

НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 1. ОРГАНИЗАЦИЯ И МЕТОДИКА
ИНФОРМАЦИОННОЙ РАБОТЫ

ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 10

Москва 2016

ОБЩИЙ РАЗДЕЛ

УДК 620.3 : 002.6

А.О. Еркимбаев, В.Ю. Зицерман, Г.А. Кобзев, М.С. Трахтенгерц

Наноинформатика: задачи, методы и технологии*

Очерчен круг проблем, решаемых наноинформатикой – недавно возникшей дисциплиной, интегрирующей методы и инструменты распространения данных по наноматериалам и основанным на них приборах и технологиях. Суммированы особенности данных по свойствам наноразмерных объектов, определяемые междисциплинарным характером и быстрой эволюцией предметной области. Представлены наиболее популярные ресурсы (базы данных, классификаторы и онтологии) по свойствам наноматериалов. Отмечены тематические диспропорции, возникшие в наноинформатике из-за преимущественного внимания к наномедицине в ущерб традиционным областям использования нанотехнологий: электронике, энергетике и др. Детально рассмотрены основные стандарты по терминологии и классификации наноматериалов, являющиеся основой проектирования новых баз данных и онтологий. В качестве наиболее перспективного и универсального подхода к решению проблем наноинформатики предложено использовать международный стандарт CODATA (комиссии по численным данным) универсального описания наноматериала.

Ключевые слова: наноматериалы, наночастицы, базы данных, онтология, систематизация данных

* Работа выполнена при поддержке Российского научного фонда, грант № 14-50-00124.

СТАНОВЛЕНИЕ НОВОЙ ДИСЦИПЛИНЫ

Появление термина наноинформатика (далее – НИ) для новой области, порожденной потребностями нано- и информационных технологий (ИТ), можно отнести примерно к 2010 г., когда была собрана рабочая группа экспертов Nanoinformatics 2010, сформулировавших ее задачи и направления деятельности. Реализуемые проекты должны были учесть междисциплинарный характер предметной области и перманентное расширение понятийного аппарата, отражающего возникновение новых материалов, устройств и приложений. График проведения семинаров с кратким изложением обсуждаемых проблем можно найти на сайте (<http://nanoinformatics.org/2010/overview>). На текущий момент наиболее обширный и детализированный план разработан в рамках дорожной карты – Nanoinformatics 2020 Roadmap [1], принятой на семинаре 2015 г.

Заметное место в этом документе заняло уточнение самого понятия «наноинформатика». По мнению авторов [1], вполне аналогично с био- или экоинформатикой, наноинформатика призвана синтезировать методы сбора, обработки и распространения данных с учетом особенностей нанотехнологий, а именно: междисциплинарного характера, многофакторности описания материалов и устройств, вариации номенклатуры свойств при переходе к новым объектами т.п. [2, 3]. Тем самым, НИ не может рассматриваться просто как применение информатики к нанотехнологиям. Один из докладчиков на семинаре 2015 г. [4] заметил, что «специалист в НИ вынужденно является посредником между двумя областями и вынужден глубоко вникать в научную суть проблем, не ограничиваясь внешними атрибутами типа баз данных (БД), онтологий и проч».

Проведенный авторами [1] анализ выполненных и продолжающихся проектов, а также уже функционирующих средств (БД, порталов, цифровых библиотек и др.) позволил выделить несколько ключевых направлений: сбор и систематизация данных (*data curation*); средства для анализа и моделирования наноматериалов (далее – НМ); интеграция данных и распространение информации. При этом, хотя разработчики дорожной карты подчеркивают междисциплинарный характер наноинформатики, подавляющее большинство рассмотренных ими проектов ориентировано на медико-биологическую тематику, включая действие наноструктур на всех уровнях (молекулярном, клеточном и всего организма), связанные с этим проблемы токсикологии и гигиены, использование наноматериалов для диагностики и терапии и т.п. По-видимому, это связано с высоким уровнем информатизации экологии и здравоохранения¹, так что включение наноматериалов лишь незначительно расширило ранее созданную инфраструктуру. Например, на портале (<http://biportal.bioontology.org/>), содержащем наиболее обширную коллекцию из 517 биомедицинских онтологий, только 4 (**NanoParticle Ontology**, **eNanoMapper**, **InterNano**, **ChEBI**) явно

представляют структуру данных, характерную для НМ. Некоторые другие (например, *Medical Subject Headings – MESH*) лишь захватывают небольшое число терминов, связанных с так называемой наномедициной, т. е. проблемами использования наноматериалов для диагностики и терапии. Процесс естественного встраивания данных по НМ в давно сложившуюся инфраструктуру медико-биологической информатики, с одной стороны, определил быстрое оформление наноинформатики как новой дисциплины со своими БД, классификаторами, онтологиями и т.п., а с другой – привел к определенному тематическому «перекоосу», отразившемуся на подборе типовых наноструктур, номенклатуре свойств, выборе инструментов и технологий работы с данными. В итоге, большинство проектов, реализуемых, согласно дорожной карте [1], не учитывают особенности данных, присущих другим областям применения наноматериалов: электроника, машиностроение, энергетика и др. В частности, наноинформатика в приложении к медицине практически никогда не имеет дела с объемным (макроскопическим) материалом, когда наноразмерная специфика проявляется только в его внутренней или поверхностной структуре.

Это положение предопределило задачи настоящей работы: проанализировав основные результаты НИ применительно к потребностям ЕНС-сообщества (медицина, токсикология и т.п.), сопоставить их с менее известными информационными ресурсами, ориентированными на такие области, как электроника, машиностроение, энергетика и др.

Как и для медицины, еще до периода активного внедрения НМ, в этих отраслях возникла адекватная инфраструктура – специальный раздел информатики для материаловедения, так называемая *material informatics* [5], с характерными подходами к представлению данных. Эти подходы предполагают расширенное описание метода измерений, состояния образца, состояния среды, условий производства и т.п. Сложность идентификации, определяемой множеством структурных и технологических факторов, радикально отличает материал от обычных веществ, чьи свойства однозначно определяют химический состав и/или структурная формула. С некоторыми изменениями эти особенности материаловедческих данных относятся и к наноматериалам, что облегчает разработку информационных ресурсов, формируя направление, расширяющее границы, очерченные в документе [1].

Анализ текущего состояния наноинформатики проведен нами на примере ряда наиболее популярных ресурсов, созданных как для медико-биологической, так и для других областей применения. Основная цель анализа – выявить, насколько объем и структура данных, равно как и функциональные возможности ресурса, соответствуют особенностям наноматериалов: многофакторное описание с особым вниманием к генезису данных, зависимость номенклатуры свойств от класса НМ, перманентное обновление структуры данных при появлении новых материалов и устройств. Требования к полноте и качеству данных, обусловленные спецификой НМ, были детально

¹ В англоязычной литературе эту область называют *environment, health and safety (EHS) practice*.

изучены в работах [2, 3] и суммированы в рекомендациях международных организаций: CODATA – Комиссии по численным данным для науки и технологии (www.codata.org) и VAMAS – Версальского проекта по высокотехнологичным материалам и стандартам (www.vamas.org).

Второй аспект в развитии наноинформатики – это объективная потребность для нанотехнологий в переходе от использования автономных информационных ресурсов, отличающихся форматами, моделями данных и семантикой, к единой инфраструктуре. Организующим началом такой инфраструктуры является представленная в сети онтология, семантически точное и машинно-обрабатываемое определение сущностей и их соотношений. Онтология позволяет строго формализовать предметную область, обеспечивая для отдельных ресурсов (БД и др.) требования к схемам описания и семантическое единство для разнородных источников. При этом ее функция значительно шире обычной таксономии (классификации), так как поддерживает логические связи между понятиями, обусловленные спецификой предметной области.

Интеграция информационных ресурсов намечается во всех дисциплинах, где работа с данными эксперимента и/или моделирования, включающая их хранение, систематизацию и распространение, составляет основное содержание деятельности, получившей название e-science. В программе [1] эта линия в развитии наноинформатики была связана с созданием так называемой *мета-онтологии*, т. е. онтологии верхнего уровня, способной охватить объекты и концепции, возникшие в различных дисциплинах, включенных в использование и/или изучение наноматериалов. Развивая эту концепцию, создатели дорожной карты наметили две ключевые задачи. Первая сводится к интеграции уже действующих таксономий или онтологий, выполненных в разных коллективах для различных объектов и приложений (например, онтологии **NPO** – *NanoParticle Ontology for Cancer Nanotechnology Research*, www.nano-ontology.org, в основном созданной для информатики в сфере клинической онкологии). Вторая задача – это реализация требований к динамическому расширению мета-онтологии так, чтобы она могла поддерживать сложную эволюцию предметной области, обусловленную открытием новых НМ, проявлением новых эффектов и закономерностей, разработкой устройств и технологий.

Соответственно, одна из целей настоящей работы – определить, насколько сегодняшнее состояние наноинформатики отвечает намеченной линии на интеграцию ресурсов путем онтологического моделирования предметной области, а также выявить проблемы, препятствующие ее реализации и возможные пути их преодоления.

БАЗЫ ДАННЫХ ПО СВОЙСТВАМ НАНОМАТЕРИАЛОВ

В наноинформатике, как и в других видах e-science, базы данных, по-прежнему, являются основным элементом инфраструктуры, несмотря на появление новых технологий управления научными данными, использующих онтологии и средства Semantic

Web [6]. На сегодня большая часть данных хранится в реляционных БД. Именно их разработка инициирует создание многообразных классификаторов и таксономий, на основе которых проектируются концептуальные схемы БД, что вносит порядок в исходный массив данных с выделением базовых сущностей и типовых структур. При этом, помимо объема и качества хранимых данных, прикладная ценность БД определяется возможностью открытого доступа к ней через сеть Интернет. Краткое описание наиболее распространенных в мире БД по свойствам и использованию наноматериалов можно найти в обзорах [7, 8]. Характерно, что за редким исключением все они охватывают сферу EHS, т. е. медицины, производственной гигиены, экологии и т.п. Например, в сводке БД по нанотехнологиям [8] из 15 упомянутых ресурсов, только один – **InterNano** (www.internano.org) не ограничен рамками так называемой *EHS practice*.

Nanoparticle Information Library (NIL), <http://nanoparticlelibrary.net/> – один из простейших по типу и структуре данных ресурсов, разработанный Национальным институтом профгигиены США (NIOSH, www.cdc.gov/niosh). Основная цель ресурса – предоставление минимального объема данных, необходимого для профессионалов в области гигиены и токсикологии. Тип наночастицы выделяет всего четверка метаданных, определяющих структуру, состав, метод синтеза/происхождение, размер. По типу структуры выделено несколько классов наноматериалов: *agglomerated spheres*; *colloids*; *crystalline*; *films*; *nano horns*; *nano rods*; *nano tubes*; *nano wires*; *other*; *quantum dots*; *spherical*. Ограниченность и фиксированный характер структуры данных компенсируются частично включением в запись контекстной информации (аннотации, гиперссылки на родственные ресурсы, графики), которая дает сведения о свойствах, использовании НМ и его токсикологических характеристиках (*health and safety info*).

Более полный объем данных по свойствам наноматериалов, с упором на оценки биологического и экологического воздействия, предоставляет БД **Nanomaterial Registry** (www.nanomaterialregistry.org), которая построена и систематизирована в соответствии с некоторым стандартом **MIAN** (*minimal information about nanomaterials*), принятым для характеристики свойств. Набор физико-химических данных охватывает 12 ключевых показателей (табл. 1) и сопутствующую информацию для оценки качества и полноты данных.

Образец записи с данными для наночастицы из серебра приведен на рис. 1.

Стандарт **MIAN** охватывает физико-химические характеристики наноматериалов, наиболее актуальные для оценки их взаимодействия с биологическими объектами или средой обитания. Отличительная особенность регистра НМ как фактографической БД – предварительная обработка исходных данных (*data curation*) с последующим присвоением специального рейтинга, характеризующего их полноту и надежность – так называемый *compliance level (CL)*. Для этих целей стандарт **MIAN** предусматривает включение, помимо измеренных показателей, информации по методам измерений, протоколам, настройке при-

боров и т.п. Полноту представления данных в источнике раскрывает специальная анкета (*best practice questions*) с вопросами о наличии в источнике первичных данных, сведений о методах контроля, калибровке инструментов и т.д. На этой основе по

специальному алгоритму оценивается *CL* и, в зависимости от его величины, данным присваивается один из 4-х уровней качества: «золотой», «серебряный», «бронзовый» и «заслуженный».

Таблица 1

Физико-химические показатели наноматериалов, включаемые в БД [7-9]

№	Показатель	№	Показатель	№	Показатель
1	Composition	5	Surface area	9	Aggregation/agglomeration state
2	Particle size	6	Surface chemistry	10	Solubility
3	Size distribution	7	Surface charge	11	Stability
4	Shape	8	Purity	12	Surface reactivity

AS RECEIVED

Time Point: 0

Nanomaterial State: aerosol

Laboratory: CDC-NIOSH-Spokane Research Laboratory

Synthesis Method: incidental nanomaterial, furnace emissions - melting and refinement of silver

?

PHYSICO-CHEMICAL CHARACTERISTICS

BIOLOGICAL INTERACTIONS

ENVIRONMENTAL INTERACTIONS

Particle Size

Mean Primary Particle Size: 5 nm

Transmission Electron Microscopy

Aggregation/Agglomeration State

State: Agglomerated

Transmission Electron Microscopy

Size: 100 nm

Transmission Electron Microscopy

Shape

Nanoscale Dimensionality: 3D

Transmission Electron Microscopy

Composition

COATING

CORE

SHELL

● CORE (hover over circle for proportion of material)

COMPONENT 1

Material Type: Metal

Molecular Identity: Ag

Purity

Impurity Of: Nanomaterial: Pb

Impurity Of: Nanomaterial: Zn

Impurity Of: Nanomaterial: Ti

Impurity Of: Nanomaterial: Se

Рис. 1. Типовая запись БД Nanomaterial Registry

Зависимость свойств наноматериала от конкретных условий эксперимента побудил разработчиков БД отслеживать историю измерений, фиксируя время и характер окружения. Такая детализация достигается включением в запись отдельных элементов под названием Instance of Characterization (IOC). Их сопоставление позволяет сравнивать разные измерения, установив идентичность двух материалов при различных методах обработки и характеристики. При наличии ряда IOCs можно отследить разброс данных для серии образцов, пошаговые изменения в ходе обработки, охарактеризовать хронологию образца. Накопленная информация позволяет делать заключения о степени близости образцов, полученных от разных производителей или подвергнутых различной обработке.

Подробное описание отдельных функций регистра НМ и технологии подготовки данных можно найти в статьях [9, 10] и в документации на сайте nanomaterialregistry.org. В процессе разработки регистра достаточно полно учитывались особенности наноматериалов, отмеченные ранее в работах авторов [2, 11]: требование многофакторного описания, обязательность контекстной информации по синтезу и процедурам измерения, оценка качества данных в зависимости от их неопределенности и полноты представления в источнике.

Важным элементом в систематизации данных, принятых в регистре НМ, является их согласование с принятым стандартом и унификация терминологии посредством контролируемого словаря (www.nanomaterialregistry.com/resources/Glossary.aspx), в значительной степени согласованного с рекомендациями ISO (Международная организация по стандартизации, см. раздел «Классификаторы и стандарты»). В частности, согласно ISO, при выделении трех типов наночастиц по их размерности за основу принято число измерений в наномасштабе. Тогда графен и нанопленки попадают в категорию 1D, нанотрубки и наностержни – 2D (два наномасштаба в сечении) и, наконец, кластеры, наносферы, квантовые точки и т.п. относят к категории 3D. Заметим, что такая классификация противоречит общепринятой, когда к 1D-структурам относят нанотрубки, к 2D-структурам – нанопленки, а к 3D-структурам – макроскопические объекты. В этом случае за основу принято число макроскопических масштабов [2, 11, 12], которое может иметь четыре значения от 0 до 3. В то же время при классификации ISO число возможных наномасштабов достигает 3-х при том, что макроскопические объекты сведены в отдельную категорию – наноструктурированных материалов.

Несмотря на охват наноматериалов различных типов (квантовых точек, нанотрубок и проч.), принятые в NIL и регистре номенклатуры характеристик стандартизованы, что противоречит одной из ключевых особенностей НМ – вариации номенклатуры свойств и структуры данных в зависимости от технологии НМ. Без учета этой особенности невозможно представить с достаточной полнотой специфику отдельных типов наноматериалов, например, нанотрубок, масштаб которых определяет так называемый индекс хиральности [2, 11], а также совместить в одной БД информацию для наночастиц и объемных

материалов, характеризующихся внутренней наноструктурой. Нереализованным остался и общий «настрой», декларированный в планах развития наноинформатики [1], на согласование ресурсов с внешними онтологиями по НМ, например NPO, когда внесение соответствующей ссылки однозначно идентифицирует объект: NPO_606 соответствует углеродной нанотрубке, NPO_395 – наночастице с внедренной в ее структуру молекулой лекарства и т.п.

В обширном списке информационных ресурсов, приведенных в обзоре [8], только один – портал **InterNano** (Resources for Nanomanufacturing, www.internano.org) не ограничен медицинской тематикой. Документы на портале охватывают все аспекты научной и производственной деятельности, связанной с наноматериалами: процессы и средства производства; нано-объекты и наноструктурированные материалы; методы характеристики; медико-биологическое использование (*EHS practice*); социально-экономическое влияние НМ; информатика и стандарты; нормативно-правовая деятельность. Многообразные типы документов (тексты публикаций, доклады экспертов, пресс-релизы, библиографические ссылки и др.) систематизированы посредством трехуровневой иерархии (таксономии), каждому из терминов которой соответствует свой URL. Например, понятию *nanoparticles* отвечает URL www.internano.org/taxonomy/term/11987, а более узкому понятию *nanospheres* – URL с тем же синтаксисом и номером термина 12189. Ссылки на URL могут использоваться в других БД или онтологиях для точной и общепонятной идентификации НМ или определенной нанотехнологии, например, технологии *multilayer film process*, которой соответствует URL с термином 12101. В то же время, сам портал InterNano не может рассматриваться как база данных, поскольку каждому из метаданных таксономии соответствуют не выбранные таблицы свойств, а произвольно подобранная коллекция документов. Кроме того, несмотря на достаточную детализацию таксономии, она не содержит ссылок на данные по свойствам, даже простейшие – форма и размер.

Значительно расширить функциональность БД, способных адекватно учесть специфику наноматериалов, удалось в недавней нашей разработке [13]. Для поддержки усложненной и эволюционирующей структуры данных, присущей НМ, было предложено отказаться от традиционной технологии реляционных БД в пользу *цифровых библиотек* – качественно новой системе хранения и интеграции электронных ресурсов [14], способной обеспечить: многоаспектную систематизацию, адекватную структуре предметной области; динамический характер состава ресурсов и концептуальной схемы; интеграцию собственных документов с документами и/или данными глобальной сети. На этой основе был создан электронный ресурс по проблеме «НМ для энергетики» с целью систематизировать данные, охватывающие как сам наноматериал (его структуру, физические процессы и свойства, методы синтеза), так и сферу его приложений с детализацией секторов энергетики, технологий, оборудования и проч. Логическая структура данных отражена в системе из 25 метаданных/полей (табл. 2). В совокупности они позволяют хранить набор призна-

ков, определяющих источник (библиографическая информация), вид НМ и энергетическую технологию/оборудование, в сочетании с массивами текстовых, табличных и графических данных.

Поля табл. 2 разбиты на 3 блока: библиография (1-3, 14-20), наноматериалы (7-10), энергетическая технология и оборудование (4-6). Правила их заполнения в деталях описаны в работе [13]. Часть полей имеет форматы *select simple/multiple*, соответствующие выбору одного или нескольких понятий из априорно составленного списка. Обеспечить главную особенность данных – вариации структуры в зависимости от типа НМ или технологии, позволяет их дополнение полями текстового формата. Так, надежную идентификацию НМ обеспечивает сочетание полей 7 и 8: поле 8 указывает тип НМ ссылкой на позицию рубрикатора, а поле 7 вводит произвольное дополнение/уточнение, расширяющее границы рубрикатора НМ. Аналогично поля 4 и 5 выделяют энергетическую технологию посредством рубрикаторов, поле 6 – расширяет их возможности включением терминов, определяющих технологию или оборудование.

Предусмотрено также хранение произвольной по объему и структуре информации о физических и эксплуатационных свойствах материала, методах его синтеза и приложениях в полях с форматом *HTML*

area (поля 11, 12, 13). Здесь возможно использование произвольных шрифтов, графики, таблиц, математических выражений для многосторонней картины в представлении сведений о конкретном материале и технологии. Таким образом обеспечено сочетание детального структурирования данных с хранением неструктурированной информации, что является необходимым элементом в сфере нанотехнологий, где быстрая эволюция предметной области не позволяет ограничиться фиксированным перечнем понятий.

Ключевым элементом обсуждаемой систематизации является рубрикатор наноматериалов (табл. 3), ссылки на позиции которого даны в поле 8. Принятая классификация ориентирована на несколько признаков. В позициях с 1.0 до 1.15 наноструктуры разбиты соответственно размерности (0D, 1D, 2D, 3D), из которых последняя группа включает объемные/макроскопические материалы с характерной для них внутренней или поверхностной наноструктурой. Существенно обогатить классификацию позволяет использование, наряду с топологическим признаком, физических особенностей НМ и сферы его приложения. Например, в рубрике **Функциональные наноматериалы** позиции 3.8-3.12 выделяют материалы, ориентируясь на их характерные свойства, а 3.3, 3.4 или 3.13 – на возможное использование.

