ISSN 0147-6882, Scientific and Technical Information Processing, 2016, Vol. 43, No. 4, pp. 199–216. © Allerton Press, Inc., 2016.
Original Russian Text © A.O. Erkimbaev, V.Yu. Zitserman, G.A. Kobzev, M.S. Trakhtengerts, 2016, published in Nauchno-Tekhnicheskaya Informatsiya, Seriya 1: Organizatsiya i Metodika Informatsionnoi Raboty, 2016, No. 10, pp. 1–18.

Nanoinformatics: Problems, Methods, and Technologies

A. O. Erkimbaev*, V. Yu. Zitserman, G. A. Kobzev***, and M. S. Trakhtengerts******

Joint Institute for High Temperatures, Russian Academy of Sciences, Moscow, 125412 Russia

**e-mail: adilbek@ihed.ras.ru*

***e-mail: vz1941@mail.ru*

****e-mail: gkbz@mail.ru*

*****e-mail: trachtengerts@yahoo.com*

Received June 15, 2016

Abstract—This paper outlines the range of problems solved by nanoinformatics, which is a newly originated discipline that combines the methods and tools for the propagation of data on nanomaterials as well as the instruments and technologies based on them. The specific features determined by the interdisciplinary character and rapid evolution of this knowledge area are summarized for the data on the properties of nanosized objects. The most-popular resources (databases, classifiers, and ontologies) on the properties of nanomaterials are presented. Some topical disproportions, which have occurred in nanoinformatics due to the predominant attention to nanomedicine at the expense of the traditional application fields of nanotechnologies, such as electronics and energetics, are pointed out. The general nanomaterial terminology and classification standards, which form a basis for the design of new databases and ontologies, are considered in detail. The CODATA (Committee on Data for Science and Technology) international standard for the universal description of a nanomaterial is proposed for use as the most advanced and universal approach to the solution of problems in nanoinformatics.

Keywords: nanomaterials, nanoparticles, databases, ontology, data systematization

DOI: 10.3103/S014768821604002X

THE ESTABLISHMENT OF A NEW DISCIPLINE

The appearance of the term “nanoinformatics” (hereinafter, NI) for the new field that has been generated by the needs of nano- and information technologies (ITs) can be dated back to 2010, when the Nanoinformatics 2010 working group of experts was convened to formulate its objectives and directions of activity. The implemented projects were obliged to take the interdisciplinary character of this knowledge area and the permanent extension of its definitions that reflect the appearance of new materials, devices, and applications into account. The schedule of the seminars with a brief description of discussed problems can be found on the website http://nanoinformatics.org/2010/overview. At the current moment, the most comprehensive and detailed plan has been developed within the framework of the Nanoinformatics 2020 Roadmap [1], which was adopted at the seminar of 2015.

An appreciable place in this document is held by the refinement of the nanoinformatics concept itself. In the opinion of the authors [1], in a similar manner to bio- and ecoinformatics, nanoinformatics is destined to synthesize methods for the collection, processing, and propagation of data with consideration for the specific features of nanotechnologies, namely, their interdisciplinary character, the multifactor description of materials and devices, variations in the nomenclature of properties upon the transition to new objects, etc. [2, 3]. Therefore, NI cannot merely be considered as an application of informatics to nanotechnologies. At the seminar of 2015, one report [4] pointed out that a specialist in nanoinformatics is involuntarily a mediator between two areas and has to obtain deep insight in the scientific essence of problems and look beyond external attributes such as databases (DB) and ontologies.

The analysis performed by the authors [1] for the completed and continued projects and functioning facilities (DB, portals, digital libraries, etc.) provided the ability to distinguish several key directions, such as data collection and systematization (*data curation*), facilities for the analysis and design of nanomaterials (hereinafter, NMs), data integration, and information propagation. Although the developers of the roadmap emphasized the interdisciplinary character of nanoinformatics, the overwhelming majority of projects they considered are oriented towards biomedical subjects, including the actions of nanostructures at all levels (molecular, cellular, and an entire organism) and the associated problems of toxicology and hygiene, as well as the use of nanomaterials for diagnostics and therapy. This seems to be due to a high level of informatization in ecology and healthcare[1]; thus the incorporation of nanomaterials only slightly broadened the earlier created infrastructure. As an example, on the portal http://bioportal.bioontology.org/, which contains the widest collection of 517 biomedical ontologies, only 4 (**NanoParticle Ontology, eNanoMapper, InterNano, and ChEBI**) explicitly represent a data structure typical for NMs. Several others (e.g., Medical Subject Headings (**MESH**)) only cover a small number of terms related to so-called nanomedicine, i.e., the problems of the use of nanomaterials for diagnostics and therapy. The process of the natural integration of NM data into the long-established infrastructure of biomedical informatics has led to the rapid formation of nanoinformatics as a new discipline with its own databases, classifiers, ontologies, etc. on the one hand, and led to a certain topical "imbalance," which has produced an effect on the selection of typical nanostructures, the nomenclature of properties, and the selection of tools and technologies for operation with data, on the other hand. As a result, most projects implemented according to the roadmap [1] do not take the specific features of data inherent in the other application fields of nanomaterials into account, such as electronics, mechanical and power engineering. In particular, nanoinformatics in application to medicine almost never deals with a bulk (macroscopic) material, where nanosized specifics is exhibited only in its internal or surface structure.

This determined the objectives of this work: after analyzing the principal results of nanoinformatics as applied to the needs of the EHS community (medicine, toxicology, etc.) we compare them with the less-known information resources oriented towards such fields as electronics, mechanical and power engineering.

For medicine, an adequate infrastructure, namely, a special branch of informatics for material science, i.e., so-called material informatics [5] with characteristic approaches to the presentation of data appeared in these fields even before the active introduction of NMs. These approaches imply the extended description of a measurement method, a sample state, a medium state, production conditions, etc. The complexity of identification due to a variety of structural and technological factors radically distinguishes a material from conventional substances, whose properties are uniquely determined by their chemical composition and/or structural formula. With some changes, these features of material-science data can also be assigned to nanomaterials; this facilitates the development of information resources, which form a direction that extends the boundaries outlined in [1].

We analyzed the current state of nanoinformatics using a number of the most-popular resources created for both biomedical and other application fields as an example. The main objective of our analysis is to reveal how completely the volume and structure of data, as well as the functionality of a resource, correspond to the specific features of nanomaterials, such as multifactor description with special attention to the genesis of data, the dependence of the nomenclature of properties on the class of a nanomaterial, and the permanent structural upgrading of data after the appearance of new materials and facilities. The requirements for the completeness and quality of data due to the specifics of nanomaterials were studied in detail in [2, 3] and summarized in the recommendations of international organizations, such as CODATA (Committee on Data for Science and Technology, www.codata.org) and VAMAS (Versailles Project on Advanced Materials and Standards, www.vamas.org).

The second aspect in the development of nanoinformatics is the objective need for nanotechnologies for the transition from the use of offline information resources with different formats, data models, and semantics to a single infrastructure. The organizing basis of such an infrastructure is a web ontology in the form of a semantically precise and computer-processable definition of entities and their relationships. An ontology enables the strict formalization of a knowledge area, thus providing the requirements for description schemes for individual resources (DB, etc.) and the semantic uniformity for dissimilar sources. In this case, its function is much broader then for a conventional taxonomy (classification), as it maintains logical relationships determined by the specifics of a knowledge area between concepts.

The integration of information resources begins to take shape in all the disciplines, where the operation with experimental and/or modeling data, including their storage, systematization, and propagation, forms a basis for the activity called e-science. In the program in [1], this area in the development of nanoinformatics was related with the creation of a so-called meta-ontology, i.e., an upper-level ontology, which is able to cover objects and concepts in different disciplines involved in the use and/or study of nanomaterials. In elaborating this concept, the developers of the roadmap outlined two key problems. The first problem is confined to the integration of previously used taxonomies and ontologies developed by different teams for different objects and applications (e.g., *NanoParticle Ontology for Cancer Nanotechnology Research* (**NPO**, www.nano-ontology.org), which was created predominantly for the informatics in the sphere of clinical oncology). The second problem is to carry out the requirements for the dynamic extension of a meta-ontology in such a fashion that it could maintain the complicated evolution of a knowledge area due to the discovery of new NMs, the manifestation of new effects and regularities, and the development of devices and technologies.

[1] In the English-language literature, this field is called *environment, health and safety (EHS) practice*.

Table 1. The physicochemical NM parameters incorporated into DBs [7–9]

No.	Parameter	No.	Parameter	No.	Parameter
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

Consequently, one of the objectives of this work is to determine how completely the current state of nanoinformatics corresponds to the outlined direction to the integration of resources by the ontological modeling of a knowledge area and to reveal the problems that hinder its implementation and possible ways to overcome them.

