

# Ontology-Based Technology for Development of Intelligent Scientific Internet Resources

Yury Zagorulko<sup>(✉)</sup> and Galina Zagorulko

A.P. Ershov Institute of Informatics Systems, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia  
{zagor,gal}@iis.nsk.su

**Abstract.** The paper discusses the main features of a technology for the development of subject-based Intelligent Scientific Internet Resources (ISIR) providing content-based access to systematized scientific knowledge and information resources related to a certain knowledge area and to their intelligent processing facilities. An important merit of ISIR is its ability to reduce appreciably the time required to access and analyze information thanks to the accumulation of the semantic descriptions of the basic entities of a knowledge area being modeled, the Internet resources relevant to this area, and the information processing facilities (including web-services) used in it directly in the ISIR content. The specific features of the technology are the use of ontology and semantic network formalisms and orientation to experts, i.e. specialists in the knowledge areas for which ISIR are built.

**Keywords:** Intelligent scientific internet resource · Ontology · Semantic network · Thesaurus · Semantic web service · Technology

## 1 Introduction

Though a great bulk of information related to various areas of knowledge is available directly on the Internet, the problem of supplying the scientific community with information on the subjects of interests has no satisfactory solution yet.

This situation is partly attributed to the specifics of scientific knowledge representation on the Internet, which is weakly formalized, insufficiently systematized and distributed over various Internet sites, electronic libraries, and archives. In addition, a major portion of the information presented on the Internet is practically inaccessible for users because of unsatisfactory operation of modern search engines, which use primitive keyword search mechanisms taking into account neither the semantics of the words contained in the query nor its context.

Another reason for this situation is that modern information systems use a rather reduced set of methods for information representation, search, and interpretation. As a rule, data and knowledge are represented in these systems as text documents (in the Enterprise Document Management System) or a set

of information resources (on the Internet catalogues or portals), though the most human-friendly form of information representation is a network of inter-related facts. Such mode of information interpretation facilitates its perception and allows content-based search and convenient navigation through it.

The problem of a convenient access to information processing means developed in various knowledge areas also remains unsolved. The methods of information processing, even those already implemented and presented on the Internet, remain inaccessible for a wide range of users because of their poor systematization and the absence of semantic information about them.

To solve the problems discussed above, we have suggested a conception of a subject-based Intelligent Scientific Internet Resource (ISIR) intended for the information and analytic support of scientific and production activity in a certain knowledge area. We call the ISIR an intelligent internet resource because not only information representation and systematization, but also all the functionality of this resource are based on the formalisms of ontology [1] and semantic networks [2].

Since the systems of this class are in great demand, we propose a technology for the development and life-cycle maintenance of a subject-based ISIR oriented directly to specialists in the knowledge areas for which such resources are created. This technology is an elaboration of the technology of building scientific knowledge portals that was earlier developed by the authors and successfully used for constructing scientific Internet resources for some knowledge areas [3].

The paper discusses the main features of the technology for the development of subject-based intelligent scientific Internet resources. The rest of the paper is structured as follows: Sect. 2 presents the ISIR conception and architecture; Sect. 3 describes the technology of ISIR development; Sect. 4 gives an example of using the technology; and Sect. 5 discusses some works that are related to the topic of the paper. The main features and merits of the technology, as well as its future evolution, are discussed in the closing section.

## 2 Conception and Architecture of the Subject-Based ISIR

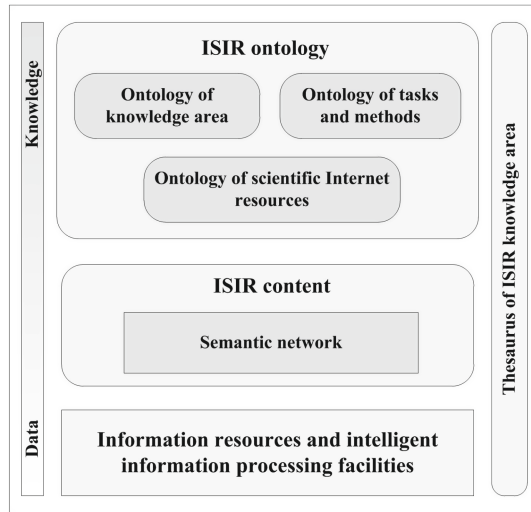
In accordance with the suggested conception, a subject-based ISIR is an Internet-accessible information system which provides systematization and integration of scientific knowledge and information resources related to a certain knowledge area, gives the content-based access to them, and supports their use for the solution of various research and production tasks supplying proper interfaces and services.

