

ВІДГУК

офіційного опонента на дисертаційну роботу **Кушнірецької Ірини Ігорівни** на тему "**Інформаційна технологія динамічної інтеграції слабоструктурованих даних у web-системах**", подану на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології.

Актуальність роботи

Під час виконання пошуку потрібної інформації у Всесвітній мережі, де міститься і постійно зростає множина різнорідних даних виникає потреба у застосуванні таких комунікаційних технологій, які б надали можливість здійснити цей процес зручно і ефективно. На сьогодні одним із рішень даного завдання є використання систем динамічної інтеграції даних. Такі системи дозволяють отримати одночасно дані із декількох джерел. В сучасних умовах набувають великої популярності системи динамічної інтеграції даних, що використовують технологію Mashup. Mashup – це такий підхід до інтегрування, що дозволяє об'єднувати дані із декількох різнорідних джерел в один інтегрований інструмент. Це нова технологія, яка порівняно недавно почала розвиватися і активно набуває широкого використання. Системи, створені на основі даної технології відкривають великі можливості використання web-ресурсів. Але поки не має загального стандартизованого підходу до проектування систем на основі Mashup-технології та процесів опрацювання інформаційних ресурсів, що призводить до виникнення ряду проблем при розробленні типової архітектури відповідних систем.

Джерелами даних для Mashup-застосунків можуть бути різного роду системи сучасного Web 2.0, зокрема особливої популярності останнім часом для джерел інформаційних ресурсів стали набувати соціальні мережі, web-форуми, різноманітні web-сайти новин, тощо. Зазвичай дані таких джерел знаходяться у слабоструктурованому вигляді, що може ускладнити процес пошуку інформації у таких web-системах із подальшою їх інтеграцією. Сучасні методи і засоби побудови Mashup-застосунків використовують досить прості способи організації інтеграції даних, наприклад, на основі переміщення деякого сегменту даних однієї web-системи до іншої (при вході в систему через соціальну мережу), не зосереджуючись на семантиці інтегрованих даних. Тому виникає потреба в застосуванні таких методів і технологій для організації динамічної інтеграції слабоструктурованих даних, які б дали змогу враховувати семантику даних на всіх етапах даного процесу.

Приймаючи до уваги вищенаведене, можна стверджувати, що тема дисертаційної роботи Кушнірецької І.І., яка присвячена створенню та вдосконаленню методів і засобів інформаційної технології динамічної

інтеграції слабоструктурованих даних, що дає змогу підвищити якість результатів пошуку даних у web-системах, є актуальною.

Основні наукові результати дисертаційної роботи, отримані особисто автором:

1. Уперше розроблено метод опису структури і змісту вхідних даних Mashup системи, який відрізняється від відомих застосувань онтологій для класифікації інформаційних ресурсів, що дало змогу виконувати процеси інтеграції даних із врахуванням їх змісту.

2. Уперше розроблено метод формування об'єданого динамічного набору даних, що відрізняється від існуючих застосувань процедур лінгвістичного аналізу змісту запиту користувача, що забезпечило підвищення релевантності та узгодженості результатів динамічної інтеграції слабоструктурованих даних.

3. Удосконалено структурну модель системи динамічної інтеграції слабоструктурованих даних на основі технології Mashup шляхом додання функціональних компонентів визначення та опису семантики вхідних та вихідних наборів даних, що дало змогу підвищити якість результатів застосування таких систем.

4. Отримала подальший розвиток Mashup технологія динамічної інтеграції даних через запровадження додаткових можливостей опрацювання семантики інформаційних ресурсів, що дало змогу підвищити якість результатів інформаційно-пошукових web-систем.

Зв'язок роботи із науковими програмами, темами.

Дисертаційна робота тісно пов'язана з планами науково-дослідної роботи Національного університету «Львівська політехніка». Результати досліджень, що відображені у дисертаційній роботі, отримані у рамках виконання науково-дослідної роботи кафедри інформаційних систем та мереж Національного університету «Львівська політехніка» за темою «Розроблення інтелектуальних розподілених систем на основі онтологічного підходу з метою інтеграції інформаційних ресурсів» (№ держреєстрації 0115U004228).

Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації, їх достовірність.

Підходи та методи, що запропоновані у дисертаційній роботі для створення нових та вдосконалення існуючих методів і засобів інформаційної технології динамічної інтеграції слабоструктурованих даних у web-системах, науково і технічно обґрунтовані. Обґрунтованість результатів дисертаційної роботи полягає у коректному використанні методу опису структури і змісту вхідних даних Mashup-системи, методу формування об'єданого динамічного набору даних та структурній моделі системи динамічної інтеграції слабоструктурованих даних на основі технології Mashup. Припущення, які

покладені в основу теоретичних досліджень, не суперечать особливостям предметної області, а внаслідок дослідження отримано можливість підвищити якість результатів пошуку даних у web-системах.

Достовірність наукових положень, висновків і рекомендацій, що сформульовані в дисертаційній роботі підтверджуються низкою експериментальних досліджень, спостережень, апробацій результатів та впроваджень.

Практичне значення та ефективність результатів дисертаційних досліджень.

Розроблено рекомендації проектування структури Mashup-системи динамічної інтеграції слабоструктурованих даних через використання підсистем динамічного семантичного опрацювання інформаційних ресурсів, що дає змогу покращити якість результатів роботи Mashup-системи.

Розроблено та впроваджено програмні засоби, які дають можливість підвищити якість даних, отриманих у результаті динамічної інтеграції, зокрема: збільшення точності до 6 %, зменшення інформаційного шуму до 6%, покращення узгодженості результатів інформаційних даних до 10% у системі, що працює за технологією Mashup.

Практичне значення дисертаційних досліджень підтверджується розробленою структурою та програмною реалізацією в проектах: News Mashup from Social Networks (система моніторингу новин із соціальних мереж), Shares and Discounts Mashup (система пошуку знижки на купівлю товару, отримання послуги тощо); системах аналізу та опрацювання даних, розроблених для КЗ ЛОР «Львівське обласне патологоанатомічне бюро» та стоматологічної клініки «Улюблений доктор», що підтверджено відповідними актами впровадження.

Результати також впроваджено в навчальному процесі НУ «Львівська політехніка» під час викладання дисципліни «Об'єктно-орієнтоване програмування».

Повнота викладення результатів у опублікованих працях, апробація роботи.

Аналіз основних наукових результатів Кушнірецької І.І засвідчує особистий внесок автора щодо розроблення методів і засобів інформаційної технології динамічної інтеграції слабоструктурованих даних у web-системах для підвищення якості результатів пошуку даних у web-системах.

Основні положення та результати дисертаційного дослідження викладено в 17 наукових працях, серед них 2 статті опубліковані у наукових періодичних виданнях інших держав, 1 стаття у виданні України, яке входить до міжнародних наукометричних баз, 6 статей у виданнях, що входять до переліку фахових видань України та 8 тез у матеріалах міжнародних науково-технічних конференцій. Опубліковані роботи у повній мірі охоплюють основні результати дисертаційного дослідження.

Оцінка основного змісту дисертації та її структури.

Дисертаційна робота є завершеною науковою працею. Вона складається з вступу, чотирьох розділів, висновків, списку використаних джерел із 160 найменувань та додатків. Оформлення дисертації відповідає основним вимогам щодо оформлення дисертаційних робіт.

У **вступі** проведено обґрунтовано актуальності теми дослідження, сформульовано мету й задачі дослідження, визначена наукова новизна та практична цінність отриманих результатів.

У **першому** розділі дисертаційного дослідження проведено аналіз існуючих інформаційних технологій та систем динамічної інтеграції даних, який показав, що у даний час набирають значної популярності системи динамічної інтеграції даних на основі технології Mashup. Ця технологія з'явилася порівняно недавно, але надає великі перспективи для використання різнорідних даних web-систем. Проте натепер відсутні стандартизовані технології побудови і опрацювання слабоструктурованих даних такого роду систем. За результатами проведеного аналізу визначено завдання дослідження, які потрібно виконати для розв'язання нової задачі, що сформульована у дисертаційній роботі.