DATABASES ON THE PROPERTIES OF NANOMATERIALS

In nanoinformatics, as well as in the other e-science branches, databases still remain as a basic infrastructure element despite the appearance of new technologies using Semantic Web ontologies and facilities for the management of scientific data [6]. At present, the majority of the data is stored in relational DBs. The development of these DBs initiates the creation of various classifiers and taxonomies used as a basis for the design of conceptual DB schemes; this brings order to an initial array of data with the definition of basic entities and typical structures. In this case, in addition to the volume and quality of stored data, the applied value of DBs is determined by the possibility of open Internet access to it. A brief description of the most commonly used worldwide DBs on the properties and application of nanomaterials can be found in the reviews [7, 8]. It is characteristic that, with rare exceptions, all of these cover the EHS sphere, i.e., medicine, industrial hygiene, and ecology. As an example, only 1 resource, **InterNano** (www.internano.org), among 15 resources mentioned in the summary of DBs on nanotechnologies [8] is not confined within the framework of so-called EHS practice.

The Nanoparticle Information Library (NIL) (http://nanoparticlelibrary.net/) developed by the National Institute of Occupational Safety and Health (NIOSH) of the United States (www.cdc.gov/niosh) is one of the simplest resources according to the type and structure of data. The main objective of this resource is to provide the minimum volume of data required for professionals in the field of hygiene and toxicology. The type of a nanoparticle is characterized only by quadruple meta-data that characterize its structure, composition, synthesis method/origin, and size. According to the structural type, nanomaterials are divided into several classes: *agglomerated spheres*, *colloids*, *crystalline materials*, *films*, *nano horns*, *nano rods*, *nano tubes*, *nano wires*, *others*, *quantum dots*, and *spherical materials*. The limited and unchangeable structure of the data is compensated by the partial supplementation of a record with some contextual information (annotations, hyperlinks to related resources, or plots), which provides data on the properties and use of a nanomaterial and its toxicological characteristics (health and safety info).

The most complete volume of data on the properties of nanomaterials with emphasis on the estimates of their biological and environmental effects is provided by the **Nanomaterial Registry** DB (www.nanomaterialregistry.org), which is built and systematized in compliance with the **MIAN** (*minimal information about nanomaterials*) standard adopted for the characterization of properties. The set of physicochemical data covers 12 key parameters (Table 1) and some auxiliary information for the estimation of the quality and completeness of data.

A record with the data for a silver nanoparticle is shown in Fig. 1.

The **MIAN** standard covers the NM physicochemical properties that are most topical for estimating the interaction of nanomaterials with biological objects or the living environment. A distinctive feature of this NM registry as a factual DB is the preliminary processing of initial data (*data curation*) with the further assignment of a special rating that characterizes their completeness and reliability, i.e., the so-called *compliance level* (CL). For these purposes, the **MIAN** standard provides the incorporation of information on measurement methods, protocols, instrument settings, etc. in addition to measured parameters. The completeness of data submission in a source is characterized by a special questionnaire (*best-practice questions*) with questions about the presence of primary data and information on control methods, tool calibrations, etc. in a source. This is used as a basis to estimate the CL by a special algorithm and, depending on its value, one of the four quality levels: “gold,” “silver,” “bronze,” and “honored,” is assigned to the data.

The dependence of nanomaterial properties on certain experimental conditions has stimulated DB developers to trace the history of measurements, fixing the time and the character of environment. Such a characterization is attained via the incorporation an individual element called an *Instance of Characterization* (IOC) into a record. Their mapping enables the comparison of different measurements by establishing

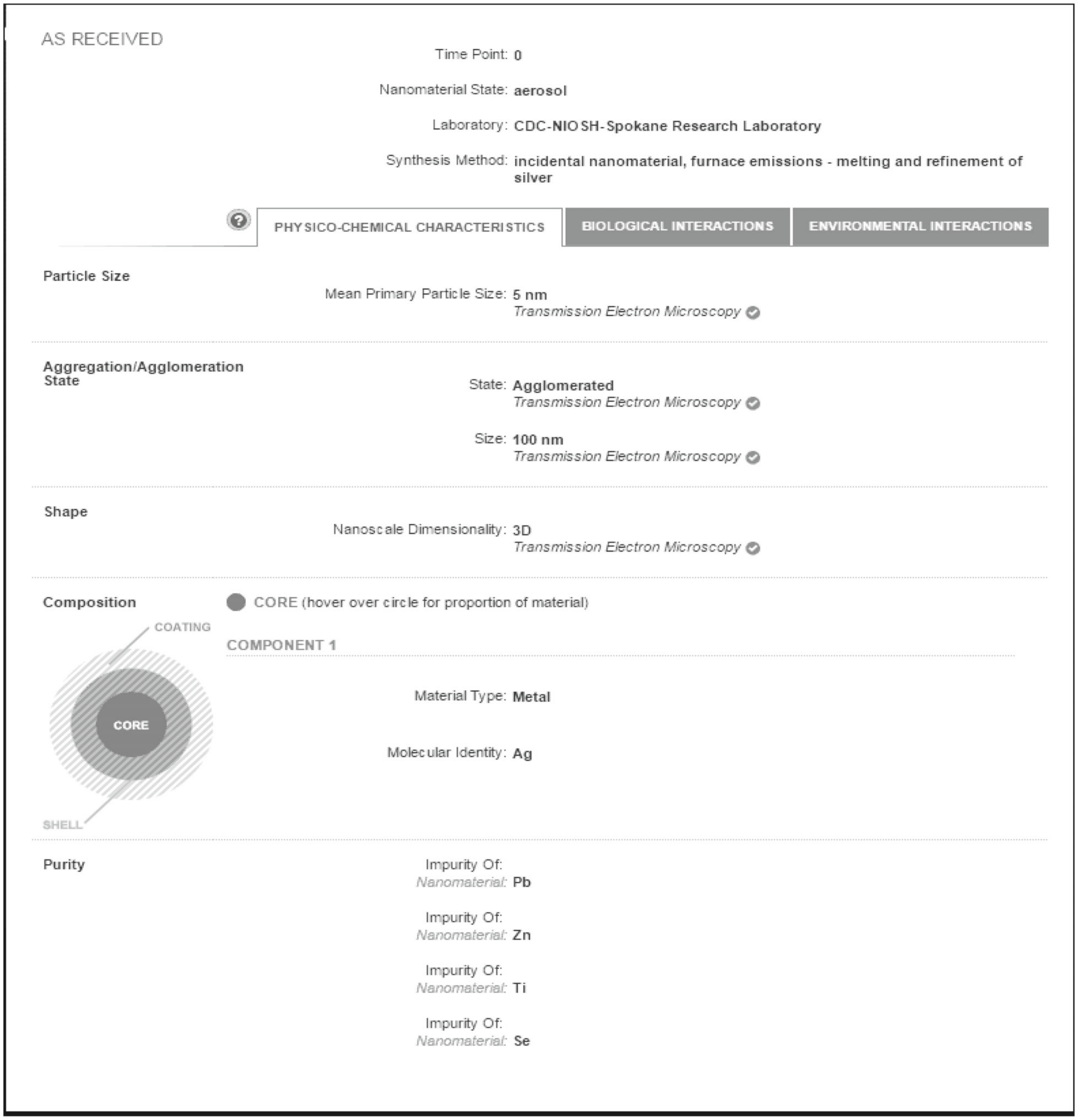

Fig. 1. A typical record in the Nanomaterial Registry database.

the identity of two materials for different treatment and characterization methods. When a number of IOCs exist, it is possible to trace the spread of data for a series of samples and step-by-step changes in the course of treatment and characterize the chronology of a sample. The accumulated information enables conclusions about the degree of similarity between samples that were received from different producers or subjected to different treatment.

A detailed description of some NM registry functions and data-preparation technology can be found in [9, 10] and the documentation on nanomaterialregistry.org. The registry was developed with sufficiently complete consideration for the NM-specific features mentioned in earlier works [2, 11], such as the requirement of multifactor description, the obligation for contextual information on synthesis and measurement procedures, and the quality evaluation of data depending on their uncertainty and completeness of presentation in a source.

An important element in the systematization of data used in the NM registry is their agreement with the adopted standard and the unification of terminology via a controlled vocabulary (www.nanomaterialregistry.com/resources/Glossary.aspx), which has been brought to agreement in large part with the International Organization for Standardization (IOS) recommendations (see the section Classifiers and Standards). In particular, according to ISO, the number of nanoscale dimensions was taken as a basis for the classification of nanoparticles into three types by their dimension. Graphene and nanofilms then enter the 1D category, nanotubes and nanorods are in the 2D category (two nanoscales in a cross section) and, finally, clusters, nanospheres, quantum dots, etc. belong to the 3D category. Let us note that such a classification contradicts the commonly accepted classification, in which nanotubes are categorized among 1D-structures, nanofilms are classified as 2D-structures, and macroscopic objects belong to 3D-structures. In this case, the number of macroscopic scales with four possible values from 0 to 3 [2, 11, 12] was taken as a basis. At the same time, the number of possible nanoscales in the ISO classification is three with the fact that macroscopic objects are placed into a particular category of nanostructured materials.

