



СКОВОРОДИНА
Ирина Сергеевна
директор департамента
агрегации научно-
образовательного
контента, группа
компаний «Контекстум»
✉ skovorodina@akc.ru



ЕФРЕМОВ
Павел Владимирович
генеральный директор,
ООО «Национальный
цифровой ресурс “Руконт”»

ТЕХНОЛОГИИ СОЗДАНИЯ, АГРЕГАЦИИ И ИСПОЛЬЗОВАНИЯ НАУЧНОГО И ОБРАЗОВАТЕЛЬНОГО КОНТЕНТА: ОТ КОНЦЕПЦИИ К ПРОЕКТАМ И ПРАКТИЧЕСКИМ РЕШЕНИЯМ

Агрегация и использование результатов научной деятельности – одно из важнейших направлений, поддерживаемых Министерством образования и науки РФ как в рамках исполнения государственных программ, так и в направлении формирования Единого информационного пространства Минобрнауки.

24 ноября 2014 г. в Вычислительном центре им. А.А. Дородницына Российской академии наук (ВЦ РАН) состоялась II научно-практическая конференция «Технологии создания, агрегации и использования научного и образовательного контента», организованная **Консорциумом «Контекстум»** совместно с **ВЦ РАН** и Национальной ассоциацией исследовательских и научно-образовательных электронных инфраструктур «**e-АРЕНА**». Конференция проводилась при поддержке Комиссии по информатизации Федерального

агентства научных организаций России и Российского фонда фундаментальных исследований.

В работе конференции приняли участие свыше 60 представителей ведущих научных центров и библиотек Российской академии наук, руководителей научных библиотек и издательских подразделений федеральных и исследовательских университетов, вузов, а также специалистов коммерческих структур, занимающихся вопросами агрегации и использования научно-образовательного контента.

Мероприятие было запланировано как следующий шаг в реализации концепции, лежащей в основе создания системы агрегации и использования результатов научной деятельности, разработанной организаторами конференции в рамках выполнения НИР «Разработка научно-методических и нормативно-организационных основ создания системы агрегации и исполь-



проект «РУАЭСТ» по количественным показателям должен превзойти все имеющиеся в России аналоги

зования результатов научной деятельности государственных академий наук». Концепция всесторонне обсуждалась участниками первой конференции, а ее функционал проектировался как развитие самостоятельных направлений и проектов. По итогам первой конференции была принята резолюция с предложениями по дополнению функционала системы сервисами, связанных с большими объемами данных и обеспечивающими интеграцию с существующими информационными системами.

Открыл конференцию **Игорь Соколов** (Главный ученый секретарь Президиума РАН). Он отметил особое значение и важность формирования системы сбора и анализа результатов научно-исследовательской деятельности не только с точки зрения разработки планов проведения научных исследований, но и технологий и методов, которые позволяют анализировать современное состояние и прогнозировать будущее науки.

Во вступительном слове **Марат Биктимиров** (зам. директора ВЦ имени А.А. Дородницына РАН) подчеркнул актуальность создания подобных систем.

Новый механизм организации совместной деятельности участников социо-экономических и научных систем внутри сообществ ученых на примере системы «Соционет» (socionet.ru) представил в своем выступлении ученый **Центрального экономико-математического института РАН**. Уже сейчас «Соционет» представляет российскому научному сообществу современную инфраструктуру взаимодействия на принципах открытой науки, интегрированную с международной Research e-Infrastructure.

Представитель открытой электронной библиотеки «КиберЛенинка» в своем докладе отметил, что ведущие страны мира уделяют особое внимание открытости результатов научных исследований. На основании опыта, полученного при реализации проекта «КиберЛенинка», был предложен план создания платформы открытого журнала.

В качестве примера модели открытой науки был представлен новый проект «РУАЭСТ» – сводный цифровой российский

архив статей периодических изданий. По мнению **Виталия Иванова** (консорциум «Контекстум»), «РУАЭСТ», не отрицая принципов и технологий открытой науки, уделяет особое внимание балансу интересов авторов, правообладателей, контент-провайдеров и пользователей, применяя технологию лицензионно-договорной работы «Контекстум».

Сегодня создан задел из 700 наименований журналов, который лег в основу действующего прототипа «РУАЭСТ» (www.ru aest.ru). В перспективе проект «РУАЭСТ» по количественным показателям должен превзойти все имеющиеся в России аналоги. Число доступных российских научных журналов с глубиной архивов с момента создания должно превысить отметку в две тысячи.

В случае отсутствия необходимой читателю научной статьи, по заявке, в короткие сроки, будет оцифрована и предоставлена заказанная публикация.

Партнерами проекта выступают консорциум «АРБИКОН» и российские библиотеки, выполняющие аналитическую роспись статей российской научной периодики.

Дальнейшие выступления были посвящены практической реализации концепции в технологиях, приложениях и сервисах.

Так, в докладе представителя **Библиотеки по естественным наукам (БЕН РАН)** был дан обзор экспертной системы централизованного комплектования БЕН РАН и создания распределенных библиографических баз данных.

Технологии управления жизненным циклом электронных публикаций в информационной системе научного журнала рассматривались ученым **Казанского (Приволжского) федерального университета** на примере функционирования платформы научных изданий science.tatarstan.ru, которая предоставляет возможности размещения научных журналов в электронном виде.

Полный цикл редакционной подготовки научной публикации в редакции научного журнала на основе информационной системы Math-Net.Ru был показан в совместном докладе сотрудников **Математического института им. Стеклова РАН**. Разработкой и апробацией системы занимались представители восьми ведущих математических научных журналов РАН.