Таблица 2

Поля цифровой библиотеки «Наноматериалы для энергетики»

№	Название / Title	Поле для поиска / Searching field	Формат ввода / Input type/format
1	Record index	Yes	Text/Textarea
2	Record type	Yes	Select simple
3	Document type	Yes	Select simple
4	Energy sector	Yes	Select multiple
5	Energy function	Yes	Select multiple
6	Object	Yes	Text/Textarea
7	Nanomaterial [free title]	Yes	Text/Textarea
8	Nanomaterial by rubricator	Yes	Select multiple
9	Chemical [free title]	Yes	Text/Textarea
10	Chemical by rubricator	Yes	Select multiple
11	Synthesis		HTML area
12	Properties		HTML area
13	Application		HTML area
14	Authors	Yes	Text/Textarea
15	Title rus	Yes	Text/Textarea
16	Title orig	Yes	Text/Textarea
17	Source		Text/Textarea
18	Year	Yes	Data
19	Language	Yes	Select simple
20	Affiliation	Yes	Text/Textarea
21	Full text		Upload file
22	WEB source		External HTML
23	More information		HTML area
24	Abstract		HTML area
25	Comments		Hidden

Примечание: заливкой разного вида отмечены поля-классификаторы, поля в формате “HTML area” и поля для связи с внешними ресурсами

Классификатор наноматериалов

1.0	Наноструктуры
1.1	0D Нанокластеры
1.2	0D Нанокристаллы
1.3	0D Фуллерены, эндофуллерены и фуллереноподобные структуры
1.4	0D Квантовые точки
1.5	1D Нанопроволоки, наностержни и нановолокна
1.6	1D Нанотрубки
1.7	2D Наноструктурированные и нанокompозитные пленки
1.8	2D Нанопористые поверхности
1.9	2D Наномембраны
1.10	2D Графен и графеноподобные структуры
1.11	3D Наноструктурированные материалы
1.12	3D Нанокompозитные материалы
1.13	3D Нанопористые материалы
1.14	3D Нанопорошки
1.15	3D Нанодисперсии (коллоиды)
2.0	MISCL
2.1	Упорядоченные ансамбли (многослойные и многополосные структуры и сетки) одинаковых твердых элементов на подложках
2.7	Твердотельные гибридные и гетероструктуры на основе полупроводников, металлов и магнетиков
3.0	Функциональные наноматериалы (ФНМ)
3.1	Катализаторы на носителях
3.2	Интеркаляционные материалы и твердые электролиты
3.3	Сенсорные нанокompозиты
3.4	Водород-абсорбирующие наноматериалы
3.5	Наноструктурированные металлы и сплавы с особыми механическими свойствами
3.6	Наноструктурированные керамические и композиционные материалы и покрытия
3.7	Наноструктурированные полимеры, волокна и композиты на их основе
3.8	Слоистые магнитные материалы и сверхрешетки
3.9	Пьезоэлектрики
3.10	Сверхпроводники
3.11	Термоэлектрики
3.12	Люминесцентные
3.13	Бионаноматериалы
4.0	Конструкционные наноматериалы (КНМ)
4.1	Техническое железо и углеродистые стали
4.2	Легированные стали
4.3	Цветные металлы и сплавы на их основе
4.4	Тугоплавкие материалы и сплавы на их основе
4.5	Неметаллические материалы

Многообразие типов и форматов данных определяет богатые возможности построенной цифровой библиотеки по систематизации данных, различающихся объемом, логической структурой и средой представления. Потенциал цифровой библиотеки значительно усиливают поля, предназначенные для гиперссылок на внешние ресурсы: полные тексты документов (статей, отчетов и т.п.), размещенных на

сервере (поле 21) или в сети (поле 22). Использование контекстной информации (данные из справочников, обзоров и т.п.) обеспечивает интеграцию с тематически родственными ресурсами.

Всемирная паутина содержит немало источников, представляющих коммерческую и производственную информацию, например БД на портале (www.nanowerk.com): Nanomaterials; Nanotechnology

Companies & Laboratories; Nanotechnology Products and Applications и т. п. Из них только БД *Nanomaterials* включает минимальные сведения о продаваемых наноматериалах (название, размеры, физическое состояние), намного уступая по объему и достоверности информации научным БД. Сведения об отечественных поставщиках продукции нанотехнологического комплекса представлены в нескольких БД, зарегистрированных на Федеральном портале www.portalnano.ru, но не имеющих открытого доступа. Одна из таких БД по нанокompозитам детально описана в литературе [15]. Схема БД отражает иерархию «статья → сущность → свойства сущности», где верхний уровень включает библиографическое описание источника, промежуточный определяет нанокompозит указанием типа матрицы и наполнителя, и самый глубокий – сведения из источника о свойствах композита в текстовой и числовой форме. Это почти единственный пример БД по свойствам объемных наноматериалов, структура и содержание которой, к сожалению, недоступны для свободного поиска. Интересно заметить, что параллельно с БД была разработана и онтология нанокompозитов [16], позволившая отразить логическую структуру данных для проектирования аналогичных ресурсов.

КЛАССИФИКАТОРЫ И СТАНДАРТЫ

Стандартизация понятий и методов с самого появления нанотехнологий заняла важное место в их инфраструктуре, обеспечивая согласованность требований к характеристике материалов, параметрам приборов, безопасности производства и эксплуатации изделий и проч. [17]. К наноинформатике непосредственное отношение имеют стандарты, касающиеся терминологии и номенклатуры объектов и свойств. Их задача – обеспечить глобальный консенсус в отношении кодирования и классификации наноматериалов, номенклатуры их физико-химических и эксплуатационных характеристик, научной и производственной лексики и т.п., без чего невозможна интеграция БД и других видов информационных ресурсов.

Ключевой элемент в терминологической унификации – это так называемый, контролируемый словарь (*controlled vocabulary*), т. е. принятый в профессиональном сообществе перечень названий и терминов, снабженных соответствующей дефиницией. Ссылки на него позволяют исключить неоднозначность и синонимию, гарантируя адекватный и однозначный отклик на поисковый запрос. В разделе настоящей статьи «Базы данных по свойствам наноматериалов» уже приведены примеры использования таких словарей, принятых в БД *Nanomaterial Registry* и на портале *InterNano*. Эти примеры иллюстрируют возможные различия в характере ссылок на термины. Так, каждый из терминов таксономии *InterNano* кодируется своим URL, что в принципе допускает компьютерную реализацию поиска и анализа термина, в то время как в регистре НМ сетевой адрес выделяет словарь в целом, не допуская отсылки компьютера к отдельным терминам. Кодирование термина собственным URL используется также в

рубрикаторе Американского института физики (**AIP thesaurus**, <http://scitation.aip.org/content/topics>), один из разделов которого посвящен нанотехнологиям, и в двух отечественных ресурсах: словаре нанотехнологических терминов (<http://thesaurus.rusnano.com/>) и тезаурусе Федерального портала наноматериалов и нанотехнологий (www.portalnano.ru). С другой стороны, многоуровневый рубрикатор того же Федерального портала не допускает кодирования рубрик.

Наряду с кодированием терминов, возможности словаря определяются ориентацией и широтой, глубиной иерархической схемы, наличием дефиниций и гиперссылок и проч. Большинство словарей или таксономий, рассмотренных в публикациях по наноинформатике [8, 18], посвящено медицине или токсикологии. Например, **NCI Thesaurus** (*US National Cancer Institute*, <https://ncit.nci.nih.gov/ncitbrowser/>) включает таксономию наномедицинских терминов с дефинициями, ссылками на внешние ресурсы, уникальными кодами, URL и проч. В тоже время таксономический словарь **InterNano** охватывает все аспекты применения и производства НМ с присвоением термину индивидуального URL, но без дефиниций и комментариев. Авторы обзора [16] разделяют словари еще по одному критерию – статусу разработчика, противопоставляя национальные или международные органы стандартизации многочисленным научным или административным структурам, обладающим достаточным авторитетом, но не наделенным статусом *SDOs*². Списки организаций – разработчиков стандартов для нанотехнологий можно найти в работах [18, 19], причем по терминологии основная деятельность проведена **ISO** (*International Organization for Standardization*) и **ASTM International** (*American Society for Testing and Materials*). Разработанные стандарты, помимо функций словаря, определяют, что считать наноматериалом, какие свойства должны его отличать, как различать их отдельные категории.

Словарь, составленный **ISO** для разных секторов нанотехнологий [20], включает 9 действующих стандартов и 4 находящихся в разработке (см. табл. 4). Базовые определения и принципы классификации наноматериалов представлены в 1, 2 и 4 томах словаря. Термином верхнего уровня признан *nanomaterial*. Множество наноматериалов разбито на два класса: нанообъекты и наноструктурированные материалы. К первому отнесены объекты, имеющие наномасштаб для 1-, 2- или 3-х измерений, ко второму – макроскопические вещества с внутренней или поверхностной наноструктурой. Сильно упрощенная классификация наноматериалов сводится к трем положениям [19]:

- 1) НМ включает нанообъекты и наноструктурированные материалы;
- 2) нано-объекты включают наночастицы (3D), нановолокна (2D) и нанопластины (1D);
- 3) нановолокна включают наностержни (твердая форма) или нанотрубки (полая форма).

² Standards developing organizations

Отдельные тома словаря ISO [20]

Том	Название	Год утверждения
1	Core terms	2015
2	Nano-objects	2015
3	Carbon nano-objects	2010
4	Nanostructured materials	2011
5	Nano/bio interface	2011
6	Nano-object characterization	2013
7	Diagnostics and therapeutics for healthcare	2011
8	Nanomanufacturing processes	2013
9	Nano-enabled electrotechnical products and systems	*
10	Nano-enabled photonic components and systems	*
11	Nanolayer, nanocoating, nanofilm, and related terms	*
12	Quantum phenomena in nanotechnology	2016
13	Graphene and other two-dimensional materials	*

Примечание: * - том в стадии разработки

Таблица 5

Классифицирующие признаки наноматериалов [21]

Химическая природа / Chemical identity	Свойства / Properties	Происхождение / Origin
Ceramics	Physical	Natural
Metallic	Mechanical	Manufactured
Semiconductors	Chemical	Engineered
Polymer [natural/synthetic]	Biological	Incidental
Carbon based	Combined [cross-phenomena]	
Organics/Inorganics		

Тот же словарь ISO (том 4) выделяет всего 5 типов наноструктурированных материалов: наноструктурированный порошок, нанокомпозит, твердая нанопена, нанопористый материал, жидкая нанодисперсия. Более детальная классификация представлена в отчете ISO [21], где заметно расширено число типов, различающихся размерностью и структурой, а также предусмотрено их выделение по другим критериям, включая химическую идентичность, физические свойства, происхождение (табл. 5).

Тем не менее, даже при учете множества признаков, словарь не способен охватить все виды вновь синтезируемых наноматериалов. Например, предложены нано-объекты, включающие ковалентно связанные нанотрубки и листы графена, которым нельзя приписать определенную размерность [22]. Проблематично отнесение к какой-либо из рубрик классификатора и, так называемых, *эндоэдральных* структур [23], полученных внедрением в полость одной из форм объектов другой природы, например, внедрением молекул фуллерена C₆₀ или C₇₀ в полость углеродной нанотрубки. Трудности возникают также приписыванием НМ определенной химической природы, поскольку она может оказаться различной для ее фрагментов. Такую возможность учли составители регистра на-

номатериалов [9, 10], приняв неидентичность состава для ядра (*core*), оболочки (*shell*) и покрытия (*coating*) наночастицы (см. рис. 1).

Другая организация – **ASTM International**, столь же широко как **ISO** вовлеченная в разработку стандартов, подготовила документ E2456-06 «Terminology for Nanotechnology» [24]. Его рекомендации сфокусированы на свойствах наночастиц, а не на словаре, детализирующем классификацию наноматериалов. В частности, введено понятие о двух категориях наночастиц (*transitive* и *non-transitive*), определяемых характером изменения свойств при переходе к наномасштабу от объемного/макроскопического материала: для первой категории наблюдается скачок при переходе к наномасштабу, для второй этот же переход в свойствах происходит плавно. Примером *non-transitive* свойств может служить удельная площадь поверхности или оптическое рассеяние. Напротив, *confinement*, определяющий свойства квантовых точек или квантовых проволок, может служить примером *transitive* свойств и частиц, им обладающих. Еще одна терминологическая новация, введенная в документе E 2456-06 – понятие об *ультрананочастице*, которая определена как частица с размером, близким или менее 30 нм.

Множество документов, регламентирующих терминологию, предложено также организациями, не имеющими статуса *SDOs*, но достаточно авторитетными в определенной предметной области, ориентируясь на которую удастся сузить и даже зафиксировать номенклатуру нанообъектов и их свойств. Примерами таких организаций могут служить *National Cancer Institute* (США), который поддерживает набор биомедицинских словарей с охватом нанотехнологий (*Enterprise Vocabulary Services*, <http://evs.nci.nih.gov/>), или *EC Scientific Committee on Emerging and Newly Identified Health Risks* (**SCENIHR**). Как правило, их рекомендации не сводятся к терминологии, а дают собственные уточнения к базовым понятиям. Например, при выборе критерия отнесения структуры к нано-объектам многие из организаций выходят за рамки типовой границы в 100 нм для одного из измерений. Проведенный в обзоре [19] анализ показал, что ориентация на проблемы медицины, фармацевтики или пищевой промышленности заставляет экспертов поднимать эту границу до 300-500 или даже 1000 нм. В одном из документов **SCENIHR** [25] принадлежность к нанообъектам определяют не линейным размером, а удельной поверхностью при наименьшей ее границе в 60 м²/кг.

Широкое применение находят тезаурусы и классификаторы наноматериалов, также не имеющие статуса стандартов, но охватывающие множество областей применения. Образцом такого классификатора служит упомянутая выше таксономия портала **Inter-Nano**, созданного в рамках Национальной нанотехнологической сети США. Таксономия предназначена, в основном, для систематизации документов и ссылок, представленных на портале. При этом классификация самих наноматериалов составлена достаточно произвольно, т. е. без соблюдения каких-либо принципов рубрикации. Например, в рубрику *наноструктурированные материалы* не занесены *нанокompозиты*, напротив – *фуллерены* произвольно отнесены к разделу *нанотрубки*, а *феррофлюиды* квалифицированы как наночастицы и т.п. Аналогичный произвол отличает **AIP thesaurus** (<http://scitation.aip.org/content/topics>), принятый Американским институтом физики, где на одном уровне иерархии представлены наноматериалы самой разной природы, например нанокompозиты и кластеры. С другой стороны, графен (в отличие от графеновой пленки) вообще не занесен в раздел НМ, а такие объекты как квантовые точки или квантовые проволоки отнесены не к НМ, а к наноструктурам, объединяющим приборы и устройства.

Ранее мы показали [26], что практически всем рубрикаторам, не имеющим статуса стандарта, в той или иной степени присуща эклектичность со смешением на одном уровне иерархии признаков разного содержания. По-видимому, наиболее обоснованной выглядит систематизация наноматериалов, принятая на Федеральном портале www.portalnano.ru (см. табл. 6). Основные виды НМ здесь разбиты на 4 категории по размерности, в качестве которой принято число макроскопических масштабов: от нульмерных (все виды наночастиц и кластеров) до трехмерных, к которым отнесены объемные материалы. В отличие от схематизации ISO, здесь размерность – это число измере-

ний, вдоль которых имеет место макроскопический масштаб, а не число наномасштабных измерений; кроме того, термин *наноструктурированные материалы* относится ко всем видам НМ, а не только к объемным. К последним относят НМ, указанные в рубрике 1.1.4 (*трехмерные наноструктуры*), а также в рубриках 3.1 и 3.2 раздела *Продукция нанотехнологий*. При всех достоинствах рубрикатора его использованию для интеграции ресурсов препятствует отсутствие англоязычной записи и кодировки в виде URL для отдельных терминов.

Несмотря на высокий уровень детализации этого рубрикатора, превосходящего аналогичные продукты, например таксономию **InterNano**, он также не способен охватить все множество наноформ. Принципиальная недостижимость законченной классификации наноматериалов, обусловленная множественностью структурных и физико-химических факторов в сочетании с непрерывным пополнением их типов, отличает задачи, решаемые наноинформатикой, от аналогичных задач в систематизации объектов химии или материаловедения. Простейший и реально практикуемый подход к преодолению этих проблем – совместное использование универсальных стандартов с классификаторами или словарями, созданными организациями, не имеющими статуса *SDOs*, ориентируясь на которые можно сузить и даже временно зафиксировать номенклатуру объектов и свойств.

Относительно недавно на решение указанной проблемы наноинформатики был направлен международный проект под названием **Uniform Description System (UDS)**, реализованный совместными усилиями рабочей группы CODATA-VAMAS³ [27]. Созданная в итоге система метаданных призвана обеспечить: достаточную универсальность в представлении данных для разных типов НМ; эволюцию схемы описания с появлением новых объектов и понятий; возможность подстройки к отраслевым особенностям применения НМ (электроника, химическая индустрия, медицина и т.п.); разработку БД, онтологий и других информационных ресурсов. Подробное описание возможностей **UDS** как основы проектирования информационных ресурсов можно найти в работе авторов [28]. Здесь мы проанализируем ее возможности в качестве кандидата на роль международного стандарта для наноинформатики, сопоставляя со стандартами **ISO** и отраслевыми рекомендациями.

Прежде всего, в системе **UDS** своеобразно решен вопрос о формальной идентификации наноматериалов, с присвоением имен и включением в иерархию. Информационная категория под названием *General Identifiers* допускает два уровня определения НМ – произвольный и согласованный с действующими стандартами (*assigned by an authority*), что предоставляет необходимую свободу, требуемую при настройке на предметную область. Например, ссылка на коды **NCI Thesaurus** (<https://ncit.nci.nih.gov/ncitbrowser/>) обеспечивает включение в созданную таксономию

³ CODATA – комиссия по численным данным для науки и технологии (www.codata.org), VAMAS – Версальский проект по высокотехнологичным материалам и стандартам (www.vamas.org).

большинства объектов и концепций наномедицины. Для других типов НМ такую же задачу (кодирование имени и включение в иерархию) может обеспечить ссылка на химическую БД **ChEBI** (*Chemical Entities of Biological Interest*, www.ebi.ac.uk/chebi), в которой НМ представлены аналогично обычным соединениям. Существенно, что код в обоих классификаторах является частью URL (например, www.ebi.ac.uk/chebi/searchId.do?chebiId=50594 для CNT), допуская компьютерную реализацию поиска термина и приписанных к нему данных. Аналогичную ссылку на стандарт **ISO** [20] с указанием номеров тома и термина трудно связать с сетевым номером, поскольку доступ в сети открыт лишь к информационной части стандарта.

Проект UDS внес еще одну новацию в сравнении со стандартом **ISO** [20]. В то время как **ISO** выделяет всего две категории наноматериалов (нано-объекты и наноструктурированные материалы), то, согласно **UDS**, НМ включают три класса: нано-объекты, их ансамбли/коллекции, объемные материалы. Понятие **ансамбля** позволяет охватить объекты сложного состава, наподобие структур типа «графен-нанотрубки» [22], которые не укладывались в систематику **ISO**. При этом ансамбль нанообъектов, как и сам нано-объект имеет от одного до трех размеров в наномасштабе. Другое изменение относится к объемным/макроскопическим материалам: они разбиты на 2 суб-

класса – материалы, содержащие идентифицируемые нанообъекты, и материалы, лишь проявляющие размерный эффект за счет внутренней или поверхностной наноструктуры. Примером первого являются наноконкомпозиты (скажем, полимер с включениями углеродных нанотрубок), примером второго – наноструктурированные стали или керамика.

В работе [28] мы провели детальный анализ иерархии, верхний уровень которой показан на рис. 2. Уже начальная категория – *Идентификация* (обеспечивает гибкость в систематике НМ, адаптируя ее к разным областям применения со своими классификаторами. Наиболее полную идентификацию дает информационная категория *Характеристики/свойства*, охватывающая все данные о свойствах НМ: форма и размеры, химический состав, внутренняя и поверхностная структура и проч. Именно их совокупность должна обеспечить однозначное выделение описываемого объекта из множества других с родственной структурой. Две последних категории (*Производство* и *Спецификация*) относятся к производству и поставкам материала на рынок. Сопоставляя номенклатуру свойств с используемой в ряде БД, например, включенной в регистр НМ в соответствии со стандартом **MIAN** (табл. 1), легко видеть их относительную близость (форма, размер, химический состав, состояние поверхности и др.) при значительно большей детализации, предусмотренной в **UDS**.

