UDC 004.896

DEVELOPMENT OF THE INFORMATION SYSTEM FOR SEMANTIC IDENTIFICATION AND DOCUMENTATION OF THE RESULTS OF NON-FORMAL AND INFORMAL EDUCATION

Serhii Pryima[a] [0000-0002-2654-5610], *Oksana Strokan*[a][0000-0002-6937-3548], *Dmitro Lubko*[a][0000-0002-2506-4145], *Yuliia Lytvyn*[0000-0001-5797-3151] , *Nadia Bilyk*[b] [0000-0003-2344-5347]

[a]Dmytro Motornyi Tavria State Agrotechnological University, 18 B.Khmelnytsky Ave, Melitopol, Zaporizhzhia obl., Ukraine

[b] M. V. Ostrogradsky Poltava Regional Institute of Postgraduate Pedagogical Education, 64, Sobornosti Str., Poltava, Ukraine

The publication describes methods and tools for development an information system (IS) for acceptance of the non-formal and informal learning results on base of computer ontologies. ESCO ontology of the multilingual skills classifier is used as a prototype of domain ontology that provides knowledge for developed IS. We describe the main elements of development of IS for validation of the results of non-formal learning that include creation of an ontology schema, process of integration of the obtained ontology into the RDF repository, development of application architecture and creation of user interface.

We consider modern approaches to design and development of knowledge-oriented distributed applications intended for functioning in the open Web environment and analyse the existing methods and software tools for knowledge representation and processing according to their suitability for solving the considered task. On base of this research we select such elements of the Semantic Web project as ontologies, Web services and software agents that can be used in development of this IS.

The paper describes the instrumental tools used in process of development of IS for validation of non-formal learning outcomes. In particular, we analyse Neo4j database management system serving the GraphDB database, the specifics of connectors and SPARQL requests to the data stored in the RDF repository and the tools used for web server creation. Comparison of PHP frameworks for web applications is performed in consideration of task requirements.

Functional modeling of IS in order to determine its main functionalities is performed, and the DFD data flow diagrams of system are designed. The benefits of Laravel software are established on base of the analysis of such criteria as security, readiness to installation of plugins and libraries, support for the MVC (Model-View-Controller) concept. User interface is developed to ensure user dialogue with IS. Authers analyse software tools oriented on development of user interface and select React framework that works efficiently with all software tools selected for IS development on the previous stages of analysis.

Key words: validation, information system, non-formal learning, GraphDB, ontology, Semantic Web, SPARQL query language, connector, RDF storage.

Introduction

Due to the rapid growth of knowledge-based systems that occupy a wide area in information technology and development of their own methods and principles, the Semantic Web initiative has undergone significant elaboration [1, 2]. Semantic Web technologies aim to define and interconnect data in a way similar to that in which traditional web technologies define and interconnect web pages. In the case of the traditional Web, each web page can be considered a unit of information or entity and pages are explicitly linked using HTML links. The main features of Semantic Web technology [3] are the expansion of machine information processing, namely: intelligent processing based on semantic technologies. In the most general sense, the semantic identification of data fragment consists in establishing its connection with the element of knowledge description of domain and explicit definition of the connection semantics. Today, ontologies are widely used for representation of domain knowledge [4] in intelligent information systems (IISs). For example, domain ontology can be used to represent learning outcomes in that domain, if individual fragments of a potential employee's resume are related to the concepts of such an ontology by relations that are also formalized by the same ontology [5].

Therefore development of information system for validation of non-formal learning outcomes is oriented on optimization of the process of combining the market of educational services and the labor market with the help of domain ontologies. In this research we take into account possibility of easy adaptation to changes in domain knowledge and use of modern standards of knowledge representation and processing.

Use of Semantic Web technologies for semantic identification of non-formal learning outcomes

One of the most popular projects related to the distributed knowledge processing is the Semantic Web. The Semantic Web conception of the was suggested by Tim Berners-Lee, one of the founders of the World Wide Web and

the head of the consortium W3C [6] in the work "Semantic Web" [7]. The Semantic Web is a Web of data. The aim of this project is to transform the whole set of available information resources, accessible through the Web, into a distributed heterogeneous knowledge base. The general idea of the Semantic Web concept is to organize such representation of data in the network that allows not only visualization, but also effective automatic data processing by software from independent manufacturers. The main components of Semantic Web are ontologies, Web services and software agents. The following open knowledge representation standards were developed for their presentation within the Semantic Web, such as the OWL ontology representation language [8], the RDF metadata description standard [9], Web services [10] and the query language SPARQL [11] for this formalized knowledge. The Resource Description Framework (RDF) is a data model for representing information about World Wide Web resources.

