

Вініченко А.М.

# Семантичні композитні веб-додатки в наукових дослідженнях



# Мета роботи

*Метою* магістерської дисертації є:

- дослідження переваг та аналіз способів побудови композитних семантичних веб-додатків,
- визначення можливих способів використання їх у наукових дослідженнях,
- реалізація прототипу такої системи для кращого розуміння переваг та недоліків технологій, які використовуються зараз.

Для цього визначено *завдання*, які вирішуються в роботі:

- Провести аналіз концепції композитних веб-сервісів(мешапів).
  - Дослідити принципи, мету, технічну сторону семантичних веб-додатків, використання і значення онтологій.
  - Запропонувати шляхи використання композитних семантичних веб-додатків у наукових дослідженнях, дослідити існуючі рішення.
  - Експериментально перевірити можливість реалізації досліджених рішень на прикладі тестового семантичного веб-додатку.
- 
-

# Mashups

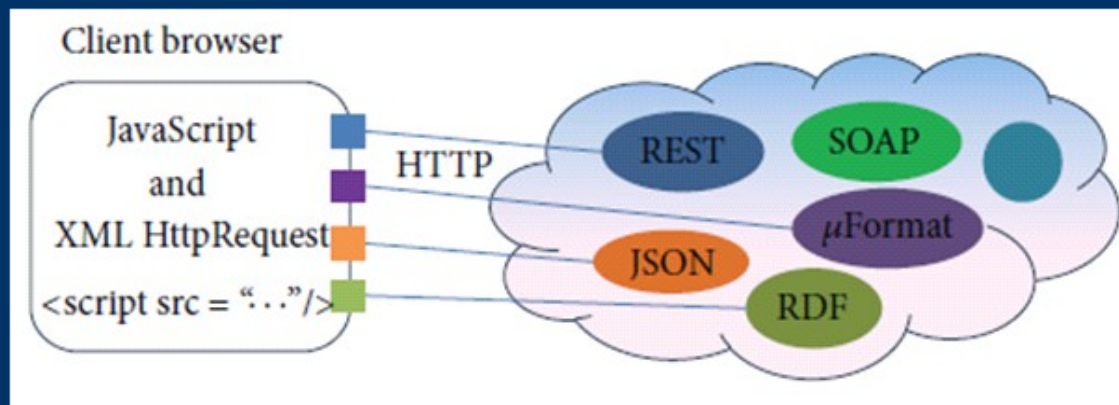
- веб-сторінка або веб-додаток, який використовує контент з більш ніж одного джерела, щоб створити єдиний новий сервіс, що використовує єдиний графічний інтерфейс.
- все більше і більше веб-додатків опубліковують свої API, які дозволяють розробникам програмного забезпечення легко інтегрувати дані і функції, а не будувати їх самостійно
- використання таких технологій як XMLHttpRequest, REST

# Поняття “Семантичний Web”

- підхід до розвитку Вебу, який полягає у збагаченні Веб-контенту спеціалізованими мета-даними і створенні засобів для автоматичної обробки цих мета-даних (інтелектуальних програмних агентів), які могли б розпізнавати значення документів з метою виконання складних пошукових завдань користувачів
  - забезпечити "розуміння" інформації комп'ютерами, виділення ними найбільш придатних за тими чи іншими критеріями даних, і вже після цього - надання інформації користувачам
  - Інший напрямок: близький до галузі штучного інтелекту, і названий онтологічним підходом. Цей підхід включає в себе засоби анотування документів, якими могли б скористатися комп'ютерні програми - Web-сервіси та агенти при обробці складних запитів користувачів.
- 
-

# Технології, що використовуються

- глобальна схема імен – URI (Uniform Resource Identifier)
- стандартний синтаксис для опису даних – RDF (Resource Description Framework)
- стандартний засіб опису властивостей даних – RDF Schema
- стандартний засіб опису зв'язків між об'єктами даних – онтології, описані за допомогою OWL Web Ontology Language
- стандартна мова запитів до даних у форматі RDF – SPARQL



