

БАЗИ ДАНИХ, БАЗИ ЗНАНЬ ТА ІНЖЕНЕРІЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

УДК 681.3

**Рогущина Ю.В.**

**Інститут програмних систем НАН  
України**

# **РОЗРОБКА МЕТОДІВ ФОРМУВАННЯ ТА ПОПОВНЕННЯ ОНТОЛОГІЧНОЇ МОДЕЛІ СЕМАНТИЧНОЇ ПОШУКОВО- РЕКОМЕНДУЮЧОЇ СИСТЕМИ**

*Запропоновано онтологічну модель інтелектуальної пошуково-рекомендуючої системи, орієнтовану на функціонування у відкритому середовищі Web, соціального Web та Semantic Web. Розглянуто напрямки здобуття знань про користувачів, проаналізовано доцільність персонального тестування для створення груп користувачів зі спільними інтересами, що надає можливість колаборативного прогнозування оцінок результатів пошуку. Розроблено методи поповнення цієї моделі новими знаннями шляхом індуктивного узагальнення досвіду взаємодії користувача з системою, що забезпечує самонавчання пошуково-рекомендуючої системи, спрямоване на вдосконалення її роботи.*

*Предложена онтологическая модель интеллектуальной поисково-рекомендующей системы, ориентированную на функционирование в открытой среде Web, социального Web и Semantic Web. Рассмотрены направления получения знаний о пользователях такой системы, проанализирована целесообразность персонального тестирования для создания групп пользователей с общими интересами, которое предоставляет возможность колаборативного прогнозирования оценок результатов поиска. Разработаны методы пополнения этой модели новыми знаниями путем индуктивного обобщения опыта взаимодействия пользователя с системой, которое обеспечивает самообучение поисково-рекомендующей системы, направленное на усовершенствование ее работы.*

*Ontological model of the intelligent retrieval and recommending system that is oriented on the operation at the open environment Web, social Web and Semantic Web is proposed. Areas of knowledge acquisition about users of this system are considered, suitability of personified testing for creation of user groups with common interests that provide collaborative of search results forecasting ratings is analyzed. Methods for supplementation of this model by new knowledge by means of inductive generalization of interaction experience between user and system for self-learning that supports its work improvement are designed.*

**Ключові слова:** онтологія, семантичний пошук, Semantic Web, індуктивне виведення, рекомендуючі системи.

## **Вступ**

Проблема підвищення ефективності семантичного пошуку з метою задоволення інформаційних потреб користувачів є сьогодні однією з ключових при розробці найрізноманітніших інтелектуальних інформаційних систем (ІС), що функціонують у відкритому середовищі [1], де доступ до різних ІР забезпечується через Web, а для їхнього представлення можуть використовуватися найрізноманітніші моделі формалізації, способи і формати збереження, умови доступу і методи обробки. При цьому

ціллю пошуку все частіше стають знання про потрібні користувачам інформаційні об'єкти. Проблема здобуття знань з Web виявляється складовою частиною найрізноманітніших ІС і в загальному випадку пов'язана з розпізнаванням інформаційних об'єктів (ІО), що мають відношення до розв'язуваної користувачем задачі. Інформаційний об'єкт може мати досить складну й певним чином формалізовану структуру. Приміром, окремим випадком такого пошуку є знаходження семантичних Web-сервісів, що потрібні для вирішення проблеми користувача, або пошук

програмного агента, що реалізує відповідні функції.

Це обумовлено тим, що більшість сучасних ІС орієнтовані на використання і перетворення знань з цікавлячої користувача предметної області (ПрО), які не закладаються в ІС при її розробці, а здобуваються динамічно з доступних інформаційних ресурсів (ІР).

У найбільш загальному виді розв'язувана в даному дослідженні проблема полягає в наступному: щоб ефективно задовольнити потребу користувача в необхідній йому інформації, необхідно отримати формалізований опис цієї потреби, що спирається на знання про те, яку проблему вирішує користувач, про його персональні особливості та переваги, які ІО для цього потрібні, та про те, як відповідні ІО характеризують інші користувачі.

При цьому користувач, як правило, не має достатні навички для створення формалізованих описів задачі і предметної області (наприклад, у вигляді онтології), а якщо і володіє, те часто не погоджується витратити багато часу і зусиль на створення таких описів. Тому необхідно формувати такі описи автоматизовано – приміром, за природномовним описом задачі чи через узагальнення прикладів, наведених користувачем.

Крім того, використання занадто складних формалізмів для таких описів, хоча і забезпечує адекватну математичну модель для пошуку, але не дозволяє користувачу самостійно зрозуміти її й помітити її невідповідність задачам, що стоять перед ним. Таким чином, проблема полягає в знаходженні деякого компромісу між складністю моделі і її наочністю.

Тому необхідно проаналізувати переваги і недоліки існуючих підходів до пошуку інформації і на цій основі побудувати модель взаємодії користувача з відкритим інформаційним середовищем Web.

Слід зазначити, що в роботі розглядається *окремий випадок* такого пошуку, при якому користувач згодний затратити певні зусилля для того, щоб пошук здійснювався більш якісно. Для цього в користувача може бути одна з низки причин, і в кожному випадку інформаційно-пошукова система (ІПС) повинна враховувати специфіку ситуації.

Найбільш характерною є ситуація, коли користувачі мають стабільні довгострокові інформаційні потреби в тих областях, де самі

вони є експертами і досить добре уявляють собі як структуру ПрО пошуку, так і структуру шуканих інформаційних об'єктів. У такому пошуку потрібно забезпечувати користувача новими відомостями та відфільтровувати ті ІР, які він вже опрацював. Такий пошук використовується, зокрема, науковими співробітниками, експертами-аналітиками і технічними фахівцями.

Також вплинути на готовність користувача витратити додаткові зусилля на пошук може те, що рішення, що він збирається прийняти за результатами пошуку, є для нього дуже важливим, а якість одержуваних результатів йому складно оцінити самостійно, тому що він сам не має достатні знання в даній ПрО. Така ситуація може бути пов'язана, наприклад, з відповідальним рішенням, із ризиком втрати великих коштів чи шкоди для здоров'я. У цьому випадку може бути корисний досвід інших користувачів, що приймали аналогічні рішення, але для того, щоб результат виявився ефективний, користувачу потрібно досить докладно описати як розв'язувану проблему, так і свої власні параметри.

Крім того, користувач змушений витратити додаткові зусилля на опис своєї інформаційної потреби у тих областях, де його переваги складно формалізувати, хоча самі оцінки тим чи іншим інформаційним об'єктам користувач може дати відносно легко. Така ситуація характерна при пошуку цікавої для користувача художньої літератури, музики і предметів мистецтва або виборі осіб, з якими зручно спілкуватися, тобто у тих ситуаціях, в оцінці яких переважають суб'єктивні фактори. У такому випадку для ефективного пошуку користувачу треба оцінити набір ІО, релевантних його потребі, а пошукова система спробує знайти узагальнені властивості «гарних» ІО.

Варто враховувати, що, хоча користувач може мати інтереси в різних областях і вирішувати різні, не пов'язані одну з одною задачі, але його власні властивості при споживанні інформації залишаються досить стабільними і залежать від психофізіологічних особливостей особи, і тому ІПС може використовувати ці відомості при рішенні нових задач. Таким чином, користувач, один раз витративши час на налаштування свого персонального профілю в ІПС, надалі може зекономити набагато більше часу, потрібного для аналізу нецікавої та некорисної інформації.

