

Информационные ресурсы образования

АНТОПОЛЬСКИЙ Александр Борисович - доктор технических наук, профессор МГУКИ, главный научный сотрудник Института научной информации и мониторинга (ИНИМ) РАО, член редколлегии журнала «Информационные ресурсы России»

тел.: +7-916-495-20-42;

e-mail: ale5695@mail.ru

ПРИНЦИПЫ СИСТЕМЫ НАУЧНОЙ ИНФОРМАЦИИ В РОССИЙСКОЙ АКАДЕМИИ ОБРАЗОВАНИЯ

Потребители научно-педагогической информации

Общее количество научно-педагогических институций в России достаточно велико и достигает трех сотен, включая около 40 государственных научных учреждений РАО и других ведомств, негосударственные научные учреждения, более 100 педагогических вузов и не менее 100 факультетов и кафедр педагогического и психологического профиля в классических университетах и других вузах.

Число студентов, обучающихся по специальностям «Педагогика» и «Психология», составило 148 тыс. из общего числа 1,4 млн. российских студентов (данные 2006 г.).

Число аспирантов по этим специальностям составило в том же году 12,3 тыс. из общего количества 140 тыс., с ежегодным выпуском 2,6 тыс. (из общего числа 30,8 тыс.). Число докторантов по тем же группам специальностей – 630 из общего числа 4,6 тыс. Таким образом, число студентов, аспирантов и докторантов, обучающимся по специальностям педагогика и психология составляет в среднем не менее 10 % от общего числа обучающихся в России по этим категориям.

На основе приведенных данных можно утверждать, что общее число ученых, занятых в этой сфере, то есть потенциальных пользователей системы научно-педагогической информации (НПИ), можно оценить приблизительно в 100 тыс. чел., что также составляет около 10 % от общего числа занятых в сфере науки в России.

Следует иметь в виду, что общее число занятых в сфере образования, то есть потенциальных потребителей образовательной информации, составляет около 7 млн. человек, при этом к числу потребителей образовательной информации следует добавить студентов (1,4 млн.), абитуриентов (до 1 млн. ежегодно), а также родителей абитуриентов. Из общего числа обучающихся в средних общеобразовательных учебных заведениях (16 млн. чел.), а также студентов средних учебных заведений (2,5 млн. чел.), значительная часть также является потенциальными пользователями информационных образовательных ресурсов

В системе НПИ отсутствуют специализированные организации потребителей и производителей информации. Участие в больших информационно-библиотечных консорциумах (НЭИКОН, АРБИКОН) для пользователей НПИ весьма ограничено и не слишком выгодно.

Общая характеристика информационных ресурсов НПИ

Так же как наука в сфере образования является частью сферы образования в целом, информационные ресурсы НПИ России являются частью гораздо более крупных по масштабу информационных ресурсов сферы образования. При этом выделение собственно НПИ среди образовательных ресурсов является трудной и не всегда выполнимой задачей. Большинство крупных владельцев НПИ такой дифференциации не проводит, существующая статистика также не позволяет выделить собственно НПИ. Поэтому данные, приводимые ниже, отчасти относятся не только к НПИ, но и ко всем образовательным ресурсам в целом.

Информационные ресурсы НПИ в настоящее время представляют значительное по объему, но плохо структурированное информационное пространство. Среди информационных систем и ресурсов НПИ наблюдается значительное дублирование.

Библиотечные фонды НПИ

Крупнейшим хранителем НПИ, как и образовательных ресурсов в целом, является сеть библиотек образовательных учреждений. К их числу относятся библиотеки вузов, которые имеются практически во всех вузах страны (около 3 тыс.) и в значительном количестве средних

специальных и общих учебных заведениях. Общее число библиотек сферы образования достигает 50 тыс. Методическое руководство сетью образовательных библиотек осуществляет НПБ им. К.Д. Ушинского РАО (для школьных библиотек и библиотек педагогических учебных заведений), библиотека МГУ (для библиотек классических университетов), ГСНМБ (для библиотек медицинских учебных заведений) и некоторые другие. Крупнейшим фондом в сфере НПИ является фонд НПБ объемом 1,6 млн. ед. хран. Данные о наличии НПИ в других библиотечных фондах отсутствуют.

Книжные ресурсы НПИ

Основным видом опубликованной НПИ являются книги. Общее число книг учебно-образовательного назначения, выходящих в России, по данным Российской книжной палаты составляет до 30 тыс. наименований в год (порядка 25 % от общего объема книгоиздания, в 2008 г. составившего около 120 тыс.).

