ISSN 0005-1055, Automatic Documentation and Mathematical Linguistics, 2012, Vol. 46, No. 4, pp. 170–176. © Allerton Press, Inc., 2012.
Original Russian Text © A.O. Erkimbaev, A.B. Zhizhchenko, V.Yu. Zitserman, G.A. Kobzev, E.E. Son, A.N. Sotnikov, 2012, published in Nauchno-Tekhnicheskaya Informatsiya, Seriya 2, 2012, No. 8, pp. 1–8.

Integration of Databases on Substance Properties: Approaches and Technologies

A. O. Erkimbaev[a], A. B. Zhizhchenko[b], V. Yu. Zitserman[a], G. A. Kobzev[a], E. E. Son[a], and A. N. Sotnikov[b]

[a] *Joint Institute for High Temperatures, Russian Academy of Sciences,
Izhorskaya ul. 13, stroenie 2, Moscow, 125412 Russia*
[b] *Joint Supercomputer Center, Russian Academy of Sciences*
e-mail: adilbek@ihed.ras.ru

Received May 26, 2012

Abstract—This paper discusses the *Data Center* project, which provides integration of the scientific electronic resources (mainly databases) that are developed and supported by the Russian Academy of Sciences. The integration technology has been verified within the framework of the Properties of Substances and Materials interdisciplinary theme, which is represented in many Institutes of engineering and natural-science profiles of the RAS. The possibilities of the XML language and ontological modeling are considered for the formalized description of the subject field of the properties of substances. Successful examples of work with databases on properties demonstrate that software engineering has achieved a high level and allows for the development of common exchange standards for heterogeneous resources.

Keywords: database, data integration, ontology, XML, substance properties, Semantic Web

DOI: 10.3103/S000510551204005X

1. INTRODUCTION. AIMS AND FORMS OF DATA INTEGRATION

The division of mathematical sciences of the Russian Academy of Sciences has started to create a distributed information system of science, education, and innovations, which is briefly called the *Data Center*. The basic elements of the design infrastructure will be represented by telecommunication networks, supercomputer systems, and information resources (IRs) [1]. The project implementation is intended to overcome the isolation and limited accessibility of IRs, such as databases (DBs), electronic issues, and information-computing means that are supported by different Institutes of the RAS. In the general form, IR integration implies the connection of resources through unified representation including possible information retrieval at a user's request. Integration assumes that the user does not have to independently select the resources and work with each resource separately. Access through a single interface means the support of the data from the set of heterogeneous sources in terms of a single data model. In addition to the unification of access to resources, integration makes it possible to solve many other problems including the discrepancies in data models (relational, object-oriented, etc.), structural and semantic heterogeneity, and so forth. Another aspect in favor of integration is the lifetime of software due to its ageing and periodic replacement with more effective software with the partial or total loss of the capability to study previous structures. Therefore, data storage for future generations requires their transformation into a certain standard form that is accessible for computer viewing when all the original means of data storage and management go out of use.

During the long history of the integration problem, which dates back to the 1970s [2], a number of solutions have been obtained whose success depended on the choice of the subject field, the properties of individual resources, and primarily, the desired level of integration. The experience from many applied projects has shown that with extremely diverse thematic areas and typologies of IRs the *bottom-up* approach appears to be the most practical. It narrows the initial concepts and structures for a limited subject field while retaining the potential of thematic area extension and inclusion of new resources. This consideration leads to the choice of the "*substance properties*" subject area when developing the basic concepts of the *Data Center* project. On the one hand, the creation of numerical databases on the properties of substances and materials, including the data on atomic—molecular constants and characteristics of nanostructures, is one of the key findings of the research that has been performed in the Institutes of natural-scientific and engineering profiles of the RAS. Therefore, the corresponding integration technologies should find application in numerous scientific groups involved in the solution of different problems of physics, chemistry, materials science, and even biology [3]. On the other hand, as for genesis and structure, these data fit best with typical DB concepts, thus making their integration relatively visible. Originally, any typical reference book with tabular data is a prototype of the relational DB where substance names act as entities and each property is considered as an attribute. As far as the properties are concerned, the specifics of scientific data imply a diversity of forms of representation (table, graphic, analytical, etc.), the complex nature of a logical structure governed by a physical model, and the variety of the data structure, which depends on the substance class and data source [4]. The complexity and variety of IR structures present natural barriers to integrating data from heterogeneous resources.

