

ОТКРЫТЫЙ ДОСТУП И ОТКРЫТЫЕ АРХИВЫ ИНФОРМАЦИИ

УДК 004.891:378.4(04/07)

В. М. Московкин

Российско-украинские усилия в продвижении идей и технологий открытого доступа к научному знанию

Рассмотрен опыт реализации программ открытого доступа к научному знанию в России и Украине. Показано влияние создания университетских репозитариев открытого доступа на вебметрические рейтинги университетов.

В последние годы на постсоветском научном пространстве стали формироваться центры по продвижению идей и технологий открытого доступа (ОД) к научному знанию. Они сосредоточились в Москве (программа «Открытый доступ» Отделения общественных наук РАН), Белгороде (Программа «Открытый доступ» Приграничного белорусско-российско-украинского университетского консорциума) и Киеве (Программа «Электронная библиотека Украины» – один из региональных центров международной программы «Открытый доступ», реализуемой в рамках сети «Электронная информация для библиотек» – eIFL.net).

Рассмотрим деятельность последних двух центров, в которых, в отличие от первой – академической – программы, лидирующую роль играют университетские научные библиотеки.

В рамках программы «Открытый доступ» Приграничного белорусско-российско-украинского университетского консорциума (ПБРУУК), в который входят десять классических университетов из Гомеля, Белгорода, Воронежа, Курска, Ростова-на-Дону, Донецка, Запорожья, Луганска, Симферополя и Харькова, 30 января 2008 г. приняты Белгородская декларация об открытом доступе к научному знанию и культурному наследию и План действий по реализации ее основных положений.*

* Русско- и англоязычные тексты этих документов размещены в «Справочнике открытого доступа» (Open Access Directory, OAD): Declaration in support of OA (1008), http://oad.simmons.edu/oadwiki/main_Page.

Члены Консорциума, подписав эту Декларацию, приняли на себя ряд обязательств, главное из которых – создание по мере готовности университетских электронных архивов открытого доступа (*open access*) – репозитариев.

Первая крупная встреча Консорциума, прошедшая в рамках заседания его Координационного совета с приглашением руководителей библиотечных и информационных служб университетов Консорциума, состоялась 28 мая 2009 г. [1] На ней фактически одобрены и приняты тексты вышеназванных документов.

Программа «Открытый доступ» ПБРУУК – вторая крупная программа на постсоветском научном пространстве после программы «Открытый доступ к результатам исследований» для институтов Отделения общественных наук РАН [2], а сама Белгородская декларация – первая на этом пространстве.

После подписания Белгородской декларации первые репозитарии ОД созданы в мае и июне 2009 г. в Белгородском государственном университете (БелГУ) и Харьковском национальном университете им. В. Н. Каразина (ХНУ им. В. Н. Каразина), которые сразу же были зарегистрированы в мировых регистрах репозиторий ОД (ROAR, Саутгемптонский университет [3]; DOAR, Ноттингемский университет [4]). В то же время интерфейс с протоколом по сбору метаданных инициативы «Открытые архивы» (OAI-PMH [5]) в ROAR для обоих репозитариев ОД был подключен только в начале 2010 г., что позволяет наблюдать в реальном масштабе времени процесс их наполнения.

Рассчитанный Испанской киберметрической лабораторией в 2009 г. июльский вебметрический рейтинг университетов показал значительный рост таких рейтингов для вышеназванных университетов [6].

Здесь следует отметить, что адекватное представление всей многогранной деятельности университета на его сайте благоприятно сказывается на упомянутом выше рейтинге, который запущен в 2004 г. – первоначально для университетов и научно-исследовательских организаций. Позднее этим рейтингом были охвачены больницы (госпитали), бизнес-школы и репозитарии ОД. Последнее для нас очень важно, так как теперь не только университеты, но и их репозитарии вошли в систему глобального вебметрического позиционирования. Пока публикуются данные по TOP-400 репозитариев ОД.