При этом не всегда ясно, какая доля книг и периодики может быть бесспорно отнесена к научным изданиям.

Периодика в сфере «Педагогика» и «Психология»

В отчете Н.К. Ханнанова достаточно подробно проанализировано состояние журналов в данной области, наличие их в интернете в свободном или платном доступе, а также их наличие в Научной электронной библиотеке и в депозитарии электронных научных изданий НТЦ «Информрегистр».

Архивы

Крупнейшим и видимо единственным специализированным архивом НПИ является Научный архив РАО, переданный в состав НПБ. Его объем определяется в 100 тыс. дел, сгруппированных в 130 фондов. Примерно 30 % фондов составляет персональная информация, остальное – архивные фонды научно-образовательных учреждений, в основном советского периода, а также личные коллекции известных ученых педагогов. В составе Научного архива имеется также около 6 тыс. диссертаций по соответствующей тематике.

Имеются отрывочные сведения о наличии НПИ в других архивах, как федеральных (ГАРФ, РГИА), так и региональных.

Так, в составе ГАРФ имеется несколько десятков фондов, содержащих информацию о деятельности общегосударственных органов управления образованием с 1917 г. по настоящее время. В том числе несколько фондов содержат документы по организации народного образования в областях, временно находившихся под управлением белых армий и правительств.

Особый интерес представляет содержание фондов Российского государственного исторического архива (РГИА). Известно, что эти фонды содержат все материалы Министерства народного просвещения Российской империи с момента его создания в начале 19-го века до Февральской революции. В настоящее время идет активная оцифровка фондов РГИА в рамках участия в создании Президентской библиотеки им. Б.Н. Ельцина. Однако доля НПИ в оцифрованной части фондов РГИА неизвестна.

Практически все региональные архивы имеют в своем составе фонды по истории народного просвещения в соответствующем регионе. Можно предполагать, что эти архивы, по мере участия в создании Президентской библиотеки также будут оцифровывать свои фонды, включая документы, посвященные истории народного образования в России.

Музеи

Специализированных музеев, посвященных истории образования в последнее время создано достаточно много, вероятно, несколько десятков. Кроме того, практически все региональные и краеведческие музеи содержат материалы, связанные с историей образования в соответствующих регионах России. Множество учебных заведений создают собственные музеи, также содержащие материалы по истории образования. Музеи различного профиля создают экспозиции, коллекции и электронные ресурсы образовательного назначения. В рамках единой коллекции ЦОР был создан ряд коллекций образовательного назначения на основе собраний различных музеев.

Все это сформировало такую область педагогики, как «Музейная педагогика». По мере создания баз данных каталогов Государственного музейного фонда возможно будет при проведении навигационных поисков и создании распределенной системы вторичных ресурсов НПИ подключать и поиск музейных экспонатов по данной тематике.

Образовательные ресурсы российского интернета

Каталог Яндекса относит к категории образовательных св. 5 тыс. сайтов и порталов (из общего числа около 100 тыс., отраженных в каталоге).

Каталог Топ 100 Рамблера относит к этой категории 6,9 тыс. из около 200 тыс. сайтов.

Каталог электронных библиотек НП ЭЛБИ относит к рубрикам «Психология» и «Педагогика» св. 350 электронных коллекций (не считая коллекций универсальной тематики, в которых образовательные ресурсы часто составляют заметную долю).

Среди электронных библиотек имеется несколько достаточно крупных, полностью или частично содержащих учебную и образовательную литературу. Например, ЭБ «Ихтик» содержит более 30 тыс. учебников и рекомендованной литературы для высшей школы. Однако в настоящее время затруднительно сказать, какая часть этих коллекций, во-первых, удовлетворяет критериям научности, а, во-вторых, пользуется существенным спросом со стороны научно-педагогического сообщества. Это предстоит определить в ходе проектирования системы НПИ

Из вторичных ресурсов самым крупным является каталог образовательных ресурсов интернета в рамках Единого окна доступа к образовательным ресурсам. По состоянию на 01.04.2009 г. он включает 45 тыс. описаний и ссылок на ресурсы образовательного назначения различных типов – как на отдельные документы различных типов, так и на целые массивы, такие как электронные издания, электронные библиотеки, сайты целиком и прочее.

Многие важные информационные ресурсы НПИ являются платными. Однако соотношение затрат на обращение к ним и реального спроса на них с учетом платежеспособности потребителей должно быть предметом специального экономического анализа. Необходимо доказать экономическую целесообразность обращения к внешним ресурсам по сравнению с созданием собственных в рамках системы НПИ. Это касается, прежде всего, зарубежных ресурсов, но и отечественных также.

