

УДК 004.89:004.423.4

*Розглянуто вдосконалення онтологічної моделі верифікаційною складовою з метою збереження та обробки інформації про процес гібридного виведення знань в розподілених інтелектуальних інформаційних системах. Запропоновано метод верифікації знань*

*Ключові слова: онтологія, база знань, верифікація, гібридне виведення*

*Рассмотрено усовершенствование онтологической модели верификационной составляющей с целью накопления и обработки информации о процессе гибридного вывода знаний в распределенных интеллектуальных информационных системах. Предложен метод верификации знаний*

*Ключевые слова: онтология, база знаний, верификация, гибридный вывод*

*Extension of the ontological model by the verification component for accumulation and processing of the information about knowledge hybrid inference in distributed intelligent information systems is considered. A method for knowledge verification is proposed*

*Key words: ontology, knowledge base, verification, hybrid inference*

# РОЗРОБКА МЕТОДУ ТА МОДЕЛІ ВЕРИФІКАЦІЇ ЗНАНЬ В ОНТОЛОГІЧНИХ СИСТЕМАХ

**М.В. Климова**

Кафедра штучного інтелекту  
Харківський національний університет радіоелектроніки  
пр. Леніна, м. Харків, 61166

## 1. Вступ

Сучасний етап розвитку теоретичних і прикладних досліджень в різних галузях науки характеризується зростанням уваги до задач моделювання інформаційних систем, що містять компоненти з інтелектуальною обробкою інформації. Широкого розповсюдження набули Web-орієнтовані інформаційні системи, зокрема, так звані портали знань [1], які використовуються з метою підвищення ефективності роботи в Web-просторі та пропонують доступ до ресурсів відповідної тематики для їх сумісного використання людьми та програмними агентами. Застосування таких систем в усіх сферах суспільного життя, а надто в процесах прийняття рішень, в тому числі в освіті та науці, зумовлює особливу увагу до методів їх верифікації, тобто забезпечення цілісності, несуперечливості та прозорості наданої системами інформації. Ця проблема залишається досі нерозв'язаною за умов розподіленої організації збереження та гібридного виведення інформації.

Потреба в забезпеченні якості систем, заснованих на знаннях на сьогодні отримала широке визнання серед світової дослідницької спільноти та здобула статус фундаментального компоненту розробки інформаційних систем.

До теперішнього часу було розроблено багато систем верифікації інформації, однак передусім під вери-

фікацією в існуючих системах розуміють лише обмежену кількість процедур, орієнтованих на конкретні предметні галузі чи задачі, для яких застосовують переважно статистичні та синтаксичні методи перевірки. Жоден з представлених системами методів не здійснює інтелектуального пояснення та доказу знань, які надає програмна система.

У зв'язку із цим досить актуальною є задача розробки методів інтелектуальної верифікації знань, що видобуваються з розподілених різноконтекстних різнорідних джерел.

## 2. Підходи до верифікації знань

Під формальною верифікацією систем розуміють демонстрацію цілісності (погодженості), повноти та точності програмної системи на кожному етапі та поміж етапами циклу розробки. Метою верифікації класичних інформаційних систем є, як правило, перевірка чітко визначених властивостей системи знань, відповідно до її специфікації. В залежності від видів властивостей, верифікація може бути обмежена до специфічних частин системи знань, таких як база знань, машина виведення або інтерфейс користувача, або зосереджуватись на специфічних аспектах функціональності системи, як наприклад, поведінка введення-виведення або дедуктивна поведінка.

Архітектура сучасних інформаційних систем передбачає модульну побудову, видобуток з розподілених джерел інформації системи, які знаходяться в різних точках фізичного простору, та розподілене збереження даних системи, можливість інтеграції та повторного використання як даних, так і елементів структури системи.

Це обумовлює структурну ізоморфність та підвищення надійності системи в цілому, але тільки за умови цілісності, правдивості та несуперечності знань, які зберігає система.

Саме тому особлива увага приділяється важливому компоненту інтелектуальної системи – базі знань – модулю знань системи, що містить абстрактні та специфічні знання про предметну галузь в машинозрозумілому вигляді [2, 3, 4].

База знань – це структура вигляду:

$$KB = (C_{KB}, R_{KB}, I, i_C, I_R), \quad (1)$$

де  $C_{KB}$  та  $R_{KB}$  – це множина концептів та множина відношень бази знань,  $I$  – множина екземплярів бази знань, функція  $i_C : C_{KB} \rightarrow \beta(I)$  ставить у відповідність екземплярам концепти, а  $i_R : R_{KB} \rightarrow \beta(I)$  – відношення.

