УДК 004.4.004.89

ОНТОЛОГІЧНЕ МОДЕЛЮВАННЯ ПОШУКУ ІНФОРМАЦІЇ З ПРОЕКТУВАННЯ В ЕЛЕКТРОННІЙ БІБЛІОТЕЦІ

Рідкокаша А. А., *к.т.н.*,
Шевченко Б. М.
Черкаський державний технологічний університет,
бульвар Шевченка, 460, м. Черкаси, Україна,
e-mail: redan@list.ru

Анотація. *Розглянуто варіант пошуку інформації, якщо додати онтологію проектної діяльності. Введено поняття потрібних при цьому онтологій, зокрема, користувача, бібліографії, менеджменту. Для них наведено схеми онтологій і приклад реалізації пошуку.*

Ключові слова: *онтологія, електронна бібліотека, проектування.*

ONTOLOGICAL MODELING OF THE SEARCH OF DESIGN INFORMATION IN ELECTRONIC LIBRARY

Ridkokasha A. A., *Ph.D.*,
Shevchenko B.M.
Cherkasy State Technological University,
Shevchenko blvd, 460, Cherkasy, Ukraine, 18006,
e-mail: redan@list.ru

Abstract. *An alternative information search by adding the ontology of project activity is considered. The notion of ontologies, needed in this case, in particular, a user, bibliography, management ones, is introduced. For them diagrams of ontologies and an example of search realization are given.*

Keywords: *ontology, electronic library, design.*

Провести повноцінний пошук інформації за задовільний час людина не здатна, а повністю формалізувати зміст текстів для пошукових систем неможливо. Вирішення проблеми пов'язано з переходом від зберігання та обробки даних до накопичення та обробки знань, зокрема, з переходом до технології Semantic Web, що базується на використанні метаданих для опису семантики інформаційних ресурсів і засобів обробки цих метаописів.

Завдання семантичного пошуку в електронній бібліотеці є спрощеним аналогом пошуку інформації в Інтернет у зв'язку з тим, що передбачається, що пошук буде здійснюватися за запитом користувача природною мовою.

На рис. 1 показана схема семантичного пошуку інформації.

Рис. 1. Діаграма потоків даних при пошуку

Введений користувачем запит, який підлягає лінгвістичному аналізу, розширюється за рахунок використання синонімів, потім перетворюється в ключові слова і надходить до пошукової машини. Пошукова машина повертає знайдені документи, які також підлягають лінгвістичній обробці, і формуються семантичні образи документів. Образи документів порівнюються з образом запиту, робиться висновок щодо релевантності кожного з документів і результати аналізу (документи, які були визнані релевантними) надаються користувачеві [1].

Центральне місце у такій моделі пошуку інформації займають онтології. Створення онтологій є складним процесом.

Є багато публікацій, де застосовується онтологія при проектуванні, наприклад, у науковому журналі «Онтология проектирования» (Росія, Самара http://agora.guru.ru/scientific_journal/), при проектуванні програмних засобів [2] та при побудові електронних бібліотек [3, 4]. Наявні публікації про розробку програмних систем із застосуванням онтологій [5, 6].

Формально під моделлю онтології Онт розуміють впорядковану трійку виду [2]:

Онт = <X, R, F>,

де X – скінченна множина понять (інакше концептів, термів) предметної області;

R – скінченна множина відношень між поняттями заданої предметної області;

F – скінченна множина аксіом і правил на поняттях чи/або відношеннях Онт.

X – завжди скінченна і непорожня множина. Якщо R і F – порожні множини, Онт трансформується в словник. Якщо R – порожня множина, а F – непорожня, маємо пасивний словник предметної області. Коли R і F – непорожні, маємо таксономію понять предметної області.

Недостатньо уваги приділено онтологіям бібліотеки, які забезпечують процес проектування. Створення такої онтології є нетривіальним і трудомістким завданням, а також залежить від того, які саме засоби використовуються для опису предметної області.

Постановка задачі побудови онтології проектування. Варіантом подолання цих труднощів є Semantic Web-серверне зберігання та надання інформації користувачеві. Основою цієї технології є онтологія, для побудови якої потрібно використовувати знання предметної області – глосарій і правила проектування конкретного засобу, що потрібен користувачу бібліотеки.

При проектуванні особливе значення має онтологія життєвого циклу – опис процесу проектування, що прояснює, які проектні рішення з яких причин прийняті, і яка їх значимість для подальших розробки та супроводу. Вона повинна описувати інформаційний потік процесу прийняття рішень і зв'язок з частинами системи і з вимогами, які мотивують ці рішення.