## 1. Інформаційна потреба користувача та її моделювання

У найбільш загальному вигляді інформаційний пошук можна розглядати як співставлення моделі інформаційної потреби користувача з інформаційними моделями різних інформаційних ресурсів (ІР), доступних ІПС. Щоб більш точно сформулювати теорію інтелектуального інформаційного пошуку, необхідно строго визначити основні компоненти їхньої семантичної моделі, їхньої властивості і зв'язки між ними.

Основними елементами цієї моделі, які можна виділити при аналізі існуючих ІПС, є:

- *користувач*;
- *запит користувача* – явно передане користувачем повідомлення про наявність інформаційної потреби;

- *інформаційна потреба* користувача, для задоволення якої і виконується пошук інформації;

- *задача*, для рішення якої користувач має потребу в задоволенні своєї інформаційної потреби;

- *група користувачів* – деяка підмножина всієї сукупності користувачів ІПС, згрупованих за якою-небудь ознакою;

- *ІР* – документи різних типів, до яких ІПС має доступ;

- *ІО* – відомості про об'єкт з певною структурою, визначеною користувачем, що містяться в одному чи декількох ІР;

- *інформаційне середовище* – сукупність усіх доступних ІР, їхніх властивостей (включаючи їхні оцінки користувачами) і зв'язків між ними;

- *ПрО* – деяка частина ІР, що зв'язуються в одну множину відповідно до якоїсь ознаки їхнього контенту (прийнято вважати, що поняття ПрО – множини всіх предметів, властивості яких і відношення між якими розглядаються в науковій теорії – у загальному випадку не може бути формалізоване як первинне поняття);

На додаток до цих елементів пропонується ввести в модель ще один елемент – *агент користувача* (АП). АП – це інтелектуальний програмний агент, що представляє інтереси користувача у взаємодії з ІПС. Для опису поведінки такого агента використовуються інтенціональні відношення, за допомогою яких можна формалізувати цілі, наміри і бажання користувача. Використання такого формалізму, як АП, дозволяє, з одного боку, уникнути приписування людині-

користувачу штучно обмеженої і формально описаної сфери інтересів, а з іншого боку – забезпечити засоби і методи прогнозування його вчинків у рамках описаної моделі взаємодії користувача і ресурсів у відкритому інформаційному середовищі.

Крім елементів, у моделі повинні відображатися зв'язки між ІР і предметними областями; інформаційними потребами і предметними областями; предметними областями і задачами користувачів; користувачами і ПрО; між користувачами з групи.

Для того, щоб формалізувати опис елементів моделі і забезпечити його автоматизовану обробку, доцільно використовувати онтологічне представлення знань про них. Таким чином, для опису інформаційної потреби користувача треба побудувати онтологічну модель, класи якої відповідають перерахованим вище елементам, а відношення – зв'язкам між ними.

У різних ІПС зберігаються різні відомості про користувачів. У загальному випадку такі відомості можна розділити на три категорії:

- отримані від користувача явно (при реєстрації чи в процесі опитування);

- імпортовані з зовнішніх джерел інформації – соціальних мереж, домашніх сторінок, запропонованих користувачем ІР, Вікіпедії і т.д.;

- здобуті з даних, накопичених ІПС у процесі взаємодії з користувачем.

На жаль, не завжди досвіду конкретного користувача досить для того, щоб визначити корисність тих чи інших ресурсів (наприклад, при звертанні до нової ПрО чи до нової групи ресурсів). У такому випадку корисним може виявитися досвід інших користувачів, що робили раніше пошук у тій же області і подібних інформаційних потребах, що мають. При колаборативному пошуку доцільно використовувати методи, застосовувані в рекомундуючих системах.

*Колаборативний пошук* – пошук, при якому використовуються знання і досвід, отримані при спільній роботі групи користувачів з ІПС. При цьому передбачається, що користувачі поєднуються в групу при наявності схожих інтересів чи у випадку, якщо вони вирішують подібні задачі.

## 2. Семантичний інформаційний пошук

Якщо ми говоримо про семантичний пошук, то передбачається, що інформація в процесі зіставлення повинна оброблятися на

семантичному рівні, з використанням знань (про користувача, ресурси, предметну область тощо).

Навіть при найпростішому варіанті знання-орієнтованого пошуку при обробці запиту враховуються не тільки формальні відомості про запит, користувача і IP, але і більш складно структуровані знання про них. Тоді при зіставленні таких знань необхідно оцінювати ступінь подоби цих знань. Зокрема, при онтологічному підході до подання знань для кожного IP може вказуватися онтологія ПрО, що характеризує його контент, а для користувача – онтологія цікавлячої його ПрО, а при зіставленні IP і запиту необхідно буде виконати зіставлення цих двох онтологій.

Крім того, наявність додаткових знань про те, що саме шукає користувач – наприклад, про структуру шуканого інформаційного об'єкта (приміром, Web-сервісу або профілю користувача соціальної мережі) – дозволяє структурувати знайдену інформацію і надавати її користувачу в більш зручному вигляді. Наприклад, якщо користувачу потрібно знайти інформацію про якийсь ученого, і користувач вказує, які саме властивості цієї людини представляють для нього інтерес, то при семантичному пошуку ПС може явно вказати прізвище людини, його спеціальність, наявність публікацій і т.п.

*Семантичний пошук* – це метод інформаційного пошуку, у якому релевантність документа запиту визначається семантично (за близькістю змісту), а не синтаксично (приміром, за частотою використання ключових слів у документі).

Те, які саме знання використовуються, як вони представлені і як вони обробляються, залежить як від специфіки розроблювальної ПС, так і від концепції, обраної її розроблювачами, але в загальному випадку результати пошуку є функцією від запиту користувача, індексу ПС та знань, що містяться в базі знань (БЗ) ПС:

$$I_s = \{i_j, j = \overline{1, n}\} = f(z, DB_{unc}, KB_{unc}).$$

### 3. Специфіка інформаційного пошуку в Web

Якщо мова йде про семантичний пошук у Web, то слід враховувати, що при цьому в Web можуть знаходитися не тільки інформаційні об'єкти, серед яких здійснюється пошук, але і зовнішні бази знань, що використовуються для пошуку. Тому при створенні таких систем варто враховувати, що такі зовнішні БЗ можуть змінювати контент,

структуру і доступність незалежно від розроблювачів ПС. Тому результати пошуку в Web є функцією ще й від вмісту зовнішніх БЗ:

$$I_{web\_s} = \{i_j, j = \overline{1, n}\} = f(z, DB_{unc}, KB_{unc}, \{KB_{web\_k}, k = \overline{1, m}\}).$$

Варто враховувати, що сьогодні багато ПС (наприклад, Google) прагнуть накопичувати і використовувати досвід взаємодії з конкретним користувачем. Але часто інформаційні потреби користувача виявляються обмеженими в часі (наприклад, накопичивши інформацію для вибору нового телефону, користувач купує його і більше не має потреби в відомостях про телефони, а ПС продовжує пропонувати їх йому) або взагалі не пов'язаними з ним (наприклад, запит виконується чийось проханням). Крім того, досить часто певну частину своїх інформаційних потреб користувач не хоче робити відкритою інформацією – приміром, запити, пов'язані з відпочинком чи здоров'ям він не хоче змішувати з запитами по роботі. Тому більш доцільно при виконанні запитів дати можливість користувачу включити його в один зі своїх профілів або взагалі не зберігати для подальшої обробки.

