

Система підтримки прийняття рішень на основі соціальних медіа

Месюра В.І., Островська М.В., Поліщук О.Д.
Вінницький національний технічний університет, leshka_profy@mail.ru

The purpose of the research is design and implementation of decision support system based on social media. The peculiarity of this system is an approach based on using the ability of microblogging that enables to form a knowledge base of users and systems messages that use social media as a medium of communication. This approach allows you to perform complex automatically created queries to regularly updated knowledge base and extract the necessary information in the form of reports.

ВСТУП

Швидкий розвиток соціальних мереж відкрив широкі можливості щодо створення систем підтримки прийняття рішень на основі соціальних медіа, які представляють собою системи семантичного аналізу інформації в мережах мікроблогінгу з інкапсуляцією складності операцій з семантичними даними.

1. ОБЛАСТІ ЗАСТОСУВАННЯ

Системи підтримки прийняття рішень на основі соціальних медіа можуть бути використані для проектів громадських установ з залученням громадських думок, для наукових проектів по вивченню навколишнього середовища, в якості користувацьких систем моніторингу, інтегрованих в системи smart house, автомобілі тощо. Ці системи можуть бути реалізовані по типу software markets, де треті сторони реалізують зв'язки до різноманітних інформаційних систем, які, як і люди, можуть постачати інформацію в мікроблогінг. Комунікаційним середовищем для системи такого типу слугують існуючі платформи мікроблогінгу, а користувачі мають доступ до неї через існуючі клієнти на РС, Mac, смартфонах у будь-якому місці з доступом до мережі Інтернет.

Загальна схема застосування СППР на основі соціальних медіа зображена на рисунку 1.

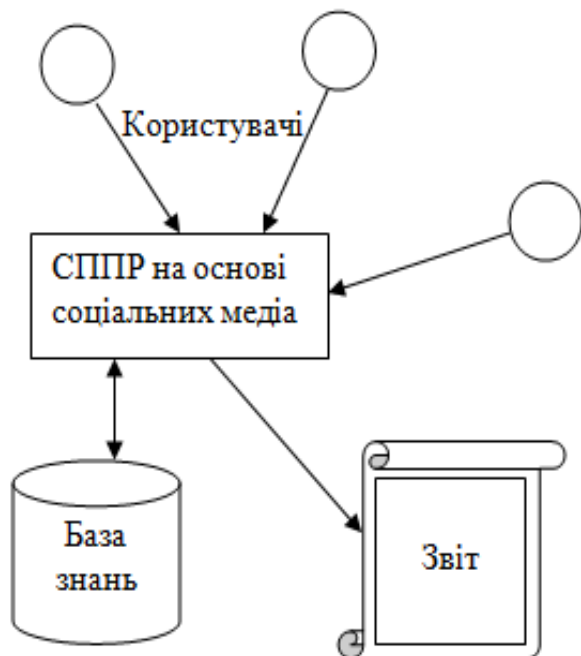


Рисунок 1 – Загальна схема застосування СППР на основі соціальних медіа

2. ПРАКТИЧНА РЕАЛІЗАЦІЯ

Семантичний аналіз інформації, що поступає від агентів-клієнтів, спрощується завдяки використанню коротких повідомлень з частковою анотацією даних відповідно до концепції TwitterData - @ для імен, # для хеш-тегів [1]. Ця концепція реалізує контрольовану мову в стилі анотацій користувачами мікроблогінгу і є сумісною з семантичними мовами OWL и RDF, тому завдяки їх використанню стає можливим без використання складних засобів обробки природної мови відшукувати семантичну інформацію та поміщати її в репозиторії RDF або онтології OWL (бази знань) [2]. На наступному етапі, використовуючи підхід побудови звітів за допомогою секвенцій SPARQL запитів, існує можливість інкапсулювати складність створення запитів до даних RDF або OWL для користувача – на екрані комп'ютера або смартфона комбінуються послідовності секвенцій простих запитів, які автоматично комбінуються в більш складні, використовуючи механізм операцій над секвенціями [3,4].

Реалізація описаної системи підтримки прийняття рішень включає в себе наступні кроки:

1. Створення початкової опорної онтології предметної області засобами GATE [5].

2. Створення додатку в середовищі GATE, який виконає первинний аналіз тексту (розбиття на речення, слова, виділить цифри, пробіли), а також анотує текст в відповідності з створеною онтологією та відповідно до розмітки Twitter. Для виконання цього пункту було створено код на мові JAPE – спеціальна мова специфікації шаблонів, що дозволяє розпізнавати регульовані вирази в анотації документів.

3. Перевірка наявності отриманих після анотації документа текстових значень суб'єкта та об'єкта в онтології, і в разі відсутності створення зразків відповідних класів.

4. Формування секвенції з набору простих запитів, що включає:

- - аналіз набору простих запитів з метою видалення запитів, що повторюються для уникнення нераціонального завантаження доступу до семантичного сховища та дублювання даних;
 - - 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 «I» та «АБО».
5. Враховуючи структуру результатів виконаного запиту до семантичних сховищ, 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.

ВИСНОВКИ

Створення систем підтримки прийняття рішень на основі соціальних медіа надає можливість забезпечити значне спрощення семантичного аналізу інформації і вирішити проблему глобальної комунікаційної інфраструктури і стандартизації процесів і протоколів зв'язку.

ЛІТЕРАТУРА

- [1] Twitter Data – a simple, open proposal for embedding data in Twitter messages [Електронний ресурс] / Todd Fast, Jiri Copsa. – 2009. Режим доступу : <http://twitterdata.org>.
- Controlled Language IE Components / Tamara Polajnar, Hamish Cunningham, Valentin Tablan, Kalina Bontcheva. University of Sheffield. – 2006. – 64с.
- Мокін В.Б. Новий метод автоматизованого формування звітів інформаційних систем на основі секвенційного підходу / В. Б. Мокін, А. Р. Ящолт // Вісник Вінницького політехнічного інституту. – 2008. - № 3. – С. 50 – 56.
- SPARQL Query Language for RDF, Tutorial [Електронний ресурс] // Lee Feigenbaum. – 2008. – Режим доступу: <http://www.cambridgesemantics.com/2008/09/sparql-by-example/>.
- Developing Language Processing Components with GATE Version 6 : інструкція користувача. / Hamish Cunningham, Diana Maynard, Kalina Bontcheva. The University of Sheffield - Department of Computer Science. – 2001-2011. – 629с.

