

Бібліотека наукових публікацій з використанням RDF- СХОВИЩ

Виконала Галушко Марія

Науковий керівник Булах Богдан Вікторович

Мета

- Виявлення ефективних способів та засобів використання RDF-сховищ у прикладних системах, таких як бібліотеки публікацій.
- Поставлена мета вимагає вирішення наступних наукових задач:
 - 1) аналіз існуючих баз знань наукових публікацій, які використовують RDF-сховища;
 - 2) аналіз способів і засобів використання RDF-сховищ у прикладних системах;
 - 3) створення бази знань наукових публікацій використовуючи проведений аналіз.

Наукова новизна

- Полягає в тому, що система обліку публікацій може здійснювати пошук не лише за бібліометричними даними, як багато існуючих прототипів, але й за контекстом самої публікації, що спрощує пошук необхідної інформації.

Діаграма цілей роботи



Вибір методів та засобів



ОСНОВНІ ПОНЯТТЯ

- **RDF** (Resource Description Framework) значить “середовище опису ресурсів”. Це модель даних, що представляє дані простими триплетами суб'єкт – предикат – об'єкт.
- **Триплет** - твердження про ресурс. Трійка: суб'єкт, предикат, об'єкт.



- **Онтологія** - концептуальна схема, яка формалізує деяку галузь знань.
- **База знань** - онтологія разом із набором індивідуальних примірників класів.
- **SPARQL** - мова запитів до даних, представлених по моделі RDF, а також протокол для передачі цих запитів і відповідей на них.
- **Семантична павутина** - нова концепція розвитку Всесвітньої павутини і мережі Інтернет, яка створена і впроваджується Консорціумом Всесвітньої павутини (англ. World Wide Web Consortium, W3C)
- **OWL** - мова опису онтологій для семантичної павутини.

СТРУКТУРА ОНТОЛОГІЇ

- **Поняття** (представники якоїсь сутності або явища, описують групу індивідуальних сутностей, які об'єднані на підставі наявності загальних властивостей)
- **Відносини** (ставлення IS-A; клас - підклас; відношення a-kind-of)
- **Аксіоми** (висловлюють ту інформацію, яка не може бути відображена в онтології за допомогою побудови ієрархії понять. Н-д, «Якщо X смертний, то X колись помре»)
- **Окремі екземпляри** (конкретні елементи будь-якої категорії (наприклад, екземпляром класу Хлопець буде Петро)

Існуючі рішення

- Google Scholar;
- Scirus;
- DBpedia .



