UDC 004.853, 004.55

EXPERIENCE OF SEMANTIC TECHNOLOGIES USE FOR DEVELOPMENT OF INTELLIGENT WEB ENCYCLOPEDIA

P.I. Andon[a][0000-0002-2204-1554], *J.V. Rogushina*[a] [/0000-0001-7958-2557], *I.Y. Grishanova*[a][0000-0003-4999-6294], *V.A. Reznichenko*[a][0000-0002-4451-8931], *A.M. Kyrydon*[b][0000-0003-4375-5350], *A.V. Aristova*[b][0000-0002-9207-778X], *A.O. Tyschenko*[b][0000-0002-9847-2991]

[a]Institute of Software Systems, National Academy of Sciences of Ukraine, 44 Glushkov Pr., Kyiv, 03680, Ukraine.

[b]State Scientific Institution "Encyclopedic Publishing House", Bohdan Khmelnytsky street, 51-A, Kyiv, 01030, Ukraine.

We consider the conceptual principles of the Great Ukrainian Encyclopedia development and analyze the main advantages of its portal version – e-VUE. This portal version requires the relevant technological platform with high-level expressiveness and efficiency. Development of e-VUE with complex knowledge structure causes the study of modern Web-technologies and matching their capabilities with the specific requirements of encyclopedia. Therefore we consider in details the project idea and purposes, the specifics of e-VUE information representation (number and volume of articles, content types and sources, relations between content elements and their properties, business processes of article publication, design requirements) and processing (navigation, requests, integration and matching means) to estimate the Wiki-technology feasibility use (namely – opportunities and restrictions of MediaWiki technological platform). Also attention is paid to the means of information security and semantic-based establishment of content access rights. The results of such analysis show that MediaWiki with semantic plug-in s can be used as a base for e-VUE portal construction but needs in extension of its functionality based on modern knowledge management tools and formats of their interoperable representation. Now we are oriented on use of ontological analysis and Semantic Web standards (OWL, RDF, SPARQL). In terms of the Semantic Web, e-VUE is a distributed database with heterogeneous types of information objects that can be used both by humans and by external software (e.g. by intelligent agents). The main principles of e-VUE knowledge base organization are grounded on Wiki categories and semantic properties of Semantic MediaWiki that are used for formal representation of the e-VUE typical information objects (TIO). Such approach allows to express TIO relations, properties and characteristics that can be used by logical inference and semantic retrieval to improve the navigation between portal elements and integration of data from various Wiki pages.

.

Great Ukrainian Encyclopedia, its purpose and features

Since 2013, the process of creating the Great Ukrainian Encyclopedia (VUE) has been actively unfolding in Ukraine. Its publication is planned for 2013–2026 in accordance with the Decrees of the President of Ukraine № 1/2013 “On the Great Ukrainian Encyclopedia” and № 7/2015 “Issues of preparation and publication of the Great Ukrainian Encyclopedia” through the State Scientific Institution “Encyclopedic Publishing House”. The Great Ukrainian Encyclopedia by the nature of information is a universal encyclopedia that represents the national idea and national culture. VUE provides the most important reference materials from almost all areas of knowledge and production activities. Universal encyclopedias are oriented on the widest range of readers with different levels of training. In this situation, the problem of providing information becomes important.

Conceptually, the Great Ukrainian Encyclopedia involves the creation not only of a modern compendium of human knowledge based on the experience of leading national universal encyclopedias, but also systematization and representation in an accessible forms (printed and electronic versions) of previous generations results. The conception of the Great Ukrainian Encyclopedia is focused on modern scientific understanding of the world picture, the history of human civilization, the contribution of the Ukrainian people. It represents the most prominent events and figures of world history, science, culture, arts and other spheres of human activity; interpreted the most important discoveries of the human mind and the creations of human hands. VUE describes social institutions and economic structures that influenced the world and peoples; the movement of political and cultural ideas and the dynamics of scientific ideas [1]. Radical changes in the political map of the world, modern teachings in the field of natural sciences, philosophy, history, political science, sociology, culturology, aesthetics, etc. are also performed in this encyclopedia.

The development of encyclopedic publications is directly related to their presentation on the Web. Online versions of encyclopedias are the most accessible to the widest range of users, they can be regularly updated and contain current facts [2]. In addition, the use of modern information technology allows to connect various multimedia resources, links to original sources and related articles. This greatly facilitates the perception of information and reduces the search time.

The peculiarity and uniqueness of the VUE project relies on the fact that two related but separate information products are created – the printed version of the encyclopedia and its intelligent portal version e-VUE [3]. This portal is

a Web-resource that differs from the printed analogue by larger volume of articles, semantically expanded functionality for navigation and search, non-linear organization of content and the presence of interactive and multimedia objects.

Portal version of VUE as an innovative information resource

One of the main goals of e-VUE development is to provide users with universal access to the content of the encyclopedia. e-VUE is built as an intelligent information system [4] that contains knowledge base (KB) and services for processing knowledge from this KB. e-VUE is an innovative information resource (IR) based of modern knowledge-oriented technologies and original software solutions. Its development requires the study of relevant models, methods of information representation and processing for implementation.

The use of modern knowledge management technologies oriented on the Web provides users with much more opportunities for information analysis and search. Unfortunately, now even the most developed and popular online encyclopedias practically do not use the methods and means of knowledge representation that already became actual standard for other types of intelligent Web applications (e.g., retrieval and recommending systems, e-learning applications). For example, the well-known Wikipedia [5] is based on Wiki-technologies [6], but does not use their semantic extension. As a result, the knowledge contained in such IRs, cannot be processed by modern data analysis tools. This significantly narrows both the functionality of the encyclopedic publications and the scope of their use. But the choice of means of semantization of Wiki-technologies is a separate scientific problem and should take into account the specifics of the task that these technologies have to solve [7].

