

ONTOLOGICAL METHODS AND TOOLS FOR SEMANTIC EXTENSION OF THE MEDIAWIKI TECHNOLOGY

J.V. Rogushina, I.J. Grishanova

Practical aspects of ontological approach to organization of intelligent Wiki-based information resources (IR) are considered. We analyze the main features, capabilities and limitations of MediaWiki as a technological platform for development of the Web-based information resource and suggest main directions of its refinement. We propose an abstract model of MediaWiki architecture that formalizes relations between the main components of this software environment and analyze the ways of its semantic extensions based on ontological representation of domain knowledge. An original algorithm of semantic Wiki pages matching with domain ontology is developed. We propose an ontological model of IR that formalizes its knowledge base structure and explicitly performs main features of typical information objects (TIO) of this IR. Such TIOs depend on domain specifics and purposes of IR, therefore their development has to involve domain experts and knowledge engineers. Use of ontology corresponding to the set of Wiki pages (either with semantic markup or without it) provides new IR functions associated with semantic search and navigation. Other important aspect of intelligent Wiki resource development deals with adaptation of user interface to the specifics of IR: enabling various tools of navigation, visualization and content analysis by processing of TIO features enriches IR functionality, reduces access time to information and makes usage of IR more efficient. Developing additional MediaWiki functionality with new requests to the MediaWiki API using TIO templates, extends data analysis and integration capabilities, and offers different, user-focused, IR content views expands the possibilities of data integration and proposes various user-oriented representations of IR content. Wiki resource semantization allows the use knowledge acquired from such IR by external application, or example, by search engines for intelligent Web retrieval. Domain ontologies based on various subsets of the Wiki pages and generated by them thesauri can be used by various Semantic Web applications, both independently or in general technological chain for personified retrieval focused on individual users and their tasks. Approbation of this approach is demonstrated by MAIPS retrieval system. We consider the use semantic similarity of concepts represented by Wiki-pages of IR as an additional way of intelligent navigation between these pages. Such approach allows to group Wiki pages according to user interests by different aspects of their content and structure. Wiki ontologies are considered as the basis for estimation of semantic similarity between domain concepts pertinent to user task. Such elements of Wiki ontology as classes, property values of class instances and relations between them are used as parameters for the quantitative assessment of semantic similarity of Wiki pages. We propose to use local similarity and generate the sets of semantically similar concepts (SSC) that takes into account some subset of page properties and categories defined by user needs. Such sets of SSCs can be considered as user task thesauri for other applications. In addition, we propose to enrich the basic tools of MediaWiki used for access management to the IR content with specialized software code that performs content classification that take into consideration separate namespaces, categories, templates and semantic properties of TIO acquired from Wiki markup. We demonstrate the software implementation of proposed solutions by developing of portal version of the Great Ukrainian Encyclopedia (e-VUE) that contains heterogeneous multimedia content with complex structure. We analyze the specifics of e-VUE knowledge system and develop its formalized TIO representation based on Semantic Web technologies and ontological analysis. Ontological model of e-VUE and original methods of its processing used for this project extend the functionality of the portal in the area of search, navigation, integration and protection of content based on background domain knowledge. In addition, original user interface of e-VUE is developed with an allowance for Encyclopedia knowledge specifics, substantially differs from the standard Wiki, meets the requirements, goals and objectives of this IR and provides a lot of additional features.

Key words: ontology, intelligent Wiki-resource, MediaWiki, semantic similarity.

Розглянуто практичні аспекти онтологічного підходу до організації інтелектуальних інформаційних ресурсів (ІР) на основі Wiki. Проаналізовано основні функції, можливості та обмеження MediaWiki як технологічної платформи для розробки Web-орієнтованих ІР та запропоновано основні напрямки її вдосконалення. Ми пропонуємо абстрактну модель архітектури MediaWiki, яка формалізує відношення між основними компонентами цього програмного середовища, та аналізуємо шляхи її семантичного розширення на основі онтологічного подання знань. Ми використовуємо онтологічний підхід до управління знаннями ІР та пропонуємо оригінальні алгоритми співставлення Wiki-сторінок та онтології предметної області (ПрО). Онтологічна модель ІР формалізує структуру його бази знань і явно подає основні властивості типових інформаційних об'єктів (ТІО) цього ІР.

Вдосконалення засобів навігації, візуалізації та аналізу контенту шляхом обробки ТІО збагачує функціональність ІР. Розробка додаткового функціоналу MediaWiki за допомогою запитів до MediaWiki API, які використовують шаблони ТІО, розширює можливості аналізу та інтеграції даних і пропонує орієнтоване на користувача подання контенту ІР. Семантизація Wiki-ресурсу дозволяє використовувати знання, що здобуваються з нього, іншими застосуваннями, наприклад, для інтелектуального пошуку у Web зовнішніми пошуковими системами. Ми пропонуємо використовувати семантичну подібність понять для інтелектуальної навігації між сторінками ІР. Wiki-онтології використовуються як основа для оцінки семантичної подібності між поняттями ПрО відповідно до задач користувача. Ми пропонуємо використовувати локальну подібність та генерувати набори семантично подібних понять (СПП), приймаючи до уваги тільки визначену користувачем підмножину їх властивостей та категорій відповідно до його інформаційних потреб. Ми демонструємо програмну реалізацію запропонованих рішень на прикладі розробки порталної версії Великої української енциклопедії (e-VUE), що містить гетерогенний мультимедійний контент зі складною структурою. Онтологічна модель e-VUE та методи її обробки, що використовуються для цього проекту, забезпечують розширення функціональності порталу в сфері пошуку, навігації, інтеграції даних та захисту інформації на основі фонових знань ПрО. Крім того, e-VUE має оригінальний користувацький інтерфейс, який суттєво відрізняється від стандартного Wiki, відповідає вимогам, задачам та цілям енциклопедичного видання і надає користувачам порталу додаткові функції.

Ключові слова: онтологія, інтелектуальний Wiki-ресурс, MediaWiki, семантична подібність.

Рассмотрены практические аспекты онтологического подхода к организации интеллектуальных информационных ресурсов (ИР) на основе Wiki. Проанализированы основные функции, возможности и ограничения MediaWiki как технологической платформы для разработки Web-ориентированных ИР и предложены основные направления ее усовершенствования. Мы используем онтологический подход к управлению знаниями ИР и предлагаем оригинальные алгоритмы сопоставления Wiki-страниц и онтологии предметной области (ПрО). Онтологическая модель ИР формализует структуру его базы знаний и явным образом представляет основные свойства типичных информационных объектов (ТІО) этого ИР. Семантизация Wiki-ресурса позволяет использовать знания, извлеченные из ИР, для интеллектуального поиска в Web внешними поисковыми системами. Онтологии ПрО, построенные на основе различных наборов Wiki-страниц такого ИР, и тезаурусы задач, генерируемые по этим онтологиям, могут

использоваться различными приложениями Semantic Web как самостоятельно, так и в составе технологического цепочки для персонализированного поиска. Кроме того, мы предлагаем обогатить основные инструменты управления доступом MediaWiki к контенту Wiki-ресурса специализированным программным кодом, который выполняет классификацию этого контента с использованием отдельных пространств имен, категорий, шаблонов и семантических свойств ТИО, извлеченных из Wiki-разметки. Мы демонстрируем программную реализацию предложенных решений на примере портальной версии Большой украинской энциклопедии (e-VUE), которая содержит гетерогенный мультимедийный контент со сложной структурой. Онтологическая модель e-VUE и методы ее обработки обеспечивают расширение функциональности портала в сфере поиска, навигации, интеграции данных и защиты информации на основе фоновых знаний Про. e-VUE имеет оригинальный пользовательский интерфейс, который существенно отличается от стандартного Wiki, соответствует требованиям, задачам и целям энциклопедического издания и предоставляет пользователям портала дополнительные функции.

Ключевые слова: онтология, интеллектуальный Wiki-ресурс, MediaWiki, семантическое подобие.

