

МОДЕЛІ ТА МЕТОДИ ВИКОРИСТАННЯ ОНТОЛОГІЙ У СЕМАНТИЧНОМУ ПОШУКУ У WEB

Ю.В. Рогушина

В роботі проаналізовано основні напрями семантизації інформаційних ресурсів Web та теоретичні засади, що можуть використовуватися для цього. Розглянуто проблеми, пов'язані з використанням та аналізом інформації з відкритого середовища на рівні знань. Обґрунтовано потреби в розробці методів та засобів семантичного пошуку в Web та доцільність використання онтологій для формалізації знань щодо інформаційних потреб користувачів та їх поточних задач. Розглянуто сучасні напрями розвитку семантичного пошуку та його персоналізації. Обґрунтовано доцільність використання семантично розмічених Wiki-ресурсів як джерела знань для побудови онтологій окремих задач, що є підмножинами більших за обсягом та складністю онтологій предметних областей. Розроблено формальну модель предметної області, розкладається її співставлення з формальним моделлю онтології предметної області. Запропоновано методи побудови онтології задачі за множиною семантично розмічених Wiki-сторінок. Проаналізовано перспективи застосування систем семантичного пошуку, здатних використовувати такі онтології. Описано програму реалізацію запропонованої моделі підходу – систему семантичного пошуку МАІПС, що забезпечує персоналізоване задоволення інформаційних потреб користувача. Така персоналізація базується на застосуванні знань з онтологій задачі та побудованих за цією онтологією тезаурусів.

Ключові слова: семантичний пошук, онтологія задачі, семантичні Wiki-ресурси, персоналізація пошуку.

В работе проанализированы основные направления семантизации информационных ресурсов Web и теоретические основы, которые могут использоваться для этого. Рассмотрены проблемы, связанные с использованием и анализом информации из открытой среды на уровне знаний. Обоснована потребность в разработке методов и средств семантического поиска в Web и целесообразность использования онтологий для формализации знаний об информационных потребностях пользователей и их текущих задачах. Рассмотрены современные направления развития семантического поиска и его персонализации. Обоснована целесообразность использования размеченных Wiki-ресурсов как источника знаний для построения онтологий отдельных задач, которые являются подмножествами больших по объему и сложности онтологий предметных областей. Разработана формальная модель онтологии задачи, рассматривается ее сопоставление с формальной моделью онтологии предметной области. Предложены методы построения онтологии задачи по множеству семантически размеченных Wiki-страниц. Проанализированы перспективы применения систем семантического поиска, способных использовать такие онтологии. Описана программа реализации предложенного в работе подхода – система семантического поиска МАИПС, которая обеспечивает персонализированное удовлетворение информационных потребностей пользователей. Такая персонализация базируется на применении знаний, содержащихся в онтологиях задачи и построенных по этим онтологиям тезаурусам.

Ключевые слова: семантический поиск, онтология задачи, семантические Wiki-ресурсы, персонализация поиска.

In this paper we show the main directions of semantization of the Web information resources and their theoretical bases. We consider the semantization problems associated with the open environment information at the knowledge level. This analysis substantiates the need for the Web semantic search methods and means. Modern principle tendencies in the personalized semantic deal with the ontology-based formalization of knowledge about the information needs of users and their current tasks. Our approach is based on use of the semantically marked Wiki-resources as a source of knowledge for constructing of ontologies of individual task. Each task ontology is considered as a small subset of domain ontology with large number of terms volume and high complexity. We develop the formal model of the task ontology as a special case of the formal model of domain ontology and propose methods for constructing of this task ontology on base of the set of semantically marked Wiki pages. The software implementation of the retrieval systems capable to process such task ontologies is analyzed on example of MAIPS – semantic retrieval system which provides personalized satisfaction of user information needs. Such personalization is based on knowledge contained into the user task thesauri that are built on base of task ontologies.

Key words: semantic retrieval, task ontology, semantic Wiki resources, personalization of retrieval.

