

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. Рассмотрены 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: электроника, энергетика 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

СТАНОВЛЕНИЕ НОВОЙ ДИСЦИПЛИНЫ

Появление термина наноинформатика (НИ) для новой области, порожденной потребностями нано- и информационных технологий (ИТ), можно отнести примерно к 2010г., 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 последующих 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 нем заняло уточнение самого понятия «НИ». По мнению авторов [1], вполне аналогично с био- или экоинформатикой, НИ 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 (НМ); интеграция данных и распространение информации. При этом, хотя разработчики дорожной карты подчеркивают междисциплинарный характер наноинформатики, подавляющее большинство рассмотренных ими проектов ориентированы 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 НМ лишь незначительно расширило ранее созданную инфраструктуру. К примеру, на портале (<http://bioportal.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 (как например *Medical Subject Headings – MESH*) лишь захватывают небольшое число терминов, связанных с, так называемой, наномедициной, т. е. проблемами использования НМ для диагностики и терапии. Процесс естественного встраивания данных по НМ в давно сложившуюся инфраструктуру медико-биологической информатики определил быстрое оформление НИ как новой дисциплины со своими БД, классификаторами, онтологиями и т.п., а с другой стороны, привел к определенному тематическому «перекоосу», отразившемуся на подборе типовых наноструктур, номенклатуре свойств,

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

выборе инструментов и технологий работы с данными. В итоге, большинство проектов, реализуемых, согласно дорожной карте [1], не учитывают особенности данных, присущих другим областям применения НМ: электроника, машиностроение, энергетика и др. В частности, НИ 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 данной работы: проанализировав основные результаты НИ применительно к потребностям EHS-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 НМ, что облегчает разработку информационных ресурсов, формируя в НИ направление, расширяющее границы, очерченные в документах [1].

В последующих разделах анализ текущего состояния НИ проведен на примере ряда наиболее популярных ресурсов, созданных как для медико-биологической, так и других областей применения. Основная цель анализа – выявить, насколько объем и структура данных, равно как и функциональные возможности ресурса, соответствуют особенностям НМ: 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 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 НИ – 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, обеспечив 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 НИ была связана с созданием так называемой *мета-онтологии*, то есть онтологии верхнего уровня, способной охватить объекты и концепции, возникшие в различных дисциплинах, включенных в использование и/или изучение НМ. 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 – 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 данной работы - определить, насколько сегодняшнее состояние НИ отвечает намеченной линии на интеграцию ресурсов путем онтологического моделирования предметной области, выявить проблемы, препятствующие ее реализации и возможные пути их преодоления.

БД ПО СВОЙСТВАМ НАНОМАТЕРИАЛОВ

В НИ, как и других видах e-science, БД, по-прежнему, являются основным элементом инфраструктуры, несмотря на появление новых технологий управления научными данными, использующих онтологии и средства Semantic Web [6]. На сегодня, бóльшая 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 НМ можно найти в ряде обзоров [7, 8]. Характерно, что за редким исключением все они охватывают сферу *EHS*, т. е. медицины, производственной гигиены, экологии и т.п. К примеру, 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 (<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 НМ: *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 НМ, с упором на оценки биологического и экологического воздействия предоставляет БД под названием **Nanomaterial Registry** (www.nanomaterialregistry.org). БД 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.

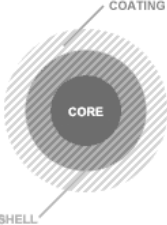
AS RECEIVED	
Time Point: 0	
Nanomaterial State: aerosol	
Laboratory: CDC-NIOSH-Spokane Research Laboratory	
Synthesis Method: incidental nanomaterial, furnace emissions - melting and refinement of silver	
<div style="display: flex; justify-content: space-around; border: 1px solid black; padding: 2px;"> PHYSICO-CHEMICAL CHARACTERISTICS BIOLOGICAL INTERACTIONS ENVIRONMENTAL INTERACTIONS </div>	
Particle Size	Mean Primary Particle Size: 5 nm <i>Transmission Electron Microscopy</i>
Aggregation/Agglomeration State	State: Agglomerated <i>Transmission Electron Microscopy</i> Size: 100 nm <i>Transmission Electron Microscopy</i>
Shape	Nanoscale Dimensionality: 3D <i>Transmission Electron Microscopy</i>
Composition	<div style="display: flex; align-items: center;"> <div style="margin-right: 10px;"> <p>● CORE (hover over circle for proportion of material)</p>  </div> <div> <p>Material Type: Metal</p> <p>Molecular Identity: Ag</p> </div> </div>
Purity	<p>Impurity Of: <i>Nanomaterial</i>: Pb</p> <p>Impurity Of: <i>Nanomaterial</i>: Zn</p> <p>Impurity Of: <i>Nanomaterial</i>: Ti</p> <p>Impurity Of: <i>Nanomaterial</i>: Se</p>

