

Развитие электронных библиотек – путь к Открытой Науке *

© С.И. Паринов

Центральный экономико-математический институт РАН
sparinov@gmail.com

Аннотация

Постоянное развитие концепции электронной библиотеки (ЭБ) формирует особую онлайн-среду для научных исследований. Успехи международной научной инициативы за открытый доступ к результатам исследований создают предпосылки для того, чтобы использование ЭБ и связанной с ними онлайн-научной среды стало повсеместным. Все вместе это порождает ряд новых обстоятельств для разработчиков ЭБ. Если в будущем ЭБ всех исследовательских организаций будут объединены в единое информационное пространство, то как научное сообщество может получить максимальные выгоды от использования возникающего системного эффекта? Как должна быть построена система фиксации и распространения результатов исследований, содержащихся в локальных ЭБ, чтобы обеспечивать их максимально возможное использование? Как мы должны организовать процесс использования результатов исследований и какие средства требуется создать, чтобы появились условия для сбора максимально полной и точной статистики о характере использования и научном влиянии результатов исследований? Как должны ЭБ и виртуальная научная среда собирать и обрабатывать статистику, чтобы показатели результативности исследований обладали существенно более высоким качеством по сравнению с текущими. Анализ возможных направлений развития ЭБ в данном контексте показал, что формируется комплекс условий для появления новой более эффективной организационной формы научно-исследовательской деятельности, которую мы назвали Открытой Наукой.

1 Введение

В данной работе рассмотрено развитие электронных библиотек (ЭБ) ограничено случаями их применения для обеспечения процесса **научных исследований**.

Современные процессы развития ЭБ создают особую виртуальную среду для научных исследований. В этих процессах можно выделить три логически связанные компоненты:

1) развитие программно-технических средств поддержки исследований (новый инструментарий);

2) использование новых инструментальных средств для совершенствования методов работы исследователей и их профессиональных взаимодействий (новые практики);

3) повышение эффективности организационных форм научно-исследовательской деятельности, использующее новый инструментарий и новые методы работы (новая форма организации сообщества исследователей).

В настоящее время формирование особой виртуальной среды для исследований наиболее заметно по многочисленным проектам интеграции содержаний ЭБ отдельных исследовательских организаций в профессиональные информационные пространства.

Наличие такого информационного пространства (как правило, на принципах федеративного объединения метаданных множества локальных ЭБ) и создаваемый им системный эффект являются отправными точками для нашего анализа.

Под профессиональным информационным пространством мы понимаем результат интеграции (на уровне метаданных) разнородных, организационно и географически распределенных источников научных данных. В минимальной конфигурации информационное пространство должно иметь единую систему навигации и поиск по накапливаемой структуре метаданных.

Кроме этого, содержание информационного пространства должно давать репрезентативное представление информационных ресурсов и деятельности соответствующего профессионального сообщества, а также иметь регулярную синхронизацию содержания

пространства с содержанием его локальных ЭБ как источников данных.

На уровне отдельных ЭБ подобные тенденции к интеграции проявляются в появлении в составе ЭБ так называемых "технических входов", предназначенных для доступа внешних программных роботов к переносу открытого контента ЭБ (как правило, только метаданные) в базы данных соответствующих информационных пространств.

В результате объединения в центральной базе метаданных достаточно большой доли производимых научным сообществом электронных материалов возникает своего рода "системный эффект", который должен быть использован в интересах данного сообщества.

Приложения, создаваемые для использования подобного системного эффекта, уже трудно классифицировать как ЭБ. Они создаются и развиваются как элементы онлайн-научной инфраструктуры во внешней по отношению к множеству ЭБ программно-технической среде.

Во втором разделе статьи рассматриваются 10 направлений наиболее важного, на наш взгляд, развития программно-технических средств поддержки исследований, которые создают и используют системный эффект профессиональных информационных пространств.

Существующее в международном научном сообществе мощное общественное движение за открытый доступ к результатам исследований популяризирует новые методы работы исследователей, прямо связанные с использованием современных ЭБ и профессиональных информационных пространств. Эти усилия постепенно формируют необходимые организационные предпосылки для того, чтобы использование ЭБ и связанной с ними онлайн-научной инфраструктуры стало повсеместным.

В третьем разделе статьи дается описание текущего состояния этих процессов.

В свою очередь, повсеместное использование учеными подобной виртуальной научной среды создает необходимые условия для реализации новой более эффективной организационной модели научной-исследовательской деятельности, которую мы назвали "Открытая Наука".

Открытая Наука возникает как результат реализации следующего комплекса программно-технических и организационных инноваций:

а) свободный доступ к результатам всех открытых научных исследований (кроме результатов, имеющих закрытый характер по коммерческим соображениям или связанных с безопасностью);

б) интеграция результатов исследований в онлайн-научную инфраструктуру, которая сконструирована для максимально широкого и полного использования этих результатов;

в) автоматический мониторинг онлайн-информационной активности ученых, формируемая

на этой основе открытая онлайн-научно-метрическая статистика и рассчитываемые на ее основе публичные показатели результативности ученых и исследовательских организаций;

г) использование онлайн-научно-метрических показателей в процедурах принятия решений о финансировании научной деятельности, включая персональные надбавки ученых.