### 4. Тенденції розвитку інформаційного пошуку

На основі аналізу існуючих ПС можна виділити основні тенденції розвитку пошуку [2]:

1. від формального – до семантичного;
2. від уніфікованому – до персоналізованого;
3. від індивідуальному – до колаборативного;
4. від закритого – до керованого;
5. від монотонного – до тематичного (з урахуванням динаміки і кінцівки інформаційних потреб).

Найбільше повно задовольнити інформаційні потреби користувача дозволяє інтегроване використання всіх цих можливостей, тобто персоналізований, Web-орієнтований колаборативний пошук, що базується на знаннях. Як правило, такий пошук є надбудовою над вже існуючими пошуковими системами (наприклад, Google), що дозволяє перевпорядкувати результати пошуку.

### 5. Постановка задачі

Щоб забезпечити для різних інтелектуальних застосувань ефективний доступ до ресурсів відкритій інформаційному середовищу Web, необхідно розробити інтегровану формальну модель, що забезпечує онтологічне представлення знань про



користувачів, ресурси і специфіку предметних областей (зокрема, враховуючи Semantic Web і соціальний Web). При розробці методів обробки представлених у цій моделі знань необхідно проаналізувати вже існуючі засоби онтологічного аналізу, можливості використання тезаурусів, методи індуктивного здобуття знань і алгоритми вироблення рекомендацій на основі накопиченого співтовариством користувачів досвіду об'єктивної користувача предметної області.

#### **6. Використання онтологічного аналізу для семантичного пошуку**

Різні дії, пов'язані з здобуттям і обробкою знань, являють собою вкрай складні і в цілому не автоматизовані процеси. Внаслідок цього знання є досить дорогим продуктом. Тому виникає потреба в забезпеченні повторного використання уже здобутих знань і в їх інтероперабельного представлення для цього. Основою для цього може послужити *онтологічний аналіз*, що базується на стандартах, технологіях і програмних засобах, розроблених у рамках проекту Semantic Web [3].

Те, що онтологія є адекватним засобом для опису різних ПрО, є на сьогодні загально визнаним фактом, а широкий вибір онтологій, доступних через Web, підтверджує популярність цього підходу серед різних груп розроблювачів і користувачів Web-додатків.

Крім того, ряд мов опису онтологій (наприклад, OWL DL) базується на дескриптивних логіках, що забезпечує досить точне прогнозування збіжності і часу роботи різних методів обробки знань, представлених у вигляді онтології.

Для того, щоб підвищити ефективність пошуку, доцільно як можна більш точно визначити дві властивості інформаційної потреби користувача – 1) до якої ПрО відноситься інформаційна потреба користувача і 2) який саме тип інформації необхідний для її задоволення. У першому випадку користувачу потрібно вибрати одну з існуючих онтологій ПрО і при необхідності модифікувати її у відповідності зі специфікою його проблеми. Тому що не тільки створення і модифікація онтологій, але й аналіз знань, що містяться у вже створених онтологіях, являє собою досить складну задачу, те такий підхід прийнятний тільки в тому випадку, якщо надалі така онтологія буде використовуватися для виконання одиничного запиту, а для досить великого числа пов'язаних з цією ПрО запитів.

Останнім часом багато розроблювачів ПС декларують застосування онтологій і підтримку семантичного пошуку. Наприклад, семантичний пошук Google включає три основних компоненти [4]: адреса URI; RDF; значення (онтологію). При цьому онтологія дозволяють зв'язувати один з одним різні дані, описані за допомогою RDF і адресовані через URI. Крім того, для підтримки семантичного пошуку в Google застосовується новий інструмент, що забезпечує семантичний пошук, – мережа знань Knowledge Graph, що забезпечує зв'язок між різними елементами проіндексованого контенту IP і дозволяє поєднувати інформацію з різних джерел.

Knowledge Graph дозволяє безпосередньо на сторінці з результатами пошуку одержувати інформацію про об'єкт пошуку і пов'язані з ним факти: на екран виводиться інформаційна панель, що відображає відомості про географічні об'єкти, людей, фільми і т.п. Це дозволяє користувачу отримати інформацію, не переходячи на сайт, що послужив джерелом інформації.

Knowledge Graph не просто ідентифікує ключові слова, але й аналізує семантику пошукового запиту і шукає інформацію в базі даних Google, що містить відомості про різні об'єкти – людей, місця, предмети тощо.

Google використовує правила виведення, щоб групувати інформацію на основі того, як користувачі працюють з даними в соціальних мережах (розповсюджуваний контент, коментарі і взаємодія), форумах і при пошуку в Google. Крім того, відслідковується поведінка користувачів на різних Web-сайтах і аналізуються послідовності їхніх дій. Важливим джерелом інформації служать і спеціалізовані Web-сервіси, що визначають географічне місцезнаходження користувачів за сигналом GPS від їхніх мобільних пристроїв і за IP-адресою.

Розроблювачі Google стверджують, що сполучення персоналізації пошуку з семантичним пошуком підвищують релевантність його результатів. Проте, слід зазначити, що зараз така інформація для запитів – не тільки українською або російською мовою, але і англійською – надається у відповідь тільки на найбільш короткі і прості запити і за охопленням значно менше, ніж набір статей у Вікіпедії. Тому залишається відкритим питання, наскільки буде корисним користувачам новий сервіс Google і наскільки насправді в ньому задіяний семантичний пошук.

Головна проблема, з яким зіштовхуються користувачі Knowledge Graph, – відсутність явної моделі пошуку і відомостей про структуру і заповнення онтології, що використовується при пошуку, що дозволяла б хоч якоюсь мірою прогнозувати його результати.

### 7. Онтологічна модель взаємодії користувачів і ресурсів у Web

Семантичний пошук – це надбудова над традиційним інформаційним пошуком, у якій для підвищення пертинентності пошуку використовується обробка знань як про самого користувача і його інформаційні потреби (персоніфікація пошуку), так і про інформаційні ресурси, серед яких здійснюється пошукова процедура. Тому для його здійснення необхідно:

– сформувані набір формальних моделей його компонентів;

– вказати зв'язки між ними: інформаційних моделей користувача, що цікавить його предметної області, задачі, що вирішує користувач, і інформаційних моделей доступних IP, що характеризують їхню семантику, і їхнього подальшого зіставлення;

– вказати засоби та джерела поповнення такої моделі знаннями.

Зараз для інтероперабельного представлення різних знань у Web широко застосовуються онтології, що забезпечують явне формалізоване представлення семантики представленої інформації й надають можливість логічного виведення на них. Тому представляється доцільним для підтримки семантичного пошуку розробляти саме онтологічну моделі взаємодії користувачів і IP в інформаційному просторі Web, а також методи їхнього зіставлення і поповнення.

