Social Network Technologies for Semantic Linking of Information Objects in Scientific Digital Library

M. R. Kogalovsky
Market Economy Institute, Russian Academy of Sciences,
Nakhimovskii pr. 47, Moscow, 117418 Russia

S. I. Parinov
Central Economics and Mathematics Institute, Russian Academy of Sciences,
Nakhimovskii pr. 47, Moscow, 117418 Russia

e-mail: kogalov@ipr-ras.ru, sparinov@gmail.com

Programming and Computer Software, 2014, Vol. 40, No. 6, pp. 314–322.

Abstract. In the last decade, scientific digital libraries were traditionally used for publishing research results and for enabling wide open access to them. Functional capabilities of digital libraries can be extended by offering users the opportunity of linking information objects of the library and providing for created linkages explicitly defined semantics based on a given ontology. Such an activity of users, which is peculiar to social networks, motivated by different reasons, and carried out on their own initiative, results in the dynamic semantic structure of the digital library content. In the environment of such a kind of a social network, certain new forms of scientific activities become possible and data sources can be created that provide more information for scientometric researches as compared to presently available ones. In this paper, we propose an approach for creating such networks and discuss results of its implementation in the Socionet environment; the Socionet is a large-scale online information space that covers information resources of a number of scientific, educational, etc. organizations. This work was supported by the Russian Foundation for Humanities, project no. 14-02-12010-v.

1. INTRODUCTION

In the last decade, active studies and developments have resulted in large-scale representative scientific digital libraries created for different knowledge domains; these digital libraries begun to play an important role in supporting scientific studies and in increasing their efficiency. Basic functions of such systems consist in publishing research results and in enabling wide open access to them. However, the functional capabilities of the scientific digital library can be extended by offering users the opportunity of linking information objects of the digital library and declaring explicitly semantics of created linkages based on a given ontology; we hereinafter refer to such linkages as *semantic linkages.*

It is supposed that the users have different motives to create, on their own initiative, binary oriented semantic linkages between information objects of the digital library content; it is also supposed that the users can receive notifications about new linkages created or about changes in properties of existing linkages and respond to such events by creating new linkages. Participants of semantic linkages may be scientific publications, sets of scientific data, profiles (metadata) of their creators and other users registered in the digital library, profiles of organizations in which they work, research programs represented in the form of digital documents, scientific reports, reviews, project descriptions, ontologies of different knowledge domains, software features and their descriptions, specifications of metadata standards, etc.

According to the linkage ontology used in a given digital library, the linkage between a user profile and a scientific publication can be an opinion on the entire publication or, in particular, its research results only (a kind of a review). The linkage between two publications may point to the fact that the method, approach, or data used in one publication are used in the other publication or, for example, may point to the plagiarism. In this work, we suppose that semantic linkages are standalone objects of the digital library. Therefore, the linkage between a user profile and a certain other linkage may express an opinion of the user on the fact or opinion expressed by this linkage. The semantics of linkages can be rather diverse and depends on types of the information objects being linked. In this paper, we confined ourselves to the above mentioned examples.

Let us only point to an important special case, namely, citation linkages between scientific publications. Such linkages contained in lists of references form a data source for traditional scientometric researches. The disadvantage of the scientometrics based on such a source is the lack of the explicit semantics of citation linkages. These linkages can naturally be called *mute* [1], since they provide no information to characterize, for example, an opinion of the author about this cited source or a motive of citing. Thus, paradoxical situations may arise when a paper is highly evaluated, due to a high citation index, because it contains gross errors and/or fundamental delusions concerning a problem under study and, therefore, stimulate active reaction of scientific community. In order to avoid such situations, we propose to endow also citation linkages with the semantics. The semantics of these linkages is defined by the users of the digital library (the authors that cite a given publication) based on the ontology used. Moreover, this can be done by other users that are experts in a given knowledge domain.

Thus, in an online digital library in the mode of a social network, the activity of users on creating semantic linkages can be ensured. Hence, certain new forms of scientific activities become possible whose results will be open to the entire scientific community, which implies a higher level of individual responsibility. In particular, these new activities involve alternative forms of reviewing scientific publications, the structurization of scientific knowledge available in a given library, and the “electronic” discussions of certain publications or other products of scientific activities.