Крупнейшими владельцами НПИ, представленной в электронном виде, являются:

- в рамках образовательных структур – ГНИИ ИТТ «Информика» и НПБ им К.Д. Ушинского, а также ряд вузов, прежде всего педагогических;
- среди федеральных информационных систем основные ресурсы НПИ имеются в РГБ, ВНИЦентре, НЭБ.

Научные учреждения РАО, с некоторыми оговорками, не являются заметными производителями электронных ресурсов НПИ.

Инициатива открытых архивов в России

Инициатива открытых архивов однозначно признается специалистами как наиболее перспективная модель распространения научно-образовательной информации. Хотя в силу многих условий, прежде всего, консерватизма отечественных издателей, руководителей научных организаций, да и многих ученых, ее распространение происходит не такими быстрыми темпами, как распространение интернета, тем не менее, успехи за последние годы достаточно впечатляющие. Нет сомнения, что эта модель должна быть положена в основу системы научно-педагогической информации в России. Далее кратко характеризуется состояние ОАИ в России

На базе ЦЭМИ РАН действует межинститутская группа «Соционет» [2] (руководитель д.т.н. С.И. Паринов), оказывающая помощь институтам Отделения общественных наук РАН (ООН РАН) в создании и развитии их собственных электронных коллекций, для их включения в ЕНИП РАН, а также в RePEc и другие международные системы распространения научной информации. Рабочая группа проводит обучение делегированных институтами ООН РАН сотрудников. Назначенные люди получают навыки создания коллекций электронных материалов и публикаций, способы ввода/редактирования информации, правила доступа к информационным ресурсам и др.

Использование сервисов Соционет и ЕНИП открывает сотрудникам институтов ООН РАН, а также другим ученым, возможность использования новых средств получения и распространения научных материалов в общероссийской и международной профессиональной среде, включая самоархивирование, электронное депонирование и открытые архивы [3].

Актуальность внедрения ОАИ в России особенно очевидна в связи с финансовым положением российской науки. Бюджеты, выделяемые российскими исследовательскими организациями на подписку на ведущие мировые научные журналы, невелики и, как правило, фиксированы. При этом подписные цены на собственно российскую научную периодику одни из

самых высоких в мире. В прошлом году Library Journal, издание Reed Elsevier, провёл исследование, сопоставив средние цены на издания разных стран, имеющиеся в трёх основных базах данных ISI — Arts and Humanities Citation Index, Social Sciences Citation Index, Science Citation Index. Выяснилось, что российские журналы занимают в этом списке совсем не почётное второе по дороговизне место. Для сравнения: в США выходит в 40(!) раз больше журналов мирового уровня, при этом американский журнал в среднем вчетверо дешевле российского. Получается, что научные журналы — важнейшие источники актуальной информации — для российских учёных с каждым годом становятся, во-первых, всё менее доступны, и, во-вторых, если говорить о публикациях на родном языке, всё менее интересны. Это, конечно, совершенно неприемлемо.

Открытые архивы можно рассматривать не только как технологическую инновацию, позволяющую сократить издержки и ускорить диффузию знаний, но и как организационно-управленческую. Поскольку созданный на базе института архив представляет результаты исследовательской деятельности всей научной организации, по сути, отражает эффективность её функционирования, значит, он может быть использован и для расчёта персональных надбавок. Добавление в систему открытого архива наукометрического и рейтингового компонентов может дополнительно стимулировать авторов к депонированию результатов научной деятельности, считает С.И. Паринов. В 2006 году он инициировал программу «Открытый доступ к результатам исследований», предполагающую создание единой инфраструктуры для оперативного распространения и использования результатов научных исследований. Сегодня «Соционет» уже объединяет архивы 22 институтов ООИ РАН.

Институты создают свои внутренние архивы, взяв за основу открытую архитектуру «Соционета», принимают положения об обязательном электронном депонировании результатов исследований в институтском архиве.

Все сотрудники институтов отлично осведомлены о принципах и метриках учёта их производительности. Автоматизация расчёта надбавок с применением стандартных, унифицированных процедур стала большим плюсом уже сама по себе. Не менее важно, что такой процесс гарантирует беспристрастность и непредвзятость оценки труда учёных.