Онтологічна модель користувача являє собою клас онтології, екземплярами якого є відомості про зареєстровані в ІПС користувачах. У ній описані наступні класи:

– онтологія  $PrO$ , що описує область, до якої відносяться інформаційні потреби користувача

$$O_{PrO} = \langle T_{PrO}, R_{PrO}, F_{PrO} \rangle, i = \overline{1, n};$$

– лексична онтологія  $PrO$ , що містить відомості про лексеми природних мов, що відповідають термінам онтології  $PrO$

$$O_{lex} = \langle T_{lex}, R_{lex}, f \rangle, i = \overline{1, n}, \text{ тобто } \forall t \in T_{PrO} -$$

така онтологія використовується для встановлення зв'язків між елементами природномовних документів та термінами онтології  $PrO$ ;

– тезаурус задачі – множина пара, першим елементом яких є терміни онтології, сукупність яких характеризує ту конкретну задачу з  $PrO$ , що у даний момент вирішує користувач, а другим – вага (позитивна чи негативна) цього терміна для даної задачі

$$Th_{ij} = \langle \{th_{k_{ij}} \in T_{PrO}, v_{k_{ij}}\}, k = \overline{1, s_{ij}}, j = \overline{1, m_i} \rangle;$$

– запит – множина ключових слів, що характеризують одну з інформаційних потреб користувача, пов'язану з конкретною задачею за допомогою тезауруса

$$z = \langle \{k_q\}, Th_{ij} \rangle, q = \overline{1, u};$$

– тема – множина запитів, пов'язаних з однією інформаційною потребою

$$thema = \langle id_{thema}, \{z_q\} \rangle, q = \overline{1, u}, \text{ що дозволяє поєднувати семантично пов'язані запити різних користувачів, які базуються на різних онтологіях і тезаурусах;}$$

– результат запиту – множина пара, першим елементом яких є посилання на IP, а другим – оцінки цих IP користувачем

$$rez = f(z, u) = \langle \{id_{ir}, rating_{ir}\} \rangle;$$

– користувач – клас, що має більш складну структуру і має наступні атрибути, які можна розділити на кілька груп:

1. реєстраційна інформація:
  - ідентифікатор користувача;
  - пароль для доступу до ІПС;
2. досвід взаємодії ІПС із користувачем:
  - список онтологій, що користувач застосовував для опису своїх інформаційних інтересів;
  - список тезаурусів, що користувач застосовував у пошукових запитах;
  - список раніше виконаних запитів;
  - список результатів виконаних запитів з оцінками користувача для знайдених результатів;

3. відомості, імпортовані з зовнішніх джерел (необов'язкові відомості, можуть бути відсутніми): ідентифікатори користувача в соціальних мережах, що дозволяють динамічно оновлювати відомості про нього; рейтинги користувача в соціальних мережах; адреса користувача у Вікіпедії й інших виках-ресурсах; адреса сайту користувача; сфера компетенцій користувача (ключові слова,

імпортовані із соціальних мереж); посилання на публікації користувача;

4. власні характеристики користувача: сфера компетенцій користувача (список ключових слів, що вводяться користувачем безпосередньо);

5. формальні дані про користувача (необов'язкові відомості, що дозволяють ІПС формувати групи користувачів зі схожими інформаційними потребами): місце проживання; вік; професія, освіта і т.д.

– група користувачів – клас, властивостями якого є ідентифікатор групи і список користувачів, за певних причин об'єднаних в одну групу (групи можуть формуватися явно шляхом вибору користувача чи автоматично на основі відповідності яким-

небудь умовам, наприклад, групи користувачів з подібними формальними даними чи виконуючих схожі запити)

$$gr = \langle id_{gr}, \{u_i\}, i = \overline{1, n} \rangle;$$

– інформаційний ресурс – відомості про знайдені раніше ресурси і їхні оцінки  $\langle U_{url}, \{z_i, m_i, q_i\}, i = \overline{1, n} \rangle$ , що включають ідентифікатор ресурсу, запити, по яких він був виявлений, оцінку користувача, якому він був наданий, і його рівень читабельності для цього користувача.

Така онтологія може бути представлена мовою OWL та використовуватися у різних застосуваннях, пов'язаних з використанням семантичного пошуку (рис.).

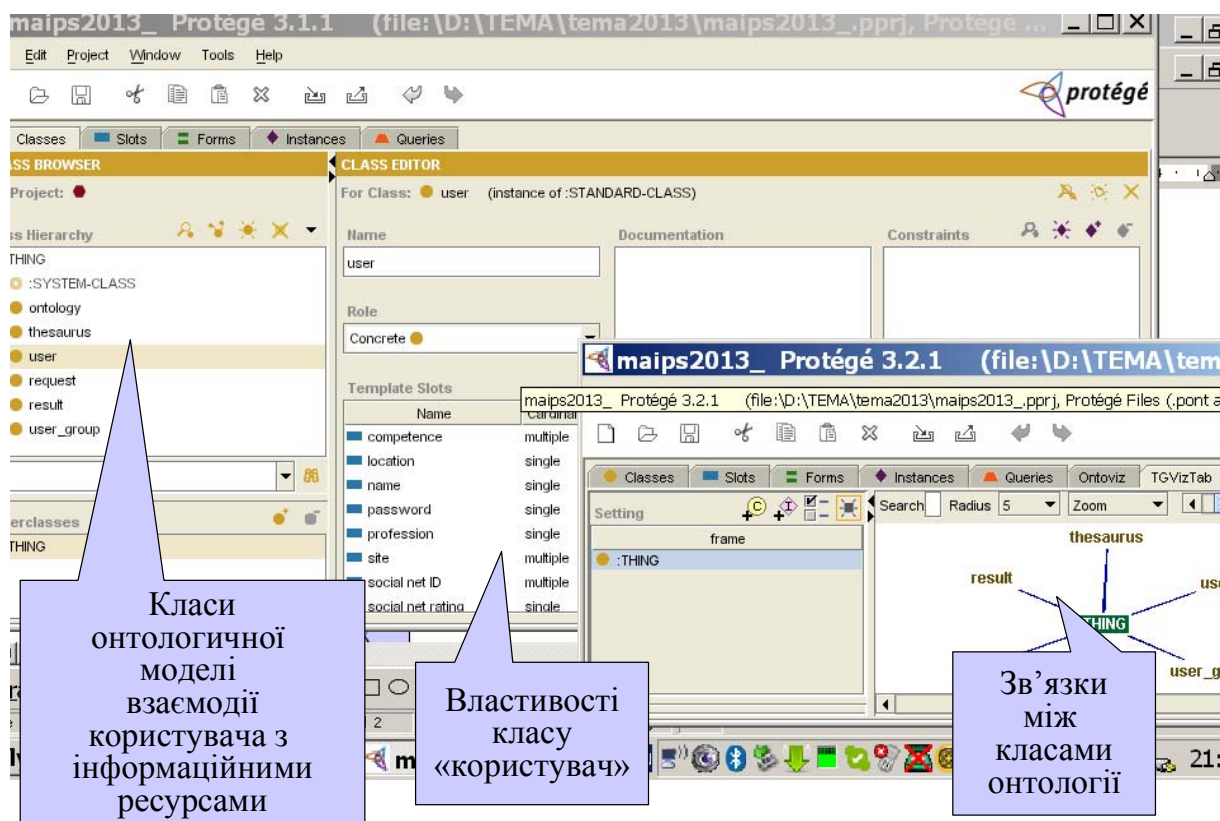


Рис. Онтологічна модель взаємодії користувачів і ІР

## 8. Використання в семантичному пошуку методів рекомендуєчих систем

Як відомо, *рекомендуючі системи* (РС) відрізняються від ІПС тем, що користувачу не треба явно формулювати пошуковий запит – система сама, на підставі наявних відомостей про користувача, пропонує йому елементи, що рекомендуються (РЕ) [5].

