

УДК 004.91:93

В.І. МЕСЮРА, М.В. ОСТРОВСЬКА, О.Д. ПОЛЩУК

СИСТЕМА ОБРОБКИ ІНФОРМАЦІЇ НА ОСНОВІ МІКРОБЛОГІНГУ ДЛЯ КІБЕРІНФРАСТРУКТУРИ ПІДПРИЄМСТВА

*Вінницький національний технічний університет,
21021, Хмельницьке шосе, 95, Вінниця, Україна,
тел.: +380 (97) 2647791, E-mail: vimes2009@yandex.ua*

Анотація. В статті запропоновано систему обробки інформації на основі мікроблогінгу для кіберінфраструктури підприємства. Оновлення бази знань системи здійснюється з використанням опорної онтології у вигляді RDF-схеми предметної області, шляхом обробки інформації, що надходить з платформ мікроблогінгу. Нові знання виводяться з використанням онтологічних правил. Формування звітів здійснюється на основі секвенційного методу. Розроблено алгоритм реалізації методу в контексті підтримки прийняття рішень. Запропоновану систему реалізовано мовою програмування java з використанням бібліотек ARQ, twitter4j, restfb.

Аннотация. В статье предложена система обработки информации на основе микроблогинга для киберинфраструктуры предприятия. Обновление базы знаний системы осуществляется с использованием опорной онтологии в виде RDF-схемы предметной области, путем обработки информации, поступающей из платформ микроблогинга. Новые знания выводятся с использованием онтологических правил. Формирование отчетов осуществляется на основе секвенциального метода. Разработан алгоритм реализации метода в контексте поддержки принятия решений. Предложенную систему реализовано на языке программирования java с использованием библиотек ARQ, twitter4j, restfb.

Abstract. In the article information processing system based on microblogging for enterprise cyberinfrastructure. The knowledge base is updated by using support ontology, as RDF-schemas of subject areas, by processing of the information, received from the platforms of microblogging. New knowledge derived using ontological rules. The reports generating is based on the sequential method. The method realization algorithm in the context of the decision support was designed. This system is developed using java programming language and ARQ, twitter4j and restfb libraries.

Ключові слова: система обробки інформації, онтологія, мікроблогінг, кіберінфраструктура підприємства, RDF-схема, Twitter, секвенції, SPARQL-запити.

ВСТУП

Сучасний стан розвитку інформаційних та комунікаційних технологій забезпечує можливість миттєвого обміну інформацією про будь-яку подію між будь-якими абонентами. Це надає принципову можливість зробити користувача безпосереднім учасником виробництва, миттєво сприймаючи його побажання та підлаштовуючи під нього виробничий процес. З іншого боку, користувач може миттєво отримувати інформацію щодо стану підприємства, його товари, послуги і перспективні плани та брати участь в формуванні асортименту та планів розвитку підприємства. Отож, актуальною є розробка програмної системи обробки інформації (COI), що надходить від користувачів, для застосування на підприємствах з кіберінфраструктурною підтримкою [1,2,3].

ПОСТАНОВКА ЗАДАЧІ

Однією з найважливіших задач при розробці кіберінфраструктури підприємства стає задача автоматичного сприйняття та обробки дуже великих потоків інформації. При цьому необхідно динамічно розпізнавати, яких саме питань стосується отримана інформація і, відповідно, структурувати її у вигляді

певної бази знань. Оскільки автоматичне розпізнавання сенсу повідомлень природної мови є надто складною задачею, при створенні COI обмежимося прийомом коротких повідомлень, по типу тих, що використовуються, наприклад, у системі Twitter [4]. При цьому, для подальшого спрощення процедури розпізнавання вмісту повідомлень можна накладати певні обмеження на їх формат. Такий підхід забезпечує можливість покласти виконання функцій комунікаційного середовища [5] для обміну інформацією на існуючі соціальні мережі, що значно знизить вартість розробки і надасть доступ до системи, що розробляється, практично необмеженій кількості користувачів та виробників.

Отже, метою даної роботи є розробка системи обробки інформації для кіберінфраструктури підприємства на основі мікроблогінгу [6], що забезпечить підвищення ефективності діяльності підприємства за рахунок динамічної адаптації виробництва під потреби користувачів.

РОЗРОБКА СТРУКТУРИ МОДЕЛІ COI

Узагальнену структуру моделі COI наведено на рис. 1. Користувачі надсилають короткі текстові повідомлення стосовно виробництва в системі мікроблогінгу, під'єднані до кіберінфраструктури підприємства. Система оновлення онтології добуває з них нові екземпляри класів, які складають опорну RDF-схему онтології, та формує знання про попит, побажання, відгуки, що відображають поточний стан середовища та зміни, які в ньому відбулись [7].

