Processing Information about Junior Specialists for Small IT-projects Teams Using Linked Data

Nafissa Yussupova
Computer Science and Robotics Department
Ufa State Aviation Technical University
Ufa, Russian Federation
e-mail: yussupova@ugatu.ac.ru

Dmitry Rizvanov
Computer Science and Robotics Department
Ufa State Aviation Technical University
Ufa, Russian Federation
e-mail: ridmi@mail.ru

Olga Smetanina
Computer Science and Robotics Department
Ufa State Aviation Technical University
Ufa, Russian Federation
e-mail: smoljushka@mail.ru

Artur Galyamov
Computer Science and Robotics Department
Ufa State Aviation Technical University
Ufa, Russian Federation
e-mail: galyamov.artur@profport.org

Konstantin Mironov
Computer Science and Robotics Department
Ufa State Aviation Technical University
Ufa, Russian Federation
e-mail: mironovconst@gmail.com

***Abstract*—In this paper, we develop an approach and Web-ontology model for globally distributed structured data processing related to junior professionals. This information is represented using linked data. The proposed approach is based on Representational State Transfer and Semantic Web technologies. The proposed approach and model were applied in a Web-application for extracting and searching information about competences of junior specialists to build teams for small IT-projects.**

Keywords- Linked data; OWL; RDFa; globally distributed data processing; web-ontologies; competences.

I. INTRODUCTION

According to some investigations [17], more than 50% of the industrial projects in the sphere of Information Technologies (IT-projects) are not successful. According to marketing research and experience of project managers [17], inadequate organizational structure of the project teams is one of the main reasons for the project's failure. Competence model can help to increase the validity of decisions taken by project managers in this situation. This paper is focused on elaboration of ontological competence model and its further use in the Web-application.

II. RELATED WORK

The area of interest is explored in a number of scientific works [1-16]. Semantic search, semantic Web, Web-ontologies, descriptive logic are studied in [18-21]. Numerous researches are dedicated to knowledge representation models and architecture of global distributed information systems [11][13][15-16].

Here, we develop a competence model for knowledge representation. A competence model is a set of skills (or competences) [4], which are required to execute tasks of the project. In this case, the competence model will be assumed in the context of management and recruitment projects.

Competences can be classified as follows [5]:

- Corporate competences, which do not depend on the specific position and are common for all employees of the entire enterprise;
- Managerial competences, which are specific to managers who carry out strategic vision, business administration and other high-level tasks;
- Professional competences, specific to concrete units.

The authors of [5] added levels for each competence to the competence model. "Level" in this model means the assessment of competence, with specific descriptions. A clear competence model is constructed in [4], however the classification of competences is very diffuse (social competences, socio-psychological competences, conceptual competences, etc.). In [6], no clear distinction between competence and skill is made. The skills are separated into knowledge and abilities. All skills may depend on strengths and weaknesses (soft). Basically, this model was constructed for students of technical colleges, but it can be adapted to other areas. It is based on the ontological approach.

Competences in the model from [7, p.4] are divided into characteristics, motives, skills, knowledge and self-esteem. If they are directly used in the taxonomy of competences, they

 ISBN: 978-1-61208-552-4

complicate the model too much and will not lead to desired results [7]. However, some of them may be useful as additional characteristics of competences, in the form of relations in the domain of competences. i.e., "negotiation with the customer" is a skill, and "overcoming stressful situations" is a feature.

In [9], IT-supported strategic competence management is considered as a part of a human resource management system. For the competence estimation, a scale from 0 (not assessed) to 9 (excellent) is used. The values are evaluated by using the period of time, when a person has used the competence.

More approaches to the solution of this classification problem, taken from the experience of HR managers [6, pp. 762], may be viewed as follows: skills, knowledge, ability, attitude. These competences are divided into 4 levels.