Принципы организации системы НПИ

Основным принципом функционирования системы НПИ является принцип открытого, свободного и бесплатного доступа к научно-педагогической информации для конечного пользователя. Доступ к коммерческим ресурсам НПИ, как правило, должен оплачиваться из государственного бюджета, общественных фондов или собственных средств научно-педагогических организаций.

Координация деятельности по созданию коллекций НПИ основывается на принципах минимизации общественных затрат на эту деятельность при согласованных параметрах роста коллекций.

Участники системы НПИ признают авторское право, и доступ к охраняемым произведениям будет предоставляться в соответствии с правилами, установленными 4-й частью Гражданского Кодекса РФ. При этом участники системы стремятся развивать альтернативные способы распространения научно-образовательной информации, в том числе в форме присоединения к Инициативе открытых архивов. Участники системы по возможности будут использовать правовую модель Common Creative License способами, не противоречащими российскому законодательству.

Участники системы могут, кроме открытых архивов и коллекций, создавать ресурсы ограниченного доступа, в том числе охраняемые авторским правом и доступные только в помещениях библиотек, а также охраняемые персональной тайной и доступные только с разрешения субъектов персональных данных.

Создание системы НПИ основывается на принципе конвергенции: участники системы должны развивать свои ресурсы и сервисы в направлении координации своих действий и созданию единого информационного пространства НПИ постепенно с минимальными изменениями своих системных решений.

Интероперабельность системных решений достигается путем принятия общих стандартов, как правило, дополняющих уже реализованные решения. Принципы интероперабельности

основываются на технологических решениях Инициативы открытых архивов и действующей в России системы «Соционет».

Участие в системе НПИ является обязательным для учреждений РАО и добровольным для других научно-педагогических институций.

В составе системы НПИ выделяются три головные организации: НПБ им. К.Д. Ушинского, ИНИМ РАО и ГНИИ ИТТ «Информика». Все системные решения принимаются на основе консенсуса этих организаций с учетом мнений остальных участников.

Функциональная структура НПИ

НПИ включает следующие два взаимосвязанных блока:

- 1) информационные системы участников системы НПИ, где они ведут свою электронную информационную деятельность, в том числе, в форме институтских открытых архивов;
- 2) сервисы онлайн-научной инфраструктуры, которые интегрируют локальные ОА и создают из них общее информационное пространство.

В составе НПИ могут создаваться функциональные блоки, которые строятся с учетом опыта создания системы «Соционет»:

- информационный хаб, функционирующий с использованием XML, RSS, и протокола OAI-PMH;
- научное информационное пространство, включающее Открытые архивы, коллекции и документы, в том числе препринты, электронные периодические издания и др.;
- онлайн-рабочие места ученых различных типов;
- персональный информационный робот;
- инструментарий для формирования профессиональной социальной сети;
- инструментарий для расчета онлайн-научомерических показателей научно-педагогических институций и отдельных ученых

Общая информационная структура НПИ

Как показано выше, ресурсы НПИ представляют собой конгломерат коллекций документов и баз данных на различных носителях, принадлежащих разным собственникам, использующих различные программно-технологические средства и метаданные, имеющие разные условия доступа (свободные, ограниченные, коммерческие). Значительная доля НПИ отсутствует в электронном виде и должна быть оцифрована или собрана в электронном виде от производителя.

Далее излагаются общие принципы конвергентного развития нынешнего состояния в единое структурированное пространство, использующее единую онлайн-инфраструктуру, применяющую технологические решения Инициативы открытых архивов и отечественной системы «Соционет».

Участники системы проводят инвентаризацию имеющихся ресурсов, а также ближайших планов развития этих ресурсов и на этой основе согласованно утверждают:

- 1) список коллекций, приоритетных для системы НПИ, которые участники системы обязуются вести собственными силами;
- 2) список внешних ресурсов, к которым необходим доступ многих участников системы.

Каждый участник преобразует имеющиеся информационные ресурсы или создает новые в виде открытого архива, состоящего в общем случае из множества коллекций. Коллекции представляют собой множество документов, объединенных по какому-либо принципу (места создания, авторства, вида документов и др), снабженных стандартным набором метаданными или базы данных. Каждая коллекция и каждый архив также имеют стандартные метаданные.

Преобразование коллекции в общем случае заключается в разработке конвертора для метаданных данной коллекции в структуру, соответствующую требованиям протокола OAI-PMH.

Новые архивы и коллекции сразу создаются с учетом требований данного протокола и других условий онлайн-инфраструктуры.

Таким образом, создается трехуровневая система информационных ресурсов НПИ: архив, коллекция (база данных), документ.