Moreover, semantic linkages created by the user form a multilayer semantic structure of the content of the digital library, i.e., its *semantic halo* [2]. Each layer of the semantic structure corresponds to a certain class of linkages defined in a given ontology. This structure can be used as a new source of data for scientometric researches that contain more information as compared to traditional data sources; in addition, it can be used for investigating properties of knowledge corpus in a particular scientific domain. By exploring layers of the semantic structure, one can discover ways in which particular schools were formed, trace the history of their development, and obtain other useful data. New scientometric indexes can be created that take into account the semantics of references [3, 4] in contrast to presently available systems (citation indexes). Note also that the users of the digital library with semantically structured content are provided with additional efficient tool for accessing the information resources available in the digital library: the visual semantic navigation through the structure of the content. To perform the activities mentioned above, the users registered in the digital library are provided with the open access to the current version of the semantically structured content, to the facilities for creating semantic linkages between information objects, and to other services.

In recent years, specific ontologies are developed for the explicit description of the semantics of linkages between information objects in information systems. These ontologies describe different classes of citation linkages, as well as linkages of different natures, like, for example, “author–publication,” “organization–author,” “publication–fragment of publication” (for example, abstract, contents, preface, references, etc.), linkages between versions of a publication or its different representations, etc. [5–9].

We mentioned the importance of using semantic linkages in scientific digital libraries in our previous work [10]. The general approach to linking digital data based on the Web technology was proposed by Berners-Lee [11]; this approach was further developed in the direction called the Open Linked Data (see, for example, [12]). In recent years, the need for standardization in this field was recognized. This resulted in the Open Annotation project initiated by Open Annotation Collaboration (OAC) based on a group of US universities. The purpose of this community is to create standards for semantic linking of digital resources in the Web environment. Basic materials developed by the OAC became a foundation for the W3C Open Annotation Community Group [14], which is intended to create general RDF-based specifications for digital resource annotation. In particular, recommendations on the annotation ontology and annotation model are being developed. Today, a draft of recommendations is available. Moreover, a number of research projects were implemented in which the explicit specification of the linkage semantics is provided on the base of ontologies; these are, for example, Nanopub.org [15] and SiteULike.org [16].

In this work, we develop and practically implement the approach proposed in [10] that is focused on working with scientific information resources and on using the corresponding technology in scientific activities. This approach is based on the ontology created by the authors using fragments of well-known ontologies. So, citation linkages, as well as linkages of other types, are taken into account. As compared to well-known projects, new functional capabilities of the proposed approach are provided, first and foremost, by representing semantic linkages as standalone information objects [1, 17–20]. Such a representation of semantic linkages is a key feature of the proposed approach. It allows also to embed their descriptions into metadata of information objects being linked. Semantic linkages can be participants of other linkages to implement additional types of relationships between information objects of the digital library.

In this paper, we discuss the proposed approach and results of its implementation in the Socionet environment [10]; the Socionet is a large-scale online information space that includes information resources of a number of scientific, educational, etc. organizations. The Socionet system is based on standards developed by the *Open Archives Initiative* [21]; the system has been used for almost a decade and a half.

In the following section, we discuss possible representations of semantic linkages (their descriptions) in the content of the scientific digital library and substantiate the advantage of their representation as standalone information objects. Section 3 is dedicated to methods for creating semantic linkages. In Section 4, we discuss the use of the linkage ontologies and ontology-based controlled vocabularies for describing semantic linkages, briefly characterize presently available ontology projects the results of which are used in this work, and present a set of semantic linkages in a digital library as a new data source for scientometric researches. Finally, we discuss results of implementing the proposed approach in the Socionet system environment in Section 5.