Таблица 6

**Классификация наноматериалов на Федеральном портале
«Нанотехнологии и наноматериалы» (верхние уровни иерархии)**

1	Объекты, относящиеся к сфере нанотехнологий, их свойства	3.	Продукция нанотехнологий		
1.1	Наноструктурированные материалы	3.1	Функциональные наноматериалы	3.2	Конструкционные наноматериалы
1.1.1	Нульмерные структуры	3.1.1	Катализаторы на носителях	3.2.1	Техническое железо и углеродистые стали
1.1.2	Линейные наноструктуры	3.1.2	Интеркаляционные материалы и твердые электролиты	3.2.2	Легированные стали
1.1.3	Двумерные наноструктуры	3.1.3	Сенсорные наноконкомпозиты	3.2.3	Медь и сплавы на ее основе
1.1.4	Трехмерные наноструктуры	3.1.4	Водород-абсорбирующие наноматериалы	3.2.7	Тугоплавкие материалы и сплавы на их основе
1.1.4.1	Объемные наноструктурированные материалы	3.1.5	Наноструктурированные металлы и сплавы с особыми механическими свойствами	3.2.9	Неметаллические материалы
1.1.4.2	Наноконкомпозитные материалы				
1.1.4.3	Нанопористые материалы				
1.1.4.4	Нанофлюиды (дисперсии)				



Рис. 2. Верхние уровни иерархии понятий в системе UDS [27]
(НМ – наноматериалы)

Помимо большего числа самих категорий, и различий в иерархии каждого из классов (нано-объекты, ансамбли, объемные материалы), глубокая детализация свойств обеспечена за счет разбиения каждой из категорий на дескрипторы, нижний уровень иерархии, заполняемый экспертом. Именно совокупность этих свойств, способных учесть тонкие детали структуры, обеспечивает многофакторный характер описания, требуемого при аттестации наноматериала [2, 3, 11]. Кроме того, добавляя или удаляя отдельные из дескрипторов, можно отразить эволюцию схемы описания, связанную с *априори* неизвестными объектами.

Для ансамбля нанообъектов идентификация включает его состав, т. е. типы и число каждого из элементов ансамбля с характеристикой каждого из элементов по той же схеме, что и для индивидуальных нанообъектов. Что же касается объемных наноматериалов, то крайнее многообразие их видов заставило разработчиков [27] отказаться от построения иерархии с детализацией свойств, ограничившись лишь выделением двух типов: Bulk Materials with Individual Nano-objects (тип 1) или Bulk Materials with Nanoscale Features (тип 2). Возможные пути развития UDS для макроскопических материалов с использованием существующих онтологий в области материаловедения рассмотрены в нашей работе [28]. В целом система UDS в сочетании с привлекаемыми словарями и таксономиями позволяет приблизиться к решению главной проблемы наноинформатики – созданию многофакторной и эволюционирующей концептуальной схемы данных.

ОНТОЛОГИИ ПО НАНОМАТЕРИАЛАМ

Онтологии способны обеспечить один из наиболее эффективных механизмов интеграции научных данных за счет сочетания предметно-ориентированного словаря с логической структурой, аналогичной концептуальной схеме БД. Словарь в виде многоуровневой таксономии обеспечивает стандартизацию понятий для целей аннотирования и поиска документов.

Логическая структура данных передается посредством использования многочисленных атрибутов, связанных ассоциативными соотношениями (типа **part_of**, **has_part**, **has_quality** и т.п.), и аксиом. Тем самым, помимо терминологического ресурса, онтология способна передать формализованную структуру знания. Среди возможных применений онтологического моделирования предметной области – проектирование БД, интеграция гетерогенных источников, запись документов в RDF-формате с их последующим включением в глобальное пространство связанных данных [29, 30]. Использование онтологий особенно актуально применительно к наноматериалам, учитывая междисциплинарный характер концепций и методов, неустойчивость терминологии и понятийного аппарата, перманентное открытие новых материалов, устройств и технологий. Особенно полезной для наноинформатики оказывается предусмотренная в онтологической инженерии возможность заимствования классов и терминов из родственных онтологий, что широко применяется, например, на био-портале (<http://bioportal.bioontology.org/>), где интегрирована обширная коллекция медико-биологических онтологий.

Одна из наиболее простых онтологий по НМ включена в **ChEBI** (*Chemical Entities of Biological Interest*) – БД и структурированный классификатор, сфокусированный на малых молекулах с исключением белков и нуклеиновых кислот [31]. По сути, это онтология химического содержания, в которой любая из нанопформ трактуется как молекулярная сущность, не отличающаяся от обычных молекул или молекулярных агрегатов. Подбор наноструктур и их систематизация в **ChEBI** не связаны с определенной дисциплиной или приложением, что облегчает заимствование ее терминов. Образец записи в БД ChEBI (рис. 3) показывает элемент иерархии, связанной с объектом *carbon nanotube*, а именно, дочерние и родительские классы, а также сущности, связанные по типу *часть-целое*.

ChEBI
The database and ontology of Chemical Entities of Biological Interest

EBI > Databases > Small Molecules > ChEBI > Main

carbon nanotube (CHEBI:50594)

Main | ChEBI Ontology | Automatic Xrefs

ChEBI Name ? carbon nanotube

ChEBI ID ? CHEBI:50594

Definition ? A molecule consisting of a graphene cylinder or two or more concentric graphene cylinders.

Stars ? ★★☆☆ This entity has been manually annotated by the ChEBI Team.

ChEBI Ontology ?

Outgoing

- carbon nanotube (CHEBI:50594) is a carbon nanostructure (CHEBI:50812)
- carbon nanotube (CHEBI:50594) is a nanotube (CHEBI:50796)

Incoming

- carbon nanorod (CHEBI:50807) has part carbon nanotube (CHEBI:50594)
- carbon nanorope (CHEBI:50813) has part carbon nanotube (CHEBI:50594)
- carbon nanotubosome (CHEBI:50806) has part carbon nanotube (CHEBI:50594)
- double-walled carbon nanotube (CHEBI:50797) is a carbon nanotube (CHEBI:50594)
- multi-walled carbon nanotube (CHEBI:50596) is a carbon nanotube (CHEBI:50594)
- single-walled carbon nanotube (CHEBI:50595) is a carbon nanotube (CHEBI:50594)

Synonyms ?

Рис. 3. Типовая запись БД ChEBI, отражающая положение класса carbon nanotube в онтологии

К последним отнесены три типа наноструктур: *carbon nanorod*, *carbon nanorope* и *carbon nanotubosome*. Тем самым фиксируются ансамбли нанообъектов (например, *nanorope* как ансамбль составленный из *nanotube*). Поскольку ChEBI, прежде всего, химическая онтология, то многие из нанообъектов рассматриваются как дочерние и по отношению к определенному классу химических форм, что расширяет определение объекта сведениями о его химической природе. Например, класс *carbon nanotube*, являясь дочерним по отношению к *nanotube*, наследует также линию органических соединений (*organic molecular entity*).

Каждому из объектов в онтологии дается строгая дефиниция и присваивается 5-значный ID (CHEBI:*****), ссылка на который в произвольном документе (или БД) дает однозначную и доступную для компьютера идентификацию. Сверх этого предусмотрено использование объектного свойства **has_role**, связывающего один из объектов таксономии с одним из классов суб-онтологии **role**, который определяет возможное приложение, а также химическую и биологическую роли объекта. Однако все ролевые сущности в ChEBI сосредоточены на обычных веществах, но не нашли применения для наноструктур. В то же время, принцип построения онтологии допускает ее естественный перенос на область нанотехнологий путем аналогичного перечисления сущностей, определяющих приложения и роли нанообъектов. Тогда структура ChEBI могла бы рассматриваться как возможный прообраз при проектировании более обширной онтологии по НМ, способной к характеристике свойств и применений. Эту же идею высказал К. Батчелор в докладе на сессии Американского химического общества [32], предложив совместить ChEBI с

малой онтологией, способной охватить свойства наноматериалов, чтобы обеспечить адекватное представление данных для произвольной наноструктуры⁴.

Значительно более полной и детализированной является онтология NPO, *NanoParticle Ontology for cancer nanotechnology researches* [33], включающая 1900 классов, распределенных по множеству уровней иерархии (вплоть до 16) и связанных 80 ассоциативными соотношениями, конкретизирующими физико-химические свойства и возможное использование наночастиц. Однако, будучи жестко ориентированной на узкий сегмент медицины, онтология в большей степени передает медико-биологическую проблематику, чем традиционные аспекты нанотехнологий. Это проявляется в доминировании соответствующих терминов (в том числе заимствованных из онтологий, собранных на биопортале), в явном выделении связей, определяющих биологическую роль наночастицы и включенных в нее химических агентов, внешних стимулов, активирующих ее действие, и других аспектов диагностики или терапии. Типовые характеристики самой наночастицы включают данные по структуре, форме, размеру, физическому состоянию (гидрогель, эмульсия и т.п.). В свою очередь химический компонент определяется его физическим положением в наночастице, молекулярной структурой, физическими, химическими и функциональными свойствами. Сверх этого включенные в онтологию понятия определяют механизм действия наночастицы, внешние стимулы, активирующие ее функцию (магнитное по

⁴ “With a small ontology of properties, and ChEBI to provide the chemistry, we can now generate arbitrary nanoparticle representation...” [29].

ле, ультразвук, pH), реакцию на действующий стимул (например, адресная доставка препарата наночастицей в ответ на включение магнитного поля). Каждая из названных категорий порождает целую иерархию взаимосвязанных понятий. В частности, онтология задает базовые характеристики наночастицы (размер, форма, масса, поверхность), детализируя отдельные элементы частицы данными по составу и другими характеристикам.

Все разновидности наноструктур сосредоточены в двух классах (*nanomaterial* и *nanoparticle*), оба из которых находятся на уровнях иерархии, порожденной классом *chemical entity*, рис. 4. Систематика наноструктур, диктуемая принятой в **NPO** иерархией классов достаточно сильно расходится с уже принятыми классификациями – к примеру класс *nanoparticle* должен был рассматриваться как дочерний по отношению к категории *nano-object* [20, 21].

Помимо наноматериалов, термины верхнего уровня, порождающие иерархию, включают молекулярные структуры, клеточные компоненты, позиции и границы в материальных сущностях (ядро, оболочка и т.п.), свойства, роли молекулярных компонентов, стимулы функций наночастиц и реакции наночастиц на стимулы, биологические процессы, химические взаимодействия и др. В то же время полные списки терминов и ассоциативных свойств на сайте (www.nano-ontology.org) допускают возможность обобщения на различные сферы приложения наноструктур, например для детализации ее типа, геометрии, химического состава и других характеристик. Каждому из терминов, включенных в онтоло-

гию, присвоен уникальный трех- или четырехзначный ID, например **NPO_126** (*nanotube*) **NPO_586** (*quantum dot*). В отличие от **ChEBI**, здесь посредством ID можно организовать ссылки на множество понятий, отражающих свойства, технологии или использование наноструктур, например: **NPO_1344** (*chemical composition*), **NPO_1445** (*atomic force microscopy*), **NPO_1344** (*emission*) и т.п. Тем самым терминологический базис **NPO** может довольно широко использоваться при построении альтернативных онтологий по наноматериалам, не связанных с медико-биологической проблематикой.

Наиболее существенное ограничение онтологии **NPO** (как и рассмотренной выше **ChEBI**) связано с тем, что она охватывает лишь такие наноформы, которые можно уподобить молекулярным сущностям (наночастицы, нанотрубки, наностержни и т.п.). За бортом остаются наиболее важные в промышленных приложениях объемные НМ (*bulk nanomaterials*), к которым по определению **ISO** [20, 21] относят макроскопические объекты (твердые тела, порошки, дисперсии), внутреннее строение которых или их поверхность отличает наличие наноструктурных единиц.

Помимо **NPO** и **ChEBI**, двух наиболее разработанных и цитируемых онтологий, в последнее время появилось еще несколько, отражающих ассортимент и характеристики нанообъектов. Одна из них, также сильно ориентированная на *EHS*-тематику, разработана в рамках Европейского проекта интеграции токсикологических данных для наноматериалов (**eNanoMapper**, www.enanomapper.net).

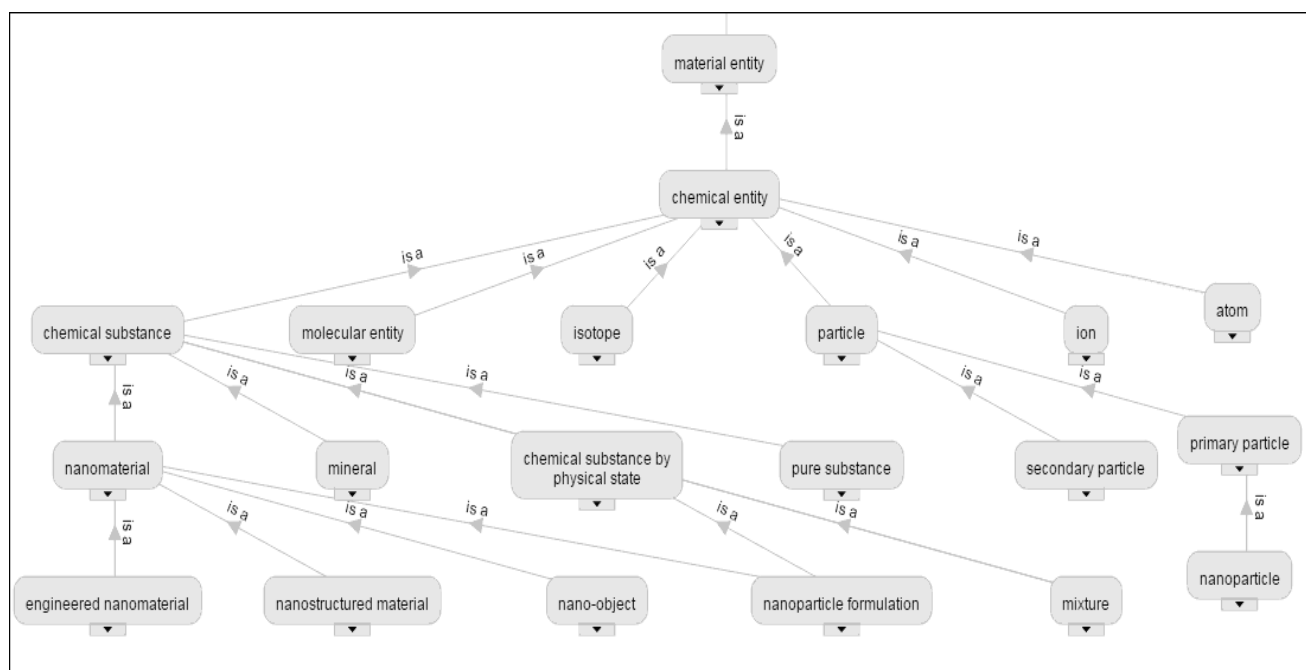


Рис. 4. Иерархия классов NanoParticle Ontology, относящихся к наноматериалам и наночастицам

Созданная для этих целей инфраструктура следует стандартам и технологиям Semantic Web [29, 30] с использованием онтологии как базового компонента. На ее основе проведена систематизация различных концепций, относящихся к биологическому эксперименту, физико-химической и экологической аттестации наночастиц, молекулярным и биологическим сущностям, включенным в оценку риска. В значительной степени использованные термины заимствованы из других онтологий, частично перекрывающих ту же тематику. Что же касается собственно наноструктур, то разработчики **eNanoMapper** импортировали их названия и физико-химические характеристики из **NPO** и **ChEBI**. Термины, относящиеся к наноструктурам, включены в иерархию, порожденную двумя классами верхнего уровня: *nanomaterial* и *nanoparticle*. Аналогично построена и их физико-химическая характеристика. Включение **eNanoMapper** в коллекцию онтологий на биопортале облегчает процесс ее интеграции, как с **NPO** и **ChEBI**, так и с обширным материалом из медико-биологических онтологий.

Онтологический подход удалось применить и к одной из ключевых особенностей наночастиц – ее форме, используя ее как основу построения таксономии [4, 34]. Различные типовые 2D и 3D структуры представлены как формальные классы, образующие многоуровневую иерархию. Тогда появляется возможность детализировать форму и морфологические особенности объекта с полнотой достаточной, чтобы оценить его функциональность. При этом, помимо геометрической таксономии, удобной при систематизации, была использована специальная техника обработки изображений, позволяющая проводить полуавтоматическую идентификацию полученных при микроскопии изображений.

Как видно, ни одна из рассмотренных здесь онтологий, начиная с **ChEBI**, по существу не способна к адекватному представлению объемных/макроскопических наноматериалов из-за крайней сложности и многообразия их типологии и структуры. В онтологиях **NPO** и **eNanoMapper** на верхнем уровне выделены две категории *nanomaterial* и *nanoparticle*, однако их содержание не соответствует принятому, согласно **ISO** [20], разделению мира НМ на нанобъекты (аналоги наночастиц) и наноструктурированные, иначе говоря, объемные наноматериалы. В рамках **NPO** класс *nanomaterial* охватывает только один, достаточно специфичный вид объемного материала – *nanoparticle formulation (NPO_868)*, вещество в виде порошка (или эмульсии), приготовленное с использованием нанотехнологий и содержащее наночастицы. Что касается класса *nanosubstructured material*, то в **NPO** он включает всего две структуры – *nanobud* (нанотрубка, снаружи связанная с фуллереном) и *nanofilms*, обе из которых, согласно классификации **ISO**, относятся к нанобъектам, ничуть не напоминая те, что подразумевались в стандартах, скажем, нанопористая или наноструктурированная керамика. Более того, с задачей адекватного представления объемных наноматериалов не удалось справиться и разработчикам **UDS**, невзирая на объявленную универсальность стандарта.

На сегодняшний день единственный образец такой онтологии для нанокompозитных материалов был создан в РХТУ им. Д.И. Менделеева [16]. Их свойства и функциональность определяются выбором матрицы (металл, керамика и т.п.) и наполнителя, в качестве которого используются нанобъекты. Заметим, что, согласно стандарту **UDS** [27], нанокompозит относится к объемным материалам с идентифицируемыми, т. е. выделяемыми в явном виде, наноструктурами. Основу онтологии составляет таксономия классов и экземпляров, содержащая такие классы, как *MaterialType*, *NanoObject*, *Nanocomposite*, а также классы, включающие химические характеристики: *ChemicalIdentity*, *ChainComposition*, *Structure* и т.д. Класс *MaterialType* имеем 5 дочерних классов (*CarbonBased*, *Ceramic*, *Metallic*, *Polymer*, *SiliconBased*), что в сочетании с довольно простой таксономией нанобъектов (*NanoFiber*, *NanoFilm*, *NanoLayer*, *NanoPowder*, *NanoSurface*) позволяет отражать формирование разных типов нанокompозита, определяемого выбором матрицы и наполнителя. К сожалению, доступно лишь краткое описание онтологии, сам OWL-файл не представлен на публичной платформе.

Кроме нанокompозитов, есть другая обширная категория материалов, включающих идентифицируемые нанобъекты, которые состоят из элементарных «нано-блоков» без разделения материала на матрицу и наполнитель. В качестве примера можно привести полимеризованные формы фуллерита, образованные за счет Ван дер Ваальсового взаимодействия молекул C_{60} или C_{70} . Другим образцом подобного НМ может служить сверхлегкий аэрогель [35] – монолитная структура, составленная из графеновых лент и углеродных нанотрубок. Следуя классификационным принципам, предложенным в [27], материал такого рода следует отнести к ансамблю (или коллекции) нанобъектов, хотя исходно не предполагалось отнесение к этой категории объектов макроскопического масштаба.

Наконец, существует обширный класс объемных наноматериалов, в которых невозможно выделить идентифицируемые нанобъекты: стали, керамика, полимеры, углеродные материалы и множество других. Приставка «нано» означает для них лишь наличие внутренней или поверхностной наномасштабной структуры. В работе [28] мы предложили достаточно естественный путь разработки онтологии, сочетая общую схему стандарта **UDS** с существующими онтологиями в области материаловедения.

ЗАКЛЮЧЕНИЕ

Результаты нашего исследования позволяют ответить на поставленный в работе вопрос: насколько сегодняшнее состояние наноинформатики отвечает объективным потребностям нанотехнологий, включая все виды используемых и вновь синтезируемых наноматериалов, возникшим на их основе приборам и технологиям, методам производства, анализа и моделирования наноразмерных объектов. Состав и возможности разработанных средств (БД, онтологий и др.) сопоставлены в работе с теми требованиями, которые возникают перед создаваемой информаци-

онной инфраструктурой, и отражают особенности возникшей системы знаний: ее междисциплинарный характер, перманентную эволюцию структуры данных, многофакторное описание при особом внимании к генезису данных и т.д. [2, 3, 11].

Подавляющее большинство средств и технологий разработано, прежде всего, для медико-биологической сферы, где они были интегрированы с давно развитой системой медицинской информатики, например со словарями и онтологиями, размещенными на биопортале (<http://bioportal.bioontology.org/>). Применительно к этим задачам особенности НМ учитывались, прежде всего, в процедурах *data curation* [36]. Наиболее полно эта технология внедрена в БД **Nanomaterial Registry** [9, 10], где предусмотрены специальные методики сопровождения данных информацией по методам измерений, протоколам, настройке приборов, изменениям материала в ходе обработки и т.п. Наряду с этим, внимание уделялось возможностям интегрирования ресурсов. Например, был разработан специальный формат **ISA-TAB-Nano** [8, 18], включающий файлы в виде электронных таблиц, структура которых была специально адаптирована к хранению и обмену результатами медико-биологических анализов, в том числе включающих информацию по наноматериалам. Вместе с тем, логическая структура большинства ресурсов была выдержана в стиле реляционной БД с жестко фиксированной номенклатурой параметров, как, например, в стандарте **MIAN**, табл. 1.