Today, the Semantic Web project is actively evolving: new languages, standards and tools are emerging, and existing ones are being improved. Therefore, it is advisable in the process of developing an intelligent applied system based on the use of Web resources, focus on these results and create semantic Web-services that can effectively use all the benefits of the new information environment. Semantic Web services expand the concept of ordinary Web-services in terms of the use of semantic information, namely ontologies and semantic markup for both application and system needs. The use of ontological analysis provides the ability to transfer knowledge to new applications, the automated export of information from semantically marked information resources (IRs) and the development of common terminological framework for interaction between various IRs and IISs. Semantic Web technologies are easily integrated with other Web technologies, such as intelligent Wiki (example – Semantic MediaWiki) [12]. Ontologies provide effective means of knowledge representation for distance learning [13].

We take into account all these factors in development of IIS for validation of non-formal learning outcomes aimed on automated generation of user passport of acquired competencies, search for vacancies and matching of competencies with vacancy requirements based on the ESCO model [14].

The following Semantic Web technologies are used in the development of this system:

- availability of personal software agents for potential employers and job seekers to personalize the interaction;
- use of semantic services for support of registration and search of vacancies and resumes, matching of vacancies and resumes at the semantic level, search of educational services capable to provide some specified qualification or education and comparison of training courses and learning programs with professions;
- ensuring the relations between professions, knowledge, skills, competencies and qualifications, as well as between their characteristics formalized with the help of ESCO ontology, various educational ontologies and domain ontologies.

Means of creating IIS for validation of non-formal learning outcomes

We use functional modeling IDEF0 [15] to determine the main functionality of proposed IIS. IDEF0 is a functional modeling methodology and graphical notation designed to formalize and describe business processes. A distinctive feature of IDEF0 is its emphasis on the subordination of objects. IDEF0 considers the logical relations between works, not their temporal sequence (workflow) [16].

Fig. 1 shows the functional model IDEF0 of IIS aimed on validation of non-formal learning outcomes.

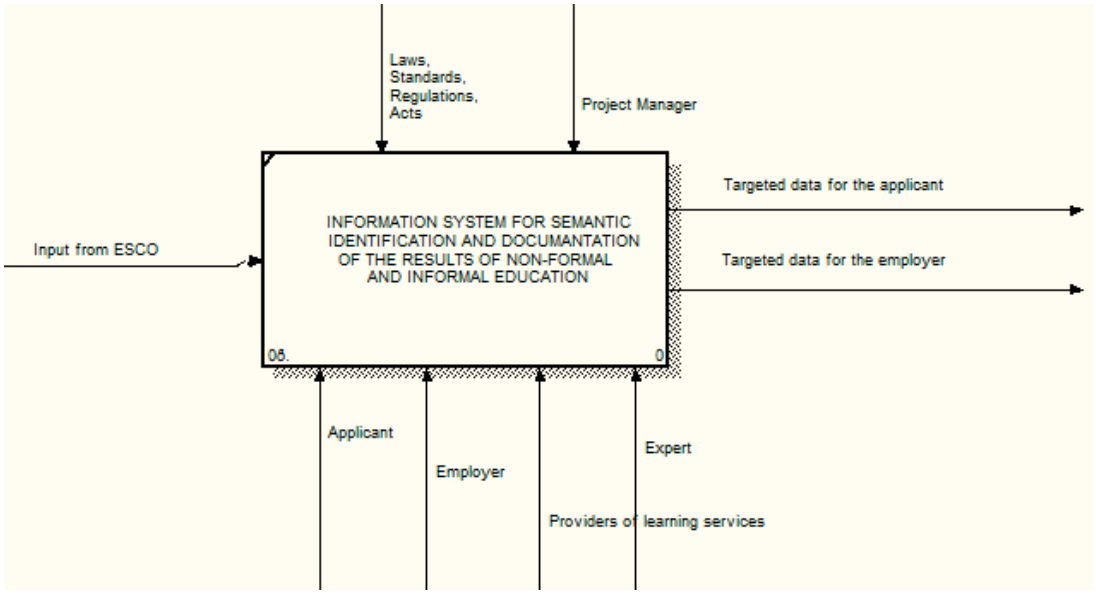

Fig. 1. Functional model of the IDEF0 system

This functional model consists of the following main components:

- *main block* is IIS based on computer ontologies. We use ESCO as a prototype of computer ontology;
- ontology of the multilingual skills classifier ESCO used as *input data*;
- *elements of management* that directly affect the development of architecture and technical task of the IIS are: "Standards, Regulations, Acts, Regulations, Standards" and "Project Manager";
- *subjects* of developed system: the applicant, the employer, the employment center, the expert and the provider of educational services;
- *output data* of the system are targeted data for the applicants and for the employers.

