

Роль библиотек в поддержке обмена научными данными

Доклад на заседании «Роль библиотек в сборе, обработке, обеспечении сохранности научных данных и обслуживании ими», организованном Секцией научно-технических библиотек в ходе 78-й Генеральной конференции ИФЛА (9–16 авг. 2012 г., Хельсинки, Финляндия).

Публикуется с разрешения автора и одобрения аппарата ИФЛА.

Ключевые слова: библиотеки, научные данные, информационное обслуживание, сбор, обработка, сохранность, архивирование, обмен, поиск, навигация, цитирование.

Большие возможности

В октябре 2010 г. группа высококвалифицированных специалистов подготовила доклад «Оседлав волну. Каким образом Европа может получить пользу от растущей волны научных данных» (The High Level Expert group on Scientific Data (2010), *Riding the Wave*, <http://www.cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>), где отмечено: растущая волна получаемых научных данных меняет природу научного исследования, позволяя учёным, которые располагают самыми разными базовыми возможностями, работать с одними и теми же массивами научных данных, что потенциально помогает решению больших общественных проблем.

Пирамида публикации данных (рис. 1) иллюстрирует растущую проблему. На рисунке показаны способы, которыми можно обеспечить доступность научных данных для их совместного или повторного использования. В основании пирамиды – то, что составляет основную часть научных данных, а именно те, которые хранятся в исходном виде на жёстких или оптических дисках, либо в шкафах лабораторий. Целый ряд причин – этических, технических, культурологических, а также проблемы безопасности – приводит к тому, что эти данные не могут использоваться совместно.

Второй уровень пирамиды – данные, которые уже хранятся в депозитариях. Они доступны для пользования, в том числе повторного. На третьем уровне – компьютерные данные, входящие в состав файлов – приложений к опубликованным статьям. На верхнем уровне – публикации, в которых могут сообщаться научные данные.

В ходе проекта «Структурный анализ» (Survey Report PARSE.Insight : http://www.parse-insight.eu/downloads/PARSE-Insight_D3-4_SurveyReport_final_hq.pdf) выявлено, что около 60% учёных хотели бы использовать научные данные, полученные другими исследователями. Интересно, что из этого же обзора следует: у более чем 40% учёных возникают реальные проблемы при попытках поделиться собственными данными.

Для того чтобы снизить нижний уровень пирамиды, увеличить объёмы совместно и повторно используемых данных, нужно рассмотреть факторы, которые затрудняют эти процессы.



Рис. 1. Пирамида данных

Некоторые позиции по совместному пользованию данными

В докладе по интеграции данных и публикаций (*Report on Integration of Data & Publications*) рассмотрены ключевые мнения учёных, издателей и архивистов относительно совместного пользования данными. Выявились возможности вовлечения каждой из отмеченных групп в этот процесс и особенно – отличные шансы для участия библиотекарей.

Учёные указали ключевые проблемы, или зоны повышенного внимания, касающиеся возможностей, которыми библиотеки уже обладают:

1) библиотеки – это место, где можно разместить и хранить данные для дальнейшего использования; 2) возможность управления обслуживанием и доступом; 3) способность интегрировать данные и публикации; 4) понимание важности данных, как первоклассного научного объекта.

Ключевая зона внимания издателей (где библиотеки могут также оказаться полезными) – представление данных в более усовершенствованных форматах с тем, чтобы облегчить их повторное использование, для чего очень важна генерация хороших метаданных.

На основе ключевых моментов, важных для каждой из участвовавших групп, можно выделить семь потенциальных возможностей, при реализации которых должно проявиться преимущество библиотек:

1. Доступность – имеется в виду помощь учёным в том, чтобы сделать их данные доступными, а также предоставить поисковые сервисы для данных.
2. Навигационная «заметность», обеспечиваемая наиболее современными способами работы с данными.

3. Возможность интерпретации посредством предоставления метаописаний и оказания помощи в подготовке к работе с ними.
4. Возможность повторного использования – с помощью консультаций по доступности тематических архивов и лицензирования для повторного пользования.
5. Возможность цитирования – распространение наиболее разумных и рациональных методов цитирования, использование постоянных идентификаторов документов.
6. Работа с данными – обработка и предоставление услуг по выработке соответствующих навыков.
7. Обеспечение сохранности – распространение передового опыта по архивированию данных.

Диапазон современных и будущих задач

Из наших исследований ясно, что у библиотек есть весьма реальные возможности для того, чтобы играть важную роль в обмене данными. Следующий шаг – придать импульс для реализации этих возможностей, или, иными словами, удовлетворения потребностей учёных в информационной поддержке со стороны библиотек.

Чтобы выявить основные направления такой поддержки, в 430 библиотек – участниц Европейской сети научных библиотек (LIBER) – были направлены анкеты. В них содержались вопросы, касающиеся востребованности библиотечной поддержки в конкретных областях. Кроме того, респонденты должны были оценить и свою готовность выполнять требования учёных, а также наличие необходимых навыков для решения поставленных задач. И, наконец, был вопрос о том, каким образом можно организовать переподготовку и обучение библиотекарей соответствующим технологиям.

Анкеты были направлены и в несколько библиотек США и Австралии, являющихся признанными лидерами в работе с научными данными. Таким образом, можно было провести интересное сравнение, особенно ответов на вопрос о том, какие новые умения необходимы для работы в этой области.