Результати роботи. Процес проектування розглядаємо як послідовність станів проекту системи. Перехід між станами відбувається через проектне рішення (діяльність, активність) проектувальника. Тому головним є поняття «проектне рішення». Перехід між станами моделюється через властивість hasPreviousDesignDecision, зв'язок між «проектним рішенням» і вимогою – через властивості isBasedOn і requires. Відповідність (властивість hasRelevance) «проектного рішення» оцінює, наскільки воно може виконати вимоги (the requirements), інша властивість дозволяє оцінити вартість реалізації проектного рішення. Властивість «hasReason» – для мотивації прийняття рішення (наприклад, про введення деякого сервісу або зміну його визначення). Онтологія ЖЦ пов'язує частини проекту з частинами специфікації і з вимогами, які зумовили виведене «проектне рішення».

При концептуальній постановці завдання онтологія бібліотеки для проектування повинна включати в себе:

– організаційні онтології, що описують організаційно-функціональну структуру електронної бібліотеки, структуру даних з декількох складових з різними типами даних, які можуть бути пов'язані з іншими матеріалами, зв'язками типу «частина/ціле» або «послідовність»;

– управлінські онтології, які описують вимоги по керуванню, процесам проектування, і записи, що характеризують просування процесів до встановлених цілей.

Інформаційні онтології в більш широкому смислі можна виразити продукційним ланцюгом: «вхідне повідомлення» —> «система знань» —> «реакція», суть якого допускає мультидисциплінарну системну інтеграцію формального логічного представлення структури та правил виведення, методів і засобів проектування програмних систем, зокрема архітектури програмної системи.

При розробці онтології вихідним для організації роботи з семантикою інформації є розробка класифікаційних схем роботи елект-

ронної бібліотеки. При концептуальній постановці проблеми кореневою частиною такої структури може бути нормативне забезпечення проектних робіт. Онтології систем управління в такій схемі стають внутрішніми онтологіями, властивості класів та зміст (записи) яких будуть основними документами, що підтверджують вимоги нормативного забезпечення електронної бібліотеки.

Згідно з таким варіантом структури онтології проектної діяльності (рис. 2) при концептуальній постановці завдання розробки для вирішення завдань в електронній бібліотеці необхідна розробка, принаймні, таких класифікаційних схем:

– управління нормативним забезпеченням бібліотеки;

– управління проектами, у тому числі їх якістю;

– портфельне управління проектами (побудова групи програмних систем);

– організаційне управління;

– управління життєвим циклом проекту;

– управління записами;

– управління бізнес-процесами на основі онтологій проектної діяльності.