Выход за рамки этого ограничения, т. е. попытки внедрить гибкую структуру данных, зависящую от класса материалов и сферы приложения, можно наблюдать одновременно с переходом от медицинской тематики к другим областям использования наноматериалов, например, к энергетике. Один из простейших приемов подстройки логической структуры к типологии НМ был применен нами в работе [13], где совокупность рубрикаторов дополнялась свободным использованием ключевых слов. Богатые возможности развертывания сложной структуры данных предоставляют разработанные для НМ онтологии, прежде всего **ChEBI** и **NPO**. Иерархия классов, включенных в онтологию, охватывает не только сами материалы, но и их физические свойства, области применения и т.п., что позволяет на уровне отдельных экземпляров обеспечивать «привязку» требуемых атрибутов к каждому из объектов, например, характеризовать размер нанотрубок иными параметрами, чем для фуллерена. Кроме того, посредством онтологии достаточно легко поддерживать эволюцию структуры данных, расширяя или модернизируя иерархию классов.

Наибольший потенциал развития наноинформатики с охватом любой из возможных сфер приложения, по-видимому, способен обеспечить недавно предложенный стандарт **UDS** [27]. Он закладывает основу логической структуры, включающей многофакторное описание всей совокупности физических и эксплуатационных характеристик, а также стадий производства и последующей работы с наноматериалами. При этом предусмотрены возможности детализации каждой из категорий стандарта, использования разнообразных классификаторов и словарей, адаптации к разным приложениям или областям зна-

ния. Можно сказать, что именно стандарт **UDS** открывает путь к решению двух главных проблем наноинформатики: требование гибкой структуры данных, адаптируемой к типам материалов и технологий, и динамическое расширение так называемой *мета-онтологии*, чтобы она могла поддерживать эволюцию предметной области. В наших работах [28, 37] была сформулирована возможная стратегия проектирования онтологии наноматериалов, в том числе объемных/макроскопических.

СПИСОК ЛИТЕРАТУРЫ

1. Nanoinformatics 2020 Roadmap. April 2011. Published by the National Nanomanufacturing Network. Amherst, MA 01003. – URL: <http://eprints.internano.org/607/>
2. Eletskii A.V., Erkimbaev A.O., Kobzev G.A., Trachtengerts M.S., Zitserman V.Y. Properties of nanostructures: data acquisition, categorization, and evaluation // Data Science Journal. – 2012. – Vol. 11. – P. 126-139.
3. Rumble J., Freiman St. Describing Nanomaterials: Developing a multi-disciplinary Framework // Nanoinformatics 2015 Workshop. – URL: <http://nanoinformatics.org/2013/agenda>
4. Maojo V. Towards a visual taxonomy of nanoparticles. Enabling Successful Discovery and Applications. Nanoinformatics 2015 Workshop. – URL: <http://nanoinformatics.org/2015/agenda>
5. Agrawal A., Choudhary A. Perspective: Materials informatics and big data: Realization of the “fourth paradigm” of science in materials science // APL Mater. – 2016. – Vol. 4. – Art# 053208.
6. Karagiannis F., Keramida D., Ioannidis Y. et al. Technological and organizational aspects of global research data infrastructures towards 2020 // Data Science Journal. – 2013. – Vol. 12. – P. GRDI1-GRDI5.
7. Thomas D.G. et al. Informatics and Standards for Nanomedicine Technology // Wiley Interdiscip Rev Nanomed Nanobiotechnol. – 2011. – Vol. 3, № 5. – P. 511-532.
8. Panneerselvam S., Choi S. Nanoinformatics: Emerging Databases and Available Tools // Int. J. Mol. Sci. – 2014. – Vol. 15. – P. 7158-7182.
9. Mills K.C., Murry D., Guzan K.A., Osteraat M.L. Nanomaterial registry: database that captures the minimal information about nanomaterial physico-chemical characteristics // J. Nanopart. Res. – 2014. – Vol. 16. – P. 2219.
10. Guzan K.A., Mills K.C., Gupta V. et al. Integration of Data: the Nanomaterial Registry project and data curation // Computational Science & Discovery. – 2013. – Vol. 6. – Art# 014007.
11. Елецкий А.В., Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А., Трахтенгерц М.С. Теплофизические свойства наноразмерных объектов: систематизация и оценка достоверности данных // Теплофизика высоких температур. – 2012. – Т. 50, №4. – С. 524-532.

12. Pokropivny V.V., Skorokhod V.V. New dimensionality classifications of nanostructures // *Physica E*. – 2008. – Vol. 40, № 7. – P. 2521-2525.
13. Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А., Трахтенгерц М.С. Цифровая библиотека вместо традиционной базы данных для нанотехнологий: опыт использования системы ABCD // *Научно-техническая информация. Сер. 2*. – 2014. – №6. – С. 12-22; Erkimbayev A.O., Zitserman V.Yu., Kobzev G.A., Trakhtenherz M.S. A Digital Library instead of a Traditional Database for Nanotechnologies: An Attempt to Use the ABCD System // *Automatic Documentation and Mathematical Linguistics*. – 2014. – Vol. 48, № 4. – P.212-223.
14. Когаловский М.Р., Новиков Б.А. Электронные библиотеки – новый класс информационных систем // *Программирование*. – 2000. – №3. – С. 3-8.
15. Порысева Е.А., Краснов А.А., Кольцова Э.М. Создание и разработка информационной системы в области нанокomпозиционных материалов // *Успехи в химии и химической технологии*. – 2012. – Т. XXVI, №11(140). – С. 38-41.
16. Карпова И., Порысева Е., Казаков Г., Кольцова Э. Разработка онтологии в области нанокomпозиционных материалов // *Информационные ресурсы России*. – 2012. – №2. – С. 5-9.
17. Nanotechnology Standards, Nanostructure Science and Technology / eds. V. Murashov, J. Howard. – New York: Springer. – 2011. – 262p.
18. D de la Iglesia, Cachau R.E., Garcia-Remesal M., Maojo V. Nanoinformatics knowledge infrastructures: bringing efficient information management to nanomedical research // *Computational Science & Discovery*. – 2013. – Vol. 6. – Art# 014011.
19. Klaessig F., Marrapese M., Abe S. Current Perspectives in Nanotechnology Terminology and Nomenclature // В монографии “Nanotechnology Standards” / eds. V. Murashov, J. Howard. – New York: Springer – 2011. – P.21-52.
20. International Organization of Standardization: Nanotechnologies – Vocabulary, ISO/TS 80004: 2010-2016. – URL: www.iso.org/obp/ui/#iso:std:iso:ts:80004:-1:ed-1:v1:en
21. ISO/TR 11360:2010. Nanotechnologies- Methodology for the classification and categorization of nanomaterials.
22. Чернозатонский Л.А., Шека Е.Ф., Артюх А.А. Графен-нанотрубные структуры: строение и энергетика образования // *Письма в ЖЭТФ*. – 2009. – Т. 89, №7. – С. 412-417.
23. Елецкий А.В. Эндоэдральные структуры // *УФН*. – 2000. – Т. 170, №2. – С. 114-142.
24. ASTM International: E 2456-06 Terminology for Nanotechnology. ASTM International, West Conshohocken, USA (2008).
25. EC Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR). Opinion on the scientific aspects of the existing and proposed definitions relating to products of nanoscience and nanotechnologies. The SCENIHR adopted this opinion at the 21st plenary on 29 November 2007. – URL: http://ec.europa.eu/health/archive/ph_risk/committees/04_scenihr/docs/scenihr_o_012.pdf
26. Буйлова Н.М., Елецкий А.В., Зицерман В.Ю., Кобзев Г.А. Систематизация источников и данных по нанотехнологиям // *Научно-техническая информация. Сер. 1*. – 2013. – №11. – С. 31-42; Builova N.M., Yeletskii A.V., Zitserman V.Yu., Kobzev G.A. The Systematization of Sources and Data on Nanotechnologies // *Scientific and Technical Information Processing*. – 2013. – Vol. 40, № 4. – P. 212-223.
27. Uniform Description System for Materials On the Nanoscale. Prepared by the CODATA-VAMAS Working Group On the Description of Nanomaterials. Version 1.0, 1 February 2015. – URL: www.codata.org/nanomaterials
28. Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А., Трахтенгерц М.С. Универсальная система метаданных для характеристики наноматериалов // *Научно-техническая информация. Сер.1*. – 2015. – №10. – С. 8-20; Erkimbayev A.O., Zitserman V.Yu., Kobzev G.A., Trakhtenherz M.S. A Universal Metadata System for the Characterization of Nanomaterials // *Scientific and Technical Information Processing*. – 2015. – Vol. 42, № 4. – P. 211-222.
29. Bizer C. Interlinking scientific data on a global scale // *Data Science Journal*. – 2013. – Vol. 12. – P. GRDI6-GRDI12.
30. Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А., Серебряков В.А., Теймуразов К.Б. Технология научных публикаций в среде «открытых связанных данных» // *Научно-техническая информация. Сер. 1*. – 2013. – №12. – С. 1-11; Erkimbayev A.O., Zitserman V.Yu., Kobzev G.A., Serebrjakov V.A., Teymurazov K.V. Publishing Scientific Data as Linked Open Data // *Scientific and Technical Information Processing*. – 2015. – Vol. 42, № 4. – P. 253-263.
31. Degtyarenko K., de Matos P., Ennis M. et al. ChEBI: a database and ontology for chemical entities of biological interest // *Nucleic Acids Research*. – 2008. – Vol. 36. – Data base issue. - D344–D350.
32. Batchelor C. Ontologies for nanotechnology. #237 - Technical Sessions // ACS National Meeting March 22-26, 2009. Salt Lake City, UT. – URL: <http://acscinf.org/docs/meetings/237nm/presentations/237nm76.pdf>
33. Thomas D.G., Pappu R.V., Baker N.A. et al. NanoParticle Ontology for cancer nanotechnology researches // *Journal of Biomedical Informatics*. – 2011. – Vol. 44. – P. 59–74.

34. Muñoz-Mármol M., Crespo J., Fritts M.J., Maojo V. Towards the taxonomic categorization and recognition of nanoparticles shapes // Nanomedicine: Nanotechnology, Biology, and Medicine. – 2015. – Vol. 11. – P. 457-465.
35. Haiyan Sun, Zhen Xu, Chao Gao. Multifunctional, Ultra-Flyweight, Synergistically Assembled Carbon Aerogels // Adv. Mater. – 2013. – Vol. 25. – P. 2554–2560.
36. Hendren C.O., Powers C.M., Hoover M.D., Harper S.L. The Nanomaterial Data Curation Initiative: A collaborative approach to assessing, evaluating, and advancing the state of the field // Beilstein J. Nanotechnol. – 2015. – Vol. 6. – P. 1752–1762.
37. Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А., Косинов А.В. Связывание онтологий с базами данных по свойствам веществ и материалов // Научно-техническая информация. Сер. 2. – 2015. – №12. – С. 1-16.

Материал поступил в редакцию 15.06.16.

Сведения об авторах

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

ЗИЦЕРМАН Владимир Юрьевич – кандидат физико-математических наук, ведущий научный сотрудник лаборатории баз данных ОИВТ РАН, Москва
e-mail: vz1941@mail.ru

КОБЗЕВ Георгий Анатольевич – доктор физико-математических наук, профессор, главный научный сотрудник, советник НИЦ электрофизики и тепловых процессов ОИВТ РАН, Москва
e-mail: gkbz@mail.ru

ТРАХТЕНГЕРЦ Михаил Самойлович – кандидат технических наук, ведущий научный сотрудник лаборатории баз данных ОИВТ РАН, Москва
e-mail: trachtengerts@yahoo.com

Информатика, её состав, содержание. Ядро и «паутина» отраслевых и прикладных информатик

Развитие ранее опубликованных положений о составе, структуре и содержании Информатики как науки. Рассмотрены вопросы об отраслевых и прикладных информатиках, о соотношении информатики и комплексов компьютерных, информационных, вычислительных наук, о связи и границах размежевания информатики и кибернетики.

Ключевые слова: информатика, научно-информационная деятельность, компьютерные науки, информационные науки, вычислительные науки, отраслевая информатика, прикладная информатика, защита информации, кибернетика, общая информатика, введение в информатику

После публикации нашей статьи с предложением договориться* о едином понимании науки и термина информатика* возникла необходимость уточнить ряд положений, в частности соотношений между информатикой и компьютерными, информационными, вычислительными науками, между общей информатикой и отраслевыми и прикладными информатиками. Напомню предложенное в указанной статье коллективно подготовленное определение информатики как науки: *Информатика (informatics) – формирующаяся в XX и XXI вв. междисциплинарная комплексная фундаментальная наука с общим объектом исследования (информационные явления, системы и процессы) и различными предметами исследования, отражающими задачи как отдельных комплексов теоретических (гуманитарных, естественных и технических) наук, так и прикладных информационных технологий.*

Необходимо разобраться в том, каковы контуры «большой системы» информатики – сложнейшей структуры научного знания, формирующейся в наше время. Проведённый нами анализ показал: наук и научных направлений, в названиях которых встречается информатика, очень много. Только в одной сфере социальной практики обнаружили экономическая информатика (в дальнейшем – и.), финансовая и., бюджетная и., валютная и., биржевая и., и. денежных потоков... Список не закрыт: желающих «открыть» собственную информатику много.

Ядро информатики, её базовое, основное место, располагается в разделе общенаучного знания, рядом с теорией систем, кибернетикой и теорией управления, семиотикой, экологией. Раздел общенаучного знания предшествует отраслевым комплексам социальных, естественных, технических наук. Структура

сложная: внутри основной части информатики, её ядра, находятся некоторые неотделимые составные части с «двойным подчинением». Так, *научно-информационная деятельность*, практически реализующая функции информатики в обществе, одновременно входит и в комплекс социальных наук, где находится рядом с родственными – библиографией и библиотечным делом, документоведением и архивоведением.

Сделаем попытку отграничить понятия *отраслевая* и *прикладная информатика*. Если речь идёт об информационных явлениях и процессах, то, как правило, мы имеем дело с отраслевой информатикой. Если речь идёт о системах (уточним: устройствах, приборах, электронном оборудовании и пр.), то о прикладной информатике. В теоретических (фундаментальных) науках сформировались (или пока ещё формируются) отраслевые информатики. В прикладных науках, как правило, мы встречаемся с соответствующими прикладными информатиками. Иногда в прикладных дисциплинах всё начинается с решения задач отраслевой информатики. Когда в теоретическом плане наступает ясность, приглашаются математики и программисты, конструкторы необходимых компьютерных устройств. Так рождается прикладная информатика.

Есть информатика историческая, археологическая и другие. Большая группа отраслевых информатик сформировалось в естественных науках (физике, химии, астрономии и др.).

В сельском хозяйстве, медицине, отраслях технических наук, производств, технологий уже появились соответствующие прикладные информатики. Однако в границах фундаментальной медицинской науки, которая изучает явления и процессы, характерные, например, для эмбриологии, анатомии и физиологии, формируется соответствующая отраслевая медицинская информатика. Когда же речь идёт о клинических медицинских отраслях, где применяются информационные приборы, устройства, системы,

* Сукиасян Э. Р. Информатика. Надо договориться о едином понимании науки и термина // Научно-техническая информация. Сер. 1. – 2016. – № 4. – С. 7–9

современная медицинская техника, мы имеем дело с прикладной медицинской информатикой.

То же самое характерно для биологической науки, в рамках которой развивается отраслевая биоинформатика. Идёт активное развитие биотехнологии. В рамках этой отрасли информатика будет прикладной. Здесь хорошо видна динамика развития. До тех пор, пока речь будет идти о теоретической науке, примерно говорить об отраслевой информатике. Но с развитием геномной инженерии мы получим раздел прикладной информатики. Чем дальше, полнее мы будем осваивать глубины науки, тем чаще будем привлекать для решения возникающих задач компьютерную технику. Конечно, без математиков и программистов мы не обойдёмся.

В такой интересной области научного знания, как метеорология, веками накапливались материалы наблюдений, регистрировались те или иные явления, анализ носил «кабинетный» характер. А носителем информации были регистрационные журналы «в бумаге». Но с развитием техники – сначала радиосвязи, затем – вычислительной, компьютерной техники, в метеорологии стала бурно развиваться прикладная информатика. Сегодня весь мир густо покрыт сетью станций метеорологического наблюдения. Обработкой информации в рамках системы занимаются специалисты по прикладной информатике. Конечно, к прикладной информатике относятся и геоинформационные системы (ГИС), обеспечивающие сбор, хранение, обработку, доступ, анализ, интерпретацию и графическую визуализацию пространственных данных.

Информатика стала «родной» и в целом ряде наук гуманитарных. В лингвистике благодаря информатике проявилось новое направление, занимающееся проблемами машинного перевода. Хочется написать иначе: автоматического, что означает: выполняемого компьютером целиком и полностью с помощью соответствующих программ. Но пока у нас машинный перевод автоматическим не получается. Тот, кто им пользуется, знает: обязательно требуется редактирование, выполнить которое может только человек. Поэтому мы имеем сегодня, по сути, автоматизированный перевод. Тем не менее, нам хочется думать, что проблема автоматического перевода разрешима. А вот в другой области лингвистики – лексикографии – компьютеры действительно произвели революцию. Традиционные для лингвистов методы работы на карточках (можно легко представить себе трудоёмкость организации и поиска материала) удалось заменить машинными. Кроме быстрого действия (что само по себе тоже очень важно), машина позволила приложить программные средства. В результате мы получили словари новых типов (например, «Обратный словарь русского языка»), о которых раньше и мечтать не могли. Специалисты знают: при разработке «большого» тезауруса без машины не обойтись. Есть и другие направления лингвистики, о которых все мы не подозреваем. Хорошо, что деньги на исследования отпускаются, учёные ведут поиск.

Порой трудно выявить предмет и объект предлагаемой авторами «информатики», так как не даётся никаких определений. Считается, что из самого названия всё настолько очевидно, что дефиниция не нужна. Между тем, понимание самого термина информатика в общест-

ве крайне неоднозначно. Авторы многих публикаций являются узкими специалистами, работают и общаются на языке, который им (и их собеседникам, окружающим) понятен. Можно порекомендовать внимательнее и требовательнее относиться к терминологии. В науке есть традиция давать чёткие определения, прилагать словарь применяемых терминов и понятий.

В классификационной структуре все отраслевые и прикладные информатики находят себе место в соответствующем иерархическом подразделении. Историческая информатика – в разделе истории, археологическая – в археологии, геологическая – в геологии, физическая – в физике и т.д. Предлагаемые названия должны быть терминологически грамотными, адекватными содержанию.

Очень важно, чтобы отраслевые и прикладные информатики нашли признание в своей науке. Но для этого надо публиковаться в отраслевой научной периодике. Сейчас значительная часть статей об отраслевых и прикладных вопросах информатики попадает на страницы журналов, так или иначе связанных с информацией и компьютерной техникой.

Рассмотрим соотношение понятий *информатика – компьютерные науки*.

Теория и проектирование (включая вопросы конструирования и расчёта) всего комплекса аппаратных средств вычислительной техники, её отдельных деталей, узлов, систем, соответствующие технологии производства, эксплуатация, ремонт, – всё это входит в компьютерные науки и технологии, но к информатике отношения не имеет. В одном из учебников по информатике для вузов встретилось, как мне кажется, правильное утверждение: «Информатика – пользовательская наука». Вузов, готовящих кадры конструкторов компьютерного оборудования, очень мало. Они занимаются не информатикой, а электронной техникой, электронным приборостроением.

Для информатики, основы которой изучаются сегодня практически везде, как в средней, так и в высшей школе, достаточен лишь самый элементарный объём знаний в области вычислительной техники. Пособия по информатике не объясняют нам, как сконструирован компьютер (сканер, принтер). Мы знакомимся лишь с функциями (назначением) отдельных узлов, устройств, оборудования в целом. Примерно на том же уровне, на котором управляем, например, телевизором, телефоном.

То же можно сказать о *вычислительной технике и программировании*. Мы используем в информатике огромное количество специально разработанных прикладных программ. Но специалисты по информатике программированием (разработкой программ) не занимаются, они лишь их используют, выбирая рекомендованные из имеющихся. Конечно, в случае необходимости, простейшую программу (например, по статистической обработке данных) можно составить и «на месте». Этому учат школьников 10-11 классов.

Сложнее обстоит дело с *информационными науками*. Именно потому, что на протяжении многих лет считалось нормальным ставить знак равенства между информатикой и информационной наукой (особенно при переводе на английский язык).

Между тем, информационные науки – очень широкий класс гуманитарного знания. Многие области прак-

тической деятельности, которые мы сегодня относим к информационным наукам, возникли в глубокой древности (библиотечное, архивное, музейное дело). Развитие теоретической составляющей каждой из указанных областей привело к рождению наук (библиотечковедения, архивоведения, музееведения). Перечень далеко не полный: легче назвать то, что не входит в информационные науки, нежели перечислить их состав.

Место комплекса информационных наук в общей (философской) классификации наук не определено. В структуре классификационных систем такого самостоятельного класса нет. Поэтому непонятно, где собирать литературу об информационных науках в целом. Философы считают, что информационные науки – часть комплекса гуманитарных наук. Однако в отечественной и зарубежной практике сложилось различное понимание самого понятия *гуманитарные науки*.

Рождение информатики и возможности компьютеризации исследований привели к ускоренному развитию ряда направлений научного поиска, таких, например, как искусственный интеллект, машинный перевод, исследования в области концентрации и рассеяния публикаций, библиометрии информационного потока, анализа цитирования, старения информации и многих других. На первых порах всеми указанными направлениями занималась теория информатики. Но со временем многими из указанных направлений стали заниматься учреждения и специалисты соответствующих областей науки. Мы привели пример с машинным переводом – это сфера лингвистики.