2. REPRESENTATION OF SEMANTIC LINKAGES

In the scientific digital library, there is a certain information metaobject (descriptor) for each publication, author, organization, and other types of objects; this metaobject contains the set of metadata that determine properties of the object. Descriptors of authors and organizations are called their *profiles*. Semantic linkages between information objects can be created. In the proposed approach, only *binary oriented semantic linkages* are used. An information object from which the linkage originates is hereafter referred to as the *source linkage object*, while an object to which the linkage is directed is called the *target object.*

If it is required to create and explicitly represent the linkage between a given and a certain other object (for example, citation linkages), then such a linkage should be explicitly described by a certain set of metadata. In known projects [15, 16], metadata of a linkage are included into the descriptor of a source linkage object. In the case of such an embedded representation of linkages, only the person with the authority of updating descriptors of information objects can create linkages. As a rule, such an authority is given to system administrators and authors of information resources but not to other users of the digital library.

An alternative method is the *autonomous* representation of linkages between information objects [1, 17, 18]; this method represents linkage metadata as *standalone information objects.* Linkage descriptors and descriptors of objects being linked (publication descriptors, user profiles, organization profiles, etc.) are used to represent linkages in the digital library. A standalone linkage descriptor contains unique identifiers of linked information objects, description of the linkage semantics, identifier of the linkage creator, creation date, textual comment, and certain other characteristics. In this case, linkages can be created not only by administrators of information resources but also by ordinary users registered in the library.

In addition to performing the descriptive function, linkage descriptors can be participants of other linkages as ordinary information objects. This cannot be done in the case of the embedded representation. Moreover, the autonomous representation offers the following advantages over the embedded one. On the base of the content of the digital library, the *repository of semantic linkages* can be created; this repository can serve as an independent data source, which is important for scientometrics. By integrating data of the repositories created in different digital libraries for a particular knowledge domain, one can build rather representative global repositories of semantic linkages for the given domain. For libraries based on the open archive technology (Open Archives Initiative, OAI) [22], the semantic repositories, as well as contents of such libraries can easily be integrated.

The autonomous representation of linkages maintains the virtual environment to provide the users with the possibility of working as experts in the mode of a social network. On their own initiative, they express their opinions and evaluations on particular publications and other scientific information resources available in the digital library, as well as on various relationships between them represented by semantic linkages. These opinions and evaluations are fixed in the system in the form of semantic linkages of a corresponding class between the expert profile and the descriptor of a particular information object. This supplements the traditional anonymous reviewing of publications, while the openness of opinions expressed by the scientific community allows one, in turn, to respond to these opinions while promoting the responsibility and objectivity of their authors. For example, such an activity is provided by the F1000Research service of the FACULTYof1000 project [23]. The proposed approach also supports structured data in semantic linkage descriptors that can be used for sophisticated scientometric researches.

Our approach supports both methods of semantic linkage representation described above. However, in this paper, we focused on the autonomous representation.

3. DESCRIPTION OF SEMANTIC LINKAGES

In recent years, linkage ontologies became a basis for implementing user tools that allow one to describe the semantics of linkages between information objects in digital libraries and other information systems. In particular, these ontologies determine the semantics of linkages between various scientific objects (publications, research topics, scientists, research organizations, etc.). In these ontologies, the hierarchy of linkage classes is defined; the linkage classes correspond to various kinds of relationships between information objects.

One of the pioneer projects in this field is *the Semantic Publishing and Referencing* (SPAR) ontology complex created by researchers from the Oxford University and University of Bologna [7, 8]. SPAR includes eight independent ontologies that describe the semantics of bibliographic objects and their relationships. The first four ontologies (FaBiO, CiTO, BiRO, and C4O) intend to description of bibliographic objects, bibliographic sources in references, citation linkages, citation contexts, and their linkages with relevant sections of cited publications. The other four ontologies (DoCO, PRO, PSO, and PWO) describe the semantics of components of documents, roles and states of publications, and workflows in publishing processes. All the ontologies are described in the OWL2 DL and RDF languages developed by the W3C group.

Another important project in the area under consideration is *Semantic Web Applications in Neuromedicine* (SWAN) [6], which is also a modular ontology complex created in the Massachusetts General Hospital of the Harvard Medical School. The aim of the SWAN ontology is to enable a *social-technical ecosystem* (a comfortable environment in the semantic Web) in which semantic context of scientific communications can be created, stored, accessed, and integrated. Also, an exchange by unstructured and semistructured digital scientific information may be executed. Ontologies included into SWAN are described in the OWL DL language.