### 2.1 Knowledge System of ISIR

As previously mentioned, an intelligent scientific Internet resource is based on the ontology and semantic network formalisms. The ontology is the core of the ISIR knowledge system containing, along with a description of various aspects

of the modeled knowledge area, a description of the structure and typology of information resources and methods of intelligent information processing facilities associated with this area. As for the semantic network, whose structure is defined by the ontology, in ISIR it plays the role of an intelligent data warehouse storing the information about the basic entities of the modeled knowledge area and relevant scientific information resources and about the web-services implementing information processing methods used in this area.

The ISIR ontology (see Fig. 1) consists of three interrelated ontologies responsible for the representation of the knowledge components mentioned above. They are the ontology of a knowledge area, the ontology of scientific Internet resources and the ontology of tasks and methods.



**Fig. 1.** Knowledge system of ISIR

The ontology of a knowledge area defines the system of concepts and relations intended for a detailed description of the ISIR knowledge area and scientific and research activity performed within the frameworks of this area.

The ontology of scientific Internet resources serves to describe information resources related to the ISIR knowledge area and presented on the Internet.

The ontology of tasks and methods includes the descriptions of tasks to be solved by ISIR and methods for their solution and descriptions of web-services implementing both the methods of task solution and information processing methods elaborated in the modeled knowledge area.

Based on the ontology and semantic network, a convenient navigation through scientific knowledge and information resources and intelligent data processing facilities (methods and web-services implementing them) is implemented, as well as the content-based search for information required.

Apart from the ontology and semantic network, the ISIR knowledge system includes a thesaurus which contains terms of the modeled area, i.e. the words and word combinations used for the representation of the ontology concepts in texts and user queries. Using various semantic relations, the thesaurus also determines the meaning of concepts by giving the correlation between concepts rather than their text definitions. Due to this fact, the thesaurus can be applied both to user queries processing and to searching for and annotating the information resources to be integrated in ISIR.

Thus, the knowledge system of ISIR not only includes a formal description of the knowledge area of ISIR, determines the typology of relevant information resources, tasks to be solved in this area and the methods of their solution (by means of the ontology), describes the meaning of concepts used in this area (by means of the thesaurus), but also provides efficient representation of information about real objects of this area, information resources, and intelligent information processing facilities (represented as web-services) to be integrated in ISIR.

### 2.2 Architecture of ISIR

ISIR has a three-tier architecture conventional for information systems (see Fig. 2). It includes a tier of information representation, a tier of information processing and a tier of information storage and access (the base tier).

The first tier is provided by a user interface. The main function of the user interface is representation of user queries and the results of search and task solutions, as well as provision of the ontology-driven navigation through the information space of ISIR. The user interface provides content-based access to

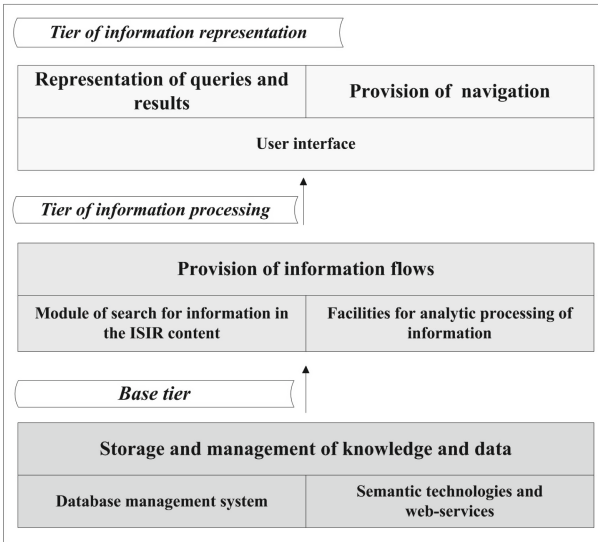


Fig. 2. Architecture of a subject-based ISIR

both the ISIR content and facilities for analytic information processing. In addition, owing to the use of the ontology and thesaurus, the user interface enables a query to be defined in terms of the modeled knowledge area.