Другий розділ дисертаційної роботи присвячений визначенню загальної динамічної структури для множини вхідних інформаційних ресурсів. На основі аналізу характеристик сучасних Mashup-систем здійснено концептуалізацію та моделювання роботи системи динамічної інтеграції слабоструктурованих даних у web-системах. Сформульовано принципи визначення вхідних даних та формування вихідних даних у Mashup-системі та встановлено, що застосування процедури структурування вхідних даних із врахуванням їх змісту може підвищити якісні показники результату процесу визначення вхідних даних, а через формування семантично узгодженого набору вихідних даних можна підвищити якісні показники результатів процесу формування вихідних даних. Автором вперше розроблено метод опису структури і змісту вхідних даних Mashup системи через застосування онтологій для класифікації інформаційних ресурсів, що дало змогу виконувати процеси інтеграції даних із врахуванням їх змісту.

У **третьому** розділі визначено принципи формування об'єднаного динамічного набору даних, що має узгоджену структуру і єдиний зміст. Описано підхід до проведення аналізу вхідних інформаційних ресурсів з метою визначення їх спільних рис та виявлення зв'язків між ними. Наведено процедуру визначення подібності схем даних для систем, що інтегруються та процедуру визначення подібності інформаційних ресурсів систем, що інтегруються за допомогою використання комплексу логістичних оцінок порівняння текстових рядків. Розроблено алгоритми створення загальної динамічної структури для множини вхідних інформаційних ресурсів із врахуванням їх змісту, що дає змогу застосувати їх для процесу визначення структури і змісту вхідних наборів даних інтегрованих web-систем. Розроблено алгоритм побудови онтологічної моделі всіх систем, що інтегруються та алгоритм отримання інформаційних ресурсів із системи, що

інтегрується, використовуючи семантично орієнтовані технології, зокрема онтології і дескрипторну логіку. Розроблено метод формування об'єданого динамічного набору даних, що відрізняється від існуючих застосуванням процедур лінгвістичного аналізу змісту запиту користувача, що забезпечило підвищення релевантності та узгодженості результатів динамічної інтеграції слабоструктурованих даних.

У **четвертому** розділі представлено організаційні принципи побудови Mashup-системи, описано розроблення інструментальних засобів динамічної інтеграції слабоструктурованих даних та наведено приклади програмної реалізації розроблених методів. Структурна модель системи динамічної інтеграції слабоструктурованих даних володіє перевагами модульного підходу та включає: адаптери для роботи з інформаційними джерелами, модуль інформації про систему, модуль даних про інформаційні джерела, підсистему формування завдання, модуль отримання доступу до інформаційних джерел, підсистему опису структури і змісту вхідних даних, підсистему формування об'єданого динамічного набору вихідних даних, модуль зберігання даних, модуль підтримки роботи системи, підсистему візуалізації результатів роботи системи. Робота програмної системи, що створена на основі інформаційної технології динамічної інтеграції слабоструктурованих даних у web-системах, полягає у представленні об'єданого узгодженого динамічного набору даних, отриманого із різних web-систем відповідно до запиту користувача. Було проведено низку експериментів тестування роботи систем із інформаційною технологією динамічної інтеграції слабоструктурованих даних у web-системах та встановлено, що застосування розробленої ІТ дає змогу покращити якість результатів пошуку даних у web-системах до 10%.

Висновки по роботі сформульовані чітко та повністю висвітлюють отримані результати.

Список літератури у повному обсязі охоплює предметну область, є достатньо інформативним та відображає опрацювання автором потрібної кількості літературних джерел.

Відповідність дисертації та автореферату встановленим вимогам.

Дисертація та автореферат дисертації за своєю структурою, об'ємом і оформленням відповідають вимогам, які встановлені до кандидатських дисертацій, зокрема пп. 9, 11, 12 «Порядку присудження наукових ступенів» (Постанова № 567, від 24 липня 2013 р.).

