Integration of an applied ontology and wiki" resources in the context of the unified knowledge base

V.S. Moshkin,
A.A. Filippov,
N.G. Yarushkina
Ulyanovsk State Technical University,
Ulyanovsk, Russia
Email: postforvadim@ya.ru
Email: al.filippov@ulstu.ru
Email: jng@ulstu.ru

***Abstract*—The article considers the architecture of the multi" agent intelligent subsystem designated for construction of the knowledge base (KB) by integrating with wiki" resources. The KB represents the storage of knowledge and contexts of different problem areas (PAs) in the form of an applied ontology with nodes used for building the structure of wiki" resources.**

***Keywords*—knowledge base, wiki" resources, ontology, context, problem area.**

INTRODUCTION

In the process of any large modern organization activity, it is necessary to make urgent management decisions timely that requires specialists to have deep knowledge of the problem area. Moreover, they should be able to use different decision support systems and tools for work with knowledge.

The desire to automate and speed" up the process of obtaining necessary knowledge about the problem area drives the need in the unified multipurpose toolkit for knowledge management that does not require a user to have some additional skills in the field of knowledge engineering and ontological analysis.

Thus, one can identify a number of scientific problems besetting modern organizations. In order to be solved, such problems require the systematic approach and include the following ones:

- the need of developing the semantic basis for representation of electronic information storage content;
- the lack of integrative conceptual models using different approaches to the storage of knowledge about the PA;
- the need of unifying the automated processing of the stored knowledge;
- the need of simultaneous use of multi" aspect contexts of the PA under consideration;
- the need of solving the problem of tracking the clarity of human reasonings.

Thereby, nowadays, the actual problem is providing specialists of a wide range of organizations with a universal tool allowing to address the knowledge management challenges [1], [10]. Furthermore, the tool should not require some extra training of users.

I. KNOWLEDGE BASE OF THE MODERN ORGANIZATION

At the moment, the ontological approach is most often used for organization of knowledge bases of expert systems. A lot of Russian and foreign researchers such as T.A. Gavrilova [3], V.N. Vagin [2], V.V. Gribova [5], Yu.A. Zagorulko [6], A.S. Kleschev [8], I.P. Norenkov, D.E. Palchunov, S.V. Smirnov [9], D. Bianchini, T.R. Gruber, A. Medche, G. Stumme and others address the problem of integration and search of information in order to provide management decision support on the basis of an ontology.

In a broad sense, ontologies are models representing knowledge within the individual contexts of the PA in the form of semantic information" logical networks of interrelated objects where the PA concepts with properties and relations between objects are the main elements.

Ontologies serve as integrators proving the common semantic basis in the processes of decision" making and data mining, and the unified platform for combination of different information systems [4], [14], [15].

However, aside from the obvious advantages of the ontological approach use, the following disadvantages arise:

- the need to involve an expert of the PA in order to construct the KB; herewith, some degree of judgment is brought into the obtained description (model) of the PA;
- conversion from the external representation of any PA object to its internal description on the knowledge representation language that requires an expert to have some certain skills in the field of knowledge engineering and ontological analysis;

- the need of adaptation of the applied ontology to the PA changing contexts that leads to the necessary use of different methods of the KB automated extension.

A lot of the problems reviewed earlier have been solved in corporate knowledge bases (CKBs) represented in different internal and public wiki" resources [13].

A wiki" resource represents a web" site with structure and content that can be changed independently by users with tools provided within the site functionality with the use of the special markup language.

Thus, corporate wiki" resources allow:

1) To construct the certain CKB fragments not requiring the expert to have some additional skills in the field of ontological analysis, knowledge engineering and using different specialized software.
2) To let a number of experts make edits into generated CKB fragments that gives a partial opportunity to avoid the problem of data subjectivity.
3) To monitor the progress dynamics of the KB content, and if needed, return to one of the previous versions of the CKB content.
4) To use the advanced set of program interfaces (API) and extensions allowing to construct and edit fragments of the KB content in the automatic or computer" aided mode.
5) To construct a framework of the PA KB ontology and update its fragments on the basis of the analysis of the content of different CKBs.