As a result of analysis of existing software tools and planned functionality of encyclopedic portal [8]the open source technological platform MediaWiki [9] and its semantic extension Semantic MediaWiki (SMW) [10] were selected for e-VUE development. This semantic extension provides formal definition and processing of the semantics of links between Wiki-pages taking into account the standards of Semantic Web [11].

Main features of e-VUE from the point of view of information technologies are:

- this IR provides verified facts and expert knowledge in the form of peer-reviewed author's articles;
- preparation of articles for publication has a clear technological algorithm;
- performers that take part in content preparation are territorially distributed;
- information represented by this IR is mostly unique and has certain licenses on the rights of use;
- content of e-VUE consists of natural language text, multimedia elements (images, videos, audio, etc.) and links to other trusted sources;
- Wiki-technologies used for content representation are expanded by specialized software solutions that use modern knowledge management technologies, Semantic Web standards, elements of artificial intelligence, ontological analysis, intelligent search, etc.

Formulation of the problem

Development of the portal version of the Great Ukrainian Encyclopedia requires the use of modern intelligent Web-technologies that provide representation and analyzes of the complex knowledge system inherent to VUE. This necessitates the study of modern Web-technologies and comparing their capabilities and expressiveness with the specific requirements arising from the e-VUE conception. We propose to develop e-VUE on base of scalable and flexible Wiki technology and to enrich open source Wiki software such as MediaWiki with the Semantic MediaWiki semantic plug-in, ontological analysis and Semantic Web technologies. The main directions of such semantic approach are substantiated in the work, and the advantages of their use are considered on examples.

Technological and scientific basis of e-VUE

Wiki-technology is a Web-oriented technology developed for building of distributed IRs. This technology allows developers of IR to add new materials and edit existing content without the use of additional software and specific competencies dealt with knowledge engineering. Changes of the IR content made by users become active immediately, while the ability to revert to previous versions is provided [12].

Wiki-text is a format of Wiki-pages that are based on simplified markup language used to indicate various structural and visual elements of content. Every Wiki-page can contain not only text, but also a variety of multimedia content (images, audio, video). Each Wiki system has its own style and syntax depending on the implementation. The use of the Wiki-environment provides such important features as joint content editing, flexible system of user rights, scalability, focus on the Web.

MediaWiki. A widely used solution for creating of Wiki-resources is the technological platform MediaWiki. It was designed for development of the most famous Wiki-resource – Wikipedia. MediaWiki is written by on PHP and MySQL It includes an editor for text and multimedia content processing and provides an interface for working with database of Wiki-pages. If users need to use graphics, audio and video on Wiki pages then MediaWiki provides download functionality of the relevant files. System supports separation of access rights for system administration and access to content. The use of various plug-ins allows users to extend the functions of the system according to their own needs – for example, show mathematical formulas, export and import of various data formats.

MediaWiki's benefits include free software, multi-user distributed support, and deployment on both the institution's local network and the Web. The architecture, basic features and limitations of MediaWiki are analyzed in detail in [13]. Categories of MediaWiki used for content structuring allow to group pages with similar themes (for

example, "Countries", "Writers", "Holidays"), origins, etc. MediaWiki templates whose content can be embedded in other pages are used to display duplicate elements of pages according to values of their parameters.

The practical use of IRs based on MediaWiki requires solving problems related to content consistency, search efficiency and knowledge reuse. Ways to do this are provided by the use of Wiki platform semantization.

Semantization of MediaWiki. Plag-in Semantic MediaWiki (SMW) is a MediaWiki semantic extension that allows the user to automatically integrate information from different Wiki pages, generate responses to complex semantic queries, build knowledge bases and visualize their results that reflect knowledge of semantic properties and categories of information objects, perform logical inference, etc., i.e. to process information at the level of knowledge. To do this, SMW uses categories, semantic properties, and queries. Semantic properties allow to describe the semantics of relations between two pages (for example, "Place of birth") and between a page and data (for example, "Year of birth"). They can be used as template parameters. Queries use a simple query language where query conditions are specified by categories, values of semantic properties and conditions dealt with these values. Such queries can be embedded in Wiki pages.

SMW provides access interfaces to Semantic Web tools. Formal descriptions of one or more Wiki pages in OWL / RDF format [14] can be obtained from the Web-interface for external use. Users can also import data from OWL ontologies [15], display Wiki annotations on existing dictionaries (such as FOAF). Because SMW follows the OWL DL standard, the exported information can be reused in various external applications. In addition, SMW provides service to support SPARQL queries.

The main technological components of e-VUE

The main components of the e-VUE portal:

- MediaWiki kernel,
- database containing the main content;
- user interface that specifies the representation of Wiki page content and navigation tools in accordance with the needs of the encyclopedia;
- additional plug-ins that extend the functionality of MediaWiki.

MediaWiki kernel is a software that provides a basic set of functions for creation and operation of Wiki-based IR. It allows to generate Wiki-pages received by the portal users using the database contents. The main access point of MediaWiki is index.php that manages the most of the queries submitted to the server and can support various actions specified by URL query parameters (for example, viewing or editing of the page content).

e-VUE database. e-VUE project uses standard MediaWiki relational database MySQL. This database contains data about content (for example, page, revision, category, recent changes), about users (user, user_groups), about multimedia files (image, filearchive), caching (objectcache, querycache), internal means (job – to perform the task queue), as well as other additional data required for the plug-ins installed in e-VUE.