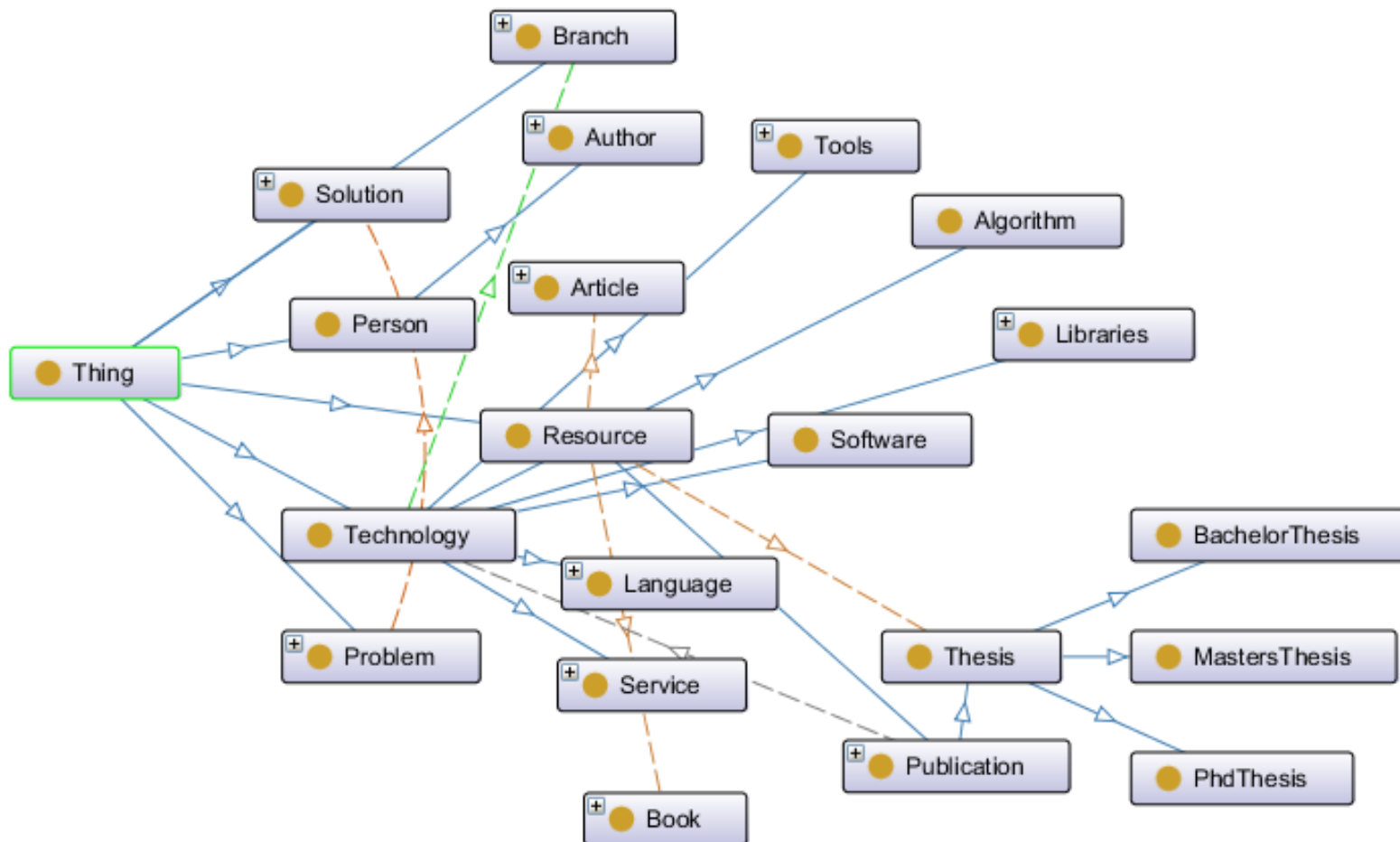
SCIRUS
for scientific information only



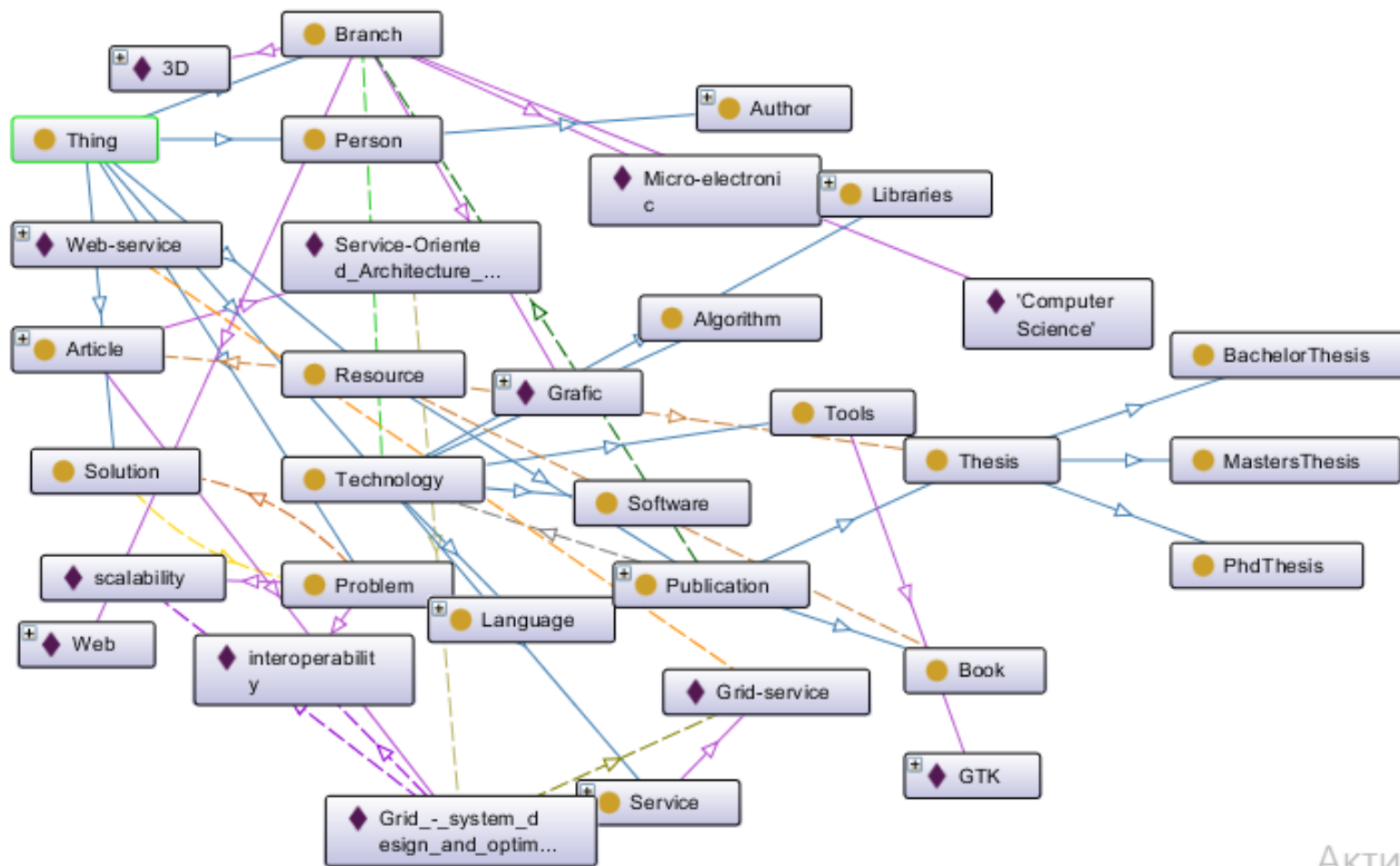
Етапи створення онтології

- 1. Визначення класів онтології:
 - Наприклад, для моєї онтології: Publication, Thesis, Article, Book, TechReport, Description, Branch, Method, Technology, Library, Problems і тд.
- 2. Створення ієрархії класів.
- 3. Визначення слотів (властивостей).
 - Наприклад, «використовує», «належить» «вирішує» (проблему), і тд.
- 4. Заповнення онтології екземплярами та встановлення зв'язків між об'єктами.

Фрагмент онтології бібліотеки публікацій



Фрагмент бази знань



Активна

Тестування прототипу



Знайти всі статті, в яких розглядаються вітчизняні системи для інженерного проектування

Запит

```
SELECT ?name
WHERE
{
  ?article lp:name ?name,
  ?article lp:subject ?subject,
  ?subject lp:use ?tool,
  ?tool lp:type ?Tool,
  ?tool lp:type ?'Native'
}
```

Результат запиту

RESULTS

Grid-systems to develop and optimization of engineering solutions