Персоналізовані рекомендації, вироблювані такими системами, – це впорядковані списки РЕ, тобто робота РС зводиться саме до ранжирування доступних РЕ. Чим вище і точніше інформованість РС про потреби користувача, тим більше ефективні результати її роботи. Але навіть у тому випадку, якщо користувач явно формулює свої інформаційні потреби, за

допомогою методів, що використовуються у РС, можна значно підвищити пертинентність пошуку, персоніфікуючи запит і використовуючи колаборативні знання про різні співтовариства користувачів, згрупованих на основі подібних інформаційних потреб [6].

Формальне створення рекомендацій у РС може бути представлена в такий спосіб: нехай  $C$  – множина користувачів РС,  $S$  – множина пропонованих РЕ (товарів, книг, фільмів, сервісів тощо).  $U$  – функція корисності, що описує інтерес користувача  $c \in C$  до РЕ  $s \in S$ , тобто  $U: C \times S \rightarrow R$ , де  $R$  – кількісна оцінка. Ціль РС – для кожного споживача  $c \in C$  вибрати такий РЕ  $s' \in S$ , що  $U(c, s') = \max_{s \in S} u(c, s)$ . Яким саме образом

визначається функція корисності, залежить від типу РС і від специфіки РЕ.

У РС при виборі РЕ використовують:

- *персональний* підхід – аналіз профілю конкретного користувача, його виявлених раніше переваг і явно виражених умов;

- *колаборативний* підхід – аналіз переваг інших користувачів, що якихось причин можуть бути поширені і на того користувача, для якого робиться вибір;

- *контент-орієнтований* підхід, при якому аналізуються контент тих РЕ, що пропонуються користувачу;

- *довірчий* підхід – аналізується якість пропонованих користувачу РЕ й ступінь довіри до них.

Слід зазначити, що в більшості реальних РС усі ці підходи реалізуються інтегровано, але їм надається різна увага.

У РС подоба між двома користувачами ґрунтується на тому, що вони дали раніше подібні оцінки тим самим РЕ. Для ефективної роботи РС треба передбачати подальші оцінки для неоцінених раніше РЕ, виходячи з невеликої кількості прикладів. Для подолання проблеми розрідженості оцінок потрібно при пошуку схожих користувачів використовувати також відомості з їхніх профілів і виявляти користувачів зі схожими профілями, наприклад, що відносяться до одного демографічного сегмента або мають подібні риси характеру.

Аналіз основних напрямків розвитку сучасних РС [7] пов'язує їх з використанням онтологій для представлення знань як про користувачів, так і про РЕ.

При персональному підході РС необхідно накопичити досить багато відомостей про користувача, щоб надалі їх узагальнювати й аналізувати. Рекомендація щодо нових РЕ для користувачів може формуватися на основі його порівняння з подібними РЕ (фільтрація на основі контенту), відгуків про РЕ в співтоваристві користувачів (колаборативна фільтрація), семантичних відношень між РЕ (евристичні рекомендації) чи поєднання цих підходів.

Колаборативна фільтрація використовує рейтинги, наданих співтовариством користувачів, щоб рекомендувати РЕ конкретному користувачу. Існують два взаємодоповнюючих підходи до колаборативної фільтрації: на основі користувача або на основі РЕ. При колаборативній фільтрації на основі користувача знаходять групи подібних користувачів, а потім конкретному користувачу рекомендують ті РЕ, що сподобалися іншим користувачам з тієї ж групи. При фільтрації на основі РЕ у групи поєднують РЕ з подібними оцінками.

На основі вищенаведеного аналізу можна запропонувати наступні підходи до роботи РС із використанням онтологій: формування моделі користувача; формування моделі РЕ; створення онтології РЕ; накопичення відомостей про екземпляри РЕ та екземпляри користувачів; накопичення оцінок РЕ користувачами; аналіз екземплярів РЕ; класифікація (чи кластеризація) користувачів на групи з подібними інтересами; формування набору стратегій, що користувач може явно вибирати для отримання рекомендації; побудова методу, що дозволяє уточнити клас необхідного користувачу РЕ.

Розглянемо також, оцінки яких саме груп користувачів доцільно застосовувати для колаборативної фільтрації. Найпростіший випадок групи – це група, що складається усього з одного користувача, для якого і здійснюється пошук рекомендацій. Можна сказати, що при цьому колаборативна фільтрація зводиться до персонального. Протилежний випадок – коли для користувача значимі оцінки РЕ всім співтовариством у цілому. Це може мати місце для тих ПрО, до яких користувач звертається вперше і ще не має власної думки не тільки про саму область, але і про критерії перебування в ній експертів. До проміжних випадків відносяться аналіз оцінок експертів у ПрО (доцільно надати



користувачу можливість явно задавати прийнятний рівень їхньої кваліфікації).

У цілому варто оцінювати обрану стратегію рекомендації по трьох напрямках – облік думки самого користувача, облік думки співтовариства, аналіз контенту самого РЕ. У такому тривимірному просторі можна розмістити більшість типових об'єктів рекомендації.

Вибір користувачем значення за кожним з трьох параметрів для стратегії рекомендації для цікавлячих його РЕ і наявність онтології (чи хоча б таксономії РЕ) дозволяє досить точно профілювати інтереси самого користувача, оцінити його власну компетентність для оцінювання РЕ (і, відповідно, значимість його думки для інших користувачів) і виявляти групи користувачів з подібними інтересами. Слід зазначити, що більш об'єктивним є визнання своєї некомпетентності в оцінюванні РЕ, чим декларування своєї високої кваліфікації.

#### 9. Тезауруси як засіб представлення онтологічних знань

Зазвичай онтологія ПрО має досить складну структуру, але при задоволенні інформаційної потреби користувача для рішення конкретної задачі ППС використовує лише частину знань, що містяться в такій онтології. Тому доцільно використовувати для моделювання знань користувача про цікавлячий його ПрО *тезаурус* – окремий випадок онтології, якому можна розглядати як проекцію онтології на задачу [ 8, 9].

Тезаурус – це  $Ts = \langle T, R \rangle$ , де  $T$  – множина термінів, а  $R$  – множина відношення між цими термінами. Множині  $T$  і  $R$  скінчені. Множина термінів тезауруса  $T$  відповідає множині концептів  $X$  онтології  $O$ . Тезауруси дозволяють моделювати знання як про користувачів, так і про ті ресурси, що вони шукають [10].

Щоб формалізувати область своїх інтересів – ПрО пошуку – користувачу треба створити тезаурус, що моделює цікавлячу його ПрО, у якій містяться основні терміни ПрО і зв'язки між ними. Тезаурус можна створити вручну чи автоматизовано. Основою для автоматичного створення тезауруса може послужити обробка набору ІР, релевантних цій ПрО, чи раніше створена онтологія ПрО, з який користувач відбирає тільки необхідні йому терміни. Різні ці підходи можуть комбінуватися один з одним.

Для створення тезаурусів ІР і РЕ пропонується використовувати спрощений алгоритм побудови тезауруса: за повним переліком слів, що використовуються в ІР, будується словник термінів, з якого відкидаються стоп-слова, що містяться в спеціально розробленому користувачем списку. Цей алгоритм застосовується тільки для тих ІР, що не супроводжуються метаописами. В іншому випадку з метаописів (у форматі RDF чи OWL) здобуваються терміни тезауруса і зв'язки між ними, що доповнюють побудований за контентом ІР словник. Аналогічно будуються тезауруси РЕ – обробляються їхній метаопис, контент, відгуки про них інших користувачів.