Despite the coverage of nanomaterials of different types (quantum dots, nanotubes, etc.), the nomenclatures adopted for characteristics in **NIL** and the registry are standardized; this contradicts one of the key features of NMs, namely, the variation of the nomenclature of properties and the structure of data depending on the NM typology. When this feature is ignored, it is impossible to provide a rather complete presentation for the specifics of individual types of nanomaterials, e.g., nanotubes, whose size is determined by the so-called *chirality index* [2, 11], and to combine the information on nanoparticles and bulk materials characterized by an internal nanostructure in the same DB. The common "striving" for the agreement of resources with external NM ontologies declared in the development plans of nanoinformatics [1], e.g., **NPO**, when the introduction of a corresponding reference uniquely identifies an object, i.e., **NPO_606** corresponds to a carbon nanotube, **NPO_395** characterizes a nanoparticle with a medical molecule embedded into its structure, etc., has also not been implemented.

In the extensive list of information resources cited in the review [8], only one resource, the portal **InterNano** (Resources for Nanomanufacturing, www.internano.org) is not confined to medical subjects. The documents on this portal cover all the aspects of scientific and production activities related to nanomaterials, such as production processes and facilities, nanoobjects and nanostructured materials, characterization methods, biomedical applications (*EHS practice*), socioeconomic effects, informatics and standards, and regulatory activities. Various types of documents (paper texts, expert reports, press releases, bibliographic references, etc.) are systematized via a three-level hierarchy (taxonomy), in which each term has its own URL. As an example, the URL www.internano.org/taxonomy/term/11987 corresponds to the *nanoparticles* concept and the URL with the same syntaxis and the terminal number of 12189 is assigned to the narrower *nanospheres* concept. URL references can be used in other DB or ontologies for the precise and easily understood identification of NMs or a certain nanotechnology, e.g., the technology *multilayer film process*, to which the URL with the terminal number 12101 corresponds. At the same time, the portal **InterNano** itself cannot be considered as a database, as a randomly selected collection of documents corresponds to each of the taxonomy meta-data instead of selected tables of properties. Moreover, despite a sufficient characterization of the taxonomy, it does not contain references to data on properties, even the simplest ones, such as shapes and sizes.

We have managed to appreciably extend the functionality of DBs that are able to adequately take the specifics of nanomaterials into account in our recent work [13]. To maintain the complicated and evolvable structure that is inherent in NMs data, it was proposed to avoid the traditional technology of relational DBs and use *digital libraries*, which are a qualitatively new electronic resource storage and integration system [14] that is able to provide multiaspect systematization that is adequate to the structure of this knowledge area, a dynamic character of the composition of resources and the conceptual scheme, and the integration of its own documents with documents and/or data from the Worldwide Web. This was used as a basis to create the electronic resource on the NMs for energetics subject for the purpose of systematizing the data that cover both a nanomaterial itself (its structure, physical processes and properties, as well as synthesis methods) and the sphere of its applications with characterization for the sectors of energetics, technologies, equipment, etc. The logical structure of data is reflected in the system of 25 meta-data/fields (Table 2). In summation, they enable the storage of a set of properties that characterize a source (bibliographic information), an NM type, and energy technology/equipment in combination with arrays of text, tabular, and graphic data.

The fields of Table 2 are divided into three blocks: bibliography (1–3, 14–20), nanomaterials (7–10), and energetic technology and equipment (4–6). The rules for filling them are given in [13]. Some fields have the formats *select simple/multiple*, which correspond to the selection of one or several concepts from an a priory composed list. A main peculiarity of these data, namely, their structural variations depending on the type of NM or the technology, is provided via their combination with fields of text format. Thus, the reliable identification of a nanomaterial is provided by the combination of fields 7 and 8: field 8 specifies the NM type via the reference to the position in the heading list, while field 7 introduces an arbitrary supplementation/refinement that extends the limits of the NM heading list. Similarly, fields 4 and 5 characterize an

Table 2. The fields of the Nanomaterials for Energetics digital library

No.	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	**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	Affilation	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

Classifier fields, HTML area fields, and external link fields are marked with bold type.

energy technology by means of heading lists, and field 6 extends their possibilities by incorporating the terms that characterize a technology or equipment.

The storage of information with a random volume and structure on the physical and service properties of a material and its synthesis methods and applications is also provided in the fields with the *HTML area* format (fields 11, 12, and 13). Here, it is possible to use arbitrary fonts, graphics, tables, and mathematical expressions for a multisided picture in the presentation of information about a certain material and technology. Hence, the combination of the detailed structure of data and the storage of unstructured information is provided as a necessary element in the sphere of nanotechnologies, where the rapid evolution of the knowledge area negates confinement to a fixed list of concepts.

A key element of the discussed systematization is a nanomaterial heading list (Table 3) with references to its positions in field 8. The accepted classification is oriented to several properties. In positions from 1.0 to 1.15, nanostructures are classified in compliance with their dimension (0D, 1D, 2D, and 3D), while the last group incorporates bulk/macroscopic materials with an internal of surface nanostructure that is typical for them. The classification can be considerably enriched using the physical features of a nanomaterial and the sphere of its application alongside with the topological property. As an example, positions 3.8–3.12 in the section "**Functional nanomaterials**" characterize materials with an orientation towards their typical properties, while positions 3.3, 3.4, or 3.13 characterize them regarding their possible applications.

A variety of data types and formats leads to the rich possibilities of the constructed digital library in the systematization of data, which differ from each other by their volume, logical structure, and presentation environment. The potentialities of the digital library are considerably strengthened by the fields designed for hyperlinks to external resources, such as the complete texts of documents (papers, reports, etc.) placed on a server (field 21) or on the Web (field 22). The use of contextual information (data from handbooks,

Table 3. The classifier of nanomaterials

1.0	**Nanostructures**
1.1	0D Nanoclusters
1.2	0D Nanocrystals
1.3	OD Fullerenes, endofullerenes, and fullerene-like
1.4	0D Quantum dots
1.5	1D Nanowires, nanorods, and nanofibers
1.6	1D Nanotubes
1.7	2D Nanostructured and composite films
1.8	2D Nanoporous surfaces
1.9	2D Nanomembranes
1.10	2D Graphene and graphene-like structures
1.11	3D Nanostructured materials
1.12	3D Nanocomposite materials
1.13	3D Nanoporous materials
1.14	3D Nanopowders
1.15	3D Nanodispersions (colloids)
2.0	**MISCL**
2.1	Ordered ensembles (multilayer and multiband structures and networks) of similar solid elements on substrates
2.7	Solid-state hybrid and hetero structures based on semiconductors, metals, and magnetics
3.0	**Functional nanomaterials (FNMs)**
3.1	Supported catalysts
3.2	Intercalation materials and solid electrolytes
3.3	Sensor nanocomposites
3.4	Hydrogen-absorbing nanomaterials
3.5	Nanostructured metals and alloys with particular mechanical properties
3.6	Nanostructured ceramic and composite materials and coatings
3.7	Nanostructured polymers, fibers, and composites based on them
3.8	Layer magnetic materials and superlattices
3.9	Piezoelectrics
3.10	Superconductors
3.11	Thermoelectrics
3.12	Luminescent materials
3.13	Bionanomaterials
4.0	**Constructional nanomaterials (CNMs)**
4.1	Technical iron and carbon steels
4.2	Alloyed steels
4.3	Non-ferrous metals and alloys based on them
4.4	Refractory materials and alloys based on them
4.5	Non-metallic materials

reviews, etc.) provides the integration with topically related resources.

The Worldwide Web contains many sources that provide commercial and production information, e.g., the database on the portal www.nanowerk.com: Nanomaterials, Nanotechnology Companies & Laboratories, Nanotechnology Products and Applications, etc. Among these, only the Nanomaterials DB contains minimum information about nanomaterials that are sold (name, size, and physical state), which his ranked slightly below the scientific DB in its volume and reliability. The information about domestic suppliers of nanotechnological products is given in several DBs registered on the Federal portal www.portalnano.ru, however, without open access. One of such DBs on nanocomposites is described in the literature. The scheme of this DB reflects the hierarchy "article → entity → entity properties," where the upper level incorporates the bibliographic description of a source, the intermediate level characterizes a nanocomposite by specifying the types of its matrix and filler, and the deepest level contains the information from a source on the properties of a composite in the text and digital forms. This is almost the only example of a bulk nanomaterial property DB, whose structure and content are unfortunately inaccessible to free search. It is interesting to note that the nanocomposite ontology [16] that enables the characterization of the logical structure of data for the design of similar resources has also been developed in parallel with the DB.

CLASSIFIERS AND STANDARDS

From the first appearance of nanotechnologies, the standardization of concepts and methods has taken an important place in their infrastructure, providing the consistency of requirements for the characterization of materials, the parameters of instruments, safety in the production and operation of products, etc. [17]. Nanoinformatics is directly related to the standards on the terminology and nomenclature of objects and properties. Their objective is to provide global consensus on the codification and classification of nanomaterials, the nomenclature of their physicochemical and service characteristics, the scientific and production vocabulary, etc., without which the integration of DBs and other types of information resources is impossible.

