

НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 1. ОРГАНИЗАЦИЯ И МЕТОДИКА
ИНФОРМАЦИОННОЙ РАБОТЫ

ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 12

Москва 2013

ОБЩИЙ РАЗДЕЛ

УДК [002:001.89] : 004

А.О. Еркимбаев, В.Ю. Зицерман, Г.А. Кобзев, В.А. Серебряков, К.Б. Теймуразов

Технология научных публикаций в среде «открытых связанных данных»*

Рассмотрена концепция и компьютерные методы реализации новой формы научной коммуникации – технологии связанных данных. Концепция выдвинута создателем Интернета Тимом Бернерс-Ли в рамках общей идеи Семантического Веб (Semantic WEB) с внедрением в интернет-страницы описаний, которые понятны компьютерам. Приведен обзор технологий и инструментов, позволяющих публиковать в Интернете открытые данные так, что происходит их автоматическое связывание с тематически родственными ресурсами. Показаны многочисленные преимущества новой формы публикаций в сфере науки – интеграция разнородных данных, доступ из публикаций к исходным массивам «сырых» данных и к программному обеспечению, стандартизация терминов и понятий предметной области за счет апелляции к размещенным в сети онтологиям и словарям

Ключевые слова: интеграция данных, связанные данные, базы данных, Семантический Веб

ВВЕДЕНИЕ

Последние десятилетия опережающее развитие информационных технологий оказывает исключительно сильное воздействие на сферу научной ин-

формации, прежде всего, на средства хранения и распространения данных и знаний. Здесь можно отметить выделение информатики как самостоятельного направления в ряде дисциплин (например, гео- и биоинформатики), всеобщее использование библиографических и фактографических баз данных (БД), переход от печатных форм издания к электронным,

* Работа выполнена при поддержке РФФИ – проект № 13-07-00218

активное использование Интернета для распространения данных и многое другое. Отражением новых тенденций стало появление специального термина **e-Science**, который означает преобладающую роль информации и её обработки в научных исследованиях. До определенного момента компьютерная экспансия носила, в основном, технический характер, резко увеличивая объем информационного потока и упрощая способы распространения и методы доступа к данным. Более того, всеобщий переход к электронным ресурсам породил новые проблемы так называемой *интероперабельности*, т.е. обеспечения соответствия форматов и структур данных в среде неоднородных источников.

Перспектива качественного скачка возникла с появлением в 2001 г. принципиально новой концепции – Семантического Веб (*Semantic WEB*), выдвинутой создателем Интернета Тимом Бернерс-Ли [1]. В это понятие вкладывается представление о будущем сети, когда в содержание страниц будет внесена определенная структура, позволяющая программным агентам понимать их смысл и выполнять предписания пользователей. Взаимодействуя в сети, агенты должны будут иметь для каждого ресурса формальное представление знания. Роль представления общей, явной и формальной спецификации знания авторы [1] возлагают на онтологию. Онтология, рассматриваемая как подлинное ядро *Semantic WEB*, это система понятий предметной области, которая представлена как набор сущностей, соединенных различными отношениями. Именно онтология представляет знания в виде формальной структуры, доступной для компьютерной обработки. В 2004 г. World Wide Web Consortium (W3C)¹ предложил универсальный стандарт для сетевого обмена онтологической информацией – Web Ontology Language (OWL). С помощью OWL эксперты предметной области и разработчики приложений могут создавать, модифицировать и соединять различные онтологии.

Хотя задуманная идея *Semantic WEB* относится к ресурсам любой тематики (бизнес, искусство, политика и т. п.), именно естественнонаучные дисциплины образуют наиболее подходящую «площадку» для отработки новых концепций. Понятийный багаж таких наук как физика, химия, астрономия и др. уже исходно достаточно формализован, чтобы лечь в основу онтологического описания. Тем самым открывается принципиальная возможность интеграции научных данных за счет спецификации содержимого разнородных источников. Среди других возможностей *Semantic WEB* – автоматизация обработки текстов, высокоточный информационный поиск, средств-

ва логического вывода и проч. В нашей статье [3] были рассмотрены уже реализованные примеры интеграции данных по физическим свойствам материалов с использованием новой концепции, что говорит о ее большом потенциале. В то же время, несмотря на отдельные успехи, нельзя сказать, что идея *Semantic WEB* нашла к настоящему времени широкое применение.

Уже в 2006 г. Тим Бернерс-Ли [4] выдвигает идею, являющуюся порождением и развитием концепции *Semantic WEB*, которая смещает главный акцент на публикацию так называемых «открытых связанных данных». Другой источник этой идеи – возникшее примерно в те же годы движение за «открытые данные», прежде всего, в области политики, административного управления, финансов и проч. [5]. Среди организаций, практикующих открытую публикацию своих данных, правительства ряда стран, Всемирный банк, ООН, Товарищество в области возобновляемых источников энергии (REEEP) и многие другие (подробный обзор см. [6]).

Данные признаются «открытыми», если они доступны обществу при соблюдении ряда принципов: полнота, свобода распространения без каких-либо ограничений в виде авторских прав, патентов и прочих механизмов контроля. Но потенциал «открытых данных» в полной мере проявляется, когда они конвертированы в систему «связанных открытых данных» путем представления их в специальном формате RDF (Resource Description Framework, см. далее) для идентификации элементов, которые они содержат. Путь от просто «открытых данных» к «связанным открытым данным» был описан Тимом Бернерсом-Ли [7] в виде «пятизвездной конструкции», где на каждом из этажей открываются новые опции для представления данных:

- ★ информация доступна на WEB в любом формате при действии открытой лицензии;
- ★★ информация доступна в виде структурированных данных;
- ★★★ допускается использование открытых форматов (скажем, CSV вместо Excel);
- ★★★★ для всех объектов используются URI идентификаторы, позволяющие выделять их посредством браузера как индивидуальные данные;
- ★★★★★ данные связываются с другими данными, образуя единый контекст.

Публикация в среде «связанных данных» представляет издателям и потребителям информации значительно больше возможностей в сравнении с простым размещением данных в традиционном **Web of Documents**. В отличие от гипертекста, где *линки* связывают отдельные документы, записанные в HTML, технология связанных данных обеспечивает связи между произвольными *вещами*, отмеченными в документе посредством URI, которые могут идентифицировать любые объекты, персоны или концепции. Среда, расширяющая возможности **Web of Documents**, получила название **Web of Data** или **Web 3.0**. Объединяя стандартным образом структурированные документы, эта среда представляет, с точки зрения пользователя, гигантскую БД со столь

¹ W3C - международный консорциум, образованный в 1994 г. в рамках CERN. Целью его создания была разработка общих протоколов, расширяющих интероперабельность ресурсов WWW, а также руководство эволюцией WEB. Консорциум разрабатывает рекомендации по новым технологиям, а также спецификации, имеющие статус стандарта; поддерживает обширный репозиторий документов о разрабатываемых и принятых стандартах, прототипах инструментальных средств, приложениях, демонстрирующих использование новых технологий. На русском языке подробное описание W3C можно найти в справочнике М.Р. Коголовского [2].

же эффективными возможностями поиска релевантной информации, как и в обычной БД.

Принципы «связанных данных», впервые выдвинутые в работе [4], обеспечили руководство для издателей, которые начали осваивать новую технологию. Появились технические документы, регламентирующие издательскую практику в среде «связанных данных», и специальный инструментарий в виде браузеров и поисковых машин, обеспечивающих при работе с **Web of Data** те же возможности, что и средства навигации в обычном **Web of Documents**. Уже сегодня в сети доступно множество связанных данных. Так называемые LOD (Linked Open Data) облака охватывают более чем 50 млрд фактов из множества областей таких, как география, масс-медиа, биология, химия, экономика, энергетика и т.д. Один из лучших и наиболее доступных примеров использования новой технологии в научной сфере – портал www.reegle.info, обеспечивающий автоматическую разметку документов, относящихся к возобновляемой энергии, энергоэффективности и проблемам изменения климата. Помимо поддержки тематического облака, портал предоставляет возможности краткого изучения технологий связанных данных, в частности, описание модели RDF, ссылки на словари и онтологии, правила связывания документов и проч. Далее мы подробно рассмотрим основные принципы и стандарты, используемые при связывании данных в сети, и приведем наиболее впечатляющие примеры возможностей новой технологии в пространстве и интеграции естественнонаучных данных.

БАЗОВЫЕ ЭЛЕМЕНТЫ – ФОРМАТЫ, СТАНДАРТЫ, ОНТОЛОГИИ

Т. Бернерс-Ли, выдвинувший идею Linked Open Data [4], предложил четыре базовых принципа, на основе которых данные должны создаваться и распространяться в сети:

1) для каждой сущности, включенной в данные, такой как персона, документ, абстрактное понятие и т.п., должен быть задан и использован уникальный идентификатор, так называемый URI (Uniform Resource Identifiers);

2) для реализации доступа к этой сущности в сети достаточно использовать обращение HTTP URI (HTTP- Hypertext Transfer Protocol);

3) при обращении по URI предоставляется полезная информация с помощью стандартов RDF, SPARQL;

4) включение ссылок на другие URI позволяет находить дополнительную информацию, не содержащуюся в исходном документе, например, сведения о принадлежности персоны определенной организации или о наличии родственной информации по тематике исходного документа.

Указанные принципы требуют идентификации каждого ресурса посредством HTTP схемы, основанной на URI, который служит не только его идентификатором, но и обеспечивает представление структурированных данных. С помощью идентификаторов

устанавливаются гиперссылки между сущностями в различных документах LOD. Тем самым происходит связывание данных в сетевом пространстве, что позволяет LOD-приложениям обнаруживать новые источники данных, неизвестные авторам на момент публикации.

Таким образом, концепция LOD базируется на трех технологиях, каждая из которых поддерживается стандартами группы W3C: HTTP, URI, и RDF. Согласно третьему из принципов [4], в качестве единой модели связанных данных предлагается RDF, которая также поддерживается стандартом W3C. По сути, RDF это некоторая модель для представления данных и метаданных, состоящая из пригодных для машинной обработки утверждений, каждое из которых имеет формальный вид «субъект – предикат – объект» и называется *триплетом*. Два примера, заимствованных из руководства к сайту [reegle](http://www.reegle.info) [8], иллюстрируют смысл и правила записи триплета:

Примеры RDF-триплетов

Субъект	Предикат	Объект
http://reegle.info/actors/2354	http://xmlns.com/foaf/0.1/name	"REEP"
actors:2354	foaf:name	"REEP"

Под субъектом в обоих примерах понимается некоторый *исполнитель* (actor), однозначно определяемый URI. Соответствующий URI также приписан и предикату, т.е. свойству субъекта. В приведенных примерах этим свойством является *имя*. Наконец, для обозначения объекта в этих примерах использован так называемый *литерал* (текстовая строка). Утверждение в обоих примерах сводится к тому, что исполнитель 2354 имеет имя "REEP" (renewable energy and energy efficiency partnership). Сокращение записи в нижней строке таблицы путем использования префиксов позволяет для субъекта и предиката использовать укороченную запись триплета. Общие правила записи в RDF-формате позволяют для объекта использовать как литерал, так и идентификатор. В этом примере вместо "REEP" мог бы стоять соответствующий URI (www.reegle.org).

В модели данных RDF направленная дуга (предикат) соединяет узлы, соответствующие субъекту и объекту. Граф на рис. 1 содержит два триплета (*с высказываниями о цвете и размере футболки*) [9], где URI определяет субъект и оба предиката, а для объектов используются как URI, так и литерал. При наличии общих идентификаторов в разных триплетах компьютер связывает их в единый граф (см. рис. 1), делая это автоматически, т.е. без участия создателя исходного документа. Связывание данных (RDF link) осуществляется за счет триплета, у которого субъект и объект относятся к разным источникам.

Как видно из рис. 2, в графической модели данных, представленной RDF [9], объединение взаимосвязанных ресурсов происходит без явного выделения какого либо одного как исходного (базового). Каждый из RDF-триплетов является элементом глобальной среды связанных данных и может служить в качестве исходной точки для ее просмотра. Таким образом, RDF является моделью данных, служащей для описания ресурсов WEB и спроектированной так, чтобы обеспечить их восприятие компьютером. При

этом RDF лишь предоставляет средства для построения модели, но никоим образом не раскрывает семантику (смысл) утверждений о ресурсах. Интерпретация содержания ресурсов, определяемых URI либо строковым *литералом*, оказывается возможной за счет отсылок к доступным в сети словарям и онтологиям, которые содержат сводки тематически ориентированных терминов. Онтология, сверх того, резервирует термины для определения концепций и их взаимосвязи. Среди онтологий с широкой сферой применения можно отметить **FOAF** (Friend of a Friend) для описания действующих агентов (персон

или организаций), **SKOS** (Simple Knowledge Organisation System) для выделения категорий, **DBpedia properties** для различных атрибутов. Запись в файле **dbpedia: document** означает, что понятие «документ» истолковано в обширной онтологии **dbpedia**, созданной для описания структурированной информации из Википедии. В пределах определенной предметной области действуют и локальные схемы, например, <http://reegle.info/schema.rdf> для возобновляемых ресурсов, так что синтаксис для записи понятия (например, «результат проекта») имеет вид: **reegle:ProjectOutput**.

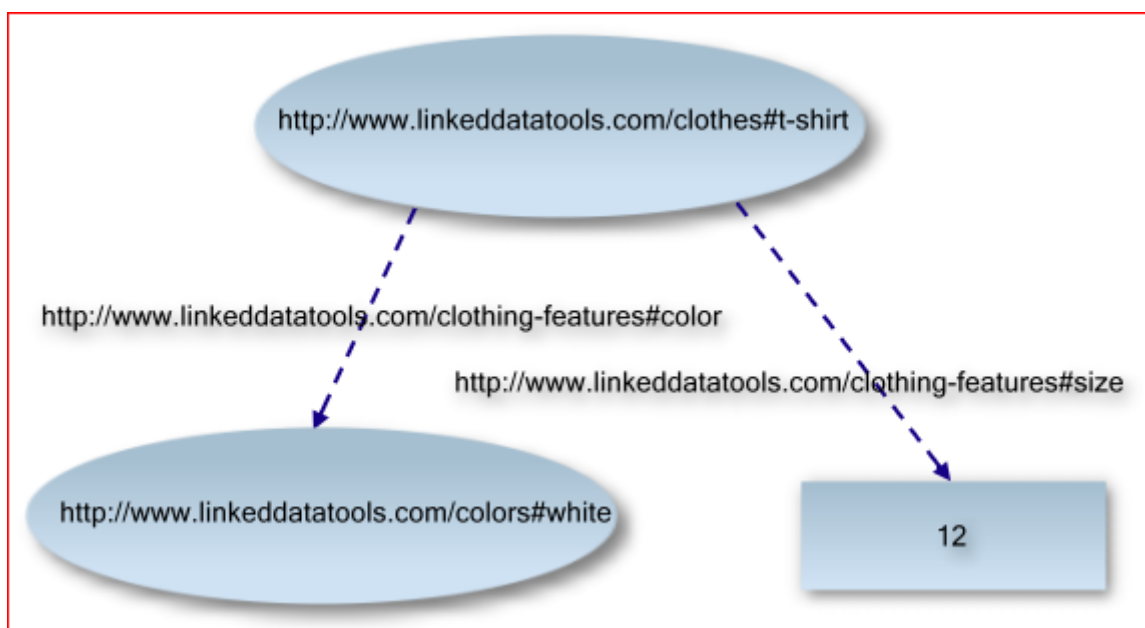


Рис. 1. Типовая запись триплетов [9].

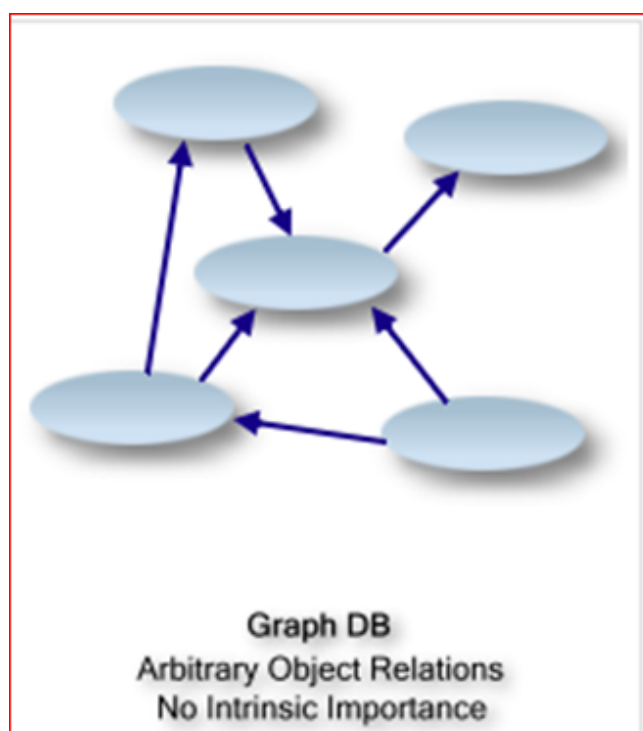


Рис. 2. Графическое представление RDF-модели [9].

Итак, соединение графов модели данных, представленной RDF, с доступными или специально созданными словарями и онтологиями обеспечивает фундамент всей технологии публикации связанных данных. Рис. 3 из презентации [10] дает простейший вид такого связывания на примере двух научных статей, проиндексированных в БД **agris** под условными номерами **CN...179** и **CN...389**. Триплеты раскрывают данные каждой статьи указанием на автора, название и т.д. При этом тематика (subject) обеих публикаций идентифицируется в некотором словаре **agrovoc** под одинаковым идентификатором **c_4416**, что и приводит к автоматическому связыванию двух ресурсов, хотя исходно это не предполагалось, и записаны они были на разных языках. Технология успешна при работе с распределёнными знаниями за счет автоматического сцепления RDF-файлов, размещённых в сети любыми авторами с последующей возможностью найти в собранном документе сведения, которых не было ни в одной из его частей.

Как отмечалось, RDF обеспечивает построение моделей данных, не касаясь самой семантики, отсылая за интерпретацией смысла данных к наличным в сети словарям и онтологиям. Реальный механизм, который позволяет использовать возможности словарей и приписывать моделям RDF-данных семантику, состоит в аннотировании семантических метаданных посредством специального синтаксиса под названием RDF Schema или язык описания RDF-словарей (RDF Vocabulary Definition Language). Соответствующие описания словарей сами записаны на RDF, представляя определения словарей в форме графов и допуская их публикацию в среде Linked Open Data. В целом RDF Schema есть не что иное как семантическое обобщение RDF, обеспечивающее структуру для описания словарей, т.е. предметно-ориентированных классов и свойств. Для отражения семантики здесь вводятся классы и свойства аналогично тому, как это делается в языках объектно-ориентированного программирования, например Java. Отличие состоит лишь в том, что вместо определения класса в терминах свойств его экземпляров, RDF Schema описывает свойства в терминах классов ресурсов, к которым она применяется. Язык, определяемый спецификацией RDF Schema, состоит из коллекции RDF-ресурсов, которые могут быть использованы, чтобы описать свойства других RDF-ресурсов в предметно-ориентированных словарях. Соответствующий документ имеет статус рекомендаций, подготовленных группой W3C в 2004 г. [11].

Еще один из основных стандартов W3C – язык запросов SPARQL (www.w3.org/TR/rdf-sparql-query). Это обозначение является рекурсивным акронимом от англ. *SPARQL Protocol and RDF Query Language*. SPARQL – это язык запросов к данным, представленным по модели RDF, точно так же, как язык SQL (Structured Query Language) обеспечивает запросы к таблицам реляционной БД, а также протокол для передачи этих запросов и ответов на них. SPARQL является рекомендацией консорциума W3C и одной из технологий семантической паутины. Предоставление SPARQL-точек доступа (*SPARQL-endpoint*) является

рекомендованной практикой при публикации данных во всемирной паутине.

Практические рецепты публикаций собственных документов как связанных данных можно найти в руководстве [12]. Помимо технических приемов, обеспечивающих подобную форму публикации, установлен ряд общих принципов, которым должна соответствовать публикация в LOD, чтобы в полной мере проявились преимущества и возможности связанных данных. В рекомендациях W3C [13] они изложены под названием «семь наилучших правил подготовки связанных данных»:

1. Моделируйте данные.
2. Присваивайте понятиям URIs.
3. Всякий раз, когда возможно, используйте словари.
4. Обеспечьте самоописание данных в форме, доступной как человеку, так и компьютеру.
5. Конвертируйте данные в RDF-формат.
6. Обеспечьте публикацию наличием лицензии для свободного и легального распространения данных.
7. Обеспечьте *хостинг* для публикуемых данных и примите меры к их широкому рекламированию.

Первые три правила включают предварительный неформальный анализ предполагаемой публикации на предмет четкого разделения собственно данных и метаданных, а также исключения из публикации нерелевантной информации, которой засорены «сырые» данные, извлекаемые из множества источников: БД, XML-файлы, CSV-таблицы² и т.п. После этого автор должен сформировать модель представления данных в RDF-формате, для чего следует создать словари и тезаурусы используемых понятий, а каждую из сущностей наделить идентификатором URI. Важным моментом в предварительной подготовке является поиск и подбор уже действующих словарей, например, с использованием каталога <http://lov.okfn.org/dataset/lov>. Правило 4 призывает автора представлять данные как «самоописываемые», когда сведения о способах кодирования информации даются в самих RDF-файлах. Доступность этой информации, как для человека, так и для компьютера, обеспечит реальное выполнение продекларированных целей Linked Open Data.