At the tier of information processing, various kinds of information search and processing, as well as its transfer between tiers, are provided. For this purpose, the tier includes a module of search for information in the ISIR content and facilities for its analytic processing implemented, among others, as web-services.

The module of search enables one to perform both the information search by keywords and extended semantic search using query representation in terms of concepts and relations of the ontology and constraints imposed on them. These facilities support navigation through the ISIR content by supplying the user interface with a semantic neighborhood of concepts and information objects to be browsed.

The facilities for filtering and visualization of ontology concepts and objects of the semantic network are used as analytic tools.

The base tier ensures the performance of the functions of knowledge (the ontology and thesaurus) and data (the ISIR content) storage and management by means of relational DBMS, Semantic Web technologies and semantic web-services [4].

The platform-independent warehouse Jena Fuseki [5] was selected to be the data warehouse, because it supports the standard query language SPARQL [6], data update and logical inference. In this warehouse, data is represented as a set of triples defining the assertions of “subject-predicate-object” kind corresponding to the well-known RDF model [7]. This data structure is highly flexible in data and knowledge representation, which enables it to store in one place both the ISIR content and specification of the ontology and thesaurus.

Program components of the ISIR base tier providing knowledge and data management are realized with the help of the SPARQL query language. Communication with the data warehouse is performed via SPARQL HTTP client, so the program components are independent of a concrete implementation of the data warehouse. In case of need, the latter can be easily replaced with another data warehouse, more efficient or suitable for a class of tasks to be solved by ISIR.

To store the housekeeping information, in particular, information about administrators and developers of the subject-based ISIR, DBMS MySQL is used.

### 3 Technology of ISIR Building

A technology for building subject-based ISIRs is under development. Its main feature is orientation to experts, i.e. specialists in a certain knowledge area. It stems from the fact that mass production of ISIR for different knowledge areas can be provided only by involving in the process the specialists from a target area. This technology allows them to collect and systematize, within the frameworks of a unified information space, an extensive knowledge and data on the required knowledge area, as well as intelligent data processing facilities used in this area.

This technology has the following basic components:

1. The methodology of ontology building together with a suit of the base ontologies;
2. The expert interface providing access to program facilities which support the construction of ontologies and thesauri and management of the ISIR content;
3. A subsystem for automatized collecting the ontological information from the Internet;
4. The user interface providing a content-based access to the ISIR content and facilities for analytical information processing;
5. A data warehouse providing universal structures for a consistent storage of the ontology, thesaurus and the ISIR content, as well as program facilities supporting access to them.

The methodology of the ontology building is the most important component of the technology suggested, since the ontology is the basis of the ISIR knowledge system. Let us consider it in detail.

### 3.1 Methodology of the ISIR Ontology Building

The ontology of a specific ISIR is built by the methodology whose basic principles are as follows:

- Structuring the ISIR ontology by dividing it into a set of relatively independent ontologies;
- Using a suit of the base ontologies including the most general concepts independent of the ISIR knowledge area;
- Building all ISIR ontologies by means of completion and elaboration of the base ontologies.

The use of this methodology considerably simplifies the construction of the ISIR ontology and its maintenance.

As mentioned above, the ISIR ontology consists of three relatively independent but interrelated ontologies: the ontology of the knowledge area, the ontology of scientific Internet resources, and the ontology of tasks and methods. All these ontologies are built from the base ontologies.

The following ontologies are proposed as the base ones: (1) the ontology of research activity, (2) the ontology of scientific knowledge, (3) the base ontology of tasks and methods, (4) the base ontology of scientific information resources, and (5) the thesaurus representation ontology.

