computing@tanet.edu.te.ua
www.computingonline.net

ISSN 1727-6209
International Scientific
Journal of Computing

Formalization of Search Context on Base of Ontologies and Multilinguistic Thesauruses

Anatoly Gladun [1], Julia Rogushina [2]

[1] International Research and Training Centre of Information Technologies and Systems,
National Academy of Sciences and Ministry of Education of Ukraine,
44 Glushkov Pr., Kiev, 03680, Ukraine; e-mail: glanat@yahoo.com
[2] Institute of Software Systems, National Academy of Sciences of Ukraine,
44 Glushkov Pr., Kiev, 03680, Ukraine; e-mail: _jjj_@ukr.net

Abstract: *For more relevant informational retrieval and matching of user request with metadata about informational recourses it is necessary to formulize the user knowledge about subject domain of search. We propose to use the ontologies and associated with them thesauri of the appropriate subject domains for representation of domain knowledge. The algorithms of formation and normalization of the multilinguistic thesauruses, and also methods of their comparison are given in this work.*

Keywords: *Ontology, Thesauruses, Informational retrieval.*

1. INTRODUCTION

The effective retrieval in the Internet becomes difficult and laborious for user that is forced to process a lot of that satisfy to formal request but don't correspond to his/her real information needs. Efficient informational retrieval has to be semantically oriented and based on knowledge of subject domain. That's why there is necessary to formulize the model of user interests domain (e.g., as ontology), link all information resources (IR) with some subject domains and then develop the algorithm for matching of IR domains with domain of user interests.

2. SEMANTICS OF THE INTERNET INFORMATIONAL RESOURCES

IR represented in the Internet can be classify on textual and multimedia ones, static and dynamic, structures and not structured etc., but every IR has some semantics and is concerned with some subject domain. In process of inaormation retrieval is very important to discover IR concerned with the domain interested to user.

Structures textual information in the Internet is mainly given in HTML and XML formats. The subject domain of textual IR can be define by two ways:

1) analyzing of IR textual content and
2) considering metadata of these IR.

There is a great deal of the widespread formats for a storing of audio and video information, 3D-scripts and images. The multimedia resources are accessible for indexation much worse than textual information. Therefore for multimedia IR only the second way is efficient. Metadata contains machine-readable information about the document that can be automatically processed by computer. Now the most perspective and common metadata model is RDF (Resource Description Framework) based on XML.

Metadata can be built in IR or be stored and updated independently of resources. With the help of RDF one can describe the structure of a IR and connect it with appropriate domain. RDF describes IR in a form of oriented marked graph - each IR can have properties that also can be IR or their collections. Most widespread set of elements for metadata specification of the Internet IR is Dublin Core Metadata Elements.

Initially World Wide Web technology was focused on work with static IR represented in the Internet. Now a lot of sites offer to the clients not only the documents, but also service (for example, sites of e-commerce). They use application servers that are ble to process the data entered by the user (queries, completed form etc.) and dynamically generate new IR depending on the parameters, specified by the user. Such dynamic component of the Internet grows much faster then static one and requires application of more complex information technologies. In this connection it is possible to consider a separate class of IR - *Web-services*. Web-service is a set of logically connected and program-accessible through the Internet functions that are based on three basic Web-standard: SOAP (Simple Object Access Protocol) - the protocol for sending of messages by the HTTP and other Internets protocols; WSDL (Web Services Description Language) - language for the description of program interfaces of Web-services; UDDI (Universal Description, Discovery and Integration) - indexing standard of Web-services.

3. STATEMENT OF PROBLEM

Efficient informational retrieval has to be semantically oriented and based on knowledge of subject domain. That`s why there is necessary to formalyse the model of user interests domain (for example, as ontology), link all IR with some subject domains and then develop the algorithm for matching of IR domains with domain of user interests.

4. THESAURUSES AND ONTOLOGIES AS MEANS OF DOMAIN KNOWLEDGE REPRESENTATION

By definition, "thesaurus" is the study of term usage in given domains associated to a human activity.

There are thesaurus for medical domain, mathematics, computer science, etc. A term is a sequence of words used in a given domain and which makes sense in this domain.

Therefore, thesaurus is on the domain knowledge side and it is used for domain description.