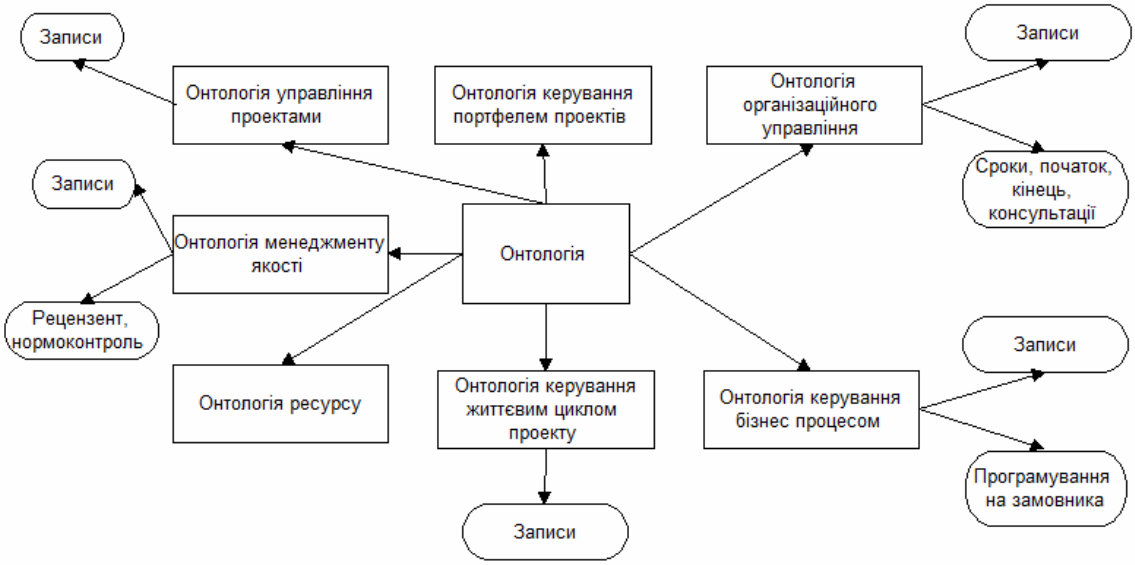

Рис. 2. Структура онтології проектної діяльності

Інструменти інженерії онтологій. Для створення і підтримки онтологій існує цілий ряд інструментів, які виконують підтримку документування онтологій, імпорт і експорт онтологій різних форматів і мов, підтримку графічного редагування, управління бібліотеками онтологій і т.д.

Для вирішення цього завдання можливе використання різних рівнів ступеня деталізації онтології (RDF Schemas, SCOS, OWL Lite, OWL DL, OWL Full). Для надання інформації за запитами користувачів у рамках, зокрема, декларативної мови запитів RDF Querry на початковому етапі розвитку, наприклад в електронній бібліотеці, розумно обмежитися найбільш простим мовним діалектом OWL Lite, який може бути реалізований на основі класифікаційної ієрархії термінів проектної діяльності та деяких простих умов узгодженості сутностей, зокрема з використанням FOL-характеристик властивостей.

Як вже було сказано вище, інструменти інженерії онтологій використовують спеціалізовані мови. Сьогодні виділяють три основні класи мов опису онтології:

– традиційні мови специфікації онтології: Ontolingua, CycL і мови, що базуються на дескрипційній логіці (такі як LOOM), також мови, побудовані на фреймах (OKBC, OCML, Flogic);

– пізніші мови, в основі яких Web-стандарти (XOL, SHOE, UPML);

– спеціальні мови для обміну онтологією через Web: RDF(S), DAML, OIL, OWL [7].

Виходячи з напряму дослідження, вибираємо OWL.

Концептуальна розробка онтології електронної бібліотеки матиме такі особливості. Властивості класів повинні включати в себе формально визначені в лінгвістичній онтології терміни діяльності з логічними відношеннями обмежувального характеру з такими об-

межувачами властивостей, як allvaluesFrom, Some ValuesFtom з кардинальними, обмеженими значеннями 0 або 1.

Необхідно правильно визначити таксономію проектної діяльності, а також властивості понять (слотів, атрибутів, ролей) і відношення між поняттями (зв'язки, залежності, функції). Наприклад, при створенні програмної системи на архітектурному етапі розробки алгоритмів розробник виявляє структуроване уявлення про програмну систему [6], тобто визначає склад і функції підсистем, а також взаємозв'язки між усіма підсистемами.

Після того, як формується набір множин проектних рішень, кожна з яких відповідає певному поняттю онтології $\forall c \in C$, $c=$, $\exists S_i=\{s_{i_1} ..., s_{i_m}\}$, ці множини можуть бути перетворені для більш ефективної обробки. Частину елементів множини $S_i=\{s_{i_1} ..., s_{i_m}\}$ користувач може ідентифікувати як одну підсистему і замінити елементом, що є загальною частиною цих елементів,

$$\exists\, l,\ l \subseteq s_{i_k},\ k=,\ p \geq 2,\ s_{i_k} \in S_i .$$

І, нарешті, для зв'язування внутрішніх онтологій між собою і з кореневою онтологією конструктивно скористатися можливістю картування онтологій із зазначенням еквівалентності між класами і властивостями, такими, як equivalent Class і equivalent Property із зазначенням того, що даний клас чи властивість в одній онтології еквівалентні класу або властивості в іншій.

Запропонований варіант базових онтологій електронної бібліотеки.

Онтологія, призначена для користувача. Як онтологія користувачів рекомендується модифікована онтологія FOAF для опису користувацьких профілів. Схема запропонованої онтології побудована з допомогою редактора Protégé і зображена на рис. 3. Базовим класом онтології є клас «Агент», один із підкласів якого – клас «Людина». У зв'язку з цим всі зв'язки класу «Агент» з іншими класами можна розглядати як зв'язок класу «Людина».