Here, we should mention the *Simple Knowledge Organization System* (SKOS) recommendation developed by the W3C [9] for support of knowledge organization systems such as thesauri, classification schemes, subject heading systems, and taxonomies in Semantic Web environment. The SKOS defines a conceptual schema called as a *common data model* for sharing and linking knowledge organization systems via the Web. The unified conceptual scheme of the SKOS simplifies the integration of existing knowledge organization systems into the semantic Web.

Note that the *modularity* of complex ontologies, such as SPAR and SWAN, simplifies their reuse. In certain applications, there is no need for the entire ontology and its individual modules can be used. Moreover, integration of ontologies is simplified. For example, the SPAR complex uses SWAN elements, while the SWAN uses SKOS elements.

An important contribution to this field is made by the euroCRIS organization, which develops the CERIF project. A main result of this project is the creation of a unified conceptual scheme called the *Full Data Model* [24]. This model is a basis for constructing the current research information systems (CRIS) meant for supporting scientific and organizational activities in different countries and various scientific organizations. The standardization of the conceptual scheme enables the interoperability of such systems. Recently, the specification of the standardized semantics for the Full Data Model was proposed in the frameworks of CERIF [5]; this is an ontology that defines the set of terms for describing entities of the Full Data Model and relationships between them.

The ontologies and combinations of their modules described above can be used in various digital libraries for creating ontologies of semantic linkages between information objects that are more adequate to the nature of information resources and functional mechanisms of a particular digital library (for example, its scientometric apparatus).

The structure of semantic linkages defined on the content of the digital library is *multilayer* [1, 17, 18]. Each layer of this structure corresponds to a certain class of linkages defined in a given ontology. As noted above, in addition to being a data source for scientometric researches, the semantic structure can be used for semantic navigation through the content of the digital library. For practical purposes, it is convenient to use the *semantic linkage taxonomy* that is based on the ontology adopted in a particular project; this taxonomy may be represented as a set of *controlled vocabularies* of classes of semantic linkages. Moreover, several taxonomies based on different ontologies can be supported. By default, the *embedded* (basic) taxonomy with its own set of controlled vocabularies is used for user requests specification. Otherwise, the taxonomy specified in the request is used.

Finally, note that linkages of certain classes only are allowed for a given pair of information objects in digital library. Permissibility of linkages of each class is defined by types of objects being linked. In the digital library, the permissibility matrix should be used in addition to the taxonomy of semantic linkages; this matrix defines what linkage classes are allowed for each pair of information objects types.

4. CREATION OF SEMANTIC LINKAGES

There are three methods for creating semantic linkages between information objects of the digital library that are represented as standalone information objects.

The first method is based on competencies of scientists acting as experts and assumes *manual* creation and description of linkages via a dedicated user interface. When creating a linkage, the expert specifies identifiers of the source and target objects in the linkage descriptor. Using the basic or other linkage taxonomy, the expert specifies a class to which the new linkage belongs. The linkage descriptor contains also identification data of the expert; linkage creation date; and, if necessary, a comment [19, 20].

The second method is used to create citation linkages based on *mute* linkages in the text of a publication. In this case, the preprocessing of the citing publication is required. The preprocessing consists in that the expert looks through the text of the publication (rather than the references) to find bibliographical references to the sources used. By analyzing the context of each reference and using the controlled vocabularies of semantic linkages, the expert performs the *ontological annotation* of the reference [26]. Of course, this can be done by the author of the publication. Then, the tagged text is processed by a special mechanism of the system to generate descriptors of the corresponding linkages for each annotated reference; these linkage descriptors are placed into a certain collection that is available in the digital library. Similar to the first method, the new linkages possess the information about their creators (expert or author of the publication) that perform the processing of the text. The problem of ontological annotation of bibliographical references in scientific publications is considered in detail in [26].

