимущественно библиографических или реферативных базах данных, часто копии печатных изданий. Однако рост числа таких документов, источники их поступления, специальное назначение и т.п. привели к необходимости осмысления их уникального места в библиотеке, политики комплектования такими документами, читательских сервисов, неразрывно связанных с их использованием, потребности в технических и программных средствах для их обслуживания. Все электронные документы в совокупности составляют **электронную библиотеку** – информационную систему, предназначенную для организации и хранения упорядоченного фонда электронных объектов, и обеспечения доступа к ним с помощью единых средств навигации и поиска [1], документально оформленную и развивающуюся в соответствии с политикой формирования электронной библиотеки внутри соответствующей организации. Однако аккумулирует ли единая электронная библиотека весь набор электронных объектов, учитываемый и обслуживаемый библиотекой научной организации? Рассмотрим на примере информационных ресурсов собственной генерации существует ли возможность обеспечить доступ к ним «с помощью единых средств навигации и поиска», учитывая при этом налагаемые на эти ресурсы корпоративные требования и пользовательские пожелания.

Информационные ресурсы, генерируемые библиотекой научной или научно-образовательной организации, представляют собой совокупность **электронных коллекций** – «части фонда электронной библиотеки, представляющей логическое объединение электронных документов, отобранных по какому-то критерию принадлежности, и предназначенной для реализации задач, которые определены политикой электронной библиотеки» [1]. По возложенным на них задачам такие электронные коллекции можно разделить на три группы:

- научные труды сотрудников – собрание документов, отражающих результаты научных исследований, проводимых организацией;
- электронно-библиотечные системы (ЭБС) [2] – собрание документов, специально предназначенных для ведения образовательной деятельности в соответствии с требованиями Рособнадзора;
- тематические коллекции – узкоспециальные собрания документов. К этой группе можно отнести мемориальные, относящиеся к научной и общественной деятельности известных

ученых и научных школ, а также специализированные профильные коллекции, созданные в процессе выполнения научно-исследовательских работ.

Научные труды сотрудников. Директивной целью формирования собрания научных трудов сотрудников является сохранение научного наследия, связанного с исследовательской деятельностью организации. Целью определяются и основные требования к этому собранию: полнота *учета* документов, точное их описание, дополнительная оцифровка печатных работ, обеспечение длительной сохранности (страховые копии). По сути это «электронный архив-система структурированного хранения электронных документов, обеспечивающая надежность хранения, конфиденциальность и разграничение прав доступа, отслеживание истории использования документа, быстрый и удобный поиск. Обычно электронный архив работает с неструктурированными документами, которые снабжаются некоторым структурированным описанием (карточка документа)» (Википедия). Перечисленные требования, за исключением оцифровки, унаследованы от фонда соответствующих печатных документов. Цифровая форма документов позволяет сформулировать дополнительные цели информационного ресурса. Наука не может существовать в отрыве от международных научных коммуникаций, и большинство ученых заинтересованы в возможно более широком распространении и всестороннем обсуждении результатов их исследований. Такое распространение носит характер *продвижения* научного исследования, служит персональной известности ученого, улучшает его показатели цитирования [3]. Возможности, предоставляемые электронной средой и глобальными сетями связи, дополнили и модифицировали требования к собранию научных трудов. Электронный архив трансформировался в «институциональный репозиторий – электронный архив для длительного хранения, накопления и обеспечения долговременного и надежного открытого доступа к результатам научных исследований, проводимых в учреждении» (Википедия). Ключевым в этом определении является *открытый доступ* к публикациям. Составляющие открытого доступа перечислены в соответствующих инициативных документах [4-5]:

«Под "открытым доступом" [к рецензируемой научной литературе] ... подразумеваем свободный доступ к ней через публичный Интернет, и право каждого пользователя читать, загружать, копировать, распространять, распечатывать, искать или делать ссылки на полнотекстовые статьи, проводить поиск роботами-индексаторами, вводить их как данные в программное обеспечение или использовать для других законных целей при отсутствии финансовых, правовых и технических преград, за исключением тех, которые регулируют доступ к собственно Интернету. Единственным ограничением на воспроизводство и

распространение и единственным условием копирайта в этой области должно быть право автора контролировать целостность своей работы и обязательные ссылки на его имя при использовании работы и ее цитировании» [4].