Пошук у Web є ключовою технологією Web, тому що дозволяє отримати доступ до його ресурсів. Результатами пошуку можуть бути різні види інформаційних ресурсів або їх елементів – природнозмінені документи та їх фрагменти, структуровані або слабкоструктуровані дані. Сам інформаційний запит, що формалізує інформаційні потреби користувача, теж може бути представлено по-різному – ключовими словами, структурованими мовними записами (SQL, SQLR тощо) або прикладами тих ІО, що потрібні користувачеві. Ціль пошуку у найбільш широкому розумінні – співставити запит з описаним тип ІР, про який відомо пошукової системи, та знайти серед них найбільш релевантні. На сьогодні збільшення ефективності пошуку пов'язують не з його повнотою, а з фільтрацією результатів, яка дозволяє з великої кількості ІР, релевантних запиту користувача, відібрати ті ІР, які стосуються його поточних інтересів, позбавивши самого користувача від рутинного аналізу запропонованої інформації.

Великий обсяг та складна структура інформаційних ресурсів (ІР) Web викликають потребу у створенні таких засобів їх автоматизованої обробки, які забезпечили б аналіз змісту цих ресурсів та дозволили б здобувати з них саме ті відомості, що потрібні користувачам для вирішення конкретних задач. Для такої обробки необхідно розробити як засоби семантичного пошуку, так і засоби семантичної розмітки ІР. Обидві ці задачі безпосередньо пов'язані із семантизацією інформації. У широкому розумінні семантизація – встановлення зв'язків між певним інформаційним об'єктом (ІО) та його змістом. Для відкритого середовища

197

Web під семантизацією ІР будемо розуміти встановлення зв'язків між ІР (або їх компонентами), доступ до яких забезпечує середовище Web, та елементами певним чином формалізованих знань, які теж є доступними через Web. Семантизація інформаційних потреб користувача – це встановлення зв'язків між формалізованим поданням цих потреб (запитом, документом-зразком тощо), наданим користувачем за допомогою Web-інтерфейсу, та елементами формалізованих знань, теж доступних через Web. Таким чином, у семантичному пошуку проблема співставлення інформаційних потреб користувача з наявними ІР зводиться до співставлення тих елементів знань, з якими вони пов'язані.

Семантичний пошук [1] – це окремий випадок пошуку інформації, в якому інформаційні потреби користувача в процесі розв'язання певної задачі задовольняються із застосуванням знань про суб'єкти і об'єкти пошукової процедури й методів аналізу цих знань [2]. Для пошуку можуть використовуватися знання, що стосуються користувача та специфіки його інформаційних потреб (персоналізація пошуку), знання про структуру та типи ІР, серед яких здійснюється пошукова процедура, та про ту предметну область (ІР), до якої відносяться ці ІР. Використання знань може бути приховано від користувача (внутрішня база знань пошукової системи, до якої користувач не має доступу), що ускладнює розуміння роботи такої системи та зменшує довіру до неї, або здійснюється безпосередньо за його вибором (база знань користувача), що вимагає від користувача певної обізнаності у пошуку та додаткових зусиль. Система семантичного пошуку (ССП) є відкритою, якщо у пошуку застосовуються знання, здобуті з відкритого середовища. Прикладами таких знань є онтології ІР, отримані із зовнішніх репозиторіїв.

Основна тенденція семантизації пошукових засобів пов'язана з переходом від виявлення документів, що містять певні ключові слова, до пошуку знань, необхідних для виконання поставленого перед користувачем завдання. Семантична розмітка ІР, серед яких виконуються пошук, дозволяє автоматизувати такий аналіз, використовуючи зовнішні бази знань, з якими пов'язуються контент ІР, для співставлення зі сферою інтересів користувача, яка теж описується певним чином формалізованими знаннями відповідної ІР. Слід відмітити, що використання онтології ІР для семантичної розмітки ІР та використання онтології семантичної розмітки ІР для вдосконалення онтології ІР розглядаються в цій роботі як два взаємопов'язані етапи онтологічного аналізу ІР, що застосовуються ітеративно до того часу, поки користувач не буде задовольнений побудованою онтологією.

