УДК 004.652
DOI: 10.15587/2312-8372.2014.31412

Кушнірецька І. І.,
Кушнірецька О. І.,
Берко А. Ю.

ВИЗНАЧЕННЯ СТРУКТУРИ І ЗМІСТУ ВХІДНИХ ІНФОРМАЦІЙНИХ РЕСУРСІВ ДЛЯ РОБОТИ МЕШАП-СИСТЕМИ

У статті представлено характеристику основних кроків визначення структури і змісту вхідних інформаційних ресурсів для роботи Мешап-системи. Наведено загальну схему діяльності Мешап-системи і описано кожен стан діяльності системи. Запропоновано можливі рішення завдань, поставлених згідно виділених кроків визначення структури і змісту вхідних інформаційних ресурсів для роботи Мешап-системи.

Ключові слова: *Мешап-система, інформаційний ресурс, стани діяльності Мешап-системи, онтології, формалізм обмежень.*

1. Вступ

На сьогодні у Всесвітній мережевій павутині міститься величезна кількість інформації різноманітного характеру і змісту, яка до того ж невпинно зростає. Ситуація різкого зростання обсягів інформаційних потоків породила ряд проблем: непропорційне зростання інформаційного шуму через слабку структурованість інформації; багаторазове дублювання інформації; поява паразитної інформації та ін. Тому потреба швидкого і якісного отримання потрібних даних стає все більш актуальною. Охоплення, узагальнення великих динамічних інформаційних масивів, що безперервно генеруються в Мережі, вимагає якісних нових підходів. Одним з рішень даної проблеми є розробка систем інтеграції даних, що працюють використовуючи технологію мешап (англ. Mash-Up). Мешап — це такий підхід для розробки додатків, який дозволяє користувачам об'єднувати дані з декількох джерел в один інтегрований інструмент [1].

Сьогодні Мешап-системи відкривають нові і широкі можливості використання інформаційних ресурсів. Тим не менш, незмінним залишається той факт, що для розроблення Мешап-системи потрібно володіти достатньо великим запасом знань методів та технологій, щоб створити справді корисну і якісну систему. І одним з найперших завдань на шляху побудови Мешапу є визначення структури і змісту вхідних інформаційних ресурсів.

2. Аналіз літературних даних і постановка проблеми

Дослідження показують, що існує високий рівень інтересу в рамках роботи Мешап-систем [1–5]. А при стрімкому розвиткові IT-технологій особливої актуальності набувають потреби в динамічній інтеграції різнорідних даних у web-серидовищі.

Архітектура Мешап-системи завжди складається з трьох частин [1]:

1. Провайдер вмісту — це джерело даних. Дані доступні через API і різні веб-протоколи, такі як RSS, REST і веб-сервіси.

Провайдер вмісту є однією з найбільш важливих частин системи, адже саме він визначає, як і яким чином будуть виходити дані з сторонніх сервісів. Вважається, що Мешапи повинні використовувати тільки ті сервіси, які надають для цього спеціальні API.

2. Мешап-сайт (рівень процесів) — це веб-додаток, що пропонує новий сервіс, який використовує джерела даних, що йому не належать.

3. Браузер клієнта (презентаційний рівень) — власне користувацький інтерфейс Мешапу. У веб-додатках вміст може бути «замешаплено» клієнтським браузером з використанням клієнтської мови програмування, наприклад JavaScript.

При створенні Мешапу найбільш хвилюючими проблемами можуть бути наступні:

— чи це не надто складно?
— чи буде ефективно?
— чи варто?

Обираючи інструменти для Мешапу потрібно враховувати інтеграцію на рівні даних (семантична інтеграція) і на сервісному рівні. При виборі Мешап-інструментів мотивацією є:

— інтеграційні інструменти для досягнення кращої ефективності;
— наявним інструментам для побудови колажів даних не вистачає функціональності.

Джерела даних можуть володіти різними властивостями, істотними для вибору методів інтеграції даних — вони можуть підтримувати представлення даних в термінах тієї чи іншої моделі даних, можуть бути статичними або динамічними і т. п. Безліч джерел інтеграції даних можуть бути однорідними або неоднорідними щодо характеристик, відповідних використовуваному рівню інтеграції [5]. Для використання Мешапом кожне джерело даних повинне бути найперше проаналізоване і змодельоване для того, щоб виконати необхідні дії пошуку та попередньої обробки [5].

