Bulletin of Electrical Engineering and Informatics
Vol. 11, No. 5, October 2022, pp. 2968~2977
ISSN: 2302-9285, DOI: 10.11591/eei.v11i5.3075

Development of methods and technologies for creating intelligent scientific and educational internet resources

Ardak G. Batyrkhanov[1,7], Zhanna B. Sadirmekova[2], Madina A. Sambetbayeva[1,3], Assel N. Nurgulzhanova[4], Zhuldyz S. Ismagulova[5], Algerim S. Yerimbetova[3,6]

[1]Department of Information Systems, Faculty of Information Technology, L. N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan
[2]Department of Information Systems, Faculty of Information Technology, M. H. Dulaty Taraz Regional University, Taraz, Kazakhstan
[3]Institute of Information and Computational Technologies, Committee of Science of the Ministry of Education and Science of the Republic of Kazakhstan, Almaty, Republic of Kazakhstan
[4]Institute of Heat Power Engineering and Control Systems, Almaty University of Power Engineering and Telecommunications named G. Daukeyev, Almaty, Kazakhstan
[5]Department of Computer Engineering, Faculty of Engineering, Khoja AkhmetYassawi International Kasakh-Turkish University, Turkestan, Kazakhstan
[6]Department of Software Engineering, Institute of Automation and Information Technologies, Almaty, Republic of Kazakhstan
[7]Atyrau State University named after H. Dosmukhamedov, Atyrau, Kazakhstan

Article Info

Article history:

Received May 24, 2021
Revised May 24, 2022
Accepted Aug 10, 2022

Keywords:

Conceptual model
Dublin core
Information system
Ontology
Research and education
Semantic web

ABSTRACT

The purpose of this work is to develop methods, technologies and tools for creating and maintaining intelligent scientific and educational internet resources (ISEIR) based on a service-oriented approach and Semantic Web technologies. The main purpose of ISEIR is to provide meaningful access to scientific and educational information resources of a given field of knowledge and integrated information processing services. According to the preliminary concept, an intelligent scientific and educational internet resource will be an information system accessible via the internet which provides ontology-based systematization and integration of scientific knowledge, data and information resources into a single information space together with a meaningful effective access to them as well as supporting their use in solving various scientific and educational tasks. ISEIR is equipped with an ergonomic web-based user interface and special editors designed to manage the knowledge integrated into it. The proposed approach to the construction of intelligent scientific and educational internet resources is the basis for the developed technology in creating and maintaining information environments for distributed learning.

Corresponding Author:

Zhanna B. Sadirmekova
Department of Information Systems, Faculty of Information Technology
M. H. DulatyTaraz Regional University
60 Tole bi Street, Taraz, 080000, Kazakhstan
Email: janna_1988@mail.ru

1. INTRODUCTION

Modern science and education need to concentrate and generalize the information on various knowledge branches. This task is complicated by versatile and dispersed of the scientific and educational information resources and can be solved by bringing all knowledge into a single information space. Support for the logical integrity of integrated resources will be provided on the basis of an ontology that provides a coherent and consistent view of the area of knowledge. To ensure such access, the concept and architecture of an intelligent information system managed by ontology will be developed.

Become more and more popular the approach to provide effective meaningful access to information resources of a certain subject and means of their intellectual processing through the creation of specialized information systems, in particular internet portals. However, the concept of such resource has not yet been developed and there is no evidence of management technology that supports the full cycle of creating and operating it.There are researches on questions of ontological modeling [1]–[8] but there are no convincing examples of ontologies and scientific and educational software services joint using for the purposes indicated by the authors.

Among similar foreign developments, the European libraries and electronic resources (EULER) project in mathematical sciences was implemented with the financial support of the European Union. Its task is to provide integrated access to library catalogues and mathematical information on the internet. And its approach is based on the Z39.50 protocol [9] and the general method of describing resources in the Dublin Core format [10]. There are also various online resources designed to support humanities research (like the English language catalog LINGUISTLINGUIST list). It is created for the exchange of knowledge between linguists and it contains information about publications, personalities, scientific institutions, grants, competitions, projects, scientific foundations, conferences, and seminars on linguistic topics. The other one is the information portal "Language Technology World" created at the German Research Center for Artificial Intelligence (http://www.lt-world.org/). Its thematic sections include in addition to information about scientific events, projects, organizations, and individuals contain more detailed information about linguistic technologies, products, and information systems in the field of natural language processing.

The projects and resources described above are essentially structured and they have annotated directories of links to internetresources, the constituent elements of which are practically unrelated, which makes it difficult to find the necessary information. The main difference in our approach is the use of the description of the modeled area of knowledge and means of intellectual processing of information resources relevant to it in the form of ontology that allows to represent knowledge and data on the subject INAIR in the form of a network of knowledge and data (semantic web) and provide the users with easy navigation and meaningful access to their accumulated knowledge and data and its processing.