Рис. 1. Типовая запись БД Nanomaterial Registry.

Стандарт **MIAN** охватывает физико-химические характеристики НМ, наиболее актуальные для оценки его 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, информацию по методам измерений, протоколам, настройке приборов 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 уровней качества: «золотой», «серебряный», «бронзовый» и «заслуженный».

Таблица 1

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

Composition	Surface area	Aggregation/agglomeration state
Particle size	Surface chemistry	Solubility
Size distribution	Surface charge	Stability
Shape	Purity	Surface reactivity

Зависимость свойств НМ 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 сравнить 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 НМ, 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 наношкале. 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 – макроскопические объекты. В этом случае за

основу принято число макроскопических масштабов [2, 11, 12], которое может принимать 4 значения от 0 до 3. В то же время при классификации **ISO** число возможных наномасштабов достигает 3, при том, что макроскопические объекты сведены в отдельную категорию – наноструктурированных материалов.

Несмотря на охват НМ 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 типологии НМ. Без учета этой особенности невозможно представить с достаточной полнотой специфику отдельных типов НМ, 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 НИ [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 НМ: 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 БД, 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 – форму и размер.

Значительно расширить функциональность БД, способных адекватно учесть специфику НМ, удалось в недавней разработке авторов [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 НМ (его структуру, физические процессы и свойства, методы синтеза), так и сферу 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 признаков, определяющих источник (библиографическая информация), вид НМ и энергетическую технологию/оборудование, в сочетании с массивами текстовых, табличных и графических данных.

Таблица 2

Поля цифровой библиотеки «Наноматериалы для энергетики»

Number	Title	Searching field	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	Syntheses	Yes	Text/Textarea
12	Properties	Yes	Text/Textarea
13	Application	Yes	Text/Textarea
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		Textarea
23	More information		HTML form
24	Abstract		HTML form
25	Comments		Hidden

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

Поля разбиты на 3 блока: библиография (1-3, 14-20), НМ (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 свойствам 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 предложенной систематизации является рубрикатор НМ (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 НМ** позиции 3.8-3.12 выделяют НМ, ориентируясь на характерное свойство, а 3.3, 3.4 или 3.13 – на возможное использование.

Таблица 3

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 (статьи, отчеты и т.п.), размещенные на сервере (поле 21) или в сети (поле 22). Использование контекстной информации (данные из справочников, обзоров и т.п.) обеспечивают 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 НМ (название, размеры, физическое состояние), намного уступая по объему и достоверности информации научным БД, рассмотренным выше. 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 авторы разработали и онтологию нанокompозитов [16], позволившую 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 НИ непосредственное отношение имеют стандарты, касающихся терминологии и номенклатуры объектов и свойств. Их задача – обеспечить глобальный консенсус в отношении кодирования и классификации НМ, номенклатуры физико-химических и эксплуатационных характеристик, научной и производственной

лексики и т.п., без чего невозможны интеграция БД и других видов информационных ресурсов.

Ключевой элемент в терминологической унификации – так называемый, контролируемый словарь (*controlled vocabulary*), то есть 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 «БД по свойствам НМ» 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 (см. выше), 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 рубрикаторм 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 рубрикаторм того же Федерального портала не допускает кодирования рубрик.