Introduction. Ontologies and Semantic Web technologies

Now the main direction of information technologies deals with problem of knowledge storing and processing . Actuality of transition from data processing to knowledge analyzing is escalated by growth of data volume (Big Data) generated by human society and heterogeneity of their representation. One of the most interesting approaches for data semantization is realised by the Semantic Web project [1]. The author of this concept promoted by the W3C Consortium is T. Bernes-Lee – the developer of the modern Web defines the Semantic Web is an infrastructure, not an application. Semantic Web technologies are focused on the processing of IR content on a semantic level. Semantic Web extends the existing Web aimed at automating the processing of Web content and transformation of the Web into the global distributed knowledge base (KB) oriented either on human users and various intelligent application such as software agents. Such transformation requires with a semantic descriptions of all IRs and means for the automated processing of these descriptions. According to the W3C definition, the main idea of the Semantic Web is to represent the Web data in such way that provides their integrated reuse in various applications and automatically generate documents from a variety of sources for various complex tasks. Moreover, the Semantic Web is designed to provide access not only to static documents that contain useful information, but also to services that enable intelligent agents to analyse relevant information. Despite the ambiguity of the Semantic Web concept, some basic directions are repeated in all its definitions and are supported by various Semantic Web application:

- description of IR semantics with use of metadata;
- the use of metadata for IR search on semantic level;
- representation of the Web as a global interoperable database;
- discovery, integration and use of the Web services on base of their semantics description;
- providing information to intelligent software agents in a form suitable for machine processing.

The core technological components of the Semantic Web are: ontologies used for interoperable and formalized representation of domain knowledge for intelligent Web applications; Web services used for processing of these ontologies in various tasks; software agents that act on the Web on behalf of and in the interests of their users by activation of corresponding Web services. According to the "Semantic Web Challenge" [2], Semantic Web applications integrate, connect and logically process the information required by the user. Such applications are oriented on the open world and have to meet the following minimum requirements:

1. Data values play a key role in their functioning, and processing of semantic information play a central role in achieving results that can not be obtained with the help of alternative technologies;
2. Applications use heterogeneous IRs with different owners that contain real-world data;
3. The retrieval is carried out in the real information space of the Web. As a rule, these applications use in some way such Semantic Web standards as RDF, RDF Schema or OWL, but this is not a prerequisite.

Ontologies are widely used now for domain knowledge representation in the intelligent Web-oriented applications. In recent years, ontologies have regained interest also within the natural language community, specifically in the context of such applications as information extraction, text mining, and question answering, in analysis of Big Data metadata etc. These works are oriented on use of ontology representation language OWL developed by W3C. However, ontology creation and analysis is very costly and complex, and their retrieval into repositories is not a simple task [3].

Knowledge reuse and access is one of the leading motivations for the Semantic Web. The main objective of using ontologies is to share knowledge between computers or computers and humans. We can consider ontology as a special type of knowledge base (KB) with semantic information about certain domain. The components of ontologies depend on the representation paradigm and approach but practically all ontological models contain certain concepts (terms, classes), properties of concepts (attributes, roles), the relationship between concepts (dependencies, functions) and constraints of their use defined by axioms. In the general case, the formal ontology model O is an ordered triple $O = \langle T, R, F \rangle$, where T is the set of domain concepts; R is the set of relations between concepts; F – set of interpretation functions of concepts and relations. This formal model can be specified in accordance with the goal and scope of the ontology. Now the most popular language of ontology representation is OWL (Web Ontology Language) [4]. Ontology built on OWL is a sequence of axioms and facts with the addition of references to other ontologies associated with it. The W3C-endorsed OWL specification includes the definition of three variants of OWL, with

different levels of expressiveness: OWL Lite, OWL DL and OWL Full. They are based on and differ by expressiveness and complexity of processing. Description logics (DL) of the SH family are used in their design: OWL DL is based on SHOIN(D), and OWL Lite is based on SHIF(D) [5]. In 2007 the W3C OWL Working Group develops OWL 2 that refines OWL and is based on the DL SROIQ(D) [6].

In this work we consider the use of the ontologies and the Semantic Web standards and technologies oriented on their representation and processing as a theoretical and technological base for intellectualization of Wiki resources. This approach expands functionality of IRs and provides sharing of their knowledge with other applications. The extension of traditional Wiki resource results transformation of Wiki-portal into the Semantic Web application that provides knowledge sharing with other intelligent Web software.

MediaWiki, its components and properties

MediaWiki is a free software with open code for development of shares Wiki resources. Now a lot of applications are based on MediaWiki [7]. The most well-known of them is Wikipedia [8]. The multilevel architecture of the MediaWiki engine is built on the LAMP architecture (server-side PHP and client-side Javascript, with MySQL database). The MediaWiki kernel provides the basic functionality after installation and configuration. This allows to create and edit pages, set their categories and links, view the content of those pages etc. In terms of information structure, MediaWiki provides templates and page categorization. The basic unit of Wiki resource on base of SMW is Wiki page. Wiki information can be stored at the informal, semi-formal or formal level, depending on the level of collaboration of the content development community. By default, there is no such thing as the required structure, but the lack of structuring makes the content very difficult to navigate. Therefore developers of Wiki-resources on the basis of MediaWiki usually build their own structure according to the domain specifics. For low level operations e-VUE uses one of the most popular and technological search tools Elastic Search. Architecture of MediaWiki at the most abstract level consist of: - DB; - core kernel containing the API; – a set of plugins (extensions); – a set of skins (templates for content displaying) [9].

MediaWiki proposes the following mechanisms for content structuring:

- *Namespaces*. All pages in MediaWiki are divided into separate namespaces provided by the storage model. Namespaces allow to reuse the base model for main content objects (for example, main e-VUE namespace is connected with pages of encyclopedia articles), as well as for multimedia objects (e-VUE uses "File", "Illustration", "Audio" and "Video" namespaces), software blocks and fragments of structured text ("Template" namespace);

- *Categories*. The set of pages located into the "Category" namespace is intended to group pages. These mechanisms provide construction of the Wiki-resource taxonomy;

- *Templates*. This pages are intended to include the reused elements of articles. This mechanism is also used to avoid conflicts in Wiki content. Templates provide the common way for data structuring. For example, e-VUE developers create templates for TIO (like persons, cities, countries, rivers etc.).

Namespaces allow to structure the content of Wiki portals and to separate its content by type. Within one namespace, pages can be grouped by categories. Some names are reserved for namespaces: User – for separating of users, MediaWiki – for official pages of MediaWiki kernel, File – for storing information about files, Special – for special pages with various services, Template – for template pages.

MediaWiki access point are `index.php` for general purposes and `api.php` for high-level programmatic content access through Application Programming Interface (API) query execution.

The MediaWiki API is a web service that provides easy access to all functions and manipulations of Wiki data and metadata. There are currently several APIs in MediaWiki – the basic MediaWiki action API [10], as well as RESTBase [11], the Wikidata request service using the SPARQL API [12], the Semantic Media Access API (Semantic MediaWiki API) [13]. Action API is supported in all MediaWiki (with some changes in functionality and related commands and options), and other APIs need to be installed separately. The API Web service is used by developers for search and integration of the Wiki content (search for category names or keyword into page titles and full text) and for edit content of IR. This API is used also in internal MediaWiki processes. This API service receives an HTTP GET request containing a set of input arguments and returns results in different formats (XML, JSON, WDDX, YAML, PHP-serialized format) that can be processed according to a task and allows to create different types of web applications for large-volume Wiki content. Each API service request has a mandatory parameter that specifies its action. The most commonly used values of this parameter are "query", "edit" and "delete".

MediaWiki works by using one of interfaces: – server-generated html code implemented by libraries written in PHP without frameworks; – HTTP API (`api.php` module) which is used both as a user interface and as an interface for external client interfaces and for external requests; – service scripts executed task-dependently.