В четвертом разделе рассматриваются предпосылки формирования Открытой Науки, исходя из логики развития ЭБ и связанного с этим развития виртуальной среды для научной деятельности.

Обсуждаемые положения, где это возможно, иллюстрируются примерами из системы Соционет (<http://socionet.ru/>) и других российских и зарубежных проектов.

2 Современный контекст применения и развития ЭБ

Современные процессы развития ЭБ вносят существенный вклад (как прямой, так и косвенный) в формировании новой виртуальной среды для научных исследований. Общий контекст и основные направления развития этой среды можно, на наш взгляд, охарактеризовать следующими девятью факторами.

2.1 Трансформация ЭБ в CRIS

Наблюдается постепенная трансформация ЭБ в CRIS (CRIS – Current Research Information System, см. подробнее в [1-2]), которая часто проявляется в форме "глубокой" интеграции ЭБ с корпоративными научными информационными системами. Главной особенностью этого направления развития является лучшее приспособление ЭБ к комплексным потребностям организации, в которых они созданы. С одной стороны, это проявляется как рост ЭБ в "вширь" за счет обработки сервисами ЭБ новых типов информационных объектов, таких как информационные профили авторов, исследовательских организаций – мест работы авторов, описания научных проектов и другие типы, имеющие те или иные содержательные связи с традиционными для ЭБ научными статьями и материалами. С другой стороны, ЭБ обмениваются данными, а иногда и становятся органичной частью комплексных информационных систем управления научно-исследовательской деятельностью и автоматизации бизнес-процессов организации.

В контексте данной статьи главным положительным следствием этой тенденции является формирование в ЭБ более полного и точного информационного представления существующей в научном сообществе структуры деятельности и действующих лиц. Например, отдельных исследователей – авторов статей, групп ученых – участников проектов, исследовательских организаций – направлений исследований и т.п.

Благодаря этому ЭБ становятся носителями информационных объектов, дающих комплексную и репрезентативную картину результатов деятельности и структуры соответствующих исследовательских организаций.

2.2 Интеграция локальных ЭБ

Наблюдается постоянный рост интереса к интеграции локальных ЭБ (т.е. принадлежащим отдельным организациям), друг с другом. Это относится, в первую очередь, к объединению контента ЭБ, а также к возможностям интеграции отдельных сервисов ЭБ. Интеграция, как правило, имеет форму конфедерации, т.е. предполагает сохранение полного контроля владельцев ЭБ над своими данными в объединенном информационном пространстве. Стимулом к объединению локальных научных информационных ресурсов является возникновение "системного эффекта", когда результирующее информационное пространство позволяет получить научному сообществу выгоды, недоступные от использования локальных ЭБ порознь. Данное направление развития проявляется в том, что в интерфейсах локальных ЭБ начинают создаваться "технические входы", предназначенные для программных роботов, которые в автоматическом режиме собирают через них метаданные об открытых для интеграции разделах ЭБ. Среди способов оформления "технических входов" в научной среде в настоящее время наиболее популярным является протокол OAI-PMH, разработанный международной инициативой Открытые Архивы [3]. Все большее количество ЭБ получают OAI-PMH интерфейс, что привело к появлению специальных названий "открытый архив" или "открытый репозиторий" (на Западе наиболее распространено название Institutional Repository или IR), обозначающих информационные ресурсы доступные в ЭБ через протокол OAI-PMH. В некоторых случаях OAI-PMH дополняется и другими подобными средствами для экспорта данных, например, RSS 2.0.

Примеры – Соционет (<http://socionet.ru/>), ЕНИП РАН (<http://enip.ras.ru/>), OAIster (<http://oaister.org/>), DRIVER (<http://www.driver-repository.eu/>), OpenDOAR (<http://www.openoar.org/>) и др.

Как следствие данных процессов, растет общая открытость существующих научных материалов, и создаются благоприятные условия для их сбора и обработки в автоматическом режиме.

2.3 Появление информационных хабов

Среди ЭБ выделился особый вид, который получил название "информационный хаб" [4]. Данная разновидность ЭБ возникла в ответ на потребность обеспечить автоматический сбор и актуализацию разнородных метаданных из распределенных локальных ЭБ в единую стандартизованную базу данных, подготовив их для разнообразного использования всеми

заинтересованными лицами. Для этого от информационного хаба требуется:

- развитая модель структуры данных, чтобы принять без потерь в центральную базу метаданных содержание большого количества разнородных ЭБ (в настоящее время, на наш взгляд, CERIF является наиболее подходящей моделью, см. <http://www.eurocris.org/cerif/cerif-releases/cerif-2008/>);

- возможность гибкой настройки на формат данных локальных ЭБ для переноса их открытого контента в центральную базу метаданных (решается созданием программных конвертеров под каждый конкретный случай);

- достаточно простой и гибкий способ экспорта (выгрузки) необходимых тематических выборок и коллекций из интегрированных стандартизованных метаданных для их последующей обработки (дать возможность выборок на основе популярных протоколов, например, RSS и OAI-PMH).