Another close approach is the semantic web concept developed by the world wide web consortium (W3C). It assumes that any document hosted on the network has an associated set of metadata (semantic annotation). To describe metadata, we use the W3C standards–resource definition framework (RDF) [1] and web ontology language (OWL) [11]. This allows us both to describe the structural properties of documents and represent their meaning in terms of domain ontologies (defined in OWL). This kind of metadata in documents makes it easier to integrate them and it favours various software applications and communities to use them. The ideology and tools of this approach have been used in the development of many applications. But they are still in the process of development and a unified methodology for creating internetresources aimed at supporting scientific research has not been created. Our approach integrates the most important components of semantic web technology, in particular, the use of ontology to represent the semantics of information resources and support their intellectual analysis. However, the considered approach doesn't offer a complete concept of intellectual scientific internet resources and a methodology for their collective construction. This approach doesn't provide methods for intelligent processing of integrated resources and it allows much less meaningful access to them.

2. THE PROPOSED CONCEPTUAL MODEL OF ISEIR

There are a large number of approaches to building intelligent internet resources based on using ontology as a conceptual model [12]–[14]. A formalized model is needed to represent the knowledge of ISEIR which provides flexible means of describing the concepts of the problem and subject areas together and variety semantic relationships between them [15]. An important requirement is the ability to set restrictions on the values of properties of objects in the domain and to describe the semantics of relations in the form of axioms [16]. Following type metaontology is proposed as a conceptual model of knowledge representation:

$$O=(K, B, T, D, S, P, A),$$

where K is a finite nonempty set of classes describing the concepts of some subject or problem domain; B is a nite set of binary relations defined on classes (concepts); T is the set of standard types; D is a set of domains (sets of values of standard type string); S is finite set of attributes describing properties of concepts K and B relationships; P is a set of restrictions on the values of attributes notions and relations; a set of axioms that define the semantics of classes and relations of the ontology.

Three types of relations are distinguished in the ontology: *BT* is asymmetric, transitive, non-reflexive binary inheritance relationon the basis of which hierarchies of concepts *K* can be built; *BP* is binary transitive inclusion relation ("part–whole") and *BA* is a finite set of associative relations [17].

The intelligent scientific and educational internet resources (ISEIR) ontology is based on the above-mentioned meta-ontology. To simplify the system configuration for the selected area of knowledge and its further maintenance, the basic ontologies that are independent of the IS domain are highlighted as well as a subject ontology that describes a specific area of knowledge (Figure 1). As the base ontology there were selected two of them. The first one describes the problem area of the system. It does not depend on the subject area. It is a top-level ontology and includes classes of concepts such as*pers on, organization, scientific activity, scientific events, publication, geographical locationand a collection of conference materials.* Such concepts are used to describe participants in ontology, organization of educational work, events (seminars, conferences), joint projectsand various types of information resources.

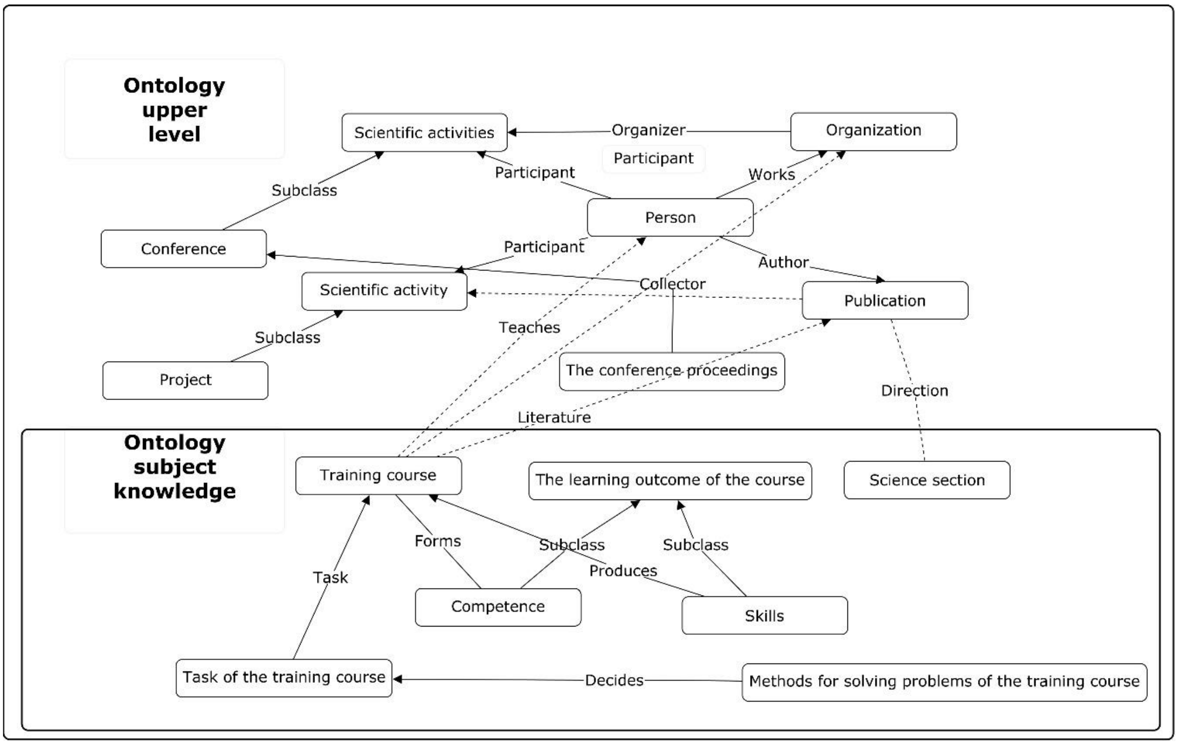

Figure 1. The basic ontology ISEIR

The concepts of basic ontologies are linked by associative relations the choice of which is made not only based on the completeness of the representation of the problem and subject areas of IS but also considering the ease of navigation through educational content and information search. The ontology built in this way allows you to describe the subject and problem area of IS, sets structures for representing real objects (including information resources) and the relationships between them. The semantics of relations between information objects is determined by the relations defined by the corresponding ontology concepts. The totality of such objects and their connections forms the information content or content of the ISEIR.