In spite of all the advantages of wiki" resources and their naive" user orientation, this type of CKB has an essential disadvantage. It is a lack of the mechanism for checking logical integrity and semantic coherence of the PA objects included in wiki" resources.

Therefore, there is a need for integration of the PA applied ontology with corporate wiki" resources in the context of the unified knowledge base.

II. CONSTRUCTING THE KNOWLEDGE BASE OF THE ORGANIZATION BY EXTRACTING KNOWLEDGE FROM WIKI" RESOURCES

One of the KB main problems is providing the mechanism for adapting the Athene technological platform [11] to the concrete PA with the use of methods of ontological analysis and data engineering.

The KB ability of taking into account the dynamic nature of processes refers to the tools existing in the KB ontology that allow to describe the process giving the permitted set of input ontology objects, constraints imposed on them, new or changed ontology objects obtained as a result of the process implementation.

The PA ontology context is a specific state of the KB content than can be chosen from a set of the ontology states. The state was obtained as a result of either versioning or constructing the KB content from different points of views.

Formally, the ontology can be represented by the following equation:

$$O = \langle T, C^{T_i}, I^{T_i}, P^{T_i}, S^{T_i}, F^{T_i}, R^{T_i} \rangle, i = \overline{1, n},$$

where n is a number of the ontology contexts; $T = \{T_1, T_2 \ldots, T_n$ is a set of ontology contexts; C^{T_i} is a set of ontology classes within the i" th context; I^{T_i} is a set of ontology objects within the i" th context; P^{T_i} is a set of ontology classes properties within the i" th context; S^{T_i} is a set of ontology objects states within the i" th context; F^{T_i} is a set of the PA functions fixed in the ontology within the i" th context; R^{T_i} is a set of ontology relations within the i" th context defined as:

$$R^{T_i} = \{R_C^{T_i}, R_I^{T_i}, R_P^{T_i}, R_S^{T_i}, R_{F_{IN}}^{T_i}, R_{F_{OUT}}^{T_i}\},$$

where $R_C^{T_i}$ is a set of relations defining hierarchy of ontology classes within the i" th context; $R_I^{T_i}$ is a set of relations defining the 'class–object' ontology tie within the i" th context; $R_P^{T_i}$ is a set of relations defining the 'class–class property' ontology tie within the i" th context; $R_S^{T_i}$ is a set of relations defining the 'object–object state' ontology tie within the i" th context; $R_{F_{IN}}^{T_i}$ is a set of relations defining the tie between F_{IN} function entry and other instances of the ontology within the i" th context; $R_{F_{OUT}}^{T_i}$ is a set of relations defining the tie between F_{OUT} function exit and other instances of the ontology within the i" th context.

Figure 1 shows the illustrative example of the knowledge base ontology that includes description of the process of the component manufacturing.

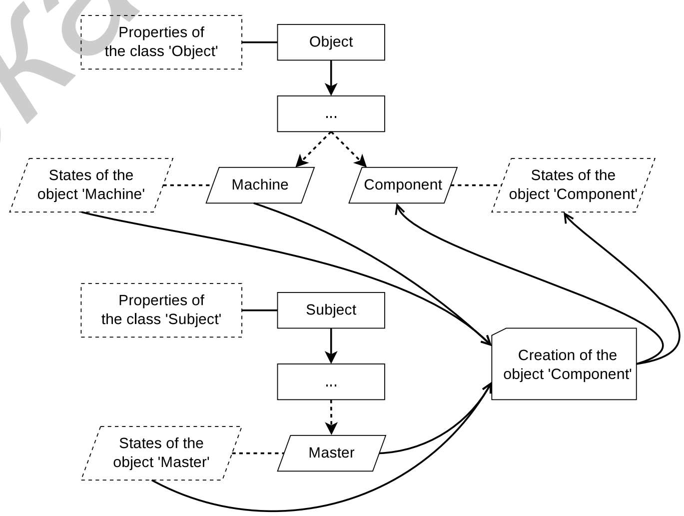

Figure 1. The illustrative example of the knowledge base ontology

The presented ontology includes 'Object' and 'Subject' classes with specific properties. These classes are the parent ones for all other ontology classes, herewith, parent properties are inherited by descendants.