After developing the functional model IDEF0, we move on to the next step, namely to design of data flow diagram of the DFD system. We do it by decomposition of the functional model IDEF0. DFD data flow diagram is a design model with a graphical representation of "data flows" into IIS. This is the so-called graphical structural analysis methodology that describes the external to the system data sources and recipients, logical functions, data flows and accessible data storages. Data flow diagrams contain four types of graphical elements:

- processes of data transformation within the described IIS;
- data storages (repositories);
- external to the system entities;
- data flows between system elements.

Decomposition of the functional model is shown in Fig. 2.

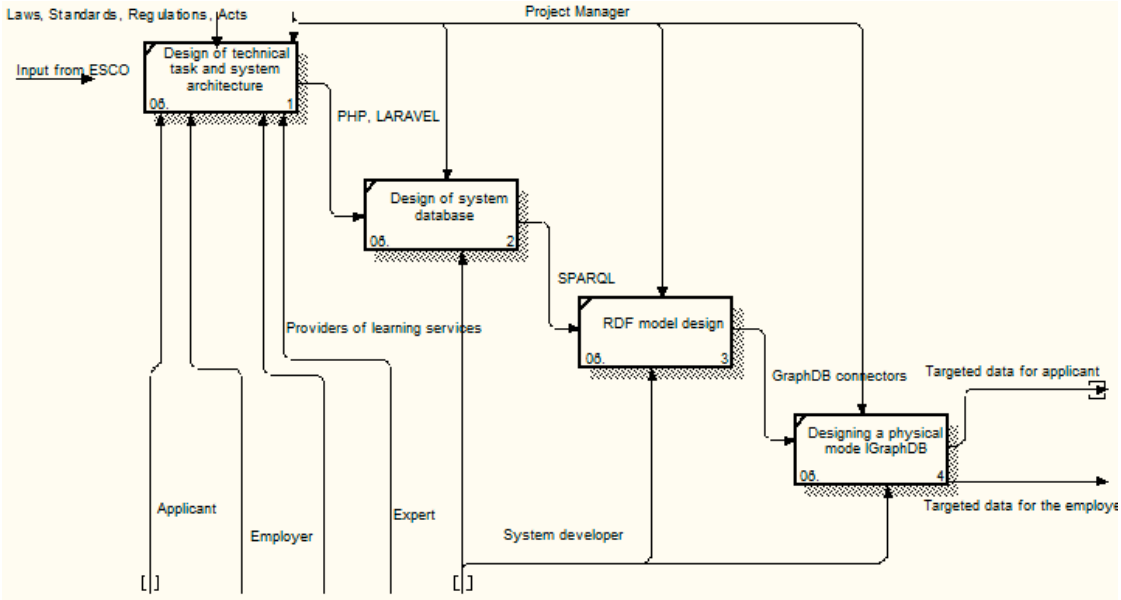

Fig. 2. Decomposition of the functional model of the system

Decomposition of the functional model of the developed system contains four blocks:

- design of technical task and system architecture;
- design of system database ;
- RDF model design;
- design of physical model GraphDB.

The implementation of the declared functions is ensured by use of small set of universal tools for domain information processing that support view information about instances (professions, skills, competencies, etc.), search of instances with the necessary properties and editing of instances.

Nowadays, domain ontologies are represented by the language OWL (Web Ontology Language) and its dialects [9-12] developed Semantic Web. OWL is based on descriptive logics and allows to describe domain concepts (classes and individuals) and the relations between them.

RDF (Resource Description Framework) is a formalism of the description of interconnected entities. It is designed by W3C to ensure metadata compatibility through common semantics, structure and syntax. RDF repositories are used in solving problems where data may have an unpredictable number of connections. RDF defines a common metadata architecture. The basis of RDF is data representation of in the form of statements (triplets) "subject-predicate-object" that describe the directed relation from subject to object. Subjects, objects and predicates are identified with use of the Uniform Resource Identifier (URI) [17] that generalizes the URL conception. The same URI can be used in different positions in different RDF triplets: to be both a subject, and a predicate, and an object; thus triplets form a kind of graph called RDF graph.

The SPARQL query language is used for processing of RDF data. It is one of the base technologies of the Semantic Web. SPARQL is a language for queries to data represented by the RDF model, as well as a protocol for

transmitting these queries and responses to them recommended by W3C consortium. Representation of SPARQL endpoints is a recommended practice for publishing data on the World Wide Web. SPARQL provides queries to mandatory and optional graph templates with conjunctions and disjunctions. The result of SPARQL queries can be represented as result sets or RDF graphs [18].