Важным требованием при распространении открытых данных является наличие лицензии, оформляющей их свободное и легальное распространение (правило 6). Хотя принцип открытых данных означает, что они доступны как «общественное достояние» без ограничений авторского права, но само «освобождение от ограничений» обеспечивается посредством публичных лицензий. Без четкого указания лицензионных условий данные не приобретают статус «открытых», даже если они выложены в машиночитаемом формате в сети. Наиболее подходят для этих целей лицензии, выдаваемые неприбыльной организацией **Creative Commons**, целью которой является легальное распространение и использование знаний и результатов творчества.

² XML – eXtensible Markup Language, CSV – Comma-Separated Values

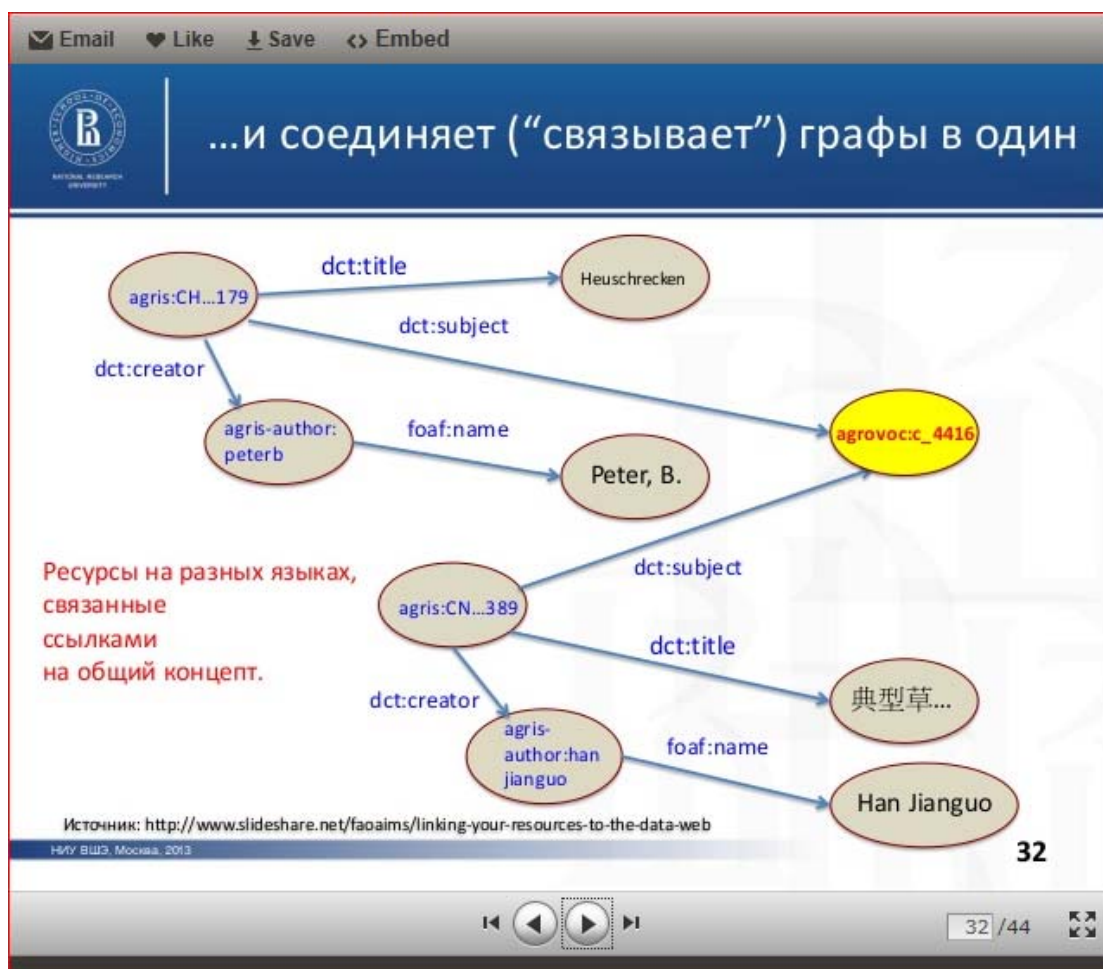


Рис. 3. Пример связывания документов в среде LOD [10].

Наконец, важная сторона публикации данных в LOD - их рекламирование с тем, чтобы они были знакомы более широкому кругу пользователей сети. Наряду с другими способами рекламы, один из лучших – добавление собственных наборов в так называемое **LOD cloud** (<http://richard.cyganiak.de/2007/10/lo/>), визуальное представление связанных наборов с поддерживаемой и обновляемой метаинформацией на некотором узле (<http://thedatahub.org>).

Обширный перечень руководств разного уровня, по которым можно изучить технологию публикаций и поиска связанных данных, представлен на сайте <http://linkeddata.org>. Там же даны рекомендации по выбору различных инструментов: издательских платформ, редакторов и средств контроля («валидаторов») RDF-файлов, специализированных браузеров и поисковых машин. Ряд авторов предпринял попытку создания упрощенных руководств, позволяющих сделать «быстрый старт» в освоении элементов новой технологии, избегая на первых порах проблем, связанных с изучением ее во всех деталях. В частности, можно указать на руководство [8] для сбора и представления материалов по возобновляемой энергетике, подготовленных по технологии LOD, а также краткий курс [9] из 5 уроков (tutorials) для знакомства с ключевыми элементами технологии: RDF, RDFS, семантическое моделирование и т.п.

ИНТЕГРАЦИЯ ЕСТЕСТВЕННОНАУЧНЫХ ДАННЫХ – ВОЗМОЖНОСТИ LINKED OPEN DATA

Применительно к науке в технологии LOD выявляется целый ряд граней, отвечающих давним потребностям сообщества ученых, включая быстрое распространение знаний, стандартизацию понятий и терминов, интеграцию текстов и структурированных данных и т.п. На рынке неоднократно предлагались инструменты, в той или иной степени способные частично удовлетворить эти потребности. В качестве примера можно привести средства, предлагаемые международной БД Web of Science, которые позволяют, отталкиваясь от определенной публикации, найти целый кластер тематически родственных (related records) за счет связывания библиографических данных и создания карты цитирования с двумя поколениями прямых и обратных цитат. Надежды на более глубокий уровень интеграции в течение ряда лет связывались с возможностями предметно-ориентированных версий XML, обеспечивающих явное представление метаданных для описания структуры и семантики информационных ресурсов Web. Была предложена даже новая концепция научной публикации в виде XML-документа, содержащего как обычный текст, так и структурированные данные

с реляционной или иерархической структурой [14]. Постепенно пришло осознание недостатков XML-технологий при их сопоставлении с возможностями Semantic WEB. XML имеет весьма ограниченные средства для интеграции ресурсов, охватывающих ряд предметных областей. В XML-документе сильнее поддерживается синтаксис, а не семантика, поскольку теги структурируют данные внутри документа, но имеют весьма слабое отношение к содержанию данных. В этом плане технологии Semantic WEB, включая LOD, открывают значительно большие возможности в представлении и интеграции данных с учетом как их структуры, так и семантики. Все проблемы с интеграцией, обычно возлагаемые на потребителя данных, в этом случае перекладываются на издателя, использующего богатый инструментарий в виде RDFS, OWL и т.п. Такая форма обеспечивает: широкое **распространение**, когда издатель предоставляет открытый доступ посредством стандартных интерфейсов, таких как SPARQL или URI; **интеграцию** данных за счет поддержки списка связей (*линков*) между различными RDF-данными, обеспечивая доступ к ним через запросы пользователя; **нормализацию** за счет использования в RDF данных общего набора словарей и онтологий.

Технологии Semantic WEB в медико-биологических науках

На сегодняшний день уже имеется определенный опыт публикации связанных естественнонаучных данных. По ряду причин наибольшая активность в освоении этих технологий наблюдается в сфере медико-биологических наук. Показателем особого внимания консорциума W3C к этой области служит создание в его структуре специальной группы – *Health Care and Life Sciences Interest Group (HCLS IG, www.w3.org/blog/hcls)*. Деятельность этой группы направлена на широкое внедрение технологий интеграции в трех направлениях: биологии, традиционной медицине и применении их методов в относительно новой области под названием **трансляционная медицина**.

В биологии главная цель использования LOD – справиться с огромным объемом новых данных, получаемых в ходе исследований. Эти исследования проводятся в широчайшем диапазоне масштабов: молекулы, клетки, клеточные структуры, ткани, органы, организмы, популяции и экосистемы. При этом используется множество экспериментальных методик, инструментов и реагентов. Можно указать ряд направлений, например, исследования генной экспрессии, фенотипов, химический скрининг, поставляющих особо большие массивы данных, на базе которых делаются заключения и выдвигаются новые гипотезы. Абсолютное большинство этих данных сосредоточено во множестве гетерогенных БД, что требует больших усилий по их интеграции путем создания интегрирующих моделей данных [2, 3].

Группа *HCLS IG* подготовила набор технических рекомендаций по конверсии исходных данных в

RDF-формат. В частности, руководство [15] дает исследователям развернутые и документированные ответы на такие вопросы, как, например: методы и инструменты отображения реляционных БД в RDF; конверсия неструктурированных данных (тексты, графика, и т.д.) в RDF; возможности для персон и программных агентов поиска и использования опубликованных RDF; лицензирование публикуемых в LOD данных. Всего руководство включает 14 развернутых рекомендаций по всем аспектам представления связанных данных, причем их сфера приложения выходит за рамки медико-биологической тематики.

Второе направление, курируемое группой *HCLS IG*, где результативность методов LOD оказалась достаточно высокой, **традиционная клиническая медицина**. Масштаб применения этих методов определяется потребностями мощной информационной инфраструктуры, объединяющей клинические данные с данными геномных исследований, биоинформатики, химической информатики и экологическими данными. Одна из основных задач, решаемых группой при внедрении технологий LOD, совмещение их с действующими на Западе стандартами сбора и обмена медицинскими данными, такими как **The Clinical Data Interchange Standards Consortium (CDISC)** или **Health Level Seven (HL7)**. Другая задача – создание формальных онтологий для клинической медицины и исследований с использованием таких средств, как OWL и RDFS.

Особенно востребованы методы связанных данных оказались в так называемой **трансляционной медицине** – относительно новой отрасли, которая активно использует данные доклинических исследований в повседневной диагностике и лечении, а также связывает разработку лекарств с непосредственной реакцией пациента для коррекции правильного выбора препарата, дозы и времени приема. Естественно, что ключевым моментом в этой связке фундаментальной и практической медицины является эффективная процедура связывания данных, полученных в лабораторных и клинических испытаниях.

Основываясь на достигнутом опыте представления связанных данных, группа *HCLS IG* разработала и рекомендовала пользователям детализацию технологического процесса (*workflow*) трансформации исходных данных в RDF-триплеты и связывания их с другими источниками (рис. 4.). В работе [16] рассмотрено несколько конкретных примеров реализации этого процесса: связывание RDF-данных, полученных с использованием технологии ДНК-микрочипов для анализа изменения экспрессии генов; связывание многочисленных данных по свойствам и испытаниям лекарств; генерация RDF-индекса концепций (который связан с БД побочных эффектов, вызванных действием лекарств) в неструктурированных клинических докладах.

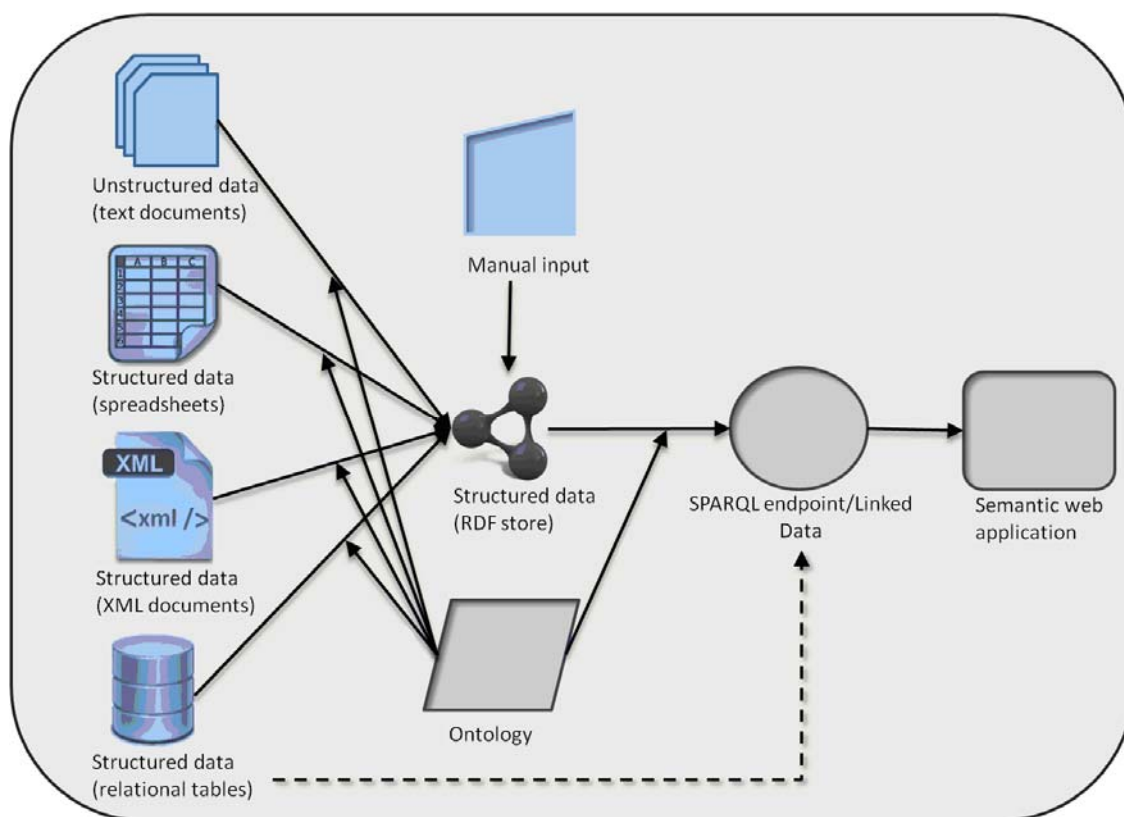


Рис. 4 Схематизация рабочего процесса при отображении неструктурированных и структурированных данных в RDF-формат [15].

Одна из проблем, с которой специалисты столкнулись при связывании медико-биологических данных, заключается в том, что исходно они в заметной степени были собраны в реляционных БД. Общее мнение, сложившееся в сообществе людей, использующих Semantic WEB, состоит в том, что данные «надо оставить там, где они лежат» и генерировать RDF синхронизированно, чтобы всегда оставалась возможность обновляемого доступа к исходным данным. Многие специалисты, использующие рекомендации группы *HCLS IG*, предпочитают формат реляционных данных (relational data base - RDB), подчас импортируя исходно неструктурированные текстовые данные именно в БД, предвзявляя все процедуры их конверсии в RDF-формат. По этой причине сложилось целое направление работ с целью построить отображения типа RDB → RDF (RDB2RDF).

По сравнению с реляционной моделью, структура RDF более выразительна, а данные, записанные в RDF, программные агенты могут интерпретировать, обрабатывать и проводить рассуждения. По-видимому, впервые идеи такого отображения высказаны еще в 1998 г. Т. Бернерсом-Ли, который рассматривал сходства и различия моделей RDF и «сущность-связь» [17]. Начиная с 2009 г. по этой теме работает специальная группа RDB2RDF Working Group Charter (www.w3.org/2001/sw/rdb2rdf/). Наиболее распространенный инструмент для представления реляционных БД в качестве связанных данных – D2R Сервер (<http://d2rq.org/d2r-server>). Используя

декларативный язык, издатель устанавливает отображение между реляционной схемой БД и целевым RDF-словарем. В итоге сервер публикует связанные данные, позволяя клиенту запрашивать данные из БД посредством SPARQL-протокола. Использование сервера включает несколько шагов: (1) скачивание и установка программных средств сервера; (2) автоматическое создание D2RQ отображения схемы БД; (3) ручная настройка отображения путем замены автоматически сгенерированных терминов на более адекватные, подобранные из известных и общедоступных RDF-словарей; (4) установка RDF-ссылок на внешние источники данных; (5) установка нескольких RDF-ссылок из существующих LOD-данных на ресурсы в новом наборе, отвечающем БД, так, чтобы поисковые роботы, обходящие страницы сети, могли обнаружить новые данные. Сервер D2R поддерживает большинство популярных БД: Oracle, MySQL, PostgreSQL, SQL Server, HSQLDB, Interbase/Firebird.

Смысл всей операции по конверсии RDB → RDF не только унификация формата, но и некоторое обогащение содержания за счет явного моделирования соотношений между сущностями, которые в реляционной БД были несущественны, и, конечно, за счет включения в данные предметно-ориентированной семантики. Указанная конверсия порождает «семантический вид» содержания БД, исходно не имеющей RDF-представления. Так, на рис. 4 показано, что, помимо генерации RDF версии БД посредством декларативного отображения, сервер D2R может прямо

обеспечить для протокола SPARQL конечную точку входа в БД. В рекомендациях группы *HCLS IG* обсуждается в деталях вся конкретика, связанная с выбором языков отображения и инструментальных средств по включению содержимого БД в среду Linked Open Data.

Технологии Semantic WEB в химии

В химии идеи Semantic WEB легли на давно подготовленную почву. С 1998 г. существует дисциплина, именуемая **хемоинформатикой**, одной из задач которой является распространение и интеграция многочисленных справочных и экспериментальных данных по свойствам и структуре соединений, константам химических реакций и т.п. В то же время, от других дисциплин химия отличается определенным консерватизмом в отношении концепций открытых данных. Как отмечает автор обзора [18], в химии сильнее, чем в других областях науки, распространение знаний передоверено коммерческим издательствам, препятствующим автоматическому выделению данных из журнальных статей. Это начинает тормозить научный прогресс, особенно в сопредельных областях (в биоинформатике, геномике, фармацевтике), и в химии постепенно начинается внедрение интеграционных технологий. В 2010 г. Американское химическое общество (ACS) провело двухдневный семинар, посвященный использованию RDF-модели, где были рассмотрены уже реализованные продукты, обеспечившие обмен данными как внутри химии, так и особенно в связке с проблемами биоинформатики. По материалам этого семинара в журнале "Journal of Cheminformatics" опубликована тематическая серия "RDF technologies in chemistry" (www.jcheminf.com/series/acsrdf2010) с подробным обзором [19]. На тот момент использование RDF носило спорадический характер, демонстрируя успешность лишь отдельных проектов. Повидимому, первое предложение использовать RDF для представления химических структур сделано в 2004 г. в работе [20], авторы которой ранее предложили для обмена данными язык CML (Chemical Markup Language). Язык запросов SPARQL начинает применяться в 2007 г. в системах аннотирования кристаллических структур. В докладах на семинаре освещались такие проекты, как Bio2RDF, Chem2Bio2RDF и OpenTox, в которых знания в области геномики, химии и фармацевтики загружались в Semantic WEB путем использования RDF. Эти проекты имели целью создание БД с химической информацией, доступной из центральной точки, объединяющей отдельные наборы данных. Реализованы также и менее масштабные проекты использования RDF, например, набор данных Open Notebook Science Solubility data.

Наиболее перспективная технология в семействе RDF (протокол SPARQL) с успехом применялась при решении задач хемогеномики - мультидисциплинарной области, определяющей соответствие между лигандами и мишенями в биологических объектах. Поиск биологически активных соединений проводился по материалам трех БД (PubChem, Uniprot, DrugBank), доступных из сервера Chem2Bio2RDF. При этом сервер обеспечи-

вал доступ и к реляционным БД, трансформируя запросы на языке SPARQL в обычные SQL.

Технология RDF связывает данные о химическом понятии (объект, концепция) с ресурсами, представляя относящиеся к этим ресурсам детали, а протокол SPARQL обеспечивает средства запроса и агрегирования данных. Следующий стандарт Semantic WEB - Web Ontology Language (OWL) связывает RDF-технологии с множеством онтологий. Подобно контролируемым словарям и тезаурусам, онтологии описывают смысл понятий, связывая термины с определениями, воспринимаемыми человеком. В то же время представление содержания в явных терминах позволяет как человеку, так и компьютеру проводить формальные рассуждения и, возможно, находить источники ошибок.

На текущий момент в химии создано не так много онтологий, особенно написанных на языке OWL. Например, Konyk et al. использовали OWL чтобы связать три крупных хранилища (PubChem, DrugBank и DBPedia), обеспечив новые способы получения знаний [21]. Имеется специальная коллекция онтологий медико-биологической направленности (OBO Foundry ontologies, www.obofoundry.org), некоторые из которых охватывают химическую проблематику, например ChEBI (Chemical Entities of Biological Interest), CHEMINF (Chemical Information Ontology), CO (A chemical ontology for identification of functional groups and semantic comparison of small molecules). Для онтологий, записанных в формате OBO, есть средства конверсии в OWL, что обеспечивает применение этого формата в виде универсального формата для интеграции химических данных.