The project pays a lot of attention to multimedia content. All multimedia files (audio, video, images, etc.) have an extended metadescriptions that contain as technical characteristics, title, description, sources and copyrights of use, as well as a meaningful descriptions of the information that is also stored in the database. Many pages are accompanied by sound, i.e. contain links to corresponding file with audio dubbing.

User interface. e-VUE has an original user interface that is developed according to the encyclopedia specifics. This interface differs significantly from the traditional Wiki-interface. e-VUE design provides ergonomic integration of text, images, audio and video elements in the screen space for holistic visualization of multimedia content. Such solution should provide that, on the one hand, e-VUE style corresponds to the style of the printed edition of the Great Ukrainian Encyclopedia, and on the other hand, e-VUE visualization of information satisfies the requirements to representation of the Web IRs (for example, access from mobile devices is supported).

e-VUE uses skin Foreground – specially configured set of design styles to define parameters of the Wiki pages design such as:

- location and form of navigation elements, information blocks and buttons,
- size, color and type of fonts for different types of content,
- background of pages and their elements.

Foreground is focused on content selection, supports customizable layouts, and has special predefined classes to support Semantic MediaWiki functionality. Use of this skin completely changes the layout of the external representation of the Wiki page and allows to add new interface elements.

In addition to the traditional elements, every Wiki-page corresponded to e-VUE article contains elements specific to the encyclopedic edition such as:

- link to the article author;
- text of article reference;
- the main semantic properties of the page specified by the corresponding template or set of templates;
- semantic markup of article content unified for all articles or for groups of similar articles;
- hyperlinks to pertinent multimedia objects;
- content structuring (categories of scientific spheres and subject domains).

All functional menus in e-VUE are carried to separate drop-down list located on the right side of the page. This list is displayed for registered users only. New items of navigation menu that allow portal navigating by taxonomies of e-VUE categories ("Personalities", "Nature", "Civilization", "Branches of knowledge") are created. Some special

references to the categories ("Authors", "Library", "Media files", “Semantic search ” are placed in the upper horizontal menu (Fig. 1). The extended basement contains links related to navigation, information on cooperation with VUE and means of communication with its developers.

With the help of Foreground the processing method of the first e-VUE page is separated from the rest of the portal pages. Such separation provides the development of unique functionality and design that demonstrates to users all important portal features (Fig. 2). The first page of the e-VUE contains information that is generated dynamically by executing of API requests. These are three blocks of new e-VUE articles – "Personalities", "Nature" and "Civilization". Each of them is represented at the first page by the name of the Wiki-page, an image, a short definition and a link to the full version of the article. Right column of the first page contains a block with dynamic banners that provide access to actual articles selected by the editorial board of VUE.

