

- Conference on Computer Science and Information Technologies, CSIT'2006, September 28th-30th, Lviv, Ukraine, p.56–58.
- [2] Серов Ю.О. Моделювання поведінки та класифікація учасників Веб-спільнот на основі нечітких множин // Вісник Національного університету «Львівська політехніка». –2008.– №610 – С. 218-228.
- [3] Серов Ю.О. Використання нечітких множин для моделювання активності учасників веб-спільнот // Р.Б. Кравець, Ю.О. Серов / Вісник Харківського національного університету радіоелектроніки "АСУ та прилади автоматики", вип. 141, 2007 р. С. 113-118.
- [4] Круглов В. В. Нечеткая логика и искусственные нейронные сети: Учеб. пособие. / Круглов В. В., Дли М. И., Голунов Р. Ю. // – М.: Издательство Физико-математической литературы, 2001. – 224 с. – ISBN 5-94052-027-8.
- [5] Brandes U Exploratory Network Visualization: Simultaneous Display of Actor Status and Connections / Ulrik Brandes, Jörg Raab, Dorothea Wagner // [web site] <http://www.cmu.edu/joss/content/articles/volume2/BrandesRaabWagner.html>
- [6] Freeman L. Visualizing Social Networks, / Linton C. Freeman, [web site] <http://www.cmu.edu/joss/content/articles/volume1/Freeman.html>
- [7] Серов Ю. О. Аналіз комунікативних процесів у Веб-спільнотах середовища Веб 2.0 / Серов Ю. О., Пелещин А.М., Слобода К. О. // Східно-Європейський журнал передових технологій №1/2 (37)/2009.- Харків.- 2009.- С.38-41
- [8] Pareto 80/20 Principle Rule <http://www.accountingissue.info/pareto-80-20-principle-rule.html>
- [9] Пелещин А. М. Позиціонування сайтів у глобальному інформаційному середовищі (монографія) / А. М. Пелещин — Львів: Видавництво Львівської політехніки, 2007. 260 с.
- [10] Page, The PageRank citation ranking: Bringing order to the Web / Page, Lawrence; Brin, Sergey; Motwani, Rajeev and Winograd, Terry // [web site] <http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf>

Підтримка запитів користувача по бібліотечних ресурсах на основі баз знань метаданих

Тарас Стисло

Науково-технічна бібліотека Івано-Франківського національного технічного університету нафти і газу, Україна, м. Івано-Франківськ, вул. Карпатська, 15, E-mail: taras_07@mail.ru

The formal logical approach for representation of user queries to library resources based on knowledgebases of metadats is introduced that allows a way of knowledge based unification for IT systems which are used in university library.

Ключові слова – digital libraries, metadatas, knowledges, knowledgebases, predicate queries.

I. Вступ

Основним питанням в процесі створення бібліотек є організація великої кількості інформації таким чином, щоб користувачі мали змогу знайти релеванту інформацію. Для цього створено складні класифікаційні схеми та правила функціонування каталогів для створення метаданих. Метадані описують інформаційні ресурси, які є в бібліотеці. Метадані надають інформацію про корисність даних ресурсів, забезпечуючи основу для пошуку та визначення інтелектуального та інформаційного контексту релевантності [1].

Цифрові бібліотеки висувають більші вимоги до метаданих, ніж традиційні бібліотеки. У даному випадку кількість інформації є набагато більшою, а доступ забезпечується різноманітними засобами, які допомагають користувачу визначити та задовольнити свої інформаційні потреби. Більшість з цих послуг надається без допомоги людини. Таким чином, важливо, щоб метадані, які використовуються в цифрових бібліотеках, були такими які можуть опрацьовуватися засобами новітніх інформаційних технологій. Цифрові бібліотеки мають надавати пояснювальну інформацію про їх склад для переформулювання запитів, при-

стосування сервісів до завдань користувача, відслідковування процесів формування зв'язків між ресурсами та користувачем і таке інше.