# Онтології

- "специфікація концептуалізації предметної області", або спрощено, документ або файл, що формально задає відносини між термінами
- словник понять предметної області і сукупність явно виражених припущень щодо змісту цих понять
- сукупність тверджень, які задають відносини між поняттями і визначають логічні правила для міркувань про них

## *Переваги:*

- можливість створення системи з елементами "штучного інтелекту" (дозволяє "виводити" нову інформацію з бази знань)
- Є "містком" для поєднання різних програмних засобів (описуючи формати даних, способи виклику, призначення), сприяючи вирішенню проблеми їх функціональної сумісності в рамках композитних додатків.

# Використання в наукових дослідженнях

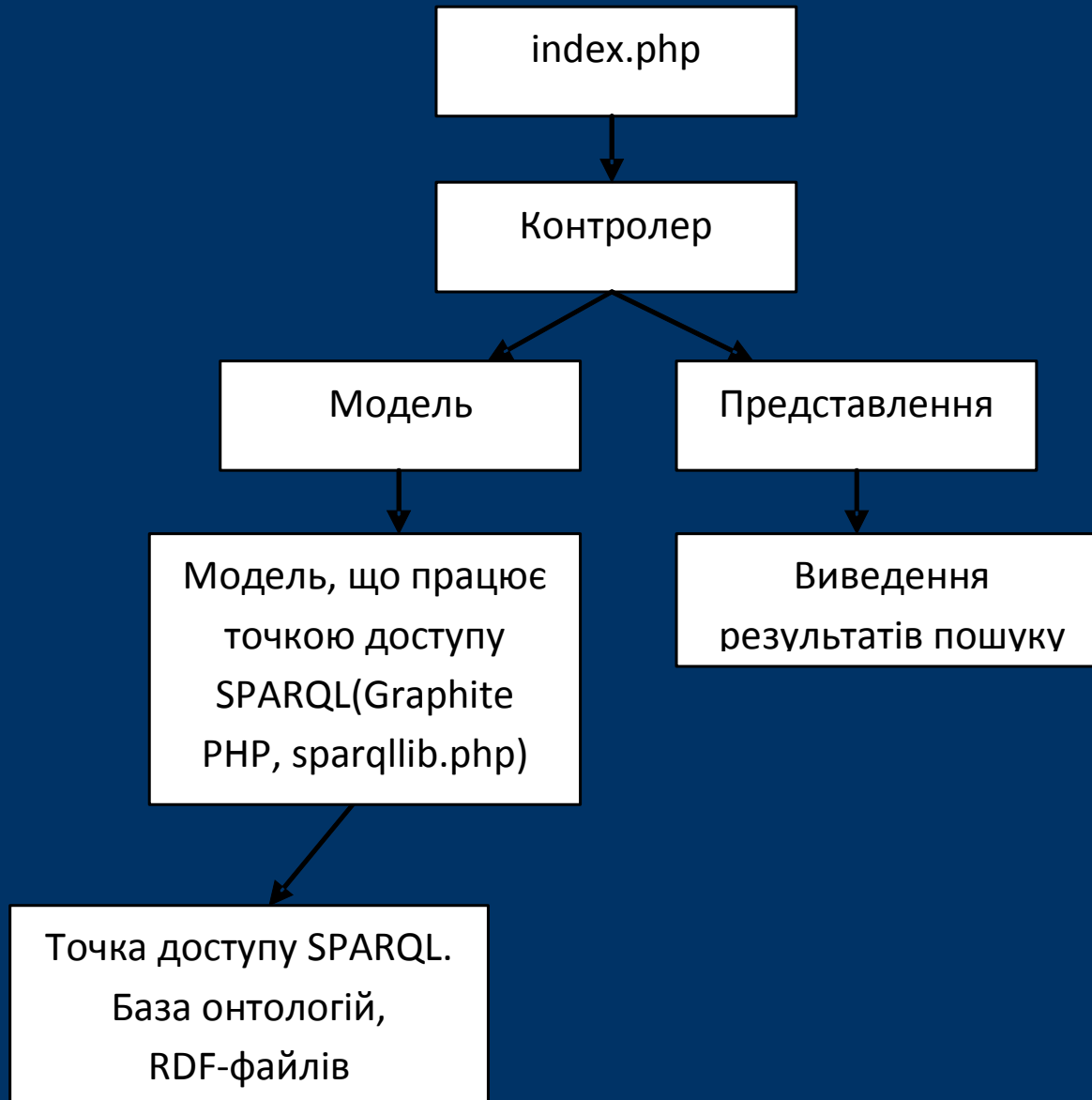
- База знань на основі наукових статей. Використання індексів цитування, які формують бібліометричну статистику.
  - Список використаної літератури, ключові слова, визначення термінів
  - Зв'язок між результатами досліджень, посилання на результати інших робіт, спростування результатів інших робіт
  - База знань термінів, законів, визначень, понять, зв'язків з іншими науками дозволять, наприклад, створювати такі цикли:
    - знання певної наукової діяльності зібрані і об'єднані в базу знань
    - дані перевіряються на цілісність, запити до них та автоматизовані міркування призводять до гіпотез
    - гіпотези можуть допомогти в розробці нових експериментів
    - експериментування генерує нові дані і, можливо, підтверджує або заперечує гіпотезу
    - дані, отримані після цього будуть інтегровані в базу знань, створюючи циклічний процес
  - Приклади: EngMath, MONET
- 
-