Semantic MediaWiki and its data model

Semantic MediaWiki (SMW) [14] is a semantic extension to MediaWiki that allows users to add relations and properties to Wiki pages. Semantic MediaWiki stores semantic data in the MySQL database in MediaWiki, which can also be exported as RDF [15]. SMW is an add-on to the MediaWiki Wiki builder tool.

The benefits of SMW are semantic-level information processing, the availability of group-based knowledge management, relatively high expressive power, robust implementation and a user-friendly interface, on-the-go documentation and user communities. SMW allows to integrate information from different Wiki pages by searching at the level of knowledge and to generate ontological structures on Wiki pages that other systems can use.

Basic semantic concepts in Semantic MediaWiki are categories that allow users to categorize Web pages (present also in normal Wiki resources), semantic properties that allow the content of Wiki content to be linked to other pages of the Wiki resource and to data (date, number, text line, geographic coordinates, etc.), defining the semantics of these relations, and semantic queries that can use categories and semantic properties. The results of semantic queries can be inserted into the corresponding Wiki pages and are automatically updated if the content of the pages from which they retrieve information changes. This ensures the integrity and relevance of the created Wiki-resource.

SMW allows users to annotate Wiki content with clearly defined information that is suitable for automatic processing (machine readable). It supports the addition of structured and semantically annotated information in Wiki using appropriate syntax. The benefits of SMW are semantic-level information processing, the availability of group knowledge management tools, relatively high expressive power, robust implementation and a user-friendly interface, the availability of documentation and user communities. SMW allows to integrate information from different Wiki pages by searching at the knowledge level and to generate ontological structures on Wiki Pages that other systems can use.

Use of SMW semantic elements solves the following major problems of Wiki technology:

- content consistency: information on different Wiki pages should be consistent;
- knowledge access: search and comparing of information from different Wiki pages on semantic level;
- knowledge reuse: semantization provides the access to Wiki content from other applications and exporting knowledge into common formats.

In addition to categories, SMW uses such mechanisms as semantic properties to structure information. They allow to semantically associate Wiki pages by relations from $R_{s_prop} = \{r_{s_prop_j}\}, j = \overline{1, m}$ with each other and with various data with types from T . The following property types are supported in Semantic MediaWiki: String, Number, Annotation, URI, Boolean, Email, Text, Geographic Coordinates, and more. Each semantic property has a type, a name and a value, as well as its own Wiki page P_{sem_prop} in a special namespace, which allows it to determine its place in the property hierarchy and to document how that property should be used. Each semantic property has its own Wiki page P_{sem_prop} where the type of values is explicitly identified by use of `[[Has type :: xxx]]` construct. "Has type" is recognized by SMW as a special property, i.e. the value of such property is a link to another Wiki page.

SMW reuses the Category namespace to define classes. For example, a Wiki page called "Category: Images" is intended to represent a class of all images, as well as a new namespace called "Properties" to define the properties of concepts. With this approach, SMW supports both binary and n-ary properties.

SMW allows to declare a "class-subclass" and "property-sub-property" relation. For example, you could declare that the "Image" category is a sub-category of "Media" by adding the annotation `[[Category: Media]]` to the page of that category. This makes it possible to create a taxonomy of categories and properties, which is very useful both for navigating the Wiki resource and for correctly creating new categories and properties. SMW also supports equivalence claims between two Wiki pages or between two classes using the redirection mechanism in MediaWiki. For example, in e-VUE, with this mechanism, reference articles are organized when there are several common formulations of the same term. SMW allows to declare a page as an instance of a class or an RDF triple entity. For example, if the Khreschatyk Image Wiki page contains annotation `[[Category: Image]]` and `[[Type :: Photo]]`, this means that the Khreschatyk Image is a digital resource of the "Photography" type. You can also create an instance of an n-ary property. For example, on the Image of Khreshchatyk Wiki page, an abstract `[[General Keywords :: Image of Khreshchatyk; Landscape]]` means that Image of Khreshchatyk belongs to the subject category of landscapes. From the point of view of functionality, the semantic annotation adds functionality for semantic representation of relation between Wiki pages. Wiki templates use semantic properties of SMW for unified representation of content.

Formulation of the problem

Development of complex IRs on base of semantic Wiki technologies requires improvement of knowledge management methods oriented on specifics of these resources. MediaWiki needs in extension of its functionality that provides more efficient processing of semantics and proposes to users more intelligent means of search, navigation and data access. An ontological model of Wiki resource represented its KB structure requires the relevant methods that process these structure elements with use of domain knowledge integrates with Semantic Web technologies.

Ontological representation of KB structure of Wiki-resource

We use practical aspects of ontological approach for organization of Wiki-based IRs. Ontological model formalizes the knowledge base structure and explicitly represents main features of IR [16]. Now ontologies are widely

used for domain knowledge representation by the Web-oriented information systems [12]. Ontology is a special kind of KB that contains semantic information [17] from a certain domain and consists of domain concept sets and statements about these concepts. Ontology formalizes their classes, objects, relations, functions and theories. Various models of ontologies usually contain concepts (classes and instances), properties of concepts (attributes), relations between concepts (dependencies, functions), and additional constraints defined by axioms [18].

The most general formal model of the O ontology can be represented by the ordered four $O = \langle T, P, R, F \rangle$, where T is a finite set of domain concepts; P is a finite set of concept properties; R is a finite set of relations between the domain concepts; F is a finite set of axioms defined on the concepts, properties and/or relations of an ontology O. Relations represent the type of interaction between the domain concepts. Axioms are used to model statements that are always true. This ontology model can be concretized depending on development aims.

Wiki ontology is a special case of ontology [19]. It formalizes the knowledge represented by IR designed on the base of Wiki technology and its semantic extensions. Expressive capabilities of Wiki ontology are limited (ontology contains only the knowledge that can be directly acquired from Wiki markup) and do not imply the use of characteristics for object properties and data properties such as equivalence and possibility of intersection. Wiki ontology and the structure of the Wiki-resource are different forms of representation of the same knowledge system. Wiki ontology can be built automatically by the semantically marked Wiki-resource (for any collection of Wiki pages), but in practice such ontology is developed before the resource development process and is used for its structuring.

Wiki resource is a copy of social software that consists of collection of commonly created and used articles that contain Wiki-structured text (oriented on human usage) and hyperlinks to other Wiki articles or external IRs. Wiki resources provide the basis for knowledge sharing. Common features of all Wiki resources are the presence of Web-based interface, simple syntax for content structuring and the ability to define hyperlink to other pages. Wiki ontology O_{wiki} represents knowledge of a semantically marked Wiki resource – a set of Wiki pages containing semantic markup. All knowledge contained in this ontology can be directly acquired from semantic markup elements of corresponding Wiki resource. Such characteristics of ontology classes, their individuals and properties as equivalence, absence of intersection, etc., are not used.

$$O_{wiki} = \langle X_{wiki_categor} \cup X_{wiki_page}, \{r_{ier_cl}\} \cup \{r_{link}\} \cup \{r_{sem_prop}\}, \{f_{equ}\}, \emptyset \rangle.$$

In this model the set of concepts X is built as combinations of Wiki pages X_{wiki_page} and categories $X_{wiki_categor}$: $X = X_{wiki_categor} \cup X_{wiki_page}$ connected by relations from $R = \{r_{ier_cl}\} \cup \{r_{link}\} \cup \{r_{sem_prop}\}$. The class set of Wiki ontology is the set of Wiki categories connected by hierarchical relations r_{ier_cl} ; the set of its instances is a set of Wiki pages that contains hyperlinks r_{link} and semantic properties $r_{sem_prop_i}, i = \overline{0, m}$; the set of data types contains a specific class "Wiki Page". This model can be refined by such Wiki elements as templates, forms, special etc.