Наряду с кодированием терминов, возможности словаря определяются также ориентацией и широтой, глубиной иерархической схемы, наличием дефиниций и гиперссылок и проч. Большинство словарей или таксономий, рассмотренных в публикациях по НИ [8, 18], посвящены медицине или токсикологии. К примеру **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. С другой стороны, 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

² Standards developing organizations

деятельность проведена **ISO** (*International Organization for Standardization*) и **ASTM International** (*American Society for Testing and Materials*). Разработанные стандарты, помимо функций словаря, определяют, что считать НМ, какие свойства должны отличать НМ, 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 НМ определены в 1, 2 и 4 томах словаря. Термином верхнего уровня признан *наноматериал*, множество которых разбито на два класса: нано-объекты и наноструктурированные материалы. К первому отнесены объекты, имеющие наномасштаб для 1, 2 или 3 измерений, ко второму – макроскопические вещества с внутренней или поверхностной наноструктурой. Сильно упрощенная классификация наноматериалов сводится к трем положениям [19]:

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

Таблица 4

Отдельные тома словаря ISO [20]

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

Примечание: * - том в стадии разработки

Тот же словарь (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).

Таблица 5

Классифицирующие признаки наноматериалов [21]

Chemical identity	Properties	Origin
Ceramics	Physical	Natural
Metallic	Mechanical	Manufactured
semiconductors	Chemical	Engineered
Polymer [natural/synthetic]	biological	Incidental
Carbon based	Combined[cross-phenomena]	
Organics/Inorganics		

Тем не менее, даже при учете множества признаков, словарь не способен охватить все виды вновь синтезируемых наночастиц. 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 наноматериалов [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 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 удается сузить и даже зафиксировать номенклатуру нано-объектов и их свойств. Примером таких организаций, может служить *National Cancer Institute (US)*, 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 нано-объектам определяют не линейным размером, а удельной поверхностью, при наименьшей ее границе в 60 м²/кг.

Широкое применение находят тезаурусы и классификаторы НМ, также не имеющие статуса стандартов, но охватывающие множество областей применения. Образцом такого классификатора служит упомянутая выше таксономия портала **InterNano**, 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, то есть без соблюдения каких-либо принципов рубрикации. К примеру, в рубрику *наноструктурированные материалы* не занесены *наноконпозиты*, 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 НМ самой разной природы, например наноконпозиты 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 проволоочки отнесены не к НМ, а к наноструктурам, объединяющим приборы и устройства.

Ранее мы показали [26], что практически всем рубрикатомам, 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 НМ, 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) вплоть до трехмерных, к которым отнесены объемные материалы. В отличие от схематизации ISO, здесь размерность – это число измерений, 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 *Производство нанотехнологий*. При всех достоинствах рубрикатора его использованию для интеграции ресурсов препятствует отсутствие англоязычной записи и кодировки в виде URL для отдельных терминов.

Таблица 6

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

1	Объекты, относящиеся к сфере нанотехнологий, их свойства	3.	Производство нанотехнологий		
1.1	Наноструктурированные материалы	3.1	Функциональные НМ	3.2	Конструкционные НМ
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 нанокompозиты	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	Нанокompозитные материалы				
1.1.4.3	Нанопористые материалы				
1.1.4.4	Нанофлюиды (дисперсии)				

Несмотря на высокий уровень детализации рубрикатора, превосходящего аналогичные продукты, например таксономию **InterNano**, он также не способен охватить все множество наноформ. Принципиальная недостижимость законченной классификации НМ, обусловленная множественностью структурных и физико-химических факторов в

сочетании с непрерывным пополнением их типов, отличают задачи, решаемые НИ, 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 которую можно сузить и даже временно зафиксировать номенклатуру объектов и свойств.

Относительно недавно на решение указанной проблемы НИ 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 анализируем ее возможности в качестве кандидата на роль международного стандарта для НИ, сопоставляя со стандартами **ISO** и отраслевыми рекомендациями.

Прежде всего, в системе **UDS** своеобразно решен вопрос о формальной идентификации НМ, 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. К примеру, ссылка на коды **NCI Thesaurus** (<https://ncit.nci.nih.gov/ncitbrowser/>) обеспечивает включение в созданную таксономию 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

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

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 НМ (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 наношкале. 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 нанокомпозиты (скажем, полимер с включениями углеродных нанотрубок), примером второго – наноструктурированные стали или керамика.