С развитием информатики стали больше говорить об *информационной безопасности и защите информации*. В учебных планах вузов появились соответствующие дисциплины (а в ряде вузов – и факультеты). Надо понимать: с информатикой связана литература о программно-технических средствах защиты информации. Проблемами правового и законодательного характера занимаются другие науки. Элементарными средствами и технологиями защиты от несанкционированного доступа при работе с компьютером (идентификация пользователя, логин, пароль) занимаются системные администраторы, а не специалисты по защите информации. Как и внедрением антивирусных программ, борьбой со Спамом. На «пользовательском уровне» вопросы информационной безопасности мы относим к информатике. Специалисты по защите информации готовятся отдельно, по специальным программам и учебным пособиям. Поскольку в последние годы издаётся много литературы, в ББК мы примем условное решение и дадим подразделение «Защита информации» в разделе информатики.

Одним из интересных является вопрос о связях и отношениях между информатикой и кибернетикой. Многие считают, что информатика выросла из кибернетики. Как показывает изучение литературы, в этом убеждении есть рациональное зерно. Кибернетика пришла в мир сразу со своим именем, в то время как информатика начала формироваться на протяжении нескольких десятилетий. И этот процесс не завершён и сегодня. Внимательно перечитывая опубликованные в печати в 50-70-х гг. XX в. работы о кибернетике и её проблемах, мы встречаемся с потоком материала о тех или иных закономерностях ин-

форматики, приписываемых кибернетике. В «Энциклопедии кибернетики» в 2-х т. (Киев, 1974) собран фактический материал не только и не столько о кибернетике, сколько об информатике. Таких примеров немало.

Между тем, в самом определении кибернетики как науки было сказано всё необходимое для того, чтобы достаточно чётко отграничить её объект и предмет: Кибернэтика (от др.-греч. κυβερνητική – «искусство управления») – наука об общих закономерностях получения, хранения, передачи и преобразования информации в сложных управляющих системах, будь то машины, живые организмы или общество. Теория управления, теория информации и теория систем – три области научного знания, каждая из которых имеет свой объект и предмет. Несмотря на тесную взаимосвязь, они являются самостоятельными. Каждая занимает собственное место в отделе ББК «Общественное и междисциплинарное знание».

Осталось рассказать о том, что входит в состав ядра информатики? Ведь учебная дисциплина под названием «Информатика» (иногда «Введение в информатику», «Информатика. Общий курс») преподаётся сегодня почти во всех вузах. В учебных заведениях общего среднего образования (школах, гимназиях, лицеях) аналогичные задачи решает курс «Информатика и информационно-коммуникативные технологии».

Мы провели анализ 27 учебников и учебных пособий, тех, которые удалось разыскать в «бумажном» виде или в Интернете. Взяв на вооружение философские категории общего, особенного и единичного, сделали попытку отделить общее (оно не должно быть связано, например, с профилем учебного заведения), особенное (по рабочей гипотезе оно должно отличаться, например, у будущих гуманитариев, математиков и инженеров), и единичное. В рамках одного учебника единичного может и не быть: такие, вероятно, весьма специальные задачи обучения в каждом вузе, будут решаться в рамках других курсов, отражающих проблематику отраслевой или прикладной информатики.

Результат позволил сделать некоторые выводы. Структура «общего» совпадает на 70-80%. Это, как правило: (1) теоретические проблемы информатики, (2) информационно-коммуникационные технологии, (3) компьютерная обработка данных, (4) технические и программные средства информатики, (5) информационные системы, (6) сетевые информационные технологии. Интернет, (7) информационная безопасность, (8) системы искусственного интеллекта.

Наша работа продолжается, уже есть проект раздела 16 *Информатика* Средних таблиц ББК. Приглашаем принять участие в обсуждении этой проблемы.

Материал поступил в редакцию 20.06.16.

Сведения об авторе

СУКИАСЯН Эдуард Рубенович – кандидат педагогических наук, доцент, зав. сектором Российской государственной библиотеки, главный редактор Библиотечно-библиографической классификации, Москва e-mail: sukiasyaner@rsl.ru

Литературоведение: потребители информации, базы данных, принципы содержательной обработки и стратегии поиска

Описывается проблемно-тематическая структура базы данных по литературоведению, формируемой в Институте научной информации по общественным наукам РАН (ИНИОН). Раскрываются принципы и особенности аналитико-синтетической переработки документов и запросов на основе комплекса лингвистических средств системы. Рассматриваются стратегии поиска информации по запросам различной сложности. Отмечается необходимость повышения качества доступа к электронным информационным ресурсам по литературоведению.

Ключевые слова: литературоведение, фольклористика, базы данных, потребители информации, лингвистические средства, стратегии поиска, электронные информационные ресурсы

ВВЕДЕНИЕ

Современное литературоведение представляет собой очень сложную и подвижную систему дисциплин. Термин «литературоведение» сравнительно недавнего происхождения; раньше более широко использовалось понятие «история литературы». Основы целостного изучения литературы начали складываться в 30-е гг. XX в. и, прежде всего, в трудах известного русского ученого М.М.Бахтина [1]. Современное литературоведение представляет собой целый комплекс дисциплин, у каждой из которых существуют свои особые внутренние границы в пределах того общего целого, которое они образуют. Выделяются три главные отрасли литературоведения: 1) теория литературы, которая исследует общие законы структуры и развития литературы; 2) история литературы, предметом которой является прошлое литературы как процесс или как один из моментов этого процесса; 3) литературная критика, которую интересует последнее, «сегодняшнее» состояние литературы; однако для неё характерна также интерпретация литературы прошлого с точки зрения современных общественных и художественных задач. В работе Д.М. Сегала [2], одного из активных участников структурно-семиотического движения в литературоведении, раскрываются основные направления в истории литературной критики и науки о литературе в России конца XIX – XX вв., большое внимание уделяется становлению и развитию русской формальной шко-

лы в литературоведении (Б. Эйхенбаум, В. Шкловский, Ю. Тынянов), сравнению идей М.М. Бахтина и «младоформалистов», русской школе фольклористики (В.Я. Пропп и О.М. Фрейденберг), а также работам основателей структурно-семиотического направления в литературоведении (Вяч. Вс. Иванова и В.Н. Топорова).

Углубление задач, стоящих перед исследователями художественной литературы, повлекло за собой усиленную дифференциацию внутри этой дисциплины. Для современного литературоведения как системы научных дисциплин важна не только тесная взаимозависимость всех её отраслей, но и возникновение дисциплин второго ряда (таких, например, как теория критики, история критики), а также различных вспомогательных дисциплин, включая литературоведческое архивоведение, текстологию, эвристику, библиографию художественной и литературоведческой литературы и пр.

В настоящее время взаимодействие литератур становится катализатором процесса рождения общечеловеческой литературы, и глобализация играет роль проявления общечеловеческой всеобщности, единения духовных усилий разных народов, обогащения мировой культуры богатством национальных культур, вбирающих в себя достижения других народов, но сохраняющих при этом свою самобытность и неповторимость [3]. В связи с этим необходимо отметить важную роль сравнительного литературове-

дения как раздела истории литературы, которое изучает международные литературные связи и отношения, сходство и различия между литературно-художественными явлениями в разных странах.

Сегодня Россия ищет объединяющее начало в стране. А ведь это русский язык и литература. Доступ к новым знаниям, идеям, духовным и моральным ценностям каждый человек получает чаще всего благодаря литературе различных стран, которая позволяет знакомиться с культурой, обычаями и традициями разных народов. В связи с этим в рамках Генеральной конференции ЮНЕСКО, проходившей в Париже в 1995 г., был учрежден Всемирный день книги и авторского права (World Book and Copyright Day), который с 1996 г. отмечается 23 апреля. Этот день является символической датой для мировой литературы, поскольку именно в этот день, например, в 1616 г. ушли из жизни Уильям Шекспир, Мигель де Сервантес и другие известные писатели и поэты. Важнейшей целью науки о литературе является воспитание квалифицированного читателя, способного адекватно интерпретировать художественный текст. В настоящее время эта задача становится особенно важной, поскольку книги не застрахованы от пожаров и происходящих в мире изменений, связанных с появлением цифровых форматов и переходом на принципы открытого лицензирования в интересах содействия обмену знаниями. Технологические изменения ставят непростые вопросы в отношении определения того, что такое книга и что есть авторство в цифровую эру (<http://www.unesco.org/new/ru/unesco/events/prizes-and-celebrations/celebrations/international-days/world-book-and-copyright-day-2015/>).

Необходимо также отметить, что 2015 г. был официально объявлен годом литературы в России. В разных регионах страны было проведено свыше ста масштабных мероприятий, среди которых Международный писательский форум "Литературная Евразия", проект "Литературная карта России", публичные чтения произведений мировых классиков литературы, творческие встречи писателей с читателями в библиотеках и книжных магазинах, проекты «Живая классика» и «ReadRussia/ Читай Россию», а также организованы литературные фестивали, конкурсы, ярмарки и выставки, позволяющие лучше узнать творчество современных писателей и поэтов. Кроме того, при финансовой поддержке Федерального агентства по печати и массовым коммуникациям был создан информационный портал ГодЛитературы.РФ, основной задачей которого стало не только освещение всех запланированных мероприятий, но и создание оригинального контента для привлечения читателей к активному участию в проводимых акциях, фестивалях, встречах с писателями и пр. (<https://godliteratury.ru>).

ПОТРЕБИТЕЛИ ИНФОРМАЦИИ ПО ЛИТЕРАТУРОВЕДЕНИЮ

Информационные потребности в области литературоведения – научных работников, исследователей, преподавателей высшей школы, аспирантов и студентов старших курсов, исследователей литературы, журналистов, писателей, поэтов, сценаристов, кинематографистов, библиотечных сотрудников, работ-

ников СМИ и Интернета, а также обычных граждан – формируются под влиянием огромного потока информации об отечественной и зарубежной литературе и, прежде всего, по теории и истории литературоведения, теории и истории литературы всех стран и народов мира – и всех исторических периодов, по теории и истории фольклора народов нашей страны и зарубежья, поступающей в Фундаментальную библиотеку по общественным наукам ИНИОН РАН.

В структуре информационных запросов по проблемам литературоведения выделяются следующие наиболее важные составляющие: проблемно-тематические, страноведческие, региональные, языковые, хронологические, а также персоналии (писатели, поэты, драматурги, прозаики, публицисты, литературные критики, историки литературы, литературоведы, филологи, исследователи литературы, переводчики и др.), школы и направления, отдельные произведения и литературные герои (персонажи).

Общее представление о современных литературоведениях дает Биобиблиографический словарь-справочник [4], который является современным аналогом изданного ранее словаря-справочника «Кто есть кто в российском литературоведении» [5]. Анализ сведений, представленных в словаре-справочнике (библиографические сведения и библиография произведений литературоведов из 80 городов страны), показывает, что за последние два десятилетия российское литературоведение значительно пополнилось работами прежних и новых исследователей.

Среди коллективных потребителей литературоведческой информации следует назвать, прежде всего, Институт мировой литературы имени А.М. Горького РАН (ИМЛИ) (<http://www.imli.ru>), который, наряду с Институтом русской литературы (Пушкинский дом), осуществляет координацию исследований в области теории и истории литературы в стране. Успех работы научных сотрудников над «Историей всемирной литературы» – главным трудом ИМЛИ во многом зависит от качества и эффективности информационного обеспечения научных исследований.

Наиболее важные информационные потребности научных сотрудников ИМЛИ связаны с исследованиями по русской литературе как органической части литературы мировой, с изучением жизни и творчества русских писателей, русской советской литературы, литератур народов СССР, а также с созданием их истории. Информационная поддержка необходима и для изучения зарубежной прозы, поэзии и драматургии, анализа творчества крупнейших мастеров мировой литературы. Важно отметить, что на сайте Института имеется раздел «Электронная библиотека ИМЛИ РАН», в которой размещено более 360 сканов научных изданий 1848–2014 гг. по следующим направлениям: «Теория литературы», «Русская литература», «Литература народов России и стран СНГ», «Зарубежная литература», «Фольклористика», «Периодика ИМЛИ», «Указатели и справочники». Издания доступны для просмотра и скачивания любому желающему без регистрации (<http://philologist.livejournal.com/8340946.html>).

Другой коллективный потребитель информации по литературоведению – Институт русской литературы РАН (ИРЛИ), который является одним из ведущих на-

учных центров, занимающихся исследованиями русской литературы (<http://www.pushkinskiydom.ru>).

Информационные потребности научных сотрудников ИРЛИ связаны с подготовкой академического Полного собрания сочинений Пушкина и Пушкинской энциклопедии, с активными исследованиями по истории русской литературы пушкинской эпохи, с изучением памятников истории русской литературы XI-XVII вв., археографии, текстологии, а также поэтики древнерусской литературы, литературных взаимосвязей, связей древнерусской книжной и народной культуры. Так, основное направление деятельности Отдела русской литературы XVIII в. – изучение русской словесной культуры XVIII столетия в ее взаимосвязях с историей, литературой и культурой других стран и эпох. Актуальным направлением научных исследований ИРЛИ является также комплексное изучение истории русской литературы XX в. и современного литературного процесса. Информационная поддержка необходима и научным сотрудникам библиотеки ИРЛИ, во-первых, для подготовки библиографических пособий литературоведческой тематики (например, о жизни и творчестве А.С. Пушкина, М.Ю. Лермонтова и М. Горького), во-вторых, для описания личных библиотек писателей.

Институт филологии Сибирского отделения РАН – это единственный академический институт за Уралом, который проводит комплексные фундаментальные исследования в области языкознания, фольклористики и литературоведения. Он осуществляет подготовку к изданию 60-томной академической серии «Памятники фольклора народов Сибири и Дальнего Востока», которая представляет собой новаторскую, единственную по своему охвату документальную публикацию произведений фольклора более чем на 30 языках, и в этом отношении она не имеет аналогов в международной фольклористической практике. В настоящее время из печати вышло 27 томов этой уникальной серии. Продолжается работа над созданием первого в отечественном литературоведении «Словаря сюжетов и мотивов русской литературы», охватывающего литературные памятники XI–XX вв. (<http://hist-phil.ru>).

Кафедры по литературоведению имеются во многих вузах страны, профессорско-преподавательский состав которых заинтересован в получении отечественной и зарубежной информации для составления учебных программ и проведения научных исследований. Важно выделить, прежде всего, такие вузы, как: Московский государственный университет им. М.В. Ломоносова, Филологический факультет, кафедра теории литературы; Московский государственный университет дизайна и технологий, кафедра языкознания и литературоведения (Институт славянской культуры); Московский государственный лингвистический университет, кафедра литературы; Российский университет дружбы народов, кафедра русской и зарубежной литературы; Санкт-Петербургская школа социальных и гуманитарных наук, кафедра сравнительного литературоведения и лингвистики.

В Литературном институте имени А.М. Горького (<http://litinstitut.ru>) подготовка студентов ведется параллельно по двум направлениям: общегуманитар-

ному (с уклоном на изучение русского языка и литературоведения) и творческому. Преподавателям и студентам этого института требуется информация по литературоведению для научных исследований и подготовки учебных курсов по русской классической литературе и славистике, по теории литературы и литературной критике, по новейшей русской литературе, по зарубежной литературе и др.

Подготовка специалистов в области литературоведения осуществляется также на кафедрах университетов в различных регионах страны.

В настоящее время важнейшие цели интеграции образовательной и научно-исследовательской деятельности в области литературоведения – это научно-методическое и информационное обеспечение научных исследований, а также инновационных образовательных программ для повышения качества подготовки студентов и аспирантов, привлечения их к проведению научных исследований на основе оперативного доступа к информационным ресурсам для поиска новых знаний, идей и концепций, необходимых для понимания современных проблем и тенденций. Так, в области литературоведения Новосибирский государственный университет совместно с Институтом истории и археологии Уральского отделения РАН работает над проектом «Литература и история: сферы взаимодействия и типы повествования», целью которого является анализ форм взаимодействия истории и литературы как различных способов описания реальности (http://www.nsu.ru/science_divisions_humanities_literary_studies_ru).

В образовательных проектах XXI в. достойное место занимает Государственная академия славянской культуры – ГАСК (<http://gask.ru>), обучение в которой строится на тесном контакте с ведущими научными и образовательными учреждениями Российской Федерации (ИМЛИ РАН, ИРЯ РАН, ИСЛ РАН, МГУ им. М.В. Ломоносова, РГГУ, МПГУ и др.) и зарубежными университетами, среди которых Софийский университет им. Климента Охридского, Велико-тырновский университет им. св. Кирилла и Мефодия, филиал Пловдивского университета, Юго-Западный университет «Неофит Рильски», Шуменский университет им. еп. Константина Преславского, Белградский университет, Остравский университет, Прешовский университет, Ягеллонский университет, Гродненский государственный университет им. Янки Купалы и др. В рамках ГАСК на базе филологического факультета функционирует научный центр «Славянские Культуры: Корни и Крона» (ЭСТРИКА), созданный с целью активизации в вузе научных исследований в области славяноведения, проведения регулярных международных и всероссийских конференций, укрепления международных научных связей с учеными из зарубежных славянских стран.

Важно отметить также уникальность магистерской программы «Теоретико-прикладное литературоведение и критика», реализованной на кафедре теории литературы филологического факультета МГУ им. М.В. Ломоносова, которая заключается в том, что студенты получают полную информацию о мировом литературном процессе, могут составить представление о литературе как явлении культуры и

виде искусства, которое складывалось разными путями в разных национальных традициях. Эта программа ориентирована на фундаментально-прикладную подготовку специалистов-литературоведов, способных максимально полно реализовать себя в образовательных, медийно-информационных, экспертно-правовых, культурно-воспитательных сферах.

ХАРАКТЕРИСТИКА ИНФОРМАЦИОННОГО ПОТОКА ПО ЛИТЕРАТУРОВЕДЕНИЮ

Информационную поддержку научных исследований и образования в области литературоведения в течение нескольких десятилетий оказывает Институт научной информации по общественным наукам РАН (ИНИОН), который с 1986 г. формирует библиографическую базу данных (БД) «Литературоведение». В настоящее время ее объем составляет почти 400 тыс. библиографических записей документов, поступающих в фонды Фундаментальной библиотеки по общественным наукам ИНИОН и ее отделений библиотечно-библиографического обслуживания (ОББО) при академических институтах РАН. Доступ для потребителей литературоведческой информации возможен в интерактивном режиме на сайте ИНИОН в Интернете (www.inion.ru) и на оптических дисках с использованием информационно-поисковой системы WebIRBIS™, особенностью которой является ее многоцелевой характер: она используется в процессе комплексной автоматизированной обработки литературы по социальным и гуманитарным наукам, в том числе при подготовке и выпуске библиографических указателей (текущих и ретроспективных).

Традиционные библиографические издания, которые выпускаются в автоматизированном режиме, применяются для поиска информации по литературоведению во многих российских библиотеках. Следует отметить, что с 1993 г. ИНИОН РАН ежемесячно издает аннотированный библиографический указатель «Новая литература по социальным и гуманитарным наукам. Литературоведение», который является продолжением библиографических указателей «Новая советская литература по литературоведению» и «Новая иностранная литература по литературоведению», выпуск которых был начат соответственно в 1953 и 1954 гг. Указатель, предназначенный для научно-исследовательской, научно-образовательной, учебной, библиографической и справочной деятельности, содержит информацию об отечественной и зарубежной литературе, по теории и истории литературоведения, о литературе всех стран и народов, всех исторических периодов, по теории и истории фольклора разных народов нашей страны и зарубежья. Он снабжен вспомогательными указателями – авторским, предметным, использованных источников. Значительный вклад в подготовку указателя, ориентированного на ученых, литературных критиков, преподавателей высшей школы, аспирантов и студентов-литературоведов, филологов широкого профиля, а также библиотечных и информационных специалистов, вносят такие специалисты-библиографы, как Е.А. Дмитренко, М.Я. Курашова, Н.В. Логутова, Е.А. Малютина, О.В. Этова.

В основу формирования информационного массива по литературоведению, поступающего в фонды Фундаментальной библиотеки по общественным наукам ИНИОН, положены такие принципы, как экспертный анализ и отбор научно значимых статей из журналов и сборников, полнота представления монографических изданий, многоаспектность содержательной переработки документов, отобранных в результате информационной фильтрации потока научной литературы, а также ее видовое и языковое разнообразие.

В информационном потоке по литературоведению представлены документы различных типов и видов: монографии, сборники статей, отдельные статьи и рецензии из сборников, альманахов, журналов и прочих периодических изданий, еженедельников газетного типа, многотомные издания, материалы научных конференций, конгрессов, симпозиумов и отчеты о них, публикации критических и художественных текстов прошлого, сопровождающиеся литературоведческим справочно-научным аппаратом, учебники и учебные пособия для высшей школы, библиографические и справочные издания, энциклопедии и словари, авторефераты диссертаций, предисловия и послесловия к художественным произведениям.

В соответствии с общесистемной технологией после завершения подготовки указателя к изданию все библиографические записи документов включаются в поисковые базы данных (сводные, отраслевые, проблемно-ориентированные). Структура и форма представления библиографической записи в них соответствуют стандартам «Системы стандартов по информации, библиотечному и издательскому делу» (СИБИД), согласованным со стандартами ИСО. Формат библиографической записи, используемой при обработке документов по социальным и гуманитарным наукам, относится к полным универсальным форматам семейства MARC и обеспечивает многофункциональное формирование различных входных и выходных форм библиографической записи, в том числе транслитерированных.