A thesaurus is a sort of terminological base: it is a collection of terms, plus a set of relations among them. In some ways a thesaurus can be a bridge from a terminological base to document indexing. It can be used as a normalization of indexing terms.

Terms of a thesaurus are used to describe a domain terms of a thesaurus are used to describe a domain Manual thesaurus building is a hard task but in this way, one can guarantee a good quality of the collected terms.

Manual thesaurus building is a hard task but in this way, one can guarantee a good quality of the collected terms. automatic Thesaurus building is quite human costless but the quality is not guaranteed. It relies on the content of document sources and also on the Natural Language treatment implemented.

Thesaurus is extracted from full text by means of syntax analysis.

The structure of thesauri is controlled by international standards that are among the most influential ever developed for the library and information field. The main three standards define the relations to be used between terms in monolingual thesauri (ISO 2788:1986), the additional relations for multilingual thesauri (ISO 5964:1985), and methods for examining documents, determining their subjects, and selecting index terms (ISO 5963:1985). ISO 2788 contains separate sections covering indexing terms, compound terms, basic relationships in a thesaurus, display of terms and their relationships, and management aspects of thesaurus construction. The general principles in ISO 2788 are considered language- and culture-independent. As a result, ISO 5964:1985, refers to ISO 2788 and uses it as a point of departure for dealing with the specific requirements that emerge when a single thesaurus attempts to express "conceptual equivalencies" among terms selected from more than one natural language [1].

Every domain has phenomena that people allocate as conceptual or physical objects, connections and situations. With the help of various language mechanisms such phenomena contacts to the certain descriptors (for example, names, noun phrases).

At present the usefulness of domain ontologies is generally recognized and is caused by their widely use. But the elements and the structure of domain ontologies are not defined standardly in different applications.

Now three main approaches to defining of a domain ontology exist. They are connected with the ways of ontological analysis application and deal with different sciences.

The first one – humanitarian approach – suggests definitions in terms understood intuitively but can't be used for solving of technical problems.

The second one – computer approach – is based on some computer languages (such as OWL, DAML+OIL) for representation of domain ontology and applied software that realized the processing of knowledge represented on these languages.

The OWL (Web Ontology Language) is being designed by the W3C Web Ontology Working Group as a revision of the DAML+OIL web ontology language. This description of OWL contains a high-level, abstract syntax for both OWL and OWL Lite, a subset of OWL. This syntax serves as part of a high-level specification for the formalism. A mapping from the abstract syntax to the OWL exchange syntax is also provided.

The description of OWL here abstracts from concrete syntax and thus facilities access to and evaluation of the language. A high-level syntax is used to make the language features easier to see. This particular syntax has a frame-like style, where a collection of information about a class or property is given in one large syntactic construct, instead of being divided into a number of atomic chunks (as in most Description Logics) or even being divided into even more triples, again for ease of readability. The syntax used here is rather informal, even for an abstract syntax - in general the arguments of a construct should be considered to be unordered whereever the order would not affect the meaning of the construct.

An OWL ontology is a sequence of axioms and facts, plus inclusion references to other ontologies, which are considered to be included in the ontology. All OWL ontologies are web documents, and can be referenced by means of a URI. Ontologies also have a non-logical component (not yet specified) that can be used to record authorship, and other non-logical information associated with a ontology.

The third one – mathematical approach – defines the domain ontologies in mathematical terms or by mathematical constructions.

We can consider that at first step of domain ontology building the humanitarian approach is used, then the mathematical model of ontology is constructed, and at last it`s software realization is developed.

Till now no generally accepted universal definition of domain ontology has been suggested. In [1] different definitions are analyzed. On the meaningful level a domain ontology will be understood as a set of agreements (domain term definitions, their commentary, statements restricting a possible meaning of these terms, and also a commentary of these statements). A domain ontology is:

- the part of domain knowledge that is not to be changed;
- the part of domain knowledge that restricts the meanings of domain terms;
- a set of agreements about the domain;
- an external approximation represented explicitly of a conceptualization given implicitly as a subset of the set of all the situations that can be represented.

All these meanings of the notion of domain ontology supplement each other.

We consider that a professional activity is a characteristic of a domain. This activity consists in solving different tasks. Task solving needs special knowledge, the same for all the tasks, that can be represented verbally. Therefore we can speak about special vocabulary of every domain that is used for specification of tasks and their solutions in this domain. A domain is considered as a set of the tasks, which are solved by specialists of this domain. When solving a task, a person uses a finite set of objects and relations among them.