На ефективність пошуку значним чином впливає те, які саме бази знань використовуються в ньому. Важливо, з одного боку, які формалізми та мови використовуються для подання знань (це можуть бути семантичні мережі та фреймові системи, таксономії тощо, мови OWL [3] та RDF [4] тощо), а з другого – наскільки цей набір знань є актуальним та відповідним потребам конкретного користувача та його поточній задачі. Якщо для рішення першого питання на сьогодні для Web-орієнтованих застосувань фактично стало стандартом використання онтологій [5] (і нинішня більшою мірою стосується того, якої складності онтології потрібно використовувати), то для другого – побудови [6] або пошуку найбільш релевантної онтології – поки що загальноприйнятні рішення ще не знайдено.

Сьогодні поширені різні підходи до семантизації пошуку в Web. Деякі з них використовують складні структуровані мови для подання запитів користувачів, інші орієнтовані на пошук за ключовими словами або на природнозмінені запити. Більшість підходів першої групи базуються на RDF і орієнтовані на технології проекту Semantic Web, тобто здійснюють пошук у вже семантично розмічених ІР. Наприклад, Swoogle [7] використовує кроуд для виявлення, індексації та запитання документів в форматі RDF. На жаль, таку розмітку на сьогодні має лише невелика частина ресурсів Web. Більш деталізовані переваги та проблеми різних моделей пошуку розглянуто в [8]. В обох напрямках пошуку широко застосовуються онтології, але ці онтології звичайно є однаковими для всіх користувачів пошукової системи.

Використання онтологій для подання знань у Web

Онтології є потужним та ефективним засобом подання знань у Web, але слід врахувати, що обробка онтологічних структур дуже складна і потребує у загальному випадку багато часу. Але досить часто для рішення задачі у певній ІР потрібно врахувати не всі поняття цієї області та зв'язки між ними, а тільки невелика їх підмножина. Тоді використання онтології ІР поділяється на два етапи – виокремлення необхідної для задачі підмножини знань онтології – ІО з простішою структурою, меншим обсягом та вимірністю (приміром, таксономії, тезаурус або словника), що значно пришвидшує їх обробку. Те, за якими умовами будуть створені ці ІО та яку структуру вони мають, залежить від специфіки задачі. Приміром, для задачі пошуку це може бути набір класів онтології, які використовуються у семантичній розмітці заданого текстового документу-зразка; підкласи цих класів та екземпляри цих класів. Інший поширений варіант побудови такого ІО – набір класів онтології, на екземпляри яких є послання у семантичній розмітці з обраної сторінки-зразку. Важливо, що, як правило, такі ІО не описують характеристик класів, відношень та екземплярів, але ці характеристики можуть бути застосовані в процесі побудови цих ІО за онтологією відповідно до кожної конкретної задачі. Приміром, у побудові ІО може бути використана симетричність певного відношення між екземплярами класів, але сам ІО не містить цю характеристику. Таким чином, це

198

дозволяє перейти від "важких" онтологій до "легких", створюючи спрощену підмножину великовогової онтології [9].

Аналіз існуючих на сьогодні засобів семантичного пошуку показує, що у переважній більшості задач для опису ІР достатньо використовувати "легкі" ("легковогові") онтології, в яких не застосовуються аксіоми, а множина відношень між класами, що описуються додатковими властивостями, звичайно є порожними. Легкі онтології – це відносно прості структури термінів з відповідними визначеннями та обмеженою множиною відношень. Такі онтології мають менші варіанти можливості, але їх значно легше та швидше обробляти. Для таких онтологій характерна наявність тільки наступних відношень:

- відношення "клас-підклас", значення якого пов'язані транзитивно;
- об'єктне відношення синонімії, значення якого пов'язані симетрично;
- об'єктні відношення, специфічні для ІР, значення яких не мають бінарних властивостей.