The creation of intelligent scientific and educational internet resources should be accompanied by the development of digital repositories to ensure long-term storage of information resources (conference materials'collections, full publications' texts, programs of training courses). The international organization for standardization (ISO) has proposed the ISO-14721:2012 open archive information system (OAIS) standard for organizing a long-term temporary storage of information resources [18]. The OAIS standard reference model is conceptual and used by organizations in order to develop metadata sets and organize the repositories. Based on this model, the concept of an "institutional repository" was created as a system for long-term storage, accumulation of information and providing reliable access to digital objects that are the result of intellectual activity of a scientific or educational institution.

Key features of the institutional repository include: i) the ability to organize a single access to information resources for the world community (including full-text indexing by world search engines); ii) unified access to metadata over standard protocols (support for interoperability); iii) saving other resourcesincluding unpublished ones (dissertations, preprints and technical reports, software, multimedia); iv) providing differentiated access to heterogeneous digital objects (publications, images); v) a system for long-term storage, accumulation and secure access in electronic form of intellectual products of a scientific or educational institution.

Institutional repositories are related to digital interoperability issues and the open archives initiative (OAI) and they partially correlate with the concept of an electronic library. They perform the functions of collecting, storing, classifying, cataloging, and providing access to digital content. The process of integrating a digital repository into an IS is based on a metadata aggregation and distribution model. The application of this model is fixed in the OAI protocol for metadata harvesting (hereinafter-OAI-PMH) [19].

2.1. Metadata of information resource

By introducing formal descriptions of domain concepts in the form of object classes and relationships between them, the system ontology sets structures for representing real relationships between elements. So, data becomes a set of different types of information objects and linkswhich form the information content of ISEIR (Figure 2). An information object (IO) is a structured set of data that represents a description of some object of the selected field of knowledge or relevant information resource. Each IO corresponds to an ontology class and it has a structure defined by this class. There may be connections between specific information objects whose semantics are determined by the relationships defined between the corresponding ontology classes. The information content of ISEIR (its content) includes both General knowledge (represented in the ontology) and specific knowledge about real objects and information resources (we call them data). Description of information resources is the most important component of ISEIR content.