Вебметрический рейтинг университетов и научно-исследовательских организаций был задуман как инструмент отслеживания влияния ОД и веб-политик на их позиционирование.

Суть вебметрической методологии оценки университетской активности состоит в использовании пяти поисковых машин – Google, Yahoo, Live Search, Exalead, Scholar Google – для измерения четырех индикаторов: SIZE (общее количество страниц), VISIBILITY (количество внешних ссылок), RICH FILES (количество файлов различных форматов), SCHOLAR (количество академических документов и их цитирование), их нормирования, взвешивания и агрегирования с получением суммарного рейтинга и ранга (места) университета.

Для расчета индикатора SIZE используются первые четыре поисковые машины; для VISIBILITY – вначале вторая, третья и четвертая, а сейчас происходит постепенный переход к одной машине (Yahoo); для RICH FILES – Google; для SCHOLAR – Scholar Google.

Для вхождения в систему вебметрических расчетов необходимо иметь только автономный домен (URL-адрес). В 2004 г. при запуске вебметрического ранжирования университетов их было 6–7 тыс. с публикацией TOP-5000; в 2010 г. эта база данных расширилась до 18 тыс. университетов с публикацией TOP-8000.

В течение двух лет мы отслеживаем динамику вебметрических рейтингов университетов ПБРУУК и видим, что они очень чувствительны к трансформации и модернизации сайтов (табл. 1, 2), а также к наличию и объему контента репозитариев ОД.

Эта работа проводится нами в рамках системы университетского вебметрического бенчмаркинга, основанной на испанском вебметрическом рейтинге, нашей методологии оценки университетской публикационной активности, базирующейся на возможностях поисковой машины Scholar Google и оценке влияния ОД и веб-политик на вебметрические показатели университетов [7–9]. Бенчмаркинг, т.е. мониторинго-сравнительная процедура, нацеленная на выработку лучшей практики вебметрического позиционирования университетов, проводится нами в рамках программы «Открытый доступ» ПБРУУК.

Динамика вебометрических рейтингов университетов ПБРУУК

Университеты	Мировые вебометрические рейтинги					Изменения рейтинга за 0,5 года			Изменения рейтинга за 2 года
	Январь 2008	Июль 2008	Январь 2009	Июль 2009	Январь 2010	06.08-01.08	01.09-06.08	06.09-01.09	
ГГУ	Ниже 5 000	Ниже 5 000	Ниже 5 000	5 805	Ниже 6 000				
ВГУ	2 370	2 377	2 625	1 752	2 024	-7	-248	873	346
ЮФУ	4 743 ¹	3 036	2 579	1 286	1 867	1 707	457	1 293	2876
БелГУ	3 802	3 660	4 108 ²	2 511	3 491 ⁴	142	-448	1 597	311
ХНУ	2 946	2 305	2 629 ³	2 570	2 342	641	-324	59	604
ВУНУ	Ниже 5 000	Ниже 5 000	Ниже 5 000	5 590	5 819			-229	

Примечания к табл. 1:

¹ Низкий рейтинг связан с реорганизацией (укрупнением) университета и трансформацией его сайта (изменение URL-адреса).

² С сентября 2008 г. произошла трансформация сайта БелГУ, при этом, несмотря на сохранение его URL-адреса, для поисковых машин стали плохо доступны ресурсы прежнего сайта, которые автоматически не перешли на новый сайт.

³ Понижение рейтинга связано с модернизацией сайта.

⁴ Понижение рейтинга связано с небольшой модернизацией сайта «наука» и изменением методики расчета индикатора «VISIBILITY»; в 2 раза увеличилась база данных по университетам мира.

Сокращения:

ГГУ – Гомельский государственный университет;

ВГУ – Воронежский государственный университет;

ЮФУ – Южный федеральный университет;

БелГУ – Белгородский государственный университет;

ХНУ – Харьковский национальный университет им. В. Н. Каразина;

ВУНУ – Восточно-Украинский национальный университет им. В. И. Даля.