В работе [28] авторы провели детальный анализ иерархии, верхний уровень которой показан на рис. 5. Уже начальная идентификация (категория *General Identifiers*) обеспечивает гибкость в систематике НМ, адаптируя ее к разным областям применения со своими классификаторами. Наиболее полную идентификацию дает информационная категория *Characterization*, 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 (*Production* и *Specification*) 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. 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 их совокупность, способных учесть тонкие детали структуры, обеспечивает многофакторный характер описания, требуемого при аттестации НМ [2, 3, 11]. Кроме того, добавляя или удаляя отдельные из дескрипторов, можно отразить эволюцию схемы описания, связанную с *априори* неизвестными объектами.

Для ансамбля нано-объектов идентификация включает: состав ансамбля, то есть типы и число каждого из элементов ансамбля с характеристикой каждого из элементов по той же схеме, что и для индивидуальных нано-объектов. Что же касается объемных НМ, крайнее многообразие их видов заставило авторов [27] отказаться от построения иерархии с детализацией свойств, ограничившись лишь выделением двух типов (рис. 2): (Bulk Materials with Individual Nano-objects или Bulk Materials with Nanoscale Features. 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 НИ – созданию многофакторной и эволюционирующей концептуальной схемы данных.



Рис. 2. Верхние уровни иерархии понятий в системе UDS [27].

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]. Применение онтологий особенно актуально применительно к НМ, учитывая междисциплинарный характер концепций и методов, неустойчивость терминологии и понятийного аппарата, перманентное открытие новых материалов, устройств и технологий. Особенно полезной для НИ оказывается предусмотренная в онтологической инженерии возможность заимствования классов и терминов из родственных онтологий, что, например, широко применяется на биопортале (<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 наноформ 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 (рис. 3) показывает элемент иерархии, связанной с объектом *carbon nanotube*, а именно, дочерние и родительские классы, а также сущности, связанные по типу *часть-целое*.

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

The screenshot displays the ChEBI database interface for the entity 'carbon nanotube' (CHEBI:50594). The page includes a navigation menu on the left, a breadcrumb trail (EBI > Databases > Small Molecules > ChEBI > Main), and a main content area with tabs for 'Main', 'ChEBI Ontology', and 'Automatic Xrefs'. The 'Main' tab is active, showing the entity's name, ID, definition, and a star rating. The 'ChEBI Ontology' tab is also visible, showing outgoing and incoming relationships.

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

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**, связывающий 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 НМ, чтобы обеспечить адекватное представление данных для произвольной наноструктуры⁴.

Значительно более полной и детализированной является онтология **NPO**, *nanoparticle ontology for cancer researches* [33], включающая 1900 классов,

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

распределенных по множеству уровней иерархии (вплоть до 16), и связанных 80 ассоциативными соотношениями, конкретизирующими физико-химические свойства и возможное использование наночастиц. В то же время, 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, физическом состоянии (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 поле, 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 др. 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].