У більш складних випадках деякі відношення ІР можуть також бути визначені як транзитивні або симетричні, і це дозволяє реалізувати більш складне логічне введення на знаннях ІР. Але частине для семантичного пошуку достатньо самої наявності відношення між екземплярами класів, щоб оцінити релевантність ІО області пошуку. Це пов'язано з тим, що аналіз, приміром, природнозміненого тексту із урахуванням усіх знань щодо ІР для визначення його семантики займає дуже багато часу, що неприйнятно для аналізу великої кількості ІР, релевантних запиту. Залежно від того, наскільки складно описує користувач потрібну ІР, залежить як час пошуку, так і його якість. Обирати компроміс між цими протилежними критеріями має сам користувач.

Така "легка" онтологія може базуватися на онтології більш широкій ІР (з більшою кількістю класів, відношень та екземплярів), яка має значно більшу виразну потужність (приміром, подана на OWL Full або OWL-DL). Але безпосередньо побудувати підмножину онтології для переважної більшості користувачів є надто складною задачею. Значно легше обробляти ІР, що містять відповідну семантичну розмітку, обираючи ті, що відносяться до поточної задачі.

Постановка задачі

Щоб реалізувати семантичний пошук інформації в Web, використовуючи знання, що містяться в онтології, доступ до яких також забезпечує середовище Web, виникає потреба у засобах здобуття з цих онтологій тих відомостей, що стосуються інформаційних потреб користувача, та методів застосування здобутих відомостей у пошукових процедурах. Пропонується будувати за онтологією довільної структури та складності більш прості інформаційні об'єкти, що містять визначену користувачем підмножину онтологічних знань, та співставити з цими інформаційними об'єктами результати, отримані від зовнішніх інформаційно-пошукових систем, відбираючи семантично близькі до потреб користувача. Для вдосконалення онтологій, що застосовуються для пошуку, пропонується застосовувати інформаційні ресурси, семантично розмічені елементами цих онтологій, а саму таку розмітку використовувати для побудови відповідних ІО.

Семантичні Wiki як джерело онтологічної інформації

На сьогодні існує велика кількість пошукових систем, орієнтованих саме на пошук структурованих ІР, приміром, для виявлення, індексації та запитання документів в форматі RDF або OWL. Вони дозволяють аналізувати вміст ІР з семантичною розміткою, але, на жаль, переважна частина ресурсів Web не є семантизованими, хоча кількість ІР, що містять семантичну розмітку та різноманітні метадані, постійно збільшується. Тому виникає потреба використовувати семантично розмічені ІР як джерело знань для обробки несемантизованих документів. В якості таких джерел пропонується використовувати бази знань, побудовані на Wiki-платформі. По-перше, такі системи застосовують стандартизовані засоби для подання семантичної розмітки (за допомогою системи категорій та властивостей). Ці елементи можна легко розпізнавати навіть у тих випадках, коли різні ІО отримані з різних ресурсів. Інша важлива перевага – у семантичних Wiki-ресурсах завжди є наявності різноманітні засоби для внутрішньої навігації та пошуку, і це дозволяє користувачу досить швидко визначити набір Wiki-сторінок, пов'язаних з його задачею. Третій важливий фактор – на сьогодні вже створено багато різноманітних семантичних Wiki-ресурсів, і їх кількість, обсяг та якість постійно збільшуються. У четвертих, у випадку, якщо інформації в семантичних Wiki недостатньо, їх досить легко доповнити відомостями з несемантизованих Wiki (приміром, з Вікіпедії різними мовами або Wiki-довідниками) – з таких ресурсів можна здобути менше семантичної інформації, але у поєднанні з семантизованими вони дозволяють досить коректно описати довільну проблему.