Finally, *the third method* can be regarded as an *automated version* of the second method. In recent years, studies on analyzing the sentiment and tonality of the text are intensively carried out [27–31]. This line of investigation is called *sentiment analysis* or *opinion mining*. The aim of such an analysis is to determine the opinion of the author of the text on a point at issue. By applying the sentiment analysis to the neighborhood (context) of an intratextual reference (as it is done in [31]), we can mine the opinion of the author on the paper being cited. Thus, the ontological annotation of references, which is performed by the expert in the second method, can be automated. Moreover, descriptors of semantic linkages can be automatically generated based on such annotations. Their creator will be the author of the text containing the annotated references. Of course, such a “diagnostics” of the linkage semantics is limited only by classes of estimation references.

5. IMPLEMENTATION OF THE PROPOSED APPROACH

The Socionet system is used as an environment for implementing the proposed approach. To date, facilities for declaration and support of linkage taxonomies were implemented; the taxonomies are represented in the form of sets of controlled vocabularies for linkage classes. Mechanisms are implemented that make it possible to create, store, modify, remove, and view semantic linkages, as well as to compose collections of linkages. Moreover, a number of scientometric services were implemented that generate new scientometric indices based on the semantics of linkages. The context-based visualization of semantic linkages is provided for single information objects with the filtration based on linkage classes; the visual semantic navigation through the content of the digital library was implemented. Let us discuss in detail basic functions of the system.

Specification of the linkage semantics. To specify the semantics of linkages, a *hybrid ontology* is used, which is an extended combination of certain fragments of CiTO [32, 33], DoCo [34], SWAN [6], SKOS [9], and CERIF [5, 24]. Based on the hybrid ontology, a two-level taxonomy of classes of semantic linkages was constructed. This taxonomy is a set of controlled vocabularies, each corresponding to a certain high-level class of the class hierarchy; values contained in the vocabularies correspond to subclasses of these classes. The basic set of controlled vocabularies covers linkages of scientific inference; linkages of using scientific results; linkages between components of a scientific product, as well as linkages between its versions or copies; estimation-based linkages; hierarchical and association linkages between publications; and linkages of “person–person,” “person–organization,” and “person–publication”. The taxonomy of semantic linkages and its controlled vocabularies are discussed in more detail in [18]. The creation and update of controlled vocabularies is supported. The users can extend taxonomies of semantic linkages by adding classes to existing vocabularies and by creating new vocabularies under supervision of the system administrator.

It is important that the created taxonomies make it possible to classify linkages not only between scientific texts: linkages can be created between scientific publications, their components. They can classify also linkages between sets of scientific data, organizations, scientists (creators of information objects and users), and other types of information objects available in the system.

Creation of semantic linkages. Sets of semantic linkages are represented in the system in the form of collections of information objects of a special type. In addition to such an “autonomous” representation, the embedded representation of semantic linkages is supported. To date, only “manual” creation of linkages by the expert user is implemented (see Section 4). All functional capabilities of managing collections of any data types available in the system can be used for collections of autonomous linkages.

From the system’s standpoint, information objects can be *internal* (included into the content) or *external*. The external objects were not included into the content of the system. They should be available in Web by their uniform resource identifiers (URIs). The permissibility of external information objects and publications presented in the system by their bibliographical descriptions enables a wider coverage of the open digital scientific information space.

Metadata of linkage collections are an integral part of the metadata repository of the Socionet system based on the *open archive* (OAI) technology [22]. This component of the repository can be constructed as an independent metadata repository to organize the set of linkage collections as an independent open archive. Such an archive can be integrated with similar linkage archives from other libraries using the OAI-PMH protocol [35]. Thus, global semantic linkage repositories can be constructed that involve more information for scientometric researches. If the archives being integrated are not based on the common standard linkage taxonomy, then the problem of taxonomy heterogeneity must be solved.

When creating a linkage in the Socionet system, its descriptor is created, which includes the following metadata: a unique identifier of the linkage, types and identifiers of its source and target objects (or URI for an external target object), the class of taxonomy (description of the linkage semantics) to which the linkage belongs, creation date, creator identifier, and comment (if any) [19, 25]. Thus, if the new linkage is a review on a certain publication, then this review can be represented in the text form as a value of the *comment* attribute in addition to the brief opinion of the user in the form of the specified class of the linkage.

Depending on the types of information objects being linked, the linkage can belong only to a particular class of the taxonomy that is allowed for a given pair of the object types. To specify the semantics of the new linkage, its creator is granted with the access to the controlled vocabularies allowed for a given pair of the information objects types being linked.