The ontology also includes object 'Machine', 'Component' and 'Master', each of them has its own set of relations. 'Creation of the object 'Component' is a description of the component manufacturing process fixed in the ontology. The process has two inputs: 'Machine' and 'Master', and one output – 'Component'. 'Component' properties directly depend on machine properties and master qualification.

In order to automate the expert work in creating the PA ontology within the context of the KB under review, the method of automated creation of the ontology structure on the basis of external wiki" resources content is used. Herewith, the ontology structure is made in the process of the analysis of the resource categories system and infoboxes (standardized tables containing the key information about the object) described in the article [16], [13].

In order to construct external wiki" resources on the basis of the KB content, the following procedures should be followed:

1) Expert points what ontology classes should be taken into account in the process of creating an external wiki" resource as a category, subcategory, and page.
2) Expert points what ontology relations describe the tie between the object and its description, for example, in the text form.
3) System on the basis of the ontology relations analysis builds the structure of the external wiki" resource.

Also, there is an alternative approach to the constructing wiki" resources (internal wiki" resources) on the basis on the KB content: the user applies the KB mechanisms for obtaining and editing data with the use of dynamically generated screen forms. The approach allows to combine advantages of ontologies and wiki" resources through client" oriented managing tools and mechanisms for checking logical integrity and semantic coherence of the KB content. Figure 2 shows the architecture of the developed knowledge base.

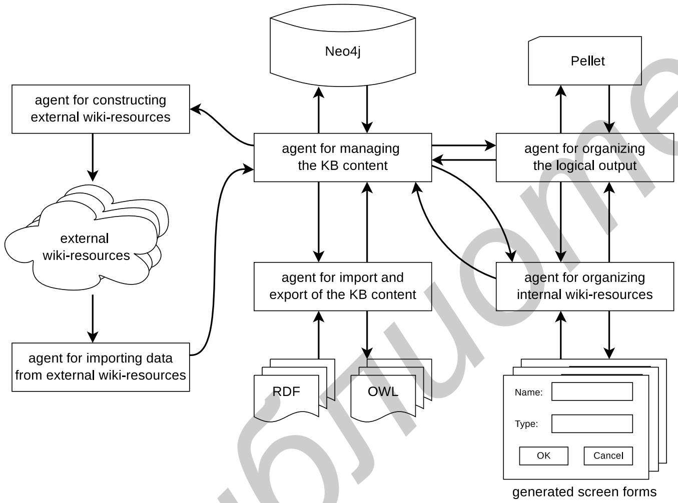

Figure 2. The knowledge base architecture

The KB consists of some agents, which interact closely among themselves:

- agent for managing the KB content;
- agent for import/export of the KB content from/to different formats of the PA ontology description (RDF, OWL, etc.);
- agent for organizing internal (on the Athene platform) wiki" resources on the basis of the KB content;
- agent for constructing external wiki" resources;
- agent for importing data from external wiki" resources to the KB;
- agent for organizing the logical output according to the KB content.

In order to develop agents of the Athene technological platform, Java programming language and Spring framework were used. Such development tools have the following advantages:

1) High development rate.
2) Existence of documentation and active community of developers.
3) Platform independence.
4) Advanced infrastructure.

Neo4j graph database is used as a storage for the managing agent ontology. It has the following advantages:

1) Native format for graph storages.
2) One database instance can serve graphs with billions of nodes and relations.
3) It can process graphs that do not have enough space in RAM.

Agents are performed in the Jetty servlets container with the modular architecture that allows to use only needed functions, thereby, it reduces the performance load on the server. Also Jetty is highly scalable for performing a lot of connections with significant downtime between the queries. It also allows to serve a lot of users.