На базе онтологии CHEMINF разработана система CHESS (Chemical Entity Semantic Specification) для представления многоатомных молекул и их составляющих с использованием средств Semantic WEB [22]. Спецификация CHESS включает три широкие категории: (1) химические сущности, т.е. реакции, комплексы, молекулы, функциональные группы, связи и атомы (возможно включение электрона и макромолекул); (2) химические дескрипторы; (3) химическую «конфигурацию», отражающую условия, при которых данные получены, равно как и источник данных. Основное требование - способность представлять понятия химии способом, не зависящим от исходных БД, программного обеспечения и конкретного раздела науки. Для этой цели применяется согласованная система кодирования атомов, связей и функциональных групп, использующая рекомендуемый IUPAC идентификатор InChI keys - линейную нотацию в виде строки символов (www.iupac.org/inchi). Другая особенность системы - ее гибкость, проявляющаяся в способности восприятия различных форматов данных и нотаций, определяющих структурную информацию (наряду с InChI, например, нотацию SMILES, Simplified Molecular Input Line Entry System). CHESS также обеспечивает хранение данных с явными указаниями на условия, при которых они получены, а также на родственные данные. Наконец, немаловажно и то, что наряду с традиционными сервисами поиска и агрегации, CHESS поддерживает и типовые инструменты хемоинформатики для поиска лекарств, хи-

мических аналогов, типовых реакций путем сопоставления с образцами (pattern matching).

Анализ возможностей CHESS [22] показывает, что, невзирая на отмеченный консерватизм, в химии также наблюдается постепенный переход к интеграции данных при достаточном универсализме этой процедуры, т.е. независимости от структуры и формата данных. Разумеется, проект [22] не является уникальным – общая картина по состоянию на 2013 г., представленная в обзоре [23], касается также технологий связывания документов, в то время как в докладах на семинаре ACS [19] акцент делался на интеграции БД. В частности, в обзоре рассмотрена форма так называемых *семантических публикаций* на примере изданий Королевского химического общества (Royal Society of Chemistry – RSC). В его структуре реализован проект RSC Semantic publishing (www.rsc.org/Publishing/Journals/ProjectProspect/index.asp), в рамках которого статьи из журналов за 2008-2010 гг. связаны (*линкованы*) с открытой БД ChemSpider (www.chemspider.com). Рукописи, представленные в RSC, подвергаются семантической разметке с целью выявить важные химические данные, особенно данные по структурам. Разметка обеспечивает «линки» текстов с дополнительными источниками данных по свойствам. Это позволяет поисковым машинам использовать разметку, в частности, чтобы выявить публикации, связанные с конкретной структурой. Подход, реализованный проектом RSC, демонстрирует преимущества публикаций в форматах, совместимых с Semantic WEB. Соответствующая функциональность, связанная с форматами RDF, теперь добавлена в интерфейс гигантской и открытой БД ChemSpider.

ЗАКЛЮЧЕНИЕ

Приведенный материал иллюстрирует далеко еще не раскрытый потенциал Semantic WEB и конкретно – технологий связанных данных в области естествознания. Хотя примеры, приведенные в статье, заимствованы из области медико-биологических наук и химии, аналогичные приложения были развиты и для других областей, например, кристаллографии, термодинамики, наук о Земле и др. Наиболее значимо то, что семантические технологии революционизируют для ученого сам процесс публикации. Научная публикация перестает быть изолированной единицей, отражаемой лишь реферативными службами, а становится элементом глобальной БД. Конверсия ее содержания в RDF обеспечивает связывание с тематически родственными публикациями и БД. Само связывание производится без участия автора, на основе использования доступных в сети словарей и онтологий. Идеальный пример – публикации Королевского химического общества (Великобритания), наделенные семантической разметкой с многочисленными отсылками к релевантным источникам, содержащими дополнительные сведения по структуре и свойствам упомянутых в исходной публикации веществ.

Идея подобного обогащения публикации дополнительными данными восходит к началу 2000-х г.г. [14, 24], когда вместо традиционной формы публикации была предложена структура, допускающая для чита-

теля прямой доступ к исходным численным файлам, полученным в результате эксперимента или численного моделирования. Подобная структура, получившая специальное название *datument* (неологизм, означающий синтез данных и документа), была построена в виде XML-документа, который по замыслу авторов концепции должен заменить обычные формы электронных публикаций, например, в виде PDF-файлов. В рамках этой концепции можно вести тонкую разметку текста с отсылками к репозиториям, где хранятся первичные данные, или к соответствующему программному обеспечению. Более того, возникла широко обсуждаемая идея, когда, наряду с открытостью *datument*'а, оказывается возможным свести научную коммуникацию к взаимодействию авторов и читателей без традиционной роли издателя, ограничив его функцию организацией peer-review и поддержкой импакт-фактора. Основные преимущества *datument* – свободное распространение табличного и графического материала, молекулярных структур, математических конструкций в виде элементов MathML.

Несмотря на большой энтузиазм, проявленный в отношении подобной формы публикации, опора на XML, по-видимому, оказалась сдерживающим фактором. Выше уже отмечалось, что XML-технология заметно слабее при интеграции множества гетерогенных ресурсов, поскольку поддерживает синтаксис, а не семантику. В то же время технология LOD, не ограничиваясь стандартизацией обмена данными, обеспечивает *подлинно семантическую* интеграцию данных путем автоматического связывания множества родственных ресурсов при согласовании терминологии и понятий с общепринятыми словарями и онтологиями.

Еще одно важное преимущество в технологиях Semantic WEB – возможность представить на одном уровне как неструктурированные (тексты статей, документы и т.п.), так и жестко структурированные данные, размещенные в реляционных БД (см., в частности, рис. 4 и описание сервера D2R из руководства [15]).

Наконец, помимо богатых технологических возможностей, вся концепция связанных данных порождает ряд идейных последствий для тех дисциплин, где она показала свои преимущества. Она стимулирует переход к открытым данным, откликаясь на постоянно возникающие инициативы, такие как Pubmedcentral, ePrints Initiative, Open Archives Initiative, Public Library of Science и др. Создается коммуникационное пространство, обладающее заметным преимуществом в сравнении с традиционными формами публикаций, распространяемых коммерчески ориентированными издательствами. Другое последствие новой технологии – повышенное внимание сообщества к словарям и таксономиям. Помимо того, что именно апелляция к общим терминологическим системам обеспечивает связывание документов, необходимость работы с онтологиями наводит определенный порядок в предметной области, стандартизируя терминологию, понятия, единицы измерения и т.п. Наконец, общие правила публикации связанных документов (см. раздел «Базовые элементы...») предъявляют и достаточно серьезные требования на предварительной стадии, как то очистка от засоряющей информации, четкое разделение данных

и метаданных, интерпретация терминов путем отсылок к общедоступным словарям или онтологиям. Тем самым уже на этапе подготовки публикации новая технология побуждает авторов обеспечивать более строгую и стандартизированную форму подачи материала.

СПИСОК ЛИТЕРАТУРЫ

1. Berners-Lee T., Hendler J., Lassila O. The Semantic Web // Scientific American. – 2001. – Vol. 284, № 5. – P. 35-43.
2. Когаловский М.Р. Энциклопедия технологий баз данных. – М.: Финансы и статистика, 2002. – 798 с.
3. Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А., Сон Э.Е., Сотников А.Н. Интеграция баз данных по свойствам вещества. Подходы и технологии // Научно-техническая информация. Сер. 2. – 2012. – № 8. – С 1-8.
4. Berners-Lee T. Design Issues: Linked Data. Online at. – URL: www.w3.org/DesignIssues/LinkedData.html.
5. Open Data – An Introduction “Today we find ourselves in the midst of an open data revolution”. – URL: <http://okfn.org/opendata>
6. Bauer F., Kaltenböck M. Linked Open Data: The Essentials. A Quick Start Guide for Decision Makers // Published by: edition mono/monochrom. – Vienna, 2012. ISBN: 978-3-902796-05-9. – URL: www.semantic-web.at/LOD-TheEssentials.pdf
7. 5 ★ Open Data. – URL: <http://5stardata.info/>
8. Reegle LOD Developer Guide. – URL: <http://data.reegle.info/developers/guide>
9. Introducing Linked Data and The Semantic Web. – URL: <http://www.linkeddatatools.com/semantic-web-basics>
10. Радченко И.А. Введение в концепцию связанных открытых данных (Linked Open Data). Цикл вебинаров Linked Open Data @ AIMS 14 февраля 2013 г. – URL: <http://www.slideshare.net/iradche/linked-open-data-16524818>
11. RDF Vocabulary Description Language 1.0: RDF Schema. W3C Recommendation 10 February 2004. – URL: <http://www.w3.org/TR/rdf-schema>
12. Bizer C., Cyganiak R., Heath T. How to Publish Linked Data on the Web. – URL: <http://sites.wiwi.fu-berlin.de/suhl/bizer/pub/LinkedDataTutorial>
13. Linked Data Cookbook. From Government Linked Data (GLD) Working Group Wiki. – URL: http://www.w3.org/2011/gld/wiki/Linked_Data_Cookbook
14. Murray-Rust P., Rzepa H.S. Scientific publications in XML - towards a global knowledge base // Data Science Journal. – 2002. – No 1 – P. 84-98.
15. Health Care and Life Science (HCLS) Linked Data Guide. – URL: www.w3.org/2001/sw/hcls/notes/hcls-rdf-guide/
16. Marshall M. S. et al. Emerging practices for mapping and linking life sciences data using RDF – a case series // Web Semantics: Science, Services and Agents on the World Wide Web. – 2012. – Vol. 14 – P. 2-13.
17. Berners-Lee T. Relational Databases on the Semantic Web, Design Issue Note, 1998-2009. – URL: www.w3.org/DesignIssues/RDB-RDF.html
18. Adams N. Semantic Chemistry. Опубликовано на сайте “The Voice of Semantic Web Technology and Linked Data Business” (semanticweb.com), 2009
19. Willighagen E.L., Brändle M.P. Resource description framework technologies in chemistry // J. Cheminformatics. – 2011. – Vol. 3, № 15. – URL: www.jcheminf.com/content/3/1/15
20. Murray-Rust P., Rzepa H.S. et al. Chemical markup, XML, and the World Wide Web. 5. Applications of chemical metadata in RSS aggregators // J. Chem. Inf. Comput. Sci. – 2004. – Vol. 44, № 2. – P. 462-469.
21. Konyk M., De Leon A., Dumontier M. Chemical Knowledge for the Semantic Web // Lecture Notes in Computer Science. – 2008. – Vol. 5109. – P. 169-176.
22. Chepelev L.L., Dumontier M. Chemical Entity Semantic Specification: Knowledge representation for efficient semantic cheminformatics and facile data integration // J. Cheminformatics. – 2011. – Vol. 3, № 20. – URL: www.jcheminf.com/content/3/1/20
23. Frey J.G., Bird C.L. Cheminformatics and the Semantic Web: adding value with linked data and enhanced provenance // WIREs Comput. Mol. Sci. – 2013. – doi: 10.1002/wcms.1127. – URL: <http://onlinelibrary.wiley.com/doi/10.1002/wcms.1127/pdf>
24. Murray-Rust P., Rzepa H.S. XML for scientific publishing // OCLC Systems & Services. – 2003. – Vol. 19, № 4. – P. 163-169.

Материал поступил в редакцию 29.07.13.

Сведения об авторах

ЕРКИМБАЕВ Адильбек Омирбекович – кандидат технических наук, зав. информационно-вычислительным центром Объединенного Института высоких температур РАН (ОИВТ РАН), Москва
e-mail: adilbek@ihed.ras.ru

ЗИЦЕРМАН Владимир Юрьевич – кандидат физико-математических наук, зав. лабораторией баз данных ОИВТ РАН, Москва
e-mail: vz1941@mail.ru

КОБЗЕВ Георгий Анатольевич – доктор физико-математических наук, профессор, зав. отделом теплофизических данных ОИВТ РАН, Москва
e-mail: gkbz@mail.ru

СЕРЕБРЯКОВ Владимир Алексеевич – доктор физико-математических наук, руководитель отдела систем математического обеспечения ВЦ им. А.А. Дородницына РАН,
e-mail: serebr@ultimeta.ru

ТЕЙМУРАЗОВ Кирилл Борисович – специалист отдела систем математического обеспечения ВЦ им. А.А. Дородницына РАН
e-mail: kbt@intring.ru

Анализ цитирования как основа для разработки дополнительного модуля в системах антиплагиата

Области использования пристатейных списков литературы расширяются с каждым годом. Являясь основой для определения взаимоотношений между авторами, публикациями и журналами, ссылки дали импульс разработкам различных показателей и индексов, что значительно продвинуло развитие библиометрии и наукометрии в целом. В нашем кратком сообщении предлагается еще один вариант использования ссылок – в качестве дополнительного модуля в системах определения плагиата; показывается, как с помощью сопоставления списков литературы, учитывая их уникальную последовательность в каждой публикации, можно выявлять прямые заимствования в научных трудах.

Ключевые слова: плагиат, анализ цитирования, модели цитирования, научный перевод, библиографические базы данных

ВВЕДЕНИЕ

С ростом количества научных публикаций и увеличением числа исследователей по всему миру параллельно увеличивается число заимствований чужого текста без ссылок на первоисточник. Для противодействия недобросовестным авторам разрабатывается специальное программное обеспечение для обнаружения этих заимствований (см., например [1]).

К настоящему моменту созданы как бесплатные программы ограниченного функционала, так и мощные дорогие системы выявления плагиата в диссертациях, закупаемые крупными университетами и центрами. Все они построены на алгоритмах автоматического индексирования текста и достаточно хорошо выявляют заимствованные фрагменты работ, а также показывают долю оригинального текста. В настоящей работе мы не ставим целью проводить обзор этих систем (они описаны в [2]), но укажем на их общий недостаток, который пока не нашел своего решения: данные системы работают только с одним языком и только с непосредственным копированием текста. Учитывая тот факт, что большая часть научной литературы написана на английском языке, авторам, владеющим английским языком, из неанглоязычных стран часто оказывается достаточным перевести публикацию или несколько публикаций для компиляции, чтобы впоследствии опубликовать работу на родном языке, выдав ее за оригинальное исследование. Укажем также, что если первый тип плагиата, не использующий языковой перевод, распространен в основном в студенческой среде, то второй, более сложный вид, – затрагивает научную сферу. Менее распространен-

ным, но также вызывающим большие затруднения в выявлении, является плагиат идей.

В связи с, как правило, более высоким научным уровнем, переводные публикации быстрее проходят рецензирование в отечественных журналах и впоследствии набирают большее количество цитирований. Переведенные работы используются также при написании диссертаций или для составления отчетов по грантам и федеральным программам, которые финансируются из государственного бюджета. Следует отметить, что нередко исследователи отказываются признавать перевод за плагиат, предпочитая называть его «обильным цитированием» [3], особенно в тех случаях, когда они в начале публикации дают общую ссылку на первоисточник. В качестве положительного аспекта опубликования подобных работ называется отсутствие в ряде научных дисциплин качественных русскоязычных публикаций. При этом не вызывает сомнений, что такие работы не дают нового знания и должны публиковаться именно как переводы, которые, безусловно, необходимы.

Выявить плагиат на основе перевода или плагиат идей в настоящий момент возможно только с привлечением к анализу высококвалифицированных специалистов, которые хорошо ориентируются в современном потоке научных публикаций по своей тематике. Очевидно, что такой анализ может применяться лишь в единичных случаях, он дорог и трудозатратен. Эта ситуация приводит к тому, что недобросовестные авторы чувствуют себя в относительной безопасности и отказываются от проведения самостоятельных и финансируемых государством исследований в пользу менее трудоемкого перевода чужих трудов. В свою

очередь, это сказывается на сокращении количества ценных отечественных разработок, на повторном финансировании исследований, которые уже проводились зарубежными коллегами и доступны в сети, на замедлении роста рейтинга российской науки в мировом пространстве.

РОЛЬ АНАЛИЗА ЦИТИРОВАНИЯ В ВЫЯВЛЕНИИ ПЛАГИАТА

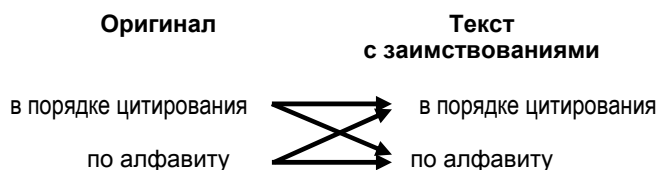
На наш взгляд, эту проблему могла бы оперативно решить автоматизация процессов сличения текстов, независимо от языков, на которых они написаны. В силу нашей совместной работы в сфере перевода и в области библиографии у нас оказался собран большой объем опубликованных русскоязычных статей, отчетов и монографий по нескольким дисциплинам (в частности, биологии, медицине и информатике), которые были созданы на основе переводов зарубежных публикаций. Анализ этих данных позволил нам заметить, что недобросовестные авторы включают в переведенный текст всю совокупность ссылок из оригинальных статей, проанализировав которую можно с достаточно большой точностью определить оригинал. Отметим, что к тем же выводам, что и мы, с некоторым опережением пришли три исследователя из Мэгдбургского университета и Калифорнийского университета в Беркли [2, 4]. В настоящее время это единственный коллектив в мире, ведущий разработку данной проблемы [5]. Интересными в этой связи представляются также более ранние работы исследователей из Калифорнийского университета в Лос-Анджелесе, в которых авторы строят алгоритмы выявления в текстах публикаций тех ссылок, которые цитирующие авторы, скорее всего, не читали [6, 7].

Создание и совершенствование алгоритмов выявления последовательностей ссылок из пристатейных библиографических списков не зависит от семантики и языка текста, а также поддается автоматической обработке, поскольку набор правил приведения ссылок в мире ограничен, а программное обеспечение работы со ссылками позволяет производить автоматическую конвертацию из одного стиля в другой. Сличение выявленных последовательностей ссылок может проводиться с использованием библиографических баз данных, как предметных (например, PubMed или GEOBASE), так и мультидисциплинарных (например, Scopus или Web of Science), которые содержат огромные массивы пристатейных списков литературы. Здесь следует сделать оговорку, что в ряде случаев для сличения последовательностей ссылок потребуется обращение к полным текстам статей – в тех журналах, где правила цитирования предполагают расположение ссылок в алфавитном порядке. В любом случае, обращение в библиографическую БД позволяет выявлять набор общих ссылок.

Таким образом, для успешного создания и эксплуатации автоматизированного модуля выявления плагиата на основе цитирований необходимо прежде всего решить две задачи:

1) разработать алгоритмы выявления моделей цитирования в заимствованном тексте. В разных источниках существуют разные правила цитирования литературы, которая может идти в алфавитном по-

рядке, в порядке цитирования в тексте, может приводиться в виде цифр или в виде «автор – год». Всего при сличении перед исследователем могут быть четыре модели:



2) на основе полученных последовательностей ссылок разработать алгоритмы создания запроса в библиографическую базу данных.

Дальнейшее сличение может проводить вручную носитель двух языков.

Первая задача стала успешно решаться два года назад указанными зарубежными исследователями, предложившими для выявления схем цитирования три различных алгоритма: силу библиографической связи (bibliographic coupling strength), наиболее длинную общую последовательность цитирований (longest common citation sequence) и жадное плиточное размещение цитирований (greedy citation tiling) [4].

Эти исследователи хорошо определили вероятность совместного появления одинаковых ссылок в двух текстах, которая зависит от времени создания цитируемых статей, их принадлежности различным дисциплинам, количества набранных ими цитирований (закон Матфея), тематики отраженных в публикациях исследований и пр., и показали ограничения в применении метода, основное из которых – малое количество ссылок. К прочим ограничениям относятся технические трудности непосредственного выявления моделей цитирования, которые могут сознательно внедряться плагиатором для сокрытия факта заимствования: изменение последовательности ссылок, их масштабирование, при котором та или иная работа цитируется большее или меньшее количество раз по сравнению с оригинальным цитированием, вставка собственных цитат.

Вторая задача только начала решаться на примере использования библиографической базы данных PubMed Central Open Access Subset, содержащей доступные полные тексты статей [8]. Также стал разрабатываться пакет программ, позволяющих в автоматическом режиме загружать для проверки текст и получать результаты на основе сличения пристатейных ссылок со ссылками из статей из БД PubMed [9]. Несмотря на значительные ограничения функционала, алгоритмы программы позволили выявить несколько случаев заимствований, конкретные примеры которых эти исследователи пока не приводили.

Между тем представляется необходимым совершенствование начатых разработок, а также их адаптация к местным условиям: создание автоматизированной системы обработки данных из отечественной БД РИНЦ, разработка русскоязычного программного обеспечения, а также алгоритмов подготовки запроса в библиографические базы данных. Наиболее трудно решаемой является задача автоматизации определения границ заимствования, когда вторичная публикация скомпонована из ряда источников. В этом случае полезным могло бы стать использование кластерного анализа.