Результаты аналитико-синтетической переработки информационного потока документов по литературоведению и фольклористике вводятся в отраслевую БД «Литературоведение». Информация о рукописях научных работ по литературоведению, депонированных в ИНИОН РАН, отражается в библиографическом указателе «Депонированные научные работы» и поступает в соответствующую базу данных.

Статистический анализ показывает, что более двух третей информационного потока по литературоведению составляют статьи из научных журналов и сборников. Общее количество монографий и многотомных изданий превышает 12%. Одной из особенностей БД по литературоведению по сравнению с другими БД ИНИОН РАН является наличие в ней большего количества рецензий. Важно отметить, что языковой состав документального массива по литературоведению очень широк: в отраслевой БД литературоведческая информация представлена почти на 100 языках, в том числе и восточных (китайском, корейском, японском). Документы на русском языке составляют чуть более половины информационного массива. Среди других языков, на которых в отрасле-

вую БД поступает литературоведческая информация, следует выделить английский, болгарский, испанский, итальянский, немецкий, польский, сербохорватский, украинский, французский. Некоторые документы представлены на двух, трех и более языках.

Содержательная переработка документов по литературоведению осуществляется в два этапа. Сначала проводится многоаспектный анализ документа, при котором особое внимание уделяется заглавию, оглавлению, прикнижной аннотации, введению, заключению, резюме, предметному указателю, с целью определения его основного информационного содержания; затем основное содержание документа представляется с помощью комплекса лингвистических средств по литературоведению и формируется структура поискового образа документа. Обязательный элемент описания содержания документов на иностранных языках – это аннотация. Для документов на русском языке аннотация является факультативной и приводится в случае, когда заглавие документа не достаточно информативно.

В процессе аналитико-синтетической переработки документов по литературоведению и создания отраслевой БД используется комплекс лингвистических средств по социальным и гуманитарным наукам, обеспечивающий формальную и содержательную переработку многоязычного потока документов с учетом специфики отдельных наук и научных дисциплин и, в первую очередь, особенностей понятийных и терминологических систем в различных языках, наличия различных научных школ, направлений, течений. Для проведения быстрого поиска элементы основных полей объединены в так называемые инверсные файлы: словари авторов, ключевых слов (дескрипторов), общий словарь, на основе которого поиск ведется по лексике всего документа, включая заглавие, аннотацию, отдельные слова и словосочетания, а также элементы рубрик отраслевого рубрикатора.

ЛИНГВИСТИЧЕСКИЕ СРЕДСТВА ПО ЛИТЕРАТУРОВЕДЕНИЮ

В настоящее время комплекс лингвистических средств, который используется в ИНИОН для аналитико-синтетической переработки документов и запросов по литературоведению, включает: 1) информационно-поисковый тезаурус (ИПТ) по литературоведению и фольклористике [6], изданный в 2010 г. и пополненный актуальной терминологией в 2014 г.; 2) отраслевой рубрикатор по литературоведению, который является составной частью Рубрикатора Автоматизированной информационной системы по общественным наукам (АИСОН) [7]; 3) списки рубрик Предметного каталога ИНИОН.

ИПТ по литературоведению и фольклористике содержит около 2000 терминов (дескрипторов, простых или сложных условных синонимов). Выбор тезауруса как основного средства представления терминологии по литературоведению обусловлен, прежде всего, актуальностью использования словарей тезаурусного типа в практике реализации информационных процессов, связанных с анализом основного содержания

документов и запросов в области социальных и гуманитарных наук, а также необходимостью фиксации понятийных связей внутри документального потока, отражаемого в отраслевой БД, для достижения требуемой полноты и точности поиска.

В настоящее время ИПТ по литературоведению выполняет функцию терминологического контроля при вводе документов в Автоматизированную информационную систему по общественным наукам (АИСОН) и обеспечивает единообразное представление в системе близких по содержанию документов. Он используется в традиционном виде (как печатное издание), или как лингвистическая база данных, а также входит как составная часть в интегрированный электронный Большой информационный словарь по общественным наукам (БИСОН) [8]. Понятия, необходимые для описания основного содержания документов по литературоведению, представлены в тезаурусе дескрипторами (более 1100 лексических единиц), которые нормализованы по принятым в АИСОН правилам.

Словарная статья каждого дескриптора представляет собой перечень лексических единиц ИПТ, семантически связанных с ним определёнными типами отношений. Для целей содержательной переработки документов и запросов в тезаурусах по социальным и гуманитарным наукам чаще всего фиксируется лишь наличие иерархических (обозначение ВТ – вышестоящих или НТ – нижестоящих дескрипторов) и ассоциативных (РТ) отношений без уточнения их особенностей. В качестве ассоциативных отношений между дескрипторами фиксируются, прежде всего, те отношения, которые полезны как для раскрытия содержания понятия или определения его места в понятийной системе науки в целом, так и для возможного расширения (или уточнения) границ информационного поиска или модификации поисковой стратегии.

Некоторые термины включены в отраслевой тезаурус как условные синонимы (простые или сложные) с отсылкой к дескрипторам, разрешенным для использования. При индексировании документов синонимы заменяются на соответствующие дескрипторы.

Например:

<i>контраст</i>	– простой синоним
USE антитеза	– дескриптор
<i>историческая драма</i>	– сложный синоним
USE драма + историческая тема	– дескрипторы.

Поиск терминов по литературоведению для индексирования документов и запросов может осуществляться по различным указателям ИПТ, включая: 1) алфавитный лексико-семантический указатель, содержащий словарные статьи дескрипторов и синонимов (недескрипторов); 2) систематический указатель, в котором дескрипторы распределены по рубрикам отраслевого рубрикатора; 3) пермутационный указатель, включающий дескрипторы и синонимы без словарных статей. Указатель литератур отдельных стран и народов, а также указатель фольклора отдельных народов в состав ИПТ не включены. Они представлены в вышедшем в 2014 г. издании «Вспомогательные фасеты АИСОН ИНИОН РАН», являющемся дополнением к ИПТ ИНИОН.

история литературы – дескриптор

BT литературоведение – вышестоящий дескриптор

NT античная литература – нижестоящие дескрипторы

всемирная литература

древняя литература

литература возрождения

литература нового времени

литература просвещения

советская литература

современная литература

средневековая литература

RT историография – ассоциативные дескрипторы

литературные архивы

литературные направления

литературные связи

литературный процесс

РУБ А170107; А1709 – коды рубрик отраслевого рубрикатора.

Приведем описание структуры верхнего уровня рубрикатора по литературоведению с указанием количества библиографических записей документов в каждой рубрике в БД:

В отраслевом рубрикаторе каждая рубрика представлена как в словесной форме (на русском и английском языках), так и с помощью буквенно-цифровой нотации (кода), что отражает иерархическую структуру рубрикатора. Например: A170721 Литература и общество = Literature and society

ния структуры библиографического указателя по литературоведению и поиска информации по данной теме (проблеме). Проблемно-тематические рубрики содержат рубрики более низшего уровня (подрубрики). Например, рубрика А1701 «Литературоведение и литературная критика» включает подрубрики:

A170107 Теория и методология литературоведения и литературной критики
A170109 История и современное состояние литературоведения и литературной критики
A170117 Международное сотрудничество в области литературоведения
A170145 Преподавание литературоведения и литературы.

Что касается подрубрики А170109 «История и современное состояние литературоведения и литературной критики», то она в процессе аналитико-синтетической переработки документов для обозначения отдельных стран (групп стран, регионов) может снабжаться буквенно-цифровыми кодами рубрик, которые отбираются из Политико-географического фасета и начинаются с латинской буквы В (например, В5990 – Россия, В6750 – СССР, В7590 – Франция, В2460 – европейские страны).

Важнейший раздел литературоведения составляет теория литературы, без которой не могут обходиться и другие разделы – критика и история литературы. Обобщая опыт мирового художественного процесса, теория литературы, как тип познания человека, исследует природу духовно-эстетических ценностей (персонаж, характер, чувства), создает правила и законы о способах построения стихотворных, прозаических и драматургических произведений, призвана объяснять сущность общественного явления, осваиваемого и постигаемого литературоведением в целом.

Структура рубрики A1707 «Теория литературы» является более сложной и включает подрубрики:

A170701 Общие вопросы
A170709 История литературно-эстетической мысли
A170721 Литература и общество
A170725 Литература и другие виды искусств
A170729 Литературный процесс
A170731 Художественные методы, литературные
 стили и направления
A170741 Поэтика художественной литературы
A170751 Языки стиль художественных произведе-
 ний
A170761 Художественный перевод
A170765 Теория и психология литературного твор-
 чества.

Следует отметить, что большинство приведенных подрубрик далее не детализируется; исключения составляют рубрики A170721, A170729, A170741, которые также имеют нижестоящие подрубрики. Так, рубрика A170741 «Поэтика художественной литературы» включает следующие подрубрики:

A17074125 Способы художественного обобщения
A17074141 Эстетические категории в литературе
A17074151 Виды и жанры литературы.

Более сложной по структуре является рубрика A17074151 «Виды и жанры литературы», которая содержит 12 подразбук:

- A1707415121 Художественная проза
- A1707415125 Поэзия
- A1707415131 Драматургия
- A1707415141 Кинодраматургия
- A1707415143 Радио- и теледраматургия
- A1707415145 Сатирические жанры
- A1707415151 Научно-фантастическая литература
- A1707415155 Научно-художественная литература
- A1707415161 Приключенческая и детективная литература
- A1707415165 Мемуарная и биографическая литература
- A1707415171 Детская литература
- A1707415175 Другие виды жанры.

Рубрика A1709 «История литературы» включает подразбук, которые определяют исторические или хронологические границы соответствующего массива документов и содержат документы, характеризующие литературный процесс не по отдельным странам, а в целом:

- A170909 Всемирная литература
- A17090905 Литература Древнего Востока
- A17090915 Античная литература
- A17090925 Литература Средних веков
- A17090927 Литература эпохи Возрождения
- A17090931 Литература 17-18 вв.
- A17090941 Литература 19 в.
- A17090955 Литература 20 в.

Отдельным литературам посвящена рубрика A170991 «Литература отдельных стран и народов». Каждая литература имеет цифровой код C. Например: C1570 – Русская литература; C1870 – Японская литература и т.д. Отдельные литературы внутри рубрики A170991 располагаются по алфавиту. Что касается русской литературы, где объем документов достаточно велик, то для уточнения исторического периода рассмотрения исследуемого произведения или писателя, или литературной ситуации применяются буквенно-цифровые коды рубрик, которые начинаются с буквы латинского алфавита D (например, D0923 – древняя литература, D0925 – литература 18 века, D0931 – литература 19 века, D0955 – литература 20 века, D0959 – литература 21 века).

Для указания наличия в документе информации по тем или иным персоналиям в библиографическую запись документа вводится буквенно-цифровой код рубрики E01: Пушкин А.С., Есенин С.А. и т.д. Благодаря такой буквенно-цифровой нотации в библиографическом указателе появляется персональная подразбука отдельного писателя. Для поиска в базе данных она не используется, поскольку в поисковом образе документа обязательно наличие ключевого слова – имени писателя: фамилия и инициалы или псевдоним. Что касается иностранных писателей, то их имя пишется в общепринятой форме и на национальном языке.

Сложной по структуре является рубрика A1771 «Народное поэтическое творчество», которая, хотя и не явно, тоже состоит из отдельных частей. Первая

часть, от подразбуки A177101 до подразбуки A177107, посвящена фольклористике как науке, вторая – подразбука A177191 «Народное поэтическое творчество отдельных стран и народов» содержит материал конкретных фольклористических исследований. По аналогии с литературами отдельных стран и народов наименование фольклора каждого народа снабжено буквенно-цифровой нотацией: C6010 – Абхазский фольклор; C6570 – Русский фольклор и т.д. и располагаются по алфавиту.

Один и тот же документ может быть представлен в отраслевой БД по литературоведению на нескольких языках. Приведем пример использования буквенно-цифровых кодов различных рубрик в библиографической записи монографии на четырех языках:

Тип документа: монография **Шифр:** 000439711
Язык: немецкий русский английский французский
Katharina II: Eine russ. Schriftstellerin / Hrsg. von Fajnstajn M., Gopfert F. – Wilhelmshorst: Gopfert, 1996. – 57, VI S. – (Frauen Lit. Geschichte; Bd 5) .
Текст на нем., рус., англ. и фр. яз.
Аннотация: Сборник статей о литературном творчестве Екатерины II.
Ключевые слова: Екатерина II; русская литература; Russian literature; Ekaterina II
Рубрики ИНИОН: A170991; C1570; D0925; E01

Важно отметить, что один и тот же документ по литературоведению может быть отнесен к нескольким рубрикам. Приведем пример библиографической записи статьи, в которой использованы коды двух тематических рубрик:

Тип документа: статья **Язык:** английский
Шифр: 22877643
Levander, C.F.; Levine, R.S.
Hemispheric American literary history // Amer. lit. history. - Oxford, 2006. - Vol. 18, N 3. - P. 397-405.
Аннотация: Применение сравнительных методов в американистике для изучения региональных, национальных и этнических проблем.
Ключевые слова: национальное и интернациональное; national and international; сравнительное изучение литературы; comparative study of literature; литературоведение; literary criticism; США литература; american literature
Рубрики ИНИОН: A170991; C1628; A170107

При составлении библиографических записей книжных изданий (монографий, сборников статей, многотомных изданий, авторефератов диссертаций, учебников, энциклопедий, словарей и пр.) используются также Списки предметных рубрик, что обеспечивает взаимосвязь документальных массивов отраслевой БД по литературоведению и предметного каталога Фундаментальной библиотеки по общественным наукам ИНИОН. Особый тип предметных рубрик – это персоналии, количество которых в области литературоведения достаточно велико. Это имена писателей, поэтов, ученых, литературных критиков, исторических деятелей, а также названия произведений и др. Приведем пример библиографической записи многотомного издания, при формиро-

вании которой использованы различные лингвистические средства – ключевые слова (дескрипторы), предметные рубрики и коды отраслевого рубрикатора:

Тип документа: многотомник

Язык: латинский английский русский немецкий

Шифр: 09474-23

Congressus nonus internationalis Fenno-Ugristarum; 7.-13.8.2000, Tartu / Red.: Nurk A. et al. – Tartu: Tartu ulikool, 2000. Pars 3 : Summaria acroasium in sectionibus et symposiis factarum. Folkloristica et ethnologia. Litteratura. Archaeologia et anthropologia et genetica. - 412 p.

Ключевые слова: библиографические указатели; bibliographic indexes; финно-угорские литературы; finno-ugric literatures; финно-угорский фольклор; finno-ugric folklore

Предм. рубрики: Финно-угорские народы; Библиография

Рубрики ИНИОН: A1794

СТРАТЕГИИ ПОИСКА ИНФОРМАЦИИ ПО ЛИТЕРАТУРОВЕДЕНИЮ

Быстрый рост и динамичное развитие информационных ресурсов, в том числе и электронных, требует от научных работников, исследователей, литературоведов применения разнообразных навыков сложного поиска, успех которого во многом зависит от умения гибко использовать отраслевые и междисциплинарные лингвистические средства. Для того чтобы быстро и продуктивно находить информацию, важно иметь представление о различных видах поиска и возможностях использования основных лингвистических средств и стратегий.

Взаимодействие пользователя с системой АИСОН осуществляется с помощью простого интерфейса на первой странице сайта ИНИОН или профессионального многооконного интерфейса, доступного там же в разделе «Ресурсы». Интерфейс позволяет формировать поисковые запросы любой сложности, редактировать результаты поиска, модифицировать поисковую стратегию, проводить эвристический поиск, а также представлять найденную информацию в различных форматах. Для просмотра библиографических записей документов можно выбрать любой формат (в том числе – транслитерированный), однако чаще всего потребители информации используют формат Biblio View (INION) или Citation. Уточнить, в какое поле библиографической записи документа включен тот или иной термин (дескриптор, слово или другой поисковый элемент), можно по формату US MARC biblio Tags.

Наиболее сложным этапом организации и проведения поиска информации по литературоведению является анализ информационных потребностей, определение основных требований к ретроспективной глубине поиска, к объему выдачи необходимой информации, а также баз данных, в которых, скорее всего, можно найти нужную информацию. В настоящее время любой потребитель информации по литературоведению имеет возможность сформировать поисковое пространство, соответствующее его информационным потребностям, путем выбора нескольких релевантных баз данных (двух, трех и даже четы-

рех) для поиска по одному и тому же запросу. Как правило, информационный поиск проводится сначала в БД текущих поступлений, затем в отраслевой БД по литературоведению и в БД по языкознанию.

Проблемно-тематический поиск реализуется по основным полям библиографических записей документов, для которых формируются свои словари. При этом для каждого элемента словаря указывается количество документов, в описание которых он включен (ключевое слово, термин, слово или словосочетание, дескриптор, код рубрики, фамилия автора и пр.). Поиск информации может проводиться как по отдельным полям библиографической записи, так и по их комбинации. При построении запросов (поисковых предписаний) необходимо использовать знак звездочки (*) – оператор правостороннего усечения поисковых элементов, который позволяет увеличить полноту поиска, а также логические операторы, круглые скобки и др. Так, для обеспечения полноты поиска по элементам общего словаря во многих случаях необходимо вводить поисковые элементы с оператором правостороннего усечения (*). Это связано с тем, что в заглавии документа и в аннотации термины (слова и словосочетания) часто представлены в разных падежах. Усечение поискового термина, например, Пушкин*, позволит найти документы, в которых встречается не только эта фамилия в различных падежах (Пушкина, Пушкину, Пушкиным), но и такие слова, как пушкинский, пушкиноведение и др.

Поиск информации по литературоведению проводится на основе «Руководства пользователя» [9], а также инструктивных материалов, доступных на сайте ИНИОН. Для формирования запросов, выбора поисковой стратегии, анализа результатов поиска, отбора релевантных документов, редактирования поискового запроса и формирования выдачи информации применяются различные технологические возможности поисковой системы WinIRBIS. При этом предусмотрены три режима поиска, из которых наиболее часто используется так называемый поиск «по образцу», обеспечивающий выбор поисковых терминов из словарей (тезаурусов) и автоматическое установление логических связей между элементами запроса. При необходимости изменения выбранной стратегии поиска можно заменить логические операторы на другие. Модели информационного поиска в контексте поисковых задач описаны в работах [10-12].

При формировании стратегии поиска чаще всего принимаются решения, основанные на результатах изучения научной проблемы (темы), и решения, базирующиеся на результатах предварительного анализа понятийных связей наиболее важных терминов и зафиксированные в литературоведческих словарях и тезаурусах. В первом случае необходимы учет и анализ максимального количества факторов, оказывающих влияние на развитие, например, той или иной теории или концепции, исторических ситуаций или процессов, которые раскрывают информационную потребность пользователя. Во втором случае важным представляется прошлый опыт и знания, полученные им при информационном поиске ранее. При выборе стратегии поиска информации по литературоведению важно иметь в виду, что процесс решения поисковой задачи чаще всего рассматривается не просто как

нахождение релевантной запросу информации, а как стремление пользователя найти новые факты, методы, идеи. В связи с этим при определении путей и способов рационального решения научной задачи представляется важным максимально полно описать информационную потребность с помощью лингвистических средств системы и указать весь набор требований к результатам поиска (проблемно-тематические и хронологические границы, ограничения на язык и вид документов и пр.).

Выбор той или иной поисковой стратегии зависит от конкретной задачи поиска. Однако при прочих равных условиях поиск по ключевым словам (дескрипторам) для обычного пользователя является более предпочтительным, чем поиск по рубрикам отраслевого рубрикатора, особенно в случаях узкой специфичности запроса.

В системе предусмотрена возможность совместного применения нескольких стратегий поиска, обеспечивающих итерационное повышение эффективности поиска путем генерации новых поисковых предписаний на основе последовательного редактирования (или комбинирования) результатов уже выполненных поисков. Так, при формировании стратегии поиска информации о литературоведческих школах представляется важным проанализировать дескрипторную статью «школы и направления» в ИПТ и отобрать те нижестоящие дескрипторы, с помощью которых можно получить более точные результаты поиска или увеличить его полноту:

- вульгарный социологизм
- деконструктивизм
- историческая поэтика
- историческая школа
- культурно-историческая школа
- мифологическая школа
- натуральная школа
- психологическая школа
- финская школа
- формальная школа.

Стратегия сплошного просмотра документов, включенных, например, в рубрики отраслевого рубрикатора, которые начинаются с буквенного кода В (страны или группы стран) или С (конкретная литература или фольклор), позволяет достаточно быстро отбирать нужные источники в случае, если к рубрике отнесено не слишком большое количество документов.

Стратегия типового поиска «по образцу» обеспечивает возможность предварительного анализа имеющихся вариантов различных поисковых признаков и включения в запрос (поисковое предписание) лишь тех из них, которые позволят получать требуемую информацию. Например, эта стратегия может использоваться при поиске рецензий по полю типов (видов) документов.