Верифікація баз знань пов'язана з низкою питань, які мають розглядатися протягом верифікації [5]:

- залежність методу верифікації від моделі подання знань;
- залежність верифікації від предметної галузі, що позначається в збереженні мета-знань про проблемну галузь в самому методі верифікації;
- критерії верифікації – визначення різних типів аномалій, що можуть бути виявленими в БЗ;
- визначення структури моделі БЗ (плоска та ієрархічна), що впливає на вибір стратегії верифікації, монотонність БЗ;
- вибір стратегії верифікації: виявлення або виявлення та виправлення помилок;
- участь експерта з галузі під час виявлення помилок та аномалій в БЗ;
- критерії якості верифікатора та інші.

Як правило верифікації підлягають три складові бази знань: структура або архітектура бази знань; синтаксис визначень без огляду семантики; семантика бази знань.

В онтологічних базах верифікацію знань передусім пов'язують з правильною побудовою онтології, тобто перевіркою того, що її визначення виконуються вірно згідно з вимогами онтології та питаннями, що їх вирішує онтологія, або функціонують вірно відповідно до представленої галузі реального світу [6]. Згідно з даним визначенням верифікація стосується правильного подання предметної галузі і вимагає перевірки всіх її визначень.

Верифікація онтології [7] складається з верифікації: кожного індивідуального визначення та аксіом, сукупності визначень та аксіом, що подані в явному вигляді в онтології, визначень, що імпортовані з інших теорій, множини аксіом, які можуть бути виведені з інших визначень та аксіом.

На сьогодні розроблено низку методів верифікації знань, в тому числі і знань, представлених онтологічною моделлю подання. Але слід відзначити суттєві не-

доліки та обмеження існуючих методів, які обумовлюють неможливість всебічної семантичної верифікації знань [8, 9, 10, 11]:

- переважне використання синтаксичних та статистичних засобів верифікації;
- зосередженість існуючих систем семантичної верифікації знань на локальних специфічних задачах;
- ефективність принципів в теорії, але неефективність реалізації на практиці;
- відсутність механізмів автоматичного тлумачення накопиченої та розподіленої інформації;
- відсутність формування елементів перевірки на несуперечливість та достовірність інформації, виходячи з явної та неявної інформації, яка може накопичуватись в процесі функціонування систем;
- застосування лише внутрішніх інформаційних моделей системи для верифікації;
- потреба в керуванні з боку експерта в предметній галузі для розв'язання проблем з розумний проміжок часу.

Необхідними є інтелектуальні методи автоматичного видобутку й аналізу отриманої інформації, що є нереалізованим у сучасних системах. Крім того, проблемою залишається автоматичне тлумачення накопиченої та розподіленої інформації, формування елементів перевірки на несуперечливість та достовірність інформації, виходячи з явної та неявної інформації, яка може накопичуватись в процесі функціонування систем, та має супроводжувати будь-який процес автоматичного отримання результату, для того, щоб клієнт, який може бути кінцевим користувачем або програмним агентом, міг довіряти та приймати отриманий результат.

### 3. Верифікаційна модель

Для того, щоб гарантувати довіру користувача, в якості якого може виступати і інтелектуальний агент, і зовнішній Web-сервіс, до результатів роботи системи, що є надзвичайно важливим для систем, що підтримують процес прийняття рішень, необхідно забезпечити прозорість системи, тобто можливість перевірки всіх її результатів.

Семантична верифікація програмних систем, представлена в цій роботі, – це процедура демонстрації користувачу точності знань, які надає система, що складається з пояснення користувачу знань, які надає йому система, та доказу того, що знання не суперечать вимогам, що накладаються на систему, цілям системи та представлений галузі реального світу. Задача ускладнюється тим, що знання, які накопичуються системами, можуть збиратися з зовнішніх для системи, розподілених джерел інформації, наприклад, Web-сервісів та формуватися в результаті використання багатьох видів виведення, які можуть комбінуватися між собою та формувати складні послідовності.

Пояснення (*Explanation*) – це демонстрація шляху семантичного гібридного виведення певного знання.