Here, the decisions that are made about a sufficient level of integration for the joint functioning of different sources and data exchange between them is significant. Very often the requirement for semantic integration, which has been highly discussed, i.e., the understanding of sense by program agents without the involvement of a person is excessive, especially with the existence of electronic issues or web portals that unite semi-structured and nonstructured data. In this case, the simplest solution is to consolidate resources at the interface level without their logical or semantic bonds. This is the most cost-effective and accessible integration method; it provides high scaling ability. We may mention the site of the US National Institute of Standards and Technologies (www.nist.gov), which provides access to numerous autonomic DBs on the physicochemical properties of substances that differ in structure and semantics, as well as the domestic IRIC DB (Information Resources of Inorganic Chemistry, http://iric.imet-db.ru/DB.asp), which is supported by the Institute of Metallurgy and Material Science of the RAS.

Another similar method is integration at the level of outer references when additional fields that store hyperlinks to other IRs that include DBs are introduced into the DB. An example of this data arrangement is shown in Fig. 1. When making a request to the TERMAL DB (IVTAN, RAS [5]) on fullerene properties one comes to the DB on carbon structures of the Ioffe Physical Technical Institute.

Both methods, while facilitating a user's work with dissimilar data, totally exclude both the generating of structured queries and the application of analytical means. For these purposes, a deeper integration level is necessary in order to allow the consistency of DB structure and semantics, which assumes the existence of a general information model and general vocabulary for determining the sense of basic concepts.