Метаданные для ресурсов любого уровня предназначены: 1) для поиска ресурса в общем информационном пространстве НИИ; 2) для администрирования этих ресурсов, включая права доступа к нему.

Каждому участнику бесплатно предоставляется программное обеспечение для создания архива организации, коллекции, организации рабочего места формирования метаданных и других сервисов, имеющихся в системе «Соционет».

Организация доступа к внешним источникам

Для доступа к внешним ресурсам НИИ, предоставляемым на коммерческих условиях, создаются консорциумы, минимизирующие затраты для участников системы НИИ. Консорциумы, как правило, функционируют на основе паевого финансирования доступа с привлечением как бюджетных, так и внебюджетных источников. Конечной целью является бесплатный, но контролируемый доступ для ученых-пользователей системы НИИ к внешним источникам. Один из участников системы НИИ берет на себя функции координатора доступа к соответствующему ресурсу и организации консорциума. Возможно также привлечение НЭИКОН.

Первоочередными отечественными внешними источниками для системы НИИ являются:

- Электронная библиотека диссертаций РГБ;
- Научная электронная библиотека;
- База данных НИОКР ВНИИЦентра;
- Реестр научных организаций ВНИИЦентра.

Перечень зарубежных источников, необходимых для системы НИИ, должен быть определен дополнительно путем анализа информационных потребностей и финансовых возможностей участников системы НИИ

Типы ресурсов (коллекций и данных) системы НИИ

В настоящее время система «Соционет» предлагает возможности создания следующих типов коллекций (этот перечень может быть принят как базовый для системы НИИ, а при желании может быть легко расширен):

- Коллекции **персональных профилей**, включающих данные о специалистах - тип данных "**person**";
- Коллекции **организационных профилей**, включающих данные о научно-педагогических институтах, - тип данных "**institution**";
- Коллекции **результатов научно-технической деятельности (РНТД)** - тип «**result**»;
- Коллекции, содержащие **описание инноваций** - тип "**innovation**";
- Коллекции с **описанием инновационных проектов** - тип "**project**";
- Коллекции для **описания предлагаемых услуг** - тип "**product**";
- Коллекции для **описания инвестиционных запросов или предложений** - тип "**investment**";
- Коллекции, содержащие "**рабочие документы**", неопубликованные материалы, препринты и другие нерцензированные научные публикации – тип "**paper**";
- Коллекции, содержащие **описание публикаций в обычных или онлайн-журналах** - тип "**article**";
- Коллекции, являющиеся **подборками описаний изданных книг, каталогами новых поступлений библиотек, планов издательств и т.п.** - тип "**book**";
- Коллекции **отдельных глав или разделов книг** - тип "**chapter**";
- В коллекциях "**тематические подборки публикаций**" хранятся списки публикаций, посвященных определенной теме. Вторичные коллекции – тип "**secondary**";
- **Новости** (например, объявления о событиях, а также любая информация, имеющая краткосрочный характер) - тип данных "**news**";
- Научные **рецензии, комментарии, обзоры**, которые обязательно посвящены рассмотрению/анализу одной или нескольких исходных публикаций, - тип данных "**comment**";
- Коллекции **диссертаций и/или их авторефератов**, включающих данные о всех видах защищенных диссертаций - тип данных "**thesis**";

- Тематические **классификаторы и рубрикаторы**, созданные как самостоятельные коллекции - тип данных "**scheme**";
- Коллекции, содержащие описание новых **типов данных** (дополнительных к выше приведенным в этом списке), - тип "**dtype**";
- Коллекции для описания новых **дисциплин и предметных областей** (дополнительно к уже включенным в систему) должны иметь тип данных "**discip**".

Распределение функций основных участников системы

В настоящем разделе излагаются принципы, по которым можно было бы распределить усилия по созданию открытых архивов и коллекций первичных ресурсов в системе НПИ, по крайней мере, среди основных участников. Эти принципы, естественно, будут скорректированы в процессе создания системы

НПБ сосредоточивает свои усилия на создании электронных коллекций на основе имеющихся у нее традиционных фондов и получения обязательного электронного экземпляра от институтов РАО:

- Электронная библиотека авторефератов диссертаций по педагогике и психологии;
- Полнотекстовая электронная библиотека классиков педагогики и психологии (в основном неохранные произведения);
- Электронная библиотека изданий РАО, передаваемых в НПБ согласно Положению об обязательном электронном экземпляре документов;
- Электронная коллекция документов на основе Научного архива РАО;
- Электронная коллекция архива педагогической периодики на основе собственных электронных журналов и архива электронной доставки документов (дополняющая массивы НЭБ);
- Организация доступа к электронной библиотеке диссертаций;
- Создание и обновление реестра педагогических полиграфических изданий организация доступа к периодическим изданиям, включая доступ к коллекции НЭБ, БЕН и т. д.;
- Создание и обновление каталога, научно-методическая экспертиза и создание коллекции электронных мультимедиа изданий образовательного назначения;
- Создание и обновление каталог электронных периодических изданий по педагогике и психологии.