Структурная динамика вебометрических рейтингов университетов ПБРУУК

Дата, университеты	Составляющие вебометрического рейтинга			
	SIZE	VISIBILITY	RICH FILES	SCHOLAR
Июль, 2008 г.				
ВГУ	1 499	4 046	2 462	660
ЮФУ	4 781	3 429	2 975	2 709
БелГУ	2 216	5 793	3 098	2 263
ХНУ	2 724	2 648	3 572	2 104
Январь, 2009 г.				
ВГУ	2 053	4 415	2 165	564
ЮФУ	4 199	2 708	2 905	2 331
БелГУ	2 851	6 475	2 347	2 501
ХНУ	3 041	3 158	3 758	1 875
Июль, 2009 г.				
ВГУ	1 958	3 970	2 295	587
ЮФУ	2 285	640	2 375	2 250
БелГУ	3 367	4 257	3 083	1 352
ХНУ	3 227	4 066	3 362	1 547
Январь, 2010 г.				
ВГУ	2 671	3 615	2 295	548
ЮФУ	2 904	2 212	1 981	1 486
БелГУ	4 604	6 577	2 151	1 498
ХНУ	2 732	3 445	2 859	1 300
Изменения составляющих рейтинга за 1,5 года				
ВГУ	-1 172	431	167	112
ЮФУ	1 877	1 217	994	1 223
БелГУ	-2 388	-784	947	765
ХНУ	-8	-797	713	804

Для дальнейшего продвижения идей ОД и создания репозитариев ОД на белорусско-российско-украинском приграничном университетском пространстве было намечено провести вторую крупную встречу под эгидой ПБРУУК с приглашением ведущих украинских специалистов и менеджеров по ОД из Киева.

Эта встреча – научно-практический семинар «Открытый доступ к научной и образовательной информации: создание институциональных репозитариев и использование мировых электронных ресурсов» – состоялась 11 марта 2010 г. в Центральной научной библиотеке ХНУ им. В. Н. Каразина [10], которая имеет более чем двухсотлетние традиции. На семинаре, по образному выражению директора этой библиотеки Ирины Журавлевой, «встретились две близкие по духу программы: «Электронная библиотека Украины» (ELibUkr; <http://www.elibukr.org>) и «Открытый доступ» ПБРУУК, что обусловило огромный интерес к этой встрече и ее эффективность.

В первом докладе на основе представленной выше методологии университетского вебометрического бенчмаркинга показано, как влияет создание репозитариев ОД на вебометрические рейтинги университетов (на примере DSpace репозитариев БелГУ и ХНУ им. В. Н. Каразина). Их создание существенно повышает значение индикаторов RICH FILES и SCHOLAR испанского вебометрического рейтинга университетов за счет полнотекстовых pdf-файлов научных работ. Запуск в мае и июне 2009 г. репозитариев ОД соответственно в БелГУ и ХНУ им. В. Н. Каразина повысил в июле 2009 г. вебометрический рейтинг этих университетов, причем наибольший вклад в это повышение внесли вышеуказанные индикаторы.

В январе 2010 г. впервые с момента запуска испанского вебометрического рейтинга университетов (2004 г.) произошла модификация методологии расчетов, что негативно сказалось на динамике вебометрических рейтингов всех университетов ПБРУУК. Наши контакты с научным директором Испанской киберметрической лаборатории Исидро Агуилло показали, что в январских (2010 г.) расчетах наиболее важный индикатор VISIBILITY (количество внешних ссылок), которому придается 50-процентный вес, рассчитывался не с помощью трех поисковых машин – Yahoo, Exalead, Live Search, как было ранее, а с помощью первых двух. Причем в следующих июльских расчетах планируется использовать единственную поисковую машину Yahoo.