These agreements are a result of understanding among members of the domain community.

5. THESAURI AND ONTOLOGIES AS MEANS OF DOMAIN KNOWLEDGE REPRESENTATION

Every domain has phenomena that people allocate as conceptual or physical objects, connections and situations. With the help of various language mechanisms such phenomena contacts to the certain descriptors (e.g., names, noun phrases). Professional activity is a characteristic of a domain. This activity consists in solving different tasks. Task solving needs special knowledge, the same for all the tasks, that can be represented verbally. Therefore we can speak about special vocabulary of every domain that is used for specification of tasks and their solutions in this domain. A domain is considered as a set of the tasks, which are solved by specialists of this domain. A domain ontology is the part of domain knowledge that restricts the meanings of domain terms, a set of agreements about the domain.

The formal model of domain ontology *O* is an ordered triple

$O = <X,R,F>$,

where

- X - finite set of subject domain concepts that represents ontology O;
- R - finite set of the relations between concepts of the given subject domain;
- F - finite set of interpretation functions of given on concepts and relations of ontology O.

An ontology is a specification of a conceptualization.

The word "ontology" seems to generate a lot of controversy in discussions about AI. It has a long history in philosophy, in which it refers to the subject of existence. It is also often confused with epistemology, which is about knowledge and knowing.

In the context of knowledge sharing, I use the term ontology to mean a *specification of a conceptualization*. That is, an ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents. This definition is consistent with the usage of ontology as set-of-concept-definitions, but more general. And it is certainly a different sense of the word than its use in philosophy.

The thesaurus can be considered as a special case of ontology.A thesaurus is a networked collection of controlled vocabulary terms. This means that a thesaurus uses associative relationships in addition to parent-child relationships. The expressiveness of the associative relationships in a thesaurus vary and can be as simple as “related to term” as in term A is related to term B [2]. The formal model of thesaurus is a pair Th = <T,R>, where T - finite set of the terms; and R - finite set of the relations between these terms.

A formal definition of a thesaurus designed for indexing is:

- a list of every important term (single-word or multi-word) in a given domain of knowledge; and
- a set of related terms for each term in the list.

Terms are the basic semantic units for conveying concepts. They are usually single-word nouns, since nouns are the most concrete part of speech.

Term relationships are links between terms that often describe synonyms, near-synonyms, or hierarchical relations.

6. USE OF THESAURUSES FOR IR RETRIEVAL

For taking into account semantics of area of user interests in process of retrieval of IR satisfying his/her informational need o it is necessary (fig. 1):

1. to generate the domain thesaurus corresponding to information needs of the user (by analysis of IR that this user considers relevant to this domain [4];
2. to construct the thesaurus for every IR known to IRS (simple dictionary without stop-words);
3. to compare the thesauruses of IR relevant to user query to IRS with the domain thesaurus and to find those ones that contain the maximum number of words in intersection.

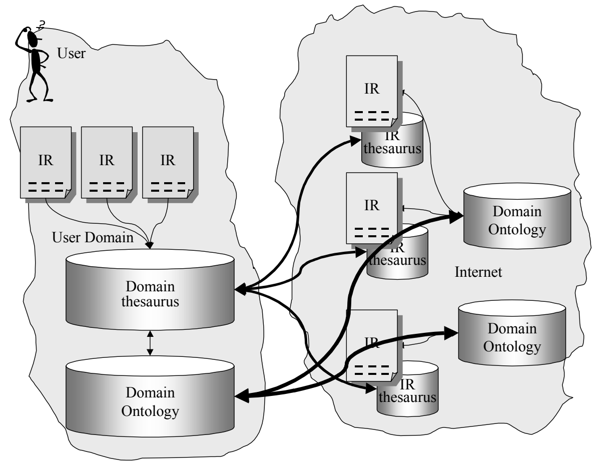


Fig.1 – Informational retrieval on base of domain thesauruses

At thesaurus construction it is necessary to use ontologies of the appropriate areas (with higher level in comparison with user domain to normalize the multilingual thesauruses). Normallization procedure is similar to stemming and provides for integrated processing of words in different morphologic forms and multilingual representations. Normalysed thesaurus contains relation between equivalent terms in different languages. As every thesaurus is constructed from the user point of view (which is reflected in user domain ontology), therefore it`s forming is the user task.