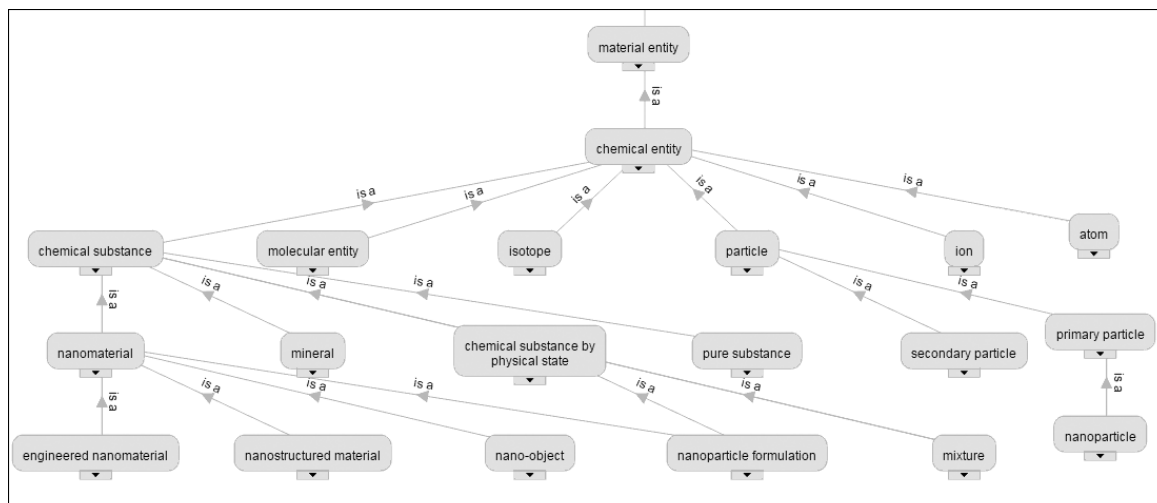


Рис. 4. Иерархия классов NPO, относящихся к наноматериалам и наночастицам.

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 др. характеристик. Каждому из терминов, включенных в онтологию, присвоен уникальный 3 (или 4)-значный 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, к примеру: **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 НМ, 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, появились в последнее время 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 НМ (**eNanoMapper**, www.enanomapper.net). 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, 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 выше онтологий, начиная с **ChEBI**, по существу не способна к адекватному представлению объемных/макроскопических НМ из-за крайней сложности и многообразия их типологии и структуры. Хотя в онтологиях **NPO** и **eNanoMapper** на верхнем уровне выделены две категории *nanomaterial* и *nanoparticle*, их содержание не соответствует принятому, согласно **ISO** [20], разделению мира НМ на нанообъекты (аналоги наночастиц) и наноструктурированные, иначе говоря, объемные НМ. 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 *nanostructured 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 нанообъектам, ничуть не напоминая те, что подразумевались в стандартах, скажем нанопористая или наноструктурированная керамика. Более того, как показано в разделе 3, с задачей адекватного представления объемных НМ 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 нанообъекты. Заметим, что согласно стандарту **UDS** [27], нанокompозит относится к объемным материалам с идентифицируемыми, то есть выделяемыми в явном виде, наноструктур. Основу онтологии составляет таксономия классов и экземпляров, включающая такие классы как *MaterialType*, *NanoObject*, *Nanocomposite*, а также классы, включающие химические характеристики: *ChemicalIdentity*, *ChainComposition*, *Structure* и т.д. Класс *MaterialType* включает 5 дочерних классов (*CarbonBased*, *Ceramic*, *Metallic*, *Polymer*, *SiliconBased*), что в сочетании с довольно простой таксономией нанообъектов (*NanoFiber*, *NanoFilm*, *NanoLayer*, *NanoPowder*, *NanoSurface*) позволяет отразить 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 нанообъекты. Они 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) нанообъектов, хотя исходно не предполагалось отнесение к этой категории объектов макроскопического масштаба.

Наконец, существует обширный класс объемных НМ, в которых невозможно выделить идентифицируемые нанообъекты: стали, керамика, полмеры, углеродные материалы и множество других. Приставка «нано» означает для них лишь наличие внутренней или поверхностной наномасштабной структуры. В работе [28] авторы

предложили достаточно естественный путь ее разработки, сочетая общую схему стандарта **UDS** с существующими онтологиями в области материаловедения.

ЗАКЛЮЧЕНИЕ

Приведенные результаты позволяют ответить на поставленный в работе вопрос: насколько сегодняшнее состояние НИ отвечает объективным потребностям нанотехнологий, включая все виды используемых и вновь синтезируемых НМ, возникшим на их основе приборам и технологиям, методам производства, анализа и моделирования наноразмерных объектов. Состав и возможности разработанных средств (БД, онтологий и др.) сопоставлены с теми требованиями, которые возникают перед создаваемой информационной инфраструктурой и отражают особенности возникшей системы знаний: ее междисциплинарный характер, перманентная эволюция структуры данных, многофакторное описание при особом внимании к генезису данных и т.д. [2, 3, 11].