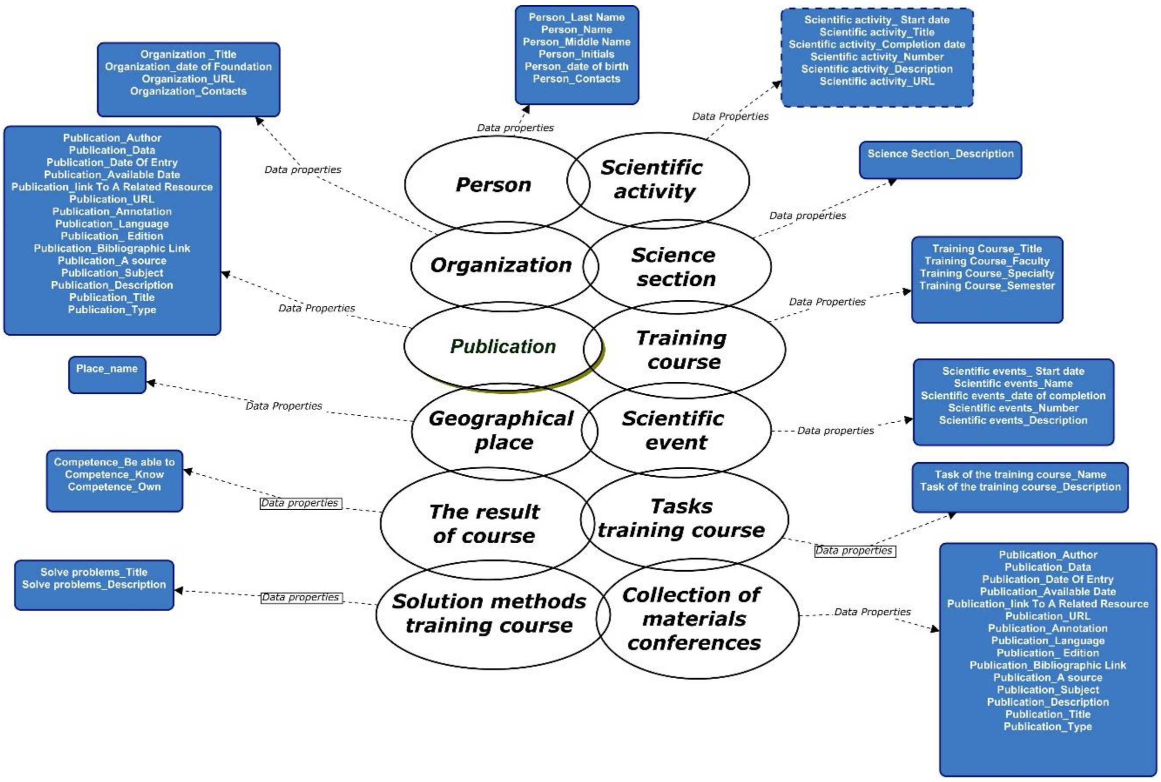

Figure 2. Metadata of information resource

The common european research information format (CERIF) standard is used to provide information about projects, a person or an organization. It is based on a data model that includes the project, organization and person entities, them relationships, and these entities attributes. The standard defines three levels of detail when describing resources [20]: i) full resource description-contains an extended set of attributes that allows you to describe various resource schemas; ii) set of attributes-designed for data exchange between different systems; iii) abbreviated set of attributes-used for meta description of resources.

In the ISEIR resource model, a project has the following properties: name, number, start and end dates. It includes the "participant" relationship between a project and the organization/person which has the «participation type» attribute. Interaction with other CERIF-based systems is quite easy. However, at the moment the standard is implemented in a limited version (for example, the concept of «project result» is not defined in ISEIR), so the project resource model will be systematically developed. The data model proposed by CERIF defines individuals and organizations as separate entities which are in good agreement with the ISEIR model and it greatly simplifies information exchange.

An organization resource has the following attributes: name, abbreviation, address, phone number, email address, type of organization, and business direction.

A *Person* resource has the following attributes: last and first names, patronymic, initials, list of WOS-publications, list of Scopus-publications, list of RSCI-publications, list of Mathnet.ru-publications. There is a "position" relationship between these resources types that have the attribute name, type, phone number, and email.

Dublin Core was chosen as the basis for the implementation of the remaining resources *(scientific events, publication, geographical location, conference proceedings, training course, competence, training course objectives, methods of solving problems, results of course development*). This choice is not accidental and is caused by the following advantages [21]: i) the set of basic semantic elements is compact but it allows you to set almost all needed attributes; ii) the semantics of each element in the standard can be refined with the help of qualifiers, both standard, known and understandable to everyone and specially designed to accurately specify the semantic meaning of a particular attribute when exchanging data within a small community; iii) the standard provides the possibility of using various semantic schemes and dictionaries; iv) it defines a mechanism for extracting information from a description by using non-standard namespace extensions; v) the standard is becoming more widespread in the world community.

The ISEIR publication data model allows you to set any basic element of the Dublin Core. It can be used qualifiers to specify the semantics of the basic elements and to facilitate the exchange of bibliographic information. But there is a serious obstacle to the interoperability of such subscheme of the ISEIR model with other systems is most of these systems consider individual publication properties like *«author»*, *«publisher»*, *«source»*, as normal text attributes while they are links to other entities (persons, organizations, publications). These models do not contradict Dublin Core but lead to a certain incompatibility with the ISEIR model and some ambiguity when integrating data into the system [22]. Each standard offers its own data model and often its own syntax for writing. The ISEIR approach is to use a single data model and syntax defined by RDF for metadata exchange. The semantics of attributes of certain resources is taken from the corresponding standard. If you can't find a suitable element in standard namespaces, you can create your own space by defining it using a URI and including elements with the required semantics in it [23].

2.2. The information content of iseir

Setting up ISEIR for the subject area and managing the system content is carried out using specialized editors (ontology and data editor), implemented as a web application and available to registered users-experts on the internet. The ontology description language and the ontology editor have to be transparent and easy to use. The knowledge representation language Semp-TAO was used as a prototype of the ontology description language [24]. The main structure for representing data in this language is a heterogeneous semantic network. The semantic network object can be any entity of the subject area identified by an expert or knowledge engineer. Each object is characterized by its own name and the values of the its attribute slots. Restrictions can be set on object slots which are logical expressions that link the values of object slots. Objects with the same properties are combined into classes. An inheritance relation is defined by classes that form a hierarchy. Their relationships are characterized by their possibility to have their own attributes to define the relationship between arguments:

$$R(Arg1, Arg\ 2, Matr),$$

where R is relation name, *Arg1, Arg 2* argument relations (classes), Matr is the set of attributes that describe additional properties of the relationship.

Mathematical properties (transitivity, symmetry and reflexivity) can be attributed to relationships. Ontologies are managed using the ontology editor. In order to ensure distributed ontology development, this editor supports a mechanism for delegating rights to experts at different levels. You can use the ontology editor to create, modify and delete any ontology elements (classes, relationships, domains) and define and modify concept hierarchies. For a more convenient representation of ISEIR information, the ontology editor includes tools for configuring knowledge and data visualization allowing us to set a template for objects visualization of this class and a template for links visualization to them for each ontology class. A class object visualization template defines the order in which all its attributes and related relationships are displayed. For clarity and meaningful representation of a reference to a specific class object, the visualization template can include both attributes of this class and attributes of classes associated with it. The attribute values included in the link template are used to build a text representation of the object reference (hyperlinks). In order to exchange ontologies with other information systems, and to integrate ontologies developed by other researchers into the ISEIR, a subsystem that performs two functions have been implemented and it is being debugged [25]. The first one is converting an ontology presented in the ISEIR format to an OWL representation. And the second is a translation of the ontology presented in OWL format to the internal ISEIR format.