The ontology of the knowledge area is built on the basis of the first two ontologies; the ontology of tasks and methods is built on the basis of the third one; the ontology of scientific Internet resources is built using the fourth one. The thesaurus representation ontology supplies a set of concepts and relations for building the thesaurus of the ISIR knowledge area. Let us describe these ontologies in detail.

The scientific knowledge ontology contains classes which define the structures for the description of the concepts of specific areas of knowledge, such as

Subdivision of science, Research method, Object of research, Scientific result, etc. The ontology also includes the relations linking the objects of these classes. Using these classes we can extract and describe divisions and subdivisions that are significant for a given knowledge area, determine classification of methods and objects of research, and describe the results of research activity.

The ontology of research activity is based on the ontology suggested in [8] for describing research projects and extended for applying it to a wider class of tasks. The ontology includes classes of concepts relating to the organization of scientific and research activities, such as Person, Organization, Event, Scientific Activity, Project, Publication, etc.

This ontology also contains relations which enable us to link its concepts not only to each other but also to the concepts of the scientific knowledge ontology. Note that the choice of these relations was based on both the completeness of the ISIR knowledge area presentation and the convenience of navigation through the information space of ISIR and information search in it.

The base ontology of scientific information resources includes Information resource as the main class. This class serves to describe the information resources relevant to the ISIR knowledge area (including the resources presented on the Internet). The set of attributes and relations of the Information resource class is based on the Dublin Core standard [9]. It has the following attributes: Title of resource, Language of resource, Subject of resource, Resource type, etc. To represent information about the sources and creator of the resource, as well as events, organizations, persons, publications and other entities associated with it, special relations are included in the ontology.

The base ontology of tasks and methods contains classes, such as Task, Method of solution, and Web-service, as well as the relations linking these classes to each other and to the classes of other base ontologies. Using these classes and relations, we can describe the tasks to be solved by a specific ISIR, methods of their solution, as well as the web-services implementing them. Besides, this ontology also describes the web-services implementing the information processing methods used in the ISIR knowledge area.

The descriptions of web-services are based on the OWL-S ontology [10] intended to describe semantic web-services. Due to this, a web-service is linked not only to the description of its interface in terms of types of input and output data, but also to a description of its semantics, i.e. what the service can do, its subject domain, constraints on the application area and service quality, etc. Besides, all its declarative properties, functionality, and interfaces are encoded in a single-valued form applicable to machine processing.

The presence of a semantic description of web-services not only facilitates their search and correct use (performance), but also makes it possible to compose from them new services in order to obtain the functionality required to solve the user problems. In addition, available semantic descriptions of the web-services predetermine their successful integration into ISIR. Besides, the content-based access to the web-services will be provided not only for software agents, but also for those who want to find intelligent information processing facilities necessary to solve their tasks.

The thesaurus representation ontology is based on the international and Russian standards regulating the structure of monolingual and multilingual information retrieval thesauri, the set and properties of the basic entities and relations between them, therefore it includes a suit of generic concepts and relations which are present in any thesaurus. In particular, it contains classes describing the following thesaurus entities: terms which are divided into descriptors (preferred terms) and ascriptors (text entries which can be replaced by the corresponding descriptors during document indexing and retrieval), sources of terms (web-resources, text documents, and collections of text documents containing terms or their definitions), and subareas of knowledge related to terms. The ontology also includes the relations that link the objects of classes listed above.

The ISIR knowledge area thesaurus is created by supplementing with specific terms the thesaurus core which is built on the basis of the thesaurus representation ontology and contains the set of terms corresponding to the names of classes and relations of the base ontologies.

### **3.2 Management of the ISIR Knowledge System**

To support the process of adjustment and management of the ISIR knowledge system, the technology provides developers with ontologies and data editors. These editors have convenient graphical interfaces implemented as web-applications and provide remote adjustment and management of the ISIR knowledge system by authorized users (experts) via the Internet. To support cooperative development of the ISIR knowledge system by a team of experts, the editors have a procedure for granting privileges to experts of different levels.

The ontology editor serves for ontology building and management. Its design enables it to be used not only by knowledge engineers, but also by experts who are not specialists in computer science and mathematics.