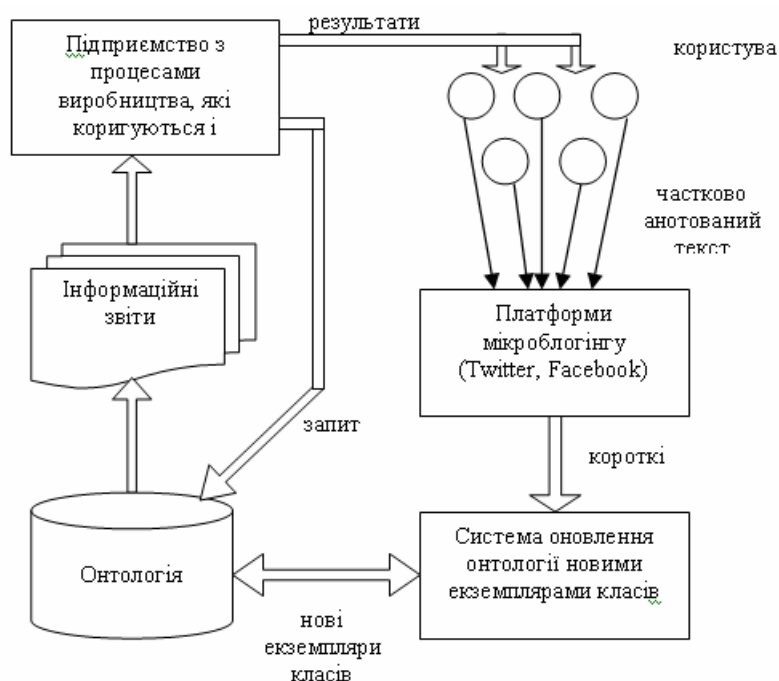


Рис.1. Структурна схема моделі COI для кіберінфраструктури підприємства

Онтологія кіберінфраструктури є центральною системою контекстних знань, на основі якої можлива реалізація системи підтримки прийняття рішень, що входить до моделі кіберінфраструктури підприємства. Фрагмент предметної області зображено на рис. 2.

Основою системи підтримки прийняття рішень є реакція на отримані знання через активні фільтри, які є наборами послідовних запитів до онтології, що виконуються з метою формування інформаційного звіту. На основі аналізу інформаційного звіту можливе генерування повідомлення (сигналу) про необхідність змін в процесах виробництва.

Прийняті рішення щодо коригування процесів виробництва дозволяють зробити результати виробництва такими, що більше відповідають потребам середовища. Таким чином покращується ефективність виробництва шляхом збільшення попиту на продукцію та послуги підприємства.

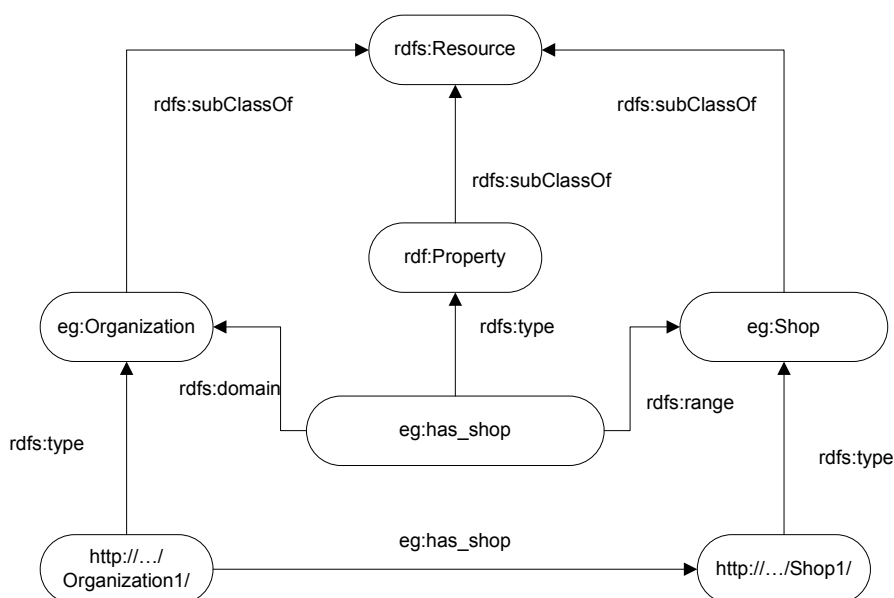


Рис.2. Графічне подання фрагменту онтології

ПІДСИСТЕМА ОНОВЛЕННЯ ОНТОЛОГІЇ

Підходи до оновлення онтологій (наприклад, підхід, що використовується в модулі CLIE [8] в середовищі GATE) використовують статичну структуру опорної онтології. Але в процесі використання онтологій може виникнути необхідність в модифікації її структури: додаванні та вилученні класів та властивостей (предикатів).