In order to develop means for interaction with agents, the REST (Representational State Transfer) mechanism was used. It this case, the remote procedure call represents a simple HTTP request (GET, POST, PUT, etc.), and necessary data are transmitted as parameters of the request. The main benefits of REST are the following ones:

- high performance due to the use of cash;
- scalability;
- integration system transparency;
- simplicity of interfaces;
- portability of components;
- modification simplicity.

All the above resources, applications, and technologies are free.

Conclusion

Consequently, enhancement of the ontology platform (by the example of the developed Athene one) via using expert knowledge control agents allows to provide specialists of different organizations with the universal toolkit for analysis of the problem area features with the ability of automated fill and extension of the database from public sources; and for problem area visualization in the form of the complexly structured material.

The most important aspect in solving the task of overall information support for specialists is the ability of data representation in the matrix of their contexts including temporary ones and also contexts of different points of view on the PA objects under review.

Acknowledgements

This work was financially supported by the Russian Foundation for Basic Research (Grant No. 16" 47" 732054 р_офи_м).

References

[1] Bova V.V., Kureichik V.V., Nuzhnov E.V. Problemy predstavleniia znanii v integrirovannykh sistemakh podderzhki upravlencheskikh reshenii [Problems of Knowledge Presentation in Management Decision Support of Integrated Systems]. Izvestiia IuFU [Izvestiya SFedU], 2010, no. 108 (7), pp. 107" 113.

[2] Vagin V.N., Mikhailov I.S. Razrabotka metoda integratsii informatsionnykh sistem na osnove metamodelirovaniia i ontologii predmetnoi oblasti [Development of a Method of Integration of Information Systems Based on a Metamodeling and the Ontology]. Programmnye produkty I sistemy [Software & Systems], 2008, no. 1, pp. 22" 26.

[3] Gavrilova T.A. Ontologicheskii podkhod k upravleniiu znaniiami pri razrabotke korporativnykh informatsionnykh sistem [The Ontological Approach to Knowledge Management in the Development of Corporate Information Systems], Novosti iskusstvennogo intellekta [News of Artificial Intelligence], 2003, no. 2(56), pp. 24" 29.

[4] Golenkov V.V., Guliakina N.A.. Semanticheskaia tekhnologiia komponentnogo proektirovaniia sistem, upravliaemykh znaniiami [Semantic Technology of Component Design of Systems, Managed by Knowledges]. Materialy V mezhdunarodnoi nauchno" tehnicheskoi konferntsii OSTIS" 2015 [Proc. of the 5th Int. Sci. and Tech. Conf. 'Open Semantic Technology for Intelligent Systems' (OSTIS" 2015)]. Minsk, 2015, pp. 57" 78.

[5] Gribova V.V., Kleschev A.S. Upravlenie proektirovaniem i realizatsiei polzovatelskogo interfeisa na osnove ontologii [The Management of User Interface Design and Implementation on the Basis of Ontologies]. Problemy Upravleniia [Control Sciences], 2006, no. 2, pp. 58" 62.

[6] Zagorulko Yu.A. Postroenie portalov nauchnykh znanii na osnove ontologii [Constructing Portals of Scientific Knowledge on the Basis of the Ontology]. Vychislitelnye tekhnologii [Computational Technologies], 2007, vol. 12, no. S2, pp. 169" 177.

[7] Karabach A.E. Sistemy integratsii informatsii na osnove semanticheskikh tekhnologii [Systems for Integration of Information on the Basis of Semantic Technologies]. Nauka, tekhnika i obrazovanie [Science, Technics and Education], 2014, no. 2 (2), pp. 58" 62.

[8] Kleschev A.S. Rol ontologii v programmirovanii. Chast 1. Analitika [The Role of Ontologies in Software Development. Part 1. Analitics]. Informatsionnye tekhnologii [Information Technologies], 2008, no. 10, pp. 42" 46.

[9] Smirnov S.V. Ontologicheskoe modelirovanie v situatsionnom upravlenii [Ontological Modeling in Situational Management]. Ontologiia proektirovaniia [Ontology of Designing], 2012, no. 2 (4), pp. 16" 24.