3. RESULTS AND DISCUSSION

Meaningful access to systematized knowledge and information resources of an area of knowledge is provided using advanced navigation and search tools provided by ISEIR. The main scenario for the user to work with ISEIR consists of selecting objects of a certain class either directly using visualization tools or using the search engine, viewing similar objects, navigating through their associations and filtering their lists.

For the end user, data on the ISEIR is represented as a set of related information objects. All information about a particular object and its relationships is displayed as an HTML page, the format and content of which depends on the class of this object and the visualization template that is created for it. In this case, objects associated with this object are represented on its page as hyperlinks that allow you to go to their detailed description [26].

The list of objects is displayed as a page containing a set of links to these objects. For large arrays of objects, a composite page is formed that includes a list of pages with navigation elements based on ISEIR data which is the process of moving from one information object to another using the links set between them. For example, when viewing information about a specific grant, we can see the values of its attributes and its relationship to other objects. Using the links provided as navigation elements, you can view detailed information about both direct links and reverse links (about grant participants, publications describing this grant). When you click on a specific link of any information object, you can get a fairly large list of objects (for example, a list of all participants in a major project or conference). In this regard, a mechanism for filtering lists of information objects was introduced which is understood as a way to select a subset of IO from the list by imposing restrictions on it, i.e., filter tasks. The filter is a set of conditions that define acceptable values for IO attributes and requirements for the existence of links with other information objects. This method allows you, for example, to filter a set of project participants by age or scientific degree (conditions for an attribute) and by the research methods they use (conditions for a related object).

Search is based on ontology which allows to set the query in terms of the ISEIR domain. The main elements of such a query are the concepts and relations of the ontology as well as the restrictions that the required data must satisfy [27]. Acceptable limits for an attribute depend on the type of its values. For example, for attributes such as «number» (integer) and «date» (data), you can set an exact value or an acceptable range of values. To set restrictions on objects that are associated by associative relations with the desired object, the user can set conditions for the values of all attributes of related objects. Conditions can also be set for the attribute values of the corresponding relationships. For example, the query "Find recommended literary works of the type «article» in a training course between 1920 and 1990" will formally look like this:

Class "Training course":
Relationship "recommended literary works":
Class "Publication"
Attribute "Type» = article"
Attribute "Start date": (>= 1920) & (<=1990)
Attribute "Expiry date": (>= 1920) & (<=1990)

Currently, search queries are a set using a special graphical interface controlled by the ISEIR ontology. When the user selects a class of information objects to search for, a search form is automatically generated. This form allows you to set restrictions on the attribute values of objects of the selected class as well as on the attribute values of objects associated with this object by associative relations.

In order to fill in the content of ISEIR, information is collected from such sources as websites of organizations, associations, projects and conferences, knowledge portals, social scientific networks. Information about projects, organizations, individuals, and conferences is extracted from these sources, i.e., all the basic classes of the ontology of scientific activity, except for information about publications. Information about publications is extracted from the repository (Dspace) which was created by the authors.

Each of these classes has its own method for extracting information including a set of templates generated based on the ontology. To improve the completeness of information retrieval, the variability of these templates is increased by using alternative terms from the thesaurus.

The information retrieval module analyzes internet resources downloaded from links. Documents on the internet can be presented in various formats (HTML, DOC, PDF). The main format for presenting information on the internet is HTML. To extract publication metadata from repositories in batch mode, data is exported in extensible markup language (XML) format (Figure 3). The proposed methods for extracting information about projects, organizations, individuals and conferences are focused on working with HTML pages while information about Publications is focused on working with XML documents.

To facilitate analysis, the HTML page and XML document of the resource are represented in the DOM tree view in accordance with the (document object model (DOM) standard which regulates the way the document content is represented (in particular, HTML pages and XML documents) by a set of objects. Based on the corresponding template, the DOM tree of each page is analyzed and the information described by this template is extracted. A template is an XML document that specifies markers for objects, relationships and attributes in the ontology that indicate the location of this object, relationship or attribute. The templates for each type of extracted information specify handlers that implement algorithms for crawling and analyzing the corresponding fragments of internet sites.