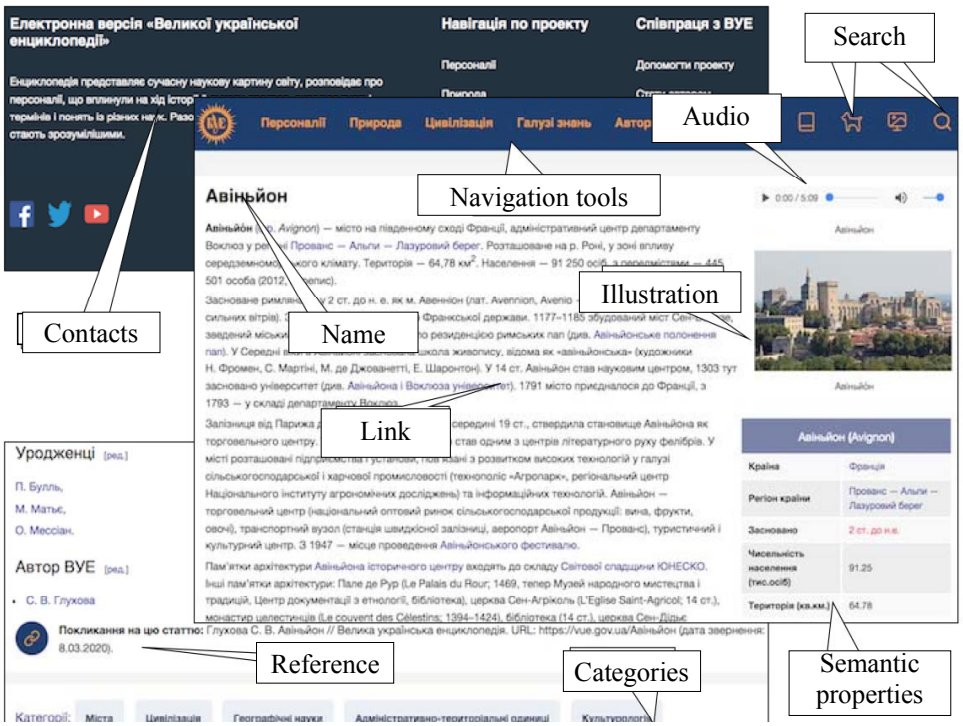

Figure 1. Base elements of e-VUE page

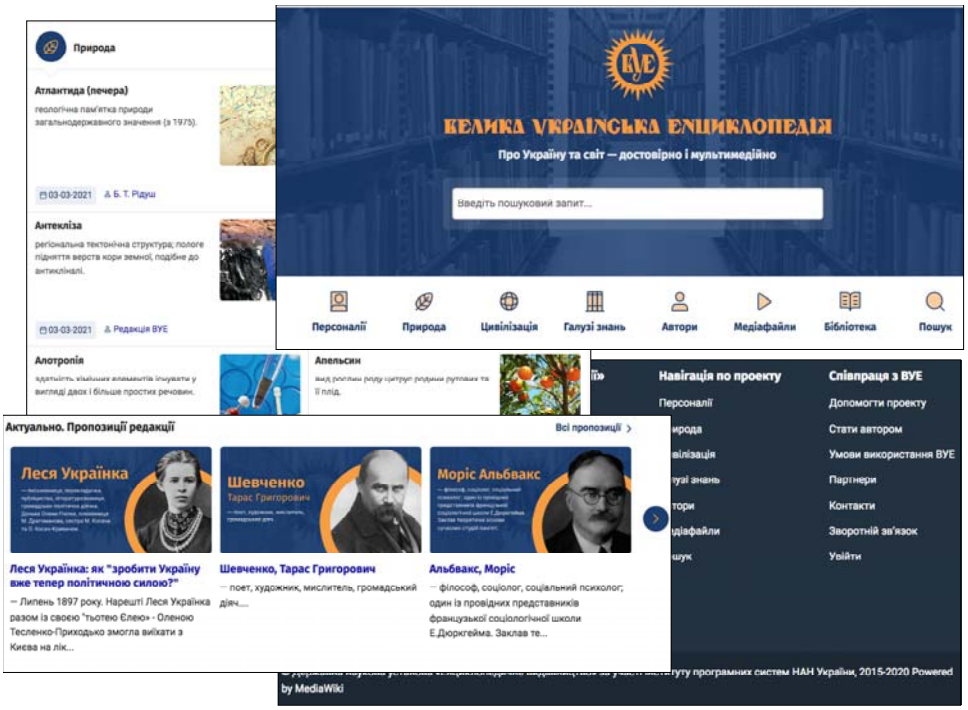

Figure 2. Functional elements of the first page of e-VUE

Extensions of e-VUE. Extensions of MediaWiki functionality used for creation of IR are provided by installation of additional plug-ins. Their composition is determined by this IR specifics [16]. The most important extensions used today in e-VUE:

- Semantic MediaWiki – a semantization plug-in that allows to define semantic properties of Wiki pages and use them for information retrieval and analysis;
- *ContactPageVUE* (original improvement of the ContactPage plug-in) – a form of feedback relevant to the needs of the e-VUE project;
- *Replace Text* – plug-in that allows administrators to perform global search and replace text on all regular Wiki-pages;
- *CategoryTree* – AJAX-component used to display the structure of Wiki categories [17];
- *EmbedVideo* – plug-in that provides the ability to insert video content [18];
- *PDF Handler* – plug-in for viewing pdf-files in image mode;
- *Shariff* – plug-in for adding of social media buttons.

Knowledge base of e-VUE

From the point of view of the Semantic Web conception [19], e-VUE can be considered as a distributed knowledge base (KB) that can be used by people and intelligent programs focused on working with the Web knowledge [20]. Therefore, it is advisable to describe in more detail e-VUE KB structure, expressiveness and specificity.

The structure of the KB takes into account the possibility of regular updating of information content (addition of new articles with various sets of structural elements, updating the previously presented materials), use of modern representation formats for multimedia illustrative material (drawings, pictures, photos, videos, audios, maps, diagrams, etc.) and expanding the functionality of the portal with further design improvements. In addition, KB of e-VUE should be used as a basis for the intelligent navigation and information retrieval on the portal, for satisfaction of the information needs of users taking into account their semantics and the specifics of physiological, technological and intelligent capabilities of different users.

Basic principles of e-VUE KB organization:

- the main tool of KB organization are the categories and semantic properties of SMW [21];
- fixed set of independent taxonomies (for example, "Branches of knowledge", "Geographical objects", "Article status") is used for the structured representation of portal information, and each Wiki-page can belong to any set of existing categories;
- typical information objects (TIO) of e-VUE are defined to integrate encyclopedic articles with similar sets of semantic properties and categories;
- the set of semantic properties of TIO in the process of improving the e-VUE KB is expanded and is not reduced (this fact provides the ability to process the previously described TIO without changes in their structure, and certain properties that are not used in practice are not removed but are not taken into account in new articles).

In the process of designing the e-VUE KB as a result of research of encyclopedic content and means of its representation the following entities are allocated:

- *categories* of Wiki-pages (the relations between categories are fixed by means of hierarchical relations “class-subclass” of different taxonomies – for example, taxonomies of knowledge branches and taxonomy of multimedia elements);
- e-VUE *typical information objects* characterized by constant set of categories and semantic properties, which values are defined for each copy of the TIO; all TIOs are divided into e-VUE *subjects* (authors of e-VUE articles, moderators of scientific directions, authors of illustrations and videos, editors, announcers, etc.) who by their actions can make changes of the encyclopedia content, and e-VUE *objects* (articles of the encyclopedia, multimedia elements, category pages and service Wiki-pages);
- *semantic properties* of TIO that connect the subjects and objects of e-VUE with semantic links;
- valid *areas of meaning* and *definition* of semantic properties that determine the scope of their use and the range of valid values;
- *characteristics* of semantic properties (such as symmetry, transitivity and reflexivity) that can be used for the logical inference of new knowledge;
- *semantic queries* that use categories and values of semantic properties as request parameters.

We develop the set of MediaWiki templates for support of e-VUE TIOs. These templates provide a unified representation of information and facilitate the input of values of semantic properties.