В последние два десятилетия происходила конкурентная борьба за модель информационного хаба между двумя основными подходами, один известен, как протокол Z39.50, второй – многим известен как подход RePEc (repec.org). К настоящему времени, на наш взгляд, подход физического копирования и синхронизации метаданных с сайтов-источников в общую централизованную базу данных, реализованный в проекте RePEc, выглядит предпочтительным и более востребованным.

Пример: система Соционет (<http://socionet.ru/>) является полноценным информационным хабом, поскольку предоставляет разработчикам средства не только для загрузки данных в базу (это делают все системы, перечисленные в примерах предыдущего подраздела), но и для выгрузки [4].

2.4 Создание информационных пространств

Профессиональные информационные пространства (см. определение в начале статьи) создаются на основе метаданных, собираемых информационным хабом через технические входы заданного множества ЭБ, и предназначены для визуализации текущей структуры и содержания объединенных информационных ресурсов. Они дают пользователям различные средства навигации и поиска по собранному контенту, а также предлагают средства для разнообразного использования собранных метаданных, включая создание новых информационных продуктов и услуг. Один из основных принципов подобных информационных пространств – создатель метаданных полностью контролирует их содержание. Посторонние лица не могут вносить изменения в структуру и содержание чужих метаданных, но могут создавать вторичные (обогатенные) метаданные на основе исходных и могут устанавливать связи от своих метаданных к любым другим (см. ниже подраздел 2.6).

Примеры: все системы, перечисленные в примерах подраздела "Интеграция локальных ЭБ" создают профессиональные информационные пространства, но с разным набором функциональности для навигации и поиска.

2.5 Формирование онлайн-научной инфраструктуры

Наличие информационного хаба и визуализация его содержимого через информационное пространство создают для разработчиков ЭБ качественно новые условия. Все большая часть существующих разнородных и исходно распределенных метаданных, представляющих содержание научных информационных ресурсов, собирается в одном месте в стандартизованном виде. Собранные метаданные открыты для использования и обработки внешними программными средствами и сервисами. Результаты обработки исходных метаданных внешними программами и сервисами, если они являются научными информационными ресурсами, также могут передаваться (только метаданные) в информационный хаб и становятся доступны конечным пользователям в составе единого информационного пространства. Поскольку как первичные метаданные, так и вторичные, полученные в результате обработки первых, аккумулируются в одном месте, возникает эффект постепенного обогащения и развития научной информационной среды. Имеет место своего рода кругооборот создания "добавочной стоимости": разработчики взяли необходимые метаданные (М), обработали их сервисами (С), создали новые метаданные (М'), которые вернулись в информационный хаб, отобразились в информационном пространстве и стали "сырьем" для следующего цикла М-С-М'.

Здесь есть два важных новых обстоятельства:

а) Упомянутые в этом пункте "внешние программные средства и сервисы" уже не имеют явной привязки к определенным ЭБ, т.к. они работают с результатом виртуальной интеграции разнородных информационных ресурсов множества локальных ЭБ в массив стандартизованных метаданных. Поэтому, на наш взгляд, такие сервисы целесообразно классифицировать как часть онлайн-научной инфраструктуры. Информационные хабы и средства создания информационных пространств также скорее являются элементами онлайн-научной инфраструктуры, чем ЭБ.

б) Поскольку в схеме обработки М-С-М' циркулируют общедоступные метаданные, то при наличии необходимых программных средств их обогащение и развитие может быть выполнено уже не только создателями этих метаданных, но любым пользователем. Средства для развития/обогащения метаданных могут быть как частью некоторой ЭБ, так и входить в состав онлайн-научной инфраструктуры. Процесс развития метаданных, если он выполняется

не их создателем, может быть организован как создание и включение в информационный хаб "улучшенных копий" (при этом оригинал улучшенных метаданных всегда сохраняется, т.к. он находится под полным управлением только своих создателей) или как создание связей между дополнительными информационными объектами и улучшаемыми метаданными.

Полноценный анализ направлений извлечения потенциального "системного эффекта" от интеграции научных метаданных еще впереди, но некоторые выгоды уже очевидны. Далее рассматриваются некоторые уже практически разрабатываемые примеры использования системного эффекта, создаваемого научным информационным пространством.

2.6 Формирование сетей связей между объектами информационного пространства

Одна из возможностей развития и "обогащения" метаданных, входящих в информационное пространство, сформированное на принципах федерации – создание инфраструктурных сервисов, позволяющих пользователям устанавливать связи между информационными объектами, собранными в центральную базу метаданных из множества различных ЭБ. Например, исследователь может создавать связи между своим профессиональным профилем, находящимся в ЭБ организации – места его работы, со своими статьями, находящимися в ЭБ других организаций (издательство журналов, конференций, ВУЗов и т.п.).

В самом общем виде это может означать создание сервисов для развития учеными своих профессиональных социальных сетей. Например, установление разнотипных сетей связей с метаданными родственных информационных объектов, развитие связей электронного цитирования, создание нового типа связей для визуализации профессионального влияния, и т.п.