На даний час вже розроблено кілька методів та засобів для вирішення проблеми визначення структури і змісту вхідних інформаційних ресурсів та автоматичного вилучення структурованої інформації з веб-сайтів і додатків, які, в силу специфіки характеру HTML мови, як правило, надають неструктуровану інформацію.

 © Кушнірецька І. І., Кушнірецька О. І., Берко А. Ю.

Ці методи, як правило, розроблені на основі певних правил видобутку даних, які можуть бути менш або більш витонченими, і визначаються або в напівавтоматичному режимі за допомогою демонстрації, або автоматично за допомогою машино-навчальних алгоритмів або інших методів штучного інтелекту. Огляд декількох засобів видобування веб-даних представлено в роботі [5].

Хоча ці підходи істотно вирішують проблему отримання даних з веб-додатків для цілей автоматичної і інтегрованої навігації по веб-сторінках або для міграції в напрямку Semantic Web, інші автори представляють свій досвід того, щоб семантика вхідних даних була виражена і збережена разом з цими даними. Ці рішення пропонують вирішувати за допомогою онтологій [6].

Онтологічні системи будують на основі таких принципів [6]:

— формалізації, тобто опису об'єктивних елементів реальності із застосуванням єдиних, суворо визначених зразків (термінів, моделей тощо);

— використання обмеженої кількості базових термінів (сутностей), на основі яких конструюють всі інші поняття;

— внутрішньої повноти;

— логічної несуперечності.

Для ефективного використання онтологій необхідна наявність інструментів, що дозволяють:

— створювати і модифікувати онтології;

— імпортувати онтології з інших джерел;

— об'єднувати онтології різного походження;

— автоматично створювати онтології для джерела даних з використанням метаданих цього джерела;

— здійснювати різні інші операції над онтологіями;

— підтримувати анотування онтології;

— підтримувати сумісність версій;

— пов'язувати концепти онтології з різними типами даних;

— забезпечувати багатомовну підтримку.

У загальному випадку формальне зображення онтології даних подають наступним чином:

$$O = \langle X, R, F \rangle, \quad (1)$$

де X — скінченна множина понять (класів, концептів) предметної області з їх властивостями (атрибутами); R — скінченна множина відношень (зв'язків, відповідностей) між поняттями; F — скінченна множина функцій інтерпретації (обмежень, аксіом) [7].

Згідно з вимогами стандарту IDEF5 [7], концепти поділяють на класи та значення класів. Зв'язки між концептами поділяють на класифікаційні — між класами і підкласами і структурні, які описують взаємодію класів.

Зазначимо, також, що використання онтологій не тільки дозволяє створювати моделі даних, адекватні реальному світу, а й відповідає загальному напрямку робіт у галузі стандартизації World Wide Web в рамках проекту семантичного Web, що дозволяє розглядати проблеми інтеграції даних та інтелектуалізації WWW з єдиної точки зору. Розроблені для цих цілей уніфікована модель даних RDF (Resource Decription Framework) і мова веб-онтологій OWL (Web Ontology Language) надають багаті можливості семантичного опису розподілених Інтернет ресурсів.

Отже, для того, щоб інтеграція даних у Мешап-системі відбувалася динамічно та при цьому можна було б зберегти семантику даних потрібно, насамперед, зрозуміти основні принципи роботи самого Мешапу та, безсумнівно, його роботу із вхідними потоками інформації. Звідси виникає проблема, що полягає у визначенні структури і змісту вхідних інформаційних ресурсів для їх подальшого використання.

3. Мета та задачі дослідження

Мета роботи полягає у дослідженні застосування якісно нових підходів подолання семантичних лабіринтів вхідної інформації, одним з яких може бути вдосконалення етапу визначення структури і змісту вхідних інформаційних ресурсів для роботи Мешап-системи.