The ontology editor allows an expert to create, modify and delete any elements of the ontology (classes of concepts, relations, and domains). When an expert describes a new class, he/she can select its parent from the set of already created classes which is represented for a user as a tree. Thereby the class inherits from the parent class not only all its attributes, but also its relations; at the same time, the parent class is linked with a new class by a "subclass" relation. For each attribute of the class, its name, the range of values, the number of possible values (one or a set), and the status of value filling (mandatory or not) are defined. When a new relation is created, its arguments are also selected from the tree of classes.

ISIR content management is implemented with the help of the data editor which operates under the control of the ontology. This allows one not only to facilitate considerably the correct insertion of information, but also to provide its logical integrity. The data editor allows one to create, modify, and delete information objects (the objects of classes defined in the ontology) and relations between them.



When a new information object is created, first of all, the expert selects the corresponding class of ontology from the tree of classes. Then, based on the description of this class presented in the ontology, a form for information insertion is automatically generated. This form contains entry fields for the values of attributes of the object and its relations with other objects. If an attribute takes its value from a domain, then the list of its possible values is displayed.

Simultaneously with the object creation, the expert can specify its connection with other objects already existing in the ISIR content. The type of these connections and classes of these objects are defined by the corresponding relations of ontology. The form for their input is automatically generated on the basis of descriptions of these relations. Based on these descriptions and the current state of the ISIR content, the data editor displays for each relation of the created (edited) object a list of objects with whom the object can be linked by a given relation.

The thesaurus building and editing its content are also implemented with the help of the data editor operating under the control of the thesaurus representation ontology. This provides the logical integrity of the thesaurus terminological system.

To be a useful resource, ISIR should have the knowledge system which contains exhaustive information about the modeled knowledge area and the scientific activity performing within its frameworks. Building and maintenance of such resource is a rather complex and labor-intensive problem requiring considerable efforts of developers.

The complexity of the problem is due to a large variety of kinds of the information collected and modes of its representation on the Internet. In particular, information about organizations, persons, projects, conferences and publications is collected from information portals, digital libraries and journals, web-sites of organizations, projects, conferences, etc. The great labour input is attributed to the large volume of information to be collected.

To solve this problem, a subsystem intended to automatize the collection of information about the basic entities of the ISIR knowledge area and Internet resources relevant to it is developed. The subsystem unites the methods of meta-search and information extraction based on ontologies and thesauri.

Information collection for ISIR consists of the following stages: (1) search for Internet resources relevant to the ISIR knowledge area, (2) extraction of information from these resources, and (3) insertion of obtained information into the ISIR content. According to this, a subsystem of information collection includes a module of a search, a module of information extraction, a module of information insertion into the ISIR content, as well as a data base for the storage of links to the Internet resources (DB LIR). Note that when ISIR is adjusted to a knowledge area DB LIR can be filled with the Internet links to the relevant (according to the experts opinion) Internet resources.

At the first stage, the search queries used by the search module for the retrieval of relevant Internet resources are generated on the basis of the ontology and thesaurus. This module addresses the search systems of Google, Yandex

and Bing via their program interfaces. It uses meta-search methods to obtain links to the Internet resources. Then this module filters duplicate and irrelevant links and adds relevant links in the DB LIR. (Note that relevance of a resource (Web page) is defined on the basis of the cosine similarity measure calculated between the vectors of weights of the terms of the search query and the Web page downloaded by its link.)

At the second stage, relevant Internet resources are analyzed and information is extracted from them. A feature of the approach to information collection implemented here is that for every type of entities (the ontology class) a specific method of information collection adjustable to the knowledge area and kinds of Internet resources is developed. Each of these methods includes a set of patterns. In these patterns, for every kind of extracted information, markers defining its position are given as well as the engines implementing the algorithm of the analysis of the corresponding fragments of Web pages and extraction of the required information from them. These patterns are also generated on the basis of the ontology. To improve the recall (completeness) of information extraction, the patterns use alternative terms from thesaurus (synonyms and hyponyms) to describe the markers.

At this stage, DB LIR can be also updated with the Internet links found in the Internet resources processed. Subsequently, these links are analyzed by the experts who decide on their relevance.