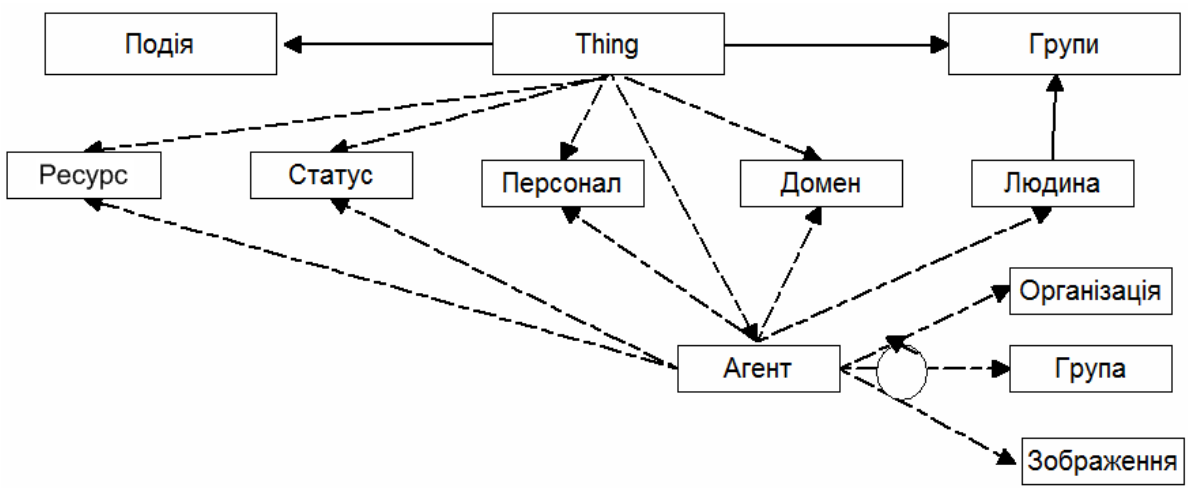

Рис. 3. Схема онтологій профілю користувача

Бібліографічна онтологія. Бібліографічна онтологія (або онтологія ресурсів) ґрунтується на бібліотечній онтології Marcont і при вирішенні завдань електронної бібліотеки буде використовуватися ця онтологія з деякими змінами. У цій онтології містяться необхідні класи і атрибути для опису електронних документів, таких як: книга, стаття, дисертація і т. д. Схема онтології зображена на рис. 4.

Клас «Ресурс» представляє всі типи електронних документів і має зв'язок з класом «Агент-Людина», в якому описуються завдання, автор, редактор, співавтор. Цей клас також зв'язується з класом «Домен», що є поняттям системи організації знань; поняттям можуть бути ключові слова, домен області документів.

Кожний документ може бути одним із типів ресурсів (книга, стаття, дисертація та ін.), а це означає, що завдання каталогізації документів можна вирішити з використанням таких типів. Завдання категоризації полягає в

анотуванні документів по своїх доменах (клас «Домен»).

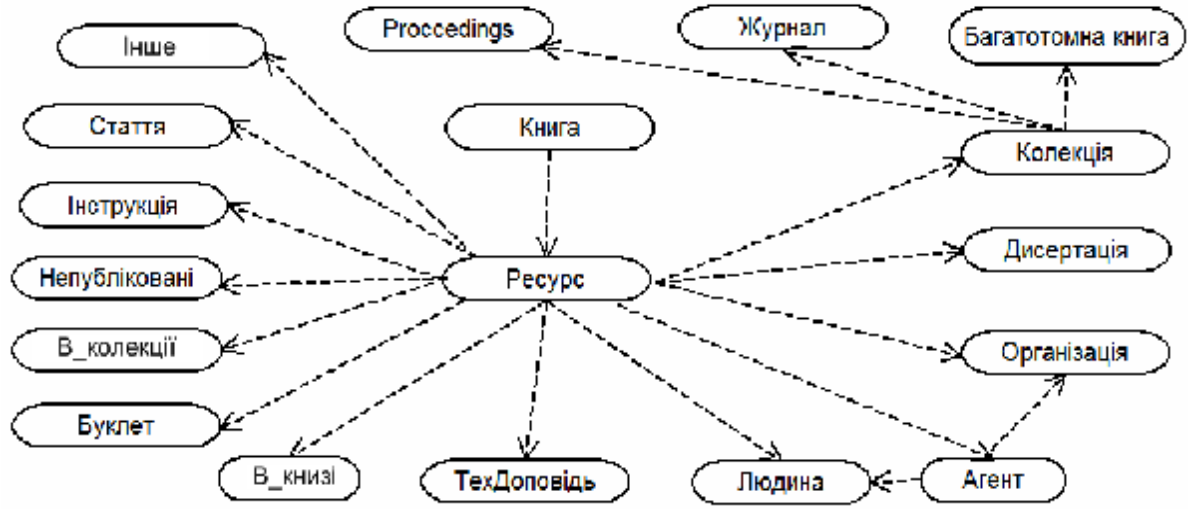

Рис. 4. Схема ресурсів

Приклад використання. Розглянемо приклад використання представлених вище онтологій для опису користувача, документа і його системних даних. Схема відношень між екземплярами класів цих онтологій показана на рис. 5, де 000000001 – Людина, 100000001 – Стаття.