Користувач вводить запит, приблизно ідентифікуючи свою інформаційну потребу за допомогою ключових слів чи вибираючи клас цікавлячого його РЕ (можливо, з набором умов і обмежень), наприклад, РЕ класу «художня література/фантастика/фентезі», видана після 2005 року. У відповідь РС формує (шляхом звертання до зовнішньої ППС чи в результаті пошуку у власній індексній БД) набір РЕ, доступних системі і відповідаючих цьому приблизному запиту –  $n$  посилань на РЕ і їхніх коротких описів  $I = \{Ref_j, D_j\}$ ,  $j = \overline{1, n}$ . Тут  $Ref_j$  – посилання на відповідний РЕ  $IP_j$  (чи його опис), а  $d_j$  – інформація щодо РЕ, доступна РС.

Якщо множина  $I$  не порожня, причому РС знайшла у відповідь на запит більш ніж один РЕ ( $n \geq 1$ ), то потрібно встановити порядок, у якому пропонувати користувачу відомості про знайдений РЕ. Тоді для всіх РЕ з цієї множини  $I = \{Ref_j, D_j\}$ ,  $j = \overline{1, n}$  формуються їхній спрощені тезауруси  $Ts(IP_j) = \langle T_j, \emptyset \rangle$ ,  $j = \overline{1, n}$  і відповідні їм словники термінів  $T_j = \{t_{j_w}\}$ ,  $j = \overline{1, n}$ ,  $w = \overline{1, q_j}$ .  $t_{j_w}$  – це слова, що використовуються в інформації про  $j$ -м РЕ, знайденому РС, тобто в.  $D_j$ ,  $j = \overline{1, n}$   $q_j$ ,  $j = \overline{1, n}$  – це кількість різних слів, використовуваних в описі  $D_j$ ,  $j = \overline{1, n}$ . Якщо слова в описі повторюються, то в словнику термінів вони фіксуються тільки один раз.

Потім користувач формує тезаурус цікавлячої його ПрО (чи вказує на раніше сформований тезаурус)  $Ts_{ПрО}$  і відповідний

йому словник термінів цієї ПрО  $T_{ПрО} = \{t_m\}$ ,  $m = \overline{1, q}$ .  $T_{ПрО}$  – це множина, що складається з  $m$  термінів, що відносяться до цікавлячої користувача ПрО. Ця множина будується аналогічно словнику термінів РЕ і звичайно формується як об'єднання словників термінів, що містяться в документах, що користувач знайшов раніше і порахував релевантними цікавлячої його ПрО (як у їхньому контенті, так і в метаописах).

Виконується порівняння  $T_{ПрО}$  і  $T_j$ ,  $j = \overline{1, n}$ , вираховується коефіцієнт їхньої

близькості  $K_j = \sum_{m=1}^q \sum_{w=1}^{w_j} f(t_{j_w}, t_m)$ ,  $m = \overline{1, q}$ ,  $w = \overline{1, w_j}$ ,

де  $f(t_1, t_2) = \begin{cases} 0, & \text{если } t_1 \neq t_2, \\ 1, & \text{если } t_1 = t_2. \end{cases}$  являє

собою кількість термінів, що зустрілися як у тезаурусі РЕ, так і в тезаурусі ПрО.

Знайдені ІР впорядковуються в залежності від значень  $K_j$ , користувачу пропонуються в першу чергу ті ІР, що мають найбільш високий коефіцієнт близькості до ПрО.

Представляється доцільним використовувати онтологію ПрО і виділяти групи слів тезауруса, що відповідають одному терміну. Для цього користувач повинний пов'язати РЕ словника термінів тезауруса ПрО з одним з термінів онтології ПрО, тобто  $\forall t_m \in T_{ПрО}$ ,  $m = \overline{1, q}$  задати функцію  $g(t_m) \in X$ , що використовується для обчислення коефіцієнта близькості  $K^O$ :

$K^O_j = \sum_{m,l} f(t_{j_w}, t_m)$ ,  $m = \overline{1, q}$ ,  $w = \overline{1, w_j}$ , де  $f(t_1, t_2) = \begin{cases} 0, & \text{если } g(t_1) \neq g(t_2), \\ 1, & \text{если } g(t_1) = g(t_2). \end{cases}$

$K^O$  являє собою кількість термінів, що зустрілися як у тезаурусі РЕ, так і в тезаурусі ПрО і при цьому посилаються на той самий термін онтології ПрО. Це дозволяє використовувати меншу кількість документів для побудови тезауруса ПрО, але вимагає більше часу для обчислень.

При створенні тезауруса ПрО, що цікавить користувача РС, необхідно явно вказати основні поняття ПрО і зв'язки між ними. На жаль, більшості користувачів досить складно це зробити (навіть маючи відповідні

знання і застосовуючи їх у своїй діяльності). На першому етапі формування тезауруса користувач може вибрати одне з наступних рішень: 1) самостійно побудувати за допомогою одного з редакторів онтологій онтологічний опис цікавлячої його ПрО; 2) знайти (наприклад, у Web) онтологію мовою OWL, що описує ПрО, близьку до області його інформаційних інтересів; 3) сформулювати множину понять ПрО, що містить найбільш характерні слова і словосполучення, що зустрічаються в цікавлячих його ІР.

#### 10. Використання індуктивного виведення для поповнення онтологій

При побудові онтологій і тезаурусів ПрО, а також при колаборативному підході до пошуку важливо визначити, які зв'язки між термінами ПрО є істотними для опису інформаційної потреби користувача. Користувачу досить складно самостійно знайти усі важливі закономірності і відкинути несуттєві.

Для виявлення таких зв'язків можна скористатися методами індуктивного і традуктивного здобуття знань з даних. Існують незалежні підходи до реалізації подібних методів: ID3, ACLS, CART і т.д. Найбільш цікавим, у зв'язку зі специфікою проведеної роботи, виявився алгоритм ID3 [11], що спеціально розроблений для здобуття корисної інформації з великих обсягів слабо структурованих даних.

Незростаючий алгоритм ID3 призначений для узагальнення досвіду експериментів, параметри і результати яких описані через якісні оцінки (лінгвістичні перемінні). Він забезпечує побудову бінарного дерева рішень, а цього недостатньо зручно для представлення закономірностей багатьох ПрО. Тому для поповнення онтологій пропонується використовувати ID3m [12] – модифікацію ID3 для довільної (скінченої) кількості рішень. Він також належить до незростаючих алгоритмів. У даному випадку, прикладами навчальної вибірки є РЕ, доступні РС, а параметрами, по яких вони описуються, є їхні властивості, описані в метаданих і в онтології РЕ, значення параметрів оцінюваних їх користувачів, а також терміни тезауруса користувача.

#### 11. Семантичний пошук у МАПС

МАПС – це мультиагентная ІПС, що забезпечує семантичний пошук різних ІО на основі знань, що містяться в онтологіях. Вона призначена для пошуку інформації в описаних користувачем відносно вузьких ПрО,

пов'язаних з їх професійними чи науковими інтересами, і рекомендує користувачу ті результати пошуку, що відносяться до цікавлячої його ПрО і відповідають його персональним інформаційним потребам. МАПС можна розглядати як рекомендуючу систему, орієнтовану на формування колаборативних рекомендацій щодо природно-мовних і мультимедійних інформаційних ресурсів, доступних через Web.