Parameters of the TIO templates are associated with the semantic properties of Wiki-pages that can be used in semantic queries (Fig. 3) [23]. It is very important to prevent the re-use of names in different meanings and the use of different names for the same meanings (for example, to introduce the category “Personal” instead of “Personality”, the category “Ukrainian cities” instead of “Cities of Ukraine”).

The expressive ability of Semantic MediaWiki is sufficient to represent knowledge structure of VUE, but for its analysis and formalization of all aspects, it is advisable to use more advanced knowledge management tools focused on ontology processing. In addition, the use of ontologies allows to import knowledge from external repositories [22]. Therefore, now the structure of the e-VUE KB is formally fixed with the help of the appropriate Wiki-ontology, which

is the result of cooperation of domain experts (in this case – moderators of knowledge spheres represented in the encyclopedia) with knowledge engineers.

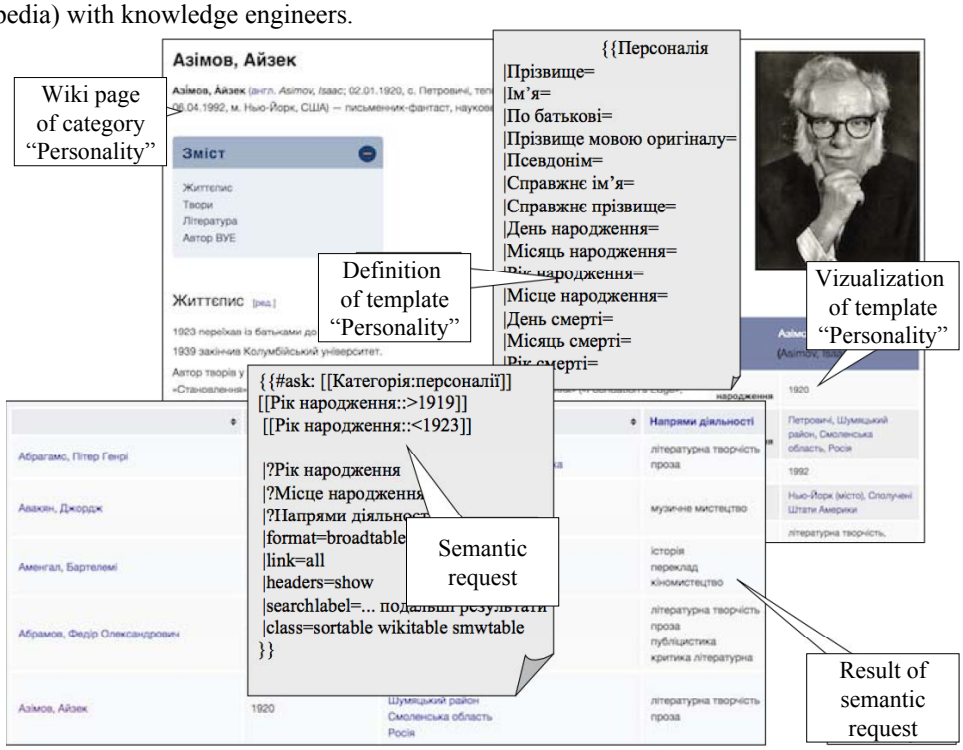

Fig. 3. Use of the "Personal" template for a semantic query

Use of ontologies for e-VUE KB development

Use of Wiki-based IR as a distributed KB requires development of Wiki ontology. Such ontology contains a formalized model of knowledge of this IR and allows to fix the main characteristics of its elements, relations, properties and links in a form suitable for automatic processing, inference and analysis. The use of Wiki-ontology elements for semantic markup provides a unified creation of hierarchically related categories, templates of TIOs, their semantic properties and queries that use them.

Wiki ontology is a special case of computer ontology that reflects the specific features of Wiki technology and its semantic extension. Expressive capabilities of Wiki ontology are limited in the field of domain axioms using and do not involve the application of characteristics for object properties and data properties [24]. Such ontology can be built automatically on base of semantically marked Wiki-resource (a set of Wiki pages containing semantic markup), but in practice usually first Wiki-ontology is developed, and on its base the Wiki-resource with semantic markup is realized.

Wiki ontology of e-VUE (Fig. 4) allows to determine more deeply and formally the presence or absence of hierarchical links between e-VUE categories, the admissibility or inadmissibility of their intersection; display links between categories and TIOs. It indicates the semantic properties of Wiki-pages that reflect the meaningful relations between encyclopedia pages relevant with various TIOs, describes properties of categorized e-VUE articles and determines the characteristics of these properties.

The structure and properties of Wiki ontologies are analyzed in more detail in [25], but it is important to note that Wiki-ontology is created for each IR separately and reflects both the specifics of the domain and purpose of such IR. This ontology allows to formalize the knowledge about the main elements of this encyclopedic edition, to clearly and unambiguously establish the relations between the semantic properties of different types of Wiki pages and their scope [26]. For example, ontology allows to indicate explicitly that the semantic property "Place of birth" takes on a meaning not just of the type "Page", but only among the pages of the categories "City" and "Country", and such a statement can not be fixed by built-in means of SMW.