ИНИМ (совместно с порталом РАО) сосредоточивает свои усилия на создании коллекций, используемых для оценки результатов научно-технической деятельности (РНТД) в сфере педагогики и психологии:

- База данных неопубликованных отчетов, переданных в ИНИМ в соответствии с Положением об обязательном электронном экземпляре документов;
- База данных ИК по НИОКР в области педагогики и психологии, приобретенная во ВНИИЦентре;
- База данных планово-финансовой информации по организации научной деятельности в РАО;
- База данных профилей научно-педагогических институций (реестр научных организаций в сфере педагогики, психологии и смежных с ними наук);
- База данных персональной информации специалистов в области педагогики и психологии (ограниченного доступа);
- Базы данных статистики НПИ и другой наукометрической информации, формирующейся на основе обработки данных Открытых архивов институтов.

ГНИИ ИТТ «Информика» сосредоточивается на ведении массивов и коллекций образовательных ресурсов, на основе имеющихся в составе системы образовательных порталов:

- Новости;
- Коллекция нормативных и методических документов, заданий, справочной информации для абитуриентов;
- Реестр образовательных учреждений;
- Коллекция картографических ресурсов в сфере образования;
- Глоссарий терминов и дефиниций в образовательной сфере;
- Коллекция правовых документов;

- Государственные образовательные стандарты;
- Сведения о мероприятиях (конференции, семинары и проч.);
- Коллекция документов по статистике российского образования;
- Сведения о фондах, финансирующих научно-образовательную деятельность;
- Сведения о конкурсах в сфере образования;
- Документы и данные по дистанционному обучению;
- Полнотекстовая электронная библиотека учебных и учебно-методических материалов;
- Единая коллекция цифровых образовательных ресурсов;
- Ресурсы для открытой мультимедиа среды;
- Всероссийская олимпиада школьников;
- Всероссийский интернет-педсовет;
- Федеральный перечень учебников для среднего образования.

Вероятно, ГНИИ ИТТ «Информика» в рамках системы НПИ сможет прекратить ведение некоторых из своих коллекций и использовать освободившиеся ресурсы более рационально.

Вторичные ресурсы системы НПИ

Самостоятельной проблемой является организация системы вторичных ресурсов для всего пространства НПИ. Один из вариантов – это создание единого распределенного каталога всех документов системы НПИ, который должен объединять вторичные ресурсы из всех существующих и вновь создаваемых источников системы НПИ (включая каталоги образовательных порталов, каталоги НПБ и Научного архива РАО, каталоги БД внешних источников и др.). Сводный каталог системы НПИ должен быть открытым и общедоступным. Его создание и функционирование на основе технологии ОАИ требует дополнительных исследований. В частности, пока неясно, должен ли он распространяться только на документы, доступные в коллекциях ОАИ или иметь более широкий характер. Также неясно, возможно ли с разумными затратами решить проблему устранения дублирования в записях сводного каталога.

В любом случае очевидно, что нынешняя система распределенной каталогизации НПИ крайне неэффективна, и участники системы в значительной степени дублируют функции друг друга.

Однако, кроме сводного каталога, могут быть весьма полезны и востребованы вторичные ресурсы (каталоги, библиографические указатели, справочно-поисковые указатели), предназначенные для поиска в отдельных коллекциях, особенно специфических по виду источника или системе хранения. Далеко не всегда целесообразно централизовать эти ресурсы в рамках единого сводного каталога.

Например, самостоятельными задачами является создание поискового аппарата для Научного архива РАО, библиографических указателей по педагогической периодике, библиографических БД по диссертациям и авторефератам диссертаций.

Одним из основных факторов, влияющих на целесообразность унификации и централизации работ по созданию вторичных ресурсов, является разнообразие средств метаданных и лингвистического обеспечения в функционирующих в настоящее время ресурсах системы НПИ. Нужно также учитывать привычку к этим средствам значительных категорий пользователей. Подробнее об этом сказано ниже.

Другим фактором, не менее существенным, является структура информационных потребностей пользователей НПИ, их дифференциация относительно видов информационных источников и систем доступа к информации.