Далее разработчиками рейтинга были исключены внешние ссылки с доменов org, net и com. По словам Исидро Агуилло, это делается для того, чтобы уменьшить брешь между академическим цитированием и анализом внешних ссылок (*inlink analysis*). Он сообщил нам, что в следующих расчетах будет сделан переход к G-фактору (<http://www.universitymetrics.com/g-factor>), основанному на TOP-1000 университетов вебометрического рейтинга. Отсюда следует, что индикатор VISIBILITY будет рассчитываться на основе внешних ссылок с доменов только университетов, входящих в TOP-

1000. Это еще в большей степени, на наш взгляд, ухудшит вебометрическое позиционирование постсоветских университетов.

В январском ранжировании 2010 г. была более чем в два раза расширена база данных по университетам мира, что привело к мощному конкурентному давлению и тем самым подорвало позиции отслеживаемых нами университетов ПБРУУК.

В качестве активизации процесса самоархивирования публикаций самими учеными нами предложено как можно быстрее принять мандаты ОД (на уровнях ПБРУУК и его университетов), предусматривающие обязательное размещение в университетских репозитариях ОД результатов НИР, выполненных за счет бюджетного финансирования.

Вице-президент Национального университета «Киево-Могилянская академия» (НаУКМА, Киев) Татьяна Ярошенко рассказала о программе «Электронная библиотека Украины» (ELibUkr) в контексте формирования центров знаний в украинских университетах, а также о проекте фонда «Возрождение» по созданию междисциплинарного открытого электронного архива (ELibrUkr – ОА-мультидисциплинарний відкритий електронний архів; <http://oa.elibukr.org/>), в котором предполагается регистрация университетов, после чего их ученые могут создавать коллекции своих полнотекстовых публикаций с размещением CV. В этом цифровом репозитарии уже зарегистрированы Харьковский национальный экономический университет, Одесская национальная юридическая академия, Волынский национальный университет, Академия сухопутных сил Украины, Академия таможенной службы Украины, Национальный технический университет «Харьковский политехнический институт», Национальная медицинская академия последипломного образования.

Менеджер программы «Открытый доступ» (eIFL.net – *Electronic information for Libraries*; <http://plip.eifl.net/sections/home>) Ирина Кучма (Киев) представила обзор новейших исследований по влиянию ОД на цитирование, рассказала о новых метриках оценки качества научных статей. Она отметила, что благодаря мощному международному движению за открытый доступ к научному знанию значительно сократился интервал времени между публикациями статей и их цитированием. Огромный вклад в это движение внесли ученые, работающие в области физики высоких энергий в Лос-Аламосской лаборатории, которые в 1991 г. создали первый и самый крупный репозитарий ОД, насчитывающий около 600 тыс. pdf-файлов с полнотекстовыми статьями (<http://arxiv.org/>).

Ирина Кучма также отметила ключевой вклад лауреата Нобелевской премии Гарольда Вармуса (бывшего директора Национальных институтов

здоровья США) в это движение – создание им публичного медицинского хранилища информации (PubMedCentral, PMC; www.pubmedcentral.nih.gov) и публичной научной библиотеки (PLOS; <http://www.plos.org/>). Последняя на основе огромного количества научной информации медико-биологического профиля разработала специальную метрику оценки качества научной статьи (<http://article-level-metrics.plos.org/>). Аналогичную метрику – «MESUR» – разработала Лос-Аламосская лаборатория, основываясь на огромном массиве физико-математической научной информации собственного электронного архива (www.mesur.org/services).

Обе метрики представляют собой синтез традиционных индексов цитирования с показателями загрузки статей в Интернете.

Отмечено, что в области физики высоких энергий 20% цитирований приходится на препринты. Это связано, на наш взгляд, с тем, что в arxiv.org очень большая доля научных публикаций представлена именно такими документами. (Напомним: Филдовская премия по математике присуждена Григорию Перельману благодаря его препринтным – нецензурированным – публикациям, размещенным в этом электронном архиве ОД.)

Ирина Кучма рассказала также о двух интересных работах, сделанных А. Сван (*A. Swan*). В первой – обобщены результаты 31 исследования по оценке влияния ОД на цитирование и показано, что в 27 из них отмечено положительное влияние на цитирование ОД [11].