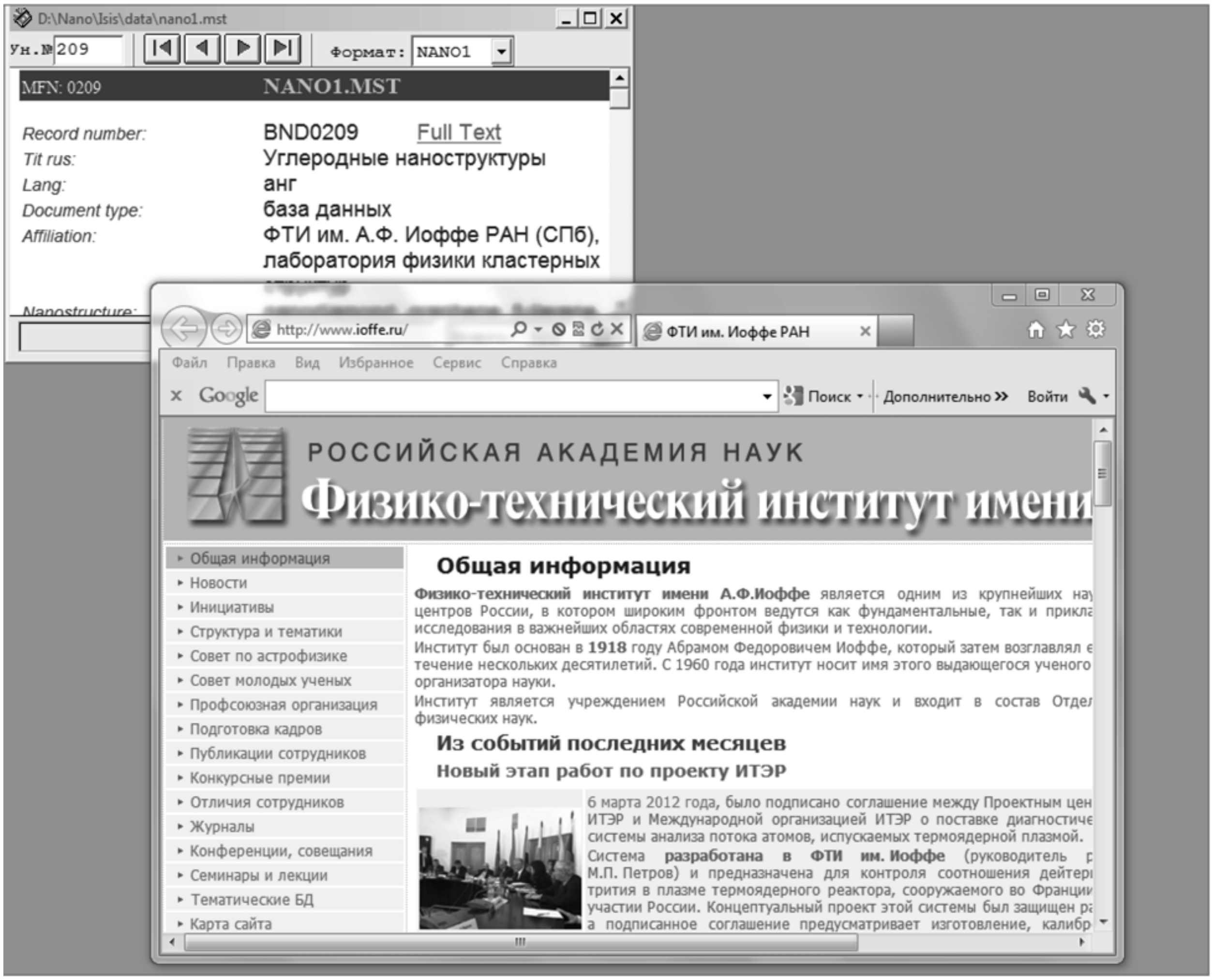
2. THE USE OF THE XML LANGUAGE AND ITS PROFESSIONAL VERSIONS

The model based on the standards the XML language, which is one the most popular formats when exchanging structured information between programs, between people, and between people and computers [6], has long been used as a data integrating model. In contrast to HTML, this language was created for description rather than reflection of data and adequately translates their structure and semantics. After dissimilar data have been transformed into a XML-formatted document it becomes accessible to various software tools. Such accessibility is based on the natural **interoperability** of XML documents, which is due to textual recording, modular nature, and total visibility when read or computer processed. Document visibility makes it **self-documented**, thus facilitating the understanding of the scheme.

Proper XML versions with their own dictionaries, support mechanisms in the form of custom browsers and programs to provide graphic representation, computing services, etc. have appeared in many disciplines [7]. Tens of such versions have recently been created whose lists can be found on the sites www.xml.com/pub/rg/Science or xml.coverpages.org/xmlApplications.html. At least two of them, ThermoML and MatML, turned out to be successful in the dissimination and exchange of properties data in thermodynamics and materials science. To a certain extent, this success is due to the fact that key knowledge in both fields has a simple structure, i.e., a set of properties in the form of constants or one-dimensional tables is assigned to an object with a characteristic name (or a set of names).

The ThermoML version, which was developed by the Thermodynamic Center of the US Institute of Standards and Technologies [8], is intended to standardize the forms of representation and exchange of physical data. The language covers the data for more than 120 properties at different forms of representations and different knowledge status (experimental, calculated, or reference). Pure substances with characteristic identification (by formula, name, or their number in the proper nomenclature) and mixtures, as determined by composition, as well as chemical and phase reactions, can be considered as objects. A detailed scheme has been developed to represent this data (http://www.trc.nist.gov/ThermoML.xsd).

The structure of a ThermoML document is a balanced combination of hierarchical and relational elements. In its explicit form, it involves all concepts, names of properties and their classes, and the parameters of state, limitations, and phases. Metadata that detail the measurement method, the sample state, and the form of uncertainty representation occupy a significant amount of room. The ThermoML version uses MathML to transfer formulas. Special means were created to capture properties data from scientific papers and generate corresponding XML documents.