Одним із ключових питань з точки зору математичного моделювання процесу побудови інформаційних інтелектуальних систем для бібліотечної справи є спосіб представлення знань, на основі якого система повинна приймати рішення в певній ситуації. Таким чином представлення знань повинно бути задано способом, який дозволяє перехід до представлення фрагментів інформації про бібліотечний ресурс в термінах структур баз знань (БЗ), зокрема баз знань метаданих як засобу підтримки логічного виведення на множині метаданих. Розглядатимемо таку базу знань з точки зору фактів і процесів, що призводять до їх зміни, тобто з погляду семантики і синтаксису такого представлення. Під синтаксисом будемо розуміти набір правил для поєднання символів в логічно коректні вирази, а під семантикою – спосіб інтерпретації виразів, що одержуються в результаті конкретних реалізацій синтаксичних правил.

Ціллю даного дослідження є розробка формально-логічного апарату підтримки діалогу користувача з інформаційно-інтелектуальними системами на основі баз знань метаданих та підтримки процесу побудови запитів користувача, що розглядаються в ході динаміки процесу оновлення і модифікації вихідної бази знань.

II. Побудова баз знань метаданих

Ми використовуємо метадані, що базуються на формальних онтологічних принципах підтримки належного

обчислювального обґрунтування. Формальні онтологічні принципи використовують логіку для визначення зв'язків з іншими концепціями. Тому множину інформаційних ресурсів бібліотеки ми розглядаємо виходячи з множини онтологічних концепцій. Це дає можливість досліджувати зв'язки між властивостями, що визначені в онтології, а також зв'язків з бібліотечними входженнями. Так, наприклад, можна визначити, коли одна послуга є специфікацією іншої, навіть якщо це співвідношення не представлено у визначенні. Тому заслуговує уваги виділення та представлення у вигляді БЗ множини метаданих цифрової бібліотеки.

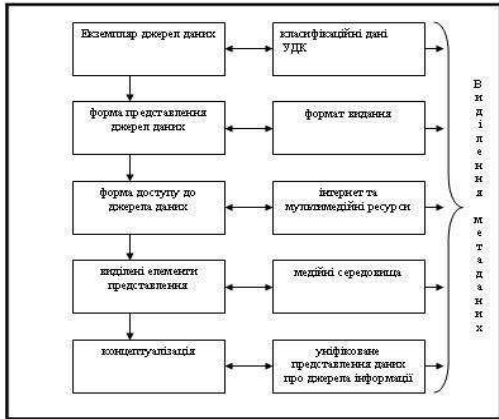


Рис. 1. Робоча ієрархія і зв'язані атрибути.

Формальна онтологія є відповідною технікою для моделювання комплексних доменів. Визначені концепції можуть утворювати мережі відносин без прив'язки до моделюючого інформаційного дерева. Концепції можуть мати багато описових вимірів (атрибутів), можуть бути також частково описані на будь-якому рівні концептуалізації (з будь-якою комбінацією вимірів) і можуть розглядатися з багатьох перспектив (оцінюватися різними послідовностями атрибутів). Вибір певного входження в онтології дозволяє підтримувати доступ з будь-якої перспективи на будь-якому рівні концептуалізації. Для порівняння, декларативні формалізми з меншою виразністю, такі як бази даних, створюють основу для побудови комбінацій та класифікацій вимірів відповідної бази знань.

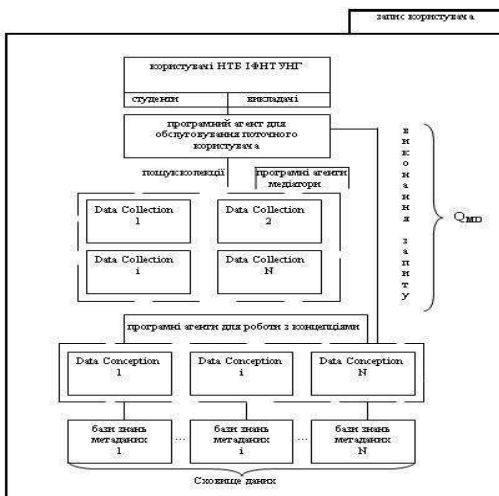


Рис. 2. Використання програмних агентів

Розглядатимемо БЗ метаданих онтології, як набір інформаційних сутностей атомарних предикатів з деякого скінченного інформаційного простору O . Всі зміни, що відбуватимуться в БЗ, будемо розглядати, як наслідок запитів користувача по бібліотечним ресурсам, що генеруються. Основою самих запитів є набір модифікаційних предикатних правил [2]. Розглядатимемо два типи правил:

$$KB_{MD+}(o) \ll KB_{MD+}(o_1), \dots, KB_{MD+}(o_l),$$

$$KB_{MD-}(p_1), \dots, KB_{MD-}(p_m)$$

$$KB_{MD-}(o) \ll KB_{MD+}(o_1), \dots, KB_{MD+}(o_l),$$

$$KB_{MD-}(p_1), \dots, KB_{MD-}(p_m)$$

де $o, o_i, p_i \in O$. Основна ідея такого запису правил полягає в тому, що $KB_{MD+}(o)$ означає, що атомарний предикат o повинен бути включений в БЗ KB_{MD} , а KB_{MD-} означає, що o – повинен бути виключений з бази знань.

Таким чином запити по метаданих задають логічні обмеження щодо БЗ. Зокрема правило (1) накладає на БЗ наступну умову: o належить БЗ, або принаймні одне з $o_i, 1 \leq i \leq l$ не належить БЗ, або принаймні одне з $p_j, 1 \leq j \leq m$ належить БЗ. Правило (2) стверджує, що для випадку, коли o і всі $o_i, 1 \leq i \leq l$ належать БЗ, а всі $p_j, 1 \leq j \leq m$ не належать БЗ, то можемо виключити o з БЗ, або вилучити одне з o_i , або додати одне з p_j .

Розглядатимемо тепер O , як деяку скінчену множину, елементи якої називатимемо атомами. Розглядатимемо також множину $\{KB_{MD_1}, \dots, KB_{MD_l}\}$ БЗ, таких, що $KB_{MD_i} \subseteq O$, де $1 \leq i \leq l$. Вирази виду $KB_{MD+}(o), KB_{MD-}(o)$, де $o \in O$ – деякий атом розглядатимемо як модифікаційні літерали. Модифікаційні літерали $KB_{MD+}(), KB_{MD-}()$ розглядатимемо як взаємообернені.

Нехай QMD – запит користувача по бібліотечним ресурсам і KB_{MD} – база знань метаданих. Введення семантики задасть специфікацію для множин БЗ, кожна з яких може бути вибрана як модифікація вихідної бази знань KB_{MD}^{noch} , після виконання запиту по метаданих.

Таку множину $\{KB_{MD_1}^m, \dots, KB_{MD_l}^m\}$ називатимемо множиною QMD – модифікацій вихідної БЗ KB_{MD}^{noch} . Кожна із QMD-модифікацій буде в свою чергу теж задовольняти всі обмеження, накладені на KB_{MD}^{noch} .

Кожний запит по метаданих можна розділити на дві частини: KB_{MD+} – правила, тобто такі, що:

$$\{\Delta \in Q_{MD} : head(\Delta) = KB_{MD+}(o),$$

для деякого $o \in O\}$

і KB_{MD-} – правила, такі, що:

$$\{\Delta \in Q_{MD} : head(\Delta) = KB_{MD-}(o),$$

для деякого $o \in O\}$

Причому

$$Q_{MD} = KB_{MD+}^+(Q_{MD}) \cup KB_{MD-}^-(Q_{MD}) \cup Q_{MD}^S,$$

де Q_{MD}^S - додаткові специфікації запиту по метада-них, які не описують жодних модифікацій стосовно вихідної бази знань. Кожна із таких модифікацій повинна бути описана в $KB_{MD+}^+(Q_{MD})$.