To assess the competences of students directly, the competences are classified as connected to computer science, business, or behavior [6]. In accordance with this classification, the score ranges from 0 to 1. Each competence is assessed by the level of knowledge (beginner, intermediate and advanced) and experience (basic, intermediate, advanced). The competence is determined by three attributes: the name (a unique identifier for use in HR-XML profile), the scale, and the traceable calibration to assess the strength. In [10], components of competence are knowledge, know-how and behavior. Knowledge is what gets people through the education system. Know-how can be obtained from personal experience and practice. Behavior is a personal characteristic, allowing knowledge to become a know-how. According to the HR-XML [10, p.762], it is necessary to admit a competence model comparison study (validity) and ensure the privacy of the information provided. The ability to compare two people in terms of possession of a competence can be achieved by describing the selection of competences from the dictionary, giving the relationship of similarity of competences and competence assessment.

III. CONTRIBUTION OVERVIEW

Within the framework of this research a new approach is proposed, which we call "Globally Distributed Processing of Weakly Structured Data" (GDPWSD). The proposed model is slightly different from the existing approaches in three aspects: it uses Internationalized Resource Identifier (IRI) instead of Uniform Resource Identifier (URI), a three-tier architecture instead of the client-server, and mediators in the form of Web ontologies to unify the process of extracting heterogeneous data sources.

A novel technique is used for ontology creation. This technique is similar to the stages of database design (conceptual, data and physical description models). Its use allows the following:

• visualizing the relationship between the concepts of the domain in the form of a conceptual model;
• formalizing relations from the conceptual model in the form of a logical model;
• constructing an ontological model on the basis of the logical model.

IV. GLOBALLY DISTRIBUTED PROCESSING INFORMATION ABOUT JUNIOR SPECIALISTS

The proposed hybrid approach is based on the Representational State Transfer (REST) approach and mediator-wrapper approach [11, p. 30; 12, p. 94], which uses Semantic Web technologies, including linked data formats.

The main concepts of the approach are the following:

- information processing is based on a three-level architecture (client, Web server, and a knowledge base / database server);
- the server does not store the view state. It applies a uniform interface to access the resources based on the use of IRI;
- three levels of abstraction are used:
    - application layer, where the search is made (managed by users);
    - intermediate level (the use of mediators), which implements data collection and aggregation; Web ontology, which brings together various descriptions, is developed at this level (managed by ontologists).
    - data sources level, using linked data (managed by site administrators).
- The project and junior specialists are associated via role and competence;
- It is taken into account that some competences are equal (synonyms) and some can be part of others (meronyms).

The features of the proposed approach allow:

- flexible and scalable data processing;
- unified access to heterogeneous data sources by implementing semantic integration via Web Ontology;
- increasing the number of relevant search results.

V. ONTOLOGY MODEL DESCRIPTION

The ontology model is aimed at describing the necessary information [1-2]. It is possible to visualize the relationship between key concepts and roles using UML (OMG ODM standard). Description logic SHOIN (D) allows to define logical axioms. Web Ontology Language with description logics (OWL DL) is used to create Web-ontology and then utilize it in Web-applications. The basic classes of the research domain and their relationships are represented in Figure 1. It is assumed that competence and competency are synonyms.