Для досягнення поставленої мети вирішувалися наступні задачі:

— дослідження та загальна формалізація роботи Мешап-систем;

— на основі виділення станів діяльності Мешап-системи формулювання основних кроків та завдань для визначення структури і змісту вхідних інформаційних ресурсів;

— надання можливих рішень для вирішення завдань, поставлених згідно сформульованим крокам визначення структури і змісту вхідних інформаційних ресурсів для роботи Мешап-системи.

4. Матеріали та методи досліджень загальної роботи Мешап-систем та розроблення принципів визначення структури і змісту вхідних інформаційних ресурсів

Для того, щоб зрозуміти основну суть роботи Мешапу і яким чином дана система працює з вхідними та вихідними даними, корисним буде використання об'єктно-орієнтованої технології моделювання загальної роботи системи, такої як UML-діаграма діяльності.

Діаграма діяльності (Activity diagram) — діаграма, на якій показане розкладання деякої діяльності на її складові частини. Під діяльністю (activity) розуміється специфікація поводження, що виконується, у вигляді координованого послідовного й паралельного виконання підлеглих елементів — вкладених видів діяльності й окремих дій (action), з'єднаних між собою потоками, які йдуть від виходів одного вузла до входів іншого [8].

Саме діаграми діяльності дозволяють реалізувати в UML особливості процедурного і синхронного управління, обумовленого завершенням внутрішніх діяльностей і дій. Метамодель UML надає для цього необхідні терміни та семантику. Основним напрямком використання діаграм діяльності є візуалізація особливостей реалізації операцій класів, коли необхідно представити алгоритми їх виконання. При цьому кожний стан може бути виконанням операції деякого класу або її частини, дозволяючи використовувати діаграми діяльності для опису реакцій на внутрішні події системи [9].

У контексті мови UML діяльність (activity) являє собою деяку сукупність окремих обчислень, виконуваних автоматом. При цьому окремі елементарні обчислення можуть приводити до деякого результату або дії (action).

На діаграмі діяльності відображається логіка або послідовність переходу від однієї діяльності до іншої, при цьому увага фіксується на результаті діяльності. Сам же результат може призвести до зміни стану системи або поверненню деякого значення [9].

Очевидно, що при використанні гетерогенних даних, які до того ж мають різну структуру або взагалі неструктуровані, які часто описують одну й ту ж проблемну область з використанням різних термінів і понять, для успішного вирішення завдання визначення структури і змісту отриманих вхідних інформаційних ресурсів їх семантика повинна бути явним чином виражена і збережена разом з цими даними. У зв'язку з цим актуально при вирішенні даної проблеми скористатися семантично орієнтованими технологіями, такими як онтології і дескриптова логіка [10].

Онтологія даних — це засіб різнобічної і детальної формалізації знань про дані за допомогою концептуальної схеми. Як правило, до складу такої схеми входить опис структури даних, що містить визначення всіх релевантних класів об'єктів, їхні взаємозв'язки та правила (теореми, обмеження), задані у предметній області набору даних [10]. Онтології даних можна описувати різними засобами, і сьогодні поширені багато мов опису онтологій. Проте, з огляду на те, що в будь-якій онтології визначають терміни і задаються логічні зв'язки між ними, точна семантика опису термінів і зв'язків в різних мовах буде однаковою.

5. Результати досліджень діяльності Мешап-систем та формування принципів визначення структури і змісту вхідних інформаційних ресурсів для роботи Мешап-системи

Визначимо основні стани діяльності Мешап-систем за допомогою використання UML-діаграми діяльності (рис. 1).

Аналізуючи діяльність Мешап-систем можна виділити наступні стани діяльності.

Перший стан — реєстрація. Якщо реєстрація пройшла успішно, переходимо до другого стану, якщо ж ні — повертаємося знову до початку реєстрації.

Другий стан — авторизація, якщо все пройшло успішно, рухаємося далі, якщо ні — повертаємося до початку авторизації.

Третій стан — формування завдання. Якщо завдання сформовано відповідно до правил системи, рухаємося далі, якщо ні — повертаємося до початку третього стану.

Четвертий стан — формування джерел для Мешапу.

П'ятий стан — пошук потрібної інформації у відібраних джерелах. Якщо результати пошуку задовільні — йдемо далі, якщо ні — повертаємося назад до пошуку.