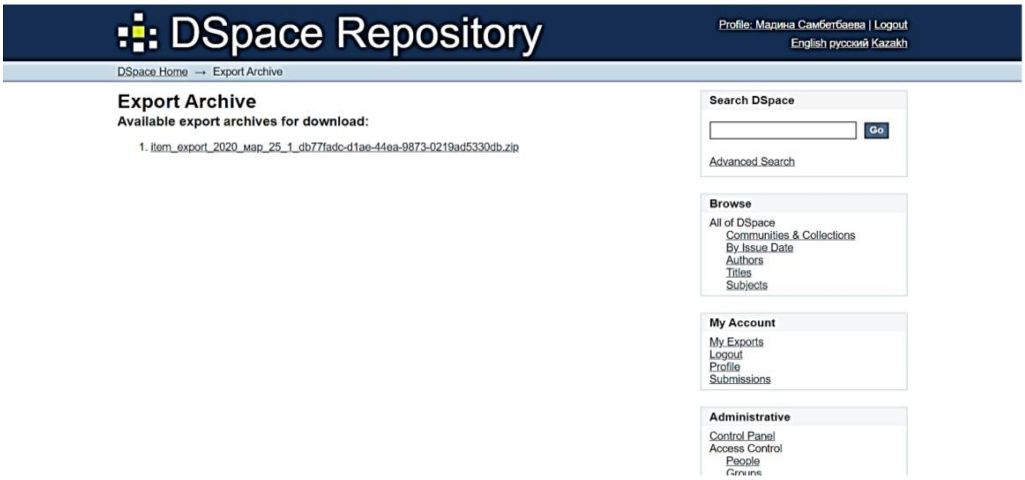

Figure 3. Exporting data from the repository

4. CONCLUSION

The information base of ISEIR consists of ontologies that, along with the traditional description of the subject area, contain a related description of the structure and typology of the corresponding data stores and network resources. In addition, the use of ontology as the basis of ISEIR, which is its declarative component, makes the system easily extensible and customizable so that it can integrate both new knowledge and new sections of information resources.

Ontology provides tools for effectively presenting a variety of needed information and it supports the systematization and integration of information resources together with meaningful access to them. Thanks to ontologies using of as an information model, ISEIR is not just another catalog of resources on a given topic but it is mainly a network of knowledge and data that allows us to maintain convenient navigation and meaningful search. Dividing the ISEIR ontology into subject-independent and subject-specific ontologies makes ISEIR customizable for any field of scientific knowledge. The possibility of declarative adjustment of the ontology during the operation of ISEIR will allow tracking the dynamics of the emergence of new knowledge and information resources on the topic and thus ensure support for its relevance and usefulness.

Based on the models and technologies listed above, a prototype of an information system for supporting scientific and educational activities has been built. ISEIR is positioned as an information system accessible via the internet, integrating and systematizing knowledge and information resources of a given subject area and providing meaningful effective access to them. It is planned to expand the information model in the near future. It is also planned to integrate several additional information systems.

ACKNOWLEDGEMENTS

This research has is funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP08857179).