 ISBN: 978-1-61208-552-4

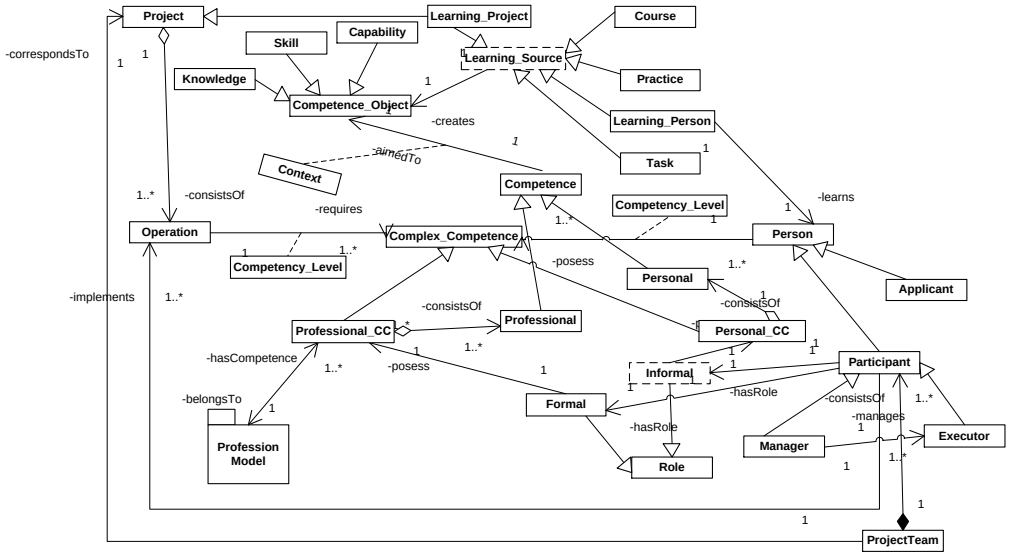

Figure 1. Representation of the Ontological Model Using UML Class Diagrams

Here, the concept of the project corresponds to *Project*, which consists of a series of operations *Operation*:

$$Project \equiv \forall \; consistsOf.Operation$$

The concept that corresponds to the notion of project team is *ProjectTeam*. In this model, it is a composition of roles in the project team:

$$ProjectTeam \equiv \forall \; consistsOf.Participant$$

In turn, the role *Role* can be formal and informal. Informal types of roles are taken from the socio-psychological models. Competence may be professional or personal. Hence, the formal role is associated with professional competence and the informal role is associated with personality:

$$Formal \subseteq \exists \; posess.Personal$$

$$Informal \subseteq \exists \; posess.Professional$$

The relation between the performer (Executor) and the Project is *Complex_Competence*, which is associated with the operation of the project *Operation*. Every complex competence consists of one or several single competences.

Each professional complex competence *Professional* is tied to a profession with the role *belongsTo*; it may be convenient to group competences, since sometimes one may need to immediately specify the profession a person should have to perform this role. For a hierarchy of occupations, the Standard Occupational Classification (SOC) was used, taking into account the specifics of the Russian employment market.

Each of the concepts belongs to its ontology. Their relationship is shown in Fig. 2.

Applying developed Web-application for searching about 100 various competences for 5 IT-projects showed increasing number of appropriate applicants for these projects by 10%.