For the given pair of objects, several linkages can be created. The same user cannot create several linkages of the same class, but can create several linkages of different classes. Different users can create several linkages of the same class.

Registration of users. To create semantic linkages and information objects, the user must be registered in the Socionet system. Upon registering, a *profile* of the user is created that is a set of metadata containing the e-mail address together with the other attributes that describe the user. These data are used for communication between the users (including the users that create information objects in the system), as well as for sending system notification messages to the user.

Notification service. The Socionet system provides a mechanism for *monitoring* the state of separate semantic linkages and semantic structure of the content as a whole; whenever necessary, the system informs the users about changes via e-mail. Moreover, if any user creates a linkage involving a particular publication, the notification is sent to the user that is the author of this publication. The system informs the users in a number of other cases, for example, when semantically inconsistent linkages created by different users are established between information objects (this function is not yet implemented). Notifications serve to stimulate the response reaction of the users that can be expressed by creating a corresponding semantic linkage.

Generation of scientometric indices. The Socionet system allows one to generate a number of new scientometric indices based on the current state of the semantic structure of the content [3, 4]. To date, only the embedded (base) taxonomy of semantic linkages can be used for this purpose.

The nature of the scientometric information generated on request is rather diverse. When calling a catalog of any collection of information objects, the total number of input and output linkages (i.e., linkages where a given object is the target and source, respectively) is generated for each information object. Thus, for a particular object selected from the catalog, we can obtain the detailed statistics (and some its components) of the number of input and output semantic linkages for each class of the taxonomy. Of course, the meaningful interpretation of the statistics obtained depends on the classes of linkages. So, we can find out how many positive and negative opinions on a given publication are expressed, in what number of works the methods or scientific data presented in the publication are used, whether there are cases of plagiarism of the publication and how many such works are presented in the system, etc.

We are intended to develop scientometric functions, as well as functions of analyzing the graph of the semantic linkage structure. So, several possible directions can be distinguished. First and foremost, statistic can be generalized over the entire set of information objects of a given type or over all linkages of the high-level class of the taxonomy (over a given vocabulary as a whole). For example, we can obtain the statistics of opinions on the entire set of monographs from a particular collection or the total number of estimate opinions on a given work. Since linkage descriptors contain their own creation date, requests of statistics on certain time intervals can be made, as well as requests of dynamic statistics (time series of certain statistical indicators). We can obtain the list of particular information objects related to a given object that is a source or target in linkages of particular classes. Meaningful interpretations of the result obtained will differ depending on the vocabulary used or its particular class of linkages. For example, we can find out the results of which publications are the basis for a given publication or, on the contrary, in which publications the results are obtained that are based on a given work. Moreover, not only direct linkages but also transitive ones can be taken into account.

Using chained linkages of the form “author–publication” + “publication–publication”, we can obtain the number of publications in which the negative or positive opinion is expressed on publications of a certain author, as well as obtain the list of such publications. Using extended chains of the form “organization–researcher (author)” + “author–publication” + “publication–publication”, we obtain similar information for an organization; this information is aggregated over all its researchers, i.e., authors of the works available in the system.

The analysis of the complete graph of linkages is of interest. Here, many problems can be solved that are related to analyzing the graph topology, to calculating subgraphs with given properties, and to visualizing the subgraphs. So, we can visualize the layer extracted from the multilayer structure of semantic linkages that corresponds to linkages of a particular class that points, for example, to the fact of using a certain publication as a basis for other publications. A subgraph can be constructed that is formed by linkages belonging to the class of scientific results development; this subgraph issue from a certain universally-recognized fundamental publication in a particular scientific domain that is available in the digital library. This subgraph characterizes the place of other publications in the development of the given scientific domain.

Of course, such information will be of interest if there is a representative set of publications in the system.

Visual semantic navigation through the linkage structure. The visualization of linkages between an information object selected by the user and other objects in the content of the system may be realized. Using filtration, the user can make visible only linkages of classes that are of interest to him. Based on the visible linkages in the current analyzed node of the linkage graph, the user can select the next node on his navigation way.