Методологія розробки онтологій

Мета - показати процес створення онтології

Опис – створення онтології, що описує Грід



Знайти статті, де описується семантичний веб

Запит

```
SELECT ?name
WHERE
{
  ?article lp:name ?name,
  ?article lp:describe ?'Semantic Web'
}
```

Результат запиту

The screenshot shows a search interface with the following elements:

- SEARCH BY CRITERIA**: A section with five input fields: Article, Autor, Theme, Problem, and Solution. The 'Theme' field contains the text 'Semantic Web'.
- Do Search**: A button located below the search criteria fields.
- RESULTS**: A section below the search button, containing the text 'Development methodology of ontologies'.

Запит

Знайти авторів, які описували базу знань Грід

```
SELECT ?author
WHERE
{
  ?article lp:description ?'Grid base of knowledge',
  ?author lp:writeArticle ?article
}
```

Результат запиту

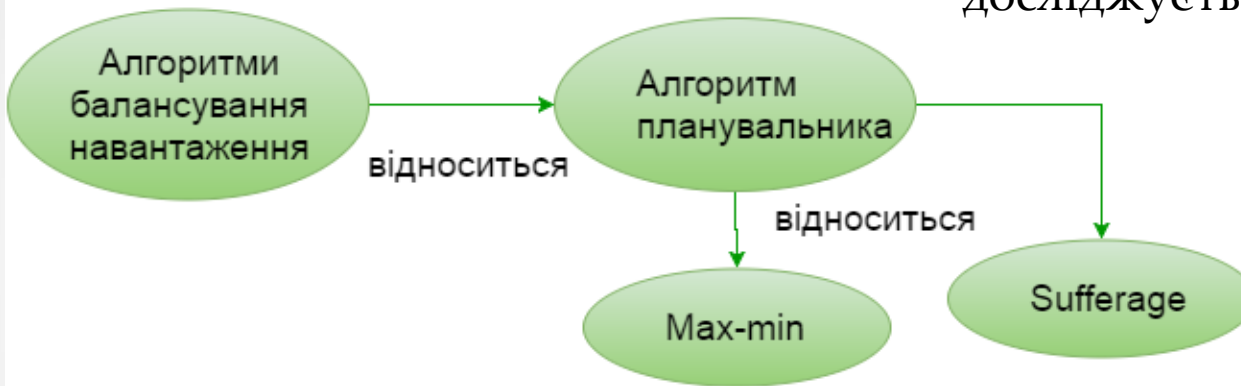
RESULTS

Petrenko A.