SPARQL queries include a set of triplet templates, which is called the main graph template. Triplet templates are similar to RDF triplets, except that each subject, predicate and object can be variable. The main graph template corresponds to subgraph of RDF data where RDF terms of this subgraph can be replaced by variables, and the result is an RDF graph equivalent to subgraph.

According to [19], the development of IIS for validation of non-formal learning outcomes that involves the use of ontological knowledge bases is performed in four stages:

1. creation of ontology scheme;
2. integration of ontology to RDF-repository;
3. development of application architecture;
4. development of user interface.

The first stage of system development deals with creation of ontology as a means of IIS knowledge representation. Ontology scheme forms a system of finite sets of concepts and statements that provide the basis for building of classes, objects and relations. This scheme determines the semantics of domain and helps to establish links between the values of its elements.

The ontology of the European classifier ESCO is used for the system operation. According to purposes of IS design this ontology becomes the source of knowledge about competencies and professions [20]. Open software tool Protege is used to modify the original domain ontology and store it. The Protege platform [21] supports two main ways to model ontologies using the Protege -Frames and Protege-OWL editors. Ne04j graph database management system is used to visualize this ontology.

Figure 3 shows an enlarged graph of the ontology loaded from the ESCO with description of its main relations. This enlarged graph contains four main classes that are directly involved into ontology processing, : "Skill", "Occupation", "NodeLiteral" and "Label". Others classes are either descriptive or conceptual. “Skill” class represents skills in the ontology, “Occupation” class represents professions, the “NodeLiteral” class describes professions and skills, and the “Label” class represents the names of professions and skills [22].

The second stage of IIS development deals with integration of ontology with RDF repository. Database of semantic graphs GraphDB is used to store the ontology obtained at the first stage. GraphDB is a database of graphs that is compatible with the RDF and SPARQL specifications. It supports open APIs based on the RDF4J project, allows to publish linked data into the Web. It is used to link data from different sources, to index them for semantic search and to enriches this information by text analysis to build large knowledge storages.

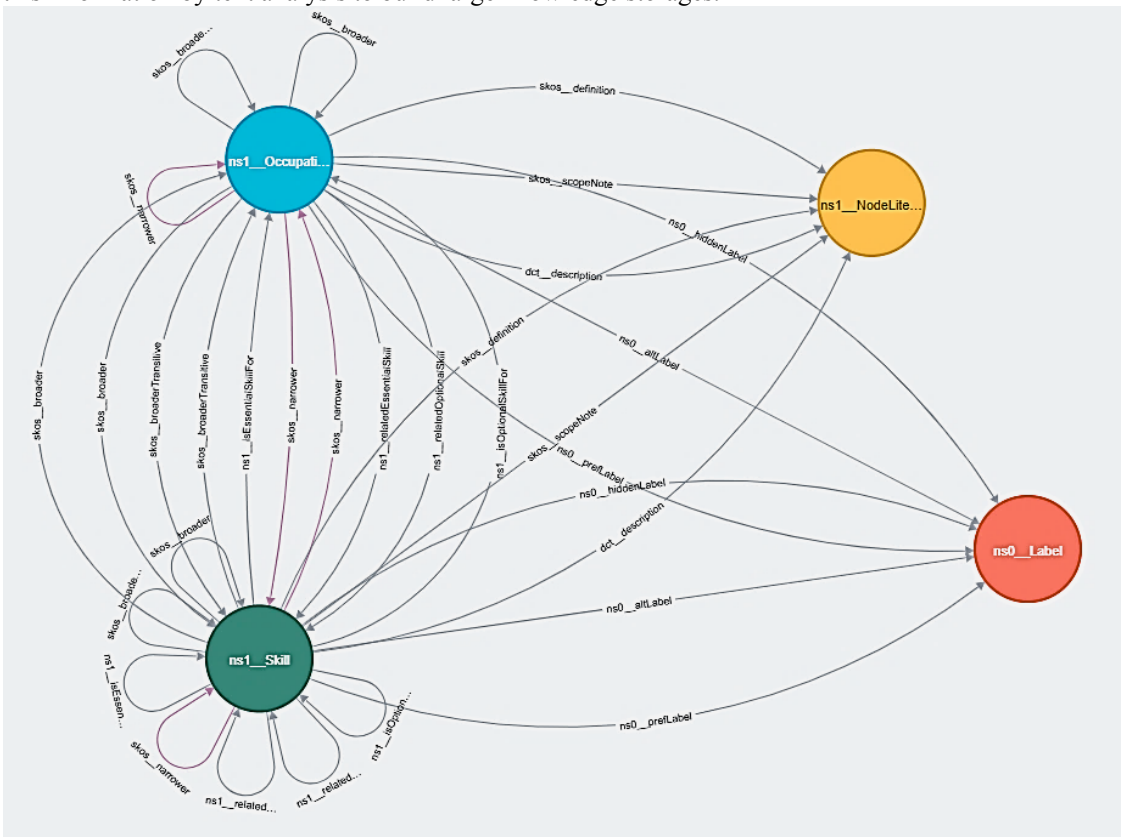

Fig. 3. The enlarged graph obtained from Protege ontology editor

Work with the RDF repository requires to access data through the SPARQL query language. Fig. 4 shows an example of a SPARQL query that allows to define a list of skills.