Такое определение, а также временные рамки в регламентах открытого доступа, как правило, не позволяет администрации научной организации в полной мере присоединиться к данной инициативе, а институциональный репозиторий может считаться репозиторием открытого доступа лишь условно, что соответствует общемировым тенденциям в этой области [6]. Содержимое институционального репозитория, предполагаемое к открытому доступу, должно пройти определенную последовательность отсева документов, полноценный открытый доступ к которым в настоящий момент времени невозможен:

- экспертное заключение о возможности опубликования, т.е. фиксация отсутствия в документе, например, сведений, содержащих государственную тайну;
- признаки, фиксирующие то обстоятельство, что документ является служебным произведением, т.е. отраженные в публикации исследования выполнялись всеми авторами в рамках их служебных обязанностей в данной исследовательской организации;
- имеются договора или свободные лицензии со стороны сотрудников сторонних организаций, являющихся соавторами произведения.

Отсеянные в результате таких шагов документы предполагают «темное» [4] размещение в репозитории, пока разрешение на открытый доступ не может быть получено, то есть в архив, сохраняющий научное наследие, такие документы занесены, но «не видны» никому, кроме администратора архива. При определении требований к доступу также должно учитываться издательское эмбарго – временные ограничения, налагаемые издателями на публикацию произведения на сайте автора или его места работы. Наиболее полный перечень продолжительности таких ограничений представлен на сайте проекта SHERPA/ROMEO, посвященного открытому доступу (<http://www.sherpa.ac.uk/romeo/PDFandIR.php?la=en>). Таким образом, требования к доступу в институциональном репозитории зависят от правового статуса документа и могут быть следующими:

- полностью открытый доступ для служебных произведений с истекшим периодом эмбарго. В случае наличия соавторов-сотрудников сторонних организаций предполагается наличие авторского договора с ними;

- отсроченный открытый доступ для документов, находящихся под действием эмбарго, либо ожидающих авторского договора с соавторами. Описание таких документов должно содержать информацию об эмбарго и сроках его действия, либо формулировку иных причин временного ограничения доступа;
- закрытый доступ для произведений вышедших в издательствах, условия публикации в которых вообще не предполагают открытого доступа. Если институциональный репозиторий будет зарегистрирован в соответствующих службах учета репозиториях открытого доступа, то, по-видимому, даже описательные метаданные таких документов должны быть недоступны анонимным пользователям;
- электронные документы, содержащие государственную тайну, не подлежат учету библиотекой, хранятся и содержатся в соответствующих подразделениях и условиях.

Имеются также некоторые особенности, связанные со спецификой институционального репозитория. Требования учета (надежность хранения) и одновременно продвижения (открытость доступа) создают некоторые технические противоречия, из-за которых программное обеспечение («коробочный» вариант) части крупных IT-компаний не пригодно к использованию для организации доступа к институциональным репозиториям. Также определенные противоречия существуют между стремлением научной организации, с одной стороны, иметь как можно больше научных публикаций, а с другой – зафиксировать свои имущественные права на публикацию, как на служебное произведение. Стремление к их разрешению побуждает идти на компромисс с издателями, выдерживая налагаемый на произведение период эмбарго, что также не способствует открытости репозитория.

Электронно-библиотечная система (ЭБС) – «автоматизированная информационная система, базы данных которой содержат организованную коллекцию электронных документов, включающую электронные издания, используемые для информационного обеспечения образовательного и научно-исследовательского процесса в образовательных организациях, обеспечивающая возможность доступа к электронным документам через сеть Интернет» [2]. В применении к библиотеке научной или научно-образовательной организации будем рассматривать только ту ЭБС, оператором которой выступает сама библиотека, т.е. генерируемую библиотекой. На первый взгляд, институциональный репозиторий и ЭБС обладают определенным сходством. Однако это мнимое сходство. Целью ЭБС является информационное обеспечение прежде всего образовательного процесса, научно-исследовательский выступает только как неразрывно связанный с образовательным (см. «... в образовательных организациях...»). Если контент институционального репозитория