Средства для поддержки целенаправленного формирования учеными профессиональных социальных сетей не являются принципиально новыми, т.к. в широком смысле для ученого это – постоянная деятельность по его профессиональной самоидентификации.

Центрами формирования такого рода связей являются следующие основные для научной деятельности виды информационных объектов: 1) карточка статьи/материала/проекта, состоящая из описательных данных (метаданные), 2) персональный профиль ученого, 3) профиль исследовательской организации или коллектива проекта.

В процедурах формирования профессиональных социальных сетей в системе Соционет между данными тремя видами объектов, а также и внутри отдельных видов создаются следующие наиболее важные конфигурации связей:

- между ученым (его персональным профилем) и организацией (профилем) – местом его

работы, это позволяет учитывать при расчете наукометрических показателей для организации соответствующие индикаторы ее сотрудников и наоборот;

- между ученым и его авторскими электронными материалами (в институтских ЭБ, научных журналах и т.п.), что позволяет при расчете наукометрических показателей ученого учитывать характеристики его материалов;

- между авторскими электронными материалами ученого и другими материалами, которые были им использованы (процитированы) при подготовке своего, что позволяет устанавливать содержательные связи между научными материалами и подсчитывать уровень их использования в научном сообществе;

- между персональными профилями ученых и статьями других ученых, как связи, отражающие индивидуальное профессиональное признание данными учеными выделенных ими статей своих коллег, что позволяет оценивать влияние (слабая форма использования) результатов исследований на научное сообщество.

Механизмы электронного депонирования материалов в системе Соционет позволяют автоматически включать в состав карточек статей/материалов (в метаданные) связи с профилем автора и его места работы. Таким образом, в процессе оформления и передачи материалов в институтскую ЭБ автоматически формируется определенная часть профессиональной социальной сети.

2.7 Развитие связей цитирования

Важной является задача развития и улучшения качественных характеристик связей между материалами, отражающими использование некоторым ученым результатов исследований, полученных другими учеными, в процессе создания им нового научного знания. В текущей научной практике эти связи устанавливаются с помощью общепринятых схем цитирования научных материалов или их фрагментов.

Одним из следствий развития в рамках ЭБ средств электронного депонирования (размещения статей и материалов в открытом доступе) является то, что традиционное научное цитирование постепенно обретает форму «электронного цитирования» с рядом новых возможностей. Схема электронного цитирования, если она реализована в ЭБ, может позволять автору электронной статьи уточнить в каком именно качестве он использовал чужие результаты исследований. Это может быть реализовано включением в связи цитирования определенных качественных атрибутов, которые в свою очередь должны быть сконструированы так, чтобы позволять их автоматическое распознавание и обработку для выявления различных характеристик использования научных знаний.

Если исходить из того, что такие характеристики должны давать данные для получения более точных

индикаторов использования результатов исследований, то возможен следующий список характеристик электронного цитирования:

1. «основание для получения моих результатов», что означает прямое научное использование автором цитируемого результата;

2. «подтверждение цитируемого результата», т.е. результат автора подтверждает цитируемый;

3. «цитируемый результат подтверждает результат, ранее полученный автором», т.е. автор утверждает о своем приоритете по отношению к независимо полученному цитируемому результату;

4. «близкий или связанный результат», означающее, что автор в определенном смысле повторил цитируемый результат;

5. «иллюстрация моих выводов», означающее определенную логическую связь между результатом автора и цитируемым;

6. «объект для критики», означает, что автор подвергает сомнению цитируемый результат.

Необходимо уточнить, что возможны случаи, когда автор при оформлении связи цитирования может одновременно отметить несколько характеристик. Для разъяснения подобных случаев автор может использовать поле «комментарии» для каждой связи цитирования.

При наличии подобной статистики о характеристиках цитирования возможно создание (в рамках онлайн-инфраструктуры) процедур для автоматического построения индикаторов, которые позволят более точно, чем известные традиционные индексы цитирования, выявлять как для отдельной статьи, так и всего корпуса научных результатов:

1. какие результаты кем используются как основа научного вывода;

2. какие результаты подтверждают или подтверждаются другими результатами;

3. какие результаты повторяют уже известные;

4. упоминание результатов в качестве общих иллюстраций;

5. какие результаты кем критикуются и имели ли данная критика позитивные для науки последствия.

Подобная модель электронного цитирования превращает практику научного цитирования в механизм формирования между учеными профессиональной социальной сети. Одна из дополнительных неожиданных причин для этого - превращение электронных научных статей и материалов в «живые документы». См. ниже подраздел 2.9 о появлении у научных статей статуса «живых документов», что в сочетании с расширенными возможностями электронного цитирования имеет для научного сообщества ряд новых и важных последствий.

2.8 Необходимость нового типа связей "оценки профессионального влияния"

При размышлении, нет ли "подводных камней" в схеме цитирования, описанной в подразделе 2.7, возникает ощущение, что предложенная модель научного электронного цитирования может нарушить в "пользу" автора баланс сил в потенциальном конфликте интересов между учеными.