Для Wiki-технологій зараз використовуються багато семантичних розширень. Сформовані на їх основі ІР можуть дивувати всім співставленням користувачів, що забезпечує актуальність інформації, мають зручну й просту для розуміння структуру, забезпечують обробку інформації на семантичному рівні, надаючи технологічну платформу для групового керування знаннями. Одним з таких

розширень є Semantic MediaWiki [10]. Ця набувона над MediaWiki має досить високу виразну потужність, надійну реалізацію і зручний інтерфейс користувачів. Сьогодні на SMW базується значна кількість сайтів, порталів та енциклопедичних видань. Прикладом ІР, що базується саме на цій платформі, є електронна версія Великої української енциклопедії [2].

Алгоритм побудови онтології задачі користувача

Формальна модель онтології визначені її основні елементи, їх структуру та зв'язки. Надалі будемо використовувати наступну формальну модель $O = (X, R, F, T)$, що складається з наступних елементів:

- $X = X_{cl} \cup X_{ind}$ – множина концептів онтології, де X_{cl} – множина класів, X_{ind} – множина екземплярів класів, таких, що $\forall a \in X_{ind} \exists A \in X_{cl}, a \in A$;
- $R = \{r_{er_cl} \cup \{r_1\} \cup \{r_2\}$ – множина відношень між елементами онтології, де r_{er_cl} – ієрархічні відношення, що можуть встановлюватися між класами онтології і властивостями класів і характеризуються такими властивостями, як антисиметричність і транзитивність, $r_{er_cl} : X_{cl} \rightarrow X_{cl}$; $\{r_1\}$ – множина об'єктних властивостей, що встановлюють стосунки між екземплярами класів: $r_1(a, a \in X_{ind}) = b, b \in X_{ind}$, $r_1 : X_{ind} \rightarrow X_{ind}$; $\{r_2\}$ – множина властивостей даних, що установлюють відношення між екземплярами класів і значеннями: $r_2(a, a \in X_{ind}) = t, t \in T$, $p_1 : X_{ind} \rightarrow Const$, такі, що усередні множин об'єктних властивостей і властивостей даних також можуть існувати ієрархічні відношення r_{er_obj} , $r_{er_obj} : \{r_1\} \cup \{r_2\}$ і $r_{er_data} : \{r_2\} \rightarrow \{r_3\}$;
- F – множина характеристик класів онтології, екземплярів класів і їхніх властивостей, що можуть застосовуватися для логічного виводу (наприклад, еквівалентність, відмінність, відсутність перетинання, область визначення й область значення);
- T – множина типів даних (наприклад, рядок, ціле).

Формально проблема побудови онтології задачі користувача полягає у наступному: за онтологією ІР O_{domain} , $O_{domain} = (X_{domain}, R_{domain}, F_{domain}, T_{domain})$, та набором Wiki-сторінок U_{user} , семантична розмітка яких базується на O_{domain} , побудувати "легкозавяжну" онтологію задачі користувача O_{user} , знання якої є підмножиною знань з O_{domain} . Потрібно відмітити, що джерела та методи побудови цієї онтології ІР знаходяться поза сферою розгляду даної роботи – вона може мати довільну структуру, високу виразну потужність та бути сформована як безпосередньо експертами ІР, так і за допомогою різноманітних засобів здобуття онтологічних знань.

$O_{user} = (X_{user}, R_{user}, F_{user}, T_{user})$, така, що:

- $X_{user} \subseteq X_{domain}$, тобто $X_{cl_user} \subseteq X_{cl_domain}$, $X_{ind_user} \subseteq X_{ind_domain}$;
- $R_{user} \subseteq R_{domain}$, тобто $r_{er_cl_user} = r_{er_cl_domain}$, $\{r_{user}\} \subseteq \{r_{domain}\}, i = \overline{0, n}, j = \overline{0, m}, m \leq n$;
- $\{r_{user}\} \subseteq \{r_{domain}\}, i = \overline{0, q}, k = \overline{0, l}, l \leq q$;
- $F_{user} = \emptyset$;
- $T_{user} \subseteq T_{domain}$.