[10] Tuzovsky A.F. Razrabotka sistem upravleniia znaniiami na osnove edinoi ontologicheskoi bazy znanii [Development of Knowledge Control System on the Basis of a United Model of Knowledge Organization]. Izvestiia Tomskogo politekhnicheskogo universiteta [Bulletin of the Tomsk Polytechnic University]. 2007, vol. 310, no. 2, pp. 182" 185.

[11] Filippov A.A., Moshkin V.S., Shalaev D.O., Yarushkina N.G. Edinaia ontologicheskaia platforma intellektualnogo analiza dannykh [Uniform Ontological Data Mining Platform]. Materialy VI mezhdunarodnoi nauchno" tekhnicheskoi konferntsii OSTIS" 2016 [Proc. of the 6th Int. Sci. and Tech. Conf. 'Open Semantic Technology for Intelligent Systems' (OSTIS" 2016)]. Minsk, 2016, pp. 77" 83.

[12] Chernyahovskaya L.R., Fedorova N.I., Nizamutdinova R.I. Intellektualnaia podderzhka priniatiia reshenii v operativnom upravlenii delovymi protsessami predpriiatiia [Intelligent Decision Support in Operative Management of Organization Business Processes]. Vestnik UGATU. Upravlenie, vychislitelnaia tekhnika i informatika [Vestnik UGATU. Management, Computer Science and Informatics]. 2011, vol. 15, no. 42 (2), pp. 172" 176.

[13] Shestakov V.K. Razrabotka i soprovozhdenie informatsionnykh sistem, baziruiushchikhsia na ontologii i Wiki" tekhnologii [Development and Maintenance of Information Systems Based on Ontology and Wiki" Technology]. Trudy 13 Vserossiiskoi nauchnoi konferentsii “Elektronnye biblioteki: perspektivnye metody i tekhnologii, elektronnye kollektsii [Proc. of the 13th All" Russian Sci. Conf. RCDL'2011 'Digital libraries: Advanced Methods and Technologies, Digital Collections']. Voronezh, 2011, pp. 299" 306.

[14] Yarushkina N.G., Perfileva I.G., Afanaseva T.V. Integratsiia nechetkikh modeleii dlia analiza vremennykh riadov [Fuzzy Models Integration in Time Series Analysis]. Izvestiia Samarskogo nauchnogo tsentra Rossiiskoi akademii nauk [Scientific Journal of 'Proceedings of the Samara Scientific Center of the Russian Academy of Sciences']. 2010, vol. 12, no. 4 (2), pp. 506" 509.

[15] Perfilieva I., Yarushkina N., Afanasieva T., Romanov A. Time Series Analysis Using Soft Computing Methods. International Journal of General Systems, 2013, vol. 42, no. 6, pp. 687" 705.

[16] Suchanek F.M., Kasneci G., Weikum G. YAGO: A Core of Semantic Knowledge Unifying WordNet and Wikipedia. Proc. of the 16th Int. World Wide Web Conf. Canada. 2007, pp. 697" 706.

ИНТЕГРАЦИЯ ПРИКЛАДНОЙ ОНТОЛОГИИ И WIKI-РЕСУРСОВ В РАМКАХ ЕДИНОЙ БАЗЫ ЗНАНИЙ

В.С.Мошкин, А.А. Филиппов, Н.Г. Ярушкина

В статье представлена модель онтологии нечеткой предметной базы знаний (БЗ), позволяющая описать проблемную область (ПрО) с учетом многообразия ее контекстов. Под контекстом ПрО понимается состояние содержимого БЗ, которое может быть сформировано из множества состояний онтологии, полученного в результате версионирования либо формирования содержимого БЗ с различных точек зрения («point of view»). Также описано применение онтологического подхода для интеграции гетерогенных корпоративных информационных ресурсов. В качестве корпоративных информационных ресурсов рассматриваются крупные корпусы специализированных текстов, непосредственно связанные с ПрО, и различные виды корпоративных БЗ в виде внутренних сайтов и wiki-ресурсов. Также представлена архитектура многоагентной интеллектуальной подсистемы.