МАПС базується на технологіях Semantic Web, зокрема, використовує мова представлення онтологій OWL і засобу його обробки. Для представлення знань об цікавлячої користувача ПрО використовуються онтології – як створені самими користувачами, так і знайдені в репозиторіях онтологій, доступних через Web [14], – і тезауруси ПрО. При цьому тезаурус будується користувачем по відповідній онтології самостійно, а онтологія обирається з набору запропонованих на сайті.

Користувач МАПС може звертатися до онтологій, створених іншими користувачами – переглядати їх, задавати по них контекст пошуку, копіювати з них потрібні фрагменти, але не має права змінювати їх. ППС може забезпечити пошук онтологій, що містять уведені користувачем терміни, а також пошук онтологій, схожих на обрану користувачем онтологію. Це дозволяє створювати групи користувачів із загальними інформаційними інтересами і запобігти дублюванню у виконанні однакових багаторазових запитів різних користувачів.

Онтологічна модель, що описує семантику взаємодії користувачів і ресурсів МАПС в інформаційному просторі Web, забезпечує знання для виконання наступних дій, пов'язаних з пошуком інформації і засновані на розглянутих вище методах.

На відміну від більшості існуючих рекомендуючих систем, МАПС дозволяє користувачу явно, безпосередньо і динамічно керувати засобами створення рекомендацій. Користувач може враховувати оцінки:

- Усього співтовариства користувачів МАПС;
- Підмножин користувачів, запити яких базуються на тих же онтологіях;
- Тих користувачів, що використовують найбільш подібні тезауруси і ключові слова для запитів;
- Запитів з обраної користувачем теми, у яку можуть входити як тільки власні запити користувача, так і запити різних

користувачів з різними онтологіями і тезаурусами;

- Явно зазначеної підмножини користувачів МАПС;

- Підмножини користувачів МАПС, побудованої по введеним користувачем формальними умовами (наприклад, за місцем проживання чи за віком);

- Підмножини запитів самого користувача, що відповідають визначеним умовам (наприклад, побудовані в зазначений інтервал часу і по використанню ключового слова).

Такі умови користувач може надавати як окремо для кожного запиту, так і для визначеної групи своїх запитів. Методи побудови рекомендацій впливають на упорядкування результатів запиту та дозволяють рекомендувати користувачу ті IP, що не зустрічаються в результатах його власних запитів, але вважаються МАПС цікавими для нього

МАПС функціонує у відкритому інформаційному середовищі Web і тому має потребу в динамічному відновленні своїх знань про це середовище. Шляхами для цього є проактивний пошук нових відомостей про своїх користувачів у Web (наприклад, пошук їхніх нових чи публікацій тих публікацій, що містять посилання на них, експорт нових відомостей з соціальних мереж) і автоматизоване відновлення онтологій ПрО (наприклад, шляхом обробки семантичні вікі-ресурсів, експорту зовнішніх онтологій).

## **12. Використання інтенціональних відношень для моделювання поведінки програмного агента користувача ППС**

Оцінити ефективність ППС і її здатність задовольняти різні інформаційні потреби користувача досить складно, тому що оцінки користувачів досить суб'єктивні, а самим користувачам звичайно складно строго вказати параметри, на підставі яких одні ІО виявилися краще інших при досить схожих формальних параметрах. Тому виникає необхідність у формальному математичному апараті, що дозволяє описувати інформаційну потребу користувача і прогнозувати поведінку інтелектуального програмного агента, що представляє перед ППС інтереси цього користувача.

*Програмні агенти (ПА)* – сучасна парадигма програмування, що дозволяє перейти на новий, більш інтелектуальний рівень взаємодії користувача з програмним і

апаратним забезпеченням [15]. Ця парадигма забезпечує підвищення ефективності роботи і дозволяє користувачам доручити ІС виконання досить складних задач – наприклад, пошук знань.

Звичайно людське поведінка прогнозується і порозумівається через такі атрибути відношення, як переконання, бажання, надії, побоювання і т.п., що називаються інтенціональними поняттями. Філософ Д.Деннет увів термін інтенціональних систем для опису сутностей, поведінка яких прогнозується шляхом приписування їм атрибутів переконання, бажання і раціональності [16], а Маккарті розглянув область застосовності таких систем [17]. Чим менше відомо про систему і її структуру, тим більше корисні інтенціональні пояснення її поведінки. Крім того, для досить складної системи (навіть за наявності повної інформації про неї) інтенціональне пояснення її поведінки часто більш практично, чим механістичне.

Інтелектуальний агент як система послідовно описується через інтенціональні стани [18] – інформаційними відношеннями (переконання, знання) і перед-відношеннями (бажання, намір, зобов'язання, здійснення, вибір, ціль і т.д.). Інформаційні відношення відносяться до інформації, яку агент має про світ, у якому він існує (у даному випадку – про користувача, про цікаву для користувача ПрО і про потрібні йому ІО, а також про інформаційний простір Web, у якому здійснюється процес пошуку інформації), тоді як перед-відношення – це те, що деяким чином керує діями агента (відомості про задачу, яку вирішує користувач, і про його переконання щодо того, що йому слід для її рішення).

*Переконання* (beliefs) агента виражають його думку про поточний стан світу і про правдоподібність способу дії, що приводить до певного ефекту.

*Бажання* (desire) описує переваги агента щодо майбутніх станів світу чи образ дій. Важливою рисою бажання є те, що агент може мати несумісні і недосяжні бажання.

*Цілі* (goals) – це несуперечлива підмножина бажань агента.

*Наміри* (intension) – несуперечлива підмножина цілей, досяжних обмеженням в ресурсах агентом, і спосіб їхнього досягнення.

Перед- і інформаційні відношення тісно пов'язані, оскільки агенти можуть робити раціональний вибір, формувати наміри тощо на основі інформації, яку вони мають щодо

світу. У відкритому середовищі Web така інформація завжди є неповною, може бути суперечною та динамічно змінюватися.

Багато інтенціональних понять – такі, як переконання і бажання – є референційно непрозорими (наприклад, істинність твердження “А вважає Х” залежить не тільки від значення Х, але і від значення А) і тому класична логіка в її стандартній формі непридатна для їхнього опису.

Питання, яка саме комбінація інформаційних і перед-відношення найкраще підходить для характеристики раціональних агентів, є предметом обговорення і в значній мірі залежить від конкретних задач. В даний час найбільш популярні різні комбінації переконань, бажань, цілей і намірів. Наприклад, Кохен і Левескье [19] використовують два базових відношення: переконання і цілі (інші відношення визначаються через них), а Рао і Джорджеф [20] – три: переконання, бажання і наміри (BDI-архітектура). Однак, усі ці формалізми фокусуються тільки на одному аспекті агентів – їх інформаційних відношеннях.

Для опису агента користувача ПС пропонується використовувати формальну модель, що включає чотири інтенціональних відношення – *знання, переконання, наміри і цілі*, а також зв'язки між цими відношеннями [21].

Розглянемо твердження типу “А (R) ρ”, де А – суб'єкт, ρ – стверджувальне твердження, а R – відношення між А та твердженням ρ. Якщо в якості відношення R розглянути переконання, то твердження, що “А вважає ρ” істинно, означає, що ρ визнається істинним для А незалежно від дійсної істинності ρ. Передбачається, що існує хтось, що стверджує істинність усього твердження “А вважає ρ”, тобто зовнішній суб'єкт К. Отже, існують дві точки зору: А – внутрішнього суб'єкта і К – зовнішнього суб'єкта, тобто в семантиці можливих світів [21] необхідно розглядати дві множини можливих світів.