7. CONSTRUCTING OF DOMAIN THESAURUS

At first user should select the set of IR that he/she considers relevant to domain of his/her interests. Every IR is described by not empty set of the textual documents connected with this IR - text of content, metadescriptions, results of indexing etc. The domain thesaurus is formed as a result of the automated analysis of these documents (the user actions of the are reduced to constructing of semantic bunches - by linking of each word of the formed thesaurus with some term of domain ontology. Algorithm of domain thesaurus construction consists from the following steps:

1. **Formation of initial set** of the textual documents relevant to domain. At the input of algorithm the set A of the textual documents describing chosen IR comes (documents from A can have the coefficients of importance and the coefficients of IR relevance that allows to define differently weight of words from these documents for the IR description).

2. **Creation of domain information space**. For every document from A, $a_i \in A, i = \overline{1,n}$, the IR thesaurus $T(a_i)$ - dictionary that contains all words occurred in the document a_i - is constructed . The IR thesaurus is formed as union of the thesauruses a_i: $T_{IR} = \bigcup_{i=1}^{n} T(a_i)$, and domain thesaurus - as association of the IR thesauruses.

3. **Clearing of the thesauri**. User should specify dictionary for every $a_i \in A, i = \overline{1,n}$ containing a stop-words $s_{j,} s_j \in Voc$. It is necessary to remove words contained in $s_{j,} s_j \in Voc$ from the thesauri. Then all service information (e.g., marking tags) is rejected. The cleared thesauri $T`(a_i), \forall p \in T(a_i) \Rightarrow p \in T`(a_i) \vee p \in s_j$, $T`(a_i) \cap s_j = \varnothing$

thus are formed. The cleared IR thesaurus is constructed as association of the cleared thesauruses

a_i: $T_{IR} = \bigcup_{i=1}^{n} T(a_i)\, T`_{IR} = \bigcup_{i=1}^{n} T`_{IR}(a_i)$, and cleared domain tesaurus - as association of the IR thesauri.

4. **Linking of thesaurus with domain ontology.** To integrate processing of words with equivalent semantics (e.g., synonyms, translations of the term on different languages, various kinds of a spelling) the domain thesaurus is associated with some domain ontology (the user can form it himself, use some ready ontology, modify it or construct it himself).

Each word from the thesaurus it is necessary to link with one of the ontological terms. User has to do it manually on base of his own expirience and knowledge in appropriated subject domain., e.g. to link words "Lada de Mandraka" and "Staffordshir terrier" with ontological term "Dog".

For each word in the list of thesaurus terms user defines the ontology name, then selects some one from the list of ontology terms and confirms the link between them.

If the relation is lacking the word is considered as a stop-word or mark-up element (e.g., HTML tag) for domain described in ontology O and should be rejected. $\forall p \in T`(a_i) \exists t = Term(p,O) \in T_O$.

If word is significant for domain then go to step for extend the domain ontology.

The group of the IR thesaurus words connected with one ontological term named the **semantic bunch** $R_j, j = \overline{1,n}$ is considered as a single unit. $\forall p \in T`_{IR} \exists R_j = \{r : r \in T`_{IR}, Term(p,O) = Term(r,O)\}$.

It allows to integrate processing of semantics of the documents written on various languages and, thus, to ensure the multilinguistic analysis of the Internet IR.

5. **Extension of ontology.** If the IR thesaurus contains words that can`t be linked with ontological terms but user considers that these words are significant than it is necessary to add the appropriate terms to domain ontology, specify their connection with other terms of ontology and return to step 4.

We use Protégé to process the ontologies in OWL. This instrumental tool supports the extension of ontology by new classes and instances.

The Protégé project has come a long way since M.Musen first built the Protégé metatool for knowledge-based systems in 1987 [5]. Protégé can be run on a variety of platforms, supports customized user-interface extensions, incorporates the Open Knowledge Base Connectivity (OKBC) knowledge model, interacts with standard storage formats such as relational databases, XML, and RDF, and has been used by hundreds of individuals and research groups.

The original goal of Protégé was to reduce the knowledge-acquisition bottleneck by minimizing the role of the knowledge engineer in constructing knowledge bases. Now Protégé is a general-purpose environment for knowledge modeling.