Record number: BND0209 Full Text
Tit rus: Углеродные наноструктуры
Lang: анг
Document type: база данных
Affiliation: ФТИ им. А.Ф. Иоффе РАН (СПб), лаборатория физики кластерных

Общая информация

Физико-технический институт имени А.Ф.Иоффе является одним из крупнейших нау центров России, в котором широким фронтом ведутся как фундаментальные, так и прикла исследования в важнейших областях современной физики и технологии.

Институт был основан в **1918** году Абрамом Федоровичем Иоффе, который затем возглавлял е течение нескольких десятилетий. С 1960 года институт носит имя этого выдающегося ученого организатора науки.

Институт является учреждением Российской академии наук и входит в состав Отдел физических наук.

Из событий последних месяцев

Новый этап работ по проекту ИТЭР

6 марта 2012 года, было подписано соглашение между Проектным цен ИТЭР и Международной организацией ИТЭР о поставке диагностиче системы анализа потока атомов, испускаемых термоядерной плазмой.

Система **разработана в ФТИ им. Иоффе** (руководитель р М.П. Петров) и предназначена для контроля соотношения дейтері трития в плазме термоядерного реактора, сооружаемого во Франции участии России. Концептуальный проект этой системы был защищен ра а подписанное соглашение предусматривает изготовление, калибр

Fig. 1. Interface of the TERMAL DB. Upon request, a reference to the Ioffe PTU is provided.

The NIST site (http://trc.nist.gov/ ThermoML.html) gives examples of the transformation of the papers from a group of physicochemical journals into respective XML files. A detailed analysis of the capabilities of ThermoML, for use during the integration of various physical data, has been reported by authors [9].

The MatMl language (XML version) is designed for the solution of similar problems with the only difference being that in this case we deal with materials whose properties depend greatly on manufacturing technology, environmental impacts, etc. rather than substances with a known stoichiometry. As a consequence, MatML does not have a fixed scheme that is suitable in all cases, as in ThermoML, allowing the user to create his own tags without limitation. The MatML elements determine a small set of typical concepts, such as <**Property-Data**>, <**Name**>, <**Units**>, etc. As well, elements are included that help one to determine the origin (the source) of the data, e.g., <**Metadata**> or <**DataSourceDetails**>. MatML was designed to meet any demand in studies or designs without strict relationships with a certain type of object or application. Thus, MatML documents are characterized by high variability with respect to applications.

Language flexibility, which is of the utmost importance for specialists, allows one to manipulate complex numerical data using MatML means when working with, e.g., multicomponent composites. It is possible to work with graphic or micrographic representations. The language extendibility feature also allows one to use other versions through dictionaries such as SVG, that operate graphics, MathML operating mathematical formulas, or femML, which is specially designed for manipulating the data of finite-element modeling. Finally, MatML can be used as a dictionary for other languages through a **namespace** in order to involve material properties within the framework of femML. Schema 3.1 has been released (www.matml.org). The site also contains the history of the development and implementation of the language; its possibilities are outlined and samples of documents transformed into

MatML format are given, e.g., evidence on ceramics from the DB or aluminum alloy properties from a printed handbook. Several successful examples of MatML document exchange in industry were discussed in [10], which is recommended on the site.

3. THE ONTOLOGIES AND CONCEPTS OF THE *SEMANTIC WEB*

Despite the great capabilities of XML as a standard in scientific data exchange, the technology itself is far from the level of integration that was suggested by Tim Berners-Lee et al. in [11] in relation to the concept of the *Semantic Web*. The authors developed the idea of the future of the web when structure is involved in page content that allows program agents to understand its sense and implement user instructions. The objective of the *Semantic Web* is to introduce computer-friendly descriptions to the Internet, which could provide the deepest level of semantic integration.

At that time, the proposed concept sounded realistic because it was based on developed technologies that allowed separate representation of the syntax and semantics of a document. At the moment of publication [11], the XML language and the resource description frameworks (RDFs) were created [12]. The XML language allows the creation of one's own tags, which impart an arbitrary structure to documents. Its syntax is expressed through RDF, which codes structure by means of triplets, each of which formalizes the fact that an object has a certain relationship with a certain meaning. This structure appears to be natural for the description of the majority of computer-processed data.