В докладе специалиста **Института системного анализа РАН** было представлено дальнейшее развитие технологии Eхastus Expert. Система предназначена для оценки оригинальности и выявления заимствований в текстах структурированных документов.

Отличительными особенностями системы являются полный лингвистический анализ текстов, включая морфологический и синтаксико-семантический анализ, нечувствительность к перефразированию и высокая скорость выявления заимствований.

Система поддерживает функции автоматического извлечения метаинформации, выявления в анализируемом научном тексте заимствованных фрагментов и источников заимствования, автоматического распознавания корректности/некорректности выявленных заимствований. На основании этого определяется степень оригинальности научного текста с учетом даты публикации.

Алгоритмы Eхastus Expert участвовали в международных соревнованиях по поиску заимствований CLEF-2014 и показали высокие результаты по качеству и скорости функционирования.

Практическое внедрение система получила в Консорциуме «Контекстум», сервисе поиска текстовых заимствований «Руко́нтекст» цифрового ресурса «Руко́нт».

Доклад **Павла Арефьева** (Национальный фонд подготовки кадров) затрагивал вопросы публикационной карьеры российского ученого. Основные проблемы, с которыми сталкивается автор, – это несоответствие системы научной коммуникации учету и оценке научной деятельности, а также тенденция, что в современной науке исследовательская работа ведется большими коллективами ученых и может привести к коллективизации автора и потере его как научного индивида.

С другой стороны, в основе неосведомленности исследователя как автора и малой доли публикаций российских авторов в международных журналах лежит не только слабое знание языков, но и отсутствие навыков академического стиля изложения и работы с библиографическими и информационными ресурсами.

Решение этих проблем видится в создании открытых ресурсов и архивов авторских публикаций (препринтов) с обеспечением процедуры рецензирования, а также развитие навыков авторов в области академического письма и работы с библиографическими и информационными ресурсами.

В выступлении специалиста ЗАО «Эвентос» были подняты вопросы использования семантических веб-технологий в качестве основы для интеграции разнородных приложений, использующих структурированные базы данных. Особое внимание было

уделено задаче связывания данных с данными из других хранилищ с целью предоставления необходимой информации.

Подходы к обеспечению доступности информационных ресурсов через Интернет рассматривались в докладе представителя **Государственной центральной научной медицинской библиотеки**. Сложность решения этой задачи кроется в конфликте между социальной ролью библиотек (доступ должен быть бесплатным) и коммерческими интересами правообладателей. Для решения этого вопроса предлагается технология электронного библиотечного абонемента (ЭБА), позволяющая использовать электронный экземпляр произведения с соблюдением условий, установленных правообладателем, включая ограничение доступа на одновременное использование правомерно приобретенных произведений. Технология ЭБА внедрена в Центральной научной медицинской библиотеке.

В заключении участники конференции высказали свои предложения по решению указанных проблем. Проект интеграции модуля «Электронная библиотека» АБИС МегаПро с сервисом поиска текстовых заимствований «Руко́нтекст» рассматривался в выступлении Владимира Грибова (генеральный директор «МегаПро»). Игорь Попов (и.о. директора издательства МГУ) предложил рассмотреть возможность создания национальной системы оперативной публикации первичных результатов научных исследований. Александр Антопольский (зав. лабораторией Института научно-педагогической информации Российской академии образования) внес предложения по организации научного информационного пространства России в условиях современной политики государства в научно-образовательной сфере.

Таким образом, на конференции были подняты вопросы агрегации и обработки больших массивов научной информации, накопленных в научных и образовательных учреждениях, и объединение их в единую информационную

в современной науке исследовательская работа ведется большими коллективами ученых и может привести к коллективизации автора и потере его как научного индивида



систему, а также создание сервисов подготовки и оценки научного контента в рамках современной организации научно-исследовательских работ. Определяя вектор дальнейшего развития науковедческих технологий и учета результатов научно-исследовательской деятельности, участники конференции отметили актуальность задачи агрегации и обработки больших массивов информации, накопленной в российских научных и образовательных учреждениях. Без предоставления эффективного доступа к этим материалам невозможна реализация государственных программ в области науки и образования.

Учитывая значимость рассматриваемых на конференции вопросов, было предложено:

- отметить важность создания научных электронных библиотек открытого доступа (Open Access) и перевода отечественной научной периодики в открытый доступ, в первую очередь журналов, финансируемых государством, а также создание инфраструктуры публикации научных открытых данных в России;

- считать целесообразным изучение возможности применения технологии «Электронного библиотечного абонемента» для предоставления открытого доступа к произведениям, защищенным авторским правом и хранящимся в библиотечных фондах, в рамках действующего законодательства;

- отметить важность размещения в открытом доступе и под открытой лицензией результатов научных исследований, выполненных с привлечением государственного финансирования;

- разработать организационно-экономическую модель деятельности информационно-библиотечных и архивных служб в сфере науки и образования для обеспечения необходимого уровня информационно-библиотечного обслуживания;

- сформировать перечень информационных ресурсов, отвечающий целям и задачам системы научно-технической информации (НТИ) и с учетом реальных потребностей научно-образовательного сообщества;

- рассмотреть возможность передачи платных ресурсов НТИ в коммерческий сектор.

- разработать организационную модель национальной системы научно-образовательной информации, исходя из межведомственного и межрегионального характера информационной деятельности, наличия инфраструктуры, кадров и других параметров.

- организаторам Конференции рассмотреть вопрос о проведении III Научно-практической конференции «Технологии создания, агрегации и использования научного и образовательного контента» в 2015 году.

