Methodical Approach and Tools for Creating and Modifying Ontology-based Information System

Elena Khayrullina 1,2

1 Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences
2 Irkutsk National Research Technical University
Irkutsk, Russia
ksevgelena@istu.edu

Abstract—The article discusses the methodological approach and tools for creating and modifying information systems based on ontologies in the field of energy. To create an information system, it is necessary to create two classes of domain ontologies, in which entities will be indicated, connections between entities are established, and the degree of influence of some ontology objects on others is indicated.

Keywords—semantic models, ontologies, information system, web-system, critical infrastructures

I. INTRODUCTION

Currently, information systems surround us everywhere - they can be found in almost any sphere of human activity. Information system is designed to store, search and process information. The main task of information systems is to meet specific information needs within a specific subject area. With the development of society, the number, complexity and size (volume) of information systems increases.

One of the approaches to solving problems associated with the complexity of describing the subject area is the use of semantic technologies. Semantics is the science that establishes the relationship between symbols and objects, which they denote [1]. In other words, semantics is a science that defines the meaning of signs. A distinctive feature of semantic information systems is the ability to process (change the presentation form, search, etc.) semantic information (expressed by signs about information on selected sides of objects).

In the course of the life cycle of an information system (IS), one may encounter the problem of changing system requirements [2]. There are several reasons for the emergence of new requirements, but the main one is the large time duration from the moment of the analysis of the subject area and the design of IP to its final implementation and implementation. During the creation of the system, new business processes may appear or existing ones will change significantly. There is also a problem of different interpretations of requirements by the customer and the contractor, due to the lack of awareness in the technologies used and the subject area, respectively. In addition, a big problem is the adaptation of the information system to changing external requirements.

Figure 1 schematically shows the model of the life cycle of IP. The first step is the study and analysis of the subject area and customer requirements. Then models of the future information system are designed. After that, the coding of the program starts according to the available models. But at any of these moments, new requirements may appear, which leads to a re-study of the subject area, the development of new models and the writing of a new program or the introduction of changes to an already created part of the information system.

II. ONTOLOGIES AS A KNOWLEDGE REPRESENTATION MODEL

Ontologies allow the automated processing of data semantics for the purpose of its effective use (representation, transformation, search). The corresponding principle of data processing is based on the presentation of the description of the subject area as a knowledge base containing concepts and interrelations, and is focused, first of all, not on the understanding of information by a person, but on the automated interpretation and processing of information. The use of ontologies in information systems allows you to reflect the real picture of the world in the form of concepts, relationships and carry out different interpretations. [1,3] Maintaining the Integrity of the Specifications.

\[ O = \langle C,R,F \rangle \]

C - a finite set of concepts (concepts) of the subject area;
R is a finite set of relationships between concepts;
F is a finite set of interpretation functions defined on concepts and / or relationships.

The most common ontological description languages are RDF, RDFS and OWL.

The Resource Description Framework (RDF) presents statements about resources in the form of subject-predicate-object triplets, suitable for machine processing, and provides the means for building information models, but does not concern the semantics of the subject domain. RDF consists of two parts: the specification of the RDF data model (oriented graphs) and the XML-based syntax for writing these models. The description of the data model is the main part of the specification. The advantage of considering an RDF data model in this way is its independence from a particular syntax. The RDF language allows you to describe metadata, but working with them also requires describing the ontology, on the basis of which metadata are built and which defines the links between the concepts and relationships used in metadata. For these purposes, it is proposed to use the RDF Schema language (RDFS).

RDFS Description Language (RDF Vocabulary Description Language or RDF Schema) is intended for simple ontologies that describe the concepts and relationships used to write metadata in RDF. The RDF Schema language has limited expressive capabilities and uses RDF to describe dictionaries of concepts and properties. Thus, any RDFS description is “ordinary” RDF data — data about classes and properties. RDFS provides only basic capabilities for describing subject dictionaries, but it can easily be extended with additional modeling primitives that more specifically and specifically describe the necessary aspects of classes and properties. Extensibility allows RDFS to serve as the foundation for more expressive domain descriptions of ontologies that allow semantic data processing.

