

РЕФЕРАТ

магістерської атестаційної роботи

на тему:

“Технології семантичного пошуку в масивах неструктурованої інформації”

Шумейка Юрія Дмитровича

Актуальність роботи

На сьогоднішній день технічні засоби дозволяють зберігати та обробляти великі обсяги інформації. Електронна інформація використовується у всіх галузях суспільної діяльності: в банках, бібліотеках, науково-дослідницьких установах, тощо. Здатність знаходження необхідної інформації в найкоротший час стає важливим та нетривіальним завданням. Для вирішення подібних завдань використовуються інформаційно-пошукові системи. Як правило, дані системи в процесі своєї роботи виходять з припущення про взаємну незалежність слів документа, що ускладнює пошук релевантних документів. Багатозначність слів, велика кількість синонімів, омонімів, тощо, призводить до того, що в результаті роботи системи, можуть бути отримані не релевантні пошуковому запиту документи. Рішенням даної проблеми може бути інформаційно-пошукова система, яка здатна розпізнавати смислове навантаження аналізує мого документа. Тому пошук шляхів підвищення ефективності роботи інформаційно-пошукових систем являється актуальним завданням в наш час.

Ціль роботи

Мета роботи полягає у виявленні недоліків моделей інформаційного пошуку та розробці шляхів їх усунення. Розробка інформаційно-пошукової системи на основі векторної моделі інформаційного пошуку, розширеною семантичною моделлю з використанням онтологій предметних галузей для роботи в єдиному пошуковому механізмі.

Задачі, що розв'язуються в роботі

1. Дослідження особливостей існуючих моделей інформаційного пошуку.
2. Визначення недоліків інформаційно-пошукових моделей та шляхів їх усунення.
3. Розробка моделі системи інформаційного пошуку з використанням онтологій.
4. Розробка інформаційно-пошукової системи на основі запропонованої моделі з використанням онтологій.

Досягнуті результати

Розв'язавши задачі, що поставлені в роботі, автор захищає:

- результати аналізу особливостей існуючих моделей систем інформаційного пошуку;
- розроблена модель інформаційного пошуку з використанням онтологій;
- розробка інформаційно-пошукової системи на основі запропонованої моделі інформаційного пошуку.

Наукова новизна роботи

Наукова новизна роботи полягає в тому, що:

- запропоновано модель інформаційно-пошукової системи з використанням онтологій предметних галузей:
 - для визначення релевантності документа пошуковому запиту, враховуючи сенс тексту документа, що досягається шляхом використанням онтологій предметних галузей;
 - використання онтологій предметних галузей дозволяє включити в запропоновану інформаційно-пошукову систему модуль автоматичної побудови відповіді на поставлене питання;

- Використання онтологій предметних галузей дозволяє тематичне структурування знайдених результатів.

Практична цінність роботи

Практична цінність роботи полягає в отриманні систематизованої теоретичної інформації про організацію інформаційного пошуку в масивах неструктурованої інформації, використовуючи розроблену модель інформаційного пошуку.

Висновки

1. Проаналізовані моделі систем інформаційного пошуку, розглянуто їх основні переваги та недоліки.
2. Запропонована модель інформаційно-пошукової системи, що використовує для визначення релевантності документів пошуковому запиту онтології предметних галузей.
3. Розроблена інформаційно-пошукова система на основі запропонованої моделі інформаційного пошуку.

Робота містить 86 с., 22 рис., 35 джерел.

Ключові слова: ПОШУКОВІ СИСТЕМИ, СЕМАНТИЧНИЙ ПОШУК, ОНТОЛОГІЇ.