Доказ точності знання (*Proof*) – це побудова обґрунтування несуперечності способів формування знань їх семантиці, що відбувається шляхом перевірки можливості побудови хоча б однієї інтерпретації, яка

є одночасно моделлю для множини елементів формування знань (*GenerationSet*) та самого знання.

На точність знання впливає спосіб формування, що може розглядатися як результат автоматичної або автоматизованої процедури в системі. Автоматична процедура формування знання представляє з себе видобуток з певного джерела знань або гібридне виведення, що застосовує виведення за аксіомами TBox онтології (*inference*), виведення за фактами ABox онтології (*reasoning*), виведення за SWRL-правилами, математичне обчислення та виведення знання у відповідь на SPARQL запит. Виведення відбувається з розподілених гетерогенних джерел – прикладань та Web-сервісів. Під автоматизованим формуванням знань розуміємо участь користувача, який явним чином вносить безпосередньо або, користуючись інтерфейсом, в базу знань інформацію, що зберігається згідно з внутрішньою структурою бази знань системи.

Верифікація знань стає можливою за умови розширення існуючої моделі подання знань (рис. 1), в якості якої в даному дослідженні обрано онтологічну модель.

Базова модель подання знання – це багаторівнева структура, заснована на дескрипційній логіці, основу якої складає множина RDF-тверджень, представлених триплетами “суб’єкт-предикат-об’єкт”, організована як позначений граф довільної розмірності, обмеження, словник та структуру для якого явним чином описує онтологія.

Розширена модель подання знань, запропонована в цій роботі, розширює існуючу триплетну форму, додаючи до неї верифікаційний компонент, перетворюючи триплети, що описують твердження, на тетради “суб’єкт-предикат-об’єкт-верифікація”, де під верифікацією розуміємо результат семантичної верифікації – семантичну структуру (*ver:Verification*), що зберігає пояснення та доказ відповідного твердження.

Верифікаційна модель (рис. 2) – це онтологія, яка описується мовою подання онтологій OWL, завдяки чому є придатною для обміну між Web-сервісами та клієнтами, що застосовують RDF/XML синтаксис.

Основними елементами верифікаційної моделі є: твердження (*rdf:Statement*), пояснення (*ver:Explanation*),

елемент підтвердження (*ver:ExplanationElement*) доказ (*ver:Proof*), дія (*ver:Action*), користувач (*ver:User*), джерело (*ver:Source*).

Ключовим поняттям в цій моделі є елемент підтвердження (*ver:ExplanationElement*) – одиниця пояснення, що зберігає опис стану середовища на різних етапах роботи системи. Елемент підтвердження генерується під час змін в середовищі, які відбуваються у відповідь на дії самої системи або користувача. Визначають різні типи елементів підтвердження для різних типів виведення знань.

Дія (*ver:Action*) в запропонованій моделі вказує на маніпуляцію, яка призвела до зміни середовища та представляє клас в верифікаційній онтології. Згідно зі способами формування знань доцільно розділити дії на прості та складні.

Простою дією є отримання знання від явного джерела знань, яким може бути користувач системи, зовнішній інформаційний ресурс, програмна система або Web-сервіс, структурування знання та збереження його в базі знань системи. Складна дія представляє виведення знання з наявних передумов шляхом застосування певного елемента виведення – правила, формули або запиту. Правилами виведення (*ver:Inference\_rule*) можуть бути SWRL- та Jena-правила, описані для даної предметної галузі, або внутрішні правила, що застосовуються машиною виведення (наприклад, модус-поненс).

Для кожного екземпляру класу «Дія», в залежності від його типу, визначається множина атрибутів, що описують даний екземпляр: дата маніпуляції (*date\_of\_action*), правило виведення (*has\_inference\_rule*), джерело дії (*has\_source*), формула, за якою відбулось обчислення (*calculated\_by\_formula*) та інші, які ставлять у відповідність даній дії множину об’єктів верифікаційної онтології.

Джерелом дії можуть бути різні об’єкти онтології: користувач, який здійснив маніпуляцію та вніс зміни у стан середовища системи, певний елемент виведення, застосування якого визначено для даного стану середовища, Web-сервіс, від якого надійшла інформація, Web-ресурс, з якого відбулося видобування інформації.