Protégé allows the developers to build inference mechanisms in an entirely separate component, a problem-solving method, which could be developed independently from the knowledge base. These problem-solving methods (PSMs) were generic algorithms that could be used with different knowledge bases to solve different real-world tasks. Protégé extended the original two-step process—generating a knowledge-acquisition tool and using it to instantiating a knowledge base—with additional steps that dealt with the problem-solving method. This methodology consisted of:

1) developing or reusing a problem-solving method,

2) defining an appropriate domain ontology,

3) generating a knowledge-acquisition tool,

4) building a knowledge base using the tool, and

5) integrating these components into a knowledge-based system by defining mappings between problem-solving methods and specific knowledge bases.

The OntoViz tab plug-in used to give an alternative visualization for the Protégé knowledge base.

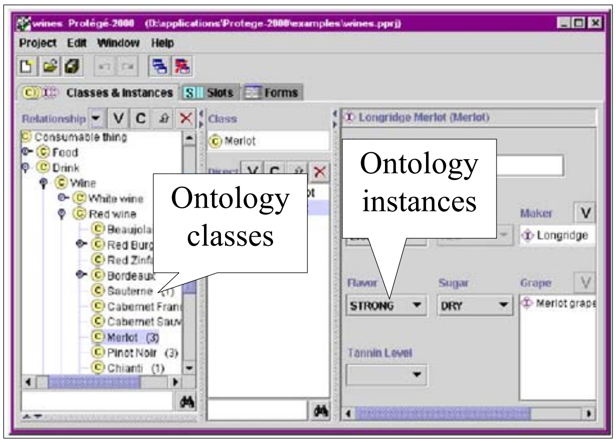


Fig.2 – The default user interface for Protégé

On base of Protégé user can create his own ontologies on base of existing ones that reflects his individual believes about subject domain. Such ontologies can not be global and widely used but they represent user knowledge and normalize his own domain thesaurus.

Relations between ontological terms and words from thesaurus are individual for every user or user's group. They reflect informational interests of user and represent his ability to information processing that is a function of his educational, cultural characteristics and experience etc.

6. **Construction of the normalized domain thesaurus**, i.e. association of all terms of domain ontology that are connected with words from the normalized IR thesaurus (Fig.3).

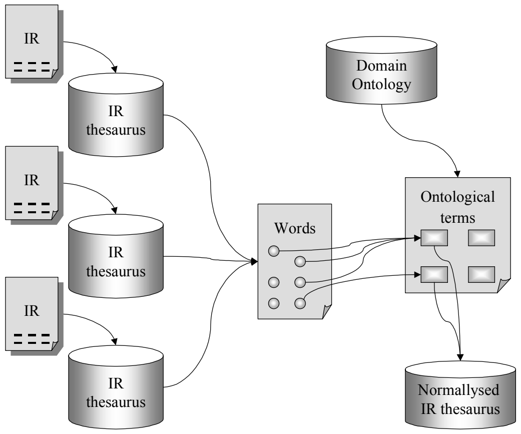


Fig.3 – Building of normalized IR thesaurus

The normalized thesaurus is a projection of set of the IR thesaurus words on set of the domain ontology terms . $L_{IR} = \{t : p \in T`(a_i), i = \overline{1,n}, t = Term(p,O) \in T_O\}$, and normalized domain thesaurus is a union of the normalized IR thesauruses (Fig.4). Informational retrieval systems (IRS) can use this set for representation of subject domain relevant with textual IR.

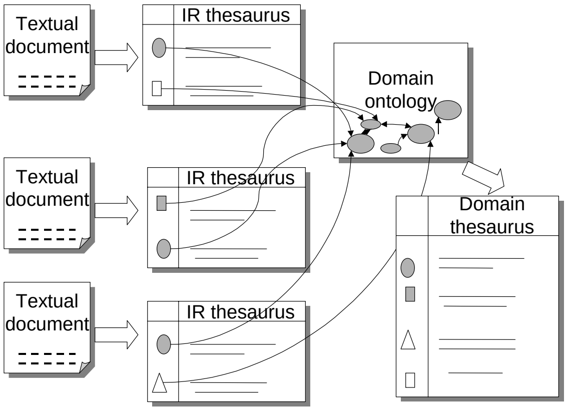


Fig.4 – Building of domain thesaurus

As result of the user query execution IRS finds a set of IR. The thesaurus of such IR is simple a dictionary that does not contain the relations between words (discovery of such connections from the text is rather difficult and in this case is not justified). IRS builds this dictionary automatically by IR content processing.