Подчеркнем: согласно первым исследованиям Стивена Харнада и его единомышленников, самоархивирование в репозиториях ОД ранее опубликованных статей увеличивает их цитируемость в 4–5 раз.

Во второй работе выделены четыре типа британских университетов и получены различные сценарии экономии средств (возврата инвестиций) от ОД [12]. Для крупного исследовательского университета (четвертый тип университетов) экономия достигает 5 млн фунтов стерлингов в год. Предложенная методика расчета прозрачна и может быть использована для тестирования отечественных университетов.

Ирина Кучма отметила также *G-Factor International University Ranking*, о котором упомянуто выше в связи с вебметрическим ранжированием университетов ПБРУУК. Она кратко рассказала о программе EOS (*Enabling Open Scholarship*; <http://www.openscholarship.org/>), в которой ректоры с сильными ОД-компетенциями помогают ректорам других университетов.

В настоящее время Ирина Кучма – модератор крупнейшего форума открытого доступа – BOAI (<http://www.soros.org/openaccess>). Она предложила всем принять участие в Международной неделе открытого доступа (18–24 октября 2010 г.; <http://www.openaccessweek.org/>).

Заместитель директора Научной библиотеки БелГУ Лариса Анищенко рассказала об этапах подготовки, принятия и реализации Белгородской декларации об открытом доступе к научному знанию и культурному наследию, которая, как отмечено выше, – первая подобная декларация на постсоветском научном пространстве.

Заместитель директора Научной библиотеки НаУКМА Оксана Бруй рассказала о политике ОД своего университета, суть которой состоит в решениях Ученого совета рекомендовать размещать научные труды в открытом доступе (в репозитории и журналах ОД) и в обязательном порядке самоархивировать результаты исследований, поддержанные грантами фонда «Возрождение» согласно его политике ОД. Она отметила, что было очень трудно убедить ученых самоархивировать их публикации в цифровом репозитории. Поэтому ставка была сделана на выявление ученых-энтузиастов.

При создании DSpace репозитория (<http://www.eknair.ukma.kiev.ua/>) были заранее прописаны все возможные его разделы (сообщества и коллекции), даже если не было желающих размещать в них публикации. После нескольких лет эксплуатации репозитория ОД установлена новая версия программного обеспечения DSpace, что положительно повлияло на вебметрический рейтинг университета, который сейчас лучший среди украинских вузов.

Пополнением репозитория занимаются три библиотечных работника; координатор этого репозитория устанавливает связи со всеми учебными и научными подразделениями университета. Представлены данные, полученные с помощью Google Analytics за период с 01.01.2010 по 09.03.2010, по посещению репозитория ОД (отмечены пользователи более 50 стран).

Проректор Харьковской национальной академии городского хозяйства Николай Пан рассказал о функционировании своего репозитория ОД – самого крупного в Украине (11 350 записей на момент проведения семинара; <http://eprints.ksame.kharkov.ua/>). К сожалению, вебметрический рейтинг этого вуза не рассчитывается, так как он не имеет автономного домена, что и отметил докладчик. Этому вузу предстоит болезненный этап перехода на автономный домен. Докладчик продемонстрировал результаты тестирования репозитория ОД с помощью Google Analytics.

Здесь, на наш взгляд, возникает интересная задача построения и анализа совмещенных временных графиков наполнения репозитория ОД (аналитический инструмент ROAR) и его посещения внешними пользователями (инструмент Google Analytics).

В заключение директор ЦНБ ХНУ им. В. Н. Каразина Ирина Журав-

лева представила результаты анкетирования пользователей этой библиотеки, в том числе и по поводу востребованности идей и инструментов ОД. Она пригласила всех участников семинара принять активное участие в крупной Международной научно-практической конференции «Библиотеки высших учебных заведений Украины на пути в международное информационное пространство» (26–28 мая 2010 г.; ЦНБ ХНУ им. В. Н. Каразина).