The query consists of two parts: the SELECT condition specifies the variables that should be displayed in the query results, and the WHERE condition provides the basic graph template that defines conditions for graph data. In this example the basic graph template consists of a single triplet template with variables "?Skill", "?O" and "?P" in object position. As a result of request execution the list of skills is generated.

```
PREFIX skos-xl: <http://www.w3.org/2008/05/skos-xl#>
PREFIX skos:  <http://www.w3.org/2004/02/skos/core#>
SELECT DISTINCT ?skill ?o ?p #?label
WHERE {
  ?skill a esco:Skill.
  ?skill skos:prefLabel ?o.
  ?skill skos:altLabel ?p.
    FILTER (lang(?o)='en' )
    FILTER (lang(?p)='en' )
```

Fig. 4. Example of SPARQL query

Connectors are used to speed up the work with the storage. GraphDB connectors provide fast normal and faceted search (aggregation) that usually is implemented by an external component or service such as Lucene.

Creation of connector in GraphDB requires to give it name and fields. The corresponding form is used for this purpose (fig. 5).

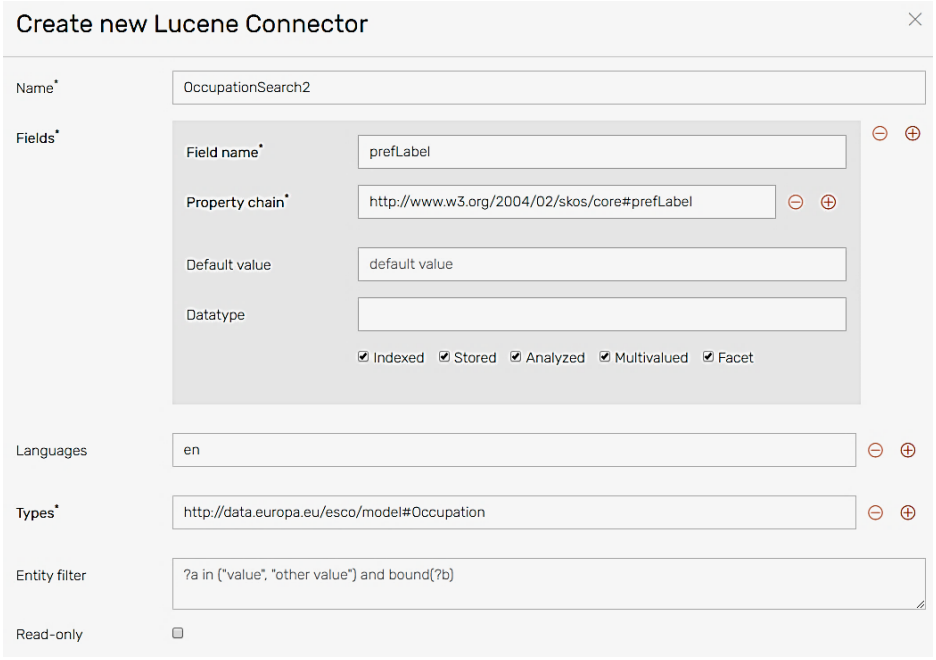

Fig. 5. Form for connector creation in GraphDB

GraphDB connectors have the additional benefit of automatically updating data with GraphDB repositories. Connectors provide entity-level synchronization where an entity is defined as something with unique identifier (URI) and a set of properties and property values. In terms of RDF, this definition corresponds to set of triplets with the same subject. In addition to simple properties (defined by one of triplet), connectors support property chains. A property chain is a sequence of triplets where the object of previous triplet is the subject of the next triplet.

This connector receives the text in the field, searches by the specified labels and returns a reference to the class that corresponds to the specified query.

GraphDB connectors have the following properties [23]:

- support for the index that is always synchronized with data stored in GraphDB;
- availability of several independent copies for each repository;
- full-text search with use of native Lucene queries;
- selection of fragments of search queries as a result of search;
- sort by any pre-configured field;
- swapping results with use of offset and limit;
- choice of analyzer;

- removal of HTML / XML tags in literals, etc.