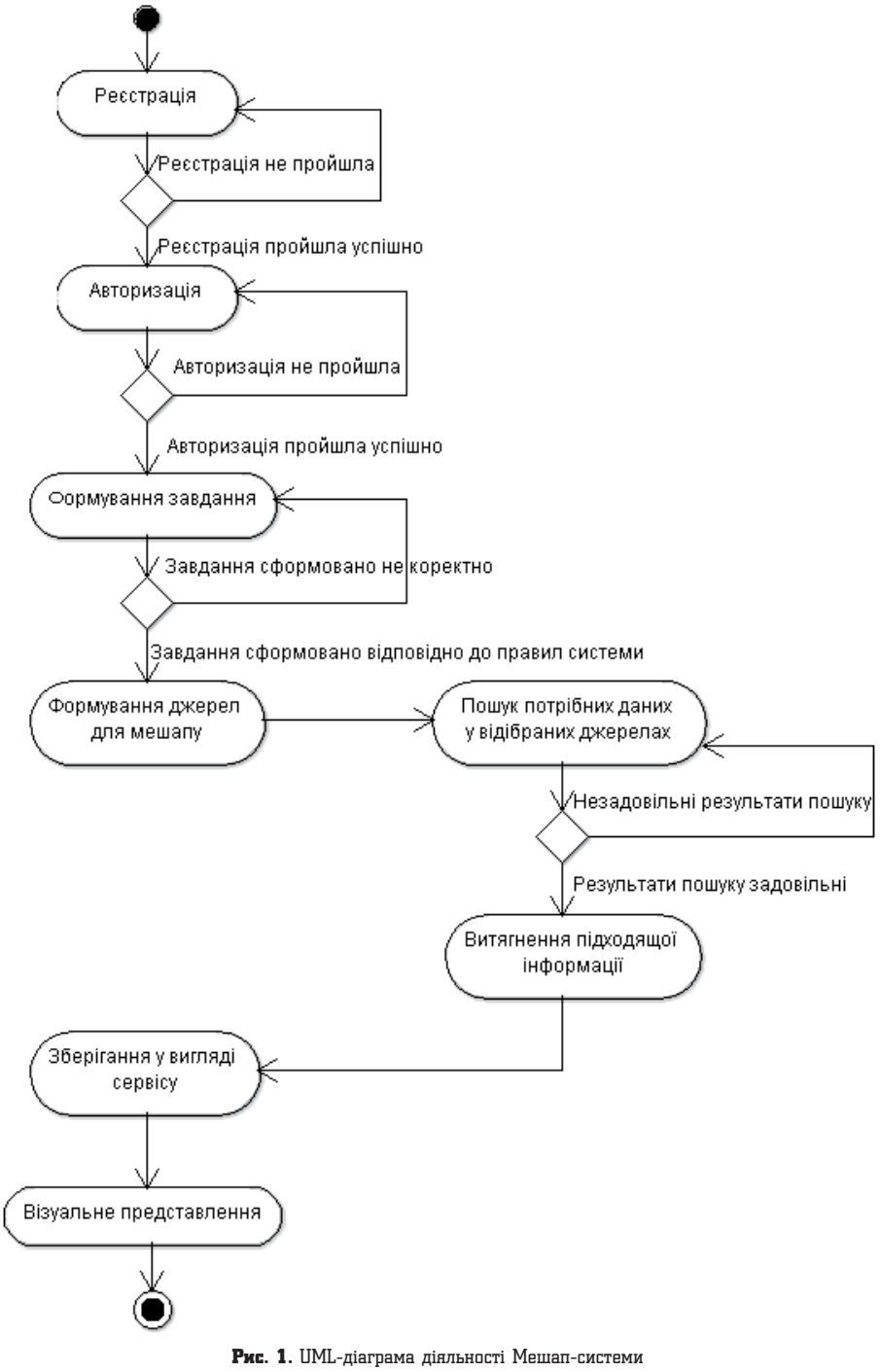

Рис. 1. UML-діаграма діяльності Мешап-системи

Шостий стан — витягнення підходящої інформації і перехід до наступного стану.

Сьомий стан — зберігання отриманої інформації у вигляді сервісу.