На «круглом столе» семинара «*Проекты открытого доступа в Украине и России как составляющие мировой науки*» Ирина Кучма сделала презентацию «*Институциональный репозитарий. С чего начать*». По завершении презентации ведущие «круглого стола» – Ирина Кучма, Владимир Московкин и Татьяна Ярошенко – ответили на все вопросы участников семинара.

На «круглом столе» отмечено, что если ранее на создание репозитария ОД уходило полгода или год, то теперь его можно создать за неделю; вся проблема заключается в его наполнении. Прозвучал вопрос о выборе программного обеспечения ОД – DSpace и Eprints, которые сейчас наиболее популярны. Ответ: последний – удобнее для ученых благодаря более понятной и привычной для них классификации разделов по научным областям. В DSpace представлены менее понятные «сообщества» (*communities*), в рамках которых следует создавать различные коллекции, не привязанные к научным областям.

Оксана Бруй отметила, что в репозитарии ОД НаУКМА реализован инструмент, позволяющий показывать новые поступления в него. Она дала информацию о новом протоколе SWORD (<http://www.swordall.org/>), который позволяет переносить материалы из одного архива в другой, однако пока процесс переноса одного pdf-файла вместе с записями занимает слишком много времени – до 20 минут. Но здесь следует отметить, что, согласно концепции испанского вебметрического рейтинга университетов, размещение чужих материалов считается плохой практикой, и университеты, злоупотребляющие этим, штрафуются. Хорошая практика для университетских репозитариев ОД, согласно Испанской киберметрической лаборатории, демонстрировалась в нашей презентации.

Все презентации семинара размещены в DSpace репозитарии ХНУ им. В. Н. Каразина (<http://dspace.univer.kharkov.ua/>).

СПИСОК ИСТОЧНИКОВ

1. **Васильева О. А.** Добро пожаловать в электронный архив БелГУ. – Режим доступа: http://www.bsud.ru/bsu/news.php?ID=46366&IBLOCK_ID=176
2. **Открытый архив результатов исследований ЦЭМИ РАН** – Режим доступа: <http://www.cemi.rssi.ru/archive/>
3. **Registry of Open Access Repositories (ROAR)**. – Режим доступа: <http://roar.eprints.org/>
4. **The Directory of Open Access Repositories (Open DOAR)**. – Режим доступа: <http://opendoar.org/>
5. **The Open Archives Initiative Protocol for Metadata Harvesting**. – Режим доступа: <http://www.openarchives.org/OAI/openarchivesprotocol.html>
6. **Ranking Web of World Universities**. – Режим доступа: <http://www.webometrics.info/>
7. **Московкин В. М.** Методология оценки и структурного анализа университетской публикационной активности с помощью поисковой машины Google Scholar (на примере Приграничного белорусско-российско-украинского университетского консорциума) // Университетское управление: практика и анализ. – 2009. – № 2. – С. 45–52.
8. **Московкин В. М.** Возможности использования поисковой машины Google Scholar для оценки публикационной активности университетов // ИТИ. Сер. 1. – 2009. – № 7. – С. 12–16.
9. **Московкин В. М.** Вебметрическая оценка публикационной активности университетов: влияние Белгородской декларации // Там же. – 2010. – № 2. – С. 12–16.
10. **Васильев О. А.** Электронный архив: выгода очевидна. – Режим доступа: <http://library.bsud.edu.ru/Events/2010/17032010.asp>
11. **Swan A.** The Open Access citation advantage: Studies and results to date. Technical Report // School of Electronics & Computer Science, University of Southampton, 2010. – Режим доступа: <http://eprints.ecs.soton.ac.uk/18516/>
12. **Swan A.** Modelling Scholarly Communication Options: Costs and benefits for universities // Project Report. UNSPECIFIED. (Unpublished), 2010. – Режим доступа: <http://ie-repository.jisc.ac.uk/442/>