After the stage of ontology integrating into the RDF repository we move on to creating the internal architecture of IIS that uses this ontological repository. The main requirements for building an application architecture are to provide access to the knowledge base. The main users of the system are employers, job seekers and experts. Users interact with the IIS (obtain access to the RDF repository) through the Web server that process requests to the repository and results of these requests. Fig. 6 shows general model of the architecture of IIS for validation of non-formal learning outcomes.

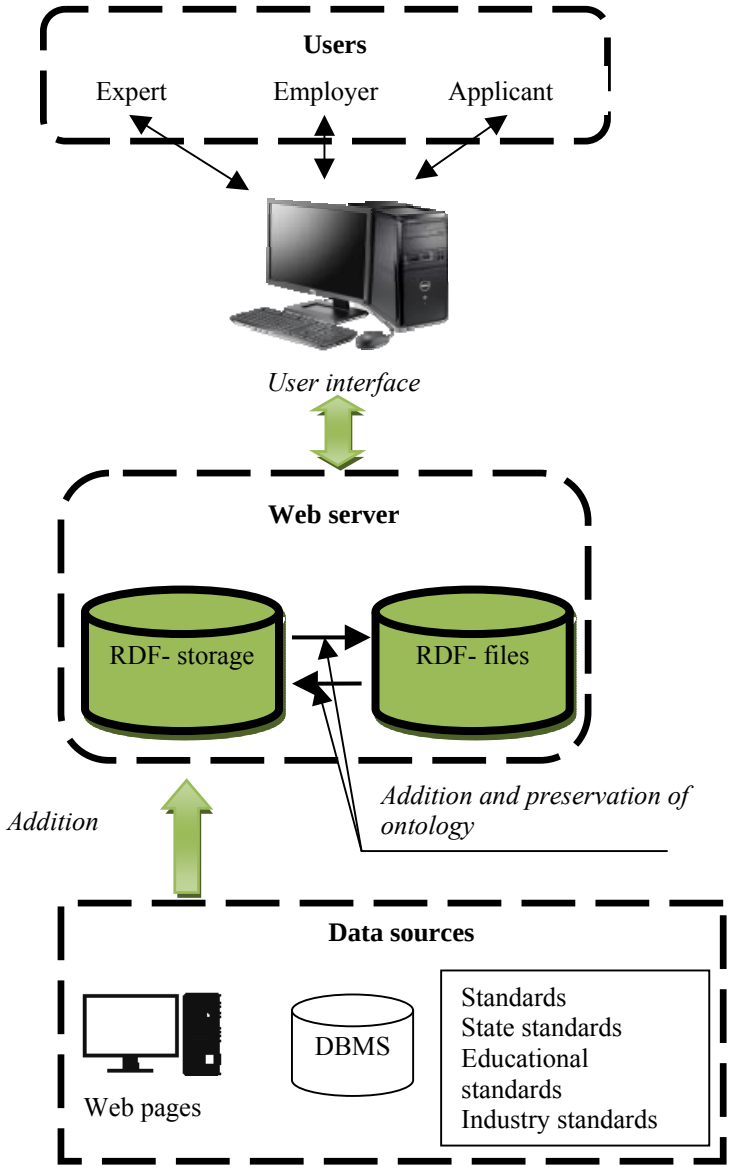

Fig. 6. General architecture of IIS

PHP and Java-Script tools can be used for creation of the Web server. PHP language provides the programmer with tools for quick and efficient solving of problems. Compared to PHP, Java Script has a reduced level of security due to free access to the source code of popular scripts and many small errors at every stage of work. Most of them are corrected easily but their presence allows to consider this language less professional than others. Therefore, PHP is preferred for this project. Laravel and Symfony PHP frameworks can be used to development of the Web applications. Table 1 shows their main features.

Table 1. Comparative analysis of Laravel and Symfony software tools

Features	Symfony	Laravel
Number of packages	7500	16 900
MVC	+	+
Installation through the composer	+	+
Full text search	+ (ElasticSearch)	+ (ElasticSearch)
Support CI, QA	-	PHPUnit

The Symfony framework provides the Web developers with built-in testing functionality, as well as the ability to work with components and reuse code. Symfony 4.2 loads the REST API in 2 ms. This makes Symfony the fastest PHP framework. Symfony adapts to any project requirements. This framework includes Event Dispatcher, thanks to it Symfony is fully configured. In addition, all Symfony components are controlled independently. But Laravel provides the Web applications with more functionality. Laravel has a good template engine that performs a lot of common tasks.

In this project we use the Laravel framework due to its increased security and readiness to install plug-ins and libraries. This framework has features such as RESTful-routing, caching, user management and authentication. Therefore Laravel speeds up the development process.