# Перспективи використання

- "проблема розмірності"
  - обробка даних відповідно до їх змісту, перехід за гіпер-посиланнями, що ведуть до визначень ключових термінів і правил логічних висновків.
  - "навчити" комп'ютерні програми, Web-служби і роботів пошукових систем і агентів "осмислено" оперувати інформацією, що вони знаходять
  - самі користувачі зможуть створювати сторінки Семантичного Web, давати власні визначення і вводити нові правила виведення, використовуючи стандартне для цієї мережі програмне забезпечення
  - єдиний стиль подання структурованої інформації, на відміну від стилю СУБД + інтерфейс, що дозволяє усі вище описані можливості керування такою інформацією в мережі, використовуючі готові веб-додатки, чи точки доступу до баз знань
  - компонування нових програм на основі набору існуючих семантично описаних бібліотек та модулів
- 
-



# Структура розробленого додатку



## Структура директорій:

- /cache
- /controllers
- /main
- /css
- /img
- /js
- /libs
- /arc
- /Graphite
- /model
- /view

```
Index::sparqlDbConnect("http://dbpedia.org/sparql");
Index::$DB->ns( "rdf","http://www.w3.org/1999/02/22-rdf-syntax-ns#" );
Index::$DB->ns( "dbpedia-owl","http://dbpedia.org/ontology/" );
Index::$DB->ns( "xsd","http://www.w3.org/2001/XMLSchema#" );

$result = Index::$DB->query('SELECT * WHERE
{
?companyURL rdf:type dbpedia-owl:Company          ;
dbpprop:companyName      ?corporation              ;
dbpedia-owl:foundedBy ?founderURL                  .
?founderURL dbpedia-owl:birthDate ?founderBirth    .
FILTER (?founderBirth > "1940-03-10"^^xsd:date )   .
{ ?founderURL rdf:type dbpedia-owl:Artist. }
  UNION
  { ?founderURL rdf:type dbpedia-owl:Actor. }
  UNION
  { ?founderURL rdf:type dbpedia-owl:MartialArtist. }
  UNION
  { ?founderURL dbpedia-owl:occupation dbpedia:Martial_arts.}
}
LIMIT 100
');
```