ПРАКТИЧЕСКИЙ СЛУЧАЙ

Опишем схему алгоритма создания запроса в библиографическую БД на конкретном примере. На рис. 1 А показана переведенная часть текста в одной из отечественных публикаций, на рис. 1 Б – фрагмент списка литературы к этой же статье. На рис. 2 приведен фрагмент оригинального зарубежного источника. По этическим соображениям мы даем ссылку только на оригинал и не приводим ссылку на отечественную публикацию, в пристатейном списке литературы которой, между тем, отсутствует ссылка на первоисточник.

Не нужно быть экспертом и знать язык оригинала, чтобы заметить также сохранение последовательности используемых латинских обозначений в обоих приведенных фрагментах. Данный пример показывает достаточно трудный для выявления плагиата вариант, при котором потребовалось обращение к полному тексту первоисточника для воспроизводства модели цитирования, так как в пристатейном списке литературы журнала с оригинальной статьей ссылки даны в алфавитном порядке. Из примера видно, что из оригинальных шести ссылок в тексте с заимствованием осталось 5 (одну, очевидно, авторы убрали как избыточную). При этом сохранилась их последовательность. Имея такие данные, можно установить границы заимствованного текста и впоследствии восстановить всю историю создания вторичной статьи.

Для запуска адаптивного иммунного ответа против ВГ необходимы профессиональные антиген-презентирующие клетки (APC), такие как дендритные клетки (DC). Уменьшение численности DC в легких в процессе инфекции приводило к снижению уровня ИФН 1-го типа и CD8⁺-клеток [53]. Альвеолярные макрофаги и ДК одними из первых активируются инфекцией ВГ типа А [54]. Во время заражения альвеолярные макрофаги производят большое количество воспалительных цитокинов, главным образом интерлейкина-6 (ИЛ-6) и фактор некроза опухолей- α (ФНО- α). Кроме того, инфицированные эпителиальные клетки легких выделяют в больших дозах моноцитарный хемоаттрактантный белок 1 (MCP-1), мобилизирующий моноциты, которые могут видоизменяться в DC и свободные макрофаги [55]. Во время активации DC созревают и мигрируют в дренирующие лимфатические узлы легких, где они представляют пептидный антиген ВГ в комплексе с молекулами МНС класса I предшественникам CD8⁺-Т-лимфоцитов. Это вызывает пролиферацию и дифференцировку Т-клеток в эффекторные клетки (цитотоксические CD8⁺-Т-лимфоциты).

В процессе эволюции ВГ приобрел механизмы, которые способны подавлять иммунитет хозяина. В частности показано, что белок NS1 ограничивает продукцию ИФН 1-го типа, а также других цитокинов и хемокинов во время заражения вирусом гриппа А [56]. Более того, ВГ способен предотвращать созревание DC. Зараженные DC имеют сниженную экспрессию молекул МНС класса I и таких коstimулирующих рецепторов, как CD80 и CD86, что препятствует представлению эндогенно экспрессируемых вирусных антигенов CD8⁺-Т-клеткам [57].

Рис. 1 А. Переведенный фрагмент текста в отечественной публикации

induction of antiviral responses in primary human dendritic and respiratory epithelial cells. J Virol. 2009;83:6849-62.

57. Young LJ, Wilson NS, Schnorrer P, Mount A, Lundie RJ, La Gruta NL, et al. Dendritic cell preactivation impairs MHC class II presentation of vaccines and endogenous viral antigens. Proc Natl Acad Sci USA. 2007;104:17753-8.

53. McGill J, Heusel JW, Legge KL. Innate immune control and regulation of influenza virus infections. J Leukoc Biol. 2009;86:803-12.
54. Nicholls JM, Chan MC, Chan WY, Wong HK, Cheung CY, Kwong DL, et al. Tropism of avian influenza A (H5N1) in the upper and lower respiratory tract. Nat Med. 2007;13:147-9.
55. Herold S, von Wulffen W, Steinmueller M, Pleschka S, Kuziel WA, Mack M, et al. Alveolar epithelial cells direct monocyte transepithelial migration upon influenza virus infection: impact of chemokines and adhesion molecules. J Immunol. 2006;177:1817-24.
56. Haye K, Burmakina S, Moran T, Garcia-Sastre A, Fernandez-Sesma A. The NS1 protein of a human influenza virus inhibits type I interferon production and the

Рис. 1 Б. Соответствующие ссылки из пристатейного списка литературы в отечественной публикации.

The quality of the initial innate response against IAVs has profound consequences for the following adaptive response. Triggering the adaptive immune response against IAVs requires professional APC-like DCs. Depletion of lung resident DCs was shown to result in reduced recruitment of pDCs (reviewed in [McGill et al., 2009a](#)), which are the main producers of type I IFN, and CD8+ cells upon IAV infection.

Residential alveolar macrophages and dendritic cells are among the first cells activated by IAV infection. Depending on the virus strain a significant proportion of infected cells are alveolar macrophages ([Nicholls et al., 2007](#)). Upon infection with IAVs alveolar macrophages produce significant amounts of inflammatory cytokines like IL-6 and TNF α , and become highly phagocytic. Additionally, infected lung epithelial cells secrete high levels of monocyte chemoattractant protein 1 (MCP-1), recruiting monocytes, which can differentiate into monocyte derived DCs and inflammatory macrophages ([Herold et al., 2006](#)).

Upon activation, DCs undergo maturation and migrate to the draining lymph nodes of the lung where they present peptide antigens to IAV-specific naïve T cells. This results in proliferation and differentiation of these T cells into effector cells. As shown for other cell types, NS1 limits the production of type I IFN as well as that of cytokines and chemokines upon IAV infection ([Fernandez-Sesma et al., 2006](#); [Have et al., 2009](#)). Moreover, IAVs are capable of interfering with the maturation of DCs. Infected DCs show reduced upregulation of MHC class II and costimulatory receptors like CD80 and CD86, essential for T cell priming. Young and colleagues demonstrated that infection of mature DCs by influenza virus interferes with the presentation of an endogenous viral antigen and can thus impact the adaptive immune response ([Young et al., 2007](#)).

Рис. 2. Фрагмент оригинальной зарубежной статьи [10], послужившей источником для заимствований.

ЗАКЛЮЧЕНИЕ

Рост плагиата в научных публикациях оказался напрямую связан с развитием компьютерных и сетевых технологий, позволивших копировать информацию в больших объемах. В то же время системы автоматического индексирования текста позволили в короткие сроки создать эффективные методы выявления прямых заимствований, а также проводить достоверный анализ соотношения оригинального и заимствованного текстов. Для эффективного и быстрого выявления более сложного вида плагиата, использующего языковые переводы и чужие идеи, должны быть задействованы автоматизированные компьютерные сервисы, которые сличают модели цитирования определенной работы с моделями цитирования других работ, а затем направляют сгенерированные соответствующие запросы в библиографические базы данных. Исследования немецких коллег и приведенные в настоящей статье примеры позволяют сделать вывод, что такая задача решается. Разработка и промышленный запуск подобной системы могли бы способствовать, на наш взгляд, значительному снижению объемов заимствований и росту числа оригинальных отечественных исследований.

СПИСОК ЛИТЕРАТУРЫ

1. Антиплагиат.ру. – URL: <http://www.antiplagiat.ru/index.aspx> (дата обращения: 15.08.2013).
2. Gipp B., Meuschke N., Beel J. Comparative Evaluation of Text- and Citation-based Plagiarism Detection Approaches using GUTTENPLAG // In Proceedings of 11th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'11). Ottawa, Canada, Jun. 2011. – ACM NY, 2011. – P. 1–4.
3. Заимствования и плагиат. Накипь и ржавчина в науке // Троицкий вариант – Наука. – 2009. – № 44. – С. 10.
4. Gipp B., Meuschke N. Citation Pattern Matching Algorithms for Citation-based Plagiarism Detection: Greedy Citation Tiling, Citation Chunking and Longest Common Citation Sequence // In Proceedings of the 11th ACM symposium on Document engineering (DocEng '11). Mountain View, CA, USA, Sep. 2011. – P. 1–10.
5. Bailey J. Using Citations to Detect Plagiarism // In Plagiarism Today. – URL: <http://www.plagiarismtoday.com/2011/08/08/using-citations-to-detect-plagiarism/> (дата обращения 15.08.2013).
6. Simkin M.V., Roychowdhury V.P. Read before you cite! // Complex Systems. – 2003. – Vol. 14. – P. 262–274.
7. Simkin M.V., Roychowdhury V.P. Do you sincerely want to be cited? Or: Read before you cite // Significance. – 2006. – Vol. 3(4). – P. 179–181.
8. Gipp B., Meuschke N., Breitingner C., Lipinski M., Numberger A. Demonstration of Citation Pattern Analysis for Plagiarism Detection // In Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval. Dublin, UK, Jul. 28-Aug. 1 2013. – P. 1119–1120.
9. Meuschke N., Gipp B., Breitingner C. CitePlag: A Citation-based Plagiarism Detection System Prototype // In Proceedings of the 5th International Plagiarism Conference. Newcastle upon Tyne, UK, Jul. 2012. – P. 1–10.
10. Schmolke M., Garcia-Sastre A. Evasion of innate and adaptive immune responses by influenza A virus // Cell Microbiol. – 2010. – Vol. 12(7). – P. 873–880.

Материал поступил в редакцию 26.08.13.

Сведения об авторах

ГУРЕЕВ Вадим Николаевич – ведущий специалист, Государственный научный центр вирусологии и биотехнологии «Вектор», Новосибирская область, Кольцово.
E-mail: gureyev@vector.nsc.ru

МАЗОВ Николай Алексеевич – кандидат технических наук, заведующий информационно-библиотечным центром, Институт нефтегазовой геологии и геофизики им. академика А.А. Трофимука СО РАН, Новосибирск.
E-mail: MazovNA@ipgg.sbras.ru

Центр обработки цифровых данных в составе информационной системы академического института

Построение и развитие академической геоинформационной среды на основе Центров цифровой обработки данных – развивающееся перспективное относительно новое направление проектирования интегральных информационных систем. Показаны преимущества развития технологий высокопроизводительной обработки цифровых данных в Центре обработки, в том числе роль информационной инфраструктуры, концентрации и распределения данных с помощью геоинформационных технологий.

Отмечены новые возможности исследований при 3D цифровом моделировании геологических объектов и обработке данных дистанционного зондирования земной коры.

Центр обработки данных в интегральной информационной системе позволяет решать задачи формирования целостного интегрального проблемно-ориентированного информационного поля для пользователя.

Ключевые слова: интегральная информационная система, информационное пространство, центр обработки данных

ВВЕДЕНИЕ

Информационная система крупного академического института (далее Института) содержит сложную разветвленную систему сбора, хранения и обработки научной информации, включающую множество элементов, различных аспектов и сфер деятельности [1]. Интегральная информационная система лежит в основе функционирования Центра информационного и телекоммуникационного обеспечения фундаментальных исследований в области наук о Земле (геология рудных месторождений, петрография, минералогия, геохимия), проходящего в настоящее время апробацию в ИГЕМ РАН (Институт геологии рудных месторождений, петрографии, минералогии и геохимии Российской академии наук). Центр обработки данных – ядро современной информационной системы.

Целью нашего исследования является разработка принципов построения и развития геоинформационной среды интегральной информационной системы с входящими в её структуру Центром обработки данных, программными пакетами и информационным полем, а также в разработке схем взаимодействия отдельных элементов среды между собой и с внешними потоками данных, с учётом структурных особенностей этой среды [2].

В основе подхода, лежит метод стратифицированного математического моделирования [3], позволяющий выделять фундаментальные аспекты функционирования системы, определять её составные блоки, раскрывать структурные связи и последовательно переходить к рассмотрению других фрагментов структуры информационной интегрированной системы.

Центры обработки данных представляют собой помещения, в которых устанавливаются компьютеры, сетевое оборудование и системы хранения информа-

ции и создаются условия для их непрерывной и надёжной работы. Основу таких Центров составляют серверные вычислительные платформы. Среди обязательных условий нормального функционирования Центра обработки данных можно назвать наличие бесперебойного питания, соблюдение температурного режима и реализацию сетевых подключений в рамках самого Центра, а также подключений к сетям поставщиков телекоммуникационных услуг.

Для оперативного информационного обслуживания пользователей необходима современная информационная инфраструктура, под которой понимается организационная система, обеспечивающая функционирование и развитие информационного пространства Института и средств информационного взаимодействия: распределённая информационно-коммуникационная среда (программно-математическое обеспечение, научный и технический персонал) и центральный банк данных. Необходимым условием высококачественного информационного обслуживания является использование быстродействующих вычислительных ресурсов.

Информационная инфраструктура создается для обработки и накопления данных как средство обеспечения пользователей информацией, содержащейся в базах и банках данных. Банки данных включают ряд баз данных, классификаторов (лингвистическое обеспечение), цифровых протоколов, определяющих форму и структуру ресурсов.

ПРЕИМУЩЕСТВА ИНФОРМАЦИОННЫХ СИСТЕМ, ВКЛЮЧАЮЩИХ ЦЕНТРЫ ОБРАБОТКИ ЦИФРОВЫХ ДАННЫХ

Информационное пространство Института содержит значительный объем информационных ресурсов: научные данные (как фактографические, так и доку-

ментальные) и знания; результаты создания основ межсетевых связей; разработанные методы математического геомоделирования; информационные технологии и системы. Перечисленные ресурсы могут быть использованы для нужд фундаментальной и прикладной науки, высшей школы, производственных отраслей и органов государственной власти. Задача Центра обработки данных – концентрация и распределение информационных ресурсов.

Перед академическими институтами, выполняющими фундаментальные научные исследования, стоит задача развития технологий высокопроизводительной обработки данных, обусловленная следующими причинами:

- быстрый рост объемов информации вследствие развития прикладных исследований, информационных технологий, спутниковых мониторинговых систем и др., вызывает необходимость использования высокопроизводительных компьютеров и вычислительных сетей нового поколения для обработки увеличивающегося потока данных;

- информационные звенья «данные – информация – знания» и интеллектуальные системы (типа экспертных, интегрированных, глобальных с включением в мировую ГРИД-систему) требуют значительных вычислительных мощностей. Феномен замещения физической реальности виртуальной также является причиной возрастающего интереса к виртуальным информационным технологиям;

- компьютерная реализация сложных математических моделей часто осложняется ограниченной производительностью вычислительных ресурсов;

- работа на современном технологическом уровне требует соответствующего оборудования, в частности сетевой инфраструктуры с высокой пропускной способностью.

- фотореалистичные трехмерные графические модели, выполненные с помощью специализированных пакетов программ, повышают требования к Центру обработки данных.

Поступающие в интегральную информационную систему данные следует подготовить и проанализировать прежде, чем их можно будет использовать в процессе принятия решений. Подготовка данных в определенной мере заключается в извлечении знания – завершающей стадии исследования. Эти работы требуют высокопроизводительного оборудования и решаются с помощью Центра обработки цифровых данных [4].

Развитие интегральной информационной системы с Центром обработки данных, включающим блок серверов с банками данных и информационной инфраструктурой (рисунок), опирается на использование интегрированных информационно-телекоммуникационных систем и сетей, а также новых методик, технологий, технических (аппаратных) средств и аналитических методов обработки информации. Формирование и обновление многоконтурной распределенной геоинформационной среды, представляющей собой совокупность информационных, технологических, коммуникационных и организационных средств, служат фундаментом современной информационной системы Института.



Принципиальная схема построения ИИС с ЦОД

В процессе нашего исследования специальное внимание было уделено вопросам организации ввода данных в интегральную информационную систему и систематизации картографической информации с применением ГИС-технологий и Центра обработки данных с использованием материалов по территории Центральной Азии, Тихоокеанского рудного пояса и Камчатского края. Другое перспективное направление – цифровое моделирование геологических объектов с использованием данных дистанционного зондирования и метода 3D цифрового моделирования.

Построение информационного пространства с использованием Центра обработки данных обеспечивает свободный доступ к информационным ресурсам, базам метаданных и распределенному Интегральному банку данных, центральной частью которого является банк данных ИГЕМ РАН, а также к информационной инфраструктуре с единой средой на базе Веб и ГИС-технологий, единой терминологической системой и развитой телекоммуникационной инфраструктурой. Внедрение Центра обработки данных обеспечит не только доступ пользователя к информационным ресурсам Института, но и возможность использования дополнительных вычислительных мощностей и инструментальных средств, заложенных в систему.

При разработке центров обработки данных учитываются требования единого открытого стандарта представления информации о ресурсах (метаданных); стандарта, обеспечивающего управляемость потоками информации (обменный формат), а также требования по организации доступа к информационным ресурсам в научных центрах страны.

ЗАКЛЮЧЕНИЕ

Процесс создания Центра цифровой обработки данных предполагает несколько обязательных этапов: изучение научно-методических материалов по созданию Центра; формирование технического задания; создание технического проекта; заказ оборудования; монтаж инженерных систем; пуско-наладка и прием Центра в эксплуатацию.

Создание Центров цифровой обработки данных – процедура сложная и длительная. Стандартизация этого процесса снижает риски работы оборудования и исключает необходимость дополнительных вложений в устранение недочетов.

Для полноценного информационного обслуживания пользователей необходима разработка целостного интегрального проблемно-ориентированного информационного поля, которое состоит из совокупности инструментальных средств, аналитических методов, описаний и баз данных, необходимых для проведения прикладных и фундаментальных исследований. Информационное поле должно быть реализовано так, чтобы пользователь мог с ним работать,

имея только компьютер и связь с сетью Института и Интернетом. Распределенный банк включает аналитические ГИС-системы, инструментальные средства поиска и отображения информации, распределенную систему баз метаданных и хранилище данных, обеспечивающие картографическую основу для конкретных геоинформационных проектов с возможностями ее аналитической обработки. В интегральной информационной системе с входящим в её состав Центром обработки данных обеспечивается автоматизированное информационно-картографическое и 3D моделирование геологических объектов на основе ГИС технологий.

СПИСОК ЛИТЕРАТУРЫ

1. Веселовский А.В. Компьютеризация геологических научных исследований РАН. с. 202 – 210. Раздел 5.4 – В кн.: Компьютеризация в геологии и недропользовании (история, состояние и перспективы развития). Вып. 3. – М.: Геокарт, ГЕОС, 2012. – 362 с.
2. Электронная Земля: использование информационных ресурсов и современных технологий для повышения достоверности научного прогноза на основе моделирования решений в интегральных информационных полях / отв. ред. акад. Ю.М. Арский, акад. В.Е. Велихов, чл.-корр. А.Б. Жижченко, акад. Н.П. Лавров, акад. Г.И. Савин – М.: ВИНТИ РАН, 2009. – 478 с.
3. Володин К.И., Гульницкий Л.Л., Пожарский И.Ф. и др. Автоматизированная система научно-технической информации – разработка и эксплуатация. – М.: Финансы и статистика, 2004. – 191 с.
4. Гитис В.Г., Ермаков Б.В. Основы пространственно-временного прогнозирования в геоинформатике. – М.: ФИЗМАТЛИТ, 2004. – 256 с.

Материал поступил в редакцию 17.09.13.

Сведения об авторах

ВЕСЕЛОВСКИЙ Александр Владимирович – доктор технических наук, профессор, заведующий лабораторией геоинформатики Института геологии рудных месторождений, петрографии, минералогии и геохимии Российской академии наук (ИГЕМ РАН), Москва
e-mail: valv@igem.ru

ГАЛЬБЕРГ Татьяна Валентиновна – ведущий инженер ИГЕМ РАН
e-mail: galberg@igem.ru

ДОКУМЕНТАЛЬНЫЕ ИСТОЧНИКИ ИНФОРМАЦИИ

УДК 025.4 : 620.3

В.М. Ефременкова

Системы классификации по нанонауке и нанотехнологиям (Обзор)

Дан аналитический обзор отечественных и зарубежных классификаторов в области нанонауки и нанотехнологий. Рассматриваются различные подходы к построению классификационных схем с ориентацией как на конкретные прикладные задачи, так и на проблему в целом. В связи с этим выделены универсальные, научные и ведомственные классификаторы.

Ключевые слова: системы классификации, универсальные классификации, научные классификации, ведомственные классификации, классификационные таблицы, нанонаука, нанотехнологии

ВВЕДЕНИЕ

Нанотехнологии – символ начала XXI в. для многих направлений развития науки, техники, технологий. Для любого специалиста нанотехнологии – это технологии и методики, в основе которых лежат операции с атомами, молекулами, отдельными наночастицами с целью создания компонентов с размерами до 100 нанометров и их последующей интеграции в функционирующие приборы и устройства с принципиально новыми качествами [1].

Сформулированное в 1959 г. Р. Фейнманом предвидение о развитии в ближайшем будущем работ в области физики - «манипуляции и должном управлении вещами совсем маленького масштаба» – начало осуществляться уже в 80-90 гг. прошлого века [2]. Это взрыв исследований в области нанонауки, а затем и нанотехнологий. Начало XXI в. ознаменовалось открытием в области пико диапазона: 4 июля 2012 г. была обнаружена теоретически предсказанная в 1964 г. элементарная частица – бозон Хиггса.

Следует добавить, что нанонаука представляет совокупность исследований в области классической и квантовой физики и химии и является разделом физики «конденсированного состояния».

Нанонаука в современном понимании «выросла как результат фундаментальных открытий в современной физике, в первую очередь, связанных с представлениями *квантовой механики*, где проявляются, часто самым неожиданным образом, квантовые эффекты в макромасштабе» (С.П. Капица) [1].