Як загальнозначущий критерій відмінності знання від переконання береться фіксована точка зору, представлена базовою системою можливих світів W. У такому випадку твердження “А вважає ρ” буде істинно у світі  $\omega \in W$  тоді і тільки тоді, коли ρ істинно в співвіднесеному зі світом  $\omega \in W$  світі  $u \in U$ . Твердження “А знає ρ” істинно в світі  $\omega$  тоді і тільки тоді, коли ρ істинно у світі



$\omega$  і в співвіднесеному з  $\omega$  світі  $u \in U$ . Останнє означає, що суб'єкт стверджує істинність  $p$ , і  $p$  дійсно істинно.

$\omega$  – класичний можливий світ, тобто в ньому виконуються закони класичної логіки і він цілком визначений: будь-яке правильно побудоване висловлення мови в ньому або істинно, або ні. Але суб'єкт  $A$  не завжди має у своєму розпорядженні повну і неспотворену інформацію про світ, тому його уявлення про світ складаються зі здогадів і припущень. Тому для опису світів з  $U$  істинностних значень “істинно” і “ложно” явно недостатньо (наприклад, існують твердження, оцінка істинності яких невідома, але суб'єкт упевнений у можливості або неможливості такої оцінки).

Вибір даної модальної системи дозволяє моделювати поведінку розумного цілеспрямованого суб'єкта, деякі припущення якого можуть виявитися помилковими, але який має здатність розрізняти, що він припускає і чого не припускає [22]. Зокрема, у МАПС у таким суб'єктом є користувач і інтелектуальний програмний агент, який представляє його інтереси в процесі семантичного пошуку. Використання модальностей дозволяє змоделювати його поведінку при оцінці різних ІО.

### Висновки

Розробка онтологічної моделі інформаційної потреби користувача при семантичному пошуку в Web забезпечує спільне використання методів створення рекомендацій, доступу до зовнішніх джерел інформації, індуктивного здобуття знань і технологій Semantic Web, що дозволяє більш ефективно забезпечити користувача необхідними відомостями, а явно обрані методи рекомендування й онтологічні описи ПрО забезпечують користувачу прогнозувати результати пошуку. Застосування парадигми інтелектуальних програмних агентів при опису елементів такої пошукової системи дозволяють користувачу розуміти поведінку такої системи.

### Список використаних джерел

1. Рогушина Ю.В. Управление знаниями на основе онтологий в дистанционном обучении / Ю.В. Рогушина – LAP LAMBERT Academic Publishing, 2013. – 92 с.
2. Рогушина Ю.В. Знание-ориентированные средства поддержки

семантического поиска в Web / Ю.В. Рогушина // Материалы IV международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» OSTIS-2014, Минск: БГУИР, 2014. – С. 339 – 352.

3. Хорошевский В.Ф. Онтологические модели и Semantic Web: откуда и куда мы идем? / В.Ф. Хорошевский // Сб. трудов симпозиума «Онтологическое моделирование», 19–20 мая 2008. – М., ИПИ РАН, 2008.

4. Amerland D. Google Semantic Search: Search Engine Optimization (SEO) Techniques That Gets Your Company More Traffic, Increases Brand Impact and Amplifies Your Online Presence / D. Amerland. – New-York: Que Publishing, 2013. – 230 p.

5. Ricci F. Recommender Systems Handbook / F. Ricci, L. Rokach, B. Shapira, P. Kantor. – New-York: Springer, 2011. – 842 p.

6. Рогушина Ю.В. Менеджмент знаний в рекомендующих системах на основе онтологий / Ю.В. Рогушина // Сборник трудов XIII международной конф. им. Т.А.Таран "Интеллектуальный анализ информации ИАИ-2013". – К.: Просвіта, 2013. – С. 14 – 20.

7. Middleton S. Ontology-Based Recommender Systems / S. Middleton, D. De Roure, N. Shadbolt // Handbook on Ontologies, Edt. by S.Staab, R.Studer. – New-York: Springer, 2009. – P. 779 – 796.

8. Браславский П.И. Тезаурус как средство описания систем знаний / П.И. Браславский, С.Л. Гольдштейн, Т.Я. Ткаченко // Информационные процессы и системы. – № 11 (2). – 1997. – С. 16 – 22.

9. Лукашевич Н.В. Тезаурусы в задачах информационного поиска / Н.В. Лукашевич. – М.: Издательство Московского университета, 2011. – 512 с.

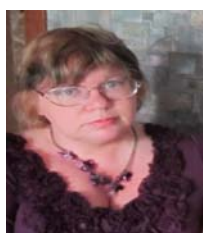
10. Гладун А.Я. Онтологии и мультилингвистические тезаурусы как основа семантического поиска информационных ресурсов Интернет / А.Я. Гладун, Ю.В. Рогушина // The Proc. of XII-th Intern. Conf. KDS'2006, Varna, Bulgaria. – P. 115 – 121.

11. Quinlan J.R. Discovery rules from large collections of examples: a case study / J.R. Quinlan // Expert Systems in the Microelectronic Age. – UK: Edinburg, 1979. – P. 87 – 102.

12. Rogushina J., Gladun A. Ontology-based competency analyses in new research domains / Rogushina J., Gladun A. // Journal of Computing and Information Technology. V.20, N. 4, 2012. – P.277-293.

13. Рогушина Ю.В. Использование критериев оценки удобочитаемости текста для поиска информации, соответствующей реальным потребностям пользователя // Проблемы программирования, № 3, 2007. – С. 76-87.
14. Рогушина Ю.В., Гладун А.Я. Репозитории онтологий как средство повторного использования знаний для распознавания информационных объектов // Онтология проектирования, № 1 (7), 2013. – С.35-50.
15. Плєскач В.Л ., Рогушина Ю.В. Агентні технології. – Монографія. К.: Київ. нац. торг.-екон. ун-т, 2005. – 338 с.
16. Dennett D.C. The Intensional Stance. The MIT Press: Cambridge, MA, 1987. – 282 p.
17. McCarthy J.. Ascribing mental qualities to machines. Technical report, Stanford University AI Lab., Stanford, CA 94305, 1978. – 84 p.
18. Konolige K. A Deduction Model of Belief. Pitman Publishing: London and Morgan Kaufmann: San Mateo, CA, 1986. – 226 p.
19. Cohen P.R., Levesque H.J. Intention is choice with commitment // Artificial Intelligence, 42, 1990. – P.213-261
20. Rao A.S., Georgeff M.P. Modeling rational agents within a BDI-architecture // In R. Pikes and E. Sandewall, eds.. Proc. of Knowledge Representation and Reasoning (KR&R-91), Morgan Kaufmann Publishers: San Mateo, CA, April 1991. – P. 473-484.
21. Hintikka J. Knowledge and Belief. Cornell University Press: Ithaca, NY, 1962. – 124 p.
22. Рогушина Ю.В. Программные агенты: определения, таксономии, модели // Управляющие системы и машины, 2001, N 5. - С.39-45

#### Відомості про автора:



**Рогушина Юлія Віталіївна** – к. ф.-м. наук, старший науковий співробітник Інституту програмних систем Національної Академії Наук України. Наукові інтереси: онтологічний аналіз на базі технологій Semantic Web, інтелектуальний інформаційний пошук, методи індуктивного вилучення знання, дослідження інтелектуальних інформаційних систем і поведінку програмних агентів.

**E-mail:** [ladamandraka2010@gmail.com](mailto:ladamandraka2010@gmail.com)