The algorithm of the IR thesaurus building consists of the following steps:

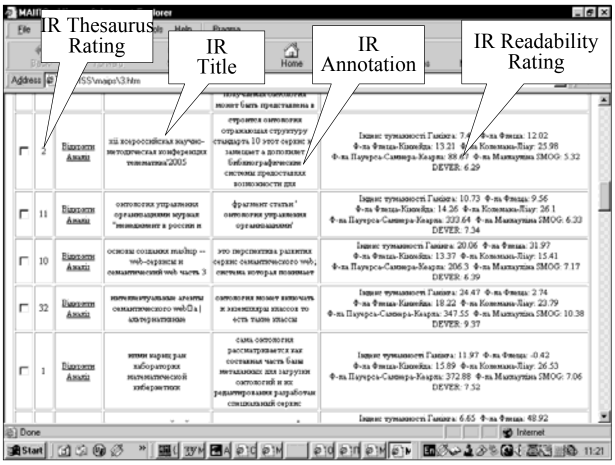
1. Formation of the initial IR set U, $U = \{IR_j, j = \overline{1,m}\}$.
2. Formation of the IR thesauruses from U. For each IR a thesaurus is formed and cleared.
3. Construction of the normalized IR thesauruses: for normalization the semantical bunches generated by the user during formation of the domain thesaurus are used.

8. ALGORITHM OF DOMAIN AND IR THESAURUSUS COMPARISON

The normalized IR thesauri L_{IR} and domain thesaurus L_{domain} are the subsets of the domain ontology terms O chosen by the user: $L_{IR} \subseteq Term(O)$, $L_{domain} \subseteq Term(O)$. If IR description contains more words linked with terms of domain interest for user (that is reflected in the normalized domain thesaurus) then it is possible to suppose that this IR can satisfy informational needs of the user with higher probability than other IR relevant to same formal query. Thus, it is necessary to find IR q satisfied the conditionst $f(q, L_{domain}) = max\, f(L_{IR}, L_{domain})$ where the function f is defined as number of elements in crossing of sets L_{IR} and L_{domain}: $f(A,B) = |A \cap B|$. If the various terms of the normalized thesauruses have for the user different importance it is possible to use the appropriate weight coefficients w_j that take into account their importance. In that case the criterion function is $f(A,B) = \sum_{j=1}^{z} y(t_j)$, where the function y is determined for all terms of domain ontology and $y(t_j) = \begin{cases} 0, t_j \notin A \vee t_j \notin B \\ w_j, t_j \in A \wedge t_j \in B \end{cases}$.

9. EXISTING AND PLANNED PROJECTS

Use of normallized thesauruses linked with domain ontologies is realyzes in intelligent IRS system MAIPS (http://progproblems.gradsoft.ua/maips-2006/) Resultes of retrieval by external IRS are filtered by indidvidual user thesauruses built on base of domain ontologies, corresponded IR and sequence of logical operations on thasauri (Fig. 5). In future we plan to transform MAIPS in the group of intelligent retrieval Web-services used the Semantic Web technologies..