The true core of the *Semantic Web* is **ontology**, or a system of the concepts of a subject field that is represented as a set of entities that are connected by various relationships. Ontology represents knowledge as a formal structure that is accessible to computer processing. In 2004, the World Wide Web Consortium (W3C) proposed a universal standard for network exchange of ontological information, the Web Ontology Language (OWL). By means of OWL, the experts of the subject field and developers of applications can create, modify, and incorporate different ontologies. Here, OWL is built on the basis of RDFs, which are based on the XML syntax. RDFs and OWL make it possible to generate classes, properties, and individual samples. Therefore, the *Semantic Web* technology provides the integration of resources in the form of a *computer-readable* card, which is called the ontology of the subject field. These cards have the purpose of describing concepts and relationships between concepts.

4. THE ONTOLOGICAL MODELING OF PROPERTY DATA

For data on substance properties, we have several successful examples of their ontological modeling. Most of these relate to materials science where the variety of data types and dictionaries is conspicuous. We may demonstrate this via the PLINUS database, which provides the study results on ceramics properties [13], ontological description of creep in constructional materials [14], and the MatONT system [15], which is designed for supporting research on new materials. The ISO 10303-235 Standard **Engineering properties for product design and verification** [16] has the same aim, however, with a significantly larger scope of information that involves industrial products along with materials. The Standard provides a single information model for determining the semantics and syntax of representation and a single dictionary for determining the data sense.

Table 1. The structure of the Semantic Web for data on the properties of materials

Layer	Content
Ontologies (OWL)	Taxonomy of materials and properties
Metadata (RDF)	Metadata used in a DB
XML Schema	Data scheme for material properties (MatML)

Table 2. The general composition of an ontology for the description of materials

Core ontologies	Peripheral ontologies
Substance	Measurement units
Process	Physical constants
Property	
Environment	
Material Data	

In its most general form, the technology for ontological data description on properties was formulated in [17]. The author focused on the *Semantic Web* concept using a laminar structure with standardized procedures of transition from layer to layer. The lower layer includes XML to determine the data schema, the middle layer (RDF) uses it to determine metadata; finally, the OWL layer uses it to present ontologies (Table 1). It is important that each of these elements alone was standardized long ago. Compared to MatML, this structure is capable of providing a higher standardization level to formalize the determination of properties, as well as the methods for handling and usage.

The ontology that covers the subject field "material properties," according to [17], must involve seven ontologies (Table 2). Core ontologies give structured definitions of terms, names, and dictionaries that represent the basic concepts for each field. Each of them is based on the taxonomy of the classes that are represented in the dictionary of concepts. Examples of the

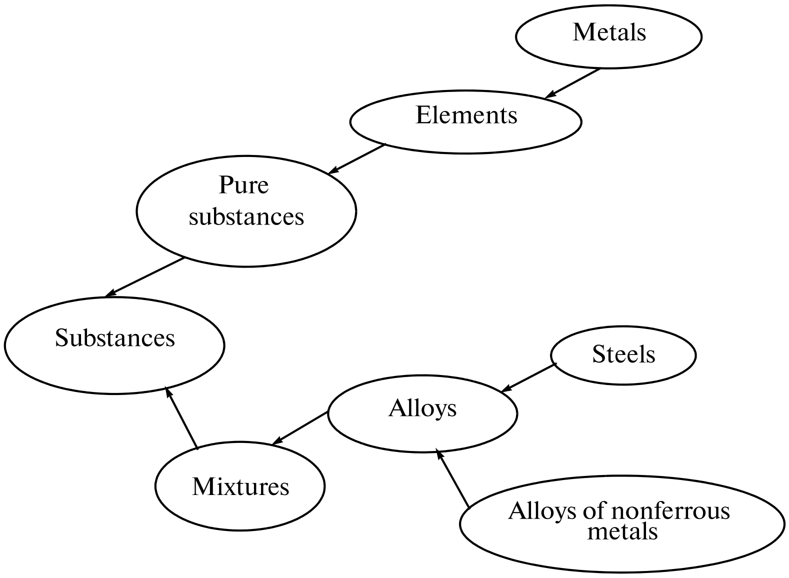


Fig. 2. A fragment of the core ontology "substance."

taxonomies for materials and their properties are given in Figs. 2 and 3. Two other core ontologies, "process" and "environment," give descriptions of the methods for manufacturing and measurement and environmental characteristics (atmosphere composition, temperature, pH, etc.), respectively (not illustrated).