Formal model of Wiki ontology O_{wiki} for non-semantic Wiki resource contains the following components:

$X = X_{cl} \cup X_{ind}$ is a set of ontology concepts, where X_{cl} is a set of classes coincides with Wiki categories represented by a set of category pages P_{categ} ; X_{ind} is a set of class instance built as a combination of Wiki pages set $P = P_{user} \cup P_{template} \cup P_{spec}$ where P_{user} is a set of pages created by users, $P_{template}$ is a set of pages describing Wiki templates, P_{spec} is a set of other special pages;

R is a set of relations $R = r_{ier_categor} \cup \{r_i\} \cup \{p_j\}$ contains hierarchical relations between categories $r_{ier_categor}$, relations between Wiki pages $\{r_i, i = \overline{1, 3}\} = \{\text{"link"}, \text{"is_a"}, \text{"use"}\}$ where "link" is a hyperlink relation of one Wiki page of this resource to another one, "is_a": $X \rightarrow P_{categ}$ defines the categorization of Wiki page, and "use": $X \rightarrow P_{template}$ relation defines the use of template into the Wiki page; $\{f_{equ}\}$ defines the equivalence of Wiki pages; $F = \{f_{equ}\}$ is a single-element set of the equivalence relation between Wiki pages that can be used for logical inference on ontology. The other elements of the ontological model for this Wiki ontology are represented by empty sets.

The formal Wiki ontology model for semantically enriched Wiki resources O_{s_wiki} is more complex in comparison with Wiki ontology of non-semantic Wiki resources (such as Wikipedia) and includes a number of elements related to semantic properties [20]: a set of Wiki pages X is enriched by the set of pages of semantic properties P_{sem_prop} (some of them are semantically defined links to other Wiki pages $P_{sem_prop_page} \subseteq P_{sem_prop}$, and others link pages to values of other data types); a set of relations $R = r_{ier_cl} \cup \{r_i\} \cup R_{s_prop}$ is enriched by relations related to the semantic properties of Wiki pages by domain-specific semantic properties $R_{s_prop} = \{r_{s_prop_j}\}, j = \overline{1, m}$ with type

“Page” that link Wiki pages by semantically defined characteristics; $\Gamma_{ier_cl} = \Gamma_{ier_categor} \cup \Gamma_{ier_property}$ is enriched by relations $\Gamma_{ier_property}$ that defines hierarchical relations of semantic properties. T is a set of types for values of semantic properties.

We propose to define typical information object (TIO) of developed IR that are characterized by similar structure, properties and elements. In ontological analysis TIOs are considered as individuals of one class. TIO templates allow to research the area of object property values of Wiki ontology, and vice versa, the characteristics of object properties of Wiki ontology are the basis for correct creation of templates of TIOs. Semantic markup of Wiki-resource allows to automate the process of Wiki ontology development, and therefore the use of semantic technologies based on ontological knowledge representation is simplified for Semantic MediaWiki-based IRs.

Wiki ontology can be generated by analysis of Wiki resource or created by domain expert before the development of this IR as a basis for its semantic markup (concepts and their attributes are used to indicate the categories and semantic properties of the Wiki resource). In practice, both variants are usually used iteratively: at first the domain ontology is generated, then it is used for markup of IR, then semantic markup of the Wiki-resource content is used for ontology improvement by new element that then are used for semantic markup.

Therefore we try to define correspondences between main elements of ontology and Wiki resource for automation of such transformations. An analysis shows that only some correspondences are one-to-one and can be detected automatically. Some others require additional clarifications and definition from the user.

Use of Wiki templates for representation of TIOs of semantic Wiki resource

Mechanisms of content templating are used in the Wiki to give content authors the ability to fill a pre-defined schema (model) of Wiki page with specific values. Templates can be used into the Wiki pages with similar elements (with homothetic structure, with the same elements and properties) that correspond to the same TIO type. There is important to understand that every TIO of Wiki resource are usually linked with one or more templates (several templates are used for TIO with very complex structure or in situations where different TIOs have the similar parts, i.e. in e-VUE TIO “Personality” uses templates “Personality” and “Award”. The template-based model provides for flexible design of the Wiki-resource scheme. Each template defines a class with freely selected attributes (equivalent to "entity type"), instances of which can be freely embedded in other objects. Template instances can be hierarchically linked. For example, e-VUE uses a hierarchically organized set of templates to provide knowledge about TIOs.

There are two types of Wiki pages involved in template processing: template page (for example, *Template:River*) is the template definition page which content is used by the pages that called this template; template instance page (for example, *Dnipro*) is a page which content is constructed using a template call with certain specific values of its parameters that can define the corresponding semantic properties for SMW.

The main important features of Wiki templates are: strong and permanent link between the template and its instances (changes made by user into the template page are immediately displayed on all pages where this template is called); different source code for templates and their instances.

Matching of Wiki ontologies as a Tool for KB structure management

Dynamics of Wiki resources is an important and very useful feature that improves the efficiency of their application. Each time, IR developers can change not only the content of Wiki pages but also the structure of their relations and properties. An important feature of Wiki ontologies deals with their use to automate the control of changes in accordance with Wiki resource structure. Instead of recording all changes in IR in natural language (NL) descriptions, it is advisable to create temporal versions of Wiki ontology that correspond to different time states of that IR.

Such changes can be caused by the following user actions: adding new semantic properties to L_{sem_prop} ; changing the value types of semantic properties to T ; adding new templates that use semantic properties to $P_{template}$; adding Wiki pages for typical IOs (by corresponding templates from $P_{typic_IO} \subseteq P_{template}$); establishing links between Wiki pages deal with typical IOs by elements of L_{sem_prop} ; modifying the structure of templates $P_{TIO} \subseteq P_{template}$ that use semantic properties. If the information about changes in KB structure is fixed in the form of NL descriptions, it takes a lot of time to search for differences between versions of complex IR with branched knowledge system. The use of Wiki ontologies corresponded to different temporal states of this resource provides automated search of such differences with the help of ontology mapping tools (mapping operations are very difficult for independently created ontologies, but they are quite simple for those ontologies that share a common basis). In addition, restrictions on Wiki ontology expressiveness also greatly speed up such processing.

Semantic similarity of Wiki pages as an instrument of domain representation for navigation in intelligent IRs

Semantically similar concepts (SSC) is a fuzzy set of concepts with the semantic distance less than the selected threshold. Now similarity measures used in various applications, such as information retrieval, unstructured data analysis [21], Big Data processing etc. Measures to determine the semantic similarity of concepts on the basis of ontologies use various semantic features of these concepts – their properties (attributes and relations with other concepts), the relative position in ontological hierarchies. Semantic similarity can be considered as a special case of semantic relation between domain concepts. In [22] the assessment of similarity in semantic networks is defined with the help of taxonomic links. Although other types of links such as “part-of” can also be used for assessment of similarity [23].

Similarity analysis of hierarchical (taxonomic) relations is probabilistic. The assessment of the similarity of concepts from different ontologies may be based on the positions of these concepts in the hierarchy of classes for which similarity has already been determined: if the superclasses and subclasses of these concepts are similar, then concepts may also be similar. The similarity of two concepts can take into account similarity estimation of: direct superclasses of these concepts; all superclasses of these concepts; subclasses of these concepts; instances of these concepts.

Similarity estimation that allows another (non-hierarchical) connections between concepts are constructed by analogy. For example, if it is known that the classes "Human" and "Person" are similar and the relation "work" and "be an employee" are also similar, and in the first ontology the class "Human" is related by the relation "work" with the class "Organization", and in the second ontology the class "Person" is related by the relation "to be an employee" with the class "Company", then concepts "Organization" and "Company" represented by the classes of different ontologies are similar.

A common and long-known way of evaluation of semantic similarity in taxonomy lies in measuring the distance between net nodes that correspond to the elements being compared – the shorter path from one node to another means their higher similarity. In general all criteria of estimation are based on the shortest path length between concepts built by various relations [24]. Some works estimate similarity of concepts in semantic networks only by taxonomic relation “is-a”, others take into account different types of relations, i.e. analyze relation “part-of-part”.