Таку можливість надає розроблена система оновлення онтології, яка включає у себе підсистему формування комунікаційного середовища, реалізовану за допомогою спеціальних API, таких як RESTFB для мережі Facebook та Twitter4j для мережі Twitter. Підсистема, що реалізує типові операції з онтологією, реалізована з використанням Jena ARQ [9]:

1. За допомогою системи оновлення онтології відбувається автоматична розмітка на об'єкт, суб'єкт та предикат у вхідному реченні.
2. Використовується динамічний словник предикатів для виявлення предиката, та RDF-схема для виявлення екземплярів класів в повідомленні.
3. Перевіряється наявність в онтології виявлених суб'єкта, предиката і об'єкта, та, в разі їх відсутності, в онтологію заносяться екземпляри класів та зв'язок між ними.
4. Однією з можливостей, які надають онтології, є виведення нових знань за умови використання онтологічних правил. Ці правила вводять в онтологію нові зв'язки між екземплярами класів, що додаються на поточному кроці алгоритму та існуючими в онтології екземплярами (рис.2 - 4).

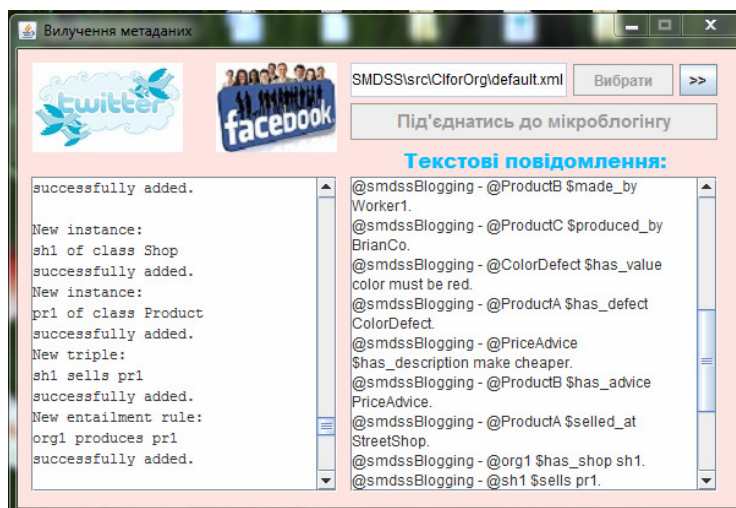


Рис.3. Головне вікно програми під час виконання



Рис.4. Вигляд онтології: а) опорної, б) оновленої

ПІДСИСТЕМА ФОРМУВАННЯ ЗВІТІВ В КОНТЕКСТІ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ

Після формування онтології використовується список введених пасивних та активних послідовних фільтрів. Перший список необхідний для отримання звіту, що демонструє результати роботи підприємства. Список активних послідовних фільтрів необхідний для отримання звіту про поточний стан підприємства та середовища. На основі цього звіту система динамічного спрацювання генерує повідомлення, що являють собою підказки або вказівки, що реалізує завершальний етап системи підтримки прийняття рішень.

Для формування фільтрів [10] використано секвенційний метод [11] для SPARQL-запитів.

Формування секвенції з набору простих запитів включає:

- аналіз набору простих запитів з метою видалення запитів, що повторюються, що дозволяє уникати нерационального завантаження доступу до онтології, скорочуючи кількість операцій зчитування, та уникати дублювання даних в результуючому наборі;
- аналіз набору вхідних та вихідних класів секвенції з метою виявлення однакових класів;
- групування вхідних даних відповідно до вхідних та вихідних класів;
- застосування до секвенції операцій «І» та «АБО».

Враховуючи структуру результатів виконаного запиту до онтології, формування звіту з результатів запиту включає такі етапи:

- групування результатів запиту по атрибутам.
- аналіз результатів запиту з метою групування даних по типам.
- відокремлення даних від типів даних.
- формування з даних звіту у певному вигляді.

В перспективі передбачається вдосконалення програмної реалізації системи формування звітів [12] за рахунок реалізації інтерфейсу побудови фільтрів за допомогою алгоритму «зведених дощок» (Dashboards). Він дозволяє послідовно виконувати проектування фільтрів за допомогою спеціальних цифрових панелей, що відображають дані про частину структури онтології та дають змогу нарощувати складність запиту та

обмеження на запит для одержання необхідної інформації за заданими критеріями. Алгоритм спрощує дії користувача до простих перетягувань панелей (дощечок) на екрані сучасних гаджетів.