A general ontology also involves material information in order to describe the data on a particular object by aggregating other classes (substance, property, etc.). Using basic ontologies, this ontology aggregates all terms and concepts that characterize a material and a particular specimen, methods and measurement conditions, criteria of data quality, and so forth. According to Table 2, the material ontology includes peripheral ontologies that govern general scientific concepts. In particular, when generating the "measurement units" ontology, the MathML syntax is used (the version designed for the translation of formulas) in order to introduce procedures required for the agreement of different measurement units.

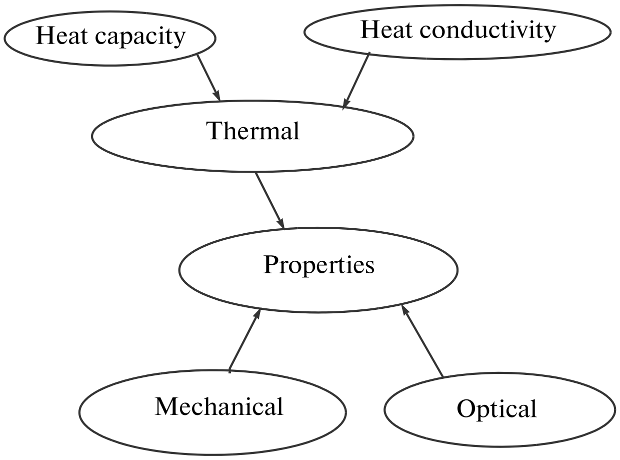


Fig. 3. A fragment of the basic ontology "property."

The material ontology has been tested by the typical data exchange procedure among the group of dissimilar DBs that contain information on physical properties. The technique is reduced to the conversion of the logical structure of each DB into a single structure that is provided by the developed ontology.

Therefore, each of the relational DBs on properties can export the data in a single format that is suitable for both exchange and a long-term archiving. XSLT templates are used to display the fields of each DB; the initial tuning could not be completely automated.

As a whole, the ontology is the structured dictionary of concepts that was adopted in materials science and embedded in the *Semantic Web*. Owing to OWL, general concepts are detected by namespaces and URI identifiers. The material ontology consists of terms and concepts that are general for all resources for the manipulation of data and knowledge. The ontology contains one more component, viz., a digital library of empirical theoretical equations written in MathML, the mathematical mark-up language. All the ontology components use the general format (XML) and can be placed on the Internet.

5. OPPORTUNITIES AND PROMISES

Even superficial analysis of the current state of IR integration reveals a contradiction, viz., the combination of potentially wide opportunities with their rare and occasional usage in practice, as far as properties of data are concerned. A similar situation was discussed earlier in [4], where it was shown that new information technologies are poorly utilized when working with substance properties and that most centers for data preparation and distribution are oriented to the traditional storage techniques of both raw and recommended data. As was noted in [4], the leading role belongs to traditions that were shaped in scientific groups and the absence of meaningful investments, as in business, where the creation of new technologies is welcome. This is relevant in full measure to IR integration, since most existing applications are involved in business and manufacturing. The ISO 10303-235 Standard mentioned above is oriented mainly to the *characterization* of a commercial item, including data representation on physical properties, as a fragment of the information model.