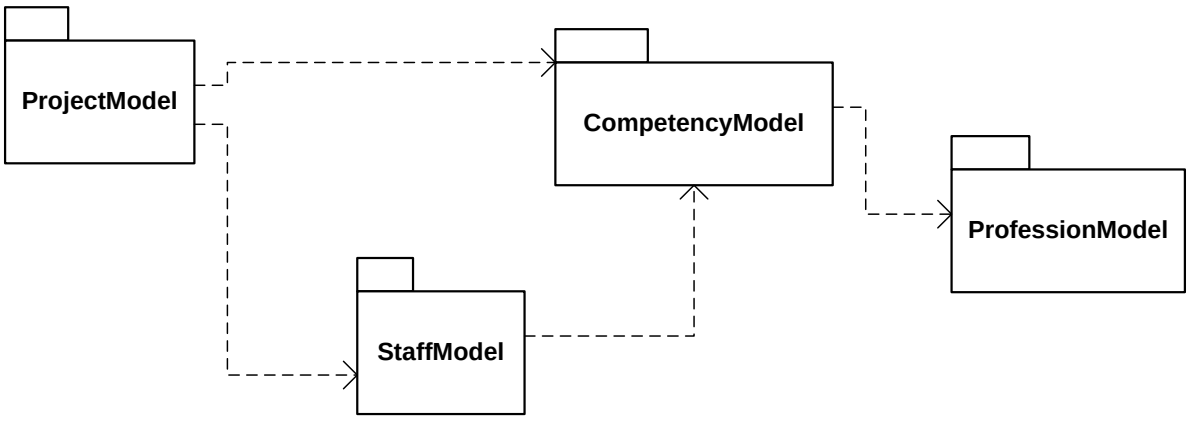

Figure 2. The Relationship of Ontology as a UML Package Diagrams

VI. WEB-APPLICATION

The Web-application based on our Web-ontology model, was implemented within the project ProfPort.org [14]. Information is extracted from the blogs and from the portfolios of the applicants and represented with RDFa, hCard microformat standard and simple HTML5/CSS3 sometimes. The Web-app is implemented with RubyOnRails framework as an application server, using document-oriented database MongoDB and RDF-store BigData. The deployment diagram and corresponding component diagram are presented in Fig. 3 and 4, respectively.

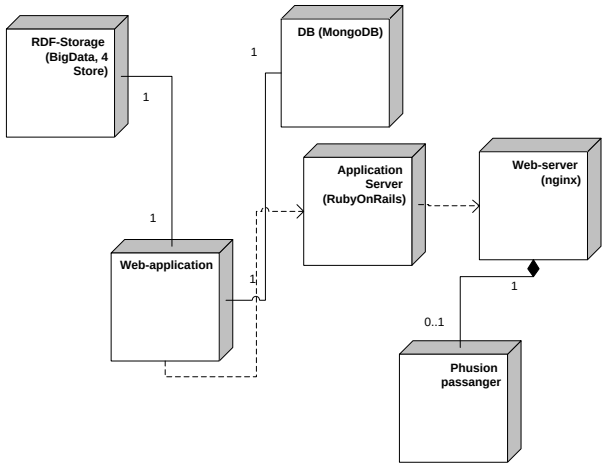

Figure 3. Deployment Diagram of the Web-App

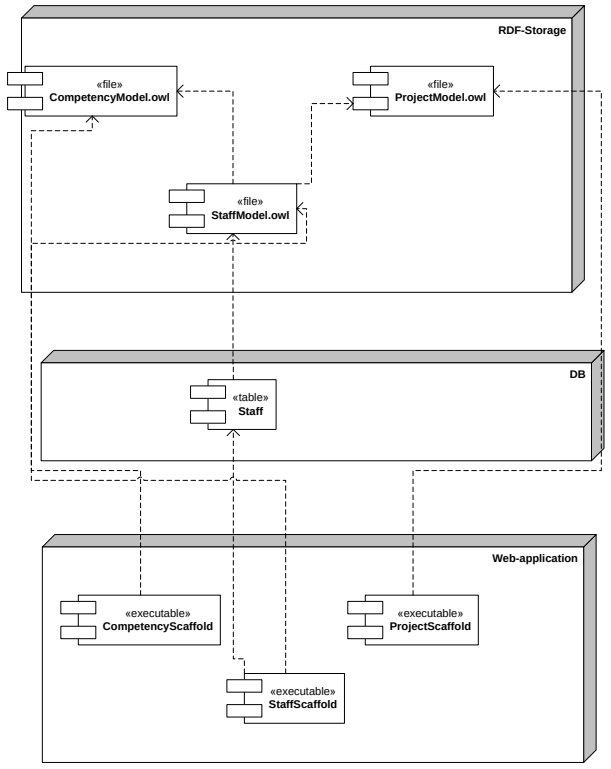

Figure 4. Component Diagram of the System Prototype

VII. EFFECTIVENESS ANALYSIS

The main purpose of the effectiveness analysis is to find out how the search results will change upon using synonymy and inheritance relations or without them. The main criteria that allows evaluating the effectiveness are completeness and relevance of the search.

Presently, 617 employment-seekers are registered in the ProfPort system; 428 competencies and 4 professions from the field of information technologies are introduced. These values allow rough designation of the search space.