Первые работы по объектам наномасштаба относятся к XVII-XIX вв. (магнитные стенки, магнитные домены, нитевидные кристаллы и др.), а расцвет исследований начался с появлением возможности увидеть атомы и манипулировать ими лишь в 1981 г. с

изобретением Э. Руска, Г. Биннингом и Г. Рорером (Нобелевская премия 1986 г.) электронных микроскопов [3].

И тогда же Г. Глейтер ввел в научный обиход термины: нанокристаллические, наноструктурные, нанофазные, нанокомпозитные и т. д. материалы [4].

В 1974 г. для обозначения процесса ультрапрецизионной обработки материалов Н. Танигути ввел термин «нанотехнология» [5]. В 1986 г. независимо от Танигути термин «нанотехнологии» предложил американский инженер Э. Дрекслер в своей книге «Машины созидания: Грядущая эра нанотехнологии» [6].

В сентябре 1989 г. Д. Эйглер стал первым, кому удалось контролировать передвижение отдельных атомов: он выложил из 35 отдельных атомов ксенона слово «IBM» в сканирующем туннельном микроскопе [7].

Бурное развитие работ в области нанонауки и нанотехнологий, с одной стороны, и нечеткость терминологии – с другой, заставила информационные службы ведущих баз данных (БД) мира разработать инструменты, облегчающие поиск информации. Одним из таких инструментов являются системы классификации.

В настоящей статье будут рассмотрены различные виды классификационных систем, предназначенных для большого круга специалистов в области нанонауки и нанотехнологий.

ОСОБЕННОСТИ ТЕРМИНОЛОГИИ В ОБЛАСТИ НАНОРАЗМЕРНЫХ ОБЪЕКТОВ

История наноразмерных объектов насчитывает несколько тысячелетий. Они существовали всегда – это вирусы, дым открытого огня, аэрозоли, коллоидные частицы серебра и золота.

«Нано» как термин или как приставка витал в умах людей с древних времен. В литературе или языкознании мы встречаемся с ним и в легендах, и в философских трактатах, а позднее – и в науке. «Атом» – максимально малый и невидимый человеку – по гречески «нанос».

Важно понимать, в какой системе координат рассматривается тот или иной объект (рисунок). Физики, уточняя определение **нанообъектов**, утверждают, что верхняя граница наночастка размерной шкалы совпадает с размером, означающим проявление квантовых эффектов. Мы живем в трехмерном мире. При уменьшении размера объекта ниже некоторой критической величины доля вклада поверхностных атомов в свойства этого объекта становится существенной и будет продолжать расти с дальнейшим уменьшением *высоты объекта*, наблюдается *квантово-размерный эффект в двух измерениях*. Такой объект называется *квантовой плоскостью*, или *квантовой ямой*. При дальнейшем уменьшении размеров в двух измерениях – это уже *одномерный квантовый объект*, *квантовая нить*, или *квантовый провод*. У *нуль-мерных объектов*, или *квантовых точек* размер объекта ниже некоторой критической величины имеет место во *всех трех измерениях*. Классифицировать такие объекты логичнее всего по степени снижения размерности: 2D — квантовые плоскости, 1D — квантовые нити, 0D — квантовые точки. Такая физическая модель позволяет более наглядно представить картину наномира и происходящих в нем процессов.

Терминология в области **наноматериалов** соотнесена со структурированием этих материалов по характерным особенностям и выполняемым функциям.

Первый подход к отбору терминов определяется геометрическими размерами структуры таких материалов. Материалы с размером микроструктуры от 1 до 100 нм называют наноструктурными (нанокристаллическими, супрамолекулярными).

Второй подход связан с огромной ролью многочисленных поверхностей раздела в наноматериалах в формировании их свойств.

Третий подход обусловлен совпадением размеров кристаллитов наноматериала с размерами, характерными для определенных физических явлений. Например, для магнитных свойств – размер однодоменного кристалла и т.д.

Весь спектр наноразмерностей можно легко понять и главное — экспериментально наблюдать на примере углеродных наночастиц. Для *биологических систем* характерна *системная организация материи на наноуровне*. Неудивительно, что как научный термин «нано» впервые появился в биологии в 1909 г. в Германии. Ханс Ломан, наблюдая микроскопические водоросли в оптическом микроскопе, назвал их «наннопланктоном». Тогда же начал издаваться журнал «Journal of nannoplankton Research».

Потребность в использовании новых наименований множителей единиц измерения связана с открытием рентгеновских лучей, развитием спектроскопии. Ученые вспомнили о приставке «нано». В 1958 г. на заседании Международного комитета мер и весов было принято предложение советского ученого Г. Д. Бурдуна именовать миллиардную долю метра (10^{-9} м) нанометром. Удвоение «н» осталось исторически в палеонтологии и частично в биологии. Спор был окончательно решен на конференции в Японии в 1974 г. в пользу термина с приставкой «нано».

Терминология по наноматериалам и нанотехнологиям в настоящее время только устанавливается. В мировой литературе четко отличают нанонауку (nanoscience) или наноразмерную науку (nanoscale science) от нанотехнологии (nanotechnology). Определяющими при отнесении объектов, явлений, методов исследования к нанонауке являются, как описано выше, приставка «нано-» (nano) или термин «квантовый» (quantum). Кроме того, необходимо принимать во внимание термины или словосочетания наномасштаба такие, как «ультрадисперсные системы», «фуллерены», «графены» «супрамолекулы» и др.

Мир низкоразмерных величин

10	– 10 ⁻² –	10 ⁻³	– 10 ⁻⁵ –	10 ⁻⁶ –	10 ⁻⁹	–	10 ⁻¹²	–	10 ⁻¹⁵ м
МАКРО		МИКРО		НАНО	ПИКО		ФЕМТО		
Человек		клетка		вирус	Атом		Протон		
		Световой		Электронный					
				Атомно-силовой					
Глаз									
				Микроскопы					

Шкала размеров, примерные величины различных объектов и методы их визуализации [8].

Приставка «нано-» оказалась очень практичной, стала брендом, позволяя разработчикам получать инвестиции в проекты. С помощью «**Нанограффити**» можно отмечать заслуги исследователей, их институтов или спонсоров, профинансировавших работу.

СИСТЕМЫ КЛАССИФИКАЦИИ В ОБЛАСТИ НАНОТЕХНОЛОГИЙ И НАНОМАТЕРИАЛОВ

Классификаторы в области нанотехнологий могут быть использованы для отбора и экспертизы проектов, оценки деятельности в области защиты прав интеллектуальной собственности, проведения статистических исследований, унификации научно-технической или иной информации в этой области. Одной из важных функций классификатора является возможность проведения аналитических и наукометрических исследований на его основе.

В области нанотехнологий разработан ряд классификаторов (рубрикаторов), в каждом из которых предложен свой подход к классификации проблем в этой области.

Среди классификаторов по нанотехнологиям, определяющих наиболее актуальные направления развития науки, техники, технологии и областей применения, можно выделить три группы: **универсальные** классификаторы, имеющие многоцелевой характер и предназначенные для классификации всех областей деятельности научных коллективов, включая ученых, разработчиков технологических циклов, оборудования и специалистов по реализации достижений ученых в различных отраслях; **научные**, отражающие научные проблемы развития рассматриваемого направления, и **ведомственные**, включающие только задачи, решаемые различными ведомствами, компаниями, институтами.

Универсальные классификаторы

Универсальный классификатор в области нанонауки и нанотехнологий должен быть многоцелевым и отражать широкий круг проблем науки, технологий и возможностей применения продукции в различных отраслях промышленности; тематическое содержание классификатора должно служить основой для разработки локальных классификаторов. Один из важных факторов успешного использования универсального классификатора – обеспечение доступа к единому portalу.

Рассмотрим следующие классификаторы:

- Единый научный рубрикатор ФЦП «Развитие инфраструктуры наноиндустрии в Российской Федерации на 2008-2010 годы».

- Рубрикатор «Нанотехнологии и наносистемы».
- Государственный рубрикатор научно-технической информации (ГРНТИ).
- Универсальную десятичную классификацию (УДК).

Научно-исследовательский институт информационных технологий и телекоммуникаций разработал **Единый научный рубрикатор ФЦП «Развитие инфраструктуры наноиндустрии в Российской Федерации на 2008-2010 годы»** [9] для федерального интернет-портала «Нанотехнологии и наноматериалы»,

предназначенный для хранения и обмена информацией участников национальной нанотехнологической сети. Иерархический шестиуровневый рубрикатор отражает потребности инновационных задач Специализированного информационно-библиографического ресурса (СИБР) и соответствует большинству заложённых в нем принципов. В рубрикации проведено четкое разделение проблем по трём аспектам:

- 1) физические свойства объекта деятельности (нанообъектов);
- 2) производство объекта деятельности;
- 3) конечная продукция отрасли с использованием объекта.

Таким образом, предложенная рубрификация отражает основные вопросы развития инновационной инфраструктуры наноиндустрии.

К существенным недостаткам этого рубрикатора следует отнести трудно воспринимаемую логику построения системы классификации, например, теоретические проблемы нанотехнологий в различных отраслях науки помещены в конце раздела «Свойства нанообъектов», а «Общий раздел», в котором должны быть сосредоточены вопросы экономики, рынка, наноинформатики, проблем развития нанотехнологий, наукометрии, информации о книгах, конференциях и т.д. находится в конце рубрикатора. В первом разделе, относящемся к структурам нанообъектов, в качестве оснований деления были выбраны модели объектов, а не сами объекты, что не корректно с точки зрения физики, так как и фуллерены, и порошки – это объемные структуры, а не точечные (нульмерные). Нечетко сформулировано содержание рубрик, относящихся к вопросам взаимодействия наночастиц с различными объектами [9].

Рассматриваемый рубрикатор стал первой попыткой создания универсального классификатора, что отмечалось в дискуссии на круглом столе в МГУ им. М.В. Ломоносова (декабрь 2009 г.) [10].

Для задач автоматизированного классифицирования огромных объемов научных публикаций и поиска информации специалистами ВИНТИ РАН был разработан принципиально новый иерархический трехуровневый рубрикатор «Нанотехнологии и наносистемы», структура и логика построения которого позволяют последовательно раскрыть содержание основных понятий в области нанонауки и нанотехнологий [11,12].

Основное содержание рубрикатора «Нанотехнологии и наносистемы» определено рубриками первого уровня:

1. Общие вопросы
2. Теоретические модели в сфере нанотехнологий
3. Объекты и материалы, относящиеся к сфере нанотехнологий
4. Кристаллическая структура, свойства наноматериалов и нанообъектов
5. Процессы, явления и эффекты в наноструктурированных материалах
6. Методы получения наноматериалов и нанообъектов
7. Методы исследования, измерения и сертификации наноматериалов и нанообъектов
8. Основные материалы нанотехнологий

9. Нанотехнологии в различных областях. Производство нанотехнологий
10. Нанотехнологии в биологии
11. Нанотехнологии в медицине
12. Нанотехнологии в военно-промышленном комплексе.

Логика построения и структура классификационной таблицы Рубрикатора «Нанотехнологии и наносистемы».

Классификатор построен по классическому принципу классификационных систем, заключающемуся в последовательности развития содержания рассматриваемой дисциплины: от общих вопросов к теоретическим моделям в области нанонауки и нанотехнологий и далее – к специальным проблемам, связанным со спецификой изучаемых объектов, их структурой, свойствами, происходящими в них процессами и явлениями, методами получения, исследования, измерения и

сертификации и с переходом к задачам материаловедения и применения в различных областях знания и отраслях экономики.

Тематика основной, ключевой рубрики рубрикатора «Нанотехнологии и наносистемы» (код **3**) – «Объекты и материалы, относящиеся к сфере нанотехнологий» сформировалась под влиянием анализа работ в области нано- и квантовых структур, но включает и объекты классической физики, и химии «конденсированного состояния вещества» типа: молекулярные кластеры, магнитные точки, тонкие пленки, нитевидные кристаллы, дендриты, гетероструктуры; биологические объекты: белки, липиды и др. Перечень наименований подразбук второго и третьего уровней этой рубрики содержит основные, устоявшиеся термины и словосочетания. Приведем фрагмент нанорубрикатора, включающего код **3** – «Объекты и материалы, относящиеся к сфере нанотехнологий» (табл. 1).

Таблица 1

Фрагмент Рубрикатора «Нанотехнологии и наносистемы» (код 3) – «Объекты и материалы, относящиеся к сфере нанотехнологий»

Код рубрики	Наименование рубрики на русском языке	Наименование рубрики на английском языке	Код рубрики ГРНТИ	Ключевые слова на русском языке	Уточняющие термины на русском языке	Ключевые слова на английском языке	Уточняющие термины на английском языке
3.3	Квантовые структуры	Quantum structures	29.19.22 29.05.15 47.09.48	квантовые структуры		quantum structures	
				квантово-размерные структуры		quantum-dimensional structures	
3.3.1	Квантовый ящик	Quantum box	29.19.22	квантовый ящик		quantum box	
3.3.2	Квантовые точки. Квантовые точечные контакты	Quantum dots. Quantum dot contacts	29.19.22	квантовые точки		quantum dots	
						quantum points	
					графеновые квантовые точки		graphene quantum dots
				квантовые точечные контакты		quantum point contacts	
3.3.3	Квантовые кольца	Quantum rings	29.19.22	квантовые кольца		quantum rings	
				нанокольца		nanorings	
3.3.4	Квантовые ямы	Quantum wells	29.19.22	квантовые ямы		quantum wells	
				наноямы		nanowells	

Терминология этого нового направления до сих пор представляет нечеткое множество слов и словосочетаний, что нашло отражение во введении на третьем уровне рубрикатора подрубрики «Другие наноструктуры», в которую включены нанообъекты типа: нанодиски, наноспираль, ноноконусы, наноканалы и др., имеющие по сравнению с терминами вышестоящих рубрик более низкую встречаемость в информационных потоках.

Введение рубрики «Кристаллическая структура, свойства наноматериалов и нанообъектов» (код 4) обусловлено тем фактом, что все свойства нанообъектов связаны с их кристаллической структурой, состоянием и дефектами. Поэтому на втором подуровне этой рубрики вначале рассматриваются вопросы «Структурной кристаллографии» (код 4.1) и только затем все свойства – от механических до оптических, включая влияние последних на свойства нанообъектов и наноматериалов (коды 4.2 – 4.13). Структура рубрики «Процессы, явления и эффекты в наноструктурированных материалах» (код 5) определялась особенностями поведения наноматериалов в каждом из трех подразделов этой рубрики.

В основу построения рубрик по методам получения, исследования и измерения нанообъектов и наноматериалов (коды 6 и 7) были положены принципы деления по областям знания: физические, химические и механические методы, которые образуют стройную систему отношений между отдельными методами получения и измерения нанообъектов. Наиболее сложным для классификации является раздел, связанный с отражением всего круга вопросов, относящихся к веществам – «Основные материалы нанотехнологий» (код 8). Это технические материалы, используемые в нанотехнологиях для создания приборов и устройств; высокочистые вещества; функциональные и конструкционные материалы для энергетики, космической техники и др.

Рубрика «Нанотехнологии в различных областях. Продукция нанотехнологий» – (код 9) отражает все области знания и отрасли промышленности, в которых используются нанотехнологии, и построена вначале по областям знания, затем по отраслям экономики. Все наиболее перспективные сферы использования нанотехнологий, например, в системах безопасности, отражены в следующих подрубниках:

9.1 Прикладная механика и машиностроение. Используя при поиске подрубрики 2-го или 3-го уровней и/или ключевые слова, можно получить информацию о механических сенсорах и датчиках, сенсорах деформации, сенсорах вибрации, волоконно-оптических сенсорах вибраций и термодатчиках

9.3.1 Химические сенсоры – различные типы химических сенсоров, включая многофункциональные сенсоры типа «электронный нос»

9.4.6 Оптоэлектроника, нанофотоника и наноплазмоника, КМОП оптические сенсоры

9.10.2 Теплоэнергетика: тепловые сенсоры

9.17.3 Нанобиосенсоры.

Рубрика «Нанотехнологии в военно-промышленном комплексе» (код 12) наименее тематически развита ввиду специфики тематики. Основное внимание

уделено аспектам использования нанотехнологий биологического или химического направлений и материаловедения.

В рубрикаторе «Нанотехнологии и наносистемы» большое внимание уделено вопросам биологической тематики (коды 10, 11) – применению достижений бионанонауки и нанобиотехнологии в различных областях медицины: проблемам терапии, системам доставки лекарств, биоматериалам, используемым для профилактики и лечения практически во всех направлениях медицины.

При построении рубрикатора и составлении перечня ключевых слов и уточняющих терминов на русском и на английском языках были также учтены все возможные синонимы и антонимы, однокоренные и разнокоренные слова, относящиеся к одному понятию. Кроме того, принималась во внимание частота употребления того или иного термина в информационных потоках. Частотность определялась путем анализа результатов последовательных поисков в БД ВИНТИ РАН и в БД SCOPUS (издательство Elsevier, Нидерланды), в поисковые предписания которых включались анализируемые термины. Такой непосредственный контроль терминов по результатам информационных запросов в итоге приводит к постоянному дополнению и модификации рубрикатора «Нанотехнологии и наносистемы» в процессе его эксплуатации.

Практика работы над построением рубрикатора «Нанотехнологии и наносистемы» и его непосредственного использования при сортировке и поиске документов привела к пониманию того, что необходимо ориентироваться не столько на соблюдение единства лингвистических средств с другими рубрикаторами по «нано», сколько на альтернативный подход – разработку таблиц соответствия рубрикаторов. Любой разрабатываемый рубрикатор должен быть сопоставлен с рубрикаторами, которые используются уже сейчас в больших государственных информационных системах и в отдельных профильных организациях, а также с зарубежными классификаторами. Другими словами, речь должна идти о системном подходе – сколько бы ни было рубрикаторов и классификаторов в данной области науки, они должны быть совместимы между собой, т.е. представлять собой систему согласованных средств поиска информации, отражающую структуру предметной области.

Государственный рубрикатор научно-технической информации (ГРНТИ) – универсальная иерархическая классификация областей знания, принятая для систематизации потока научно-технической информации [13]. ГРНТИ является частью общесистемных средств лингвистического обеспечения государственных органов и сетей научно-технической информации России наряду с комплексом информационно-поисковых тезаурусов, номенклатурой грамматических средств информационно-поисковых языков, правилами представления данных в коммуникативных форматах, методиками индексирования документов и запросов для обмена информационными материалами в режиме сети.

В области нанонауки и нанотехнологий в ГРНТИ предусмотрены рубрики:

29.19.03 Теория конденсированного состояния

29.19.16 Физика тонких пленок. Поверхности и границы раздела

29.19.22 Физика наноструктур. Низкоразмерные структуры. Мезоскопические структуры

29.19.24 Электронная структура твердых тел

29.29 Физика атома и молекулы

29.35.37 Электронная и ионная эмиссия

29.35.43 Электронная и ионная микроскопия

45.09.39 Углеродные материалы, включая наноструктурированные

47.09.48 Наноматериалы для электроники

47.13.07 Технология и оборудование для производства приборов и устройств нанoeлектроники.

При сопоставлении рубрик рубрикатора «Нанотехнологии и наносистемы» с рубриками ГРНТИ решение о тематическом соответствии принималось на основании максимального точного совпадения содержания понятий, выраженных рубриками независимо от их положения в иерархической структуре рубрикаторов. Рубрике верхнего уровня одного рубрикатора могла быть сопоставлена рубрика нижнего уровня другого рубрикатора

Однако в ряде случаев соответствия рубрик на верхних уровнях найти не удалось в силу отсутствия в рубрикаторе ГРНТИ информации о новых ветвях развития науки, техники и технологий. Не всегда понятия сопоставленных рубрик совпадают достаточно точно. В ряде случаев определенная тематика нанотехнологий может быть отнесена к разным рубрикам ГРНТИ, где она присутствует в контексте разных аспектов рассмотрения, тогда указываются две и более рубрики ГРНТИ. Наблюдаются и противоположные случаи: широкие рубрики ГРНТИ могут быть указаны как соответствия к разным нанотехнологическим рубрикам.

Универсальная десятичная классификация (УДК). В 2006 г. в УДК появился код, относящийся к нанотехнологиям – «Нанотехнологии» – код 620.3.

Однако, как отмечалось выше, в различных разделах УДК можно найти информацию о процессах, явлениях, материалах в сфере нанотехнологий, добавив для уточнения наномасштабов общий определитель свойств – **022.513.2**. Основные коды в разделах математики, где содержится информация по теоретическим задачам, решаемым математическими методами, код **519**; в астрономии – коды **52** и **53**, например, «плазменные кристаллы в пылевой плазме» межзвездного пространства; в химии – коды **544, 546, 548** [14].

Научные классификаторы

Для структурирования, хранения и обмена информацией между учеными и научными коллективами в узкотематической области знания создаются научные классификаторы, отражающие только научные направления исследований в области нанонауки.

В настоящее время к научным классификаторам относят «Научный нанорубрикатор (Н)», разработанный ГК «Роснанотех» [15], Рубрикатор информационных изданий ВИНТИ РАН [16] и «Нанорубрикатор» (научный классификатор) [17].