Обзор показал, что научные библиотеки ощущают очень высокую востребованность со стороны учёных в организации работы с данными. Более 80% респондентов нужна поддержка в работе с данными. Когда мы углубились в суть проблемы и сравнили потребность с уровнем оказания поддержки, оказалось, что такая поддержка действительно не высока. Это не удивительно, если учесть, что технология совместного пользования данными всё ещё находится на относительно ранней стадии в большинстве научных дисциплин.

Наибольшая потребность ощущается в архивировании данных, но реализуется примерно половина всех потребностей, а наибольший разрыв между потребностью и реальной помощью отмечен при формировании планов работы с данными, поскольку только 19% библиотек оказывают в этом помощь (рис. 2).

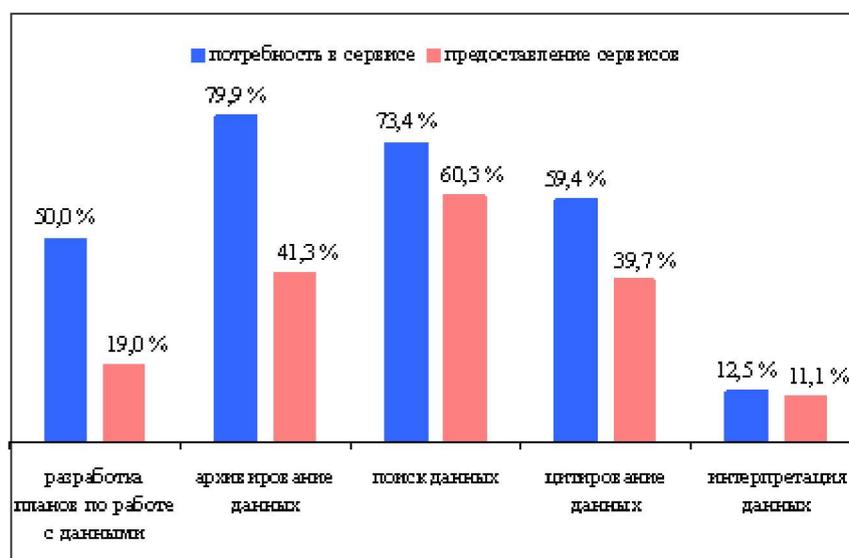


Рис. 2. Результаты опроса европейских научных библиотек о наличии спроса на сервис и его предоставлении (доля библиотек, ответивших на вопрос положительно)

Только 12% библиотек считают, что они обладают достаточным умением, чтобы удовлетворять потребности пользователей. Обнадёживает, что 56% библиотек сейчас работают над формированием тех навыков, которые, по их мнению, необходимы для оказания поддержки в работе с данными.

Повышение квалификации

Итак, какими дополнительными умениями и навыками должен обладать библиотекарь, чтобы обеспечить потребность учёных в работе с данными? Ответы респондентов выстроились в следующем порядке:

1. Умение работать с информационными технологиями.
2. Умение обрабатывать данные и архивировать их.
3. Практический опыт исследований по данной тематике.
4. Умение оценить научные данные.
5. Умение провести визуализацию массивов данных.

В ответах примечательно то, что экспертные сообщества (т.е. те библиотеки, где уже накоплен достаточный практический опыт) отдают приоритет необходимости владеть навыками обработки данных и архивирования. Умению работать с информационными технологиями библиотечные специалисты отводят лишь четвёртое место. Означает ли это, что из практического опыта следует, что навыки работы с информационными технологиями вовсе не столь важны, как полагают в библиотеках, активно работающих с научными данными? Наличие опыта исследований по этой тематике также занимает важное место среди дополнительных умений, названных библиотеками, уже имеющими опыт работы с данными.

Наилучшее средство для выработки всех навыков, по мнению библиотек, – это непрерывное профессиональное развитие. Оно возможно во всех областях, поэтому следует интегрировать технологии обработки данных в состав курсов по профессиональной подготовке.

Заключение

Независимо от того, в каком порядке выстраиваются приоритеты необходимых умений и способов, с помощью которых они приобретаются, ясно, что потребуются значительные вложения в наработку этих умений.

Совместное пользование научными данными и технология электронной науки (*e-science*) меняют природу исследования, и при этом возникают новые благоприятные возможности для библиотек. Библиотекам следует как можно быстрее воспользоваться ими, но при этом – действовать достаточно осторожно в отношении инвестиций для выработки умений. Такие вложения должны соответствовать наиболее востребованным областям, а также тем, которые могут оказать наибольшее воздействие в условиях ограниченного финансирования.

Библиотекари заинтересованы в постоянном профессиональном развитии, и оно должно быть им доступно, чтобы строить новые знания на основе уже наработанных.

Список источников

(в авторском варианте)

ParseInsight Survey (2009) http://www.parse-nsight.eu/downloads/PARSE-Insight_D4-GapAnalysisFinalReport.pdf

Reilly et al. (2011) ODE report on the integration of data and publications:
<http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2011/10/ODEReportOnIntegrationOfDataAndPublications.Pdf>

Survey Report PARSE.Insight : http://www.parse-insight.eu/downloads/PARSE-Insight_D3-4_SurveyReport_final_hq.pdf

Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, et al. (2011) Data Sharing by Scientists: Practices and Perceptions. [PLoS ONE 6\(6\):e21101.doi:10.1371/journal.pone.0021101](https://doi.org/10.1371/journal.pone.0021101)

The High Level Expert group on Scientific Data (2010), Riding the Wave,
<http://www.cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>