Будувати онтологію задачі доцільно тільки в тому випадку, якщо користувач планує працювати над великою та досить складною проблемою, рішення якої буде потребувати інформації на протязі досить значного часу, значно більшого, ніж час, потрібний на побудову власної онтології (приміром, плануючи дослідження на кілька років, доцільно витратити кілька годин на те, щоб надіати отримувати семантично відфільтровані відомості).

Основні етапи побудови онтології задачі

Етап 1. Знайти семантичний Wiki-ресурс W , що за тематикою співвідноситься із задачею користувача або перекриває більш широку ІР. Найпростіше використовувати неспеціалізовані енциклопедії та довідники

(такі, як електронна версія Великої української енциклопедії [11]), але, якщо користувач має відомості до більш спеціалізованих ресурсів, то їх застосування може збільшити ефективність роботи.

Етап 2. У цьому Wiki-ресурсі відібрати множини $IP_{W_{user}}$, що безпосередньо пов'язані із задачею користувача, $W_{user} \subseteq W$. Почати цей відбір можна із пошуку Wiki-сторінок, назви яких співпадають із ключовими словами цієї задачі, а надалі за допомогою вбудованих засобів навігації Wiki-ресурсом переходити до сторінок, сполучених із цими сторінками семантичними властивостями (усіма або тільки тими, що цікавлять користувача) або відносяться до тих самих категорій. На цьому етапі користувач має виконати певну кількість роботи самостійно, щоб охарактеризувати ту інформацію, що йому потрібна, та відкинути ту, що стосується його поточної задачі (це може бути цінна інформація, важлива для ПроО в цілому, але не потрібна саме для рішення поточної проблеми). Від того, наскільки точно буде виконано відбір, залежить ефективність виконання семантичного пошуку: відсутність потрібної інформації не дозволить знаходити відповідні ресурси, а наявність зайвих сторінок збільшить час обробки.

Етап 3. Здійснюється аналіз інформації з $W_{user} = \{w_{user_i}, i = \overline{1, s}\}$.

Для кожної сторінки виконуються наступні дії:

- інформація про класи сторінки (усі або відібрані користувачем) дозволяє поповнити $X_{cl_{user}}$.

Ієрархічні відношення між цими класами можна визначити за допомогою сторінок цих категорій;

- ім'я самої сторінки заноситься до $X_{ind_{user}}$;
- імена тих семантичних властивостей сторінок, що використані на даній сторінці та область визначення яких відносяться до типу "Сторінка" (усі або відібрані користувачем), заносяться до $\{f_{user_j}\}$;
- імена сторінок, на які дана сторінка посилається за допомогою семантичних властивостей типу "Сторінка" (усі або відібрані користувачем), також заносяться до $X_{ind_{user}}$;
- імена сторінок, на які дана сторінка посилається за допомогою гіперпосилань (усі або відібрані користувачем), також заносяться до $X_{ind_{user}}$;
- якщо дана сторінка відсилає до іншої сторінки, тоді ім'я такої сторінки розглядається як синонім поточної сторінки, заноситься до $X_{ind_{user}}$ та пов'язується відношенням синонімії із іменем поточної сторінки.

Якщо обробляються несемантизовані Wiki-сторінки, тоді алгоритм їх обробки значно скорочується:

- інформація про класи сторінки (усі або відібрані користувачем) дозволяє поповнити $X_{cl_{user}}$.

Ієрархічні відношення між цими класами можна визначити за допомогою сторінок цих категорій;

- ім'я самої сторінки заноситься до $X_{ind_{user}}$;
- імена сторінок, на які дана сторінка посилається за допомогою гіперпосилань (усі або відібрані користувачем), також заносяться до $X_{ind_{user}}$;
- якщо дана сторінка відсилає до іншої сторінки, тоді ім'я такої сторінки розглядається як синонім поточної сторінки, заноситься до $X_{ind_{user}}$ та пов'язується відношенням синонімії із іменем поточної сторінки.