«Научный нанорубрикатор (Н)» – трехуровневый иерархический классификатор, содержит три рубрики первого уровня:

- объекты, относящиеся к сфере нанотехнологий
- получение, диагностика и сертификация наноразмерных систем
- продукты нанотехнологий [15].

В ВИНТИ РАН для индексации информационных продуктов БД/РЖ по нанонауке и нанотехнологии используются три рубрикатора 3 – 7 уровней:

Рубрикатор БД «Физика нанообъектов и нанотехнологии» – код рубрики 3-го уровня 291.19.22 с подрубриками 4-го уровня, отражающими тематическое содержание: свойства, структура, методы получения и диагностика нанообъектов [16].

Рубрикатор БД «Спинтроника» – код рубрики 3-го уровня 291.19.36 с подрубриками 4-го уровня, отражающими тематическое содержание: материалы и технология их получения, основные эффекты, приборы спинтроники [16], в рубрикаторе БД «Биотехнология. Нанобиотехнологии. Нанобиоматериалы», рассматриваются проблемы нанобиологии и нанобиотехнологии [16].

Ведомственные классификаторы

Конкретные прикладные нужды отдельных узкотематических направлений технологии, техники, промышленности, экономики, менеджмента и др. находят отражение в ведомственных классификаторах:

Классификатор Международной организации по стандартизации

Классификатор РФФИ

Международная патентная классификация

Классификатор нанотехнологий «8+»

Классификатор Международной организации по стандартизации – International Organization for Standardization (ISO) [18] предусматривает на верхнем уровне семь направлений; в классификациях статистических служб Канады и Австралии их, соответственно, девять и четырнадцать. Варианты, предложенные Европейской патентной организацией – European Patent Office (EPO) и Центром исследований нанотехнологий Японии (NRNC), включают по шесть направлений. В России ключевым документом в сфере нанотехнологий, является Федеральная целевая программа (ФЦП) «Развитие инфраструктуры nanoиндустрии в Российской Федерации на 2008–2013 годы» [19]. Эти классификации предусматривают девять позиций, пять из которых можно объединить в категорию наноматериалов. К числу обязательных направлений для всех рассматриваемых классификаторов относятся также наноэлектроника, нанобиотехнологии и наномедицина (табл. 2).

Примеры группировок основных направлений нанотехнологий

№ п/п	Статистическая служба Канады	ЕРО	ISO	NRNC	Статистическая служба Австралии	ФЦП «Развитие инфраструктуры наноиндустрии в Российской Федерации на 2009-2013 гг.»
1	Нанопотоника	Нанобиотехнологии	Нанобиотехнологии	Электроника	Нанотехнологии для окружающей среды	Нанoeлектроника
2	Нанoeлектроника	Нанотехнологии для обработки, хранения и передачи информации	Нанoeлектроника	Оптоэлектроника	Молекулярная и органическая электроника	Наноинженерия
3	Нанобиотехнологии	Нанотехнологии для материалов и науки о земле	Наномедицина	Медицина и биотехнологии	Нанобиотехнологии	Функциональные наноматериалы и высокочистые вещества
4	Наномедицина	Нанотехнологии для распознавания, взаимодействия и манипулирования	Нанометрология	Измерение и производство	Нанoeлектромеханические системы	Функциональные наноматериалы для энергетики
5	Наноматериалы	Нанооптика	Нанооптика	Охрана окружающей среды и энергетика	Нанoeлектроника	Функциональные наноматериалы для космической техники
6	Квантовые вычисления	Наномагнетизм	Нанопотоника	Наноматериалы	Выращивание, сборка и производство наноструктур	Нанобиотехнологии
7	Самосборка		Нанотоксикология		Производство нанопроductии	Конструкционные наноматериалы
8	Инструменты				Наноматериалы	Композитные наноматериалы
9	Прочее				Наномедицина	Нанотехнологии для систем безопасности
10					Нанометрология	
11					Нанопотоника	
12					Нанодиагностика	
13					Нанотоксикология, здоровье и безопасность	
14					Прочее	

Классификатор РФФИ предназначен для систематизации заявок, отчетов и других материалов, связанных с выделением грантов на проведение инициативных научных работ по линии Российского фонда фундаментальных научных исследований (РФФИ) [20]. Тематика наноисследований проиндексирована кодами практически всех представленных в классификаторе разделов, но основные коды классификатора относятся к области физики 02 (кроме 02.01 –

ядерная физика – это пикто диапазон), наиболее точно раздел 02.2 – физика конденсированных сред. Для химии и наук о материалах предусмотрен код: 03-450 – Наноструктуры и кластеры. Супрамолекулярная химия. Коллоидные системы и др. Проблемы формирования новых материалов в области биологии и медицинских наук – разделы 04.2 Физико-химическая биология; в разделе 07 – в областях инфокоммуникационных технологий и вычислитель-

ных систем предусмотрены коды: 07.7.30 – Микро- и нанoeлектромеханические устройства; 07.7.40 – Перспективные технологические процессы микро- и нанoeлектроники; для раздела 08 Фундаментальные основы инженерных наук: 08.1.06 – Проблемы механики в проектировании новых материалов, 08.2.06 – Нано- и мембранные технологии.

В настоящее время рассматриваются две системы **патентных классификаций** в области нанотехнологий: патентная классификация США и Международная патентная классификация (МПК) [21, 22].

Верхний уровень таблицы **классификации США**: класс 977 – Нанотехнологии с подклассами: 700 Наноструктуры; 839 Математические алгоритмы особым образом предназначенные для моделирования конфигурации или свойств наноструктур; 840 Производство, обработка или обнаружение наноструктур; 902 Специфическое применение наноструктур.

В **международной патентной классификации** (МПК) нанотехнологии выделены в отдельный класс B82B – Наноструктуры; их изготовление или обработка.

Наиболее репрезентативна **Европейская патентная база** [23], классификатор которой включает следующие разделы: Y01N2/00 – Нанобиотехнологии; Y01N4/00 – Нанотехнологии для передачи и хранения информации; Y01N6/00 – Нанотехнологии для науки о веществе и поверхности; Y01N8/00 – Нанотехнологии для измерения свойств; Y01N10/00 – Нанооптика; Y01N12/00 – Наномагнетики.

Структурирование процесса глобального технико-экономического развития в области нанотехнологий было проведено компанией «Science Global Management», разработавшей **Классификатор нанотехнологий «8+»** [24] – это бизнес-ориентированный классификатор технологий, в котором выбраны 8 направлений с высокой перспективой применения продуктов нанотехнологий:

- 1) твердотельные устройства на основе поверхностных и многослойных структур с заданным электронным спектром
- 2) фотонные кристаллы
- 3) управляемый рост нанофрактальных систем: селективных катализаторов, спецкоксов и т.д.
- 4) функциональные нанокомпозиты, в том числе на основе полимеров и на основе фуллереноподобных структур
- 5) устройства молекулярной электроники
- 6) устройства микроэлектромеханики
- 7) медицинские нанотехнологии
- 8) нанодисперсные порошки.

Кроме того, имеются проекты классификации научных и образовательных специальностей в сфере нанотехнологий.

ЗАКЛЮЧЕНИЕ

Цели всех разработанных в настоящее время классификаторов в области нанонауки и нанотехнологий – структуризация информации, обеспечение поиска, хранения, обработки и демонстрации достоверных аналитических и статистических данных. В каждом из рассмотренных ведомственных или науч-

ных классификаторов предложен свой подход к классификации в наноразмерной области с ориентацией на прикладные нужды каждого конкретного направления нанонауки или нанотехнологии. У каждого классификатора – своя специфика, но каждый из них наиболее хорош именно в той узкой области, для которой он предназначен. Роль универсального классификатора в области нанонауки и нанотехнологий – отражать широкий круг проблем технологий и возможностей применения продукции в рассматриваемой и других отраслях науки и промышленности и служить основой для разработки локальных классификаторов.

Если классификаторов несколько и они не связаны между собой, то никакие ведомства или структуры, уполномоченные правительством в качестве ответственных за развитие данного направления, не смогут получать в режиме реального времени оперативную и достоверную информацию о состоянии дел и перспективах в области нанотехнологий в России и в мире.

СПИСОК ЛИТЕРАТУРЫ

1. Нанонаука и нанотехнологии. Энциклопедия систем жизнеобеспечения. – М.: Изд-во ЮНЕСКО, Изд-во EOLSS: МАГИСТЕР-ПРЕСС, 2009. – 992 с.
2. Feynman R.P. There's plenty of room at the bottom // Engineering and Science. (Calif. Inst. Technol.). – 1960. – Vol. 23, № 2. – P.22-29. – URL: <http://www.zyvex.com/nanotech/feynman.html>
3. Руска Э., Биннинг Г., Рорер Г. Нобелевская премия по физике за 1986 год. – URL: www.nobelprize.org/nobel_prizes/physics/laureates/1986
4. Gleiter H. Deformation of Polycrystals // Proc. of 2nd RISO Symposium on Metallurgy and Materials Science / eds. N. Hansen, T. Leffers, H. Lithold. – Roskilde, RISO Nat. Lab., 1981. – P. 15-21.
5. Taniguchi N. On the Basic Concept of «Nano-Technology» // Proc. Intl. Conf. Prod. Eng.. Part II. — Japan Society of Precision Engineering, 1974.
6. Drexler E. Engines of creation. The coming of nanotechnology. – URL: e-drexler.com.
7. Эйглер Д. Мы построим наномир. – URL: <http://www.almaden.ibm.com>; URL: http://elementy.ru/images/elpub/nanotech_nanoscience
8. Пиотровский Л.Б., Кац Е.А. Нанотехнология, нанонаука и нанообъекты: что значит «нано»? – URL: <http://elementy.ru/lib/431265>
9. Единый научный рубрикатор ФЦП «Развитие инфраструктуры наноиндустрии в Российской Федерации на 2008-2010 годы». – URL: <http://www.portalnano.ru/rubricator/>
10. Рубрикаторы по нанотехнологиям – ФГАУ ГНИИ ИТТ "Информика" – URL: www.informika.ru/events/main/365
11. Кузнецов А.Ю., Борисова Л.Ф., Ефременкова В.М., Кириллова О.В., Проница Т.А. Разработка рубрикатора для «Специали-

- зирования информационно-библиографического ресурса» (СИБР) в области нанотехнологий // Российские нанотехнологии. – 2011. – Т. 6, № 5-6. – С. 160-173
12. Рубрикатор «Нанотехнологии и наносистемы» «нано». – URL: <http://rubric.neicon.ru/>
13. Государственный рубрикатор научно-технической информации. – М.: ВИНТИ, 2001. – 391 с.
14. УДК. Универсальная десятичная классификация – URL: <http://forum.udcc.ru/>
15. Научный нанорубрикатор (Н). – URL: www.RUSnano.com
16. Рубрикатор информационных изданий ВИНТИ РАН. – URL: <http://scs.viniti.ru/rubtree/main.aspx?tree>
17. Нанорубрикатор (научный классификатор) – URL: <http://expertcorps.ru/static/cms/Nanorubricator>
18. ISO Business plan ISO/TC 229. Nanotechnologies. Draft. 23.04.2007 – URL: www.nanorf.ru/Attachment.aspx?Id=5818
19. Федеральная целевая программа «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2013 годы». – URL: www.programs-gov.ru
20. Классификатор РФФИ. – URL: <http://www.rfbr.ru>
21. Обоснование классификации отраслей наук в области нанотехнологий. Международная классификация изобретений. Восьмая редакция. – М.: ФИПС, 2010. – URL: http://www.nanometer.ru/2008/12/16/nanotehnologii_i_nanomateriali
22. Патентный классификатор по нанотехнологиям США. – URL: <http://nanorf.ru>
23. EPO - Espacenet - European Patent Office. – URL: www.epo.org/searching/free/espacenet.html
24. Глазьев С.Ю. Нанотехнологии как ключевой фактор нового технологического развития. – URL: www.glazev.ru

Материал поступил в редакцию 28.06.13

Сведения об авторе

ЕФРЕМЕНКОВА Валентина Макаровна – кандидат физико-математических наук, научный сотрудник ВИНТИ РАН,
E-mail: efrem@viniti.ru

СПРАВОЧНО-ИНФОРМАЦИОННЫЙ РАЗДЕЛ

УДК 005.745(470):[5:3]

В.В. Арутюнов

О всероссийской научно-практической конференции «Математика, информатика, естествознание в экономике и в обществе»

Рассматриваются итоги работы состоявшейся в конце 2012 г. конференции в Московском финансово-юридическом университете (МФЮА), на которой было представлено 60 докладов и функционировало семь секций: математики; информатики и информационных технологий; физики, техники и электроники; биологии и экологии; математического моделирования экономических и социальных процессов; методологии учебного процесса; социально-экономических проблем. Приводится краткий обзор основных докладов на секции информатики и информационных технологий, а также на секции методологии учебного процесса.

Ключевые слова: информационные технологии, математика, информатика, физика, биология, социальные процессы, математическое моделирование, экономические процессы, экология, методология учебного процесса, информационная безопасность, социально-экономические проблемы

30 ноября 2012 г. в Москве в МФЮА (ныне Московском финансово-юридическом университете) проводилась конференция по актуальным проблемам математики, информатики и естествознания (МИЕСЭКО). В ней приняли участие более 100 специалистов из 48 организаций.

На конференции, где было представлено 60 докладов, функционировало семь секций: математики; информатики и информационных технологий; физики, техники и электроники; биологии и экологии; математического моделирования экономических и социальных процессов; методологии учебного процесса; социально-экономических проблем.

Приведем краткий обзор основных докладов, представленных на секции информатики и информационных технологий, а также на секции методологии учебного процесса.

В докладе А.Н. Алфимцева, Д.А. Локтева, А.А. Локтева «Реализация пользовательского интерфейса системы видеомониторинга на основе онтологического подхода» рассматривалась задача проектирования и реализации пользовательского интерфейса для управления мультikomпонентной системой видеомониторинга, в которую входят модули получения качественного изображения, система управления базами данных, система управления внешним освещением, система визуального распознавания образов, система интерфейсов для разработчиков, пользователей и взаимодействующих

программ, система создания сетевых шаблонов, которые могут масштабироваться в зависимости от структуры сети, система адаптации разрабатываемого комплекса для нужд организаций и задач различных предметных областей. Разработанные авторами программные модули учитывают возможность влияния оператора-пользователя на эксплуатацию комплекса видеомониторинга на различных этапах, а также возможность встраивания системы безопасности предприятия или отдельно взятого объекта в более крупную систему обеспечения безопасности на транспорте, объектах инфраструктуры, зданиях и сооружениях.

Доклад Г.М. Антоновой «Имитационная модель сети связи на транспортном уровне передачи» посвящен описанию имитационной модели, позволяющей учесть динамический характер сети связи для повышения степени адекватности модели процессам, происходящим при передаче сообщений в реальной среде на транспортном уровне протоколов передачи данных. Приводится алгоритм моделирования и результаты моделирования фрагмента сети. Результаты моделирования показывают, что предлагаемый алгоритм позволяет организовать взаимодействие имитационных моделей и процедур подготовки данных о переменной структуре фрагмента сети, представленного ориентированным графом, а также отразить динамические изменения в структуре ветвей и узлов графа.

Расширение возможностей предлагаемого метода моделирования динамических свойств сети связи можно выполнить, если сделать изменения модельного времени передачи информации и изменения структуры фрагмента сети связи независимыми, т.е. подчинить закон изменения элементов фрагмента сети внешним обстоятельствам, как это и происходит в реальных системах.

В докладе В.В. Арутюнова **«О современных мобильных компьютерных платформах»** анализировались возможности около десятка современных операционных систем для смартфонов [Symbian OS; Android; iPhoneOS (iOS); Windows Mobile, Windows CE и Windows Phone 7; BlackBerry OS и др.]; рассматривались и сравнивались их технико-экономические показатели, особенности, преимущества и недостатки; приводится также краткая информация о фирмах-производителях. Отмечается, что современные мобильные компьютерные платформы в основном базируются на одной из четырех ОС: Symbian, Windows Mobile, Android, iPhoneOS (iOS). В заключение констатируется, что в 2012 г. российские и зарубежные специалисты значительно продвинулись в сфере создания специального сорта стекла (который возможно использовать в том числе для смартфонов), практически не загрязняющегося, не имеющего бликов и на котором невозможно нанести царапины. Значимость использования такого стекла для экрана современных мобильных компьютерных платформ трудно переоценить.

В докладе А.К. Соловьева, Д.А. Карасева, М.А. Коржова **«Проблемы информационного обеспечения принятия решений по управлению государственными пенсионными обязательствами»** отмечалось, что управлять пенсионными обязательствами государство должно через формирование пенсионных прав застрахованных лиц. Для этого необходима система показателей и адекватная ей система учета и отчетности; также должна быть разработана отчетность по формированию пенсионных прав и государства, и местных органов, и работодателей – всех сторон пенсионной системы. По мнению авторов, в ней должны отражаться:

- суммы текущих начисленных взносов и численность застрахованных лиц, за которых они вносились в отчетном году;
- суммы страховых взносов, уплаченные по начислениям предыдущих лет, и численность застрахованных лиц, за которых они вносятся;
- суммы расчетного пенсионного капитала и пенсионных накоплений лиц, за которых взносы уплачивались по начислениям отчетного года;
- суммы расчетного пенсионного капитала и пенсионных накоплений лиц, которые не осуществляли трудовую деятельность и не вносили взносы в отчетном году (т.е. тех лиц, у которых права формируются исключительно за счет индексации/инвестирования средств).

Доклад А.Г. Юферова **«Кодификация знаний и учебный процесс»** посвящен рассмотрению некоторых направлений в этой сфере инженерии знаний, которые могут стать инструментальным средством в

процедурах кодификации, в том числе применительно к учебному процессу; обозначается новая проблемная точка роста - потребность в дальнейшем развитие теории информации не только как средства оптимального кодирования сигнала для передачи данных, но именно как аппарата для описания преобразования "информационных" структур, которые используются или могут использоваться для представления знаний. Отмечается необходимость разработки специфических средств учета и «оптимизации» психологических аспектов восприятия обучаемыми информации в системах дистанционного обучения, гарантирующих овладение навыками теоретической и практической деятельности в соответствующей специальности.

В докладе М.А. Николаева **«Цели и задачи создания веб-комплекса теплофизической оптимизации ядерных энергетических установок»** анализировались теоретические и практические вопросы создания вычислительного веб-комплекса (отраслевого или корпоративного сайта), предназначенного для решения задач теплофизической оптимизации ядерных энергетических установок (ЯЭУ). Теоретический аспект рассматриваемой задачи заключается в разработке унифицированной архитектуры и технологии (программных шаблонов) накопления и интеграции различных вычислительных кодов как веб-приложений и веб-сервисов, обеспечивающих удаленный доступ к вычислительным ресурсам; инфологическом анализе структуры данных и алгоритмов в задачах теплофизической оптимизации ЯЭУ для проектирования соответствующих унифицированных баз данных и программных шаблонов; разработке и реализации посредством указанной технологии ряда методик многопараметрической оптимизации ЯЭУ.

Практическая цель разработки состоит в организационно-технологическом усовершенствовании процедур теплофизической оптимизации ЯЭУ, обеспечивающем в том числе организацию вариантных проектных расчетов с сохранением исходных данных и результатов в базе данных; удаленном доступе специалистов к вычислительным ресурсам, методическим материалам и стандартам; дистанционном обучении методикам теплофизической оптимизации ЯЭУ и ряде других аспектов.

Процессы сохранения в различных форматах (TIFF, JPEG, GIF, BMP, PSD) растровых изображений трех видов (фотографии, рисунки векторного происхождения, коллажи из нескольких слоев) рассмотрены в докладе А.М. Подорожного **«Сохранение растровой графики различного типа в пространственных форматах»**. Размеры файлов сопоставлены с расчетом растровых массивов. Это позволяет наглядно проиллюстрировать особенности сохранения графики в различных форматах и сделать ряд практических выводов, включая следующие: снижение качества при экранном просмотре наблюдается только для форматов GIF и JPEG, при этом для формата JPEG можно подобрать режим, в котором снижение качества визуально незаметно; в коллажах каждый слой растрового изображения увеличивает размер файла, но не пропорционально числу

сложений, а значительно меньше; для программ обработки растровой графики (в частности, для изучения процессов трассировки растровой графики) лучше всего подходит формат BMP и др.

В докладе И.С. Романовой «**Анализ посещаемости образовательных математических сайтов сети Интернет**» исследуется посещаемость студентами образовательных математических сайтов. В качестве методов исследования использовались опрос, анкетирование студентов и статистическая обработка полученной информации. На основе полученного уравнения регрессии автор делает следующий вывод: посещаемость специального математического образовательного сайта www.exponenta.ru зависит от периода обучения (обычное и каникулярное) и практически не зависит от месяца в периоде обучения. Это объясняется стабильностью аудитории сайта (преподаватели, заинтересованные математикой студенты и др.), слабой (по мнению автора) «раскруткой» сайта и недостаточным вниманием к изучению дисциплины «математика» в школьных курсах.