Прежде всего, подавляющее большинство средств и технологий разработано 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, изменений 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, то есть попытки внедрить гибкую структуру данных, зависящую от класса материалов и сферы приложения, можно наблюдать одновременно с переходом от медицинской тематики к другим областям использования НМ, например к энергетике. Один из простейших приемов подстройки логической структуры к типологии НМ был применен авторами 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** (раздел 4). Иерархия классов, включенных в онтологию, охватывает не только сами материалы, но и физические свойства, области применения и т.п., что позволяет на уровне отдельных экземпляров обеспечить «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 НИ 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, стадий производства и последующей работы с НМ. При этом предусмотрены возможности детализации каждой из категорий стандарта, использования разнообразных классификаторов и словарей, адаптации к разным приложениям или областям знаний. Можно сказать, что именно стандарт **UDS** открывает путь к решению двух главных проблем НИ, о которых говорилось во Введении: требование гибкой структуры данных, адаптируемой к типам материалов и технологий, и динамическое расширение так называемой *мета-онтологии*, чтобы она могла поддерживать эволюцию предметной области. Возможная стратегия проектирования онтологии НМ, в том числе объемных/макроскопических НМ, была сформулирована в работах авторов [28, 37].

Литература

1. Nanoinformatics 2020 Roadmap. April 2011.
Published by the National Nanomanufacturing Network.
Amherst, MA 01003.
<http://eprints.internano.org/607/>
2. A.V. Eletsii, A.O Erkimbaev, G.A Kobzev, M.S Trachtengerts, V.Y. Zitserman. Properties of nanostructures: data acquisition, categorization, and evaluation// Data Science Journal.- 2012. - V. 11. - P. 126-139.
3. J. Rumble, St. Freiman.
Describing Nanomaterials: Developing a multi-disciplinary Framework.
Nanoinformatics 2015 Workshop.
<http://nanoinformatics.org/2013/agenda>
4. V. Maojo. Towards a visual taxonomy of nanoparticles.
Enabling Successful Discovery and Applications.
Nanoinformatics 2015 Workshop.
<http://nanoinformatics.org/2015/agenda>

5. A. Agrawal, A.Choudhary. Perspective: Materials informatics and big data: Realization of the “fourth paradigm” of science in materials science//APL Mater. – 2016- V. 4. – Art# 053208.
6. F. Karagiannis., D. Keramida, Y. Ioannidis, et al. Technological and organizational aspects of global research data infrastructures towards 2020//Data Science Journal. – 2013- V. 12. – P. GRDI1-GRDI5.
7. D.G. Thomas et al. Informatics and Standards for Nanomedicine Technology// Wiley Interdiscip Rev Nanomed Nanobiotechnol. – 2011. - V. 3. – No 5. – P. 511-532.
8. S. Panneerselvam, S. Choi. Nanoinformatics: Emerging Databases and Available Tools// Int. J. Mol. Sci. – 2014. – V. 15. – P. 7158-7182.
9. K.C. Mills, D. Murry, K.A. Guzan, M.L. Ostraat.
Nanomaterial registry: database that captures the minimal information about nanomaterial physico-chemical characteristics// J. Nanopart. Res. – 2014. – V. 16. – P. 2219.
10. K.A. Guzan, K.C. Mills, V. Gupta et al. Integration of Data: the Nanomaterial Registry project and data curation// Computational Science & Discovery. – 2013. – V. 6. – Art# 014007.
11. А.В. Елецкий, А.О. Еркимбаев, В.Ю. Зицерман, Г.А. Кобзев, М.С. Трахтенгерц.
Теплофизические свойства наноразмерных объектов: систематизация и оценка достоверности данных//Теплофизика высоких температур. – 2012. – Т. 50. - №4. – С. 524-532.
12. V.V. Pokropivny, V.V. Skorokhod. New dimensionality classifications of nanostructures// Physica E. – 2008. – V. 40. - No 7. - P. 2521-2525
13. А.О. Еркимбаев, В.Ю. Зицерман, Г.А. Кобзев, М.С. Трахтенгерц.
Цифровая библиотека вместо традиционной базы данных для нанотехнологий: опыт использования системы ABCD//НТИ. Сер. 2. - 2014. - №6. – С. 12-22.
14. М.Р. Когаловский, Б.А. Новиков. Электронные библиотеки- новый класс информационных систем// Программирование. - 2000.-№3.- С. 3-8.
15. Е.А. Порысева, А.А. Краснов, Э.М. Кольцова. Создание и разработка информационной системы в области нанокomпозиционных материалов//Успехи в химии и химической технологии. 2012.- Т. XXVI. - №11(140). – С. 38-41.
16. И. Карпова, Е. Порысева, Г. Казаков, Э. Кольцова. Разработка онтологии в области нанокomпозиционных материалов// Информационные ресурсы России. – 2012. - №2. - С. 5-9.
17. V. Murashov and J. Howard (eds.), Nanotechnology Standards, Nanostructure Science and Technology. © Springer Science+Business Media, LLC 2011
18. D de la Iglesia, R.E. Cachau, M. Garcia-Remesal, V. Maojo.
Nanoinformatics knowledge infrastructures: bringing efficient information management to nanomedical research//Computational Science & Discovery. - 2013. – V. 6. – Art# 014011.
19. F. Klaessig, M. Marrapese, S. Abe. Current Perspectives in Nanotechnology Terminology and Nomenclature. В монографии “Nanotechnology Standards”. Murashov, V.; Howard, J. (Eds). 2011, XIV, 262 p.
20. International Organization of Standardization: Nanotechnologies- Vocabulary, ISO/TS 80004: 2010-2016; 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. Л.А. Чернозатонский, Е.Ф. Шека, А.А. Артюх. Графен-нанотрубные структуры: строение и энергетика образования// Письма в ЖЭТФ. – 2009. – Т. 89. - №7. -С. 412-417.
23. А.В. Елецкий. Эндоздральные 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.