Восьмий стан — візуальне представлення готового Мешапу.

Отже, визначено 8 основних станів, які, зазвичай, наявні у будь-якій Мешап-системі. Розглядаючи четвертий стан має місце вирішення проблеми ідентифікації інформаційних ресурсів або деякого інформаційного об'єкту (далі — об'єкту), яка може включати в себе кілька етапів, серед яких можна виділити основні:

— завдання умови виділення деякого об'єкту серед всієї сукупності об'єктів (визначення пошукового образу);

— процес розпізнавання об'єкту — перевірка задоволення умовами пошукового образу характеристик об'єкту сканування.

Найважливішими станами при роботі Мешап-системи є пошук потрібної інформації (п'ятий стан), витягнення знайденої інформації (шостий стан) та її зберігання у вигляді сервісу (сьомий стан). При роботі з інформаційними ресурсами на цих станах досить часто може виникнути ряд проблем, які можуть вплинути ненайкращим чином на результати роботи системи. Від коректної роботи системи на даних трьох станах залежить безпосередньо кінцевий результат системи. Звичайно, що кожен наступний стан залежить від попереднього, звідси — якщо на п'ятому стані знайдено потрібну інформацію, на шостому стані її витягнуто з джерел, на сьомому стані ми повинні зберегти отриману інформацію. І тут виникає потреба, що полягає у визначенні структури і змісту отриманої вхідної інформації.

Отже, розглянемо детальніше ту роботу на сьомому стані, що стосується визначення структури та змісту отриманої вхідної інформації. І для цього сформуємо основні кроки для визначення структури і змісту вхідних інформаційних ресурсів та опишемо завдання, які потрібно вирішити на кожному з них:

Перший крок. Визначення форми представлення вхідних даних: структуровані, слабоструктуровані чи неструктуровані дані.

Другий крок. Класифікація вхідної інформації відповідно до предметної області із виділенням і збереженням семантики даних.

Третій крок. Виділення у вхідному інформаційному ресурсі набору реквізитів (атрибутів), що відображають його основні характеристики і аспекти зазначеної предметної області.

Четвертий крок. Встановлення чітких границь та задоволення основних обмежень, що стосуються вхідного потоку web-інформації.

П'ятий крок. Формування моделі вхідного інформаційного ресурсу загальної структури, визначеної предметною областю.

Отже, автори даної статті описали роботу Мешап-системи на сьомому стані діяльності, що відповідає за визначення структури і змісту отриманої вхідної інформації. Для цього ними було сформовано 5 основних кроків для визначення структури і змісту вхідних інформаційних ресурсів.

6. Обговорення результатів досліджень формування принципів визначення структури і змісту вхідних інформаційних ресурсів для роботи Мешап-системи

Проаналізуємо детально кожен із сформованих основних кроків визначення структури і змісту вхідних інформаційних ресурсів для роботи Мешап-системи.

Згідно першого кроку — першочерговим завданням роботи Мешап-системи для визначення структури і змісту вхідних інформаційних ресурсів є визначення форми представлення вхідних даних: структуровані, слабоструктуровані чи неструктуровані дані.

Відповідно до другого кроку потрібно використовувати такі методи і технології класифікації вхідної інформації відповідно до предметної області, щоб можна було виділити і зберегти семантику даних. Відповідно до даного принципу пропонуємо скористатися семантично орієнтованими технологіями, такими як онтології.

Аналізуючи третій крок — виділення у вхідному інформаційному ресурсі набору реквізитів (атрибутів), що відображають його основні характеристики і аспекти зазначеної предметної області, пропонуємо здійснювати за допомогою онтологій.

Щодо четвертого кроку — встановлення чітких границь та задоволення основних обмежень, що стосуються вхідного потоку web-інформації, пропонуємо визначення такого набору значень змінних вхідного інформаційного ресурсу у прикладній онтології, який повністю відповідає рішенню задачі задоволення обмежень, відповідно до формалізму об'єктно-орієнтованих мереж обмежень, запропонованого А. В. Смирновим [11].