The similarity of concepts is also related to their content. The more information is shared by two concepts, the more similar they are, and the information co-shared by two concepts is determined by the information content of the concepts included in them into taxonomy. Concept similarity can be defined by similarity of strings and words. Feature vectors are widely used for knowledge representation, especially in case-based reasoning and machine learning. They can be applied for representation of words. In real task, there is great variability of distances covered by a single taxonomic relation, especially if some subsets of taxonomies (such as biological categories) are much denser than other ones. Therefore, it is advisable to take into account the semantics of relations between concepts for different taxonomic relations and to consider the number of instances in subclasses. All these researches use such aspects of ontological knowledge representation (hierarchical and domain-specific relations of classes and individuals, properties of class individuals) that can be represented by semantic Wiki resources. Therefore we consider the use of Wiki resources as a source for generation of sets of semantic similar concepts for intelligent systems [25].

We propose to extend e-VUE navigation and search functionality by retrieval of semantically similar IOs – both *globally similar* (similarity is estimated by the full set of features – either categories and values of semantic properties) and *locally similar* (similarity is estimated only by certain subset of these features). Concepts similar for the current Wiki page are defined by its characteristics. We analyzed the advantages and disadvantages of scientific researches deal with of the semantic similarity of concepts and considered the possibility of scaling them to the big volumes of data that are inherent to the set of e-VUE articles and categories. The use of ontologies for knowledge structure formalization provides an effective establishing semantic similarities between the concepts. We pay particular attention to handling special cases of ontological relations (e. g. “is-a” relations, synonymy, hierarchical relations “Class-subclass” and “Class-class instance”) implemented by the basic means of the SMW technological environment.

The structure of e-VUE Knowledge Base

Now Ukraine is actively working to create the Great Ukrainian Encyclopedia. The quality of the information represented in this encyclopedia is ensured by focusing on peer-reviewed author articles with original content prepared by specialists in relevant fields. It provides verified facts and theories recognized by the scientific communities.

e-VUE (vue.gov.ua) is a portal version of the Great Ukrainian Encyclopedia. In addition to textual content, e-VUE pages may contain other types of information (images, maps, video, audio, etc.) and links to trusted sources. This is an innovative project to create a national encyclopedia based on modern means of knowledge representation. The creation of this IR is based on modern knowledge-oriented technologies (in particular, on the Semantic Web) and intelligent Wiki software. The portal version of the Great Ukrainian Encyclopedia uses the free software MediaWiki version 1.29.1. and its semantic extension Semantic MediaWiki version 2.5.5. e-VUE uses Foreground style set that focuses on content selection, supports adaptive layouts and has special predefined classes to support Semantic MediaWiki functionality. This set is built on the Zurb's Foundation Framework (v4.3.2) focused on supporting mobile devices and an extended adaptive front-end framework. Foreground style set completely changes the layout of the Wiki page representation and allows to add new interface elements.

Specifics of e-VUE causes the development of original methods and software solutions. This IR contains multi-domain knowledge with very complex structure represented by heterogeneous elements. Therefore we use e-VUE to demonstrate the effectiveness of the described above approaches and methods.

The advantages of e-VUE via other online encyclopedias based of traditional Wiki technology are:

- the explicit establishment of meaningful links between the pages of articles and their elements;
- search for information by meaning - by categories and values of semantic properties of pages;
- the ability to integrate information from different articles and automatically generate integral content;
- the ability to export knowledge in the formats of modern Web-technologies.

e-VUE becomes a knowledge source not only for humans, but also for other Semantic Web applications that can use knowledge exported from e-VUE in generally accepted presentation formats.

The basic principles of the e-VUE knowledge management are as follows:

- the main tools for KB organizing are categories and semantic properties provided by Semantic MediaWiki;
- an extensible collection of independent taxonomies connected with the sets of corresponding categories is used for structured representation of information (each article can be assigned to the arbitrary set of categories, and their hierarchical relations don't cause conflicts in categories processing);
- TIOs represent the sets of reused semantic properties (articles that belong to some IO type have a specific set of categories and properties with fixed names and values);
- the sets of TIO semantic properties in the process of refinement of the e-VUE database can be expand, but can't be decreased (non-used properties cannot be deleted for correct processing of prior articles);
- the estimation of the similarity score of article take into account the knowledge from the relevant domain ontologies (the number of links between concepts and the level of significance of these relations), taxonomies of categories and the category volume.

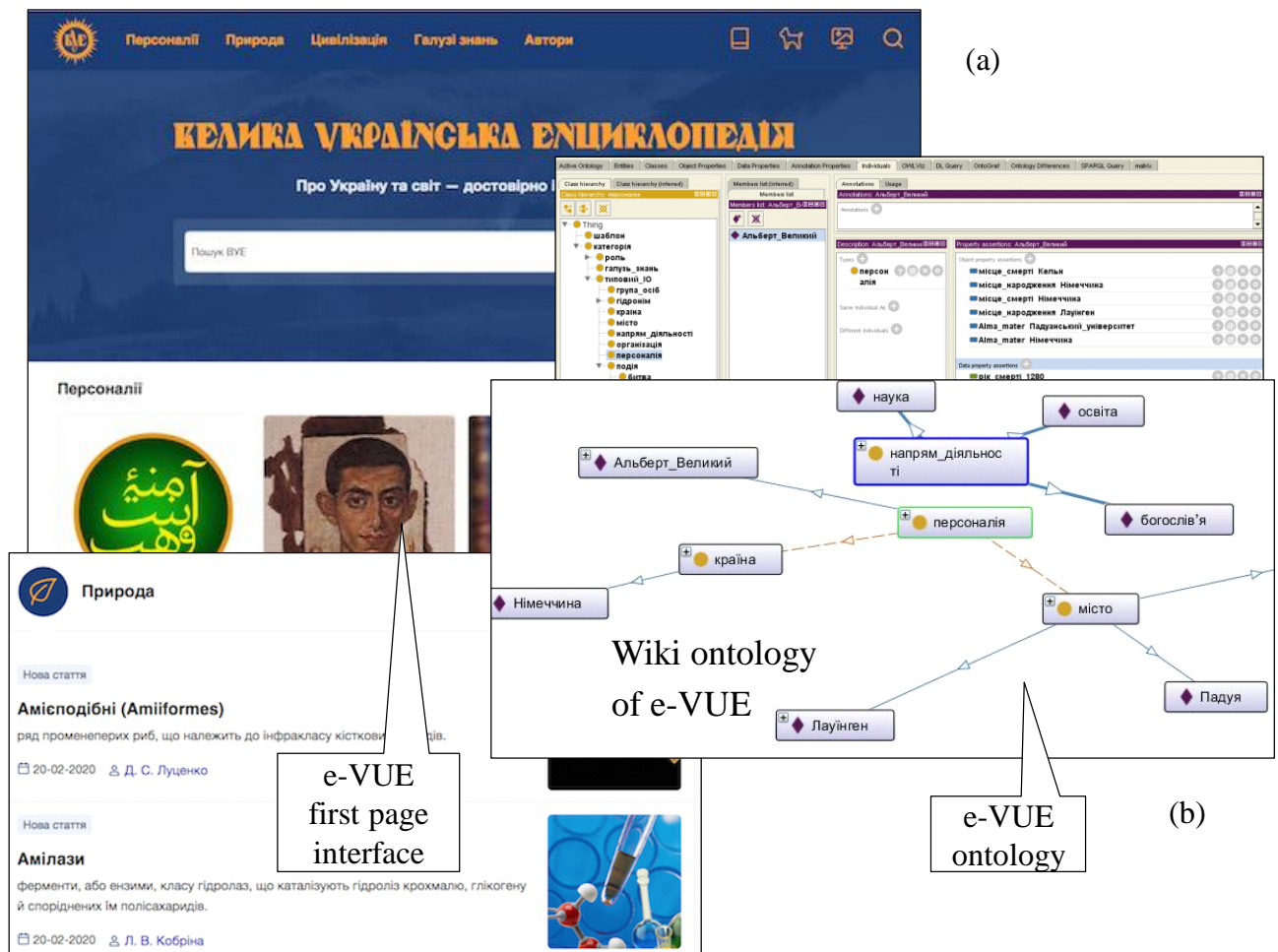


Fig. 1. e-VUE ontology (a) and user interface (b)

The main e-VUE elements described by its Wiki ontology are Wiki pages and semantic properties of these pages. In this ontology Wiki pages of e-VUE articles and research subjects (such as authors of articles, moderators of

scientific branches etc.) are represented by the individuals of classes, and Wiki pages of e-VUE categories and templates are represented by ontology classes (Fig. 1a). Semantic relations of e-VUE are represented depending on their domain of values by object properties (for relations with other pages) and data properties (for all others). For classification of e-VUE pages several independent taxonomies are used as a source of category hierarchies. Now e-VUE use taxonomies that classify articles by scientific topics; by sets of structural elements represented by templates of TIOs; by the publication form and degree of readiness etc. Categories can be applied to Wiki pages individually or together. Elements of e-VUE ontology are used in development of original user interface with various templates and semantic queries for data integration (Fig. 1b).