REFERENCES

[1] T. Gruber, "Toward Principles for the Design of Ontologies Used for Knowledge Sharing," *International Journal Human-Computer Studies*, vol. 43, no. 5–6, pp. 907–928, 1993.
[2] N. Guarino, "Formal Ontology in Information Systems," in *Proceedings of FOIS'98*, 1998, pp. 3–15.
[3] T. Gavrilova and V. Khoroshevsky, *Knowledge bases of intelligent systems*. St. Petersburg, 2000.
[4] Y. Zagorulko, G. Zagorulko, and O. Borovikova, "Pattern-based methodology for building the ontologies of scientific subject domains," *Frontiers in Artificial Intelligence and Applications*, vol. 303, pp. 529–542, 2018, doi: 10.3233/978-1-61499-900-3-529.
[5] A. Kalinichenko Leonid, Y. K. Dmitry, A. Kovaleva Dana, and Y. Malkov Oleg, "Methods and tools for hypothesis-driven research support: A survey," *Informatics and Applications*, vol. 9, no. 1, pp. 28–54, Jan. 2015, doi: 10.14357/19922264150104.
[6] M. R. Kogalovskij, "Metadata, Their Properties, Functions, Classification and Presentation Tools," 2012.
[7] A. N. Bezdushnyi, A. B. Zhizhchenko, M. V. Kulagin, and V. A. Serebryakov, "Integrated information resource system of the Russian Academy of Sciences and a technology for developing digital libraries," *Programming and Computer Software*, vol. 26, no. 4, pp. 177–185, Jul. 2000, doi: 10.1007/BF02759466.
[8] A. F. Tuzovsky and V. Z. Yampolsky, "Integration of information using Semantic Web technologies," *Problems of Informatics*, vol. 2, pp. 51–58, 2011.
[9] NISO, *Information retrieval (Z39.50): application service definition and protocol specification*, vol. 2003. Maryland: National Information Standards Organization Bethesda, 2002.
[10] Z. B. Sadirmekova, J. A. Tusupov, M. A. Sambetbayeva, A. N. Nurgulzhanova, and K. Z. Doshtaev, "Development of the Methodology for Metadata Extraction from Documents in the Course of Information System Integration Using the Ontological Model of the Subject Field," *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 9, pp. 6231–6239, Sep. 2020, doi: 10.30534/ijeter/2020/214892020.
[11] G. Antoniou and F. van Harmelen, "Web Ontology Language: OWL," in *Handbook on Ontologies*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 67–92.
[12] V. A. Vasenin, S. A. Afonin, A. A. Zenzinov, K. V. Lunev, and D. A. Shachnev, "Mechanisms of «ISTINA» system for intelligent state analysis and progress stimulation for projects in the sphere of science and higher education," in *Proceedings of 21th Scientific Conference "Scientific Services & Internet – 2019,"* 2019, pp. 210–221, doi: 10.20948/abrau-2019-48.
[13] A. S. Adzhiyev, "Design and Implementation of Scientific Information Systems Based on the Object Data Model: Math-Net Information System," Technical Sciences, 2005.
[14] Z. B. Sadirmekova, O. L. Zhizhimov, D. A. Tusupov, and A. S. Madina, "Requirements for information systems to support scientific and educational activities," *CEUR Workshop Proceedings*, vol. 2569, pp. 44–47, 2020, doi: 10.25743/ict.2019.77.63.026.
[15] A. M. Fedotov *et al.*, "Classification model and morphological analysis in multilingual scientific and educational information systems," *Journal of Theoretical and Applied Information Technology*, vol. 86, no. 1, pp. 96–111, 2016.
[16] S. S. Kurmanbekovna, S. Z. Bakirbaevna, B. A. Gabitovich, S. M. Aralbaevna, and Y. A. Sembekovna, "Development of Technology to Support Large Information Storage and Organization of Reduced User Access to this Information," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, pp. 493–503, 2021, doi: 10.14569/IJACSA.2021.0120757.
[17] A. D. Solovjov and B. V. Dobrov, *Ontology and Thesaur: learning materials*. Moscow: KazSU and MSU, 2006.
[18] T. C. C. for S. D. S. CCSDS, *Reference Model for an Open Archival Information System (OAIS): Recommended Practice*, no. 650.0-M–2. 2012.
[19] S. Amin, "The Open Archives Initiative Protocol for Metadata Harvesting: An Introduction." pp. 1–13, 2003, [Online]. Available: http://www.openarchives.org/OAI/2.0/.
[20] B. Jörg and A. Asserson, "CERIF 2008 - 1 . 1 Full Data Model (FDM) Introduction and Specification," pp. 1–43, 2010.
[21] Z. B. Sadirmekova, D. A. Tusupov, and M. A. Sambetbaeva, "Metadata of the information system for supporting scientific and educational activities: definition and use," in *Proceedings of the International scientific and practical conference "Auezov readings-18: the spiritual heritage of the great Abai*, 2020, pp. 264–268.
[22] Z. Sadirmekova, J. Tussupov, M. Sambetbayeva, and Z. H. Altynbekova, "Data integration methods in information systems," *Herald of The Kazakh-British Technical University*, vol. 17, no. 3, pp. 165–172, 2020.
[23] Z. B. Sadirmekova, M. A. Sambetbayeva, J. A. Tussupov, and Z. T. Altynbekova, "Building an Information System To Support Scientific and Educational Activities Based on the Ontological Model of the Subject Area," *Series Physico-Mathematical*, vol. 6, no. 334, pp. 45–52, Dec. 2020, doi: 10.32014/2020.2518-1726.96.
[24] Z. B. Sadirmekova, J. A. Tussupov, M. A. Sambetbayeva, and Z. T. Altynbekova, "Development of Integrated Information Systems to Support Scientific Activity," *SIST 2021 - 2021 IEEE International Conference on Smart Information Systems and Technologies*, 2021, doi: 10.1109/SIST50301.2021.9465964.
[25] E. Abad-Segura, M. D. González-Zamar, J. C. Infante-Moro, and G. R. García, "Sustainable management of digital transformation in higher education: Global research trends," *Sustainability (Switzerland)*, vol. 12, no. 5, 2020, doi: 10.3390/su12052107.
[26] L. Guo, D. Wang, F. Gu, Y. Li, Y. Wang, and R. Zhou, "Evolution and trends in intelligent tutoring systems research: a multidisciplinary and scientometric view," *Asia Pacific Education Review*, vol. 22, no. 3, pp. 441–461, Sep. 2021, doi: 10.1007/s12564-021-09697-7.
[27] V. Palanisamy and R. Thirunavukarasu, "Implications of big data analytics in developing healthcare frameworks – A review," *Journal of King Saud University - Computer and Information Sciences*, vol. 31, no. 4, pp. 415–425, 2019, doi: 10.1016/j.jksuci.2017.12.007.

BIOGRAPHIES OF AUTHORS

Ardak G. Batyrkhanov is a Master of Engineering Sciences, Senior Lecturer at the Department of Information Systems, Faculty of Information Technology, L.N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan and at H. Dosmukhamedov Atyrau State University, Atyrau, Kazakhstan. Scientific interests: distributed information system, thesaurus, information retrieval, digital library, ontology. Has more than 30 publications, including: 1 academic book; 8 articles in SCOPUS base journals, 3 articles in Web of Science base, 6 articles in the journals of the COXON list and VAK RF. Scopus H-index - 2, Web of Science H-index - 1. He can be contacted at email: batyrkhanov.a@rambler.ru.

Zhanna B. Sadirmekova (4.06.1988) Doctor of Philosophy (PhD) in 6D070300 - Information Systems. She successfully defended her thesis on "Development of technology for integration of information systems to support scientific and educational activities based on metadata and ontological model of the subject area. Sadirmekova J.B. is currently pursuing postgraduate studies at the Federal State Autonomous Educational Institution of Higher Education "Novosibirsk National Research State University" in the field of 03.06.01 - Physics and Astronomy. Research interests: information technology in education, artificial intelligence, search, digital library, ontology. Has more than 30 publications, including: 4 articles in SCOPUS base journals, 2 articles in Web of Science base, 7 articles in COXON list journals; H-index - 1. She can be contacted at email: janna_1988@mail.ru.