преимущественно самостоятельно генерируемый, то ЭБС состоит, главным образом из агрегируемого привлеченного контента. В данном контексте научная и научно-образовательная организации значительно разнятся. Несмотря на то, что в редком научном институте нет аспирантуры, ординатуры или структур дополнительного образования, эта деятельность не является основной деятельностью по государственному заданию, а лишь вспомогательной. В научной организации не существует штатных преподавателей. Как следствие, создание учебников, учебных и методических пособий и иных, связанных с образовательной деятельностью документов, не является частью должностных обязанностей сотрудников, а электронные документы не являются служебными произведениями. Таким образом, электронные документы, входящие в состав ЭБС, являются объектами авторского или смежного права или предоставляются издателями на лицензионных условиях. Права доступа к каждому конкретному электронному документу определяются условиями авторского либо лицензионного договора и могут быть весьма разнообразны. Доступ к ресурсу является ограниченным – требующим авторизации, регистрации или предоставляется на иных условиях информационного посредника. А ЭБС, в отличие от институционального репозитория, является локальным информационным ресурсом, с ограниченным контингентом пользователей.

Выбор программной платформы. Выше были определены основные различия между двумя наиболее распространенными информационными ресурсами, самостоятельно генерируемыми библиотеками научных и научно-образовательных организаций. Тематические коллекции, в зависимости от входящих в них документов, могут тяготеть по условиям доступа к одному из двух рассмотренных ранее ресурсов. Все они, по сути, являются полнотекстовыми базами данных, предполагающими поиск не только по описательным метаданным документа, но и по его полному тексту. Полнотекстовый поиск, характер содержимого информационных ресурсов и требования к условиям доступа определяют выбор программной платформы, на которой этот ресурс будет реализован. Традиционно в библиотеках описательные метаданные документов формировались с помощью системы автоматизации библиотеки (САБ), а полнотекстовый поиск обеспечивался ее специализированными модулями. Однако ни одна из отечественных САБ не обеспечивает интеграцию электронных документов в среду Интернет, с помощью которой может быть обеспечено эффективное продвижение публикаций в мировое информационное пространство. Указанный факт побуждает научные организации формировать институциональные репозитории в технологиях и на открытом программном обеспечении Интернет-сообщества.

В таблице 3 приведен сравнительный анализ требуемой для полнотекстовых ресурсов функциональности наиболее распространенного в России открытого программного обеспечения для институционального репозитория DSpace и недавно разработанного модуля САБ ИРБИС, предназначенного для формирования ЭБС, – ИРБИС64+. Функциональные характеристики, присутствующие и недостающие в каждой из сравниваемых платформ, выделены цветом.

Обе программных платформы имеют свою область уникальных функций. У ИРБИС64+ – это набор функций, связанных с российской корпоративной каталогизацией, осуществляемой как через ИРБИС-корпорацию, так и через иные центры, такие как ЛИБНЕТ или региональные корпорации, связанные по протоколам Z39.50. Стоит также отметить, что в САБ ИРБИС довольно давно интегрированы средства заимствования данных о цитированиях публикации из индексов научного цитирования, что является полезным дополнением для такого информационного ресурса, как институциональный репозиторий. DSpace без дополнительного программирования способен обеспечить заимствование только по протоколам Z39.50 в формате Дублинского ядра. САБ также обеспечивает комфортную для сотрудников библиотеки среду формирования библиографического описания документа в привычных библиографам MARC-форматах, обеспеченную лингвистическими средствами и справочниками. Немаловажным для формирования ЭБС является также интеграция ИРБИС64+ с модулем книговыдачи, позволяющим вести индивидуализированный учет статистики использования электронных документов в ней.

Таблица 3

Сравнительные характеристики программного обеспечения

Требования	ИРБИС64+	DSpace
Интеграция с системой книговыдачи	+	-
Комфортный и привычный интерфейс библиографа	+	-
Поддержка российской корпоративной каталогизации	+	-
Полнотекстовый поиск	+	+
Представление результатов поиска в порядке убывания релевантности	+	+
Постраничный поиск и просмотр	+	+
Система ограничения доступа к содержимому на уровне документа	+	+
Развитая система ограничения прав доступа к документам на стороне пользователя	+	+

Система учета обращений пользователя к полным текстам	+	+
Учет морфологии русского языка при поиске	+	+
Хранение запросов	+	+
Интеграция с поисковыми системами Интернет	-	+
Самостоятельное присвоение уникальных идентификаторов URI каждому электронному объекту	-	+
Поддержка свободного лицензирования	-	+
Интеграция с Web-аналитикой поисковых систем	-	+
Интеграция с другими открытыми архивами в Интернет	-	+
Развитие в соответствии с развитием Интернет-технологий	-	+