Знання проблемної області описуються в прикладній онтології. Дана онтологія створюється експертами. Вона об'єднує в собі дві складові: концептуальні знання та знання про завдання, що вирішуються в проблемній області. Тому розглянемо концепцію подання знань у прикладній онтології, де використовується формалізм об'єктно-орієнтованих мереж обмежень, запропонований А. В. Смирновим [11]. Відповідно до даного формалізму знання описуються як множина класів, атрибутів класів, доменів атрибутів і обмежень. Сумісність з завданням задоволення обмежень — знання, представлені за допомогою цього формалізму, можуть бути інтерпретовані як завдання задоволення обмежень.

Об'єктно-орієнтована множина обмежень має наступний вигляд:

$$A = \langle O, Q, D, C \rangle, \tag{2}$$

де O — множина класів об'єктів («класи»); Q — множина атрибутів класів («атрибути»); D — множина доменів — областей допустимих значень атрибутів («домени»); C — множина обмежень (приналежність атрибутів Q класам O і доменів D атрибутам Q, приналежність вхідних (вихідних) аргументів Q методам O і доменів D аргументам Q):

$$C = \langle O, Q, D \rangle. \tag{3}$$

Таксономічне і ієрархічне відношення, використовувані для моделювання таксономії та ієрархії класів, виглядають наступним чином:

$$C^{\text{II}} = \langle O_1, O_2, 1 \rangle, \quad (4)$$

де C^{II} — таксономія класів (клас O_2 є підкласом класу O_1).

$$C^{\text{II}} = \langle O_1, O_2, 0 \rangle, \quad (5)$$

де C^{II} — ієрархія класів (клас O_2 є частиною класу O_1).

Обмеження на сумісність (True) (несумісність (False)) класів O_1 і O_2, що використовується для моделювання ситуації, коли два або більше примірника класів не можуть бути частинами одного і того ж класу:

$$C^{\text{III}} = \langle O_1, O_2, True/False \rangle. \quad (6)$$

Асоціативні відношення між класами O_1 і O_2:

$$C^{\text{IV}} = \langle O_1, O_2 \rangle. \quad (7)$$

Потужність множини значень атрибутів (кардинальність класів), що використовується для завдання кількості значень, що може приймати даний атрибут, $N_{\max}$ — максимально можлива кількість підкласів класу O:

$$C^{\text{V}} = \langle O_1, O_2, N_{\max} \rangle. \quad (8)$$

Функціональні обмеження (атрибути Q_1 і Q_2 пов'язані функціональним відношенням):

$$C^{\text{VI}} = \big((O_1, O_2), (O_1, O_2) \big). \quad (9)$$

Визначивши можливий ряд обмежень системи, сформуємо задачу задоволення даних обмежень. Задачу задоволення обмежень (Satisfaction Constraint Task) будемо розглядати як трійку множин: множину змінних, множину відповідних областей змінних (домени змінних) і множину обмежень на змінні. Рішенням задачі задоволення обмежень називається набір значень змінних, що задовольняють всім обмеженням на змінні:

$$SCT = (V_{SCT}, D_{SCT}, C_{SCT}), \quad (10)$$

де V_{SCT} — множина змінних, D_{SCT} — множина доменів змінних, C_{SCT} — множина обмежень на змінні.

Виділимо основні операції створення прикладної онтології:

1. Формування множини ключових слів.
2. Переклад термінології:
— переклад ключових слів на мови, використовувані в початкових онтологіях.
3. Пошук:
— виявлення онтологій, в яких представлені ключові слова;
— маркування елементів вихідних онтологій, імена яких збігаються з ключовими словами.
4. Вибірка:
— маркування елементів вихідних онтологій, пов'язаних в логічному висновку з ключовими словами.
5. Добування:
— побудова зрізів початкових онтологій на основі маркованих елементів.
6. Переклад термінології:
— переклад словників зрізів в словник прикладної онтології.
7. Перетворення представлень:
— перетворення форматів представлення зрізів в формалізм системи.
8. Інтеграція:
— об'єднання зрізів в єдину онтологію.
9. Доопрацювання експертами:
— встановлення відсутніх обмежень, визначених у формалізмі системи;
— включення в прикладну онтологію онтології завдань і методів.