При наличии в научной среде потенциального конфликта интересов, возможности публичной качественной оценки чужих результатов, применяемые отдельным ученым по своему субъективному разумению, в обязательном порядке должны быть сбалансированы возможностями научного сообщества дать встречную публичную профессиональную оценку субъективных мнений отдельного ученого. Если научная истина и новое научное знание «рождаются в процессе борьбы мнений», то средства ЭБ по обеспечению научной деятельности должны давать равные возможности для всех участников этой «борьбы».

В статье [5] предлагается следующий набор качественных характеристик для субъективного оценивания учеными степени влияния отдельного научного результата/статьи на развитие соответствующих областей науки:

- очень интересный результат (particularly interesting);
- поворотный пункт для развития науки (landmarks);
- новаторский/революционный результат (groundbreaking).

На наш взгляд, этот список должен быть сбалансирован еще, как минимум, следующими характеристиками:

- результат, основанный на заблуждении;
- ненаучный подход к получению результата;
- результат с возможным опасным влиянием.

Процедуру формирования подобных оценок пользователями информационного пространства для выбранных ими статей и материалов предлагается организовать в виде создания нового типа связей. При создании таких связей ученый не только выбирает качественные характеристики (для одной статьи/результата можно одновременно указать несколько характеристик), отражающие его понимание степени влияния исходной статьи/результата на развитие науки, но и должен иметь возможность прокомментировать свое решение. Созданные таким образом профессиональные оценки, с одной стороны, связаны с профилем ученого, который их сделал (можно будет увидеть сводный список: какие статьи и как оценил соответствующий ученый), а с другой – с научной статьей/материалом, для которой сделана оценка (для каждой статьи можно будет увидеть какие профессиональные оценки она имеет и кому они принадлежат).

2.9 Мониторинг и поддержка сетей связей

Если ученый (или его представитель) депонирует в ЭБ свою статью, то у него остается возможность постоянно (т.е. на протяжении всей своей профессиональной жизни) редактировать и изменять текст этой статьи. Научные электронные статьи по определению превращаются в "живые" документы (в одном европейском проекте их назвали "текущими публикациями" - liquid publication [5]).

Выгоды и потенциальные проблемы от этого очевидны:

Выгода - научный результат, над которым долгое время работает ученый, локализован в одном и том же эволюционирующем информационном объекте. Его будет легче найти, а его цитирование гарантирует ссылку на текущую обновленную версию научного результата. Сохраняя старые версии статей, можно проследить историю развития научной идеи. И т.д. и т.п.

Потенциальные проблемы - очередное редактирование "живого" документа может нарушить цитаты из этой статьи, уже сделанные учеными. С учетом возможно развитой сети связей цитирования между "живыми" документами, нарушение связей цитирования с одним документом, может поставить под сомнение содержание и множества других документов в данной сети цитирования.

Решение - онлайн-овая научная инфраструктура должна иметь сервис мониторинга всех связей, существующих между объектами информационного пространства, включая связи цитирования. Подобный сервис, например, будет уведомлять:

- авторов исходной статьи - о том, что и кем процитировано из его статей, а также какие связи цитирования автор нарушает, когда вносит в свою статью изменения;
- авторов статей, цитирующих другие статьи - о том, что цитируемая статья была изменена и сделанные цитаты требуют проверки;
- читателей - о наличии или отсутствии обновления цитат в читаемой статье, если цитируемые статьи изменялись уже после "выхода" читаемой статьи.

2.10 Мониторинг развития информационного пространства и активности действующих лиц

Другое важное направление использования "системного эффекта" от интеграции метаданных локальных ЭБ в центральной базе связано с разработкой инфраструктурных сервисов для автоматического мониторинга изменения структуры и расширения информационного пространства, а также отслеживания параметров информационной активности ученых в рамках этого пространства. В первую очередь это – сбор качественной и количественной статистики о результативности работы ученых и исследовательских организаций (учет всех видов связей, отражающих

использование и профессиональное влияние их результатов исследований, подсчет количественных показателей и т.п.), а также сбор статистики о востребованности результатов исследований в открытом доступе.

Процедуры сбора статистики могут работать в непрерывном режиме, а формируемая ими статистическая база может наращиваться, например, ежедневными порция данных. Это означает формирование и обновление статистического портрета отслеживаемых процессов практически в реальном времени.

Подобный мониторинг открывает возможность автоматического формирования открытой онлайн-научомерической статистики. На этой основе можно организовать автоматизированное построение публичных индикаторов результативности ученых и научных организаций. Сконструировать новую профессиональную сигнальную систему, улучшающую ученым ориентировку в текущих тенденциях и направлениях развития науки, а также упрощающую определения ими своего места в научном сообществе.