Оценка результатов научно-технической деятельности (РНТД)

В целях повышения точности измерения научного вклада и влияния ученого на развитие науки нужно дать ему возможность излагать научный результат в максимально "сухом" и формализованном виде, как, например, заполнение специально сконструированной онлайн-формы (возможно, похожей на форму 1 для описания РНТД из приложения к постановлению Правительства РФ от 04 мая 2005 г. № 284 «О государственном учёте результатов научно-исследовательских, опытно-конструкторских и технологических работ гражданского назначения», уже реализованной в «Соционете» как тип данных «result»).

Для отдельно взятого РНТД (в виде статьи или в другой форме), который путем электронного депонирования помещен в институтский ОА, включенный, в свою очередь, в онлайнную научную инфраструктуру, возможен автоматический сбор следующих данных:

- 1) статистика использования РНТД в виде подсчета связей цитирования данного РНТД с учетом качественных характеристик связей;
- 2) статистика профессионального признания РНТД как подсчет рецензируемых научных журналов и изданий, опубликовавших материалы с изложением этого РНТД с учетом их импакт-фактора, алгоритм расчета которого также может быть существенно улучшен;
- 3) статистика востребованности РНТД как подсчет количества просмотров его карточки (описания) и скачиваний полного текста.

Указанная функция дает следующие возможности:

- Ученому –
 - лучшие возможности профессиональной самореализации,
 - оперативную сигнальную систему об отношении научного сообщества к результатам исследований ученого;
- Научно-педагогической институции –
 - усиление правильных мотиваций для сотрудников,
 - реалистичные показатели результативности работы подразделений и отдельных ученых;
- Научному сообществу –
 - повышение эффективности и «прозрачности» расходования общественных фондов,
 - детальную наукометрическую картину структурных и динамических изменений в науке.

Метаданные и средства лингвистического обеспечения системы НПИ

Состав метаданных и средств лингвистического обеспечения системы НПИ в целом определяется:

- 1) наличием разнообразных систем метаданных в уже созданных ресурсах;
- 2) требованиями к единой системе метаданных в рамках протокола ОАИ-РМН.

Перечислим некоторые из средств, используемых в настоящее время в коллекциях и системах, которые предполагается включить в систему НПИ.

НПБ им К.Д. Ушинского

Для организации библиографических метаданных используется формат MARC в версии ИРБИС. Часть электронного каталога представлена в графическом формате. В качестве классификационных языков используется ББК (основной язык), ГРНТИ и УДК (дополнительные), для вербального поиска используется Тезаурус по педагогике и психологии, а также свободная лексика.

Научный архив РАО – для организации метаданных используется стандартная система полей описи архивных фондов, для тематического поиска локальные классификаторы для каждого фонда

ГНИИ ИТТ «Информика» для организации метаданных в некоторых каталогах и коллекциях применяет RUSLUM. Многие коллекции представлены в виде файловых систем и метаданных в стандартных форматах не имеют. В качестве классификационных языков применяются ГРНТИ, ОКСО и другие классификаторы сферы образования. Для лексического поиска применяется, как правило, свободная лексика. Однако на образовательных порталах размещено очень большое количество коллекций и можно найти примеры и фрагменты различных лингвистических средств, разных классификаций и др.

РГБ для Электронной библиотеки диссертаций применяет формат MARC в версии MARC-21, а в качестве тематической классификации – ББК.

ВНТИЦ для организации метаданных в БД НИОКР и БД научного реестра применяет формат МЕКОФ, в качестве классификационного языка – ГРНТИ, а для лексического поиска Тезаурус ВНТИЦцентра, предусматривающий, однако, свободное пополнение лексики и потому неотличимый от ключевых слов.

НЭБ использует для организации метаданных собственный формат, а для тематического поиска – ГРНТИ. Для лексического поиска используется свободная лексика.

В других ресурсах, потенциально представляющих интерес для НПИ, можно найти некоторый вариантный состав метаданных, как правило, не слишком развитых, а также различные локальные классификаторы для тематической организации коллекций. В электронных библиотеках, размещенных в интернете, доминируют простейшие системы навигации, дополненные средствами лексического поиска.

Наконец, Инициатива открытых архивов предусматривает для организации метаданных систему Дублинского ядра и свободное применение любых тематических классификаций в любом архиве и коллекции.