6. CONCLUSIONS

An approach to semantic structuring of the content of a scientific digital library by using a dedicated ontology of semantic linkages is proposed. Basic elements of the proposed approach were implemented in the Socionet system. The advantage of this approach is that it considerably increases the value of information resources of the digital library. For the scientific community, the approach offers the prospect of involvement in certain new forms of scientific activities in the environment of a kind of a social network and provides a new data source for nontraditional multiaspect scientometric researches. Some ideas used in the proposed approach were published almost simultaneously in foreign literature and in works of the authors of this paper. However, our approach seems more developed in terms of its functional capabilities. Largely, these advantages are ensured by the use of the autonomous description of semantic linkages, the advanced ontology of linkages, and the technology of social networks. The development of the proposed approach is under way.

REFERENCES

1. Kogalovsky, M.R. and Parinov, S.I., Semantic structuring of the content of scientific digital libraries based on ontologies, in *Sovremennye tekhnologii integratcii informatsionnykh resursov: sbornik nauchnykh trudov* (Modern Technologies for Integration of Information Recourses: Collection of Papers), St. Petersburg: Boris Yeltsin Presidential Library, 2011, pp. 26–45.

2. Dix, A., Levialdi, S., and Malizia, A., Semantic halo for collaboration tagging systems, in *Social Navigation and Community-Based Adaptation Technologies Workshop*, 2006.

3. Parinov, S., Kogalovsky, M., and Lyapunov, V., A challenge of research outputs in GL circuit: From open access to open use, *15th Int. Conf. on Grey Literature, The Grey audit: A Field Assessment in Grey Literature*, Bratislava, 2013. http://www.textrelease.com/gl15callforpapers.html.

4. Parinov, S. and Kogalovsky, M., Semantic linkages in research information systems as a new data source for scientometric studies, *Scientometrics*, 2014, vol. 98, no. 2, pp. 927–943.

5. CERIF 1.3 semantics: Research vocabulary, CERIF Task Group, euroCRIS. http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3 Semantics.pdf.

6. Semantic Web applications in neuromedicine (SWAN) ontology, W3C Interest Group Note. http://www. w3.org/TR/2009/NOTE-hcls-swan-20091020/.

7. Shotton, D., Introduction the semantic publishing and referencing (SPAR) ontologies. http://opencitations. wordpress.com/2010/10/14/introducing-the-semantic-publishing-and-referencing-spar-ontologies/.

8. Shotton, D. and Peroni, S., Semantic annotation of publication entities using the SPAR (Semantic Publishing and Referencing) ontologies. http://imageweb.zoo.ox.ac.uk/pub/2010/Publications/Shotton&Peroni_semantic_annotation_of_publication_entities.pdf.

9. SKOS Simple Knowledge Organization System reference, W3C recommendation. http://www.w3.org/TR/skos-reference/.

10. Kogalovsky, M.R. and Parinov, S.I., Use of citation linkages for scientometric measurements in the Socionet system. http://socionet.ru/publication.xml?h=repec:rus:rssalc:web-32.

11. Berners-Lee, T., Linked Data, World Wide Web Consortium. http://www.w3.org/DesignIssues/Linked-Data.html.

12. Jorg, B., Ruiz-Rube, I., Sicilia, M., Dvorak, J., Jeffery, K., Hoellrigl, T., et al., Connecting closed world research information systems through the linked open data web, *Int. J. Software Engineering Knowledge Engineering*, 2012, vol. 22, no. 3, pp. 345–364.

13. Open Annotation Collaboration. http://www.openannotation.org/.

14. Open Annotation Collaboration Group. http://www. w3.org/community/openannotation.

15. Groth, P., Gibson, A., and Velterop, J., The anatomy of a nano-publication. http://iospress.metapress.com/ content/ftkh21q50t521wm2/.

16. Shotton, D., Use of CiTO in CiteULike. http://opencitations.wordpress.com/2010/10/21/use-of-cito-in-citeulike/.

17. Kogalovsky, M.R. and Parinov, S.I., Classification and use of semantic linkages between information objects in scientific digital libraries, *Informatika i ee primeneniya*, 2012, vol. 6, no. 3, pp. 32–42.