To evaluate the speed of query execution, we use a standard personal computer with the following configuration:

- Processor: Intel Core i7-3770 @ 3.40GHz
- Motherboard: ASUSTek P8Z77-V LK
- Memory: DDR 3 Kingston 2xDIMM 4096 MB 800 MHz
- OS: Windows 8.1 Professional 64 bit
- SSD: OCZ Vertex3 224 GB
- Local server: Endels 1.64 Freeware

The fragment of the algorithm associated with the SPARQL query is used for the longest time using ARC2.The results of the experiment are shown in Table. 1.

TABLE I. SPEED OF EXECUTION OF VARIOUS REQUESTS

Experiment No. / Phrase	Programming in PHP (Separate Phrase)	Programming (Specific Competence)	Programming (Search Phrase)
1	58.366 s.	1.0568 s.	12.2957 s.
2	58.2956 s.	1.0506 s.	12.2085 s.
3	58.3636 s.	1.0572 s.	12.3267 s.
4	58.6011 s.	1.0502 s.	12.2845 s.
5	58.3122 s.	1.0415 s.	12.2504 s.
6	58.1308 s.	1.0635 s.	12.3185 s.
7	58.224 s.	1.0549 s.	12.3308 s.
8	58.1068 s.	1.0357 s.	12.2582 s.
9	58.1809 s.	1.0489 s.	12.2565 s.
10	58.2108 s.	1.0633 s.	12.3031 s.

Now, let us look at how the completeness is changed by using synonymy and inheritance relations. Coding is synonymous with the competence of Programming, which is the parent for several competencies, such as "Programming with PHP", "Programming in C#", "Programming with JAVA", and so on. Due to the use of ontology, these relationships can be taken into account in the search. Table 2 clearly shows the results of comparing the completeness of the results of the search for the keyword "Coding".

 ISBN: 978-1-61208-552-4

TABLE II. COMPARISON OF THE COMPLETENESS OF THE SEARCH

Inquiry	Number with synonymy and hierarchy	Number with synonymy only	Number with hierarchy only	Number without synonymy or hierarchy
Coding	10 competencies, 305 employment seekers	1 competency, 53 employment seekers	0 competencies, 0 employment seekers	0 competencies, 0 employment seekers

Based on the results shown in this table, it can be concluded that using an ontological approach for creating an information model that supports synonymy and inheritance relations, it has been possible to achieve a significant increase in the completeness of search results.

Let us now consider how the secondary search affects the relevance of the search results. For example, find employment seekers by searching the words PHP and git. By this query, quite a lot of results were found, about 10 competencies and 63 carriers of this competence. Let us assume the person conducting the search would want to clarify the results. To do this, it would be only necessary to enter those who are found in PHP. Accurate results are then obtained, with 6 competencies and 26 employment seekers, which more closely match search requests. After that, if there is a desire to find from the identified employment seekers only the ones with the competence of "Application of the basic PHPUnit testing techniques," our search would result in one specific competency left from the initial 6, and 6 employment seekers who own it. The search results and their quantitative estimates are summarized in Table 3.

TABLE III. THE IMPACT OF SECONDARY SEARCH ON THE RELEVANCE OF SEARCH RESULTS

Inquiry	"PHP Git"	"PHP"	Specific Competence
Results	10 competencies, 63 employment seekers	6 competencies, 26 employment seekers	1 competency, 6 employment seekers

The data given in Table 3 allows drawing the following conclusion: the use of detailed search for the results obtained has made it possible to increase relevance due to a gradual narrowing of the search area.

VIII. CONCLUSIONS

In this paper, an approach for processing information about junior specialists is proposed. It allows extraction and gathering of necessary information taking into account "synonyms" and "meronyms".

A Web-ontology model is created using the concepts of the proposed approach. It allows storing all necessary information about junior specialists and their competences.

A Web-application is developed using RubyOnRails framework. The created ontology is used for information representation. It allows a search for necessary specialists and building teams for projects.