At the third stage, the information extracted at the previous stage is inserted in the ISIR content.

#### 4 Use Case: ISIR in Decision-Making Support

The technology of ISIR development was used for creating a resource in the area of decision-making support. This resource contains systematized information about the knowledge area “decision support” and methods for solving the problems specific for this area, and provides the content-based access to them.

In accordance with the methodology, the ontology of the resource in question is built from base ontologies. The ontology of the knowledge area has been extended by such concepts as Decision-making process, Decision-making stage, Situation, Problem situation, Alternative, and the relations between them. The base ontology of tasks and methods was supplemented with the entities concretizing the concepts and relations contained in it. For example, for the class Task, the subclasses Structuring of knowledge area, Situation analysis, Objectives formulation, Criteria formulation, Alternatives development, Criteria evaluation, and others were introduced. These subclasses describe the tasks to be solved at different stages of decision-making.

A structured description of the entities of this resource gives an idea about all aspects of decision-making support and the use of specific methods. It allows one to obtain answers to the questions like “What problems are solved at certain stages of decision-making?”, “What methods are used to solve this problem?”, “What input data are required for this method?”, and “What solvers, frameworks or web-services are available for the implementation of the method?”.

It is also possible to get both informal and formal descriptions of a method, information about groups and teams developing this method, and links to Internet resources relevant to the method.

Figure 3 shows the user interface of the resource as a page with ontology and description of the interior point method belonging to the class of linear programming methods.

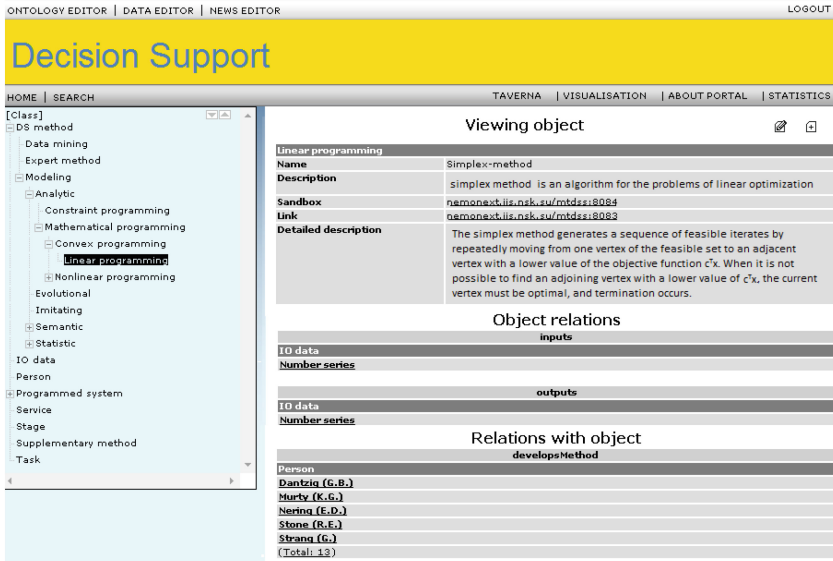


Fig. 3. User interface of the resource on decision-making support

The basic functionality of the resource is implemented using the services described in Sect. 2.2. Let us consider some particular services created specifically for this resource.

Since the description of decision support methods often requires a complex mathematical description, a service has been developed that provides advanced text edition functions including mathematical formulas editing.

The methodology of web-services development was tested on linear programming methods. A web-service was designed in the C# language using the data transfer protocol SOAP running under HTTP protocol and the WSDL language [11]. This web-service can be used in two ways: to explore its work and to embed it into third-party applications. For these purposes, it is supplied with appropriate user interfaces. In the first case, the interface is a kind of a “sandbox” which allows one to set the input data, run the chosen method and look through the results of its work. In the second case, the interface enables one to obtain a method specification, a description of the input data, as well as the examples of requests to the service and its responses.

The user interfaces of the web-service are accessible from the page containing the description of the method. For example, a page with the description of the interior point method shown in Fig. 3 contains a detailed text description of this method, references to the “sandbox” and method specification.