A key element in terminological unification is the so-called *controlled vocabulary*, i.e., the list of names and terms accepted in a professional community with a corresponding definition. The references to this vocabulary exclude uncertainty and synonymy and ensure an adequate and univocal response to a search query. The Databases on the Properties of Nanomaterials section of this paper contains some examples of application for such vocabularies adopted in the **Nanomaterial Registry** database and on the **InterNano** portal. These examples illustrate the possible distinctions between the characters of references to terms. Thus, each term in the **InterNano** taxonomy is encoded by its own URL; this in principle allows computer-aided search and analysis for a term, whereas the network address in the NM registry selects a vocabulary as a whole without referring a computer to individual terms. The encoding of a term with its own URL is also used in the heading list of the American Institute of Physics (**AIP thesaurus**, http://scitation.aip.org/content/topics), in which one section is devoted to nanotechnologies, and in two domestic resources: the vocabulary of nanotechnological terms (http://thesaurus.rusnano.com/) and the thesaurus of the Federal Portal of Nanomaterials and Nanotechnologies (www.portalnano.ru). On the other hand, the multilevel heading list of the same Federal Portal does not allow the encoding of headings.

Along with the encoding of terms, the possibilities of a vocabulary are determined by the orientation, breadth, and depth of its hierarchical scheme, the availability of definitions and references, etc. Most vocabularies or taxonomies considered in the papers on nanoinformatics [8, 18] are devoted to medicine or toxicology. As an example, **NCI Thesaurus** (*United State National Cancer Institute*, https://ncit.nci.nih.gov/ncitbrowser/) incorporates the taxonomy of nanomedical terms with definitions, references to external resources, unique codes, URL, etc. At the same time, the **InterNano** taxonomic vocabulary covers all the NM application and production aspects with the assignment of an individual URL to a term, however, without definitions and comments. The authors of the review [16] classify the vocabularies according to another criterion, namely, the developer status, setting national or international standardization organs against multiple scientific or administrative structures, which have sufficient authority, but no reliable SDO[2] status. The lists of organizations that are developers of standards for nanotechnologies can be found in [18, 19] and the main activity on terminology has been performed by *International Organization for Standardization* (**ISO**) and *American Society for Testing and Materials* (**ASTM International**). In addition to the vocabulary functions, the developed standards define what can be considered as a nanomaterial, what properties must characterize it, and how their individual categories can be distinguished.

The vocabulary composed by the **ISO** for different sectors of nanotechnologies [20] incorporates nine current standards and four standards under development (see Table 4). The basic NM definitions and classification principles are given in vocabulary volumes nos. 1, 2, and 4. The upper level term is recognized as *nanomaterial*. The variety of nanomaterials is divided into two classes: nano-objects and nanostructured materials. The first class incorporates objects with a nanoscale for one, two, or three dimensions,

[2] Standard-developing organizations.

Table 4. Some volumes of the ISO vocabulary [20]

Vol.	Name	Year of approval
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	*

* Stands for a volume under development.

Table 5. Classification characteristics of nanomaterials [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		

while the second class contains macroscopic substances with an internal or surface nanostructure. A strongly simplified classification of nanomaterials is confined to the three conditions [19]:

(1) nanomaterials including nano-objects and nanostructured materials;

(2) nano-objects including nanoparticles (3D), nanofibers (2D), and nanoplates (1D); and

(3) nanofibers including nanorods (solid form) or nanotubes (hollow form).

The same **ISO** vocabulary (volume 4) distinguishes only five types of nanostructured materials: a nanostructured powder, a nanocomposite, a solid nanofoam, a nanoporous material, and a liquid nanodispersion. A more detailed classification is given in the **ISO** report [21], where the number of types with different dimensions and structures is appreciably increased and their classification by other criteria, including chemical identity, physical properties, and origin, is also provided (Table 5).

Nevertheless, the vocabulary cannot cover all the types of newly synthesized nanoforms, even taking the variety of characteristics into account. As an example, nano-objects that incorporate covalently bonded graphene nanotubes and sheets, to which a certain dimension cannot be assigned, have been proposed [22]. It is also problematic to assign so-called *endohedral* structures [23] obtained via the embedding of objects of different natures into the cavity of one of the forms, e.g., the C_{60} or C_{70} fullerene molecules into the cavity of a carbon nanotube, to any heading of the classifier. Some problems are also encountered when assigning a certain chemical nature to a nanomaterial, as it can prove to be different for its fragments. Such a possibility has been taken into account by the authors of the NM registry [9, 10] by accepting nonidentity between the compositions of the core, shell, and coating of a nanoparticle (see Fig. 1).

Another organization, **ASTM International**, which is as profoundly involved in the development of standards as **ISO**, has prepared the E2456-06 Terminology for Nanotechnology document [24]. Its recommendations are focused on the properties of nanoparticles instead of the vocabulary that details the classification of nanomaterials. In particular, it introduced the concept of two nanoparticle categories (transitive and non-transitive) as determined by the character of change in properties upon the transition from a bulk/macroscopic material to the nanoscale: a jump upon the transition to the nanoscale is observed for the first category, while the same transition in properties is smooth for the second category. An example of non-transitive properties is the specific surface area or the optical scattering. In contrast, the confinement that characterizes the properties of quantum dots or quantum wires can serve as an example of transitive properties and particles that have this property. Another terminological innovation introduced in document E 2456-06 is the *ultra-nanoparticle* concept, which is defined as a particle with a size close to or less than 30 nm.

Many documents the specify terminology were also proposed by organizations that have no SDO status but are rather authoritative in a certain knowledge area, whose orientation provides the ability to reduce and even fix the nomenclature of nano-objects and their properties. Examples of such organizations are the *National Cancer Institute* (United Sates), which maintains a set of biomedical vocabularies with coverage of nanotechnologies (Enterprise Vocabulary Services, http://evs.nci.nih.gov/), or the *Scientific Committee on Emerging and Newly Identified Health Risks* (**SCENIHR**) of the European Community. As a rule, their recommendations are not confined to terminology and give their own refinements for basic concepts. As an example, many organizations step over the typical limit of 100 nm for one dimension when selecting a criterion for the classification of a structure among nano-objects. The analysis performed in the review [19] shows that the orientation to the problems of medicine, pharmaceutics, or the food industry leads experts to increase this limit to 300–500 and even 1000 nm. In one of the **SCENIHR** documents [25], the affiliation to nano-objects is determined by a specific surface area with its lowest limit of 60 m^2/kg instead of the linear size.

Wide application has also been found by the NM thesauri and classifiers, which also do not have the status of a standard, but cover many application fields. An example of such a classifier is the above-mentioned taxonomy of the **InterNano** portal created within the National Nanotechnology Infrastructure Network of the United Sates. The taxonomy is generally designed for the systematization of documents and references presented on the portal. However, the classification of nanomaterials themselves is composed in a rather arbitrary fashion, i.e., without regard for any principles in constructing the system of headings. As an example, *nanocomposites* have not been included in the section *nanostructured materials* and, on the contrary, *fullerenes* were arbitrary assigned to the section *nanotubes*, *ferrofluids* were classified as nanoparticles, etc. Similar arbitrariness characterizes the **AIP thesaurus** (http://scitation.aip.org/content/topics) adopted by the American Institute of Physics, where nanomaterials of most diverse natures, i.e., nanocomposites and clusters, are placed at the same hierarchical level. On the other hand, graphene (in contrast to a graphene film) has not been included in the section *nanomaterials* at all, and such objects as quantum dots or quantum wires are classified among *nanostructures,* straddling instruments and devices instead of nanomaterials.

It was shown in our earlier work [26] that eclecticity with a shift of characteristics with different contents at the same hierarchical level is inherent to one degree or another in almost all the heading lists without the status of a standard. It is likely that the NM systematization adopted at the Federal Portal www.portalnano.ru (see Table 6) is the most substantiated one. The basic types of nanomaterials are divided here into four categories by their dimension taken as the number of macroscopic scales from zero-dimensional (all the types of nanoparticles and clusters) to three-dimensional, among which bulk materials are classified. In contrast to the **ISO** systematization, the dimension here is the number of measurements along which a macroscopic nanoscale takes place, instead of the number of nanoscale measurements; moreover, the term *nanostructured materials* concerns all the types of nanomaterials rather than bulk materials. Among the latter are NMs specified under heading 1.1.4 (three-dimensional nanostructures) and under headings 3.1 and 3.2 of the section *Products of Nanotechnologies*. With all the advantages of this heading list, its use for the integration of resources is hindered by the absence of English-language recording and encoding in the form of a URL for individual terms.