Nevertheless, resource integration on substance properties is actually demanded; therefore, the migration of the corresponding technologies (in particular, the *Semantic Web*) in natural sciences is just a matter of time. We may cite several examples from Russian science where the concepts and tools of integration have penetrated deeply into the practical activities of researchers. Thus, in the Center of integrated information system functions of the Institute of Atmospheric Optics (Tomsk, Russia) applied ontology for the "quantitative spectroscopy" subject area is under development. Using this, extensive information from dissimilar DBs is processed, where the results of the solutions of the direct and reverse problems of finding wave numbers for molecules that are of interest in aerophysics (H_2O, CO_2, NH_3, and H_2S) [18, 19]. The knowledge layer of this system contain the knowledge used for semantic search, integration, and systematization of IRs on molecular spectroscopy. The knowledge base is represented in the form of applied ontologies by which basic problems are solved. The system includes an electronic library of publications from which facts that are related through *concepts*, relationships, and limitations concerning molecular spectroscopy are retrieved. The XML language was used to build the data model for the subject area.

Another successful example of DB integration is the program infrastructure for uniting user interfaces of their own DBs with the tools that were designed for the computer design of unstudied substances and prediction of their properties developed in the Institute of Metallurgy of the RAS. This mainly deals with inorganic substances and materials and is oriented largely to the materials of electronic engineering [20]. The developed technology has provided opportunities for international integration as well, in particular, through conjugating with the Japanese AtomWork DB.

The theme "substance properties" is partially represented in biological DBs, including typical data on the physicochemical properties of protein molecules, various biopolymers, and low-molecular weight compounds. Here, the integration techniques should satisfy severe demands because of the large volume, complexity, and the variety of the data structure in biological DBs. The technology that was developed by the authors in [3] assumes the creation of a virtual DB (MetaBase) for user work under the condition that the data can be stored both in it and in external sources. An approach that is related to preliminary data indexing was chosen where all the data from the DB under integration is indexed and an index is generated with a sufficient quantity of data for satisfying inquiries. An inquiry takes place only within the index frameworks without referring to external bases and the search results are displayed in the form of object profiles and references to external sources. It is also significant that the DB that is under development has a flexible schema to support the possible development of DBs and specify the set of concepts of this subject field during its evolution.

Therefore, the achievements in a number of fields show that the entire set of technologies and tools for BD integration has essentially been developed and is accessible to developers. In the design of the **Data Center** [1], three basic components can be distinguished in its architecture: (1) mechanisms for the reduction of data from a particular information resource to models of the subject field using profession-oriented XML versions (ThermoML, MathML, femML, MatML, etc.) accompanied by the addition, development, and generation of new XML standards; (2) arrangement of an intermediate layer for processing the reduced data to a universal ontologically formed state; (3) organization of the mechanisms for working with universal data (storage, access, exchange, search, etc.).

Thus, the creation of the **Data Center** as a single academic network that unites the information resources in a wide field of knowledge is fairly realistic. Considering the theme of "substance properties" as a suitable test field, it is possible to develop respective standards of integration of data methods from the set of DBs that have been developed and used in the Russian Academy of Sciences.

REFERENCES

1. Zhizhchenko, A.B. and Sotnikov, A.N., Formation of Integrated Distributed Information System of Science, Education and Innovation, *Trudy VII Tverskogo sotsial'no-ekonomicheskogo foruma "Informatsionnoe obshchestvo"*, (Proc. 7th Tver. Social-Economic Forum 'Information Society'), 2011, vol. 2.
2. Kogalovskii, M.R., Methods of Data Integration in Information Systems, Inst. Probl. Rynka Ross. Akad. Nauk. http://www.cemi.rssi.ru/mei//articles/kogalov10-05
3. Miginskii, D.S., Labuzhskii, V.V., Lavrent'ev jr., M.M., et al., Databases Semantic Integration Technology in System Biology, *Computational Technologies*, 2008, vol. 13, no. 6, pp. 102–120.

4. Zitserman, V.Yu., Kobzev, G.A., and Fokin, L.R., Possibilities and Perspectives of Informational Technologies in Preparation and Distribution of Reference Data: Properties of Substances and Materials, *Nauchn.-Tekhn. Inform. Ser. 1*, 2004, No. 2, pp. 7–14.

5. Trakhtengerts, M.S., New Effective Tool for Textual Data Bases, CDS/ISIS for Windows, *Nauchn.-Tekhn. Inform. Ser. 2*, 2006, No. 6, pp. 30–33.