Если исследователи массовым образом используют описанные выше элементы виртуальной научной среды, то порождаемая в результате онлайн-статистика позволяет формировать статистический портрет ученого как набор, например, следующих данных:

- Персональные сведения об ученом, и история их изменений;

- Динамика роста количественных показателей активности ученого, в т.ч. числа статей, результатов исследований, материалов других типов и т.п.;

- Динамика количеств и структуры цитирования результатов данного ученого, а также цитирований чужих результатов, сделанные данным ученым;

- Параметры распределения качественных характеристик для цитирований, выполненных данным ученым, а также для результатов данного ученого, выполненных другими учеными;

- Динамика количеств, структуры и распределение качественных характеристик оценок профессионального влияния работ данного ученого, сделанных научным сообществом, а также чужих работ, сделанных данным ученым;

- Распределение цитирования и оценок влияния между различными результатами данного ученого.

Аналогичный статистический портрет может быть составлен и для исследовательских организаций на основе агрегирования статистических портретов ученых, которые работают в соответствующих организациях.

Онлайн-наукометрия безусловно является научным информационным ресурсом, и поэтому также должна быть в составе информационного пространства. Рассчитываемые на ее основе показатели использования, влияния и востребованности научных статей могут служить для обогащения/развития метаданных

соответствующих информационных объектов. Технически это может решаться созданием дополнительного вида связей между метаданными определенного объекта и сформированным для него статистическим портретом. Примером подобного решения в Соционет являются ссылки "График суммарной статистики просмотров", которые автоматически включаются в веб-страницы, визуализирующие метаданные научных материалов. Подробнее эта тема обсуждается в публикациях [6, 7].

При определенных условиях собираемая таким образом онлайн-статистика будет содержать репрезентативные статистические портреты всех действующих лиц (ученых, проектов, организаций и т.п.) научного сообщества, а также фиксировать все значимые для научного сообщества процессы и их результаты. Фактически, это означает создание, и обновление практически в реальном времени своего рода информационной проекции или модели научного сообщества, объединенного соответствующим информационным пространством.

Вся собираемая онлайн-статистика является открытой и, следовательно, создаваемая на ее основе информационная модель сообщества также является открытой для всех членов научного сообщества.

С учетом того, что в реальной жизни научное сообщество представляет собой систему в существенной мере скрытую для прямого наблюдения и изучения, то появление информационной модели, которая при определенных условиях может быть достаточно точной, может оказаться ценным приобретением.

Представляется перспективным использование подобной информационной модели в сервисах, создаваемых разработчиками ЭБ для своих пользователей, т.к. это означало бы определенный учет "обратных связей". Такие адаптивные сервисы могли бы предоставлять пользователям услуги, которые учитывали бы текущее состояние научного сообщества, реакцию ученых на определенные события и другие параметры информационной модели сообщества.

Выводы к разделу: Современный контекст функционирования и развития ЭБ характеризуется стремлением к объединению содержания локальных ЭБ. Это проявляется в виде создания информационных хабов и профессиональных информационных пространств. Разработка "системного эффекта" от объединения метаданных в одной базе идет по двум направлениям: а) создание возможностей для улучшения содержательной связанности родственных информационных объектов (например, автора со всеми своими статьями, исходно размещенных в различных ЭБ, развитие связей цитирования, визуализация связей профессионального влияния, поддержка актуальности связей при превращении статей в "живые" документы и т.д.); б) комплексный

мониторинг информационной активности ученых, формирование на этой основе онлайн-оценки научометрии, создание индикаторов научной результативности ученых и исследовательских организаций, создание комплекса статистических индикаторов, выполняющих роль профессиональной сигнальной системы для научного сообщества.

3 Электронные библиотеки и открытый доступ к результатам исследований

Параллельно и отчасти независимо от потока программно-технических инноваций, являющихся развитием парадигмы ЭБ, в настоящее время в международном научном сообществе сформировалось мощное общественное движение за открытый доступ к результатам исследований [8].

Инициатива открытого доступа призывает все исследовательские организации создать открытые электронные репозитории, библиотеки или архивы для размещения в публичном доступе всех законченных результатов открытых исследований, проводимых в соответствующей организации. Именно откликом на эту инициативу объясняется рост количества ЭБ, имеющих "технический вход" для выгрузки метаданных внешними программными роботами (см. выше подраздел "Интеграция локальных ЭБ"). Масштабы и динамику этих процессов в реальном времени иллюстрирует "Реестр репозитория открытого доступа" (ROAR, <http://roar.eprints.org/index.php?action=browse>).

Для создания мотивации у ученых к размещению своих результатов исследований в открытом доступе данная инициатива рекомендует организациям, финансирующим исследования, обязывать исследователей, которым они платят деньги, оперативно выкладывать в открытые институтские репозитории результаты соответствующих исследований. Отслеживание количества научных организаций, включая научные фонды, уже обязавших своих ученых депонировать все законченные результаты открытых исследований ведет еще один международный регистр Registry of Open Access Repository Material Archiving Policies (ROARMAP, <http://www.eprints.org/openaccess/policysignup/>).

Еще один важный аспект инициативы открытого доступа - увеличение количества научных коммерческих журналов, официально объявивших о согласии с размещением авторами в открытом доступе своих статей, которые были переданы на рассмотрение или уже опубликованы в таких журналах. Отслеживание ситуации в этой области ведется проектом ROMEO (<http://romeo.eprints.org/publishers.html>). Согласно этому источнику на апрель 2009 г. 97% зарегистрированных научных издательств объявили о согласии с этим положением (как правило, с определенными оговорками).