Despite the high level of characterization in this heading list that is superior to similar products, e.g., the **InterNano** taxonomy, it also can not cover the entire variety of nanoforms. The principal unattainability of an accomplished classification of nanomaterials due to a variety of structural and physicochemical factors in combination with the continuous supplementation of their types distinguishes the problems solved by nanoinformatics from similar problems in the systematization of objects in chemistry or material science. The simplest and really flexible approach to overcoming these problems is the combined application of universal standards with classifiers or vocabularies that were created by organizations without SDO

Table 6. Classification of nanomaterials on the Nanotechnologies and Nanomaterials Federal portal (upper hierarchical levels)

1.	Objects related to the sphere of nanotechnologies and their properties	3.	Products of nanotechnologies		
1.1	**Nanostructured materials**	3.1	**Functional NMs**	3.2	**Constructional NMs**
1.1.1	Zero-dimensional structures	3.1.1	Supported catalysts	3.2.1	Technical iron and carbon steels
1.1.2	Linear nanostructures	3.1.2	Intercalation materials and solid electrolytes	3.2.2	Alloyed steels
1.1.3	Two-dimensional nanostructures	3.1.3	Sensor nanocomposites	3.2.3	Copper and copper-based alloys
1.1.4	Three-dimensional nanostructures	3.1.4	Hydrogen-absorbing nanomaterials	3.2.7	Refractory materials and alloys based on them
1.1.4.1	Bulk nanostructured materials	3.1.5	Nanostructured metals and alloys with unique mechanical properties	3.2.9	Non-metallic materials
1.1.4.2	Nanocomposite materials				
1.1.4.3	Nanoporous materials				
1.1.4.4	Nanofluids (dispersions)				

status and may be followed to reduce and even temporarily fix the nomenclatures of objects and properties.

The relatively recent international **Uniform Description System (UDS)** project [27] as carried out by the joint efforts of the CODATA–VAMAS[3] working group, was aimed at solving this problem of nanoinformatics. The meta-data system created as a result is meant to provide sufficient universality in the presentation of data for different NM types, the evolution of a description scheme with the appearance of new objects and concepts, the possibility of matching industry-specific features in the application of NMs (electronics, chemical industry, medicine, etc.), and the development of DBs, ontologies, and other information resources. A detailed description of the possibilities of **UDS** as a basis for the design of information resources can be found in [28]. Here, we will analyze its possibilities as a candidate for the role of an international standard for nanoinformatics by comparing it with the **ISO** standards and industry-specific recommendations.

First of all, the problem of the formal identification of nanomaterials has been solved in the **UDS** system in quite an original fashion with the assignment of names and incorporation into a hierarchy. The *General Identifiers* information category allows two NM definition levels: arbitrary and consistent with current standards (*assigned by an authority*), thus providing required freedom to match to a knowledge area. As an example, a reference to **NCI Thesaurus** codes (https://ncit.nci.nih.gov/ncitbrowser/) provides the incorporation of most objects and concepts of nanomedicine into the created taxonomy. For other NM types, the same problem (encoding of a name and incorporation into the hierarchy) can be solved via the reference to the **ChEBI** chemical DB (*Chemical Entities of Biological Interest*, www.ebi.ac.uk/chebi), which presents nanomaterials in the same fashion as ordinary compounds. It is essential that the code in both classifiers is a part of a URL (e.g., www.ebi.ac.uk/chebi/searchId.do?chebiId=50594 for CNT), thus allowing a computer-aided search for a term and the data attributed to it. A similar reference to the **ISO** standard [20] with the specified volume and term numbers is difficult to relate with a network number, as network access is open only to the information part of this standard.

The **UDS** project has introduced another innovation in comparison with the **ISO** standard [20]. While **ISO** distinguishes only two categories of nanomaterials (nano-objects and nanostructured materials), NMs in **UDS** include three classes: nano-objects, their ensembles, and bulk materials. The *ensemble* concept provides the ability to cover objects of complicated composition such as "graphene–nanotube" structures [22] that did not fit into the **ISO** systematics. In this case, an ensemble of nano-objects, as well as a nano-object itself, has from one to three dimensions on a nanoscale. Another novelty concerns

[3] CODATA is the Committee on Data for Science and Technology (www.codata.org); VAMAS is the Versailles Project on Advanced Materials and Standards (www.vamas.org).

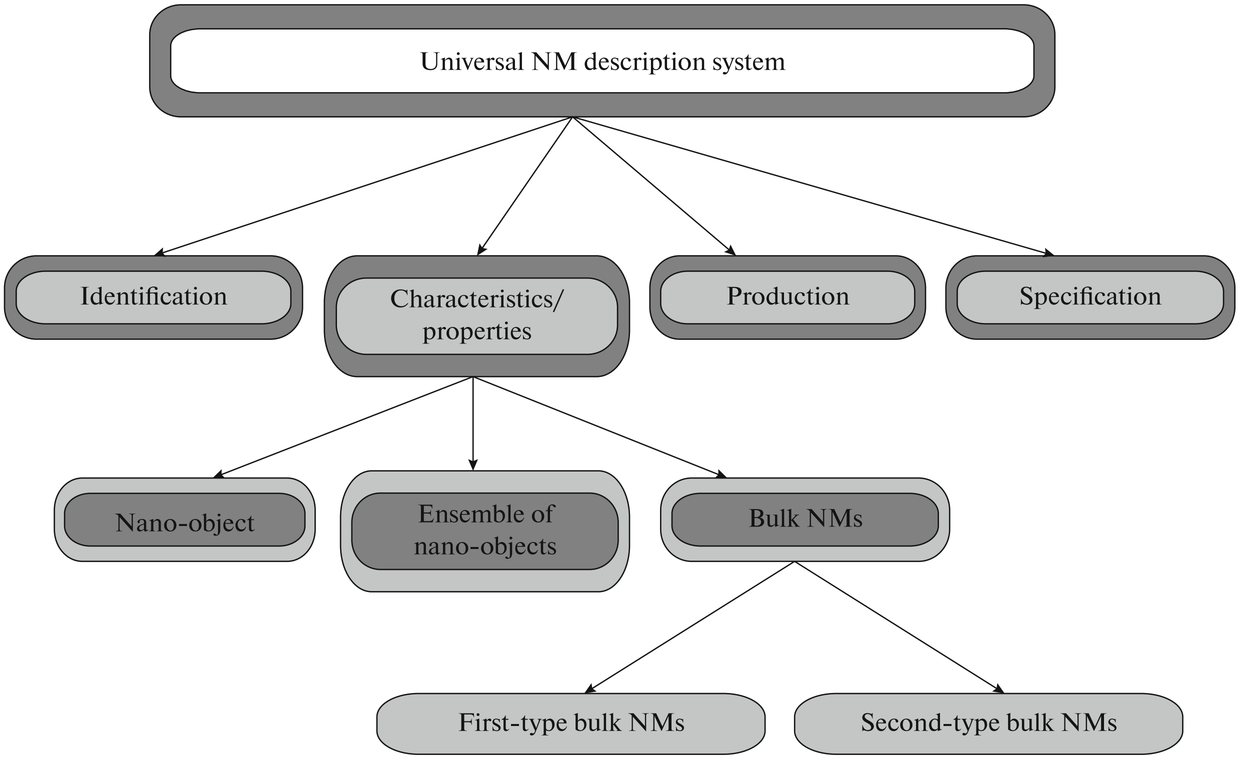

Fig. 2. The upper levels in the hierarchy of concepts in the **UDS** system [27].

bulk/macroscopic materials: they are divided into two subclasses: materials containing identifiable nano-objects and materials that only exhibit a dimensional effect due to their internal or surface nanostructure. An example of the former is nanocomposites (for instance, a polymer with inclusions of carbon nanotubes): nanostructured steels or ceramics can serve as an example for the latter.

A detailed analysis of the hierarchy, whose upper level is shown in Fig. 2, has been performed in our work [28]. Even the initial *Identification* category provides flexibility in the systematics of NMs, adapting it to different application fields with their own classifiers. The most complete identification is given by the *Characteristics/Properties* information category, which covers all the data on the properties of NMs, such as their shape and size, chemical composition, or internal and surface structure. The set of these data must provide the unambiguous distinction of a described object from a variety of others with a related structure. The two last categories (*Production* and *Certification*) concern the production and delivery of a material to the market. Comparing this nomenclature of properties with the nomenclatures used in a number of DBs, e.g., the nomenclature incorporated into the NM registry in compliance with the **MIAN** standard (Table 1), it is possible to see their relative similarity (shape, size, chemical composition, surface state, etc.) with the much better characterization provided in the **UDS**. In addition to the large number of categories themselves and distinctions in the hierarchy of each class (nano-objects, ensembles, bulk and materials), deep characterization of properties is provided by dividing each category into descriptors that are the lower hierarchical level filled by an expert. This set of properties makes it possible to take fine structural details into account and provides a multifactor character of the description that is required for the certification of a nanomaterial [2, 3, 11]. Moreover, by adding or removing some descriptors, it is possible to reflect the evolution of the description scheme in connection with a priory unknown objects.