ВИСНОВКИ

В роботі запропоновано систему обробки інформації для кіберінфраструктури підприємства на основі мікроблогінгу, яка надає користувачам можливість шляхом простого обміну короткими текстовими повідомленнями в соціальних мережах формувати онтологію, що може в подальшому бути використана як користувачами, так і кінцевими споживачами для добування необхідної інформації в зручному вигляді, її систематизації та пошуку. Розроблено програмне забезпечення системи обробки інформації мовою програмування JAVA з використанням бібліотек ARQ, Twitter4j, RestFB. В програмі реалізовано алгоритми оновлення онтології та секвенційного підходу до формування звітів, які є базою для автоматизації налаштування процесів виробництва на підприємстві у відповідності до динамічних змін навколишнього середовища.

СПИСОК ЛІТЕРАТУРИ

1. Hsu, Ch. Models of cyberinfrastructure-based enterprises and their engineering. An evolutionary journey. In books: Service enterprise integration. An Enterprise Engineering Perspective. [Електронний ресурс] / Under edition: Hsu, Ch. // New York, US-NY: Department of Decision Sciences and Engineering Systems, Rensselaer Polytechnic Institute, 2007. – Режим доступу: <http://viu.eng.rpi.edu/publications/SEIBookHsuCh08.pdf>
2. A Market Mechanism for Participatory Global Query: A First Step of Enterprise Resources Self-Allocation [Електронний ресурс] / Cheng Hsu, Christopher D. Carothers, David M. Levermore. : Information Technology and Management, 2004. – Режим доступу: <http://viu.eng.rpi.edu/publications/EnterpriseResourcesMarket.pdf>.
3. Bilodray Yu.V., Mesyura V.I. Using popular microblogging in Subject-Environment Inter-action Model of CyberInfra-structure-Assisted Enterprises // Грузинский технический университет. Труды. Автоматизированные системы управления». -№1(10), Тбилиси.- 2011. - С. 449-455.
4. Further Use of Controlled Natural Language for Semantic Annotation of Wikis / Brian Davis and Siegfried Handschuh and Hamish Cunningham and Valentin Tablan. Digital Enterprise Research Institute, National University of Ireland, Galway. – 2006.
5. Білодрай Ю.В., Месюра В.І. Використання соціальних медіа для менеджменту групової цілеспрямованої діяльності. - Вісник ВПІ. 2011. – №2 (89). – С. 58-64.
6. Розробка моделі соціальної кіберінфраструктури підприємства на основі мікроблогінгу. / Островська, М.В., Поліщук О.Д., Білодрай В.М., Месюра В.І. : Матеріали другої Міжнародної конференції студентів і молодих науковців – Одеса: ОНПУ, 2012. – С.87 – 88.
7. Островська М.В. Система підтримки прийняття рішень на основі соціальних медіа / В.І.Месюра, М.В.Островська, А.Д.Поліщук // Сучасні інформаційні системи та технології (AIST 2012): матеріали міжнародної науково-практичної конференції, Суми, СумДУ, 17 – 18 травня 2012 р. / Міністерство освіти і науки, молоді та спорту України, Сумський державний університет – Суми, СумДУ, 2012. – С. 95-97.
8. Controlled Language IE Components / Tamara Polajnar, Hamish Cunningham, Valentin Tablan, Kalina Bontcheva. University of Sheffield. – 2006.
9. Островська М.В. Інтелектуальний модуль оновлення семантичного сховища/ М.В. Островська, В.І. Месюра // Свідоцтво про реєстрацію авторського права на твір № 43115: комп'ютерна програма. – Дата реєстрації 06.04.2012.
10. Bilodray Yu.V., Mesyura V.I. Automated sequential approach of building semantic data filters // Сучасні інформаційні технології 2011 (MIT-2011) / Матеріали першої Міжнародної конференції студентів і молодих науковців – Одеса: ОНПУ. – т.1, 2011. – С.157 – 158.
11. Мокін В.Б., Яшолт А.Р. Новий метод автоматичного формування звітів інформаційних систем, оснований на секвенційному підході. – 2008/3 – Вісник Вінницького політехнічного інституту.
12. Поліщук О.Д. Інтелектуальний модуль формування звітів на основі секвенційного підходу/ О.Д. Поліщук, М.В. Островська, В.І. Месюра // Свідоцтво про реєстрацію авторського права на твір № 43118: комп'ютерна програма. – Дата реєстрації 06.04.2012.

Надійшла до редакції 20.06.2012р.

МЕСЮРА В.І. – к.т.н., професор кафедри комп'ютерних наук, Вінницький національний технічний університет, Вінниця, Україна.

ОСТРОВСЬКА М.В. – студентка кафедри комп'ютерних наук, Вінницький національний технічний університет, Вінниця, Україна.

ПОЛІЩУК О.Д. – студент кафедри комп'ютерних наук, Вінницький національний технічний університет, Вінниця, Україна.