These principles provide a fundamental distinction between e-VUE and non-semantic Wiki resources, such as well-known Wikipedia: e-VUE supports arbitrary meaningful relations between concepts, while Wikipedia can use only two types of relations – a hierarchical relation of categorization “Class-Class Element” and page hyperlink. The use of relations with explicitly defined semantics significantly expands both the expressiveness of knowledge representation and the search and navigation ways. The structure of e-VUE knowledge base is formally fixed by Wiki ontology based on the described above formal model for semantically enriched Wiki resources [26]. This ontology is the result of the collaboration of domain experts that moderate the knowledge branches of encyclopedia with knowledge engineers. This formalization makes possible to describe complex knowledge structure more clearly and to prevent the reuse of category names with different meanings. e-VUE ontology allows visualization of encyclopedia categories relations, describe their characteristics and provide annotations. Wiki ontology is implemented in OWL using the Protégé ontology editor. Ontology provides the basic set of tags for semantic markup of the NL content of e-VUE articles.

We identify the following e-VUE TIOs (https://vue.gov.ua/Help:Basic_Templates_VUE) that are supported by Wiki templates: City; Continent; Personality (Fig. 2a) ; Battle; Country; Group of persons; Historical event; Lake; Mineral; Ocean; River; Sea; Taxon; Organization etc). Moreover, the special templates are used to represent other typical e-VUE objects distinct from encyclopedia articles (authors, moderators, media content, etc.): e-VUE Author (Fig. 2b); Moderator; Illustration; Literature source; Video etc. Other additional TIO templates allow to structure content of special elements of the articles (e. g. Citation, Publication, Awards, etc.).

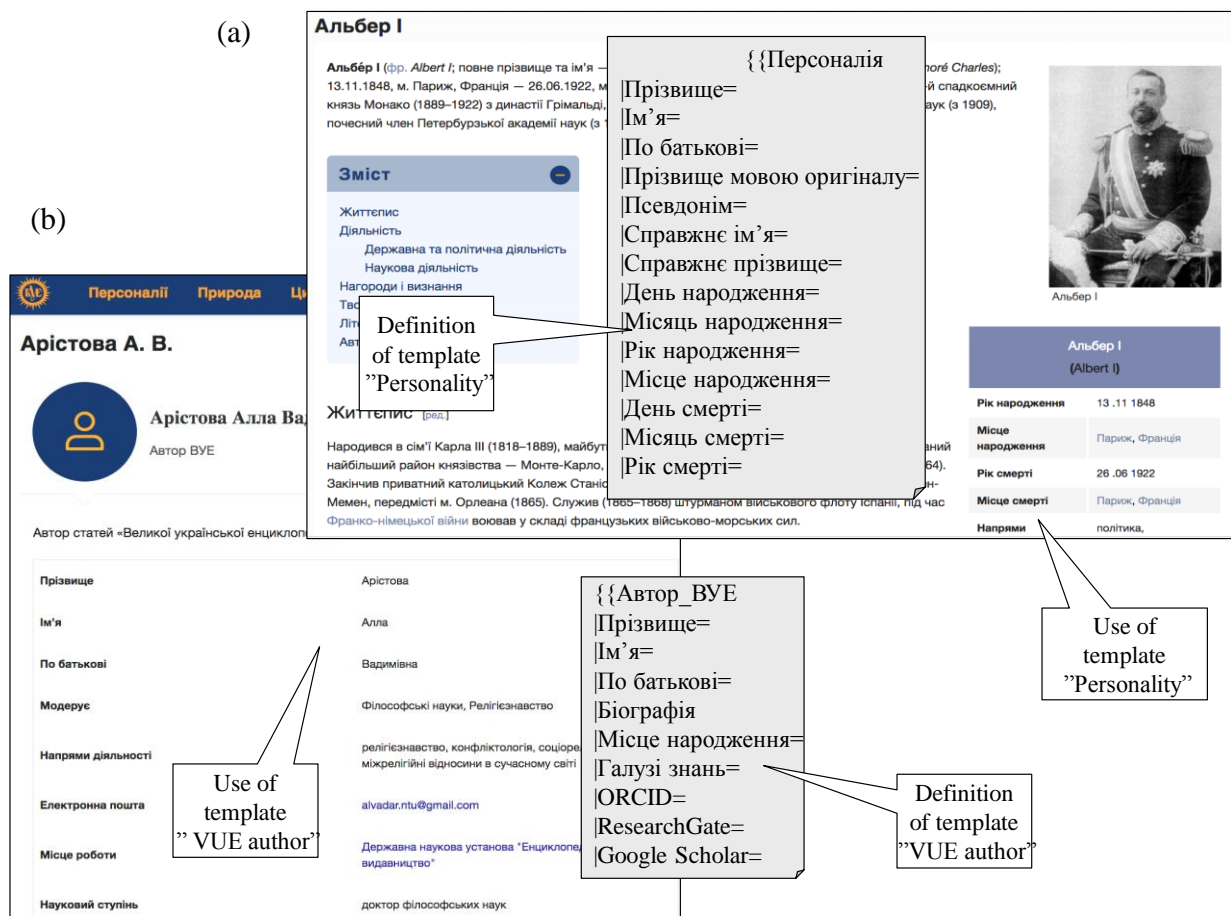


Fig. 2. Templates of TIO “Personality” (a) and “VUE author” (b) in e-VUE

Semantic search on e-VUE

One of the most important factors in the efficiency of IR use is the organization of search procedures that allows

the user access to information. Access efficiency depends on both access time and user-friendly interface. Use of semantics for search organization substantially optimizes this process and increases the possibilities of requests.

Content of Wiki pages from the viewpoint of information retrieval is a semi-structured data where content is marked up by Wiki tags defined by categories and semantic properties.

e-VUE search procedures are classified by search parameters into the following groups by: keywords placed in the page title and NL content; domain topics represented by Wiki categories; TIOs with specific sets of semantic properties defined by corresponding TIO templates; semantic similarity of Wiki page that is defined locally by the arbitrary collection of categories, semantic properties and their values. Search by keyword (for encyclopedias by the article title or by the initial letters of the title and words in natural language content) is traditional, practically implemented by all electronic encyclopedias. But it requires users to know the exact name of what they are looking for. This search provides the fastest access to the information. Search by domain topics is more flexible and uses the common classifiers (it makes search more formalized but complicates queries for ordinary users). The third and fourth ones take into account the semantics of the information needs of the user and allow to find semantically related articles. The e-VUE search by the type of IO is supported by the large number of templates for TIOs. Each template is associated with corresponding e-VUE category, and the user can apply these categories and their hierarchy for navigation. Such search is more convenient for users that know exactly what kind of information they need.