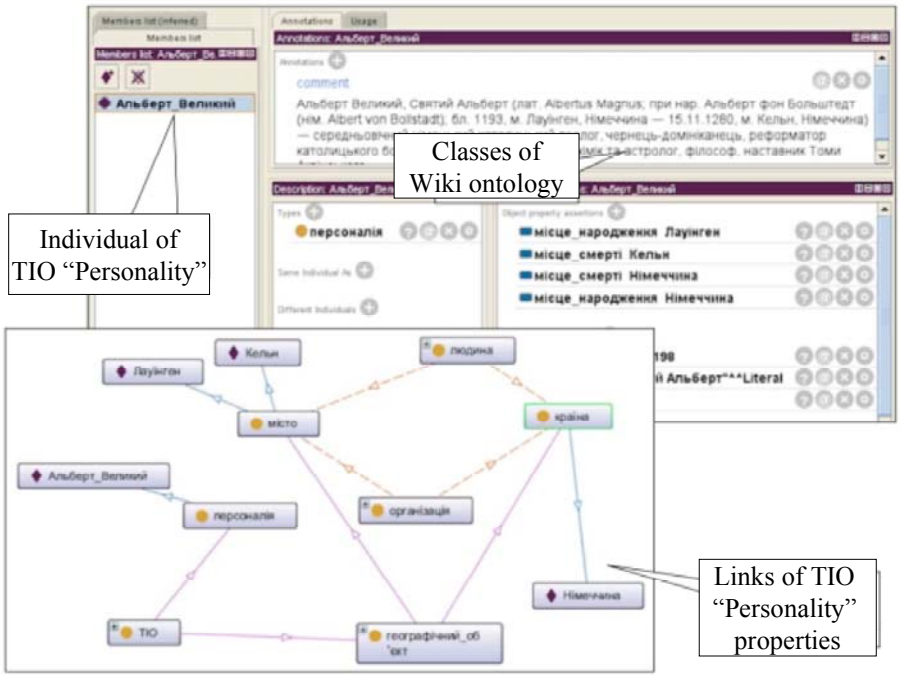

Fig.4. Fragment Element of e-VUE Wiki-ontology that formalizes TIO "Personality"

TIO of e-VUE and use of Semantic MediaWiki templates for their creation

The main stage of information system development is to determine its TIOs and methods used for their processing. The set of such TIO can be considered as an information model of domain that formalizes structure, parameters, their types and possible values, relations with other TIOs, etc.

e-VUE domain is complex, and the encyclopedia itself should uniformly reflect the knowledge of different fields. Therefore, the process of portal development needs in solving of the following tasks deals with the construction of the e-VUE KB and the creation of methods for analyzing the knowledge contained in this KB:

1. Definition of TIO groups with common (or similar) semantic properties and structure (e.g., e-VUE contains TIOs “Sea”, “City”, “Organization”).

2. Search of correspondences between the TIO and categories (or sets of categories). For example, e-VUE TIO "Personality" corresponds to the category "Personalities").

3. Definition of semantic properties typical for each TIO group, research of their intersections and types, especially – for connections between semantic properties and TIOs of their values (e.g., e-VUE TIO "Personality" has semantic property “Place of birth ” with values from TIOs “ City ”and“ Country ”).

4. Analysis of the characteristics of semantic properties specific to each group of TIOs – whether it is mandatory to determine their values for all TIOs, whether they can have more than one value (e.g., e-VUE semantic property "Place of birth" can have one or multiple values or not be defined but semantic property "Year of birth" can have unique property).

5. Analysis of what information about the TIO should be reported to users and in what form (e.g., semantic property "Place of birth" should be accompanied by text explanation, and the date, month and year of birth should be displayed together, and the value of the property "Sex" should not be displayed – it is required only for automated processing).

6. Creating of TIO templates that take into account the results of previous stages of analysis (for example, now we create e-VUE templates for TIO "Personality", "City", "Country", "Organization", etc. that are used in the articles of various fields of knowledge).

7. Development of semantic queries based on TIO information.

8. Creating templates that use these semantic queries (for example, we develop e-VUE templates with semantic queries that find moderators of knowledge directions for the current page, generate list of e-VUE articles for authors pages, search famous residents of cities and countries, etc.).

The creation of a system of TIOs relevant to domain of IR is the subject of system analysis, and the reflection of this system in the form of an ontological structure is the subject of ontological analysis. It should be noted that the creation of such models can be based both on the skills of the knowledge engineer and experience of domain experts, and on Data Mining tools that use machine learning methods.

The main stages of e-VUE article creation

The process of e-VUE articles creating requires a clear technological algorithm that consists of the main stages where various actors are involved in encyclopedia development [27]:

Stage 1. Preparation the article for uploading to the portal (moderator of scientific direction, author of the article).

Stage 2. Editing of the article (literary editor, scientific editor, art editor, and bibliographer).

Stage 3. Control of the article (moderator of scientific direction, head of the editorial department, author of the article).

From the point of view of Wiki technology and its semantic extension the 2nd stage of this process is connected with execution of the following steps of Wiki-page construction:

1) formation of industry dictionaries of articles that determine the complete set of Wiki pages of e-VUE (these dictionaries are necessary for defining of cross-links between Wiki pages);

2) creation a short definition of the article and its categorization;

3) generation of Wiki-page contained name and definition of the article;

4) preparation of encyclopedic article by author who is a specialist in the relevant scientific field;

5) reviewing and checking the article by scientific and literary editors;

6) publication the text of the article on the Wiki-page on the portal;

7) Wiki-markup on the article page for content structuring;

8) clarification and adding of the list of categories of article;

9) addition of semantic templates-integrators (e.g., “Reference”, “Moderator”, “Author” and, if necessary, templates of relevant TIO);

10) supplementing the article with pertinent multimedia material (relevant files are uploaded to the portal, pages with meta-descriptions are created for them, links to these files are added to the Wiki-page of article);

11) if necessary, audio file with postsynchronization of article text is created and uploaded to the portal, and the player of this file is added to the page;

12), the date of publication of the final version of the article is fixed after the last verification and approval of the content, and the Wiki-page becomes available to all users (in the process of Wiki page creating its content is available only to registered users of e-VUE who develop this IR).