III. Висновок

В даній статті виконано побудову представлення формального логічного підходу опису запитів користувача по бібліотечних ресурсах на основі баз знань метаданих,

що дозволить уніфікувати використання знань в інформаційних бібліотечних системах. Подальші дослідження даного напрямку будуть зосереджені на розширенні введених означень і дослідженні їх властивостей.

- [1] P. Weinstein, W. Birmingham "Creating Ontological Metadata for Digital Library Content and Services". - International Journal on Digital Libraries.- 1998.
- [2] В.І. Шекета "Модифікаційні предикатні запити, як інструмент підтримки діалогу з користувачем в інформаційних системах на основі баз даних і знань" // Вісник Тернопільського державного технічного університету / Технічні науки – 2003.- Том 8-№4. - С.113-119.

Метод реферування множини документів на основі просторів даних та онтологічного опису

Наталя Шаховська, Василь Литвин, Володимир Крайовський

Кафедра інформаційних систем та мереж, Національний університет "Львівська політехніка", УКРАЇНА, м.Львів, вул.С.Бандери, 12, E-mail: natalya233@gmail.com

Abstract – The architecture and refer algorithm are described. There are described mining of refer and his main characteristics and functions.

Ключові слова – реферат, реферування, простір даних, онтологія.

I. Вступ

У різних галузях науки спостерігається експоненційний ріст обсягів експериментальних даних. Складність використання таких даних виникає внаслідок їхньої природної різноманітності (зберігання у різних системах, призначення для різних задач, різні методи опрацювання та зберігання тощо).

Розрив, який збільшується між джерелами даних і сервісами, приводить до необхідності пошуку нових шляхів організації рішення задач над множинними розподіленими колекціями даних і програм, які концентруються в спеціалізованих центрах даних і обчислювальних ресурсах.

Традиційно при рішенні певних задач фахівці використовують звичні для них джерела інформації і формують завдання з огляду на лише такі джерела. Очевидна неповнота інформації, яку вдається охопити при такому підході. Безліч джерел даних і сервісів, що існують в Інтернеті, їхня розмаїтість, викликають потребу в радикальній зміні такого традиційного підходу. Сутність цієї зміни полягає в тому що задачі повинні формуватися незалежно від існуючих джерел інформації, і лише після такого формулювання повинна здійснюватися ідентифікація релевантних завданню джерел, приведення їх до виду, необхідному для розв'язання задачі, інтеграція, ідентифікація сервісів, які дозволяють реалізувати окремі частини абстрактного процесу рішення завдання. Тільки після цього повинне здійснюватися конструювання конкретного процесу рішення на основі відібраних джерел даних і сервісів.

Цей підхід, що називається орієнтованим на проблему, повинен значно більше опиратися на використання техніки подання знань, ніж традиційний. Так, формулювання задачі повинне бути засноване на визначенні її проблемної області, термінології й системі понять, визначенні методів, алгоритмів і процесів рішення задачі. Однією з проблем при такому підході залишається та, що у всіх названих вище інфраструктурах дотепер відкритим є питання інтегрованого подання множинних джерел інформації. Інший підхід передбачає створення предметних посередників, що підтримують взаємодію між дослідником і джерелами за допомогою опису предметної області класу завдань (у термінах понять, структур даних, функцій і процесів рішення завдань). При цьому передбачається, що інформаційні джерела опубліковані в колективних сховищах, а операції ідентифікації потрібних джерел є їхніми основними операціями. Для цього в колективних сховищах підтримуються метадані для опису збережених у них інформаційних джерел [1]. Така інтеграція інформації надзвичайно важлива у задачі реферування, адже тут йде мова про опрацювання неструктурованих або частково структурованих джерел, поданих у різних форматах. Тому метою статті буде розроблення методів реферування множини неоднорідних джерел даних.

II. Основний матеріал

Реферування – це процес видобування найважливішої інформації з одного або декількох джерел для складання їхньої скороченої версії для потреб певних користувачів або задач [2]. **Реферат** – це семантично адекватний виклад основного змісту первинного документа, що відрізняється ощадливим знаковим оформленням, сталістю лінгвістичних і структурних ха-