Идеи открытого доступа получили государственную поддержку в ряде стран в виде появления национальных программ по созданию электронных репозиториях открытого доступа, развитию открытых архивов, созданию на их основе научного информационного пространства, сбора онлайн-научометрической статистики и ее использования при определении уровня финансирования науки.

Так, в Великобритании запущен специальный проект в поддержку репозиториях (Repositories Support Project, <http://www.rsp.ac.uk/>). Более важно, что в этой стране действует государственная программа Research Assessment Exercise (RAE, <http://www.rae.ac.uk/>), в которой в числе прочих показателей результативности ученых и исследовательских организаций используются элементы онлайн-оценки (индексы просмотров и скачиваний статей в Интернете и т.д.).

В Австралии действует государственная программа поддержки университетов в создании репозиториях открытого доступа Australian Scheme for Higher Education Repositories (ASHER), по которой выделяется 25.5 миллиона долларов на три года для поддержки создания и обновления цифровых репозиториях.

Выводы к разделу: Исходно идеи открытого доступа отталкивались от достижений в развитии ЭБ, поскольку ЭБ являются основным инструментом для обеспечения доступа к электронным публикациям. В настоящее время ситуация, на наш взгляд, несколько поменялась: инициативы открытого доступа к результатам исследований являются стимулом к дальнейшему развитию ЭБ и создают необходимые организационные предпосылки и мотивации, чтобы использование ЭБ и основанной на них новой виртуальной среды для исследований стало общедоступным и повсеместным.

4 Открытая Наука как следствие развития электронных библиотек

Повсеместное использование исследователями виртуальной научной среды, одним из основных элементов которой являются ЭБ, формирует необходимые условия для внедрения в практику научной деятельности новой модели ее организации, которую мы назвали "Открытая Наука". Для того чтобы оценить преимущества Открытой Науки перед существующим в настоящее время способом организации деятельности научного сообщества, рассмотрим ее основные положения.

Система организации научной деятельности может быть названа Открытой наукой, если она удовлетворяет следующим принципам:

1. Реализован повсеместный свободный доступ к результатам исследований через электронное депонирование учеными (или их представителями) всех результатов открытых исследований в ЭБ. Модель организации ЭБ

приближается к CRIS, что обеспечивает надлежащее информационное представление в ЭБ исследователей, организаций и других информационных объектов (см. подраздел 2.1).

2. Ученые используют электронную форму фиксации результатов своих исследований и научного приоритета, специально сконструированную в составе ЭБ для их максимально широкого и полного использования (см. подразделы 2.2 – 2.6).

3. Использование результатов исследований в виде их электронного цитирования или оценки их профессионального влияния выполняется по усовершенствованной схеме, включающей спецификацию качественных характеристик использования/влияния (см. подразделы 2.7 и 2.8).

4. Действует онлайн-научная инфраструктура, которая собирает содержание ЭБ, научных архивов и репозиториях отдельных организаций в единое научное информационное пространство (см. подразделы 2.2 – 2.5).

5. Действуют нормы по формированию в научном информационном пространстве репрезентативных информационных образов исследователей и научных организаций, а также правила по их поддержанию в актуальном виде и формированию на их основе профессиональных социальных сетей (подраздел 2.6). Действует система контроля правильности установленных связей с учетом превращения научных материалов в "живые" документы (подраздел 2.9).

6. В составе онлайн-научной инфраструктуры созданы средства электронного автоматизированного мониторинга за изменением качественных и количественных параметров научных архивов и информационных образов действующих лиц, обеспечено формирование на этой основе открытой наукометрической базы данных с ежедневным обновлением (подраздел 2.10).

7. Реализован автоматизированный расчет на основе наукометрической статистики индикаторов и показателей функционирования науки, включая комплекс показателей результативности работы ученых и научных организаций, размещение этих ежедневно обновляемых показателей в открытом доступе (подраздел 2.10).

8. Обеспечено использование этих показателей в принятии решений о финансировании научной деятельности, включая определение персональных надбавок ученым.

Из данного набора требований, если представить, что он полностью реализован, логически вытекают следующие преимущества Открытой Науки по сравнению с действующей системой научной деятельности:

- полный корпус современных результатов публичной науки в открытом онлайн-доступе, т.к. все результаты исследований, для которых нет ограничений доступа, своевременно помещаются в

ЭБ, открытые архивы и репозитории исследовательских организаций;

- лучшие условия для доступа к научному знанию, и использования научных результатов, как следствие - ускорение научного кругооборота и повышение эффективности научной деятельности;

- лучшие условия для мониторинга результативности научной деятельности для отдельных ученых и исследовательских организаций;

- возможность создания прозрачной и проверяемой системы публичных индикаторов результативности ученых и исследовательских организаций, а также выявление и визуализация тенденций развития науки;

- работающие мотивации для ученых выкладывать все свои результаты в ЭБ, т.к. их персональные надбавки за научную деятельность зависят от соответствующих показателей.