The identification of an ensemble of nano-objects incorporates its composition, i.e., the types and number of each ensemble element, with the characterization of each element by the same scheme as for an individual nano-objects. For bulk nanomaterials, the great variety of their types led the developers [27] to avoid the construction of a hierarchy with the characterization of properties and confine themselves to classification into only two types: *Bulk Materials with Individual Nano-objects* (type 1) or *Bulk Materials with Nanoscale Features* (type 2). Some possible ways to develop the **UDS** for macroscopic materials using existing ontologies in the field of material science are considered in our work [28]. On the whole, the **UDS** system in combination with these vocabularies and taxonomies allows us to approach the solution of the main problem of nanoinformatics, namely, the cre-

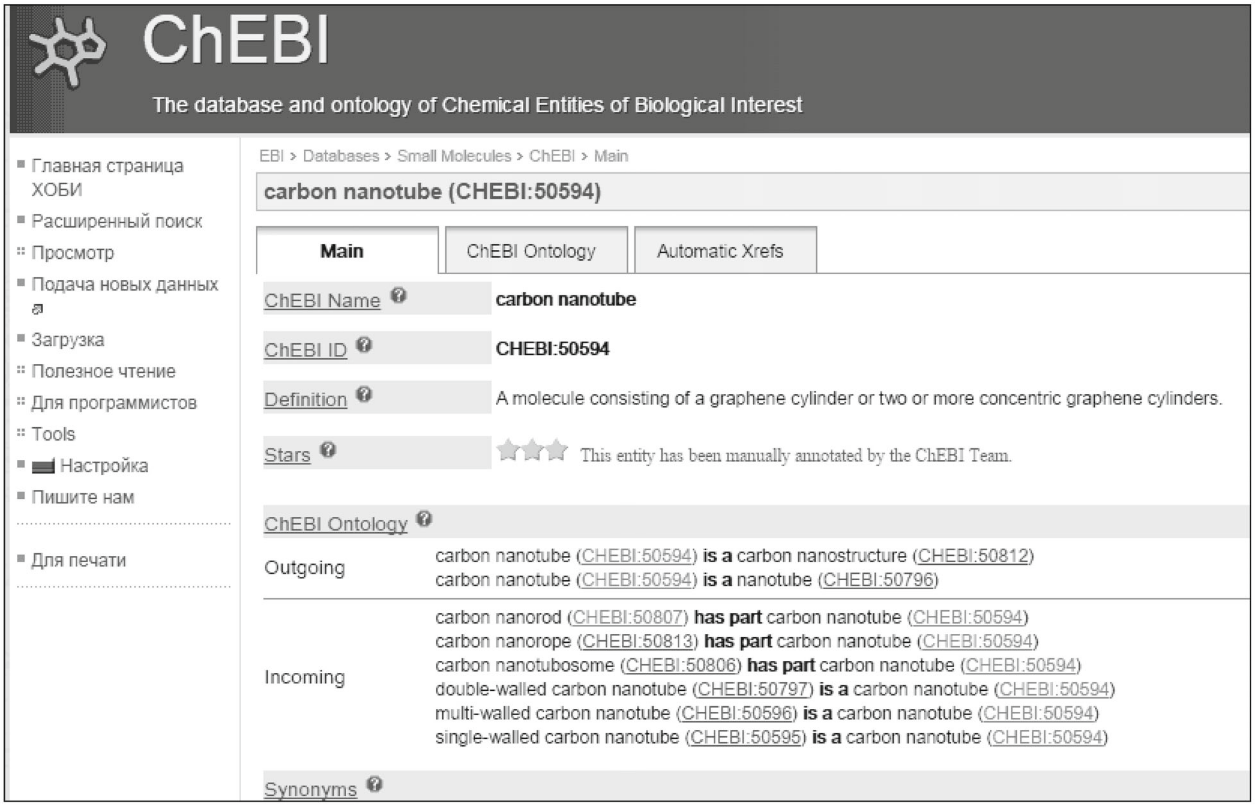

Fig. 3. A typical record characterizing the position of the class *carbon nanotub*e in the ChEBI database ontology.

ation of a multifactor and evolving conceptual data scheme.

ONTOLOGIES ON NANOMATERIALS

Ontologies can provide one of the most efficient mechanisms for the integration of scientific data via the combination of a subject-oriented vocabulary with a logical structure that is similar to a conceptual DB scheme. A vocabulary in the form of a multilevel taxonomy provides the standardization of concepts for the annotation of and search for documents. The logical structure of data is reproduced via the use of numerous attributes linked by associative relationships (e.g., **part_of, has_part, and has_quality**) and axioms. Thus, in addition to a terminological resource, an ontology can reproduce a formalized structure of knowledge. Among the possible applications of the ontological modeling of a knowledge area are the design of databases, the integration of heterogeneous sources, and the recording of documents in the RDF format with their further incorporation into the global space of linked data [29, 30]. The use of ontologies is especially topical for nanomaterials, taking the interdisciplinary character of concepts and methods, the instability of terminology and definitions, and the permanent discovery of new materials, devices, and technologies into account. The ability to borrow classes and terms from related ontologies in ontological engineering is especially useful for nanoinformatics and is widely applied, for example, on the bioportal http://bioportal.bioontology.org/, where a large collection of biomedical ontologies is integrated.

One of the simplest ontologies on nanomaterials is included into **ChEBI** (*Chemical Entities of Biological Interest*); this is a database and a structured classifier focused on small molecules with the exception of proteins and nucleic acids [31]. In essence, it is an ontology of chemical contents, in which any nanoform is interpreted as a molecular entity that does not differ from ordinary molecules or molecular aggregates. The selection of nanostructures and their systematization in **ChEBI** are not related with a certain discipline or application; this facilitates the borrowing of its terms. The example of a record in the ChEBI database in Fig. 3 illustrates an element of a hierarchy related with a *carbon nanotube*, namely, child and parent classes and entities related by the whole–part type.

The latter includes three types of nanostructures: *carbon nanorods*, *carbon nanoropes*, and *carbon nanotubosomes*. This fixes ensembles of nano-objects (e.g., a nanorope as an ensemble composed of nanotubes). Since **ChEBI** is first of all a chemical ontology, many nano-objects are considered as child classes with respect to a certain class of chemical forms, thus sup-

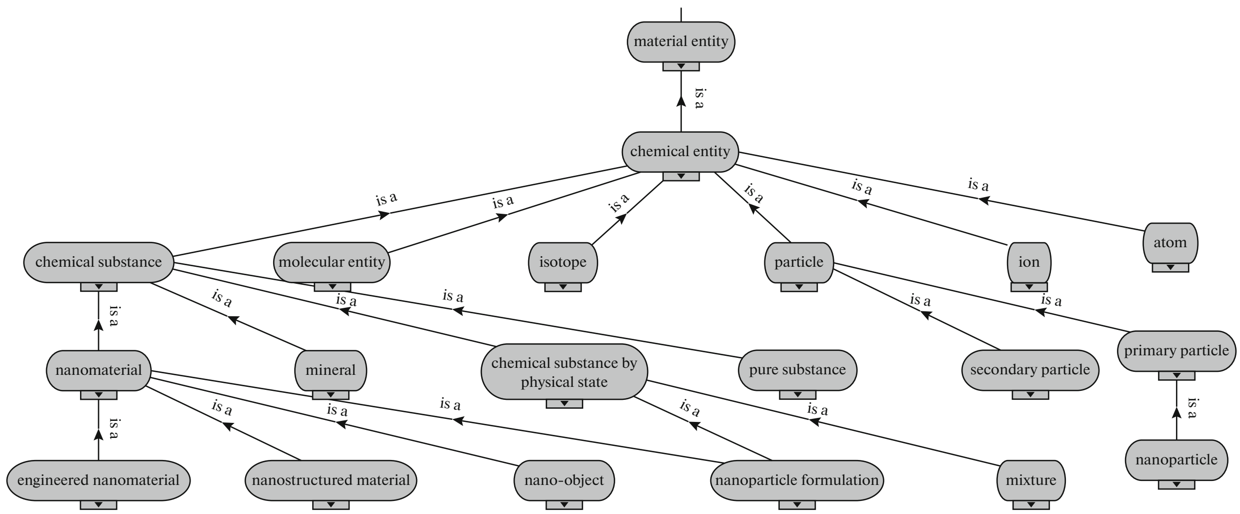

Fig. 4. The hierarchy of the classes of nanomaterials and nanoparticles in the NanoParticle Ontology.

plementing the definition of an object with the data on its chemical nature. As an example, the class *carbon nanotube*, which is a child with respect to *nanotube*, also inherits the area of organic compounds (*organic molecular entity*).

The ontology gives a strict definition to each object and labels it with a five-digit ID (**CHEBI:*****); reference to this in an arbitrary document (or DB) provides unique and computer-comprehensible identification. In addition, the use of the object property **has_role** that relates one of the taxonomic objects with one of the subontology classes, **role,** which characterizes the possible application and chemical and biological roles of an object, is also provided. However, all the role-based entities in **ChEBI** are focused on ordinary compounds, but have not found application for nanostructures. At the same time, the principle of construction for an ontology allows its natural extension to the field of nanotechnologies using a similar list of entities that specify the applications and roles of nano-objects. The **ChEBI** structure could be then considered as a possible prototype when designing a more extensive NM ontology that is able to characterize properties and applications. C. Batchelor [32] suggested the same idea in a report at an American Chemical Society session; he proposed to combine **ChEBI** with a small ontology that is able to cover the properties of nanomaterials to provide an adequate presentation of data for an arbitrary nanostructure[4].