Madina A. Sambetbayeva Doctor of Philosophy (Ph.D.) in 6D070300 - Information Systems, Associate Professor. Her research interests are related to information systems, information retrieval systems, information retrieval processes, decision support systems, linguistic support of information systems. Currently, she is working on research and development of the fundamentals of the technology for creating models and software for building distributed information systems to support scientific and educational activities, taking into account the morphology of the Kazakh language, identified in the form of electronic libraries that are consistent with international standards and trends in the development of national and international information infrastructure. Project Manager of the grant financing project No.AP05132046 “Development of the fundamentals of the technology for creating models and software for building distributed information systems to support scientific and educational activities (taking into account the morphology of the Kazakh language)” for 2018-2020. No.211 dated 03/19/2018 for 2018-2020. She is also involved in interstate projects, working together with scientists from Novosibirsk State University and the Institute of Computational Technologies SB RAS.She has over 50 publications, of which: 1 educational-methodical manual; 11 articles in SCOPUS database magazines, 16 articles in the journals of the list of Higher Attestation Commission of the Republic of Kazakhstan and the Russian Federation, 5 certificates of official registration of computer programs used in pedagogical and scientific practice. H-index - 3. (1 wage-rateasLRA). She can be contacted at email: madina_jgtu@mail.ru.

Assel N. Nurgulzhanova born in 1979, higher education, Master, Candidate of Technical Sciences, Associate Professor (associate professor). In 2000, she graduated from Al-Farabi Kazakh National State University with a degree in Applied Mathematics, awarded the academic degree of Bachelor of Applied Mathematics. In 2002, she graduated from the Master's degree of the Kazakh National State University named after Al-Farabi in the specialty "Applied Mathematics", awarded the academic degree of Master of Applied Mathematics. In 2009, she defended her PhD thesis on "Improving the management of dynamic systems with distributed parameters in electric transport", specialty 05.22.08 - Transportation process management. In 2020, he was awarded the title of associate professor (associate professor) in the specialty 05.13.00 – Computer Science, computer engineering and management. She worked at KazATC named after M.Tynyshpaeva from 2002 to 2020 as an assistant, teacher, senior lecturer, associate professor, associate professor, head of the department. Since 2020, he has been working at the AUEC named after G.Daukeev as an associate professor of the Department of Automation and Management.The total scientific and pedagogical experience is 20 years. She is constantly working to improve her qualifications, has taken advanced training courses in the field of IT, has certificates, is the author of 8 teaching aids, 3 monographs, more than 71 scientific and methodological works, including in publications with a non-zero impact factor in the Web of Science database, Thomson Reuters, Scopus -4. Nurgulzhanova A.N. in

2019 was awarded the honorary title - “Ulagatty ustaz”, has letters of thanks from the President-Rector of KazATC named after M.Tynyshpaeva (2014), the party "Nur-Otan" (2019). Is the holder of the title "Best University teacher" of 2019. She can be contacted at email: nurgulzhanova@mail.ru.

Zhuldyz S. Ismagulova received a BA degree in Mathematics from the Kazakh National Pedagogical University named after Abai, Almaty, Kazakhstan, in 1989. Between 1993 and 2001 she worked as senior lecturer of “Information systems” department at the Kyzylorda State Univeristy, named after Korkyt Ata, Kyzylorda, Kazakhstan. In 2001 she continued her career in “Computer Engineering and information systems” department at the Kazakh Academy of Transport and Communications named after M. Tynyshpaev. In 2010 she defended her dissertation for the degree of candidate of technical sciences on the topic "Development and research of methods for distributing resources on network models under uncertainty", specialty 05.13.01 - System analysis, management and information processing”. 2010-2021 - Associate Professor of the Department of IT Engineering at the Kazakh Academy of Transport and Communications named after M. Tynyshpaev. Since September 2021, she has been an Associate professor at the Department of Computer Engineering at the International Kazakh-Turkish University named after Khoja Akhmet Yassawi. During the period of her scientific activity, she wrote about 45 scientific publications, she also took part in international scientific and practical conferences. She is the author of 3 monographs. She can be contacted at email: zhulduz66_66@mail.ru.

Algerim S. Yerimbetova (1988) – PhD, Candidate of Technical Sciences, Associate Professor. Currently, she is engaged in the research and creation of linguistic algorithms for semantic search and retrieval of information using Link grammar, the Link Grammar Parser software package based on it, and mathematical and logical methods. In this case, in the end, prestigious data structures are finite models (in the sense of model theory, such as part of mathematical logic) and partial assignments between them, ensuring the truth of certain formulas. She is the author of more than 70 articles on computational linguistics and the development of information systems: 1 textbook, 2 monographs, 19 articles in the journal SCOPUS, 14 articles in the journals of the CQAES list; 1 patent, 4 certificates of entry in the State Register of Copyright Rights. She can be contacted at email: aigerian@mail.ru.