18. Parinov, S.I. and Kogalovsky, M.R., Technology for semantic structuring the content of scientific digital libraries, *Trudy XIII Vserossiiskoi nauchnoi konferencii "Elektronnye biblioteki: perspektivnye metody i tekhnologii, elektronnye kollektcii"* (Proc. XIII All-Russian Sci. Conf. "Digital Libraries:

Advanced Methods and Technologies, Digital Collections”), Voronezh, 2011, pp. 94–103.

19. Parinov, S., Open repository of semantic linkages, in *Proc. 11th Int. Conf. on Current Research Information Systems e-Infrastructure for Research and Innovations (CRIS 2012)*, Prague, 2012. http://socionet.ru/publication.xml?h=repec: rus:mqijxk:29.

20. Parinov, S., Towards a semantic segment of a research e-infrastructure: necessary information objects, tools and services, *Metadata and Semantics Research, Communications in Computer and Information Science*, Dodero, J.M., Palomo-Duarte, M., and Karampiperis, P., Eds., Springer, 2012, vol. 343, pp. 133–145. http://socionet.ru/pub.xml?h= RePEc:rus:mqijxk:30.

21. Parinov, S.I., Lyapunov, V.M., and Puzyrev, R.L., Socionet system as a platform for developing scientific information resources and online services. http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2003/part1/PLP.

22. Open Archives Initiative. http://www.openarchives.org/.

23. F1000Research, The first Open Science journal for life scientists. http://f1000research.com/.

24. CERIF 1.3 Full Data Model (FDM): Introduction and Specification, euroCRIS. http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_FDM.pdf.

25. Parinov, S., Towards a semantic segment of a research e-infrastructure: necessary information objects, tools and services, *Metadata and Semantics Research, Communications in Computer and Information Science*, Dodero, J.M., Palomo-Duarte, M., and Karampiperis, P., Eds., Springer, 2012, vol. 343, pp. 133–145. http://socionet.ru/pub.xml?h= RePEc:rus:mqijxk:30.

26. Kogalovsky, M.R., Ontological annotation of references in scientific publications and its use in scientometrics, *Informatsionnye resursy Rossii*, 2013, no. 5, pp. 5–13.

27. Klekovkina, M.V. and Kotel’nikov, E.V., Method for automated classification of texts by tonality based on the vocabulary of emotional lexis, *Trudy 14-i Vserossiiskoi nauchnoi konferencii “Elektronnye biblioteki: perspektivnye metody i tekhnologii, elektronnye kollekcii”* (Proc. 14th All-Russian Sci. Conf. “Digital Libraries: Advanced Methods and Technologies, Digital Collections”), Pereslavl-Zalessky, 2012, pp. 118-123.

28. Feldman, R., Techniques and applications for sentiment analysis, *Communications of the ACM*, 2013, vol. 56, no. 4, pp. 82–89.

29. Galassini, C., Malizia, A., and Bellucci, A., An approach for developing intelligent systems for sentiment analysis over social networks, *Intelligent Systems and Control /742: Computational Bioscience*, Whidborne, J.F., Willis, P., and Montana, G., Eds., Cambridge, 2011.

30. Pang, B. and Lee, L., Opinion mining and sentiment analysis, *Foundations and Trends in Information Retrieval*, 2008, vol. 2, nos. 1-2., pp. 1–135. http://dl.acm.org/citation.cfm?id= 1454712.

31. Small, H., Interpreting maps of science using citation context sentiments: A preliminary investigation, *Scientometrics*, 2011, vol. 87, no. 2, pp. 373–388.

32. Shotton, D., CiTO: The citation typing ontology. http://www.jbiomedsem.com/content/1/S1/S6.

33. Shotton, D. and Peroni, S., CiTO: The citation typing ontology, version 2.0. http://purl.org/spar/cito/.

34. Shotton, D. and Peroni, S., DoCO: The document components ontology. http://speroni.web.cs.unibo.it/cgi-bin/lode/req.py?req=http:/purl.org/spar/doco.

35. The Open Archives Initiative protocol for metadata harvesting, version 2.0. http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm.

Translated by Yu. Kornienko