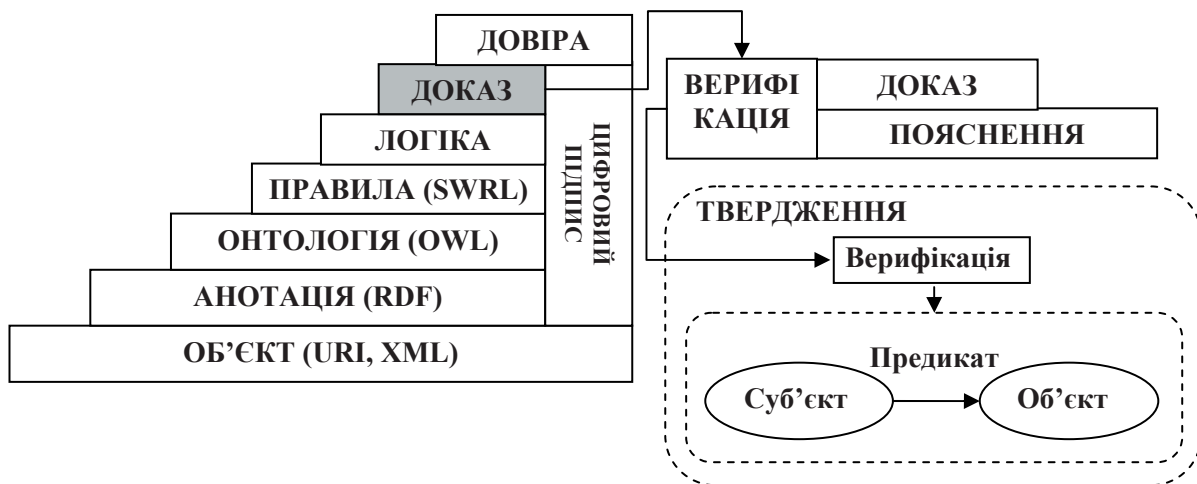


Рис. 1. Розширена модель подання знань

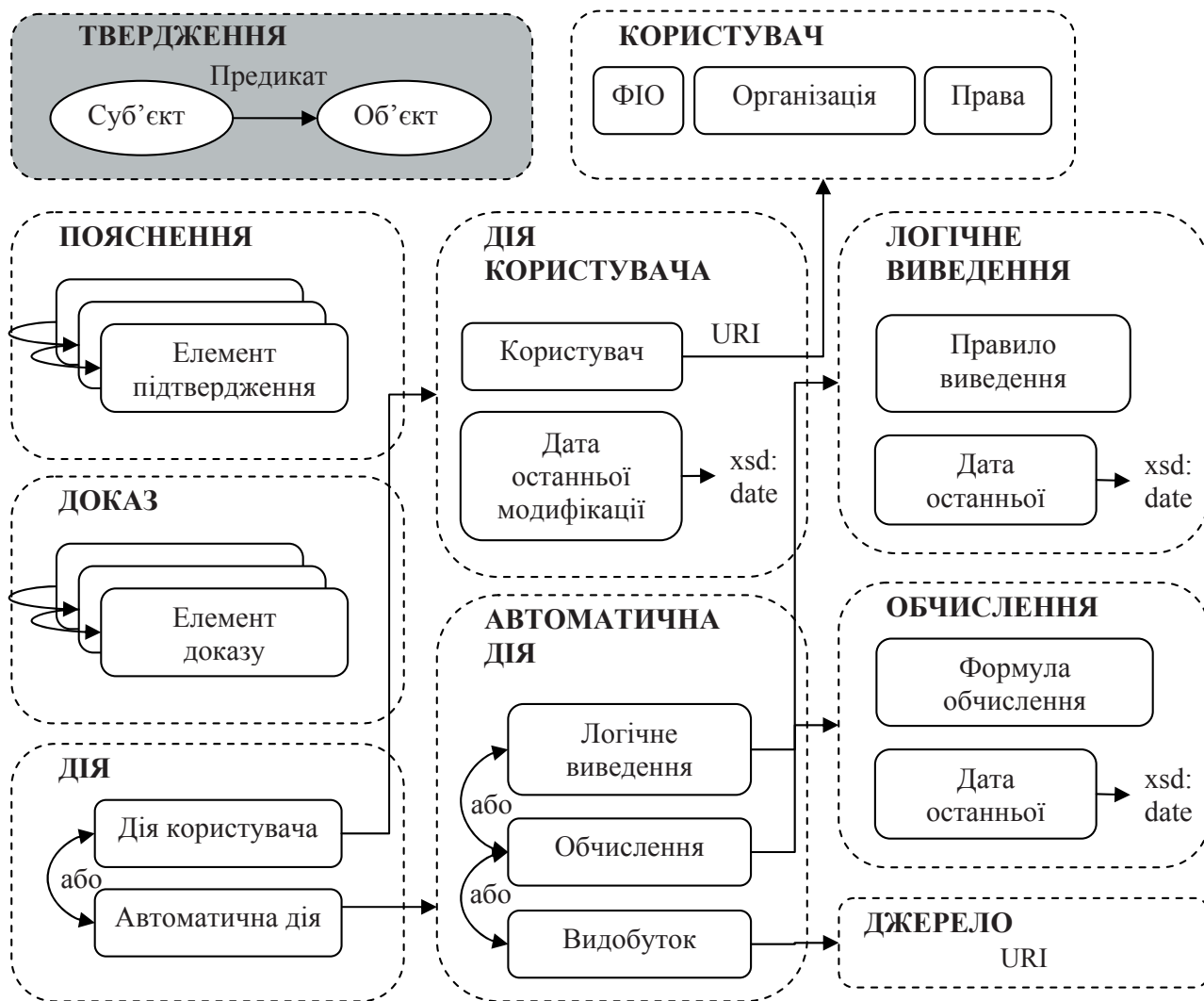


Рис. 2. Схема класів онтології та відношення між ними, що задають структуру верифікаційної моделі

Завдяки розширенню моделі та внесенню до неї елементів верифікації, можемо говорити про створення нового типу виведення знань в онтологічних системах, що дозволяє:

- створювати пояснення та доказ точності знань;
- виявляти зміни в механізмах виведення знань (порівняння структури пояснень);
- у випадку внесення змін до механізмів виведення знань в системі, перевіряти коректність існуючих знань;
- перевіряти ідентичність пояснень та доказів знань.

#### 4. Розробка методу верифікації знань

Метод верифікації знань, що містяться в онтологічних базах знань, зводиться до методів пояснення знань та доказу знань.

Метод пояснення знань (рис. 3) полягає у покроковому виконанні генерації елементу підтвердження (*ver:ExplanationStep*) для кожного твердження знання та формуванні пояснення, що об'єднує всі

згенеровані елементи підтвердження в одну структуру (*ver:Explanation*). Генерація елементу відбувається шляхом автоматичного створення об'єктів верифікаційної онтології, які зберігають мета-дані, що описують твердження, та процедуру їх формування, під час кожної зміни, що фіксується в онтологічній базі знань.

Пояснення твердження відбувається у відповідь на зовнішній запит (від користувача, зовнішнього Web-сервісу або прикладання).

В результаті застосування методу пояснень формується багатовимірне деревоподібна структура, вузлами якої є твердження, що описують стан середовища на різних етапах роботи системи: коренем є твердження, що потребує пояснення шляху виведення, листям (вузлам рівня n) – твердження, які визначають стан середовища до виконання початкової дії, що призвела до активації виведення твердження, що потребує пояснення шляху виведення, вузлами рівня n-1 – твердження, що описують початковий стан середовища, зміна якого відбулась в результаті початкової дії користувача; гілками якої є дії, що призводили до зміни середовища.