Search by semantics is more complicated for users and demands skills in SMW-QL query language and knowledge about specific properties of retrieved IOs. But the results of this search are much more personified and pertinent to individual user needs. Semantic MediaWiki supports SMW-QL query language for semantic search within Wiki resources. Such queries allow to filter pages by specified criteria and to display as a result of a request only the user-interested information, not the entire text of a Wiki-page. The results of such requests dynamically integrate the current information from various Wiki pages and display it in user-friendly form of diagrams, maps, tables and charts, etc. Modification of pages that satisfy to query conditions causes automatic update of query results.

One of the significant advantages of e-VUE as a semantic portal is the ability to find *semantically similar concepts* (SSC). Now these possibilities are realized in test version. Wiki page categories, semantic properties and their meanings, as well as natural-language content of e-VUE can be used to SSCs search. Such information cannot be generated by Semantic MediaWiki's built-in tools [27] and therefore this search is implemented through separately designed API requests. We demonstrate the capabilities of the described above approach by search of such local SBPs (by a combination of categories and semantic properties) as domain experts (fig. 3) and compatriots. This search is usually related to some TIO. In details this method is described in [25].

SemanticSimilarPages

Aviация
Семантично близькі сторінки

- 1 - Aviaцијна промисловiсть
- 2 - Автожир
- 3 - АН
- 4 - AviaВНIТО
- 5 - Aеронавтика
- 6 - Aероплан
- 7 - Антей (лiтак)
- 8 - Aеростат
- 9 - Aviаремонтний завод

Search persons by page categories

Avиация
Семантичний пошук персоналiй в множинi

Технiчнi науки
Aviацiйна та ракетно-космiчна технiка
Проектування, виробництво та випробування лiтальних апаратiв

Антонов, Олег Костянтинович	Сiкорський, Iгор Iванович	Адер, Клеман Агнес	Абрамович, Всеволод Михайлович	Авдеев, Сергiй Васильович
Айвинс, Марша Сью	Айзлi, Донн Фултон	Адлер, Георгiй Петрович (1886–1965)	Акiяма, Тойохiро	Анатра, Артур Анатольович (1913–1943)

Категорiї: Цивiлiзацiя | Технiчнi науки | Avиацияна та ракетно-космiчна технiка
Проектування, виробництво та випробування лiтальних апаратiв

SSCs of current page

SSC search options

Categories of current Wiki page

Personalities with the similar categories

Fig. 3. Search of experts for domain of selectes wiki page

Use of SSCs for information retrieval by MAIPS

MAIPS (Certificate of copyright registration for the work # 32015 "Computer program" Multi-Agent Information Search System "MAIPS". Authors – Grishanova I.Yu., Rogushina J.V. Join Date 09-02-2010) is an information retrieval system that uses the ontological knowledge representation, agent paradigm and Semantic Web technologies to information search on a semantic level [27]. MAIPS is an example of metasearch information system that processes domain ontologies as an instrument of representation of user professional or scientific interests. Ontologies provide domain knowledge about user interests and are used in MAIPS for personification of search results. MAIPS uses thesauri based on domain ontologies for user task representation. The system is focused on users with continual and structured information interests that require a regular supply of relevant information from the Web (functionally MAIPS is aimed at the implementation of complex reusable queries in a fairly narrow areas associated with professional or scientific interests of users; user requests can be repeated from session to session or be changed but deal with the same domain of user expertise).

MAIPS uses two types of ontologies – internal and external. Internal ontologies are created directly by MAIPS developers and can be replenished by new individuals (not by classes) during the interaction with MAIPS users. The main feature of such ontologies is the structure fixed by developers that helps in prediction of the algorithm convergence and calculation time. External ontologies allow to integrate into MAIPS the dynamic and distributed knowledge accessible by the Web. The search of such ontologies can be performed in the various ontological repositories [3], or they can be formed in the process of the user interaction with other intelligent applications. In general, we have no information about the complexity of the ontology structure, about DL on which they are based, completeness of their knowledge, and this lack prevents to predict the work time of algorithms for deep analysis and processing of these ontologies. Therefore MAIPS handles external ontologies by specially simplified algorithms that use only the most general properties of ontologies (e. g., processing only of taxonomic relations “Class-subclass”) for generation of task thesaurus. That is enough for the majority of information search tasks oriented on TIOs. But generation of thesaurus by domain ontology and NL task description requires from user some simple skills in operation with ontology components.

Structure of Wiki resource KB is much more intuitive and understandable for users. They can select the initial set of task-pertinent Wiki pages and define the criteria of similarity with the help of TIO characteristics. Therefore we propose to use the sets of SSCs generated by API requests to Wiki resources that take into account user-defined properties of concepts as task thesauri into MAIPS. Such requests exempts users from manual analysis of domain ontologies. We test this approach on examples from various spheres of sciences and find out that its effectiveness depends of IR pertinence to user information needs.

Conclusions

We consider practical aspects of development of intelligent Wiki-based IRs by MediaWiki technological platform and analyze the ways of its semantic extensions based on ontological representation of domain knowledge. We propose an ontological model of IR that formalizes its knowledge base structure that is based on explicitly represented system of domain-specific typical information objects (TIO) of this IR. Use of ontology corresponding to the set of Wiki pages (either with semantic markup or without it) provides new IR functions associated with semantic search and navigation. The ways of user interface adaptation to the IR specifics are developed. Proposed intelligent means of navigation, visualization and content analysis by processing of TIO features enriches IR functionality, reduces information access time and makes IR usage more efficient. We demonstrate the software implementation of proposed solutions by developing of portal version of the Great Ukrainian Encyclopedia (e-VUE) that contains heterogeneous multimedia content with complex structure. Wiki ontology of e-VUE and original methods of its processing used for this project extend the functionality of the portal in the area of search, navigation, integration and protection of content based on background domain knowledge.

References

1. Davies J., Fensel D., van Harmelen F. Towards the Semantic Web: Ontology-driven knowledge management. John Wiley & Sons Ltd, England. 2002. 288 p.
2. Semantic Web Challenge. <http://challenge.semanticweb.org/>.
3. Gladun A., Rogushina J. Ontology repositories as a mean of knowledge reuse for recognition of information objects. *Ontology for design*. 2013. N 1 (7). P. 35–50.
4. OWL Web Ontology Language. Overview. W3C Recommendation: W3C, 2009. <http://www.w3.org/TR/owl-features/>.
5. Horrocks I., Patel-Schneider P., van Harmelen F. From SHIQ and RDF to OWL: The making of a Web Ontology Language. *Web Semantics: Science, Services and Agents on the World Wide Web*. 2003. Vol. 1. P. 7–26.
6. OWL 2 Web Ontology Language Document Overview. W3C. 2009, <http://www.w3.org/TR/owl2-overview/>.
7. MediaWiki. – <https://www.mediawiki.org/wiki/MediaWiki>.