Безусловным преимуществом программной платформы DSpace является ее неразрывная связь с информационной средой Интернет, использование ее поисковых механизмов, стандартов, протоколов и методов программирования, интеграция с поисковыми системами Интернет и их аналитическими инструментами. Все это обеспечивает прекрасные возможности для открытого доступа к документам в институциональном репозитории, чему в немалой степени также способствует поддержка свободного лицензирования Creative Common и присваивание каждому объекту в репозитории уникального идентификатора URI. Открытость программного обеспечения и непрерывный процесс его доработки кооперативными усилиями Интернет-сообщества позволяет репозиторию модифицироваться от версии к версии в соответствии с современными тенденциями развития информационных технологий, которые медленно, но верно продвигаются в направлении семантического поиска и связанных данных [7-8]. Программная платформа DSpace является свободно распространяемой, что стимулирует сотрудников отдела автоматизации к включенности в мировое информационное пространство, давно и активно осваиваемое зарубежными библиотеками [9].

Выводы. В целом современный этап развития информационных ресурсов, генерируемых библиотеками научных и научно-образовательных организаций можно охарактеризовать как нуждающийся в двух разнонаправленных программных платформах. А задача библиотеки в этом случае состоит в том, чтобы наладить взаимодействие между ними, взаимообмен данными и интеграцию поиска. Такой порядок работы позволяет библиотекам, с одной стороны, сохранить комфортную среду

формирования описаний электронных документов, с другой, обеспечить эффективное продвижение и интеграцию результатов научных исследований родной организации в общемировое научное пространство [10].

СПИСОК ЛИТЕРАТУРЫ

1. ГОСТ Р 7. 0. 96 — 2016. Электронные библиотеки. Основные виды. Структура. Технология формирования. — Введен 1.07.2017 г. — М.: Стандартинформ, 2016. — 17 с. — Режим доступа: http://www.rsl.ru/datadocs/doc_7667ve.pdf.
2. Электронно-библиотечные системы. Проект ГОСТ Р. — М.: Стандартинформ, 2017. — 19, [5] с. — Режим доступа: http://apoer.ru/media/doc/ebs_gost_20170214.pdf.
3. The effect of open access and downloads ('hits') on citation impact: a bibliography of studies // The Open Citation Project. — 2013. — Июнь. — Режим доступа: <http://opcit.eprints.org/oacitation-biblio.html/>.
4. Десять лет с Будапештской инициативой открытого доступа: устанавливая открытость. — 2012. — Режим доступа: <http://www.budapestopenaccessinitiative.org/boai-10-translations/russian>.
5. Bethesda Statement on Open Access Publishing. — 2003. — Режим доступа: <http://web.archive.org/web/20101211030116/http://www.earlham.edu/~peters/fos/bethesda.htm>.
6. Storer, Robert. Advantages and disadvantages of open access in libraries / Robert Storer. — 2015. — Апрель. — Режим доступа: <http://www.sirsiidynix.com/news/advantages-and-disadvantages-of-open-access-in-libraries>.
7. Warner, Simeon. Linked Data for Libraries: Experiments between Cornell, Harvard and Stanford / Simeon Warner. — 2015. — Режим доступа: <https://www.slideshare.net/simeonwarner/linked-data-for-libraries-experiments-between-cornell-harvard-and-stanford>.
8. UNIMARC и связанные данные / пер. В.Скворцова. — Режим доступа: http://www.nilc.ru/text/Other_publications/Other_publications17.pdf. — Пер. статьи: UNIMARC and linked data / Gordon Dunsire, Mirna Willer // 187 — Advancing UNIMARC: alignment and innovation — IFLA UNIMARC Programme (UNIMARC).
9. Godby, Carol Jean. From Records to Things: Managing the Transition from Legacy Library Metadata to Linked Data / Carol Jean Godby, Karen Smith-Yoshimura // Bulletin of the Association for Information Science and Technology. — 2017. — Vol.43. — № 2. — P. 18-23. — Режим доступа: https://www.asist.org/files/bulletin/dec-16/Godby_Yoshimura.pdf.
10. Enis, Matt. Open Future / Matt Enis // Library Journal. Library Systems Landscape 2017. — 2017. — Апрель. — Режим доступа: <http://lj.libraryjournal.com/2017/04/technology/open-future-library-systems-landscape-2017/>.