Результатом виконання цих етапів є онтологія задачі користувача. Хоча існує можливість їх автоматизації, через те, що більшість операцій потребує втручання користувача та прийняття рішення щодо кожного поняття та відношення, доцільніше виконувати ці дії безпосередньо за допомогою редактору онтологій (наприклад, Protégé [12]). Таким чином побудована онтологія представлена мовою OWL та може використовуватися семантичними пошуковими системами для фільтрації результатів запитів [13]. Основна ідея такої фільтрації полягає у тому, що контент знайдених IP (або певних їх елементів, визначених користувачем) порівнюється із елементами онтології. За результатами порівняння знайдених IP впорядковуються відповідно до кількісного значення коефіцієнту близькості. Алгоритм порівняння та міри близькості залежать від специфіки конкретних пошукових систем та від обраних користувачем налаштувань. Приклад такої системи семантичного пошуку, що використовує онтології користувачів, – металюшкова система МАПС, що призначена для пошуку інформації, пов'язаної з професійними чи науковими інтересами користувачів, – більш детально описано в [9]. Формалізація інформаційних потреб користувача МАПС базується на представленні знань щодо його інтересів через онтології ПроО та похідні від них тезауруси задал.

201

Персоналізація семантичного пошуку в МАПС базується на тому, що користувач явно визначає, які саме знання (онтології та базовані на них тезауруси задал) використовуються для формалізації його інформаційних потреб. Це є характерною рисою цієї ЦСП та її основою відмінністю від інших систем, що підтримують пошук інформації на семантичному рівні.

Висновки

Враховуючи складність створення та обробки онтологій, пропонується використовувати їх для семантизації пошуку в Web наступним чином: в онтологіях накопичувати в інтероперабельному вигляді знання ПроО, поповнюючи їх відомостями з тих інформаційних ресурсів, для семантичної розмітки яких застосовувалися елементи цих онтологій, а для рішення прикладних задач будувати за онтологіями ПроО похідні інформаційні об'єкти (з простішою структурою та меншим обсягом, в яких аксіоми ПроО використовуються на етапі побудови, але не містяться безпосередньо), які надалі використовувати як джерело знань, необхідних для задачі користувача. Запропонований у роботі алгоритм дозволяє будувати онтології, що характеризують поточну задачу користувача і можуть застосовуватися для семантичного пошуку інформації.

Література

1. Baeza-Yates R., A. Raghavan R. Next generation Web search. S. Ceri and M. Brambilla, editors, Search Computing, Springer, 2010. P. 11–23.
2. Rogushina J. Analysis of Automated Matching of the Semantic Wiki Resources with Elements of Domain Ontologies. *International Journal of Mathematical Sciences and Computing (IJMSC)*. 2017. Vol. 3, N 3. P. 50–58.
3. Antoniou G., Van Harmelen F. Web ontology language: Owl. *In Handbook on ontologies*. Springer Berlin Heidelberg, 2004. P. 67–92.
4. Cyganiak R., Wood D., Lanthaler M. RDF 1.1 Concepts and Abstract Syntax. W3C Recommendation 25 February 2014. <http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>.
5. Gruber T.R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies*. 1995. Vol. 43. Issues 5–6. P. 907–928.
6. Gladun A., Rogushina J., Valencia-García R., Martínez-Béjar R. Semantics-driven modelling of user preferences for information retrieval in the biomedical domain. *Informatics for health and social care*. 2013. Vol. 38, N 2. P. 150–170.
7. Finn T.W., Ding L., Pan R., Joshi A., Kolar P., Java A., Peng Y. Swoogle: Searching for knowledge on the Semantic Web. Proc. AAAI-2005, AAAI Press, MIT Press, 2005. P. 1682–1683.
8. Fazzingaa B., Lukaszewicz T. Semantic search on the Web. *Semantic Web – Interoperability, Usability, Applicability*. 2010. N 1. P. 1–7. https://www.researchgate.net/profile/Thomas_Lukaszewicz/publication/220575552_Semantic_search_on_the_Web/links/0046351e94ce8994bd000000.pdf.
9. Рогошина Ю.В. Семантичний пошук у Web на основі онтологій: розробка моделей, засобів і методів. МДПУ ім. Богдана Хмельницького. 2015. 291 с.
10. Krotzsch M., Vrandečić D., Volkel M. Semantic MediaWiki. <http://c.unik.no/images/6/6d/SemanticMW.pdf>.
11. Рогошина Ю.В. Разработка распределенных интеллектуальных систем на основе онтологического анализа и семантических wiki-технологий. *Онтология проектирования*. 2017. Т. 7, № 4(26). С. 453–472.
12. Protégé. <http://protege.stanford.edu/>.
13. Rogushina J.V. The Use of Ontological Knowledge for Semantic Search of Complex Information Objects. *Open semantic technologies for intelligent systems*. 2017. P. 127–132.