---

---

## Artists list:

Company url	Corporation	Founder URL	Founder Birth
<a href="http://dbpedia.org/resource/Basement_Comics">http://dbpedia.org/resource/Basement_Comics</a>	Basement Comics	<a href="http://dbpedia.org/resource/Budd_Root">http://dbpedia.org/resource/Budd_Root</a>	1958-01-01+02:00
<a href="http://dbpedia.org/resource/Mam_Tor_Publishing">http://dbpedia.org/resource/Mam_Tor_Publishing</a>	Mam Tor Publishing	<a href="http://dbpedia.org/resource/Liam_Sharp">http://dbpedia.org/resource/Liam_Sharp</a>	1968-05-02+02:00
<a href="http://dbpedia.org/resource/HandMade_Films">http://dbpedia.org/resource/HandMade_Films</a>	HandMade Films	<a href="http://dbpedia.org/resource/George_Harrison">http://dbpedia.org/resource/George_Harrison</a>	1943-02-25+02:00
<a href="http://dbpedia.org/resource/Eio_Books">http://dbpedia.org/resource/Eio_Books</a>	Eio Books	<a href="http://dbpedia.org/resource/Ki_Loingfellow">http://dbpedia.org/resource/Ki_Loingfellow</a>	1944-12-09+02:00
<a href="http://dbpedia.org/resource/Pop!_Goes_the_Icon">http://dbpedia.org/resource/Pop!_Goes_the_Icon</a>	Pop! Goes the Icon	<a href="http://dbpedia.org/resource/Pj_Perez">http://dbpedia.org/resource/Pj_Perez</a>	1976-08-04+02:00
<a href="http://dbpedia.org/resource/Pottermore">http://dbpedia.org/resource/Pottermore</a>	Pottermore	<a href="http://dbpedia.org/resource/J._K._Rowling">http://dbpedia.org/resource/J._K._Rowling</a>	1965-07-31+02:00
<a href="http://dbpedia.org/resource/Artists_Music_Guild">http://dbpedia.org/resource/Artists_Music_Guild</a>	Artists Music Guild, Inc.	<a href="http://dbpedia.org/resource/David_L._Cook">http://dbpedia.org/resource/David_L._Cook</a>	1968-11-11+02:00
<a href="http://dbpedia.org/resource/Broadway_Comics">http://dbpedia.org/resource/Broadway_Comics</a>	Broadway Comics	<a href="http://dbpedia.org/resource/Jim_Shooter">http://dbpedia.org/resource/Jim_Shooter</a>	1951-09-27+02:00
<a href="http://dbpedia.org/resource/Defiant_Comics">http://dbpedia.org/resource/Defiant_Comics</a>	Defiant Comics	<a href="http://dbpedia.org/resource/Jim_Shooter">http://dbpedia.org/resource/Jim_Shooter</a>	1951-09-27+02:00
<a href="http://dbpedia.org/resource/HGGC">http://dbpedia.org/resource/HGGC</a>	HGGC	<a href="http://dbpedia.org/resource/Greg_Benson">http://dbpedia.org/resource/Greg_Benson</a>	1968-01-23+02:00

# Висновки

- Можна виділити наступні переваги застосування семантичних композитних веб-додатків у наукових дослідженнях:
    - Системи пошуку наукових даних “за смислом”, а не за ключовими поняттями. “Розумний пошук” по базах експериментальних даних, базах публікацій, термінів понять, законів, теорем і т.д.
    - Можливість побудови автоматизованого циклічного процесу здобуття нових знань з існуючих в наукових дослідженнях.
    - Спрощення інтеграції наукових мережевих додатків та баз знань в розподіленому Веб-просторі. Створення спеціалізованих композитних додатків з існуючих, семантично описаних мережевих компонентів
  - Досліджені технології були використані для створення тестового семантичного веб-додатку (PHP), що використовує множину точок доступу SPARQL для отримання інформації з RDF-баз знань. Даний підхід продемонстрував більшу гнучкість, аніж традиційний підхід “СУБД + графічний інтерфейс”, дозволяючи переносити частину логіки до бази знань.
  - Ці напрацювання можуть бути використані при подальшій розробці спеціалізованих композитних наукових додатків, а також в навчальному процесі.
- 
-

Дякую за увагу!