These steps are performed manually by the appropriate specialists, but some of them can be automated with the help of specially designed software tools. For example, the problem of mass publication of articles with definitions on the portal based on the text of the industry dictionary requires such a solution. Specialists of knowledge directions form separate industry dictionaries that contain lists of articles with short definitions. Each dictionary is represented by a pdf-file containing 3-5 thousand articles. We need to automate the process of matching articles in dictionaries with a set of portal articles: if the articles is not entered on the portal (i.e. there is no corresponding Wiki page), we need to create such page, place definition and with use of appropriate templates define semantic properties dealt with status of article.

The problem is complicated by the fact that some of the articles have already been entered on the portal, some of them have already been submitted with completed articles, while others do not even have a definition or this definition should be replaced by definition from the dictionary. Analysis of the dictionary text showed that some articles in the dictionary and on the portal differ in spelling. In addition, some difficulties are caused by recognizing paragraphs of text that correspond to the article (usually one article with a definition is represented by one paragraph of text, but some articles are represented by several paragraphs).

For automated processing due to the lack of standardized separation of the article string from the definition string we decided to create an utility that would implement semi-automatic input of missing articles, based on the article of the first word of the paragraph, which is clearly separated by a space, but terminal decision about article representation is a function of human – suitably trained operator.

Due to the large volume of files, the process of dictionary analysis takes a lot of time, so in order to avoid the server overload we propose to do this work divided in parts. Utility input is a text file that contains articles with definitions and database of the e-VUE portal. As a result of the search, the operator receives a Web-page with a table with three columns (Fig. 5) where the first one contains information about the article from the dictionary, the second one shows founded pages from the e-VUE database for this article, and the third one contains a form where the operator can edit the text of the article and definition attached to the e-VUE and choose the action – skip or publish.

This utility executes the database searches and adds articles to the e-VUE by calling the MediaWiki API. To protect content changes against unauthorized interference MediaWiki uses the API for authorization and record the person who performed the operation.

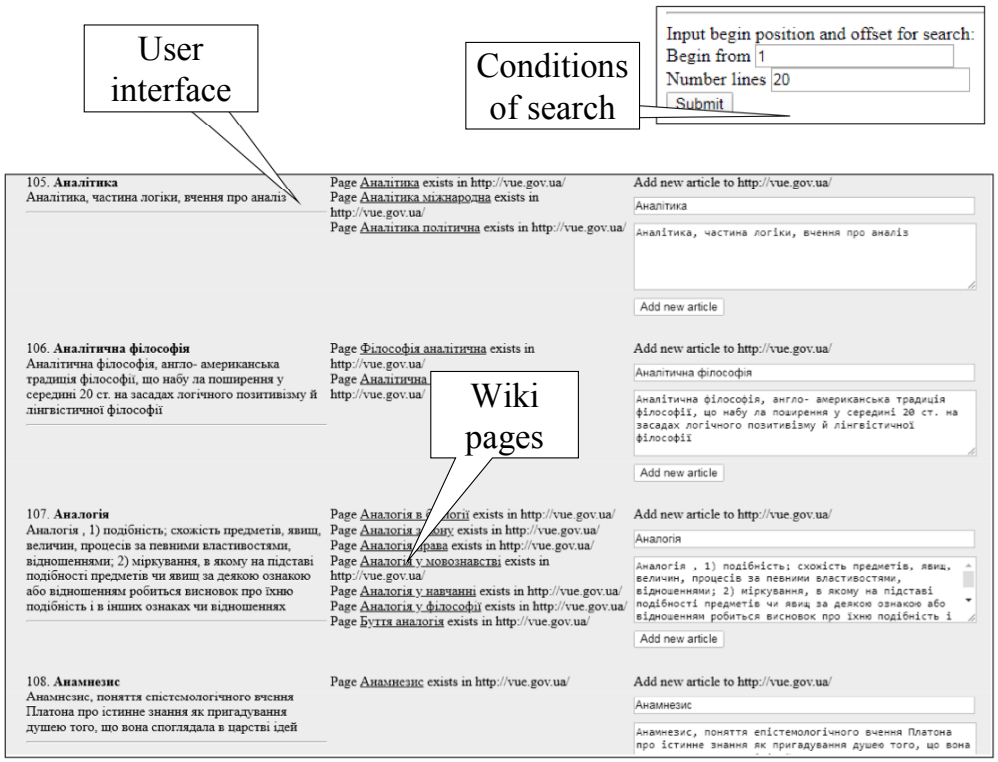

Fig. 5. Utility for entering dictionary articles to the portal (on the example of philosophy dictionary processing)

Prospects for e-VUE functionality expanding

As a result of the research, the development of the general structural scheme of the e-VUE KB with main elements and relations between them is completed. Specifics of e-VUE knowledge substantiates the use of ontological analysis and development of ontological model of e-VUE for formal and unambiguous representation of the structure and content of the knowledge base of the online version of the Great Ukrainian Encyclopedia.

e-VUE is an innovative information product that provides technological support for intelligent information processing. This is caused by several fundamental advantages of this IR:

- possibility of automated and integrated content updating to ensure the information actuality;
- combination of different means of information representation for different audiences through cross-references and integration of multimedia elements at the semantic level;
- ability of knowledge sharing with other intelligent systems based on Semantic Web formats;
- flexible and intelligent system of search and navigation that takes into account the semantic links between the articles of the encyclopedia.