6. Kogalovskii, M.R., Standards of XML Platform and Data Bases. A Review, *Trudy 3-ei Vseross. konf. "Elektronnye biblioteki: perspektivnye metody i tekhnologii, elektronnye kollektsii" 2001* (Proc. 3 rd All-Russ. Conf. "Electronic Libraries: Perspective Methods and Technologies, Electronic Collections"), Petrozavodsk: Karel'skii Nauchn. Tsentr Ross. Akad. Nauk, 2001.

7. Erkimbaev, A.O., Zitserman, V.Yu., and Kobzev, G.A., XML Language Versions in Problems of Storage and Distribution of Scientific Data, *Trudy vseross. nauch. shkoly-seminara molodykh uchenykh, aspirantov i studentov "Intellektualizatsiya informatsionnogo poiska, skantekhnologii i elektronnye biblioteki"*, (Proc. All-Russ. Sci. School–Semin. of Young Scientists, Post-Graduated Students and Students "Intellectualization of Informational Search, Scan-Technologies and Electronic Libraries"), *Taganrog: Taganrog. Tekhnol. Inst. Yuzhn. Federal. Univ.*, 2011.

8. Frenkel, M., Global Communications and Expert Systems in Thermodynamics: Connecting Property Measurement and Chemical Process Design, *Pure Appl. Chem.,* 2005, vol. 77, no. 8, pp. 1349–1367.

9. Erkimbaev, A.O., Zitserman, V.Yu., Kobzev, G.A., and Fokin, L.R., The Logical Structure of Physicochemical Data: Problems of Numerical Data Standardization and Exchange, Russ. J. Phys. Chem. A, 2008, vol. 82, no. 1, pp. 15–25.

10. Kaufman, J.G. and Begley, E.F., MatML. A Data Interchange Markup Language, Adv. Mater. Proc., 2003, vol. 161, no. 11, pp. 35–36.

11. Berners-Lee, T., Hendler, J., and Lassila, O., The Semantic Web, *Sci. Am.*, 2001, vol. 284, no. 5, pp. 35–43.

12. http://www.w3.org/RDF/

13. van der Vet, P.E., Speel, P.-H., and Mars, N.J.I., Ontologies for Very Large Knowledge Bases in Materials Science: A Case Study, *Proc. 2nd Int. Conf. on Building and Sharing Very Large-Scale Knowledge Bases*, University of Twente, 1995, pp. 73–83.

14. Ashino, T. and Fujita, M., Definition of Web Ontology for Design-Oriented Material Selection, *Data Sci. J.*, 2006, vol. 5, pp. 52–63.

15. Cheung, K., Drennan, J., and Hunter, J., Towards an Ontology for Data-Driven Discovery of New Materials, *Proc. AAAI Workshop on Semantic Scientific Knowledge Integration*, Stanford University, 2008, pp. 26–28.

16. Swindells, N., The Representation and Exchange of Materials and other Engineering Properties, *Data Sci. J.*, 2009, vol. 8, pp. 190–200.

17. Ashino, T., Materials Ontology: An Infrastructure for Exchange Materials Information and Knowledge, *Data Sci. J.*, 2010, vol. 9, pp. 54–61.

18. Privezentsev, A.I. and Fazliev, A.Z., Knowledge Bases for Description of Informational Resources in Molecular Spectroscopy. I. Basic Conceptions, *Elektronnye Biblioteki*, 2011, vol. 14, no. 1.

19. Lavrent'ev, N.A., Privezentsev, A.I., and Fazliev, A.Z., Knowledge Bases for Description of Informational Resources in Molecular Spectroscopy. II. Data Model in Quantitative Spectroscopy, *Elektronnye Biblioteki*, 2011, vol. 14, no. 2.

20. Stolyarenko, A.V., Kiseleva, N.N., and Podbel'skii, V.V., Mechanisms of Data Bases and Analytical Tools Integration, *Biznes-Informatika*, 2010, No. 4(14), pp. 60–66.