Стратегия выборочного поиска заключается в просмотре и анализе определенной совокупности источников по определенному признаку (например, по хронологическому периоду – последние 5 лет, последнее десятилетие, определенный год издания, тип издания и пр.) и применяется в случае тематических запросов, по которым требуется найти небольшое количество документов (поиск библиографической информации по определенной теме, предмету, вопро-

су, а также о лице, организации, культурно-историческом объекте и т.п.).

Для увеличения полноты поиска система позволяет использовать найденный при поиске релевантный документ (или совокупность документов) в качестве нового запроса, рассматривая статистически значимые термины релевантных документов в качестве поисковых элементов для автоматического выполнения эвристического поиска. Применение эвристического поиска позволяет сократить время решения поисковых задач, однако получаемые при этом результаты часто не являются наилучшими – точными и полными, а относятся лишь к множеству допустимых решений.

Механизмы, заложенные в стратегию эвристического поиска, направлены на установление таких отношений в проблемной ситуации, которые позволяют, например, исключить выдачу нерелевантных документов путем отсека неперспективных ветвей в дереве поисковых вариантов по тому или иному запросу, или обеспечивают нахождение новых релевантных документов. При этом потребитель не только получает новые навыки научно-исследовательской работы, но и учится ориентироваться в информационном пространстве, управляя его изменением путем модификации исходного запроса на основе анализа результатов поиска.

Наконец, для поиска информации по литературоведческим проблемам часто приходится привлекать информацию из таких смежных дисциплин, как история, социология, антропология, психология, языкознание, философия. Еще двести лет назад в дисциплинарной структуре филологии исторически оформился союз науки о литературе с лингвистикой. Кроме того, во многих странах теория литературы уже около полувека все более тесно связывается с такими разделами философии, как эстетика, герменевтика, онтология и др. Анализ данных о количестве документов по различным проблемам литературоведения, включенным в отраслевые БД ИНИОН, показывает, что общий объем документов по рассматриваемой тематике превышает 50 тыс. Больше всего литературоведческой информации содержится в БД по языкознанию, историческим наукам, философии и социологии. Приведем пример библиографической записи монографии на английском языке, включенной в БД ИНИОН по языкознанию, в которой рассматривается взаимосвязь философии языка и структурной лингвистики с теорией литературы:

Тип документа: монография Язык: английский Шифр: 064480812 Norris, Chr. Fiction, philosophy and literary theory: Will the real Saul Kripke please stand up?. – L. ; N.Y.: Continuum, 2007. – [5], 265 p. Библиогр. в примеч. Ind.: p.261-265. Аннотация: В частности, взаимосвязь философии языка и структурной лингвистики с теорией литературы. Ключевые слова: структурная лингвистика; structural linguistics; философия языка; philosophy of language; языкознание; linguistics; литературоведение; literary criticism Предм. рубрики : Литература; Теория; Философия языка; Структурная лингвистика; Аналитическая философия Рубрики ИНИОН : A163161; A16212141; A16210755
--

Информация по литературоведению может быть найдена не только в отраслевых, но и в сводных базах данных, которые позволяют проводить ретроспективный поиск документов на заданную глубину (например, поступивших в библиотечные фонды за последние три года) и обеспечивают возможность поиска документов без обращения к отраслевому БД (например, в массиве текущих поступлений). Так, поиск информации о жизни и творчестве Н.М.Карамзина, включенной за последние тридцать лет в 5 баз данных ИНИОН, позволил найти свыше 1250 документов, посвященных изучению его произведений, анализу его писем и пр. Полученные результаты информационного поиска позволили по-новому взглянуть на роль карамзинского наследия, понять его значение для национальной культуры развития. Н.М. Карамзин по праву считается реформатором русского литературного языка. Он одним из первых в России стал употреблять букву «Ё», а также использовать множество новых слов, среди которых «ответственность», «благотворительность», «достопримечательность» и многие другие.

Приведем пример библиографической записи монографии, найденной при поиске в БД по литературоведению:

Тип документа: монография **Язык:** русский
Шифр: 100881112

Карамзин: pro et contra: Личность и творчество Н.М. Карамзина в оценке рус. писателей, критиков, исследователей: Антология / Сев.-Зап. отд-ние Рос. акад. образования, Рус. Христиан. гуманист. акад.; Отв. ред.: Бурлака Д.К.; Сост.: Сапченко Л.А. – СПб.: Изд-во Рус. Христиан. гуманист. акад., 2006. – 1080 с., портр. – (Рус. Путь)
Библиогр.: с. 864-866 и в коммент. Указ. имен: с. 1060-1072.

Аннотация: Сапченко Л.А. Карамзин в движении времени, с. 7-24.

Ключевые слова: Карамзин Н М; Karamzin N M; русская литература; Russian literature; литературная критика; literary criticism; Россия; Russia

Предм. рубрики: Карамзин НМ

Рубрики ИНИОН: А170991; С1570; D0925; E01

Для повышения качества научно-информационного обеспечения потребителей информации в области литературоведения в ИНИОН РАН создается база данных реферативной информации, которая доступна на оптических дисках с 2003 г. и на сайте Научной электронной библиотеки (www.elibrary.ru). Важно также отметить, что одним из научных изданий ИНИОН, представляющих интерес для отечественных и зарубежных литературоведов, является «Литературоведческий журнал», который выпускается с 1993 г. В этом журнале публикуются научные статьи по истории отечественной и зарубежной литературы, по теории литературы, а также хроника литературной жизни и библиография по литературоведению. Как правило, номера журнала организованы по тематическому принципу и связаны с одним автором или темой. Выпуски нередко привязаны к значимым литературным датам: например, к 750-летию рождения Данте, 450-летию рождения У.Шекспира, 200-летию рождения А.С. Пушкина. Среди публикаций этого

журнала, представленных в № 28 за 2011 г., особый интерес представляет статья М. Альтшуллера, посвященная описанию споров Н.М. Карамзина с А.С. Шишковым о судьбах России [13]. Очередной номер журнала приурочен к 250-летию со дня рождения Н.М.Карамзина.

ЗАКЛЮЧЕНИЕ

Основными направлениями развития информационных ресурсов по литературоведению является стремление их создателей к реализации универсальности этих ресурсов и соединению информации для обычного читателя и научной информации, необходимой исследователям-литературоведам, литературным критикам, преподавателям высшей школы. Новые технологии хранения, распространения и использования информационных ресурсов по литературоведению определяются, с одной стороны, их всеобщей значимостью, а с другой – необходимостью высокой степени их доступности в сети Интернет и высокого уровня удобства этих ресурсов для пользователя. Особенности информационных ресурсов в области художественной литературы и литературоведения рассматриваются в работе [14]. В соответствии с библиографической традицией литературоведческие и литературно-критические информационные ресурсы, доступные в сети Интернет, предлагается разделить на три группы: научные литературоведческие, литературно-критические и методические.

Современный литературный Интернет включает, прежде всего, полнотекстовые электронные библиотеки художественных произведений, авторские страницы писателей и поэтов, а также сайты литературно-художественных журналов. Самая известная в Рунете электронная библиотека художественной литературы – это «Библиотека Максима Мошкова» (<http://lib.ru>).

Многие информационные ресурсы литературоведческой направленности содержат свод самих художественных текстов, приобретая тем самым универсальный характер. Так, крупнейшая в российском сегменте Интернета гуманитарная электронная библиотека представлена на сайте E-Lingvo.net (<http://e-lingvo.net/>). В ней находятся как сами художественные тексты на русском, английском, немецком, французском и других языках, так и научные работы известных филологов и литературоведов.

Информационный портал Philology.ru, который создан более 15 лет назад, содержит библиотеку текстов работ, имеющих печатные аналоги, включая произведения классиков, монографии, хрестоматийные тексты, статьи, методические пособия; литературоведение представлено следующими разделами: общее литературоведение (теория литературы, история литературы, стиховедение, фольклористика и пр.), исследования по русской литературе (от фольклора и древнерусской литературы до литературы начала XXI в.), литературе Европы и Америки (эпоха античности – XX в.) и литературе Азии и Африки. Раздел Personalia включает публикации о крупнейших литературоведах (<http://www.philology.ru/literature.htm>).

Особый интерес для литературоведов представляет Фундаментальная электронная библиотека (ФЭБ) «Русская литература и фольклор», которая обеспечивает доступ к полным текстам произведений русской

словесности, а также библиографии, научным исследованиям и историко-биографическим работам; она содержит информацию различных видов (текстовую, звуковую, изобразительную) в области русской литературы XI-XX вв. и русского фольклора, а также по истории русской филологии и фольклористики. Специфика ФЭБ состоит в том, что, с одной стороны, это репозиторий текстов (источников, исследовательской и справочной литературы), а с другой – эффективный инструмент для их анализа (www.feb-web.ru).

Одним из наиболее полных мировых источников информации по литературоведению является информационная система EBSCO, в которой содержатся четыре базы данных: MLA International Bibliography, Academic Search Complete, MasterFile Complete, and ATLA Religion Database with ATLA Serials. Необходимо отметить, что поиск литературоведческой информации может проводиться во всех базах данных одновременно (<https://www.ebscohost.com>).

В заключение следует отметить, что информационное обеспечение в области литературоведения составляет важную часть информационной политики по формированию информационной культуры постоянно растущей российской аудитории, интересующейся литературой и фольклором. Реализация политематического подхода к поиску информации по литературоведению требует использования новых информационно-коммуникационных технологий и доступа к полнотекстовым электронным ресурсам (отечественным и зарубежным).

СПИСОК ЛИТЕРАТУРЫ

1. Махлин В.Л. Бахтин и современное литературоведение // Вопросы литературы. – 1996. – № 3. – С.65-67.
2. Сегал Д.М. Пути и вехи: Русское литературоведение в двадцатом веке. – М.: Водолей, 2011. – 280 с.
3. Багиров А.М. Глобализация и диалог литератур // Актуальные проблемы гуманитарных и естественных наук. – 2014. – № 4. – С. 332-336.
4. Кто есть кто в российском литературоведении: Библиогр. словарь-справочник / РАН. ИНИОН. Центр гуманит. науч.-информ. исслед. Отд. литературоведения; Гл. ред. А.Н. Николюкин; Науч. ред.: А.А. Ревякина; Сост.: Т.Н. Красавченко и др. – М., 2011. – 406 с.
5. Кто есть кто в русском литературоведении [Текст] : справочник / АН СССР, Ин-т науч. информ. по обществ. наукам, Ин-т мировой лит. им. А. М. Горького [Сост.: А.В. Дранов и др. ; редкол.: А.Н. Николюкин (гл. ред.) и др.]. – М.: ИНИОН АН СССР
Ч. 1. А-И. – 1991. – 211 с.
Ч. 2. К-О. – 1992. – 231 с.
Ч. 3. П-Я. – 1994. – 262 с.
6. Информационно-поисковый тезаурус ИНИОН по литературоведению / РАН. ИНИОН. Центр информатизации. Фундаментальная библиотека [Сост.: М.Я. Курашова, В.Б. Смиренский, Г.П. Швырева, О.В. Этова при участии Р.Р. Мдивани] (Классификационный указатель по ББК) /

ред. Е.В. Магай; отв. ред. Р.Р. Мдивани. – М.: ИНИОН РАН, 2010. – 112 с.

7. Рубрикатор Автоматизированной информационной системы по общественным наукам (АИСОН) / РАН. ИНИОН. Фундам. Б-ка, Отд. науч.-библиограф., Отдел каталогизации и электронных каталогов. Центр информатизации / сост. Г.С. Антонюк и др.; отв. ред. А.И. Слива, В.Л. Лейбович, М.Б. Шнайдерман. – М., 2012. – 284 с.
8. Шемберко Л.В., Шнайдерман М.Б., Слива А.И. Лингвистический навигатор по социальным и гуманитарным наукам: назначение, структура и принципы применения // Научно-техническая информация. Сер. 1. – 2014. – № 11. – С.26-37; Shemberko L.V., Shnaiderman M.B., Sliva A.I. A Linguistic Navigator for the Social Sciences and Humanities: Purpose, Structure, and Application Principles // Scientific and Technical Information Processing. – 2014. – Vol. 41, № 4. – P. 244-253.
9. Библиографические базы данных. Руководство пользователя = Bibliographical Databases. User's Manual / РАН. ИНИОН – М., 2001. – 34 с. (IRBIS Software).
10. Голицына О.Л., Максимов Н.В. Модели информационного поиска в контексте поисковых задач // Научно-техническая информация. Сер. 2. – 2011. – № 2. – С.1-13; Golitsyna O.L., Maksimov N.V. Information Retrieval Models in the Context of Retrieval Tasks // Automatic Documentation and Mathematical Linguistics. – 2011. – Vol. 45, № 1. – P.20-32.
11. Xie H.I. Shifts in information-seeking strategies in information retrieval in the digital age. A planned-situational model // Information Research. – 2007. – Vol. 12, № 4. – URL: <http://www.informationr.net/ir/12-4/colis/colis22.html>
12. Xie H. Patterns between interactive intentions and information-seeking strategies // Information Processing & Management. – 2002. – № 38(1). – P. 55-77.
13. Альтшуллер М. Шишков и Карамзин в споре о судьбах России // Литературовед. журнал. – 2011. – № 28. – С. 43-57.
14. Буранбаев А.М. Информационные ресурсы в области художественной литературы и литературоведения в сети Интернет: опыт анализа // Современная техника и технологии. – 2016. – № 2 (54). – URL: <http://technology.snauka.ru/2016/02/9493>

Материал поступил в редакцию 19.06.16.

Сведения об авторах

ШЕМБЕРКО Людмила Винцентовна – зав. сектором Института научной информации по общественным наукам (ИНИОН) РАН, Москва
e-mail: irichem@mail.ru

СЛИВА Алексей Иванович – кандидат исторических наук, зав. отделом Института научной информации по общественным наукам (ИНИОН) РАН
e-mail: aslivainion@mail.ru

Цифровые музейные коллекции и представление объектов естественнонаучного музейного хранения в электронной библиотеке «Научное наследие России»

Рассматриваются общие вопросы отражения музейной информации в сети Интернет и формирования цифровых музейных коллекций, виртуальных музеев и выставок. Обсуждаются проблемы создания и предоставления пользователям Интернета мультимедийных коллекций естественнонаучных музеев как части интегрированной электронной библиотеки «Научное наследие России». Приводятся примеры успешного включения ресурсов ряда музеев в эту библиотеку.

Ключевые слова: научная музейная информация, цифровые коллекции, виртуальный музей, виртуальные выставки, электронная библиотека «Научное наследие России»

ВВЕДЕНИЕ

Развитие сетевых технологий открывает широкие перспективы для популяризации музейных сокровищ. Сегодня музеи мира раскрывают свои фонды не только для посетителей, пришедших в музей в часы его работы, но и для значительной части населения Земного шара 24 часа в сутки 7 дней в неделю. Заинтересованный пользователь Интернета может получить представление не только об экспонатах музея, демонстрируемых в данный момент, но и о его запасах, представить которые в экспозиции не позволяют различные ограничения (по площадям, ресурсам, персоналу и т.д.). И хотя реальное знакомство с музейными экспонатами оказывает на посетителя музея существенно более сильное эмоциональное воздействие, нежели просмотр на экране «оцифрованных» объектов, востребованность музейной информации, представленной в сетевом доступе, непрерывно возрастает. Музеи размещают на своих сайтах не только информационные материалы (общее описание фондов, историю создания музея, каталоги фондов и коллекций), но и отсканированные копии документов, цифровые фотографии хранящихся объектов, мультимедийные материалы о музее и его фондах.

ЦИФРОВЫЕ КОЛЛЕКЦИИ МУЗЕЕВ

Наибольшее распространение в мире получили сайты художественных музеев, представляющие интерес для значительного числа людей, независимо от национальности, уровня образования и профессии. Все крупнейшие художественные музеи мира предоставляют в свободном доступе свои цифровые коллекции, которые содержат детальные описания музейных объектов, их цифровые образы. На основе

этих коллекций формируются виртуальные тематические выставки, виртуальные туры по музею. Это позволяет привлечь большее количество посетителей сайта музея, способствует популяризации его экспозиции.

Современные программные средства позволяют создавать и демонстрировать трёхмерные модели объектов, не только существующих в настоящее время (на основе обработки фотографий), но и полностью утраченных (на основе обработки их сохранившихся фрагментов и документов). В качестве примера можно привести Государственную Третьяковскую галерею, где хранится созданная Александром Ивановым серия акварелей и графических этюдов («Библейские эскизы») для росписи строившегося тогда в Москве храма Христа Спасителя. Во время проведения одной из выставок, посвященной творчеству А.Иванова, с помощью 3D-проекции посетителям были показаны «Библейские эскизы» приблизительно в том размере, в каком они должны были бы быть в храме Христа Спасителя. В результате появилось новое качество восприятия этих произведений.

Достижения информатики открывают новые возможности по формированию «виртуальных музеев».

Виртуальный музей – это тип веб-сайта, адаптированный для экспозиции музейных материалов – от предметов искусства и исторических артефактов до новейших технологических достижений. За счёт применения интернет-технологий виртуальный музей позволяет решать проблемы экспонирования большего (по сравнению с количеством экспонатов, ограниченным физическим помещением музея или выставочного зала) количества предметов, хранящихся в запасах. Виртуальный музей обеспечивает быстрый и лёгкий доступ к цифровым образам му-

зейных объектов широкому кругу пользователей. В отличие от простых коллекций фотографий или электронных книг, виртуальный музей обладает расширенными поисковыми возможностями для нахождения и классификации экспонатов. Данные об экспонатах могут быть представлены на нескольких языках. Цифровые технологии создания и поддержки виртуальных музеев позволяют создавать мультимедийное описание экспоната и помещать его в «виртуальную реальность».

Виртуальная выставка – это частный случай виртуального музея. Виртуальная выставка, как правило, посвящена конкретной тематике, событию или персоне. Создавая виртуальные выставки, музеи решают такие задачи, как:

- сохранение экспозиции на более длительный срок, чем это предусматривает традиционная выставка;
- предоставление возможности удаленному пользователю воспользоваться ресурсами музея.

На сегодняшний день большая часть современных музеев предоставляет посетителям своих сайтов электронные каталоги имеющихся коллекций, виртуальные туры по музею, электронные буклеты или путеводители. Некоторые художественные музеи, такие как Метрополитен музей (США), Лондонская национальная галерея, предлагают пользователям тысячи изображений предметов искусства, доступных для скачивания в высоком разрешении.

Отдельную «нишу» в музейном мире занимают научные музеи. Среди них музеи, посвященные отдельным разделам науки (геологии, ботанике, зоологии и т.д.), отраслям науки и техники (Политехнический музей, Музей космонавтики, Музей естествознания и т.п.), отдельным ученым (музей К.Э. Циолковского, музей Д.И. Менделеева, музей Н. Теслы и др.).

В отличие от художественных музеев, научные музеи хранят в своих фондах значительное число рукописных и печатных материалов, образцы техники, модели, демонстрирующие научные открытия и т.п.

Например, Нью-Йоркский Музей естествознания (Музей естественной истории), обладающий самой большой в мире коллекцией в области естественных наук, хранит библиотеку, в которой собрано более 450 тыс. документов по естествознанию. Собрание Музея включает также киноматериалы, видеозаписи и отчеты об экспедициях в различные точки Земли и многое другое. Цифровые коллекции библиотеки этого Музея представляют собой базы данных оцифрованных архивных документов, фотографий, изображений предметов искусства и памятников, иллюстраций редких книг.

Наряду с Нью-Йоркским Музеем естествознания, одним из крупнейших научно-технических музеев мира является московский Политехнический музей. Здесь собраны устройства и предметы, иллюстрирующие этапы развития технической мысли. На сайте музея можно ознакомиться не только с цифровыми образами некоторых экспонатов, но и с описанием работы отдельных механизмов, а также с биографиями ученых, чьи работы составляют коллекции музея. Кроме того, Политехнический музей предлагает пользователям своего сайта «посетить» тематические виртуальные выставки.

Роль научных музеев в жизни общества трудно переоценить – они необходимы для целей образования, популяризации науки, для развития научных исследований в области истории науки. Предоставление возможности ознакомления с научными музеями через Интернет существенно расширяет аудиторию их «посетителей» и повышает уровень значимости музеев для общества. Однако создание отвечающих современным требованиям виртуальных научных музеев несравненно более сложная задача по сравнению с музеями художественными. Для получения максимального эффекта современный виртуальный научный музей должен содержать не только электронную библиотеку (оцифрованные печатные, рукописные и фотодокументы, снабженные метаданными и развитым поисковым аппаратом), но и модели, демонстрирующие природные явления, научные открытия, технические решения. В виртуальном естественнонаучном музее может быть показано не только то, что где-то можно реально увидеть человеческим глазом, но и полностью смоделированные процессы, например, «мир глазами стрекозы» или ускоренный в миллионы раз процесс эволюции растений от ранних водорослей до цветковых растений и т.п.

Музейная научная информация, хотя и ценна сама по себе, наиболее востребована в комплексе с другими видами научной информации – публикациями и архивными материалами. В современном мире идет активное движение в направлении развития институтов социальной памяти, что предусматривает интеграцию информационных ресурсов и формирование единого информационного пространства библиотек, музеев, архивов [1]. Поэтому, наряду с задачами, решаемыми в интересах отдельных музеев, возникает задача отражения музейной информации в интегрированных научных ресурсах. Примером работы по формированию виртуального «объединенного» института социальной памяти может служить электронная библиотека «Научное наследие России» (ЭБ ННР) [2-4].