Відповідно до п'ятого кроку завершальним етапом роботи Мешап-системи для визначення структури і змісту вхідних інформаційних ресурсів є формування моделі вхідного інформаційного ресурсу загальної структури, визначеної предметною областю. Для вирішення даної задачі хочеться зазначити, що Мешапи працюють, використовуючи API додатків для інтеграції між кількома додатками, дозволяючи вилучення даних і обмін даними між додатками. API допомагають отримати доступ і використовувати ресурси, не зосереджуючись на своїй внутрішній організації. Оскільки у всіх ресурсів API є унікальні URI з HTTP-доступом, створювати математичну модель інформаційного ресурсу потрібно із фігуруванням URI ресурсу [12].

7. Висновки

У результаті проведених досліджень:

1. Досліджено та формалізовано роботу Мешап-системи, через побудову загальної UML-діаграми діяльності Мешап-системи та опису її роботи, відповідно до виділених станів діяльності.

2. Наведено та охарактеризовано 5 основних кроків для визначення структури і змісту вхідних інформаційних ресурсів у роботі Мешап-системи.

3. Розглянуто онтологічний підхід до моделювання та концепцію подання знань у прикладній онтології, де використовується формалізм об'єктно-орієнтованих мереж обмежень, запропонований А. В. Смирновим, як одні з можливих рішень завдань, наведених у принципах щодо побудови Мешап-системи.

Литература

1. Kushniretska, I. I. Application of MashUp Technology for Dynamic Integration of Semi-Structured Data [Text] / I. I. Kushniretska, A. Y. Berko // Матеріали VI Міжнародної конференції молодих вчених CSE-2013. — Львів: Видавництво НУ «Львівська політехніка», 2013. — P. 220–221.
2. Altinel, M. Damia — A Data Mashup Fabric for Intranet Applications [Text] / M. Altinel, P. Brown, S. Cline, R. Kartha, E. Louie, V. Markl, L. Mau, Y.-H. Ng, D. Simmen, A. Singh // VLDB'07, September 23–28, 2007, Vienna, Austria. — VLDB Endowment, 2007. — P. 1370–1373.

3. Ennals, R. J. Mashmaker: mashups for the masses [Text] / R. J. Ennals, M. N. Garofalakis // Proceedings of the 2007 ACM SIGMOD international conference on Management of data — SIGMOD'07. — ACM Press, 2007. — P. 1116–1118. doi:10.1145/1247480.1247626
4. Huynh, D. F. Potluck: Data Mash-Up Tool for Casual Users [Text] / D. F. Huynh, R. C. Miller, D. R. Karger // The Semantic Web. — Springer Science + Business Media, 2007. — P. 239–252. doi:10.1007/978-3-540-76298-0_18
5. Wong, J. Marmite: Towards End-User Programming for the Web [Text] / J. Wong // IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 2007). — IEEE, 2007. — P. 1541–1546. doi:10.1109/vlhcc.2007.40
6. Baader, F. The Description Logic Handbook: Theory, Implementation and Applications [Text] / F. Baader, D. Calvanese, D. McGuinness, D. Nardi, P. F. Patel-Schneider. — Cambridge University Press, 2003. — P. 132–136.
7. Lenzerini, M. Data Integration: A Theoretical Perspective [Text] / M. Lenzerini // Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems — PODS'02. — ACM Press, 2002. — P. 233–246. doi:10.1145/543613.543644
8. Bock, C. UML 2 Composition Model [Text] / Conrad Bock // Journal of Object Technology. — November — December 2004. — Vol. 3, No. 10. — P. 47–73. — Available at: \www/ URL: http://www.jot.fm/issues/issue_2004_11/column5
9. Леоненков, О. Самовчитель UML [Електронний ресурс] / О. Леоненков. — Київ, 2007. — Режим доступу: \www/URL: http://kniga.scienceontheweb.net/samouchitel-uml-2.html.
10. Бездушный, А. А. Математическая модель интеграции данных на основе дескриптивной логики [Текст]: автореф. дис. канд. физ.-мат. наук / А. А. Бездушный. — М., 2008. — С. 21.
11. Смирнов, А. В. Подход к построению распределенной системы интеллектуальной поддержки принятия решений в открытой информационной среде [Text] / А. В. Смирнов, М. П. Пашкин, Н. Г. Шилов, Т. В. Левашова; Санкт-Петербургский институт информатики и автоматизации РАН. — Санкт-Петербург: СПИИРАН, 2007. — С. 145–153.
12. Кушнірецька, І. І. Методика «чорної скриньки» та агентно-орієнтованого підходу для забезпечення Mash-Up інтеграції інформаційних ресурсів [Text] / І. І. Кушнірецька, А. Ю. Берко // Матеріали 3-ї Міжнародної наукової конференції «Інформація, комунікація, суспільство 2014» (ICS-2014). — Львів — Славське: Видавництво НУ «Львівська політехніка», 2014. — С. 62–63.