Алгоритми балансування навантаження
вГрід-системах

Мета - дослідження
Предмет - алгоритми балансування
навантаження

Знайти всі статті, де
досліджується алгоритм Max-min



Запит

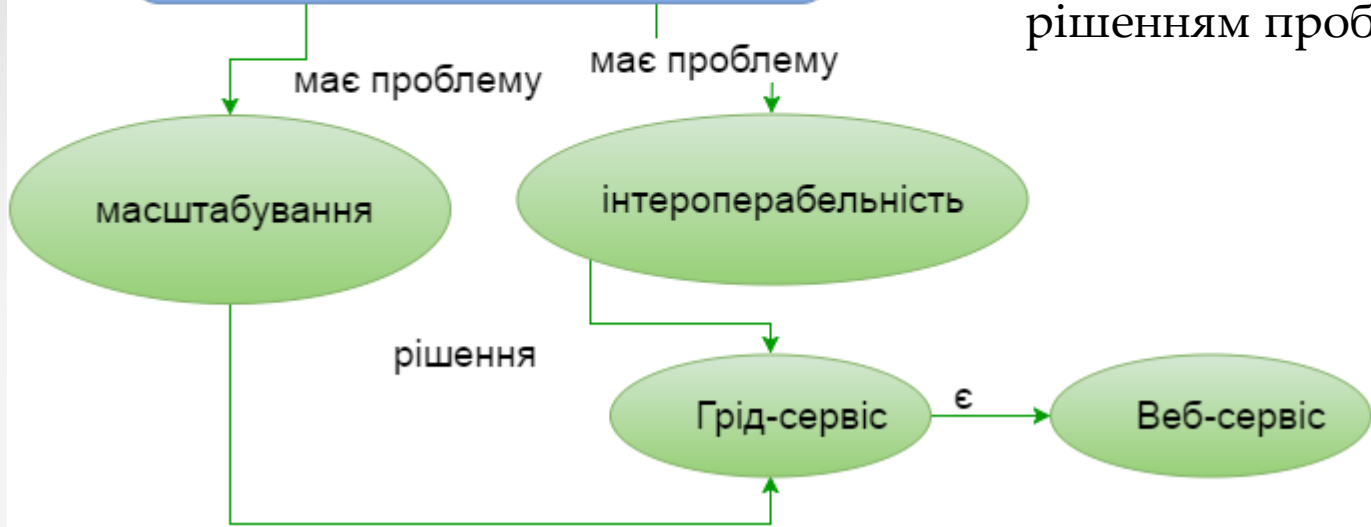
Результат запиту

RESULTS

Load balancing algorithms in Grid-systems

```
SELECT ?name
WHERE
{
  ?article lp:name ?name,
  ?article lp:object ?algorithm,
  ?algorithm lp:name ?'Max-min'
}
```


Предмет дослідження – програмне забезпечення Грід



Знайти усіх авторів публікацій, де рішенням проблеми є веб-сервіс

Результат запити

Petrenko A.
Bulakh B.
Kiselev H

Запит

```
SELECT ?authors
WHERE {
  ?authors pb:property ?Author.
  ?problem pb:type ?Problem.
  ?service pb:type ?'Web-Service'.
  ?problem pb:resolvesBy ?service
}
```

Висновки

- Проаналізовано існуючі системи обліку публікацій чи інформаційних ресурсів, запропоновано підхід, що дозволить створити бібліотеку публікацій з пошуком по контексту;
- Досліджено інструментарій розробки семантичних веб-додатків;
- Було розроблено базу знань наукових публікацій, яка робить можливим пошук по контексту;
- Було створено сайт, за допомогою якого можна здійснювати пошук по базі знань;
- Було протестовано базу знань і показано можливість виконувати «розумний пошук».

Перспективи подальшого розвитку

- Заповнення бази знань більшою кількістю публікацій;
- Застосування бази знань для інших прикладних задач;
- Дослідження методів та засобів автоматизації заповнення бази знань.

Дякую за увагу!