ПРЕДСТАВЛЕНИЕ ЦИФРОВЫХ КОЛЛЕКЦИЙ ЕСТЕСТВЕННОНАУЧНЫХ МУЗЕЕВ В ЭЛЕКТРОННОЙ БИБЛИОТЕКЕ «НАУЧНОЕ НАСЛЕДИЕ РОССИИ»

Электронная библиотека «Научное наследие России», разработка которой ведется с 2008 г. в Российской академии наук, а затем в Федеральном агентстве научных организаций (ФАНО) России, представляет собой общедоступную через Интернет (<http://e-heritage.ru>) электронную библиотеку, основными задачами которой являются:

- информационная поддержка научных исследований и образовательных программ в области истории науки;
- формирование и сохранение развернутой информации о выдающихся ученых, внесших вклад в развитие российской науки;
- предоставление свободного доступа широкому кругу пользователей Интернета в России и за рубе-

жом к научному наследию российских ученых, внесших значительный вклад в развитие отечественной и мировой науки;

- популяризация научного знания.

В основу технологии формирования этой Библиотеки положен принцип распределенного наполнения и централизованной поддержки. ЭБ ННР – это стандартизированная среда управления информацией электронных фондов, разработанная для интеграции ресурсов библиотек, архивов и музеев, содержащих объекты научного наследия и связанные с ними метаданные из различных источников, обеспечивающая обмен информацией и ее совместное использование посредством Интернета [2].

Основными участниками проекта по формированию и поддержке ЭБ ННР являются: Межведомственный суперкомпьютерный центр (МСЦ) РАН, Библиотека Академии наук (БАН), Библиотека по естественным наукам РАН (БЕН РАН), Институт научной информации по общественным наукам (ИНИОН РАН), Государственный геологический музей им. В.И. Вернадского РАН, Государственный минералогический музей им. Ферсмана РАН, а также ещё более 20 организаций, входящих в структуру ФАНО.

ЭБ ННР активно взаимодействует с Российским государственным архивом, Государственным биологическим музеем имени К. А. Тимирязева, Российским государственным архивом кинофотодокументов и др.

ЭБ ННР ориентирована на широкий круг пользователей. Её ресурсы полезны школьникам, студентам, научным работникам, историкам науки.

В информационном пространстве Библиотеки размещены сведения об ученых, в том числе их развернутые биографии со ссылками на источники, электронные копии печатных и архивных документов (как с авторством самих ученых, так и документов, посвященных научным проектам, научным событиям, научным школам и т.д.), а также электронные образы связанных с учеными объектов музейного хранения [5].

Отбор материалов для включения в ЭБ ННР осуществляется в соответствии с принципами, устанавливаемыми Советом электронной библиотеки, в который входят представители руководства организаций – участников проекта.

На начальном этапе создания Библиотеки было принято решение формировать контент, начиная с первых российских ученых и их публикаций, постепенно «продвигаясь» по временной «шкале». Это обусловлено тем, что оцифровка материалов XVIII – начала XX вв. не требует заключения каких-либо договоров с правообладателями, предусмотренных законодательством об охране авторских прав, что существенно упрощает процесс наполнения Библиотеки.

Другим принципиальным решением на данном этапе является отказ от распознавания текста отсканированных материалов. Это обусловлено проблемами распознавания естественнонаучных текстов, текстов, представленных в старой русской орфографии и на иностранных языках (на этапе становления российской науки научные работы публиковались на латыни, французском, немецком языках, что отраже-

но в электронной библиотеке). В то же время для повышения уровня сервиса при работе с электронными изданиями в ЭБ ННР имеется возможность навигации по тексту с использованием распознанного (или введенного вручную) оглавления изданий.

В соответствии с требованиями максимально точного воспроизведения печатных изданий при их сканировании используется разрешение 600 точек на дюйм, а обрезка и очистка изображений выполняется специально разработанными средствами, исключающими возможность потери информации (в том числе замечок на полях изданий).

Конкретные материалы для включения в ЭБ ННР предлагают участники проекта, предложения рассматриваются и принимаются редакционной группой Межведомственного суперкомпьютерного центра РАН. Вся технология формирования предложений, передачи их в редакционную группу и утверждения осуществляется в автоматизированном режиме в рамках специального сетевого технологического блока Библиотеки [6].

Разработанное участниками проекта программное обеспечение ЭБ ННР позволяет создавать и поддерживать коллекции, хранящие архивные и музейные ресурсы, а также иметь любые вложенные наборы объектов.

Сотрудничество с музеями – важная и необходимая задача ЭБ ННР. Это сотрудничество основано на работе с научными сотрудниками музеев, определяющими, какие именно музейные коллекции и предметы должны быть оцифрованы и представлены в Библиотеке в первую очередь.

Материалы музея могут быть представлены в ЭБ ННР в виде цифрового образа объекта, сопровождаемого либо развернутой информацией о нем (включающей все его характеристики, используемые в музейной технологии), либо ограниченным набором ранее введенных метаданных, обеспечивающих поиск и связь музейного объекта с объектами других типов. Во втором случае полная информация о музейном объекте хранится на сайте музея, а в ЭБ ННР вводится ссылка на нее. Более подробно технология передачи и представления данных о музейных объектах описана в [7].

Не у всех музеев есть возможность создавать свою базу электронных коллекций. Если оцифрованных материалов нет, то разработчики ЭБ ННР могут оказать свою помощь в подготовке метаданных и формировании цифровых образов объектов.

Материалы, представленные в ЭБ ННР, по логике ее построения связаны с персоналией, т.е. все элементы Библиотеки, в том числе, музейные предметы и коллекции, должны быть связаны с конкретным ученым. В то время как электронные коллекции, создаваемые самими музеями, в первую очередь, объединены тематикой или «научным направлением», т.е. формируются по тематическому принципу. В этом состоит основное отличие музейных коллекций, отображаемых на портале ЭБ ННР, от коллекций на сайтах естественнонаучных музеев. Рассмотрим это на примере сайта партнера ЭБ ННР – Государственного биологического музея имени К.А. Тимирязева. На сайте музея (рис. 1) посетителю предлагается ознакомиться

ся с тематическими коллекциями музея, совершив виртуальный тур по его залам.

С основной страницы пользователь, перейдя, например, по вкладке «Происхождение и становление человека» (в левой части рис. 1), попадает в соответствующий виртуальный зал музея, где расположены витрины с экспонатами, в частности, антропологические реконструкции, созданные отечественным учёным М.М. Герасимовым, вошедшим в историю мировой науки как автор метода пластической реконструкции. Воспользовавшись гиперссылками на предлагаемые экспонаты, пользователь может подробно ознакомиться с ними.

Информация о М.М. Герасимове и его скульптурных антропологических реконструкциях, хранящихся в Государственном биологическом музее им. К.А. Тимирязева, отражена в ЭБ ННР. Ознакомиться с ней, пользуясь интерфейсом Библиотеки (<http://e-heritage.ru>), можно двумя путями: первый предполагает переход к музейным объектам от связанного с ними ученого, второй – путем задания известного атрибута одного из музейных объектов.

Воспользовавшись первым путем, пользователь должен, перейдя по вкладке «Атрибутный поиск ученых», задать фамилию «Герасимов» и (при желании) имя «Михаил».

Затем по активной ссылке «Герасимов Михаил Михайлович», являющейся результатом поиска, пользователь получит страницу Библиотеки, относящуюся к М.М.Герасимову (рис. 2).

С этой страницы (см. рис. 2) можно перейти на страницу с подробной биографией М.М. Герасимо-

ва, на перечень его публикаций, имеющихся в ЭБ ННР, на страницу «музейная информация», содержащую наименования коллекций, относящихся к М.М. Герасимову, или на страницу, содержащую список музейных предметов, связанных с М.М. Герасимовым (рис. 3)

Каждое наименование предмета из списка, представленного на рис.3, является активной ссылкой, перейдя по которой можно получить подробную информацию о предмете. На рис. 4 представлена информация об антропологической реконструкции (Мужчина из Ла-Шапел-о-Сен. Франция. Неандерталец). Этот экспонат присутствует и на виртуальной витрине сайта Музея, однако его описание там не приведено. Интерфейс ЭБ ННР позволяет увеличить и просмотреть отдельно изображение экспоната.

Если пользователю известны элементы наименования интересующего объекта, то он может обратиться к вкладке основной страницы ЭБ ННР «Атрибутный поиск музейных предметов», задать известный фрагмент и получить список наименований объектов, удовлетворяющих запросу, каждое наименование является активной ссылкой на информацию об объекте, аналогичную приведенной на рис. 4.

Таким образом, взаимодействие музеев с электронной библиотекой может носить «взаимовыгодный характер» – музеи могут использовать технологию ЭБ ННР и ее платформу для решения музейных задач, например, для создания виртуальных выставок и электронных коллекций, а ЭБ ННР, в свою очередь, – использовать музейный контент для расширения информационного пространства по научному наследию.

Виртуальный визит
1. Природа и человек
2. Мир животных
3. Растения, грибы, бактерии
4. Царство грибов
6. Питание. Пищеварение. Обмен веществ
7. Кровь и кровообращение
8. Нервная и эндокринная система
9. «Прозрачная наука»
10. Жизнь растений
11. Смотри в оба!
12. Основы эволюционной теории
13. Происхождение и становление человека
14. Подводный мир в волшебных шарах
15-16. Основы генетики
17. Развитие жизни на Земле

[Главная](#) / [Экспозиция](#)

Виртуальный визит

Основная экспозиция Биологического музея им. К. А. Тимирязева, фотографии лета 2010 года.



Рис. 1. Страница сайта музея имени К.А. Тимирязева со списком виртуальных экспозиций



Рис. 2. Страница Электронной библиотеки, относящаяся к М.М. Герасимову

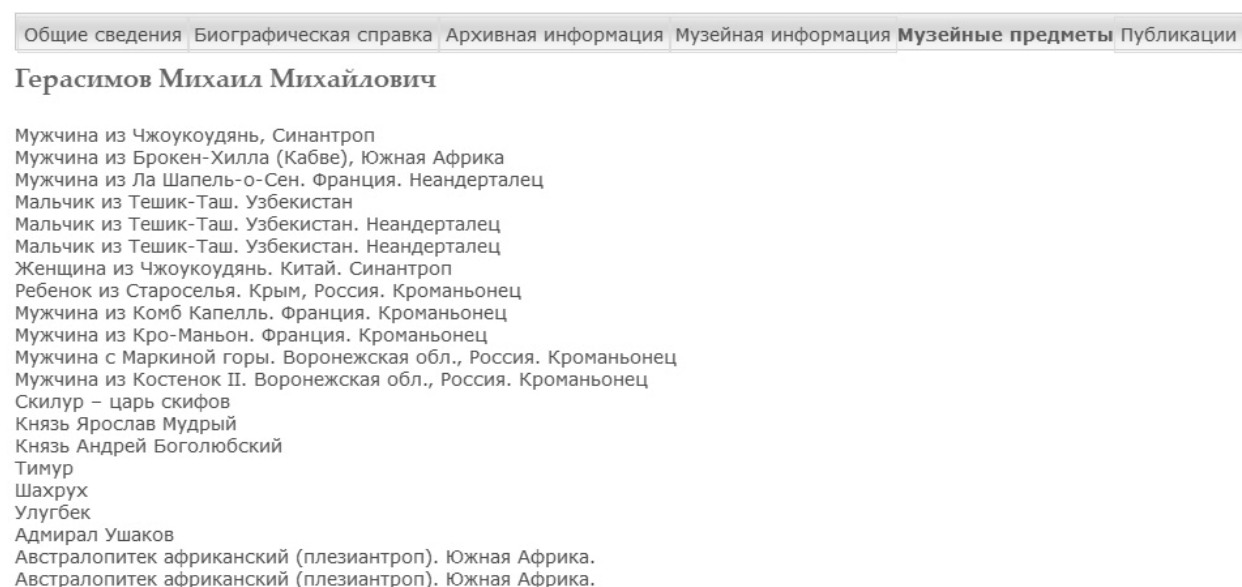


Рис. 3. Страница Электронной библиотеки, содержащая список музейных предметов, связанных с М.М. Герасимовым

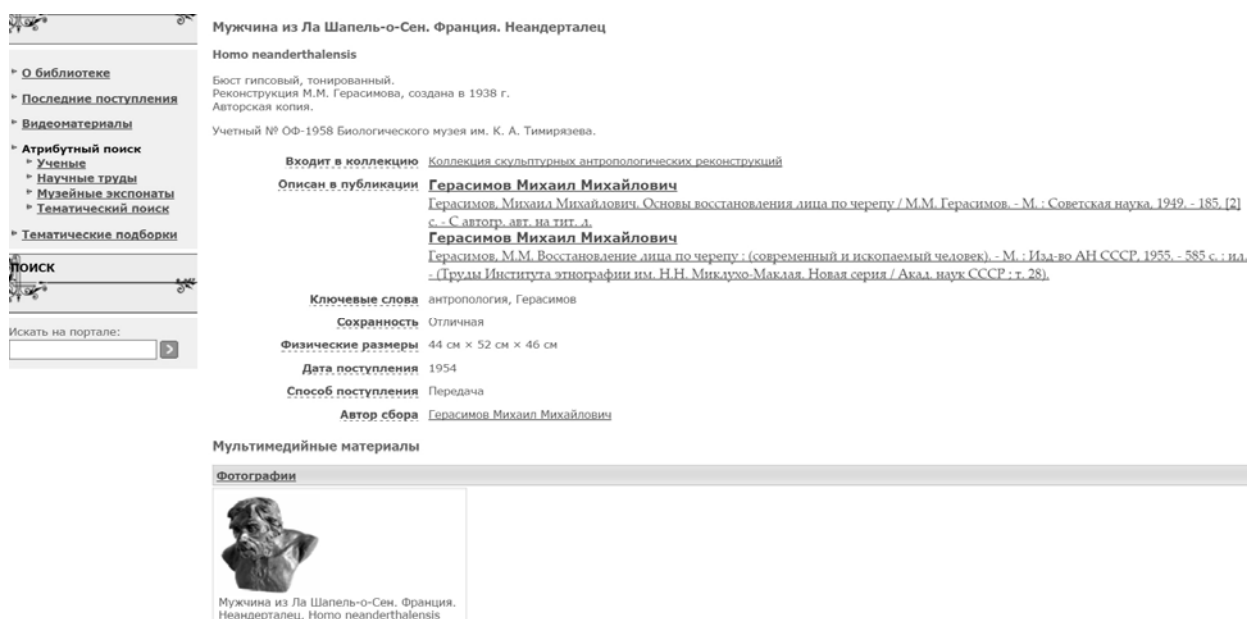


Рис. 4. Мужчина из Ла-Шапель-о-Сен. Франция. Неандерталец

В настоящее время в электронной библиотеке «Научное наследие России» содержатся данные о более чем 5000 российских ученых. Лишь очень небольшая часть из них сопровождается музейной информацией. Для дополнения контента ведется активная работа с различными музеями на предмет выявления объектов, связанных с той или иной персоналией, последующей их обработки и включения в Библиотеку. Эти работы в настоящее время поддерживаются грантом РФФИ (проект 16-07-00765а).

СПИСОК ЛИТЕРАТУРЫ

1. Цветкова В.А., Кочукова Е.В. Библиотека как институт социальной памяти // Библиография и книговедение. – 2015. – № 2. – С. 46-49.
2. Каленов Н.Е., Савин Г.И., Серебряков В.А., Сотников А.Н. Принципы построения и формирования электронной библиотеки «Научное наследие России» // Программные продукты и системы. – 2012. – Т. 4, № 100. – С. 30-40.
3. Каленов Н.Е. Интегрированная электронная библиотека «Научное наследие России» // Информационный бюллетень РБА. – 2014. – № 71. – С. 92-94.
4. Антопольский А.Б., Атаева О.М., Серебряков В.А. Среда интеграции данных научных библиотек, архивов и музеев «LibMeta» // Информационные ресурсы России. – 2012. – № 5. – С. 8-12.
5. Каленов Н.Е. Соболевская И.Н. Сотников А.Н. О взаимодействии электронной библиотеки «Научное наследие России» с естественнонаучными музеями // Информационные ресурсы России. – 2015. – № 148. – С. 2-6.
6. Якшин М.М. Платформа SciRus – основа технологического комплекса электронной библиотеки «Научное наследие России» // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: XVI Всероссийская научная конференция RCDL-2014 (Дубна, 13-16 октября 2014 г.): труды конференции. – М., 2014. – С. 362-368.
7. Каленов Н.Е. Соболевская И.Н. Сотников А.Н. Интеграция музейной информации в электронную библиотеку «Научное наследие России» // Роль музеев в информационном обеспечении исторической науки: сб. статей / авт.-сост. Е.А. Воронцова; отв. ред. Л.И. Бородин, А.Д. Яновский. – М.: Этерна, 2015. – С. 445-450.

Материал поступил в редакцию 20.07.16.

Сведения об авторах

КАЛЕНОВ Николай Евгеньевич – доктор технических наук, профессор, директор Библиотеки по естественным наукам РАН, Москва
e-mail: nek@benran.ru

СОБОЛЕВСКАЯ Ирина Николаевна – кандидат физико-математических наук, научный сотрудник Межведомственного суперкомпьютерного центра РАН – филиала ФГУ ФНЦ НИИСИ РАН, Москва
e-mail: nik_first@mail.ru

СОТНИКОВ Александр Николаевич – доктор физико-математических наук, профессор, заместитель директора по научной работе Межведомственного суперкомпьютерного центра РАН – филиала ФГУ ФНЦ НИИСИ РАН
e-mail: asotnikov@jscc.ru

Центр (Отдел) научно-информационного обслуживания (ЦНИО) ВИНИТИ РАН

Информационные услуги, предоставляемые ЦНИО ВИНИТИ РАН:

- проведение тематического поиска и консультации поисковых экспертов;
- подготовка списков научной литературы;
- подбор, копирование полнотекстовых материалов из первоисточников на бумажном носителе и в электронном виде;
- библиометрическая оценка публикационной активности исследователей и научных организаций с использованием российских и зарубежных баз данных;
- информационное обеспечение информационно-аналитической деятельности по подготовке и предоставлению аналитических обзоров и других научных материалов.

ВИНИТИ РАН располагает следующими информационными ресурсами:

- фондом НТЛ, включающим более 2,5 млн. отечественных и иностранных журналов, книг, депонированных рукописей, авторефератов диссертаций и другой научной литературы, ретроспектива – с 1991 года;
- базами данных и Интернет-ресурсами: БД ВИНИТИ (разработка ВИНИТИ), БД SCOPUS, БД Questel (патенты) и другими реферативными ресурсами;
- полнотекстовыми электронными ресурсами (статьи, патенты, материалы конференций).

Ознакомиться с информацией о доступных полнотекстовых и реферативных ресурсах можно на сайте ВИНИТИ www.viniti.ru

К услугам пользователей – **Электронный Каталог ВИНИТИ** <http://catalog.viniti.ru>
и **служба электронной доставки документов.**

Осуществляется платное информационное обслуживание по разовым заказам и на договорной основе с предоставлением всех необходимых финансовых документов.

Проводится индивидуальное обслуживание пользователей в читальном зале ЦНИО ВИНИТИ.

Обращаться в ЦНИО ВИНИТИ:

- адрес: 125190, Россия, г. Москва, ул. Усиевича, 20;
- телефоны: 8(499) 155 -42 -43, 8(499) 155 -42 -17;
- эл. почта cnio@viniti.ru, fdk@viniti.ru;
- факс 8(499) 930 -60 -00 (для ЦНИО).

База данных (БД) ВИНИТИ РАН

Федеральная база отечественных и зарубежных публикаций по естественным, точным и техническим наукам, генерируется с 1981 г., обновляется ежемесячно, пополнение составляет около 1 млн документов в год. Тематическое наполнение соответствует реферативному журналу ВИНИТИ. Для поиска одновременно по всем или нескольким тематическим фрагментам генерируется единая Политематическая БД.

БД ВИНИТИ РАН в сети INTERNET

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

На основе БД ВИНИТИ 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 ВИНИТИ** выполняют тематический поиск по разовым или постоянным запросам, а также окажут **консультационные услуги**.

БД ВИНИТИ РАН на CD-ROM

Любые наборы тематических фрагментов БД ВИНИТИ или их разделов за любой период с 1981 г., а также **проблемно-ориентированные выборки** из БД ВИНИТИ по актуальным направлениям научных исследований могут быть предоставлены на договорной основе **в поисковой системе (ИПС) "Сокол"**, работающей под управлением Microsoft Windows и обеспечивающей следующие возможности:

- **Чтение** документов в режиме последовательного просмотра или выборочно по оглавлению за весь период заказанной ретроспективы
- **Поиск** документов по автору, заглавию, источнику, ключевым словам или словосочетаниям, реферату, рубрикам, году издания, стране, языку и т.д. (всего более 20 признаков)
- **Словарь** системы поможет правильно подобрать термины для поиска и выбрать глубину их усечения.
- Для **уточнения поиска** можно дополнительно использовать год издания документа, язык текста документа, рубрики, шифры тематических разделов БД.
- Выполненные **запросы можно сохранять** для их последующего использования и/или редактирования.

125190, г. Москва, ул. Усиевича, 20, БД ВИНИТИ РАН.

Отдел взаимодействия с потребителями – (499) 155-45-25, (499) 152-58-81

E-mail: csbd@viniti.ru, sales@viniti.ru

WWW: <http://www.viniti.ru>