				...может быть представлена в	
□	2	Відкрити Аналіз	xii всероссийская научно-методическая конференция телематика'2005	строится онтология отражающая структуру стандарта 10 этот сервис и замещает и дополняет библиографические системы предоставляя возможности для	Індекс туманності Ганінга: 7.4 Ф-ла Флеша: 12.02 Ф-ла Флеша-Кінкейда: 13.21 Ф-ла Колемана-Ліау: 25.98 Ф-ла Пауерса-Самнера-Кеарла: 88.67 Ф-ла Маклафліна SMOG: 5.32 DEVER: 6.29
□	11	Відкрити Аналіз	онтология управления организациями журнал "менеджмент в россии и	фрагмент статьи ' онтология управления организациями'	Індекс туманності Ганінга: 10.73 Ф-ла Флеша: 9.56 Ф-ла Флеша-Кінкейда: 14.26 Ф-ла Колемана-Ліау: 26.1 Ф-ла Пауерса-Самнера-Кеарла: 203.64 Ф-ла Маклафліна SMOG: 6.33 DEVER: 7.34
□	10	Відкрити Аналіз	основы создания mashup -- web-сервисы и семантический web часть 3	это перспектива развития сервис семантического web; система которая понимает	Індекс туманності Ганінга: 20.06 Ф-ла Флеша: 31.97 Ф-ла Флеша-Кінкейда: 13.37 Ф-ла Колемана-Ліау: 15.41 Ф-ла Пауерса-Самнера-Кеарла: 206.3 Ф-ла Маклафліна SMOG: 7.17 DEVER: 6.39
□	32	Відкрити Аналіз	интеллектуальные агенты семантического web[]a | альтернативные	онтология может включать и экземпляры классов то есть такие классы	Індекс туманності Ганінга: 24.47 Ф-ла Флеша: 2.74 Ф-ла Флеша-Кінкейда: 18.22 Ф-ла Колемана-Ліау: 23.79 Ф-ла Пауерса-Самнера-Кеарла: 347.55 Ф-ла Маклафліна SMOG: 10.38 DEVER: 9.37
□	1	Відкрити Аналіз	имени марчук рам лаборатория математической кибернетики	сама онтология рассматривается как составная часть базы метаданных для загрузки онтологий и их редактирования разработан специальный сервис	Індекс туманності Ганінга: 11.97 Ф-ла Флеша: -0.42 Ф-ла Флеша-Кінкейда: 15.89 Ф-ла Колемана-Ліау: 26.53 Ф-ла Пауерса-Самнера-Кеарла: 372.88 Ф-ла Маклафліна SMOG: 7.06 DEVER: 7.52
					Індекс туманності Ганінга: 6.65 Ф-ла Флеша: 48.92

Fig 5 – MAIPS user interface.

10. CONCLUSION

The proposed approach to use of domain ontology for creation and normalization of the IR thesauruses allows to fulfi informational retrieval at a semantic level abstracting from language of the IR description. The application of thesaurus measure of the information allows to offer to the user only understandable to him/her items of information that provides pertinence of information retrieval.

11. REFERENCES

[1] B.M. Matthews , K. Miller , M.D. Wilson "A Thesaurus Interchange Format in RDF". – http://www.limber.rl.ac.uk/External/SW_conf_thes_paper.htm

[2] A. Kleshchev, I. Artemjeva "A Structure of Domain Ontologies and their Mathematical Models". – http://www.iacp.dvo.ru/es/.

[3] The differences between a vocabulary, a taxonomy, a thesaurus, an ontology, and a meta-model. – http://www.metamodel.com/article.php?story=20030115211223271.

[4] A. Gladun, J. Rogushina, and V. Shtonda "Ontological Approach to Domain Knowledge Representation for Informational Retrieval in Multiagent Systems", in Information Theories and Applications, V.13, N.4, 2006, pp.354-362.

[5] J.H. Gennari, M.A. Musen, R.W. Fergerson, W.E. Grosso, M.Crubezy, H. Eriksson, N.F. Noy, and S.W. Tu "The Evolution of Protege: An Environment for Knowledge-Based Systems Development" - http://smi.stanford.edu/smi-web/reports/SMI-2002-0943.pdf.

__Dr. Anatoly Gladun__ is a senior researcher at the International Research and Training Centre of Information Technologies and Systems (National Academy of Sciences and Ministry of Education of Ukraine) – IRTC). He holds a PhD in Department of Computer Sciences at the Electrotechnical University (Saint-Petersburg, Russia). His research interests include the Intelligent Software Agents (models, architectures, methodologies of development) and their Application for Information Retrieval and E-commerce; Network Management and Semantic Web, Semantic Web Services, Ontologies, Wireless Technologies.

He is an Associate Professor at the Department of Computer Science (half-time) at University "Kiev-Mogyla Academy".

__Dr. Julia Rogushina__ is a senior researcher at the Institute of Software Systems, National Academy of Sciences of Ukraine. She received her PhD degree in Computer Science in Glushkov`s Institute of Cybernetics, Kyiv.

Her research interests include the development and application of intelligent information systems; theory of software agents behavior, inductive knowledge acquisition, intelligent information retrieval, ontological analysis, Semantic Web technologies.

She is an Associate Professor at the Department of Information Systems (half-time) of Kyiv Slavistic University.