ОПРЕДЕЛЕНИЕ СТРУКТУРЫ И СОДЕРЖАНИЯ ВХОДНЫХ ИНФОРМАЦИОННЫХ РЕСУРСОВ ДЛЯ РАБОТЫ МЕШАП-СИСТЕМЫ

В статье представлена характеристика основных шагов определения структуры и содержания входных информационных ресурсов для работы Мэшап-системы. Приведена общая схема деятельности Мэшап-системы и описано каждое состояние деятельности системы. Предложены возможные решения задач, поставленных соответственно выделенных шагов определения структуры и содержания входных информационных ресурсов для работы Мэшап-системы.

Ключевые слова: Мэшап-система, информационный ресурс, состояния деятельности Мэшап-системы, онтологии, формализм ограничений.

Кушнірецька Ірина Ігорівна, *аспірант, кафедра інформаційних систем та мереж, Національний університет «Львівська політехніка», Україна,* ***e-mail: presty@i.ua****.*

Кушнірецька Оксана Ігорівна, *аспірант, кафедра інформаційних систем та мереж, Національний університет «Львівська політехніка», Україна,* ***e-mail: presstail@i.ua****.*

Берко Андрій Юліанович, *доктор технічних наук, професор, завідувач кафедри загальної екології та екоінформаційних систем, Національний університет «Львівська політехніка», Україна,* ***e-mail: berkoandriy@yandex.ua****.*

Кушнирецка Ирина Игоревна, *аспирант, кафедра информационных систем и сетей, Национальный университет «Львовская политехника», Украина.*

Кушнирецка Оксана Игоревна, *аспирант, кафедра информационных систем и сетей, Национальный университет «Львовская политехника», Украина.*

Берко Андрей Юлианович, *доктор технических наук, профессор, заведующий кафедрой общей экологии и экоинформационных систем, Национальный университет «Львовская политехника», Украина.*

Kushniretska Irina, *Lviv Polytechnic National University, Ukraine,* ***e-mail: presty@i.ua****.*

Kushniretska Oksana, *Lviv Polytechnic National University, Ukraine,* ***e-mail: presstail@i.ua****.*

Berko Andriy, *Lviv Polytechnic National University, Ukraine,* ***e-mail: berkoandriy@yandex.ua***

УДК 651.2:004.89
DOI: 10.15587/2312-8372.2014.29990

Осідач А. О.

РОЗРОБЛЕННЯ АРХІТЕКТУРИ СИСТЕМИ КОРПОРАТИВНОГО ДОКУМЕНТООБІГУ

На основі проведеного дослідження ключових вимог, характеристик, існуючих проблем систем документообігу, аналізу ринку загалом та конкретних програмних рішень запропоновано архітектурну модель системи електронного корпоративного документообігу. Особлива увага звертається на модульність системи та головні функціональні вузли архітектури. Описана архітектура відрізняється від існуючих рішень наявністю інтелектуальної складової — аналізатора виконання документів, формування його життєвого циклу та системи кластеризації.

Ключові слова: *корпоративний документообіг, інтелектуальна система, електронний документ, система кластеризації, аналізатор виконання.*

1. Вступ

Автоматизація документообігу — ключовий процес, що сприяє підвищенню ефективності управління. Це досягається шляхом скорочення часу, необхідного для обробки документів, їх класифікації, зберігання, підвищення швидкості, простоти пошуку та одержання інформації. Основними цілями впровадження системи