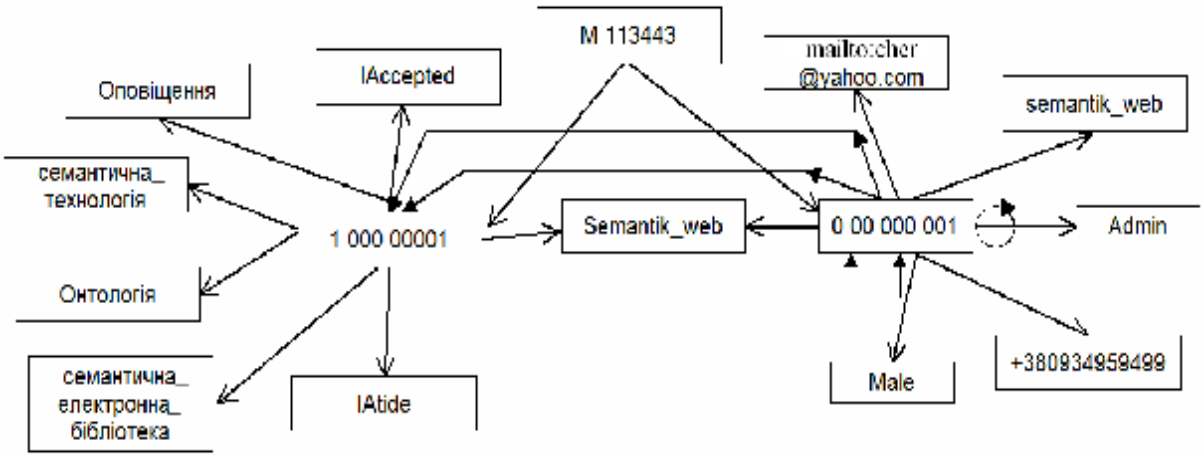

Рис. 5. Схема використання онтологій

Властивості, представлені текстом OWL:

```
<Person rdf:about="&users;0000000001">
<rdf:type rdf:resource="&owl;NamedIndividual"/>
<phonerdf:resource="&FloadUser;+380934959499"/>
<user:interestAgent rdf:resource="& users;000000001"/>
<user:interestResource rdf:resource="& Resource;100000001"/>
<users:gender rdf:resource="& users; Male"/>
<inbox rdf:resource="users;mailto:cher@yahoo.com"/>
<topic interest rdf:resource="&resource;Semantic_Web"/>
<structure is rdf:resource="&structure;Admin"/>
</Person>
<owl:NamedIdividual rdf:about="&resource; 100000001">
<rdf:type rdf:resource="http://www.marcont.org/ontology/2/0#Article"/>
```

<hasPage rdf:datatype=”&xsd;int”>&</hasPages>
<has Abstract rdf:datatype=”&xsd;string”> В статті розглядаються варіант онтології для розробки електронних бібліотек, їх специфікації для вирішення основних задач у системі управління електронними документами, а також проводиться вивчення методології створення онтологій.</hasAbstract>
<resource:hasTitle rdf:datatype=”&xsd;string”> Розробка онтологій для електронних бібліотек </Resource:hasTitle>
<hasYear>2011</hasYear>
<hasAutor rdf:resource=”&users;000000001”/>
<hasDomain rdf:resource=”&resource;semantic_Web”/>
<hasPublisher rdf:resource=”&resource;Повідомлення”>
<hasKeyword rdf:resource=”&resource;Онтологія”>
<hasKeyword rdf:resource=”&resource; семантична_технологія”/>
<hasKeyword rdf:resource=”&resource; семантична _електронна_бібліотека”>
<rdf:type rdf:resource=”&ontology;Resource”/>
<isUploadedBy rdf:resource=”&users;0000000001”>
<hasSubmissionStatus rdf:resource=”http://semDl.tpu.ru/ontology/structure/1.0/IAccepted”/>
<hasResourceType rdf:resource=”http://semDl.tpu.ru/ontology/structure/1.0/IArticle”/>
</owl:NamedIndividual>

Висновки. У роботі був проведений аналіз існуючих засобів і методів побудови онтологій. Розглянуто концептуальний підхід до онтологічного моделювання при проектуванні, створено і розглянуто на прикладі нову модель онтології. Результати роботи будуть використані в електронній бібліотеці. Розроблена комп'ютеризована підсистема з можливістю автоматизованого розширення онтологій і коригування тверджень, що автоматично додаються.