8. Grishanova I., Rogushina J. Development of access management methods to information from wiki resources. *Problems in programming*. 2020. N 9. P. 33–46. (in Ukrainian).
9. Wikipedia – <https://www.wikipedia.org>.
10. MediaWiki action API online documentation, <https://www.mediawiki.org/wiki/API:Query>.
11. MediaWiki RESTBase online documentation, <https://www.mediawiki.org/wiki/RESTBase>.
12. Wikidata Query Service online documentation, https://www.mediawiki.org/wiki/Wikidata_Query_Service.
13. Semantic MediaWiki API online documentation, <https://www.semantic-mediawiki.org/wiki/Help:API>.
14. Krötzsch M., Vrandečić D., Völkel M. Semantic mediawiki. *International semantic web conference*. 2006. P. 935–942. https://link.springer.com/content/pdf/10.1007/11926078_68.pdf.
15. Resource Description Framework (RDF) Model and Syntax Specification. W3C Proposed Recommendation. <http://www.w3.org/TR/PR-rdf-syntax>.
16. Rogushina J. Problems of ontological analysis use for knowledge representation in Wiki resources. *Problems in programming*. 2019. N 2. P. 17–37. (in Ukrainian)
17. Ushold M., Gruninger M. (1996) Ontologies: Principles, Methods and Applications, *Knowledge Engineering Review*. Vol. 11, N 2.
18. Guarino N. (1998) Formal Ontology in Information Systems. Formal Ontology in Information Systems, Proc. of FOIS'98. P. 3–15.
19. Rogushina J. (2017) Analysis of Automated Matching of the Semantic Wiki Resources with Elements of Domain Ontologies // *International Journal of Mathematical Sciences and Computing (IJMSC)*, Vol. 3, N 3. P. 50–58. <http://www.mecs-press.org/ijmsc/ijmsc-v3-n3/IJMSC-V3-N3-5.pdf>.
20. Rogushina J. (2016) Semantic Wiki resources and their use for the construction of personalized ontologies // *CEUR Workshop Proceedings* 1631. P. 188–195.
21. Rogushina J. (2019) Use of Semantic Similarity Estimates for Unstructured Data Analysis. Information technology and security. Proc. of the XIX International Scientific and Practical Conference ITB-2019. Kyiv, IPRI NAS of Ukraine. P. 118–126.
22. Rada R., Mili H., Bicknel E., Blettner M. (1989) Development and application of a metric on semantic nets. *IEEE Transaction on Systems, Man, and Cybernetics*, 19(1). P. 17–30.
23. Richardson R., Smeaton A.F., Murphy J. (1994) Using WordNet as a knowledge base for measuring semantic similarity between words. Working paper CA-1294, Dublin City University, School of Computer Applications, Dublin. <ftp://ftp.compapp.dcu.ie/pub/w-papers/1994/CA1294.ps.Z>.
24. Collins, A., Loftus, E. (1975) A spreading activation theory of semantic processing. *Psychological Review*, 82. P. 407–428.
25. Rogushina J., Grishanova I. (2019) Use of ontologies for search and navigation in online version of Great Ukrainian Encyclopedia. *Problems in programming*. N 4. P. 28–52. (in Ukrainian)
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. Kurydon A.M., Kyiv. P. 69–74. (in Ukrainian)
27. Rogushina J. (2016) Use of the Ontological Model for Personification of the Semantic Search // *International Journal of Mathematical Sciences and Computing (IJMSC)*, Vol. 2, N 1. P. 1–15. <http://www.mecs-press.org/ijmsc/ijmsc-v2-n1/IJMSC-V2-N1-1.pdf>.

Література

1. Davies J., Fensel D., van Harmelen F. Towards the Semantic Web: Ontology-driven knowledge management. John Wiley & Sons Ltd, England. 2002. 288 p.
2. Semantic Web Challenge. <http://challenge.semanticweb.org/>.
3. Гладун А.Я., Рогушина Ю.В. Репозитории онтологий как средство повторного использования знаний для распознавания информационных объектов. *Онтология проектирования*. 2013. № 1 (7). С. 35–50.
4. OWL Web Ontology Language. Overview. W3C Recommendation: W3C, 2009. <http://www.w3.org/TR/owl-features/>.
5. Horrocks I., Patel-Schneider P., van Harmelen F. From SHIQ and RDF to OWL: The making of a Web Ontology Language. *Web Semantics: Science, Services and Agents on the World Wide Web*. 2003. Vol. 1. P. 7–26.
6. OWL 2 Web Ontology Language Document Overview. W3C. 2009. <http://www.w3.org/TR/owl2-overview/>.
7. MediaWiki. – <https://www.mediawiki.org/wiki/MediaWiki>.
8. Гришанова І.Ю., Рогушина Ю.В. Розробка методів керування доступом до інформації у wiki-ресурсах. *Проблеми програмування*. 2020. № 1. С. 33–46.
9. Wikipedia – <https://www.wikipedia.org>.
10. MediaWiki action API online documentation, <https://www.mediawiki.org/wiki/API:Query>.
11. MediaWiki RESTBase online documentation, <https://www.mediawiki.org/wiki/RESTBase>.
12. Wikidata Query Service online documentation, https://www.mediawiki.org/wiki/Wikidata_Query_Service.
13. Semantic MediaWiki API online documentation, <https://www.semantic-mediawiki.org/wiki/Help:API>.
14. Krötzsch M., Vrandečić D., Völkel M. Semantic mediawiki. *International semantic web conference*. 2006. P. 935–942. https://link.springer.com/content/pdf/10.1007/11926078_68.pdf.
15. Resource Description Framework (RDF) Model and Syntax Specification. W3C Proposed Recommendation. <http://www.w3.org/TR/PR-rdf-syntax>.
16. Рогушина Ю.В. Проблеми використання онтологічного аналізу для подання знань у wiki-ресурсах. *Проблеми програмування*. 2019. № 2. С. 17–37.
17. Ushold M., Gruninger M. Ontologies: Principles, Methods and Applications. *Knowledge Engineering Review*. 1996. Vol. 11, N 26.
18. Guarino N. Formal Ontology in Information Systems. Formal Ontology in Information Systems. *Proc. of FOIS'98*. 1998. P. 3–15.
19. 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. <http://www.mecs-press.org/ijmsc/ijmsc-v3-n3/IJMSC-V3-N3-5.pdf>.
20. Rogushina J. Semantic Wiki resources and their use for the construction of personalized ontologies. *CEUR Workshop Proceedings*. 2016. Vol. 1631. P. 188–195.
21. Rogushina J. Use of Semantic Similarity Estimates for Unstructured Data Analysis. *Інформаційні технології та безпека*. Матеріали XIX Міжнародної науково-практичної конференції ІТБ-2019. К.: ІПРИ НАН України. 2019. P. 118–126.
22. Rada R., Mili H., Bicknel E., Blettner M. Development and application of a metric on semantic nets. *IEEE Transaction on Systems, Man, and Cybernetics*. 1989. 19(1). P. 17–30.
23. Richardson, R., Smeaton, A. F., Murphy, J. Using WordNet as a knowledge base for measuring semantic similarity between words. Working

- paper CA-1294, Dublin City University, School of Computer Applications, Dublin. <ftp://ftp.compapp.dcu.ie/pub/w-papers/1994/CA1294.ps.Z>.
24. Collins, A., Loftus, E. A spreading activation theory of semantic processing. *Psychological Review*. 1975. 82. P. 407–428.
 25. Рогошина Ю.В., Гришанова І.Ю. Використання онтологій для пошуку та навігації в онлайн-версії «Великої Української Енциклопедії». *Проблеми програмування*. 2019. № 4. С. 28–52.
 26. Рогошина Ю.В., Гришанова І.Ю. Онтологічна модель бази знань онлайн-версії «Великої української енциклопедії» та методи її застосування для семантичного пошуку та навігації. Енциклопедичний контент і виклики сучасного світу: Збірник матеріалів наукової конференції / За ред. Киридон А.М. – К.: Державна наукова установа «Енциклопедичне видавництво». 2019. С. 69–74.
 27. Rogushina J. Use of the Ontological Model for Personification of the Semantic Search. *International Journal of Mathematical Sciences and Computing(IJMCS)*. 2016. Vol. 2, N 1. P. 1–15. <http://www.mecs-press.org/ijmsc/ijmsc-v2-n1/IJMCS-V2-N1-1.pdf>

Received 28.02.2020

About the authors:

Rogushina Julia Vitaliivna,

Candidate of Physical and Mathematical Sciences,

Senior Research Fellow,

Institute of Software Systems of the National Academy of Sciences of Ukraine.

Publications in Ukrainian journals – 150.

Publications in foreign journals – 31.

<http://orcid.org/0000-0001-7958-2557>,

Grishanova Irina Yuriivna,

Researcher of Institute of Software Systems of the National Academy of Sciences of Ukraine.

Publications in Ukrainian journals – 18.

Publications in foreign journals – 3.

<http://orcid.org/0000-0003-4999-6294>.

Authors' place of work:

Institute of Software Systems NAS of Ukraine,

03181, Kyiv-187,

Akademician Glushkov Avenue, 40.

Phone: 066 550 1999.

E-mail: ladamandraka2010@gmail.com,

i26031966@gmail.com.