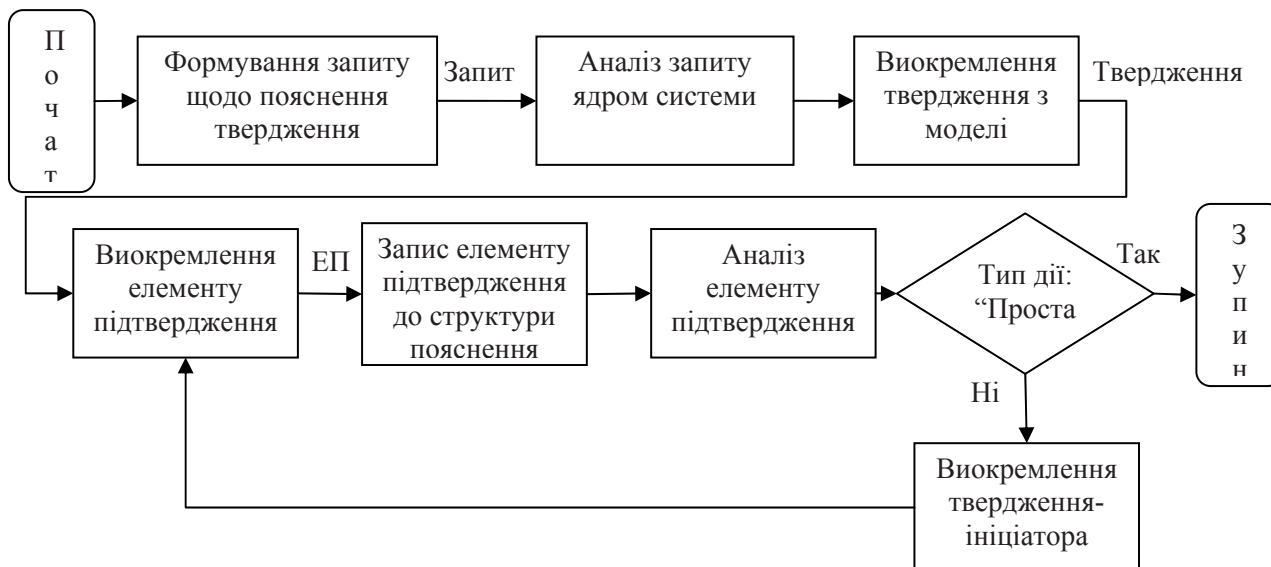


Рис. 3. Послідовність дій методу пояснення знань

### 5. Висновок

У статті розглянуто вдосконалення онтологічної моделі верифікаційною складовою з метою збереження та обробки інформації про процес гібридного виведення знань в розподілених інтелектуальних інформаційних системах.

Запропоновано новий тип виведення знань на основі елементів верифікації, що дозволяє створювати пояснення та доказ точності знань, виявляти зміни в механізмах виведення знань, у випадку внесення змін до механізмів виведення знань в системі, перевіряти коректність існуючих знань відповідно до нового виведення, перевіряти ідентичність пояснень та доказів знань.

Описано метод верифікації знань, що зводиться до методів пояснення знань та доказу їх точності.

Дослідження виконано із застосуванням підходів Semantic Web, що дозволило підвищити ефективність та надійність інтелектуальної обробки інформації в розподілених інформаційних системах.

### Література

1. Semantic Portals – Requirements Specification [Електронний ресурс] / W3 Consortium. – Режим доступу: <http://www.w3.org/2001/sw/Europe/reports/requirements-demo-2>. – Загол. з екрану.
2. Brachman, Ronald J. Knowledge Representation and Reasoning / Ronald J. Brachman, Hector J. Levesque // Morgan Kaufmann Publishers. – 2004. – 381 pp.

3. Fagin, Ronald. Reasoning about Knowledge / Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi // Journal of Logic, Language and Information. - Springer Netherlands. – 1999. Volume 8, Number 4. – 469-473 pp.
4. Levesque, Hector J. The Logic of Knowledge Bases / Hector J. Levesque, Gerhard Lakemeyer // MIT Press, Cambridge. – MA. – 2000. – 282 pp.
5. Zhang, Du. Knowledge base verification: issues and approaches // AAAI Technical Report. – AAAI. – 1993. – 148-149 pp.
6. Gómez Pérez, A. Ontological Engineering / A. Gómez Pérez, M. Fernández López, Chorcho O // Springer Verlag. – London, UK. – 2004.
7. Gomez-Perez A. Criteria to Verify Knowledge Sharing Technology // Knowledge Systems Laboratory. – WebMaster. – January, 1995.
8. Lozano-Tello, Adolfo. ONTOMETRIC: a method to choose the appropriate ontology / Adolfo Lozano-Tello, Asuncion Gomez-Perez // Journal of Database Management. – 2004.
9. McGuinness, Deborah L. The Chimaera Ontology Environment / Deborah L. McGuinness, Richard Fikes, James Rice, and Steve Wilder // Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI 2000). – Austin, Texas. – July 30 - August 3, 2000.
10. McGuinness, Deborah L. Explaining Answers from the Semantic Web: The Inference Web Approach / Deborah L. McGuinness and Paulo Pinheiro da Silva. // Journal of Web Semantics. – Vol.1 No.4., pages 397-413. – October 2004.
11. Horrocks, Ian. Description Logics in Ontology Applications. // Proc. of the 9th Int. Conf. on Automated Reasoning with Analytic Tableaux and Related Methods (TABLEAUX 2005). – Springer. – 2005. – pages 2-13.