## 5 Related Work

A large variety of research works are devoted to the solution of the problem of systematization and integration of information resources related to a certain domain or community and supporting convenient access to them.

To solve the problem of integration of information resources which are heterogeneous in structure and content as well as in data access technology, a scientific Resource Space Management System (sRSMS) has been developed [12]. The sRSMS aims at providing a homogeneous view over and access to a space of scientific resources sourced from the Web and accessible via a variety of different heterogeneous technologies. Using this system it is possible to develop applications which enable the users to operate the scientific resource space via domain-specific, intuitive instruments. This is provided by abstracting the various kinds of scientific knowledge into a uniform conceptual model, by abstracting the operations supported by the services, by providing access to scientific knowledge (from merely accessing paper data/metadata to extracting and tagging content, crawling citations, submitting for review, etc.), and by hiding the technical details of accessing heterogeneous platforms/resources.

The approach, discussed in [13], is aimed at the development of structured community portals that extract and integrate information from Web pages and present it in a unified view of entities and relationships accepted in the community. This approach is top-down, compositional, and incremental. At first, experts select a small set of data sources highly relevant to the community. In order to facilitate the selection, a special tool to search a data source and rank it using the PageRank and TF-IDF metrics for measuring their relevance is provided. Next, plans extracting data from these sources and representing them in the form of entities and relationships are composed from a given set of extraction/integration operators. These plans are built for every entity. Executing these plans yields an initial structured portal (in the form of an entity-relationship graph). Then this portal can be incrementally expanded by detecting and adding new sources.

Note that the first approach involves the programmer in the development of applications (for building adapters to the resources integrated in them). In the second approach, only experts (for the selection of relevant data sources) and a knowledge engineer (for the definition of an ER schema and composition of plans extracting information) are needed. Though in outward appearance our approach is similar to the first approach, in spirit it is closer to the second one. In contrast to it, however, our approach provides a more detailed description of the modeled knowledge area, supplying the user with a set of base ontologies containing a wide spectrum of entities for representing research activity.

As for the collection of information from the Internet, many researchers deal with this problem. However, as shown in the survey [14], the majority of such

studies are aimed at the extraction of information needed for the solution of the tasks of electronic commerce or analysis of social networks [15], and only a minor part of this research collects information for the needs of research activity [16]. In particular, in the study [16] information about the scientific content of Knowledge-Based Systems journal for the period 1991–2014 extracted from the bibliometric database ISI Web of Science is used to study the knowledge field of the journal and the community formed around it with the help of performance analysis and science mapping methods [17].

## 6 Conclusion and Future Work

The paper presents the main features of the technology for the development of subject-based intelligent scientific Internet resources (ISIR) providing content-based access to the systematized scientific knowledge and information resources related to a certain knowledge area and to intelligent processing facilities used in this area.

An important merit of ISIR is that it allows the researchers to reduce appreciably the time required to access and analyze information thanks to the accumulation of the semantic descriptions of the basic entities of a knowledge area being modeled, the Internet resources relevant to this area, and the information processing facilities used in it directly in the ISIR content.

The use of ontologies as a conceptual basis of ISIR makes it a convenient and efficient tool for the representation and systematization of all kinds of information necessary for the researcher about the knowledge area modeled.

The use of ontologies also raises the level of the ISIR development technology and creates prerequisites for its use by experts, i.e. specialists in the modeled knowledge area. On the one hand, basing the facilities for development and maintenance of the ISIR knowledge system on ontology makes them available for experts since knowledge representation in the form of objects (classes) and relations between them accepted in the ontology is most natural for humans. On the other hand, ontology is a convenient means to formalize and fix common knowledge, shared by all experts-developers, about the area modeled. In addition, ontology enables knowledge reuse, which simplifies and speeds up the development of new applications. Moreover, the availability of a representative suit of the base ontologies and methodology for building all ISIR ontologies on their basis, together with knowledge and data editors convenient for experts, makes ISIR building much easier and less labour-consuming.