Thus, the use of such semantic technologies as RDF, RDFS, ontologies inside information systems gives them new intellectual properties and the ability to quickly adapt to changing external conditions of functioning [1,2].

The main advantage of the ontology is that they allow the automated processing of data semantics for the purpose of its effective use (representation, transformation, search). The corresponding principle of data processing is based on the presentation of the description of the subject area as a knowledge base containing concepts and relationships, and is focused primarily on automated interpretation and information processing [3].

OntoMap toolkit [4] will be used to build ontologies (see Figure 2).

III. METHODOLOGICAL APPROACH FOR CREATING AND MODIFYING INFORMATION SYSTEMS BASED ON AN ONTOLOGICAL MODEL

It is proposed to allocate some variational (relatively easily changeable) part, which would be associated with the information system model [5]. This model should adequately describe the subject area and the tasks facing the information system (see Figure 3).

Due to this decision, when the model changes, the system behavior will also change. This will allow you to quickly adapt to changing environmental conditions and greatly simplify the implementation of new requirements for the information system. Ontologies can be used as such information system model.

To create web-systems is necessary to determine two ontology classes. To create web-systems is necessary to determine two ontology classes. The first class of ontologies (ODATA) answers the question “What we keep?” It is required for a static display of system structure and allows viewing the information from the database corresponding to this structure.

The second class of ontologies (O\text{ACT}) answers the question “How does it work?” It is needed so that the system becomes dynamic, that is, it can not only display information, according to the first class ontology, but also allows you to work with data, that is, edit, add or change [5].

For example, create a model of collective distributed energy activities. The ontology presented in Figure 4 highlights the essence of the project manager who creates the project and appoints researchers and power engineering experts. According to the focus and objectives of the project, it identifies...
design research and energy technology and research, and researchers and energy experts are respectively appointed as managers. Also, as a result of the work of the research team on the project, scientific publications are created, which are entered into the system by a bibliographer.

After creating the ontology of the domain, it is necessary to add the attributes corresponding to each entity.

For the construction of the ontology of the second class from the ontology of the subject area, those objects that can influence information are singled out. Such objects are “leader”, “researcher”, “energy expert”, “bibliographer”. And such objects as “project”, “project of research”, “scientific publication” and “energy technology” cannot influence other objects.

The second step in building the ontology of the second class is the selection of strong and weak links. Strong bonds are a bond between two objects, one of which can change the other. For example, “supervisor” - “project”, “researcher” - “project of research”, “bibliographer” - “scientific publication”. Weak bonds are formed when there is a connection between objects, but one object cannot change another. For example, the “project” - “scientific research.” In Figure 5, strong links are highlighted with a red line.

And the last step (see Figure 6) is to form the rights of objects with a strong connection or, in other words, choose the type of connections. In total there are 7 types of communication:

- delete;
- insert;
- update;
- delete, insert;
- delete, update;
- insert, update;
- insert, update, delete.

Thanks to this ontology, the system will be able to “understand” who and what actions can be performed with the data within the framework of the information web-system being developed.

IV. SYSTEM DEVELOPMENT TECHNOLOGY OF WORK WITH THE SYSTEM

The technology of working with the system is shown in Figure 7.

The first step to create an information system is to model its architecture. In this step, we need to create two classes of ontologies described above: the data storage structure and the data processing model. It should be noted that these classes of ontologies are rigidly interconnected, since the data processing model is based on the storage model.
ontology is changed automatically (dynamic change). The second way - after adjusting the ontology, the system must be given a command to edit the data structure.

The third step is filling the system with data. This step begins after the creation of the information system and the data is filled in the working mode by the users of the system.