Выводы к разделу: Различные направления развития, существующие в настоящее время в научном сообществе вокруг ЭБ, имеют заманчивую перспективу сформировать новую высокоэффективную модель научной деятельности под названием Открытая Наука.

5 Заключение

Описанные выше направления развития ЭБ способствуют, на наш взгляд, формированию некоторых важных характеристик научного сообщества:

а) открытость его информационного пространства для надлежащего представления как действующих лиц науки, так и результатов их исследований, или деятельности в общем случае;

б) открытость онлайн-научной инфраструктуры для создания разработчиками новых информационных продуктов и сервисов в интересах данного научного сообщества на основе содержания информационного пространства;

в) открытость информационной модели сообщества, которая позволяет всем желающим получить картину текущего состояния и происходящих в сообществе процессов, а также может быть использована для принятия более обоснованных решений всеми действующими лицами науки.

С точки зрения отдельного исследователя эти три перечисленные выше характеристики означают не только лучшие условия для его профессиональных взаимодействий с научным сообществом, но и новый уровень "прозрачности" или публичности результатов его профессиональной деятельности. С одной стороны, Открытая Наука дает ученому эффективную саморазвивающуюся среду для его профессиональной деятельности, ориентированную на максимальное использование полученных им результатов. С другой, деятельность ученого, полученные результаты, степень их использования

и влияния на науку, место ученого в науке и отношение к нему научного сообщества - все эти характеристики профессиональной деятельности ученого являются публичными в Открытой Науке.

Парадигма электронных научных библиотек переживает в настоящее время важное обновление. Развитие как самих ЭБ, так и связанных с ними программно-технических и организационных систем, составляющих комплексную среду для научной деятельности, проходит этап качественных изменений. На горизонте – Открытая Наука, возможность которой прямо вытекает из тенденций развития ЭБ и которая обещает ученым большое количество изменений в привычном научном мире. Надеемся, что данная статья будет способствовать формированию в научном сообществе убеждения в том, что отмеченные изменения олицетворяют собой прогресс.

Литература

- [1] Current Research Information System (CRIS) – <http://www.eurocris.org/>
- [2] Kulagin M. V., Lopatenko A. S.. Current Research Information System and Digital Libraries. Needs for Integration. Сборник трудов конференции RCDL'2001, электронная версия сборника - <http://rcdl2001.krc.karelia.ru:8002/papers/contents.ru.shtml>, электронная версия статьи - http://rcdl2001.krc.karelia.ru/papers/papers/kulagin_lopatenko/kulagin_lopatenko_paper.rtf
- [3] The Open Archives Initiative Protocol for Metadata Harvesting, <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>
- [4] Паринов С.И.. Информационные хабы. Соционет: электронный депонент, 2006, <http://socionet.ru/publication.xml?h=repesc:rus:mqijxk:9>
- [5] Fabio Casati, Fausto Giunchiglia, Maurizio Marchese. Publish and perish: why the current publication and review model is killing research and wasting your money, ACM Ubiquity 8 (3), Feb 2007. http://www.acm.org/ubiquity/views/v8i03_fabio.html (в тексте цитируется более новая версия этой статьи: Liquid Publications: Scientific Publications meet the Web, Version 2.3, October 1, 2007, <http://liquidpub.org/attachment/wiki/WikiStart/LiquidPub%20paper-latest.pdf>)
- [6] Когаловский М.Р., Паринов С.И. Метрики онлайн-информационных пространств. // Экономика и математические методы, 2008, т. 44, №2, с. 108-120
- [7] Паринов С.И. e-Science - онлайн-будущее науки. // Информационные технологии, №9, 2007, приложение.
- [8] Stevan Harnad. The Implementation of the Berlin Declaration on Open Access. D-Lib Magazine, March 2005, V. 11 N. 3. Текущее состояние см.

на <http://www.eprints.org/openaccess/> и <http://www.sherpa.ac.uk/>

Digital libraries development is a way to Open Science

Sergey Parinov

Step by step permanent development of digital libraries (DL) concept is forming a particular virtual research environment. Successfulness of international initiative for open access to research is creating conditions to use DL and related research environment as common research practice. It creates some important challenges for DL developers. If, in the future, DL of all research organizations are federated as universal research data and information space (DIS), how can research community exploit the emerged systemic effect to produce maximal benefits? How should we construct, in form and function, a system for shaping and sharing research results from local DL so as to provide maximal usage? How should we organize a process of research results usage, and design necessary tools to provide maximally comprehensive and accurate statistics on the uptake, usage and impact of research results? How should the DL and the research environment accumulate and process statistics to generate new online metrics sufficient for research assessment of higher quality, sensitivity, breadth, accuracy, reliability, and validity than current metrics? Analysis of possible in such context DL development trends shows an ability to appear a new efficient organization form for research activities. We called it as Open Science.

* Автор благодарен руководству ЦЭМИ РАН за поддержку данных исследований, а также М.Р. Когаловскому за многочисленные комментарии и пожелания, которые позволили улучшить текст этой статьи. Частично данные исследования поддержаны грантом РФФИ № 09-07-00378.