The fourth stage consists in development of user interface to ensure the user's dialogue with IIS. The most relevant tools for solving the tasks are React, Angular [25], Vue.js [26].

React is an open JavaScript library for creating user interfaces designed to solve the problems of partial updating of the content of the Web page for development of one-page applications. React allows developers to create large applications that use data that changes over time without reloading the page. React framework has the following advantages: speed, simplicity and scalability. React conception is built on the representation of user interface by separate self-contained parts – components that are quite simple to maintain and extend.

Angular is open-source front-end framework developed by Google's Angular Team, as well as the community of private developers and corporations. It is written in TypeScript. The Angular framework supports the MVW conception and has the following advantages: fast code writing, fast testing of any software part and two-way data binding.

Vue.js is a JavaScript framework that uses the MVVM template to create user interfaces based on data models through reactive data binding. Vue.js implements two-way data binding, server-side visualization, Vue-cli (scaffolding tool for quick start) and optional JSX support. After analysis of these tools and comparison of their features with specifics of solving problem, we choose React framework for development of user interface due to its effective work with the tools selected in the previous stages.

User interface of IIS for validation of non-formal learning outcomes consists of a functional panel and a working window. The work form offers the user to search by the set of parameters: the name of the profession, basic and additional skills.

The list of search parameters is specified in the user electronic portfolio. Such e-portfolio is a "business card" of a searcher. It contains data on various aspects of his/her activity, professional development, educational activity and personal data. The list of added parameters is stored in the ontology and can be supplemented.

The user also has the opportunity with the help of developed form to add new profession according to the specified parameters. To add a new profession, the user enters the profession name, selects basic and additional skills and clicks button “add a profession”. The list of parameters for skill search is stored in the ontology. As a result, a request to the connector that makes changes to the ontology is generated.

Conclusions

Analysis of the development of modern knowledge-based IIS shows the feasibility of use the Semantic Web technologies in the development of software tools aimed at combining the market of educational services with the labor market. The main advantage of the Semantic Web initiative is an expansion of machine information processing executed on semantic level.

In this work we propose an information system for validation of non-formal learning outcomes that uses ontology of the multilingual skills classifier ESCO as an external source of knowledge. We use functional modeling of IDEF0 to determine the main functional capabilities of this IIS. Decomposition of proposed model provides identification of its main blocks: design of the technical task and system architecture; database design; RDF model design; design of physical model GraphDB.

Development of this IIS includes the creation of ontology scheme, integrating the resulting ontology into the RDF repository, development of application architecture and user interface. We compare existing methods and software tools for knowledge representation and analysis and their suitability for solving of considered problem. As a result of analysis, database management system Neo4j that supports the graph database GraphDB; operation of connectors and SPARQL queries to the data placed in the RDF repository are used for development of proposed system. Task-aimed comparative analysis of PHP frameworks (criteria as security, readiness to install plug-ins and libraries, support for the concept of MVC (model-view-controller)) shows the advantages of Laravel. React framework is selected for user interface development due to its effective work with previously selected tools.

References

1. Semantic Web. www.w3.org/standards/semanticweb/.
2. Semantic Web technology. https://www.sciencedirect.com/topics/computer-science/semantic-web-technology.
3. Kifer, M., De Bruijn, J., Boley, H., & Fensel, D. (2005, November). A realistic architecture for the semantic web. In International Workshop on Rules and Rule Markup Languages for the Semantic Web (pp. 17-29). Springer, Berlin, Heidelberg. https://www.researchgate.net/profile/Michael-Kifer/publication/44086469_A_Realistic_Architecture_for_the_Semantic_Web/links/0deec52aa0e361c7b7000000/A-Realistic-Architecture-for-the-Semantic-Web.pdf

4. Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology?. In Handbook on ontologies (pp. 1-17). Springer, Berlin, Heidelberg.
5. Pryima S., Rogushina J. Development of methods for support of qualification frameworks transparency based on semantic technologies. *Information Technologies and Learning Tools*. 2017. Vol. 59. N 3. P. 201–210.
6. W3C Semantic Web Activity. https://www.w3.org/2001/sw/.
7. Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. Scientific American, 284(5), 34-43.
8. OWL Web Ontology Language. Overview. W3C Recommendation: W3C, 2009. http://www.w3.org/TR/owl-features/.
9. Lassila O., Swick R. Resource Description Framework (RDF) Model and Syntax Specification, W3C Recommendation. URL: http://www.w3.org/TR/REC-rdf-syntax.
10. Sheila A. & Mcilraith Tran Cao Son & Honglei Zen. Semantic Web Services, Stanford University. URL: http://www.ksl.stanford.edu.
11. Pérez, J., Arenas, M., & Gutierrez, C. (2006, November). Semantics and Complexity of SPARQL. In International semantic web conference (pp. 30-43). Springer, Berlin, Heidelberg. https://link.springer.com/content/pdf/10.1007/11926078_3.pdf.
12. Krötzsch, M., Vrandečić, D., & Völkel, M. (2006, November). Semantic mediawiki. In International semantic web conference (pp. 935-942). Springer, Berlin, Heidelberg.
13. Gladun, A., & Rogushina, J. (2010). Distant control of student skills by formal model of domain knowledge. International Journal of Innovation and Learning, 7(4), 394-411.
14. ESCO (European Skills, Competences, Qualifications and Occupations) URL: https://ec.europa.eu/esco/portal/home.
15. Dašić, V. Š. P., & Labović, R. J. D. (2009). Functional and Information Modeling of Production Using IDEF Methods.https://link.springer.com/content/pdf/10.1007/11926078_3.pdf
16. Dorador, J. M., & Young, R. I. (2000). Application of IDEF0, IDEF3 and UML methodologies in the creation of information models. International Journal of Computer Integrated Manufacturing, 13(5), 430-445.
17. Masinter, L., Berners-Lee, T., & Fielding, R. T. (2005). Uniform resource identifier (URI): Generic syntax. Network Working Group: Fremont, CA, USA. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.731.6302&rep=rep1&type=pdf.
18. Huang, J., Abadi, D. J., & Ren, K. (2011). Scalable SPARQL querying of large RDF graphs. Proceedings of the VLDB Endowment, 4(11), 1123-1134.
19. Shenchenko YE. & Shenchenko A. & Bogdan. Features of the construction of applied software systems based on ontological knowledge bases. ACS and automation devices: all-in. inter. scientific and technical Sat. Kharkiv: Publishing House of KNURE. 2011. Vol. 157. P. 36–41. (in Ukrainian).
20. Pryima S., Rogushina J. Ontological approach to qualifications matching on base of competences: model and methods. *Науковий вісник НГУ*. 2017. N 6. P. 162–168. Available from: http://www.nvngu.in.ua/jdownloads/pdf/2017/06/06_2017_Rogushina.pdf.
21. Protégé. W3C Semantic Web. URL: http://www.w3.org/2001/sw/wiki/Protégé.
22. Rogushina J., Priyma S. Use of competence ontological model for matching of qualifications. *Chemistry: Bulgarian Journal of Science Education*. 2017. Vol. 26. N 2. P. 216–228.
23. About GraphDB. https://graphdb.ontotext.com/documentation/free/about-graphdb.html
24. Pryima S.M., Rogushyna Yu.V., Strokan O.V. Use of semantic technologies in the process of recognizing the outcomes of non-formal and informal learning. *CEUR Workshop Proceedings*. 2018. Vol. 2139. P. 226–235. http://ceur-ws.org/Vol-2139/226-235.pdf
25. Angular.URL:https://uk.wikipedia.org/wiki/Angular_(%D1%84%D1%80%D0%B5%D0%B9%D0%BC%D0%B2%D0%BE%D1%80%D0%BA) (дата звернення: 22.02.2020).
26. Vue.js. URL: https://uk.wikipedia.org/wiki/Vue.js.

Про авторів:

Serhii M. Pryima,
Doctor of Pedagogical Sciences, Professor,
Professor at the Department of Computer Science.
Number of scientific publications in Ukraine – 43.
Number of international scientific publications – 15.
http://orcid.org/0000-0002-2654-5610,

Oksana V. Strokan,
Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Computer Science.
Number of scientific publications in Ukraine – 20.
http://orcid.org/0000-0002-6937-3548,

Dmitro V. Lubko,
Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Computer Science.
Number of scientific publications in Ukraine – 24.
http://orcid.org/0000-0002-2506-4145,

Yuliia O. Lytvyn,
Research Assistant at the Department of Computer Science.
Number of scientific publications in Ukraine – 16.
http://orcid.org/0000-0001-5797-3151,

Nadia I. Bilyk,
Doctor of Pedagogical Sciences, Associate Professor,

Professor of кафедри педагогічної майстерності та інклюзивної освіти
M. V. Ostrogradsky Poltava Regional Institute of Postgraduate Pedagogical Education, Ukraine
Number of scientific publications in Ukraine – 45.
Number of international scientific publications – 13.
http://orcid.org/0000-0003-2344-5347.

E-mail: pryima.serhii@tsatu.edu.ua,
oksana.strokan@tsatu.edu.ua,
dmytro.lubko@tsatu.edu.ua,
litwin.yu1994@gmail.com,
bilyk@poippo.pl.ua