The application of this approach to support the tasks of scientific and technological forecasting of energy research, research on cybersecurity problems of energy facilities, as well as in other areas, will significantly simplify the creation and support of information systems.

V. TOOL ARCHITECTURE

The architecture of tools for creating and modifying information systems based on an ontological model is presented in Figure 8.

The construction and modification of the ontological model, as already described, occurs with the help of the OntoMap toolkit [5]. Further, the obtained ontologies of the first and second class in a special OXM format are transmitted to the developed plug-in “ModelToXML”, which extracts the metadata and on their basis generates an XML file. Based on the XML file, the CreateRepository component automatically creates a database and a set of tables corresponding to the entities and their attributes specified in the ontological model, and also assigns access rights.

And the final stage is the configuration of the information system and connection to the database using the component “CreateSystem”. After completing the configuration, the web information system is ready to be entered by users.

If it is necessary to make changes to the system, it is necessary to modify the domain ontologies created and then reconfigure the information system.

The implementation of the components of a web application configuration is performed using technologies such as JavaScript [7], Bootstrap [8], HTML5 [9], server technologies are implemented using Java and PHP.

VI. CONCLUSION

The paper considers a methodical approach and tools for implementing and modifying informational Web-systems based on semantic domain models, presented in the form of ontologies. This approach allows to obtain an information system that is completely ready for operation based on user-built ontologies. The result is achieved by interpreting the ontologies of the first and second class subject areas.

Application of the proposed approach greatly simplifies the creation and modification of support
systems for the tasks of scientific and technological forecasting of energy research and research on the problems of cybersecurity of energy facilities.

ACKNOWLEDGMENT

The results were obtained during the implementation of the basic research project of the fundamental research programs of SB RAS III.17.2.1, reg. № АААА-17-117030310444-2, III.17.1.4, reg. № АААА-А17-117030310436-7, with partial financial support of RFBR grants № 17-07-01341, № 18-37-00271.

REFERENCES

[1] Rationale for the development of electric power systems: Methodology, models, methods, their use / N.I. Voropay. Novosibirsk: "The science". 2015 p.448
[2] Cagnin C. Future-Oriented Technology Analysis. Strategic Intelligence for an Innovative Economy. – Springer, 2008. – p 170.
[3] Mikheev A.V. Semantic-based approach for energy technology forecasting // Proceeding of Internation Workshop “Contingency management, intelligent, agent-based computing and cyber security in critical infrastructures” (CM/IAC/CS/CI-2016) – Russia, Irkutsk: MESI, 2016. – pp. 41-42.
[4] Kopaygorodsky A.N. Information support of collective expert activity for forecasting of innovative development of power engineering// Proceeding of Internation Workshop “Critical Infrastructures: Contingency Management, Intelligent, Agent-based, Cloud Computing and Cyber Security”.
[5] Khayrullina E.P., Kopaygorodsky A.N. Application of ontologies in the design and implementation of information systems // Informatization and visualization of economic and social life, 2015. – pp. 200-204
[6] Gaskova D.A. Analysis of cybersecurity violations in the energy sector// “System research in the energy” / Proceedings of young scientists ESI SB RAS, №. 47. – Irkutsk: ESI SB RAS, 2017. – pp. 101-107.
[7] Manoj K. S., Dileep K. G. Effective big data management and opportunities for implementation // Information Science Reference, America, 2016 – p.440
[8] Dumbill E. Planning for Big Data // O'Really Radar Team, America, 2012 – p.78
[9] Woon, Wei Lee, Andreas Henschel and Stuart Madnick. “A Framework for Technology Forecasting and Visualization.” IEEE, 2009. – pp. 155-159.
[10] P. Willett. “The Porter stemming algorithm: then and now” // Program: Electronic Library and Information Systems, 2006. – pp. 219-223.
[11] Huiolong Fan, Yongbin Qin. “Research on Text Classification Based on Improved TF-IDF Algorithm”// The International Conference on Network, Communication, Computer Engineering, 2018. – pp.501-506.