Список літератури

1. Марченко О. О. Моделювання семантичного контексту при аналізі текстів на природній мові / О. О. Марченко // Вісник Київського університету. – 2006. – № 3. – С. 230–235. – (Серія : фіз.-мат. науки).
2. Тарасов А. Ф. Онтологический подход к построению логических моделей программных систем / А. Ф. Тарасов, О. А. Лябик // Вісник Донбаської державної машинобудівної академії. – 2006. – № 1Е (6). – С. 51–54.
3. Соколова Н. В. Формирование электронной среды информационного обеспечения учебного и научного процессов на основе интеграции электронных коллекций [Электронный ресурс] / Н. В. Соколова, А. В. Кедрин, Д. Р. Храмцов // Электронные библиотеки. – 2009. – № 2. – Режим доступа : www.elbib.ru/index.phtml?page=elbib/rus/journal/2009/part2/SKH.
4. Привезенцев А. И. Базы знаний для описания информационных ресурсов в молекулярной спектроскопии. 3. Формирование базовой и прикладной онтологии [Электронный ресурс] / А. И. Привезенцев, Д. В. Царьков, А. З. Фазлиев // Электронные библиотеки. – 2012. – № 2. – Режим доступа :www.elbib.ru/index.phtml?page=elbib/rus/journal/2012/part2.
5. Дорошенко А. Е. Средства автоматизации разработки параллельных программ на основе онтологий и алгебр алгоритмов / А. Е. Дорошенко, Е. А. Яценко // Проблемы программирования. – 2008. – № 4. – С. 94–103.
6. Акуловский В. Г. Некоторые аспекты формализации архитектурного этапа разработки алгоритмов / В. Г. Акуловский // Проблемы программирования. – 2009. – № 2. – С. 3–11.
7. Овдей О. М. Обзор инструментов инженерии онтологий [Электронный ресурс] / О. М. Овдей, Г. Ю. Проскудина // Электронные библиотеки. – 2004. – № 4. – Режим доступа : http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2004/part4/op.

References

1. Marchenko, O. O. (2006) Modelling of semantic context in the analysis of texts in natural language. *Visnyk Kyivs'koho Universitetu, Phys.-Math. Science,* (3), pp. 230-235 [in Ukrainian].
2. Tarasov, A. F. & Lyabik, O. A. (2006) An ontological approach to the construction of logical models of software systems. *Visnyk Donbas'koji derzhavnoji mashinobudivnoji akademiji,* 1E (6), pp. 51-56 [in Russian].
3. Sokolova, N. V., Kedrin, A. V. & Hramtsov, D. R. (2009) The formation of electronic information environment for educational and scientific processes through the integration of digital collections. *Electronnye biblioteki,* (2), available at: www.elbib.ru/index.phtml?page=elbib/rus/journal/2009/part2/SKH [in Russian].
4. Privezentsev, A. I., Tsar'kov, D. V. & Fazliev, A. Z. (2012) Knowledge bases for the description of information resources in molecular spectroscopy. 3. Formation of basic and applied ontology. *Electronnye biblioteki,* (2), available at: www.elbib.ru/index.phtml?page=elbib/rus/journal/2012/part2 [in Russian].
5. Doroshenko, A. E. & Yatsenko, E. A. (2008) Automation means for the development of parallel programs based on ontologies and algebra algorithms. *Problemy programmirovaniya,* (4), pp. 94-103 [in Russian].
6. Akulovsky, V. G. (2009) Some aspects of the formalization of architectural phase of algorithms design, *Problemy programmirovaniya,* (2), pp. 3-11 [in Russian].
7. Ovdey, O. M. & Proskudina, G. Yu. (2004) Overview of Ontology Engineering Tools. *Electronnye biblioteki,* (4), available at: http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2004/part4/op [in Russian].