NanoParticle Ontology for Cancer Nanotechnology Research (**NPO**) [33], which incorporates 1900 classes that are distributed over a set of hierarchical levels (up to 16) and linked by 80 associative relationships that characterize the physicochemical properties and possible uses of nanoparticles, is much more complete and detailed. However, being strictly oriented to a narrow segment of medicine, the ontology better characterizes the biomedical range of problems than the traditional aspects of nanotechnologies. This is shown in the predominance of corresponding terms (including those borrowed from the ontologies collected on the bioportal) and the explicit characterization of relationships that specify the biological role of a nanoparticle and chemical agents incorporated into it, the external stimuli that activate its function, and other aspects of diagnostics or therapy. The typical characteristics of the nanoparticle itself include the data on its structure, shape, size, and physical state (hydrogel, emulsion, etc.). In turn, a chemical compound is characterized by its physical position in a nanoparticle, molecular structure, and physical, chemical, and functional properties. In addition, the concepts incorporated into the ontology characterize the mechanism of action for a nanoparticle, the external stimuli that activate its function (magnetic field, ultrasound, and pH), and its response to an acting stimulus (e.g., the targeted delivery of a preparation by a nanoparticle in response to the actuation of a magnetic field). Each of these categories gives rise to an entire hierarchy of interrelated concepts. In particular, the ontology specifies the basic characteristics of a nanoparticle (size, shape, mass, and surface area) by describing the individual elements of a nanoparticle with the data on its composition and other characteristics.

All the types of nanostructures are grouped in two classes (*nanomaterial* and *nanoparticle*), both of which are at the hierarchical level induced by the class *chemical entity* (Fig. 4). The systematics that are dictated for nanostructures by the **NPO** hierarchy of classes rather strongly differ from the already accepted classifications: for example, the class *nanoparticle* must be con-

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

sidered as a daughter class with respect to the category *nano-object* [20, 21].

In addition to nanomaterials, the upper-level terms that shape the hierarchy include molecular structures, cell components, positions and boundaries in material entities (core, shell, etc.), properties, roles of molecular components, stimuli of nanoparticle function and responses of nanoparticles to stimuli, biological processes, and chemical interactions. At the same time, the full lists of terms and associative properties on the www.nano-ontology.org site allow the possibility of extension to different application spheres of nanostructures, e.g., for the characterization of its type, geometry, chemical composition, and other characteristics. Each term incorporated into the ontology is labeled with a unique three- or four-digit ID, e.g., **NPO_126** (*nanotube*) or **NPO_586** (*quantum dot*). In contrast to **ChEBI**, it is possible here to organize references to a variety of concepts that reflect the properties, technologies, or application of nanostructures by an ID, e.g., **NPO_1344** (*chemical composition*), **NPO_1445** (*atomic force microscopy*), or **NPO_1344** (*emission*). Thus, the terminological basis of an **NPO** can be rather widely used for the construction of alternative NM ontologies that are not related with biomedical problems.

The most essential limitation of the **NPO** ontology (as well as the above-considered **ChEBI**) is that it covers only the nanoforms that can be assimilated into molecular entities (nanoparticles, nanotubes, nanorods, etc.). Bulk nanomaterials, which are the most important in industrial applications and, according to the **ISO** definition [20, 21], include macroscopic objects (solids, powders, and dispersions) whose internal structure or surface is characterized by the presence of nanostructured units, remain outside its scope.

In addition to **NPO** and **ChEBI,** which are the best developed and cited ontologies, several other ontologies that reflect the assortment and characteristics of nano-objects, have appeared in recent years. One of these, which is also strictly oriented to EHS topics, was developed within the European Project for the Integration of Toxicological Data on Nanomaterials (**eNanoMapper,** www.enanomapper.net). The infrastructure created for these purposes follows the Semantic Web standards and technologies [29, 30] with the use of an ontology as a basic component. It was used as a basis to perform the systematization of different concepts related to biological experiments, the physicochemical and ecological certification of nanoparticles, and molecular and biological entities that are involved in the evaluation of risk. The terms were borrowed to a considerably degree from other ontologies that partially overlap the same subject. For nanostructures as such, the developers of **eNanoMapper** imported their names and physicochemical characteristics from **NPO** and **ChEBI**. The terms related to nanostructures are incorporated into the hierarchy due to the two upper-level classes, *nanomaterial* and *nanoparticle*. Their physicochemical characterization is built in a similar fashion. The incorporation of **eNanoMapper** into the collection of ontologies on the bioportal facilitates the process of its integration with both **NPO** and **ChEBI** and the extensive material from biomedical ontologies.

The ontological approach was also successfully applied to one of the key features of nanoparticles, i.e., their shape, using taxonomies as a basis for construction [4, 34]. Different typical 2D and 3D structures were presented as formal classes that compose a multilevel hierarchy. The ability to characterize the shape and morphological features of an object with completeness sufficient to estimate its functionality is a result. In this case, a special image-processing technique that enables the semi-automatic identification of images obtained by microscopy was used in additional to a geometric taxonomy that is convenient for systematization.

As can be seen, none of the ontologies considered here, beginning with **ChEBI,** can provide an adequate presentation of bulk/macroscopic materials due to the ultimate complexity and variety of their types and structures. In the **NPO** and **eNanoMapper** ontologies, two categories, *nanomaterial* and *nanoparticle,* are distinguished at the upper level, but their content does not correspond to the adopted division of NMs into nano-objects (analogues of nanoparticles) and nanostructured, in other words, bulk nanomaterials according to **ISO** [20]. Within the framework of the **NPO**, the class *nanomaterial* covers only one rather specific bulk material type, such as *nanoparticle formulation* (**NPO_868**), i.e., a substance in the form of a powder (or emulsion) that was prepared with the use of nanotechnologies and contains nanoparticles. For the class *nanostructured material*, in the **NPO** it incorporates only two structures: a *nanobud* (a nanotube bonded to fullerene outside) and a *nanofilm*, both of which, according to the **ISO** classification belong to nano-objects and do not at all resemble the structures implied in the standards, for example, nanoporous or nanostructured ceramics. Moreover, the **UDS** developers have not solved the problem of the adequate presentation of bulk nanomaterials despite the claimed universality of this standard.

The only example of such an ontology for nanocomposite materials has been created in the Mendeleev Russian University of Chemical Technology [16]. Their properties and functionality are specified by the selection of a matrix (metal, ceramic, etc.) and a filler, for which nano-objects are used. Let us note that, according to the **UDS** standard [27], a nanocomposite is classified among bulk materials with identifiable, i.e., explicitly defined nanostructures. The ontology is based on the taxonomy of classes and instances of such classes as *MaterialType*, *NanoObject*, *Nanocomposite*, and the classes *ChemicalIdentity*, *ChainComposition*, *Structure*, etc. that incorporate chemical characteristics. The class *MaterialType* has five daughter classes (*CarbonBased*, *Ceramic*, *Metallic*, *Polymer*, *SiliconBased*), which provide the ability to reflect the formation of different types of nanocomposites as governed by the selection of a matrix and a filler in combination with a rather simple taxonomy of nano-objects (*NanoFiber*, *NanoFilm*, *NanoLayer*, *NanoPowder*, and *NanoSurface*). Unfortunately, only a brief description of this ontology is available and the OWL-file itself is not presented on the public platform.

In addition to nanocomposites, another wide category of materials that contain identifiable nano-objects exists that consist of elementary "nanounits" without division of a material into a matrix and a filler. Polymerized fullerites formed due to the van der Waals interaction between C_{60} or C_{70} molecules are one example. Another example of a similar NM may be a superlight aerogel [35], which is a monolithic structure built of graphene ribbons and carbon nanotubes. Following the classification principles proposed in [27], the materials of this kind should be classified as an ensemble (or collection) of nano-objects, although the assignment of macroscopic objects to this category was not initially implied.

Finally, a wide class of bulk materials exists in which no identifiable nano-objects can be distinguished, such as steels, ceramics, polymers, carbon materials, and many others. For these materials the prefix "nano" means only the existence of an internal or surface nanoscale structure. In our work [28], a rather natural ontology development that combines the general **UDS** standard scheme with the existing ontologies in the field of material science was proposed.

CONCLUSIONS

The results of our study allow us to answer the question posed in this work: how completely does the present state of nanoinformatics correspond to the objective needs of nanotechnologies, including all the types of used and newly synthesized materials and the instruments and technologies based on them, as well as the methods for the production, analysis, and modeling of nanosized objects. The composition and possibilities of these facilities (DBs, ontologies, etc.) have been compared in this work with the requirements for the created information structure and reflect the specific features of the originated knowledge system: its interdisciplinary character, permanent evolution in the structure of data, multifactor description with special attention to the genesis of data, etc. [2, 3, 11].