References

1. Baeza-Yates R., A. Raghavan R. Next generation Web search. S. Ceri and M. Brambilla, editors, Search Computing, Springer, 2010. P. 11–23.
2. Rogushina J. Analysis of Automated Matching of the Semantic Wiki Resources with Elements of Domain Ontologies // *International Journal of Mathematical Sciences and Computing (IJMSC)*. 2017. Vol. 3, N 3. P. 50–58.
3. Antoniou G., Van Harmelen F. Web ontology language: Owl. *In Handbook on ontologies*. Springer Berlin Heidelberg, 2004. P. 67–92.
4. Cyganiak R., Wood D., Lanthaler M. RDF 1.1 Concepts and Abstract Syntax. W3C Recommendation 25 February 2014. <http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>.
5. Gruber T.R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies*. 1995. Vol. 43. Issues 5–6. P. 907–928.
6. Gladun A., Rogushina J., Valencia-García R., Martínez-Béjar R. Semantics-driven modelling of user preferences for information retrieval in the biomedical domain. *Informatics for health and social care*. 2013. Vol. 38, N 2. P. 150–170.
7. Finn T.W., Ding L., Pan R., Joshi A., Kolar P., Java A., Peng Y. Swoogle: Searching for knowledge on the Semantic Web. Proc. AAAI-2005, AAAI Press, MIT Press, 2005. P. 1682–1683.

202

8. Fazzingaa B., Lukaszewicz T. Semantic search on the Web. *Semantic Web – Interoperability, Usability, Applicability*. 2010. N 1. P. 1–7. https://www.researchgate.net/profile/Thomas_Lukaszewicz/publication/220575552_Semantic_search_on_the_Web/links/0046351e94ce8994bd000000.pdf.
9. Rogushina J.V. Semantic retrieval for Web on base of ontologies: design of models, tools and methods. Melitopol: Bogdan Hmelnytsky MDUPU. 2015. 291 p. [in Ukrainian]
10. Krotzsch M., Vrandečić D., Volkel M. Semantic MediaWiki. <http://c.unik.no/images/6/6d/SemanticMW.pdf>.
11. Rogushina J.V. Design of distributed intelligent systems on base of ontological analysis and semantic Wiki technologies. *Ontology of designing*, Vol. 7, N 4(26) 2017. P. 453–472. [in Russian]
12. Protégé. <http://protege.stanford.edu/>.
13. Rogushina J.V. The Use of Ontological Knowledge for Semantic Search of Complex Information Objects. *Open semantic technologies for intelligent systems*. 2017. P. 127–132.

Про автора:

Рогошина Юлія Віталіївна,

кандидат фізико-математичних наук,
старший науковий співробітник Інституту програмних систем НАН України.
Кількість наукових публікацій в українських виданнях – 140.
Кількість наукових публікацій в зарубіжних виданнях – 30.
Індекс Хірша – 10.
<http://orcid.org/0000-0001-7958-2557>.

Місце роботи автора:

Інститут програмних систем
НАН України,
03181, Київ-187,
проспект Академіка Глушкова, 40.
Тел.: 066 550 1999.
E-mail: ladamandraka2010@gmail.com

203