Добавим, что в качестве государственных стандартов (включенных в систему стандартов по информации, библиотечному и издательскому делу – СИБИД) до сих пор действовали стандарты на МЕКОФ и ГРНТИ, в настоящее время разрабатывается ГОСТ на систему Дублинского ядра.

Описанное исходное состояние метаданных и лингвистического обеспечения практически однозначно диктует следующие решения.

Дублинское ядро в некотором минимальном варианте будет единым средством организации метаданных. Для действующих коллекций можно будет использовать конверторы, впрочем, из форматов MARC и RUSLOM в формат Дублинского ядра они уже давно разработаны.

ГРНТИ – непосредственно или через таблицы соответствия – будет основным языком тематического поиска и тематического структурирования коллекций. Для коллекций, в которых использование иной тематической классификация является принципиальной (примером может служить Единая коллекция ЦОР, где необходима классификация школьных предметов), используется эта иная классификация и тематический поиск в этой классификации осуществляется с переходом на локальный классификатор. Для некоторых коллекций, например, авторефератов и полных текстов диссертаций, использование ГРНТИ целесообразно параллельно с классификацией ВАК. Следует иметь в виду, что таблицы соответствия для ГРНТИ уже разработаны для классификаций ВАК и УДК и завершаются для ББК.

Для некоторых коллекций в качестве единого нормативного словаря (авторитетного файла) для нормированного лексического (понятийного) поиска можно применять Тезаурус по педагогике и психологии. Именно так предполагается организовать сквозной тематический поиск в Электронном каталоге НPB и справочно-поисковом аппарате Научного архива РАО.

В перспективе имеет смысл рассмотреть для НПИ модель организации обобщенного тематического и нормированного лексического поиска с использованием тезауруса и различных классификаторов и словарей, интегрированных при помощи средств OWL (Ontology Web Language), в рамках концепции семантического веба. При этом может быть использован опыт проектирования ЕНИП РАН. Однако необходимость такого проекта, который, скорее всего, будет достаточно трудоемким, должна быть серьезно обоснована потребностями пользователей системы НПИ.

В качестве еще одной перспективы может быть рассмотрена возможность применения для НПИ средств автоматизации лингвистической обработки текста (АЛОТ) Университетской информационной системы (УИС) РОССИЯ, опыт применения которой имеется в «Соционете». Возможность применения АЛОТ также облегчается, поскольку в УИС РОССИЯ применяется тот же ГРНТИ, а также Тезаурус, концептуально весьма схожий с Тезаурусом по педагогике и психологии.

Таким образом, для системы НПИ могут применяться следующие лингвистические средства:

Для атрибутивного поиска

Средства Дублинского ядра - для поиска по всем архивам и коллекциям, полные средства MARC и RUSLOM – для некоторых коллекций.

Для тематического поиска

ГРНТИ - для поиска по всем коллекциям (возможно, с небольшими исключениями).
Другие классификации – для поиска по отдельным коллекциям.

Для лексического поиска

Свободная лексика – для поиска по всем коллекциям. Тезаурус и некоторые другие словари и авторитетные файлы – для поиска по отдельным коллекциям.

Организационная модель системы НПИ

Координирующим и совещательным органом системы НПИ является Межведомственный совет по информационным ресурсам образовательной сферы при НПБ им. К.Д. Ушинского РАО. Состав Совета утверждается Президентом РАО. В состав Совета по должности входят ответственные представители НПБ, ИНИМ и ГНИИ ИТТ «Информика».

Функции Совета определяются Положением о Совете, также утверждаемом Президентом РАО. Решения Совета, имеющие организационно-экономический характер, подлежат утверждению руководством РАО и Рособразования.

Функцию оперативного управления системой выполняют НПБ им. К.Д. Ушинского, ИНИМ и ГНИИ ИТТ «Информика», распределяющие между собой на основе консенсуса конкретные функции администрирования системой, включая:

- ведение информационного хаба;
- консультирование участников системы;
- ведение словарей и форматов метаданных;
- разработку и поддержку новых сервисов;
- заключение договоров с внешними владельцами ИР НПИ и организация коллективного пользования этими ресурсами.

Функции администратора системы выполняет лицо, назначенное согласованным решением трех головных организаций системы.

Отдельные функции администрирования системой могут быть переданы по взаимному соглашению администрации системы «Соционет».

Литература:

1. Ханнанов Н.К. *Аналитический отчет «Базы данных в области педагогических, психологических и смежных с ним наук: современное состояние». Часть I. М.: ИНИМ, 2008. - 3 п.л.*
2. <http://socionet.ru/>
3. <http://ct.eurocris.org/CRIS2006/>