Testing the developed Web-app showed an increased number of relevant junior specialists for IT-projects.

REFERENCES

[1] A.F. Galyamov, "Ontology knowledge base for decision making support to IT-project implementation," Proc. All-Russian winter school-seminar of postgraduates and young scientists «Actual problems of science and technique», Ufa, USATU, 2008, pp. 117-128.

[2] A.F. Galyamov, A.G. Abaytullin, and D.V. Popov, "Ontological model for decision support in IT consulting," Vestnik St. Petersburg University. Series "Information Science, Computer Science and Management, vol. 1, 2010, pp. 49-55.

[3] K. Zimin and M. Sukhanov, "Market research and prospects for the IT outsourcing," http://www.iemag.ru/researches/detail.php?ID=20002 (accessed Dec, 2011).

[4] E. Grebenyuk, "To assist HR-specialist: competence model," http://www.c-culture.ru/magazines/culture/2006-3-55 (accessed Dec, 2011).

[5] N. Volodina, "Competence model – it's easy," http://www.rhr.ru/index/rule/employees_certification/15320.html (accessed Dec, 2011).

[6] J. Dorn and M. Pichlmair, "A Competence Management system for Universities," Proc. European Conference on Information Systems, St. Gallen, 2007, pp. 759-770.

[7] K. Tucker and M. Cofsky, "Critical Keys to Competence-Based Pay," Compensation & Benefits Review, Nov.-Dec. 1993, pp. 46-52.

[8] Sh. Fletcher, "Competence-Based Assessment Techniques," Kogan Page, 2001.

[9] Hustad and Munkvold, "IT-Supported Competence Management: A Case Study at Ericsson," Information Systems Management, 2005.

[10] J. Dorn, M. Pichlmair, "Ontology development for human resource management," Proc. European Conference on Information Systems, 2009.

[11] R. T. Fielding, "Architectural Styles and the Design of Network-based Software Architectures," Dissertation. University of California, Irvine, 2000.

[12] Gio Wiederhold, "Mediators in the architecture of future information systems," Computer, 25(3), 1992, pp. 38–49.

[13] G. Kovacs, D. Bogdanova, N. Yussupova, and M. Boyko, "Informatics Tools, AI Models and Methods Used for Automatic Analysis of Customer Satisfaction," Studies In Informatics And Control (SIC). National Institute for R&D in Informatics, Vol. 24 (3), 2015, pp. 261–270.

[14] Portal for searching information about competences of junior specialists. Available at: http://profport.org/ (accessed: 12 December 2016)

 ISBN: 978-1-61208-552-4

[15] N. I. Yussupova, D. A. Rizvanov, K. R. Enikeeva, and O. N. Smetanina, "Knowledge representation models for decision making in complex systems management in uncertainty and resource restrictions," Proc. Information decision making technologies, Ufa, USATU, Vol. 2, 2016, pp. 24-27.

[16] N.I. Yussupova, D.R. Bogdanova, and M.V. Boyko, Processing structured information using artificial intelligence methods, Innovatsionnoe mashinostroenie, 2016.

[17] The Standish Group Report Chaos - Project Smart https://www.projectsmart.co.uk/white-papers/chaos-report.pdf

[18] L.R. Chernyahovskaya, V.N. Kruzhkov, and F.A. Dikova, "Ontological approach to development of decision support system," http://www.aselibrary.ru/datadocs/doc_917re.pdf.

[19] F. Baader, D. Calvanese et al. The Description Logic Handbook. Theory, implementation, and applications, Cambridge University Press, 2003, 340 p.

[20] V.A. Vittikh, "Ontological models of situations in the process of a collegial decision making," Proc. Problems of control and modeling in complex systems, Samara, 2009, pp. 405-410.

[21] T.A. Gavrilova, "Ontological engineering [in Russian]," http:// www.kmtec.ru/publications/library/authors/ontolog_engeneering.shtml.

 ISBN: 978-1-61208-552-4