However, our experience of applying the technology suggested has shown that though it provides convenient facilities for building ISIRs, it remains complicated for experts. Therefore, the future development of the technology will aim at eliminating this shortcoming. For this purpose, within the frameworks of the technology specialized program shells oriented to particular classes of knowledge areas and tasks will be built. The shells will differ in base ontologies sets and, possibly, program components. Such shells will be virtually “empty” ISIRs, i.e. they will include all the necessary conceptual and procedural components, but

the lower levels of their ontologies will be undeveloped and their contents will be empty. We suppose that the development of such specialized shells will make the technology of ISIR building completely applicable for experts.

**Acknowledgments.** The authors are grateful to the Russian Foundation for Basic Research (grant No. 13-07-00422) for financial support of this work.

## References

1. Guarino, N.: Formal ontology in information systems. In: Proceedings of FOIS 1998, Trento, Italy, pp. 3–15. IOS Press, Amsterdam (1998)
2. Deliyanni, A., Kowalski, R.A.: Logic and Semantic Networks. *Comm. ACM* **22**(3), 184–192 (1979)
3. Borovikova, O., Globa, L., Novogrudska, R., Ternovoy, M., Zagorulko, G., Yu, Z.: Methodology for knowledge portals development: background, foundations, experience of application, problems and prospects. In: Joint NCC&IIS Bulletin, Series Computer Science, vol. 34, pp. 73–92 (2012)
4. McIlraith, S.A., Son, T.C., Zeng, H.: Semantic web services. *IEEE Intell. Syst.* **16**(2), 46–53 (2001)
5. Fuseki: serving RDF data over HTTP. [http://jena.apache.org/documentation/serving\\_data/](http://jena.apache.org/documentation/serving_data/)
6. SPARQL Query Language for RDF. W3C Recommendation, 15 January 2008. <http://www.w3.org/TR/rdf-sparql-query/>
7. Resource Description Framework (RDF). <http://www.w3.org/2001/sw/wiki/RDF>
8. Benjamins, V.R., Fensel, D.: Community is Knowledge! in (KA)2. In: Proceedings of 11th Banff Knowledge Acquisition for Knowledge-based Systems Workshop, KAW 1998, Banff, Canada, April 1998. SRDG Publications, University of Calgary, Calgary (1998)
9. Hillmann, D.: Using Dublin Core. <http://dublincore.org/documents/usageguide/>
10. OWL-S: Semantic Markup for Web Services. <http://www.w3.org/Submission/OWL-S/>
11. Chinnici, R., Moreau, J.-J., Ryman, A., Weerawarana, S.: Web Services Description Language (WSDL) Version 2.0 Part 1: Core Language. <http://www.w3.org/TR/wSDL20/>
12. Parra, C., Baez, M., Daniel, F., Casati, F., Marchese, M., Cernuzzi, L.: A scientific resource space management system. Technical report DISI-10-013. University of Trento, Ingegneria e Scienza dell'Informazione (2010)
13. DeRose, P., Shen, W., Chen, F., Doan, A.H., Ramakrishnan, R.: Building structured web community portals: a top-down, compositional, and incremental approach. In: VLDB 2007, Vienna, Austria, 23–28 September 2007, pp. 399–410 (2007)
14. Ferrara, E., De Meo, P., Fiumara, G., Baumgartner, R.: Web data extraction, applications and techniques: a survey. *Knowl.-Based Syst.* **70**, 301–323 (2014)
15. Bernabe-Moreno, J., Tejada-Lorente, A., Porcel, C., Fujita, H., Herrera-Viedma, E.: CARESOME: a system to enrich marketing customers acquisition and retention campaigns using social media information. *Knowl.-Based Syst.* **80**, 163–179 (2015)

16. Cobo, M.J., Martinez, M.A., Gutierrez-Salcedo, M., Fujita, H., Herrera-Viedma, E.: 25 years at knowledge-based systems: a bibliometric analysis. *Knowl.-Based Syst.* **80**, 3–13 (2015)
17. Cobo, M.J., Lopez-Herrera, G., Herrera-Viedma, E., Herrera, F.: Scimat: a new science mapping analysis software tool. *J. Am. Soc. Inf. Sci. Technol.* **63**(3), 1609–1630 (2012)