Автореферат дисертації за змістом відповідає основним положенням, які викладено в дисертації та не містить інформації, яка не відображена в основній роботі. Стиль викладення матеріалів досліджень та наукових положень і рекомендацій забезпечує їх адекватне і належне сприйняття. Наукова новизна відповідає паспорту спеціальності 05.13.06 – інформаційні технології, зокрема за напрямками досліджень «Розроблення інформаційно-пошукових і експертних систем обробки інформації для прийняття рішень, а також знання-орієнтованих систем підтримки рішень в умовах ризику та

невизначеності як інтелектуальних інформаційних технологій» та «Дослідження, розроблення та впровадження Інтернет-технологій для побудови сервіс-орієнтованих систем, а також для організації та реалізації систем розподіленої обробки інформації».

Матеріали дисертації викладені у чіткій логічній послідовності, на належному науковому рівні та повністю відповідають поставленій меті і задачам дослідження. Висновки роботи відповідають її змісту, обґрунтовані і підтверджені результатами дослідження.

Недоліки та зауваження до роботи:

1. У першому розділі пункту 1.1.4. описано стандарти для систем інтеграції даних, але стандарт, який використано для отриманих у роботі результатів не вказано.

2. Рисунок 2.1 - Діаграма переходів станів Mashup-системи та рисунок 2.2 - UML-діаграма діяльності Mashup-системи містять тривіальні результати, причому не зрозуміло, як вони далі у роботі застосовуються.

3. Запропонований формальний опис Mashup-системи динамічної інтеграції даних на основі теорії множин недостатньо використовується при описі кінцевих алгоритмів та засобів, що реалізують запропоновану інформаційну технологію.

4. Для розробленого методу опису структури і змісту вхідних даних було детально описано всі кроки методу та вирішення завдань кожного із них, але недостатньо уваги приділено оцінюванню повноти та адекватності розробленого методу, не доведено його суттєвих переваг над іншими аналогічними методами опису вхідних даних.

5. У опису розроблених алгоритмів створення загальної динамічної структури для множини вхідних інформаційних ресурсів із врахуванням їх змісту замість блок-схем алгоритмів наведено тільки діаграми діяльності, що зменшує розуміння логіки обробки даних.

6. Запропонований у 3.4 метод формування об'єданого динамічного набору даних зведено до 4-х кроків з відповідним обґрунтуванням кожного із них. Проте у описі недостатньо уваги приділено оцінюванню отриманих результатів методу.

7. Рисунок 4.3, який зображає структурну модель системи динамічної інтеграції слабоструктурованих даних, що працює на основі Mashup, має заплутану структуру, його було б доцільно винести у додатки з більшим масштабуванням.

8. Наведені у додатках акти впровадження занадто лаконічні – мають декларативний характер і не містять даних, що обґрунтовують кількісні показники отриманих переваг від впровадження.

9. Робота не вільна від редакційних та стилістичних недоліків, зустрічаються, на жаль, описки та русизми.

Вказані зауваження не знижують загальної позитивної оцінки дисертаційної роботи та можуть розглядатись як рекомендації до подальших

наукових досліджень та впровадження отриманих результатів в практику розробки систем динамічної інтеграції слабоструктурованих даних.

Загальні висновки

Дисертаційна робота Кушнірецької Ірини Ігорівни на тему «Інформаційна технологія динамічної інтеграції слабоструктурованих даних у web-системах» є завершеною науковою працею, яка розв'язує актуальне наукове завдання розроблення методів і засобів інформаційної технології динамічної інтеграції слабоструктурованих даних у web-системах для підвищення якості результатів пошуку даних у web-системах. Робота відповідає паспорту спеціальності 05.13.06 – інформаційні технології.

За змістом, оформленням і науково-практичними результатами дисертаційна робота задовольняє вимогам щодо дисертацій на здобуття наукового ступеня кандидата технічних наук відповідно до пунктів 9, 11, 12 «Порядку присудження наукових ступенів», а її автор, Кушнірецька Ірина Ігорівна, заслуговує на присвоєння їй наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології.

Офіційний опонент
декан факультету комп'ютерних
систем і автоматики Вінницького
національного технічного
університету МОН України
доктор технічних наук, професор



О.В. Бісікало