Further research aimed at improving the development of e-VUE portal are closely associated with its further intellectualization and improvement of its services. Development of the portal's functionality is planned in the following areas:

- use of elements of artificial intelligence and machine learning to acquire knowledge from the e-VUE content and maintaining its integrity;
- development of means of automatic content generation using the e-VUE knowledge base and its ontological model;
- analysis and extension of the TIO structure to create additional ways of navigation between semantically related articles;
- improving the means of semantic search to find information objects of a certain type that satisfy the conditions defined by the user;
- improving of the publishing process of portal information;
- personification and content access control;
- further integration of portal with Semantic Web applications for import and export of knowledge.

References

1. Scientific basis, theoretical and methodological principles of modern encyclopedias creation: a collective monograph (2015) Ed. Kiridon A.M. K.: State Scientific Institution "Encyclopedic Publishing House". 160 p. (in Ukrainian)
2. Encyclopaedia Britannica. https://www.britannica.com.
3. The Great Ukrainian Encyclopedia. https://vue.gov.ua.

4. Andon F.I., Yashunin A.E., Reznichenko V.A. Logical models of intellectual information systems. K.: Naukova Dumka. 1999. 396 p. (in Russian)
5. Wikipedia – https:// www. wikipedia.org.
6. Leuf B., Cunningham W. The Wiki way: collaboration and sharing on the Internet, 2001, http://www.citeulike.org/group/13847/article/7659081.
7. Rogushina J.V. Semantic Wiki-resources and their use for the construction of personalized ontologies. *Problems in programming.* (2-3), 2018, P. 188–195.
8. Rogushina J. (2018) Modern Wiki Technologies as a Basis for Knowledge Management in Electronic Encyclopedic Editions // Traditions and Modern Concepts of Encyclopedic Affairs in Ukraine: A Collective Monograph / Ed. Kiridon A.M. K.: State Scientific Institution «Encyclopedic Publishing House». P. 225–233. https://ev.vue.gov.ua/wp-content/uploads/2019/11/Traditions.pdf. (in Ukrainian)
9. MediaWiki. – https://www.mediawiki.org/wiki/MediaWiki.
10. Krötzsch M., Vrandečić D., Völkel M. Semantic MediaWiki. I*nternational semantic web conference.* 2006. P. 935–942. https://link.springer.com/content/pdf/10.1007/11926078_68.pdf.
11. Davies J., Fensel D., van Harmelen F. Towards the Semantic Web: Ontology-driven knowledge management. John Wiley & Sons Ltd, England, 2002. 288 p.
12. Foley B., Chang T. Wiki as a professional development tool. Society for Information Technology & Teacher Education International Conference. 2008. P. 2959–2966.
13. Grishanova I. & Rogushina J. Development of methods for controlling access to information in wiki resources. *Problems in programming.* 2020. N 1. P. 33–46. (in Ukrainian)
14. Resource Description Framework (RDF) Model and Syntax Specification. W3C Proposed Recommendation. http://www.w3.org/TR/PR-rdf-syntax.
15. OWL Web Ontology Language. Overview. W3C Recommendation: W3C, 2009. http://www.w3.org/TR/owl-features/.
16. Manual:Extensions. https://www.mediawiki.org/wiki/Manual:Extensions.
17. Extension:CategoryTree. https://www.mediawiki.org/wiki/Extension:CategoryTree.
18. Extension:EmbedVideo. https://www.mediawiki.org/wiki/Extension:EmbedVideo.
19. W3C Semantic Web Activity. http://www.w3.org/2001/sw/Activity/.
20. Gladun A. & Rogushina J. (2016) Semantic technologies: principles and practices. K.: ADEF-Ukraine LLC, 308 p. (in Ukrainian)
21. Rogushina J. (2017) The use of semantic properties of Wiki-resources for enhancing the functionality of the Great Ukrainian Encyclopedia // Encyclopedic publications in the modern information space: a collective monograph / Ed. Prof. Kiridon A.M. – K.: State Scientific Institution "Encyclopedic Publishing House". P. 104–115. (in Ukrainian)
22. Rogushina J. (2014) Knowledge-oriented tools for semantic search supporting in the Web // OSTIS-2014, Minsk. P. 339–352. (in Russian)
23. Gladun A. & Rogushina J. (2013) Ontology repositorias as a mean of knowledge reuse for recognition of information objects. *Ontology for design*, N 1 (7). P. 35–50.
24. Rogushina J. Analysis of Automated Matching of the Semantic Wiki Resources with Elements of Domain Ontologies. *International Journal of Mathematical Sciences and Computing* (*IJMSC*). 2017. Vol. 3. N 3. P. 50–58.
25. Rogushina J. Semantic Wiki resources and their use for the construction of personalized ontologies. *CEUR Workshop Proceedings* 1631, 2016. P. 188–195.
26. Rogushina J., Grishanova I. (2019) Ontological model of knowledge base of online version of Great Ukrainian Encyclopedia and methods of its use for semantic search and navigation. *Encyclopaedic content and challenges of modern world* / Ed. Kyrydon A.M., Kyiv, P. 69–74. (in Ukrainian)
27. Methodical recommendations for the preparation, editing and design of articles for the Great Ukrainian Encyclopedia (2015) / Ed.. Kiridon A.M. K.: State Scientific Institution "Encyclopedic Publishing House". 120 p. (in Ukrainian)