The overwhelming majority of facilities and technologies were developed first of all for the biomedical sphere, where they were integrated with the long-developed system of medical nanoinformatics, e.g., the vocabularies and ontologies placed on the bioportal http://bioportal.bioontology.org/. In application to these problems, the specific features of NMs were taken into account first of all in data-curation procedures [36]. This technology has been most completely introduced in the **Nanomaterial Registry** database [9, 10], where some special methods are provided for the supplementation of data with the information on measurement methods, protocols, instrument settings, material transformations under treatment, etc. Along with these problems, attention was paid to the possibilities of integration for resources. As an example, the special format **ISA-TAB-Nano** [8, 18], which incorporates files in the form of spreadsheets and whose structure was specially adopted to the storage and exchange of biomedical analysis results, including the information on nanomaterials, has been developed. At the same time, the logical structure of most resources is kept within the style of a relational DB with a strictly fixed nomenclature of parameters, e.g., in the **MIAN** standard (Table 1).

Steps outside the framework of this limitation, i.e., some attempts to introduce a flexible data system that depends on the class of materials and the sphere of their application can be observed simultaneously with the transition from medical topics to other NM application fields, e.g., energetics. One of the simplest techniques for the matching of a logical structure to the typology of NMs has been applied in our work [13], where a set of heading lists was combined with the free use of keywords. Rich possibilities for deploying a complicated data structure are provided by the ontologies developed for nanomaterials, first of all, **ChEBI** and **NPO**. The hierarchy of classes incorporated into an ontology covers not only materials as such, but also their physical properties and application fields, etc., thus providing the "binding" of the required attributes to each object at the level of individual instances, e.g., the characterization of the size of nanotubes with parameters other than those for fullerene. Moreover, an ontology provides a rather easy means for maintaining the structural evolution of data, extending or upgrading the hierarchy of classes.

The recently proposed **UDS** standard [27] seems to have the greatest potential for the development of nanoinformatics with the coverage of any possible application sphere. It forms a basis for a logical structure that incorporates a multifactor description for the entire set of physical and service characteristics, as well as the stages of production and further operation with nanomaterials. This also enables the characterization of each category of the standard, the use of various classifiers and vocabularies, and the adaptation to different applications and knowledge areas. It is possible to say that the **UDS** standard opens the path to the solution of two main problems of nanoinformatics: the requirement for a flexible data structure that is adaptable to the types of materials and technologies and the dynamic extension of the so-called *meta-ontology,* such that it could maintain the evolution of a knowledge area. A possible strategy for the design of an ontology of nanomaterials, including bulk and macroscopic ones, has been formulated in our works [28, 37].

ACKNOWLEDGMENTS

This work was supported by the Russian Scientific Foundation (grant no. 14-50-00124).

REFERENCES

1. *Nanomanufacturing 2020 Roadmap,* Amherst, MA: National Nanomanufacturing Network, 2011. http://eprints.internano.org/607/.
2. Eletskii, A.V., Erkimbaev, A.O., Kobzev, G.A., Trachtengerts, M.S., and Zitserman, V.Y., Properties of nanostructures: Data acquisition, categorization, and evaluation, *Data Sci. J.,* 2012, vol. 11, pp. 126–139.
3. Rumble, J. and Freiman, St., Describing nanomaterials: Developing a multi-disciplinary framework, *Nanoinformatics 2015 Workshop.* http://nanoinformatics.org/2013/agenda.
4. Maojo, V., Towards a visual taxonomy of nanoparticles. Enabling successful discovery and applications, *Nanoinformatics 2015 Workshop*. http://nanoinformatics.org/2015/agenda.
5. Agrawal, A. and 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 Sci. J.,* 2013, vol. 12, pp. GRDI1–GRDI5.
7. Thomas, D.G., et al., Informatics and standards for nanomedicine technology, *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.,* 2011, vol. 3, no. 5, pp. 511–532.
8. Panneerselvam, S. and Choi, S., Nanoinformatics: Emerging databases and available tools, *Int. J. Mol. Sci.,* 2014, vol. 15, pp. 7158–7182.
9. Mills, K.C., Murry, D., Guzan, K.A., and Ostraat, 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, *Comput. Sci. Discovery,* 2013, vol. 6, art. 014007.
11. Eletskii, A.V., Erkimbaev, A.O., Zitserman, V.Yu., Kobzev, G.A., and Trakhtengerts, M.S., Thermal properties of nanoscale objects: Systematization and evaluation of data reliability, *Teplofiz. Vys. Temp.,* 2012, vol. 50, no. 4, pp. 524–532.
12. Pokropivny, V.V. and Skorokhod, V.V., New dimensionality classifications of nanostructures, *Phys. E,* 2008, vol. 40, no. 7, pp. 2521–2525.
13. Erkimbayev, A.O., Zitserman, V.Yu., Kobzev, G.A., and Trakhtenherz, M.S., A digital library instead of a traditional database for nanotechnologies: An attempt to use the ABCD system, *Autom. Doc. Math. Linguist.,* 2014, vol. 48, no. 4, pp. 212–223.
14. Kogalovskii, M.R. and Novikov, B.A., Electronic library as a new class of information systems, *Programmirovanie,* 2000, no. 3, pp. 3–8.
15. Poryseva, E.A., Krasnov, A.A., and Kol'tsova, E.M., Creation and development of an information system in the field of composite materials, *Usp. Khim. Khim. Tekhnol.,* 2012, vol. 26, no. 11 (140), pp. 38–41.
16. Karpova, I., Poryseva, E., Kazakov, G., and Kol'tsova, E., Development of the ontology in nanocomposite materials, *Inf. Resur. Ross.,* 2012, no. 2, pp. 5–9.
17. *Nanotechnology Standards, Nanostructure Science and Technology,* Murashov, V. and Howard, J., Eds., New York: Springer, 2011.
18. de la Iglesia, D., Cachau, R.E., Garcia-Remesal, M., and Maojo, V., Nanoinformatics knowledge infrastructures: Bringing efficient information management to nanomedical research, *Comput. Sci. Discovery,* 2013, vol. 6, art. 014011.
19. Klaessig, F., Marrapese, M., and Abe, S., Current perspectives in nanotechnology terminology and nomenclature, in *Nanotechnology Standards,* Murashov, V. and Howard, J., Eds., New York: Springer, 2011, pp. 21–52.
20. International Organization of Standardization: Nanotechnologies – Vocabulary, ISO/TS 80004: 2010–2016. http://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. Chernozatonskii, L.A., Sheka, E.F., and Artyukh, A.A., Graphene-nanotube structures: Structure and energy of formation, *Pis'ma Zh. Eksp. Teor. Fiz.,* 2009, vol. 89, no. 7, pp. 412–417.
23. Eletskii, A.V., Endohedral structures, *Usp. Fiz. Nauk,* 2000, vol. 170, no. 2, pp. 114–142.
24. *ASTM International: E 2456-06 Terminology for Nanotechnology,* West Conshohocken: ASTM International, 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 21st Scenihr Plenary on November 29, 2007.* http://ec.europa.eu/health/archive/ph_risk/committees/04_scenihr/docs/scenihr_o_012.pdf.
26. Builova, N.M., Yeletskii, A.V., Zitserman, V.Yu., and Kobzev, G.A., The systematization of sources and data on nanotechnologies, *Sci. Tech. Inf. Process.,* 2013, vol. 40, no. 4, pp. 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, February 1, 2015. http://www.codata.org/nanomaterials.
28. Erkimbayev, A.O., Zitserman, V.Yu., Kobzev, G.A., and Trakhtenherz, M.S., A universal metadata system for the characterization of nanomaterials, *Sci. Tech. Inf. Process.,* 2015, vol. 42, no. 4, pp. 211–222.

29. Bizer, C., Interlinking scientific data on a global scale, *Data Sci. J.,* 2013, vol. 12, pp. GRDI6–GRDI12.

30. Erkimbayev, A.O., Zitserman, V.Yu., Kobzev, G.A., Serebrjakov, V.A., and Teymurazov, K.V., Publishing scientific data as linked open data, *Sci. Tech. Inf. Process.,* 2015, vol. 42, no. 4, pp. 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 Res.,* 2008, vol. 36, pp. D344–D350.

32. Batchelor, C., Ontologies for nanotechnology. #237 - Technical Sessions, *ACS National Meeting March 22–26, 2009,* 2009, Salt Lake City, UT. 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, *J. Biomed. Inf.,* 2011, vol. 44, pp. 59–74.

34. Muñoz-Mármol, M., Crespo, J., Fritts, M.J., and Maojo, V., Towards the taxonomic categorization and recognition of nanoparticles shapes, *Nanomed.: Nanotechnol., Biol., Med.,* 2015, vol. 11, pp. 457–465.

35. Haiyan Sun, Zhen Xu, and Chao Gao, Multifunctional, ultra-flyweight, synergistically assembled carbon aerogels, *Adv. Mater.,* 2013, vol. 25, pp. 2554–2560.

36. Hendren, C.O., Powers, C.M., Hoover, M.D., and 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, pp. 1752–1762.

37. Erkimbaev, A.O., Zitserman, V.Yu., Kobzev, G.A., and Kosinov, A.V., Associating the ontologies with databases by properties of substances and materials, *Nauchn.-Tekhn. Inform., Ser. 2. Protsessy Sist.,* 2015, no. 12, pp. 1–16.

Translated by E. Glushachenkova