- http://ec.europa.eu/health/archive/ph_risk/committees/04_scenihp/docs/scenihp_o_012.pdf
26. Н.М. Буйлова, А.В. Елецкий, В.Ю. Зицерман, Г.А. Кобзев. Систематизация источников и данных по нанотехнологиям//НТИ. Сер. 1. – 2013. - №11. - С. 31-42.
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 www.codata.org/nanomaterials
28. А.О. Еркимбаев, В.Ю. Зицерман, Г.А. Кобзев, М.С. Трахтенгерц. Универсальная система метаданных для характеристики наноматериалов//НТИ. Сер.1. – 2015. - №10. – С. 8-20.
29. С. Bizer Interlinking scientific data on a global scale//Data Science Journal. – 2013. – V. 12. – P. GRDI6-GRDI12.
30. А.О. Еркимбаев, В.Ю. Зицерман, Г.А. Кобзев, В.А. Серебряков, К.Б. Теймуразов. Технология научных публикаций в среде «открытых связанных данных». Научно-техническая информация. Сер. 1. - 2013. - №12. - С. 1-11.
31. K. Degtyarenko, P. de Matos, M. Ennis et al. ChEBI: a database and ontology for chemical entities of biological interest// Nucleic Acids Research. – 2008. -V. 36. - Database issue. - D344–D350.
32. С. Batchelor. Ontologies for nanotechnology. #237 - Technical Sessions. ACS National Meeting March 22-26, 2009. Salt Lake City, UT. <http://acscinf.org/docs/meetings/237nm/presentations/237nm76.pdf>
33. D.G. Thomas, R.V. Pappu, N. A. Baker et al. NanoParticle Ontology for cancer nanotechnology research// Journal of Biomedical Informatics. – 2011. – V. 44. – P. 59–74.
34. M. Muñoz-Mármol, J. Crespo, M.J. Fritts, V. Maojo. Towards the taxonomic categorization and recognition of nanoparticles shapes// Nanomedicine: Nanotechnology, Biology, and Medicine. - 2015.- V. 11. – P. 457-465.
35. Haiyan Sun , Zhen Xu Chao Gao. Multifunctional, Ultra-Flyweight, Synergistically Assembled Carbon Aerogels// Adv. Mater. – 2013. –V. 25. – P. 2554–2560./
36. С. О. Hendren, С. М. Powers, М. D. Hoover, S. L. Harper. The Nanomaterial Data Curation Initiative: A collaborative approach to assessing, evaluating, and advancing the state of the field//Beilstein J. Nanotechnol. – 2015. – V. 6. – P. 1752–1762.
37. А.О. Еркимбаев, В.Ю. Зицерман, Г.А. Кобзев, А.В. Косинов. Связывание онтологий с базами данных по свойствам веществ и материалов// НТИ. Сер. 2. – 2015. - №12. - С. 1-16.