Математическая модель, описывающая динамику популярности других образовательных веб-сайтов, показывает, что посещаемость сайта изменяется резкими случайными всплесками. Некоторые сайты показывают стабильный рост посещаемости, но большая часть сайтов сохраняет небольшую посещаемость; при этом отмечается то её рост, то падение. Большой рост посещаемости наблюдается во время сессий в учебных заведениях и при сдачах контрольных семестровых работ.

В докладе А.П. Титова «**Модель сети связи E-GOVERNANCE: проблемы и решения**» проанализирован способ организации государственной власти с помощью систем локальных информационных сетей и сегментов глобальной информационной сети. Используя данные Федеральной службы информации и Единого государственного реестра, где на конец 2011 г. зарегистрировано ~ 4,2 млн юридических лиц, автор смог оценить объем передаваемой информации. Разработаны рекомендации по проектированию сети электронного государства; в частности, сибирскую часть России и крайний Север, по мнению автора, целесообразно к этой сети присоединять сетями на основе СВЧ-антенн, реализованных на микрополосковых линиях.

Предложена топологическая реализация E-GOVERNANCE, позволяющая максимально экономично покрыть территорию Российской Федерации цифровым пространством и снизить влияние фактора географических масштабов.

Доклад Н.В. Гришиной «**Анализ проблем обеспечения информационной безопасности субъекта информационных отношений**» посвящён описанию нового подхода к обеспечению информационной безопасности субъекта информационных отношений как элемента единого информационного поля, в котором данный субъект реализует свои бизнес-процессы. Автор делает вывод, что на современном этапе необходимо оценивать информационную безопасность субъекта информационных отношений как элемента единого информационного пространства.

При этом имитационное моделирование процессов взаимодействия субъектов единого информационного пространства позволяет проанализировать процессы взаимодействия субъектов, определить риск потери информации, риск получения неполной или недостоверной информации, риск получения дезинформации, риск использования информации другим субъектом, а также выявить и исследовать влияние различных факторов, определяющих качество организации и функционирования механизмов, обеспечивающих информационную безопасность субъектов.

В докладе К.Н. Колюцкого «**Возможный подход к оценке информативности признаков распознавания**» рассмотрена возможность оценки информативности признаков распознавания на основе нейросетевого решения. Приведены структура нейронной сети и особенности алгоритма, оценивающего информативность. Процедура оценки информативности заключается в формировании выходного слоя в результате однократного последовательного предъявления сети всех образцов обучающей выборки. Входной слой и слой образцов образуют полносвязную структуру. Размерность входного слоя определяется количеством признаков распознавания, а размерность слоя образцов - количеством образцов в обучающей выборке. Весовые значения связей, входящих в каждый элемент слоя образцов, устанавливаются равными значениям признаков соответствующего образца.

И.И. Медведева в докладе «**Программно-технологические методы защиты компьютерной информации**» анализирует специализированные программно-технологические методы защиты компьютерной информации от несанкционированного просмотра и прослушивания. В их число входит окрашивание текста в различные цвета (в основном в белый); установление минимального размера букв; размещение невидимого (белого) текста в рисунках, размеры которых затем следует уменьшать до минимально возможных, и некоторые другие. Для защиты звуковой компьютерной информации указываются следующие приёмы: уменьшение громкости звука на 50 – 60 дБ с использованием специальной функции соответствующего звукового редактора; размещение в начале звуковой дорожки какой-либо звуковой информации, доступной для прослушивания, а затем защищаемой информации с уменьшенной громкостью на 50 – 60 дБ и ряд других.

Для защиты текстовой, табличной информации, а также аудио- и видеoinформации автор предлагает специальную «игру» с расширениями файлов.

В докладе Г.Е. Шепитько «**Зарубежный опыт защиты информации в национальных доменах Интернета**» обсуждается возможность использования зарубежного опыта для введения регулирования сети Интернет в России. Анализируя зарубежный опыт по противодействию сексуальной эксплуатации детей и детской порнографии в Интернете, автор приходит к выводу, что в этой области в части регулирования национальных доменов Россия отстает от развитых демократических стран из-за отсутствия необходимого правового обеспечения.

Для защиты российских пользователей от угроз в Интернете предлагается многоуровневая контентная фильтрация:

- первый уровень (внешний периметр Интернета) – международные и отечественные поисковые системы;
- домены первого уровня (РФ и RU) с использованием web-архивов для защиты авторских прав;
- домены второго уровня и ниже (сайты ведомств, организаций, предприятий);
- рабочие места пользователей.

Доклад В.В. Арутюнова «Об интернет-рисках для пользователей сетей» посвящён анализу статистики компьютерных преступлений за последние десять лет, современных видов угроз для пользователей со стороны сети Интернет (с классификацией интернет-рисков для детей), различных типов интернет-зависимости. Отмечаются следующие основные современные угрозы от сети для физических лиц и организаций, приводящие к возникновению интернет-рисков при работе пользователей в Интернете: использование классических компьютерных вирусов, троянских программ, интернет-червей, вредоносного программного обеспечения; неправомерные действия со стороны специалистов с использованием ИС; спам; DDos-атаки; финансовое мошенничество с использованием ИС. Среди интернет-рисков для детей выделяются следующие категории: коммуникационные (связанные с общением в Интернете); компьютерные вирусы; контентные (порнография, распространение наркотиков, пропаганда ненависти, жестокости и экстремизма); потребительские риски, когда под различными предложениями детей вынуждают подключаться к платным услугам; интернет-зависимость как новый вид психического расстройства. Автор отмечает, что психиатры выделяют при этом пять типов интернет-зависимости: киберсекс (любовь к порносайтам), виртуальные знакомства, игры на деньги, ролевые онлайн-игры и интернет-сёрфинг (бездумное перелистывание страниц). Для снижения интернет-зависимости автор считает целесообразным предусмотреть преподавание в школах или на начальных курсах вузов специального курса по современным методам и средствам, применяемым для снижения интернет-рисков при работе в сети.

В докладе А.А. Афониной «Формирование конкурентоспособного специалиста в условиях современной рыночной экономики» рассматривается необходимость формирования конкурентоспособного специалиста в сложившихся условиях рыночной экономики; анализируется человеческий капитал как существенный фактор роста и развития экономики. В качестве основного средства формирования конкурентоспособного человеческого капитала в процессе профессиональной подготовки рассматривается метод проектов. Формулируются преимущества и недостатки метода проектов, а также его основные идеи, включающие: укрепление связи обучения с актуальными потребностями участников общественно-экономических и социально-культурных отношений; развитие индивидуальных способностей обучающихся; развитие умения самостоятельно планировать

свою работу и выполнять её, доводя до намеченного результата; формирование системы контроля и самоконтроля.

Из недостатков проектного подхода можно выделить: малый опыт работы с применением данного метода обучения; локальное применение данного метода (только в крупных учебных заведениях). К основным преимуществам метода проектов относятся следующие: метод проектов позволяет наименее ресурсозатратным способом создать «естественную среду», т. е. условия деятельности, максимально приближенные к реальным ситуациям для формирования компетентностей учащихся; он предполагает непрерывное решение учащимися какой-то актуальной проблемы. Метод проектов может быть реализован с помощью различных средств обучения, в том числе и с использованием новейших информационных технологий.

Л.И. Шаталова в докладе «Геометрический проект учащихся в математическом образовании» акцентирует внимание на актуальной для современной педагогической науки проблеме организации учебной деятельности учащихся. В нем рассматривается геометрический проект – экскурсия для развития математических способностей учащихся 5-6 классов. Автор раскрывает содержание фазы проектирования, включающей поисковый, организационный и информационный этапы, и технологической фазы, содержащей этапы технологии реализации проекта и оформления результатов; при этом автор описывает содержание каждого этапа.

В докладе А.Г. Юферова «Вопросы обобщения универсальной десятичной классификации» рассматривается применение в УДК оснований деления с произвольным числом значений для решения задач классификации технико-экономической и социальной информации. В этом случае единственное отличие от традиционного применения УДК состоит в том, что кодовая последовательность УДК (индекс УДК) может обозначать не только понятие, но и материальный объект. Такой подход с учетом того факта, что система УДК обеспечивает неограниченный рост классификатора «в глубину» детализации, с одной стороны, обеспечивает унификацию различных разрабатываемых классификаторов, а с другой – решает задачу развития УДК путем включения непосредственно в её структуру существующих классификаторов.

Тезисы докладов конференции изданы в трех номерах «Вестника Московского финансово-юридического университета» (№ 1-3 за 2013г.). С электронными версиями тезисов докладов можно ознакомиться на портале МФЮА по адресу: www.mfua.ru.

Сведения об авторе

АРУТЮНОВ Валерий Вагаршакович – доктор технических наук, профессор Российского государственного гуманитарного университета, Москва
e-mail: awagar@list.ru

Указатель статей, опубликованных в сборнике «Научно-техническая информация», и Авторский указатель за 2013 год*

Указатель статей

**МАТЕРИАЛЫ ВСЕРОССИЙСКОЙ
НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ С
ИНОСТРАННЫМ УЧАСТИЕМ «МЕТОДИКА
ИСПОЛЬЗОВАНИЯ И ВЕДЕНИЯ
УНИВЕРСАЛЬНЫХ КЛАССИФИКАЦИЙ
НАУЧНОГО ЗНАНИЯ – УДК, ББК, ГРНТИ»
Москва, ВИНТИ РАН, ноябрь 2012 г.**

Антошкова О.А. Универсальная десятичная классификация для библиотечного и книжного дела	4 (1) 3
Смирнова О.В. Составление индексов УДК. Методика и практика	4 (1) 5
Астахова Т.С., Дмитриева Е.Ю., Сурикова Н.Г. Эталонный массив УДК на русском языке	4 (1) 7
Ивановский А.А. Применение УДК как основы лингвистического обеспечения сводного тематико-типологического плана комплектования ЦБС БЕН РАН: новые принципы	4 (1) 9
Сукиасян Э.Р. Проблемы классификационной практики. Ситуационный анализ организационных и технологических решений	4 (1) 12
Сухоруков М.К., Гребенникова А.А. Тематическая рубрикация электронных изданий средствами УДК: практика применения, проблемы и перспективы	4 (1) 15
Белоозеров В.Н., Шапкин А.В., Шабурова Н.Н. Табличное и тезаурусное сопоставление классификационных систем	4 (1) 18
Шабурова Н.Н. Проблемы систематизации библиотечных фондов академического института	4 (1) 21
Антопольский А.Б. Классификации информационных ресурсов	4 (1) 25
Старых В.А., Башмаков А.И. Классификация и каталогизация информационных ресурсов сферы образования	4 (1) 27
Белоозеров В.Н., Старых В.А. Онтология классов информационных ресурсов образования	4 (1) 31

Митрович Божидар. УДК: исследования истории славян до VI века	4 (1) 37
Антошкова О.А., Белоозеров В.Н., Дмитриева Е.Ю. Стандарты, используемые в информационной деятельности	4 (1) 38

ОБЩИЙ РАЗДЕЛ

Семенюк Э.П. Информационный подход и реалии современности	1 (1) 1
Двоеносова Г.А. Функции документа	2 (1) 1
Столяров Ю.Н. Онтология документа: дополнительные пояснения	2 (1) 8
Берестова Т.Ф. Свойства информации как потенциал ее иерархического функционирования и видового многообразия	3 (1) 1
Двоеносова Г.А. Статус документа	5 (1) 1
Лобанов А.С. Основные понятия квалиметрии	5 (1) 11
Полтавская Е.И. Информация: к уточнению понятия	5 (2) 1
Галявиева М.С. О становлении понятия «Информетрия» (Обзор)	6 (1) 1
Сютюренко О.В. Макропроблемы информационной безопасности человека и общества	6 (1) 11
Зацман И.М. Информационно-компьютерная наука: предпосылки становления	7 (1) 1
Штеренберг М.И. Исходные понятия энтропии, порядка, организации, информации, знания и смысла	7 (1) 13
Успенский В.А. В.К. Финн на фоне зарождения семиотики в ВИНТИ	7 (2) 2
Стародубов В.И., Куракова Н.Г., Цветкова Л.А., Арефьев П.Г., Еремченко О.А. Соответствие предметной структуры науки России приоритетам ее финансирования	8 (1) 1
Нестерович Ю.В. Предпосылки формирования терминосистемы документоведения (Обзор)	8 (1) 15

* 4 означает номер сборника, (1) – серию, 3 – страницу

Белоногов Г.Г., Гиляревский Р.С., Селетков С.Н., Хорошилов А.А. О путях повышения качества поиска текстовой информации в системе Интернет	8 (2) 1	Захарова О.С., Шогин А.Н., Старцева О.Б., Олейникова О.Л. Автоматизация верстки периодических реферативных изданий ВИНИТИ РАН в системе Adobe InDesign	5 (1) 31
Полтавская Е.И. Информация субъективная, социально-определенная и документ	8 (2) 12	Маркусова В.А., Аллахвердян А.Г. Факторы, влияющие на подготовку нового поколения	5 (1) 35
Плешкевич Е. А. Документальность как атрибутивное свойство документа	9 (1) 1	высококвалифицированных научных кадров в Европе и США: сравнительный анализ	
Финн В.К. Эпистемологические основания ДСМ-метода автоматического порождения гипотез. Часть I	9 (2) 1	Арутюнов В.В. О практике государственной аттестации студентов вуза по специальности «Организация и технология защиты информации»	6 (1) 15
Нестерович Ю.В. Базисная теоретическая схема информологии и экспликация понятия документальной информации (Обзор)	10 (1) 1	Ковалев А.И. Информационное обеспечение качества деятельности предприятия	6 (1) 21
Плешкевич Е. А. Философские проблемы идентификации документа как объекта познания	11 (1) 1	Мартirosян З.Г., Саркисян Д.Б. Применение интернет-технологий в учебном процессе	6 (1) 31
Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А., Серебряков В.А., Теймуразов К.Б. Технология научных публикаций в среде «открытых связанных данных»	12 (1) 1	Алешин И.М., Корягин В.Н., Сухорослов О.В., Холодков К.И., Шогин А.Н. Инверсия сейсмических данных: высокоуровневый веб-интерфейс к инструментарию Globus Toolkit	7 (1) 19
Финн В.К. Эпистемологические основания ДСМ-метода автоматического порождения гипотез. Часть II	12 (2) 1	Арутюнов В.В. СПАМ: прошлое, настоящее, будущее	8 (1) 24
ОРГАНИЗАЦИЯ ИНФОРМАЦИОННОЙ РАБОТЫ		Бапиева М.К., Чултурова Д.Ш., Кульевская Ю.Г., Улезько Г.Г. Мониторинг реализации научно-технических программ в Казахстане	9 (1) 8
		Шемберко Л.В., Слива А.И., Денискин А.А. Информация по религиоведению в базах данных ИНИОН РАН по социальным и гуманитарным наукам: принципы обработки и выбор стратегии поиска	9 (1) 17
Мельникова Е.В., Мельников О.А., Саркисян Д.Б. Система НТИ Великобритании: современное состояние и роль в инновационном развитии страны	1 (1) 12	Раевская Е.Г. Особенности интеграции научных изданий некоторых стран исламского мира в международные базы данных	9 (1) 28
Харыбина Т.Н., Бескаравайная Е.В., Мохначева Ю.В., Слащева Н.А. Тенденции развития научных школ в Пушинском научном центре РАН	2 (1) 14	Лукашевич А.В., Лукашевич Н.Л., Седякина А.Н. Технология перекрестных ссылок в Реферативном журнале ВИНИТИ РАН	9 (1) 34
Орлова М.М. Стандартизация управления ИТ-услугами: исторический аспект	2 (1) 20	Арутюнов В.В. О результативности научной деятельности в области приоритетных направлений развития науки, технологий и техники	10 (1) 12
Егоров В.С. Особенности доступа к научной журнальной литературе в электронном обществе	3 (1) 8	Брежнева В.В., Гиляревский Р.С. О содержании учебной и научной дисциплины «Информационный менеджмент»	10 (1) 20
Шогин А.Н., Алпатова М.Г., Журавлева И.П., Олейникова О.Л. Реализация словаря как инверсной структуры для информационно-поисковой системы в БнД ВИНИТИ РАН	3 (1) 19	Берёзкина Н.Ю., Сикорская О.Н., Хренова Г.С. Использование баз данных цитирования для оценки научной деятельности организаций Беларуси	10 (1) 25
Арский Ю.М., Быков В.А. Деятельность ВИНИТИ РАН – Базовой организации государств – участников СНГ по межгосударственному обмену научно-технической информацией	5 (1) 23		

Ковалев А.И. Логико-содержательное проектирование системы информационного обеспечения качества деятельности предприятия	11 (1) 10	Буйлова Н.М., Елецкий А.В., Зицерман В.Ю., Кобзев Г.А. Систематизация источников и данных по нанотехнологиям	11 (1) 31
Боровик М.А., Шемберко Л.В. Информационное обеспечение исследований по проблемам украиноведения на основе электронных ресурсов ИНИОН РАН	11 (1) 20	Ефременкова В.М. Системы классификации по нанонауке и нанотехнологиям (Обзор)	12 (1) 19
Василенко Е.А., Панфилов В.И., Жуков Д.Ю., Сивуха Д.В. Разработка наукометрической базы данных публикационной активности ученых РХТУ им. Д.И Менделеева	11 (1) 26	ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ	
Гуреев В.Н., Мазов Н.А. Анализ цитирования как основа для разработки дополнительного модуля в системах антиплагиата	12 (1) 12	Старичкова Ю.В. Структурная сложность орграфов: некоторые математические модели и их приложения	2 (2) 1
Веселовский А.В., Гальберг Т.В. Центр обработки цифровых данных в составе информационной системы академического института	12 (1) 16	Михеенкова М.А., Волкова А.Ю. Спецификация интеллектуальной системы типа ДСМ	7 (2) 5
		Бетин В.Н., Супрун А.П. Ситуационные центры: методика грамматического обучения интеллектуальной аналитической системы	9 (2) 30
		Волкова А.Ю. Опыт создания интеллектуальной ДСМ-системы для исследования данных различных предметных областей	11 (2) 12
ДОКУМЕНТАЛЬНЫЕ ИСТОЧНИКИ ИНФОРМАЦИИ		ИНФОРМАЦИОННЫЕ СИСТЕМЫ	
Московкин В.М. Построение университетских публикационно-терминологических структур с помощью поисковой машины Google Scholar: на примере экологических терминов и классических университетов Харькова и Скопье	1 (1) 26	Голицына О.Л., Максимов Н.В., Окропишина О.В., Строгонов В.И. Онтологический подход к идентификации информации в задачах документального поиска: практическое применение	3 (2) 1
Ставинский Е.Н., Романова М.С., Ситникова И.С. Документальное обеспечение научно-исследовательских работ в академическом институте	1 (1) 32	Соловьёв А. А., Пескова О. В. Об архитектурах программных систем вопросно-ответного поиска	10 (2) 1
Маркусова В.А., Крылова Т.А., Либкинд А.Н., Зиновьева Н.В., Миндели Л.Э. Библиографические показатели федеральных и национальных университетов России по БД Web of Science	2 (1) 24	ИНФОРМАЦИОННЫЙ ПОИСК	
Овченкова Е.А. Журналы по нанотехнологиям в системе научной периодики России	2 (1) 38	Кузнецов Б.А. Опыт применения поисковой оболочки с автоматическим созданием терминологических комбинаций из текста запроса для поиска в реферативных массивах на примере БД «Медицина»	1 (2) 1
Нестеров А.В. О востребованности публикаций исследователей	3 (1) 24	Ильвовский Д.А., Климушкин М. А. Выявление дубликатов объектов в прикладных онтологиях с помощью методов анализа формальных понятий	1 (2) 10
Мдивани Р.Р. Тезаурусы ИНИОН РАН по социальным и гуманитарным наукам	7 (1) 23	Ковалев И.В., Зеленков П.В., Распопин Н.А., Демиш А.В., Бахмарева К.К. Дискретная декомпозиция поискового запроса с учетом семантической связи структур данных	10 (2) 8
Штовба С.Д., Штовба Е.В. Индекс цитирования, учитывающий скрытую диффузию научных знаний	7 (1) 28	Калафати Ю.Д., Козловский П.А., Старков С.О., Тельных А.А. Поиск «похожего» в графических базах	10 (2) 12
Егоров В.С. Научная книжная литература в электронном обществе	8 (1) 33		
Гуреев В.Н., Мазов Н.А. Тематика публикаций организации как основа формирования объективного и оптимального репертуара научной периодики	10 (1) 30		

Нефедов О.М., Трепалин С.В., Королева Л.М., Бессонов Ю.Е. Быстрый поиск точных химических структур в больших базах данных с использованием InChI Key-кодировки структур	12 (2) 27	Вагин В.Н., Фомина М.В., Антипов С.Г. Моделирование алгоритмов индуктивного формирования понятий в «зашумленных» базах данных	7 (2) 20
--	-----------	--	----------

ИНФОРМАЦИОННЫЕ ЯЗЫКИ

Кирков А.Ю., Павловский В.Е. Акустический тональный язык коммуникации роботов	2 (2) 8	Гриняк В.М., Девятисильный А.С. Классификация движущихся объектов типа «надводный-воздушный» в лингвистических переменных	8 (2) 20
Грибова В.В., Клещев А.С., Крылов Д.А. Контекстно-свободные грамматики искусственных языков	4 (2) 9	Яшков И.Б. Сравнение методов дискретизации и отбора непрерывных параметров для логико-комбинаторной классификации	8 (2) 26
Грибова В.В., Клещев А.С., Крылов Д.А. Контекстно-зависимые грамматики искусственных языков	6 (2) 1	Осипов А.Л., Бобров Л.К. Прогнозирование свойств химических соединений на основе структурно-неаддитивных моделей с учетом парциальных вкладов структурных элементов	9 (2) 35
Грибова В.В., Клещев А.С. Онтологическая парадигма программирования	11 (2) 27	Волкова Г.А. Принципы семантического анализа предметной области при создании информационно-аналитических систем	10 (2) 19
Клышинский Э.С., Кочеткова Н.А., Логачева В.К. Метод кластеризации слов с использованием информации об их синтаксической связности	11 (2) 36	Бузмаков А.В. Узорные структуры для анализа сложных последовательностей	10 (2) 27

ИНФОРМАЦИОННЫЙ АНАЛИЗ

Карпов А. А. Машинный синтез русской дактильной речи по тексту	1 (2) 20	АВТОМАТИЗАЦИЯ ОБРАБОТКИ ТЕКСТА	
Либкинд А.Н., Маркусова В.А., Либкинд И.А., Янц М. (Нидерланды), Иванов К.Н. Моделирование динамики процесса сохранения журналов в качестве наиболее авторитетных научных изданий	3 (2) 9	Гусев С.В., Чеповский А.М. Автоматическая идентификация текстов на славянских языках, пользующихся кириллицей, записанных латинским алфавитом	1 (2) 27
Ярушкина Н.Г., Афанасьева Т.В., Тонерян М.С. Применение нечетко-гранулярного моделирования числовых временных рядов	3 (2) 35	Шумарин С.И. Адаптация иноязычных аббревиатур в современном русском языке (на материале терминов мобильной связи)	2 (2) 16
Бауман К.Е., Корнетова А.Н., Топинский В.А., Хакимова Д.А. Оптимизация прогноза вероятности посещения контекстной рекламы в поисковой системе «Яндекс»	4 (2) 1	Якушевич И.В. Лингво-когнитивная модель символа	2 (2) 24
Рыжков О.Ю., Бобров Л.К. Формализованная запись нормативных документов для компьютерной обработки	5 (2) 7	Падучева Е.В. Онтологические категории имен речи: проблема категориальной неоднозначности	4 (2) 18
Блинова В.Г., Решетникова В.В. База данных «ДСМ-эксперименты по фармакологии»	5 (2) 15	Хайруллин В.И. Корпусный словарь фобиотерминологии	4 (2) 30
Ревенко А.В. Нахождение ошибок в бинарных таблицах данных	6 (2) 10	Васильченко К.А. Терминологические словосочетания и проблема их определения	4 (2) 33
Логинов Е.Л. Информационная платформа, объединяющая телематические, вычислительные и информационные сервисы в ЕЭС России	6 (2) 19	Кузнецов И.О. Автоматическое извлечение двусловных терминов по тематике «Нанотехнологии в медицине» на основе корпусных данных	5 (2) 25
		Яцко В.А. Алгоритмы распознавания собственных имен	5 (2) 34

Акинина Ю. С., Кузнецов И. О., Толдова С. Ю. Сравнение двух методов автоматического извлечения участников события из неструктурированных источников	6 (2) 24	Захаров В.П. Новая книга Ю. Д. Григорьева и Г. Я. Мартыненко	2 (2) 37
Алексеев А.В. К вопросу о символическом семиозисе слова в истории языка	8 (2) 32	Плешкевич Е.А. История архивоведческой мысли: размышления в связи с выходом учебника [Рец. на кн.]	3 (1) 29
		Плющ М. А. Об использовании Интернета для сбора информации по истории книги и библиотек: на примере сведений о реестре А.И. Богданова	7 (1) 32
СПРАВОЧНО-ИНФОРМАЦИОННЫЙ РАЗДЕЛ		Шелов С.Д., Цумарев А.Э. Белорусская терминология и профессиональная лексика в словарях и в компьютере [Рец. на кн.]	11 (1) 43
Михеенкова М.А., Дружинина Е.Г. О состоянии и перспективах российских исследований в области искусственного интеллекта (по материалам конференции КИИ-2012)	2 (2) 30	Арутюнов В.В. О Всероссийской научно-практической конференции «Математика, информатика, естествознание в экономике и обществе»	12 (1) 28

Авторский указатель

Акинина Ю. С.	6 (2) 24	Буйлова Н.М.	11 (1) 31	Еремченко О.А.	8 (1) 1
Алешин И.М.	7 (1) 19	Быков В.А.	5 (1) 23	Еркимбаев А.О.	12 (1) 1
Алексеев А.В.	8 (2) 32			Ефременкова В.М.	12 (1) 19
Аллахвердян А.Г.	5 (1) 35	Вагин В.Н.	7 (2) 20		
Алпатова М.Г.	3 (1) 19	Василенко Е.А.	11 (1) 26	Жуков Д.Ю.	11 (1) 26
Антипов С.Г.	7 (2) 20	Васильченко К.А.	4 (2) 33	Журавлева И.П.	3 (1) 19
Антопольский А.Б.	4 (1) 25	Веселовский А.В.	12 (1) 16		
Антошкова О.А.	4 (1) 3	Волкова А.Ю.	7 (2) 5		
	4 (1) 38		11 (2) 12	Забежайло М.И.	7 (2) 33
Арефьев П.Г.	8 (1) 1	Волкова Г.А.	10 (2) 19	Захаров В.П.	2 (2) 37
Арский Ю.М.	5 (1) 23	Гальберг Т.В.	12 (1) 16	Захарова О.С.	5 (1) 31
Арутюнов В.В.	6 (1) 15	Галаяниева М.С.	6 (1) 1	Зацман И.М.	7 (1) 1
	8 (1) 24	Гиляревский Р.С.	8 (2) 1	Зеленков П.В.	10 (2) 8
	10 (1) 12		10 (1) 20	Зиновьева Н.В.	2 (1) 24
	12 (1) 28	Голицына О.Л.	3 (2) 1	Зицерман В.Ю.	11 (1) 31
Астахова Т.С.	4 (1) 7	Гребенникова А.А.	4 (1) 15		12 (1) 1
Афанасьева Т.В.	3 (2) 35	Грибова В.В.	4 (2) 9		
			6 (2) 1		
Бапиева М.К.	9 (1) 8		11 (2) 27	Иванов К.Н.	3 (2) 9
Бауман К.Е.	4 (2) 1	Гриняк В.М.	8 (2) 20	Ивановский А.А.	4 (1) 9
Бахмарева К.К.	10 (2) 8	Гуреев В.Н.	10 (1) 30	Ильвовский Д.А.	1 (2) 10
Башмаков А.И.	4 (1) 27		12 (1) 12		
Белоногов Г.Г.	8 (2) 1	Гусев С.В.	1 (2) 27	Калафати Ю.Д.	10 (2) 12
Белоозеров В.Н.	4 (1) 18			Карпов А. А.	1 (2) 20
	4 (1) 31	Двоеносова Г. А.	2 (1) 1	Кирков А.Ю.	2 (2) 8
	4 (1) 38		5 (1) 1	Клещев А.С.	4 (2) 9
Берёзкина Н.Ю.	10 (1) 25	Девятисильный А.С.	8 (2) 20		6 (2) 1
Берестова Т.Ф.	3 (1) 1	Демиш А.В.	10 (2) 8		11 (2) 27
Бескаравайная Е.В.	2 (1) 14	Денискин А.А.	9 (1) 17	Климушкин М. А.	1 (2) 10
Бессонов Ю.Е.	12 (2) 27	Дмитриева Е.Ю.	4 (1) 7	Клышинский Э.С.	11 (2) 36
Бетин В.Н.	9 (2) 30		4 (1) 38	Кобзев Г.А.	11 (1) 31
Блинова В.Г.	5 (2) 15	Дружинина Е.Г.	2 (2) 30		12 (1) 1
Бобров Л.К.	5 (2) 7			Ковалев А.И.	6 (1) 21
	9 (2) 35				11 (1) 10
Боровик М.А.	11 (1) 20	Егоров В.С.	3 (1) 8	Ковалев И.В.	10 (2) 8
Брежнева В.В.	10 (1) 20		8 (1) 33	Козловский П.А.	10 (2) 12
Бузмаков А.В.	10 (2) 27	Елецкий А.В.	11 (1) 31	Корнетова А.Н.	4 (2) 1

Королева Л.М.	12 (2) 27	Орлова М.М.	2 (1) 20	Теймуразов К.Б.	12 (1) 1
Корягин В.Н.	7 (1) 19	Осипов А.Л.	9 (2) 35	Тельных А.А.	10 (2) 12
Кочеткова Н.А.	11 (2) 36			Толдова С. Ю.	6 (2) 24
Крылов Д.А.	4 (2) 9			Тонерян М.С.	3 (2) 35
	6 (2) 1	Павловский В.Е.	2 (2) 8	Топинский В.А.	4 (2) 1
Крылова Т.А.	2 (1) 24	Падучева Е.В.	4 (2) 18	Трепалин С.В.	12 (2) 27
Кузнецов Б.А.	1 (2) 1	Панфилов В.И.	11 (1) 26		
Кузнецов И.О.	5 (2) 25	Пескова О. В.	10 (2) 1		
	6 (2) 24	Плешкевич Е.А.	3 (1) 29	Улезько Г.Г.	9 (1) 8
Кульевская Ю.Г.	9 (1) 8		9 (1) 1	Успенский В.А.	7 (2) 2
Куракова Н.Г.	8 (1) 1		11 (1) 1		
		Плющ М. А.	7 (1) 32		
		Полтавская Е.И.	5 (2) 1	Финн В.К.	9 (2) 1
Либкинд А.Н.	2 (1) 24		8 (2) 12		12 (2) 1
	3 (2) 9			Фомина М.В.	7 (2) 20
Либкинд И.А.	3 (2) 9				
Лобанов А.С.	5 (1) 11	Раевская Е.Г.	9 (1) 28		
Логачева В.К.	11 (2) 36	Распопин Н.А.	10 (2) 8	Хайруллин В.И.	4 (2) 30
Логинов Е.Л.	6 (2) 19	Ревенко А.В.	6 (2) 10	Хакимова Д.А.	4 (2) 1
Лукашевич А.В.	9 (1) 34	Решетникова В.В.	5 (2) 15	Харыбина Т.Н.	2 (1) 14
Лукашевич Н.Л.	9 (1) 34	Романова М.С.	1 (1) 32	Холодков К.И.	7 (1) 19
		Рыжков О.Ю.	5 (2) 7	Хорошилов А.А.	8 (2) 1
				Хренова Г.С.	10 (1) 25
Мазов Н.А.	10 (1) 30				
	12 (1) 12	Саркисян Д.Б.	1 (1) 12		
Максимов Н.В.	3 (2) 1		6 (1) 31	Цветкова Л.А.	8 (1) 1
Маркусова В.А.	2 (1) 24	Седякина А.Н.	9 (1) 34	Цумарев А.Э.	11 (1) 43
	3 (2) 9	Селетков С.Н.	8 (2) 1		
	5 (1) 35	Семенюк Э.П.	1 (1) 1		
Мартиросян З.Г.	6 (1) 31	Серебряков В.А.	12 (1) 1	Чеповский А.М.	1 (2) 27
Мдивани Р.Р.	7 (1) 23	Сивуха Д.В.	11 (1) 26	Чултурова Д.Ш.	9 (1) 8
Мельникова Е.В.	1 (1) 12	Сикорская О.Н.	10 (1) 25		
Мельников О.А.	1 (1) 12	Ситникова И.С.	1 (1) 32		
Миндели Л.Э.	2 (1) 24	Слащева Н.А.	2 (1) 14	Шабурова Н.Н.	4 (1) 18
Митрович Божида	4 (1) 37	Слива А.И.	9 (1) 17		4 (1) 21
Михеенкова М.А.	2 (2) 30	Смирнова О.В.	4 (1) 5	Шапкин А.В.	4 (1) 18
	7 (2) 5	Соловьёв А. А.	10 (2) 1	Шелов С.Д.	11 (1) 43
Московкин В.М.	1 (1) 26	Ставинский Е.Н.	1 (1) 32	Шемберко Л.В.	9 (1) 17
Мохначева Ю.В.	2 (1) 14	Старичкова Ю.В.	2 (2) 1		11 (1) 20
		Старков С.О.	10 (2) 12	Шогин А.Н.	3 (1) 19
		Стародубов В.И.	8 (1) 1		5 (1) 31
Нестеров А.В.	3 (1) 24	Старцева О.Б.	5 (1) 31		7 (1) 19
Нестерович Ю.В.	8 (1) 15	Старых В.А.	4 (1) 27	Штеренберг М.И.	7 (1) 13
	10 (1) 1		4 (1) 31	Штовба Е.В.	7 (1) 28
Нефедов О.М.	12 (2) 27	Столяров Ю.Н.	2 (1) 8	Штовба С.Д.	7 (1) 28
Нешитой В.В.	11 (2) 1	Строгонов В.И.	3 (2) 1	Шумарин С.И.	2 (2) 16
		Супрун А.П.	9 (2) 30		
		Сухорослов О.В.	7 (1) 19	Якушевич И.В.	2 (2) 24
Овченкова Е.А.	2 (1) 38	Сукиасян Э.Р.	4 (1) 12	Янц М. (Нидерланды)	3 (2) 9
Окропишина О.В.	3 (2) 1	Сурикова Н.Г.	4 (1) 7	Ярушкина Н.Г.	3 (2) 35
		Сухоруков М.К.	4 (1) 15	Яцко В.А.	5 (2) 34
Олейникова О.Л.	3 (1) 19	Сюнтюренко О.В.	6 (1) 11	Яшков И.Б.	8 (2) 26
	5 (1) 31				

УВАЖАЕМЫЕ ЧИТАТЕЛИ!

ЦЕНТР НАУЧНО-ИНФОРМАЦИОННОГО ОБСЛУЖИВАНИЯ ВИНИТИ РАН

ПРЕДОСТАВЛЯЕТ КОПИИ ПЕРВОИСТОЧНИКОВ

ВИНИТИ РАН осуществляет обслуживание копиями первоисточников, хранящихся в фонде научно-технической литературы ВИНИТИ, в фондах других библиотек, а также в доступных ВИНИТИ электронных ресурсах.

Фонд научно-технической литературы ВИНИТИ включает более 2 млн изданий по точным, естественным и техническим наукам, в т.ч.:

- отечественные и иностранные периодические и продолжающиеся издания – с 1987 г.;
- отечественные книги – с 1987 г.;
- иностранные книги – с 1991 г.;
- рукописи, депонированные в ВИНИТИ, – с 1962 г.

Заказы на бумажные или электронные копии первоисточников принимает Центр научно-информационного обслуживания (ЦНИО) ВИНИТИ. ЦНИО ВИНИТИ обслуживает коллективных (организации и учреждения) и индивидуальных пользователей.

Формы обслуживания:

- абонементная (на основе договоров и предоплаты);
- разовые заказы (с предоплатой заказа по счету);
- индивидуальная форма обслуживания в читальном зале ЦНИО ВИНИТИ.

На сайте ВИНИТИ (<http://www.viniti.ru>) представлен полный Электронный каталог научно-технической литературы (<http://catalog.viniti.ru>), зарегистрированной в ВИНИТИ с 1994 г. Доступ для просмотра и поиска по Каталогу свободный. Постоянные абоненты ЦНИО ВИНИТИ, имеющие логин и пароль для работы с Каталогом, могут делать заказ копий непосредственно через Каталог.

Услуги по изготовлению копий первоисточников из фондов других библиотек предоставляются только постоянным абонентам. Место хранения первоисточников указывается в Электронном каталоге.

За подробной информацией обращаться по адресу:

125190, Россия, Москва, ул. Усиевича, 20, ВИНИТИ РАН, ЦНИО

Телефоны: 8 (499) 155-42-43, 155-42-09, 152-54-59

Факс: 8 (499) 943-00-60

E-mail: cnio@viniti.ru; **URL:** <http://www.viniti.ru>

БАЗА ДАННЫХ ВИНИТИ РАН

ВИНИТИ предлагает к использованию через WWW-сервер (<http://www.viniti.ru>) крупнейшую Федеральную базу отечественных и зарубежных публикаций по естественным, точным и техническим наукам. БД ВИНТИ РАН генерируется с 1981 г., обновляется ежемесячно, пополнение составляет около 1 млн документов в год. БД ВИНТИ представлена ретроспективными тематическими фрагментами и единой политематической БД (ретроспектива с 2001 г.), объединяющей все тематические фрагменты БД ВИНТИ.

БД ВИНТИ РАН в сети INTERNET

Сервер ВИНТИ – <http://www.viniti.ru> – обеспечивает on-line доступ к Базе данных ВИНТИ РАН круглосуточно без выходных.

На основе БД ВИНТИ РАН предоставляются следующие услуги:

- Диалоговый поиск научно-технической информации в **режиме on-line**;
- **Демо-версия**, позволяющая ознакомиться с основными функциями поисковой системы, составом данных, формами представления документов и получить навыки работы с системой;
- **Поисковые эксперты ВИНТИ** выполняют тематический поиск по разовым или постоянным запросам, а также окажут **консультационные услуги**.

БД ВИНТИ РАН на CD-ROM

Любые наборы тематических фрагментов БД ВИНТИ или их разделов могут быть предоставлены на **CD-ROM в поисковой системе (ИПС) "Сокол"**, обеспечивающей все поисковые функции, доступные в режиме on-line:

- Поиск можно вести в годовом или ретроспективном массиве (за несколько лет сразу) в одном или нескольких тематических фрагментах .
- Поиск по словам и любым словосочетаниям из заглавия, реферата, ключевых слов.
- Использование года, языка, рубрик, шифров тематических разделов БД для уточнения поиска.
- Поиск по словарю, выполняющему функции многоаспектного указателя, в том числе авторского, предметного, источников, индексов МПК, номеров патентных документов и депонированных рукописей и т.д.
- Возможность запоминания запросов для последующего их использования и/или редактирования.
- Чтение документов не только как в РЖ (последовательный просмотр документов одного номера за другим), но и чтение документов нужных тематических фрагментов (разделов) по оглавлению за весь период заказанной ретроспективы.

ИПС "Сокол" является прикладной программой Microsoft Windows.

Любые наборы тематических фрагментов БД ВИНТИ или их разделов могут быть подготовлены в **коммуникативных форматах** ISO-2709, МЕКОФ, txt на любых видах электронных носителей.

Продукты предоставляются на договорной основе.

Информационная служба БД ВИНТИ: 125190, Москва, ул. Усиевича 20, ВИНТИ

Телефон: (499) 155-45-01, 155-45-02, **Факс:** (499) 152-62-31 **e-mail:** csbd@viniti.ru

РОССИЙСКАЯ АКАДЕМИЯ НАУК
Федеральное государственное бюджетное учреждение науки
ВСЕРОССИЙСКИЙ ИНСТИТУТ НАУЧНОЙ И ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ
РОССИЙСКОЙ АКАДЕМИИ НАУК

предлагает научным работникам, аспирантам и другим специалистам в области естественных, точных и технических наук, желающим быстро и эффективно опубликовать результаты своей научной и научно-производственной деятельности, использовать способ публикации своих работ через *систему депонирования*.

«Депонирование (передача на хранение) – особый метод публикации научных работ (отдельных статей, обзоров, монографий, сборников научных трудов, материалов научных мероприятий – конференций, симпозиумов, съездов, семинаров) узкоспециального профиля, разрешенных в установленном порядке к открытому опубликованию, которые нецелесообразно издавать полиграфическим способом печати, а также работ широкого профиля, срочная информация о которых необходима для утверждения их приоритета. Депонирование предусматривает прием, учет, регистрацию, хранение научных работ и обязательное размещение информации о них в специальных информационных изданиях».

Подготовка и передача на депонирование научных работ происходит в соответствии с «Инструкцией о порядке депонирования научных работ по естественным, техническим, социальным и гуманитарным наукам» (М., 2003).

Результатом депонирования является публикация информации о депонированных научных работах в информационных изданиях ВИНТИ РАН – Реферативном журнале и аннотированном библиографическом указателе «Депонированные научные работы».

В соответствии с “Положением о порядке присуждения ученых степеней”, утвержденным Постановлением Правительства Российской Федерации от 30.01.2002 № 74 (в ред. Постановлений Правительства РФ от 20.04.2006 № 227, от 02.06.2008 № 424, от 20.06.2011 № 475), научные работы, депонированные в организациях государственной системы научно-технической информации, признаны публикациями, учитываемыми при защите кандидатских и докторских диссертаций.

Подать научную работу на депонирование можно обратившись в Отдел депонирования ВИНТИ РАН по адресу:

125190, Москва, ул. Усиевича, 20.
ВИНТИ РАН, Отдел депонирования научных работ.
Тел.: 8 (499) 155-43-28, Факс: 8 (499) 943-00-60.
e-mail: dep@viniti.ru

С инструкцией о порядке депонирования можно ознакомиться на сайте ВИНТИ РАН:
<http://www.viniti.ru>