

КОНТЕКСТНЕ СПІВСТАВЛЕННЯ ВЕБ-СЕРВІСІВ У ВИРІШЕННІ ЗАДАЧІ ВИЯВЛЕННЯ. ОНТОЛОГІЧНІ ПІДХОДИ

О.В. Захарова

Співставлення є невід'ємною складовою всіх задач веб сервісів. Ефективна реалізація процесу співставлення сервісів та пошукового запиту, або сервісів один з одним гарантує успішне вирішення задачі виявлення сервісів, побудови найкращого покриття пошукового запиту та складного композитного сервісу, що забезпечує реалізацію поставленої бізнес-задачі. Всі веб сервіси мають власні особливості, вирішують різні задачі, використовують різні методи, але всі вони також мають певні характеристики: функціональну та процесну моделі представлення, мову формального опису, протокол взаємодії, спільні типи елементів, такі як ідентифікатори, повідомлення, параметри тощо. Це дозволяє класифікувати їх певним чином, а аналіз цих характеристик – виділити елементи та аспекти процесу співставлення. Процес знаходження сервісів повинен враховувати структурну, синтаксичну та семантичну відповідність сервісів, містити порівняльний аналіз сервісів за максимально можливою кількістю характеристик, що є значимими для співставлення. Він має бути покроковим та побудований таким чином, щоб кожний крок скорочував множину сервісів-кандидатів, залишаючи сервіси з більшим ступенем відповідності. Для цього необхідні алгоритми, які повертають кількісні оцінки відповідності при виконанні кожного кроку, кожного типу співставлення, та алгоритми, що в змозі ефективно узагальнити ці оцінки для виведення кінцевих значень близькості сервісів та запиту. Дані дослідження присвячені питанням використання апарату дескриптивних логік (ДЛ) для співставлення веб сервісів за їх контекстними описами, що, як правило, містять інформацію про призначення сервісів, їх область застосування, бізнес-функції тощо. Тобто це інформація, яка представляє семантику сервісу, але в текстовій формі, не зручній для автоматизованої обробки. В даній роботі пропонується онтологічний підхід до співставлення веб сервісів за контекстом. Визначено розширення раніше запропонованої ДЛ онтології сервісу верхнього рівня. Даний підхід також передбачає використання спеціальної онтології загального текстового опису сервісу, фрагмент таксономії якої наведений у роботі. Ця онтологія повинна охоплювати всі важливі семантичні аспекти контекстних описів, і не лише сприяти визначенню відповідності сервісу та запиту, але й дозволяти семантично категоризувати наявні сервіси: за предметною областю, за функціями, що реалізуються тощо.

Ключові слова: веб-сервіс, задача виявлення, процес співставлення, типи співставлення, методи встановлення ступеня відповідності, семантичне співставлення, контекстне співставлення, онтологічні підходи, онтологія сервісу верхнього рівня, категоризація та класифікація веб сервісів, таксономія загального текстового опису веб сервісів, встановлення відповідності сервісу за контекстом.

Сопоставление (matching) – неотъемлемая составляющая фактически всех задач веб сервисов. Эффективная реализация процесса сопоставления сервисов с поисковым запросом, или друг с другом гарантирует успешное решение задач нахождения нужных сервисов, построения наилучшего покрытия поискового запроса и сложного композитного сервиса, реализующего поставленную бизнес-задачу. Все веб сервисы имеют свои особенности, решают разные задачи, используют разные методы, однако они все также имеют определенный набор характеристик: функциональную и процессную модели, формальный язык описания, протокол взаимодействия, общий набор типов элементов, таких как идентификаторы, служебные сообщения, параметры и т. д. Это позволяет классифицировать их определенным образом, а анализ этих характеристик – выделить элементы и аспекты процесса сопоставления. На сегодняшний день существует огромное количество сервисов, и их число постоянно растет, а процесс нахождения сервисов, соответствующих запросу, очень сложный и многогранный. Он должен учитывать структурное, синтаксическое, но, прежде всего, семантическое соответствие сервисов, обеспечивать их сравнительный анализ по максимально возможному числу характеристик, значимых для сопоставления. Процесс должен быть пошаговым и построенным таким образом, чтобы каждый шаг сокращал множество сервисов-кандидатов, оставляя сервисы с большей степенью соответствия. Для этого нужны алгоритмы, которые возвращают количественные оценки соответствия при выполнении каждого шага, каждого типа сопоставления, и алгоритмы, эффективно обобщающие эти оценки для вывода конечных значений близости сервисов и запроса. Данные исследования посвящены вопросам использования аппарата дескриптивных логик для сопоставления веб сервисов по описаниям их контекстов, которые, как правило, содержат информацию о назначении сервисов, про область их использования, бизнес-функции и т. д. То есть это как раз информация, представляющая семантику сервиса, но в текстовой форме, не удобной для автоматизированной обработки. Существует немало исследований, которые пытаются решить эту проблему, применяя к контекстным описаниям сервисов стандартные методы текстового анализа. В данной работе предлагается онтологический подход к сопоставлению веб сервисов по контексту. Определено расширение ранее предложенной ДЛ онтологии сервиса верхнего уровня. Данный подход предусматривает также использование специальной онтологии общего текстового описания сервиса, фрагмент таксономии которой представлен в работе. Эта онтология должна охватывать все важные семантические аспекты контекстных описаний, и не только способствовать определению соответствия сервиса и запроса, но и позволять семантически категоризовать имеющиеся сервисы: по предметной области, реализуемым функциям и т. д.

Ключевые слова: веб-сервис, задача поиска, процесс сопоставления, типы сопоставления, методы определения степени соответствия, семантическое сопоставление, контекстное сопоставление, онтологические подходы, онтология сервиса верхнего уровня, категоризация и классификация веб сервисов, таксономия общего текстового описания веб сервисов, установление соответствия сервиса по контексту.

Matching is integral part of all web services tasks. The effective implementation of the process of matching services with a search query, or with each other, ensures successful resolving of the problems of service discovery, building the best coverage of the search query and a complex composite service that implements the business task. All web services have own characteristics, solve different problems, use different methods, but they all also have a certain set of characteristics: functional and process models, formal description language, communication protocol, a common set of element types, such as identifiers, service messages, parameters and etc. This allows to classify services in a certain way, the analysis of these characteristics permits to define the elements and aspects of the comparison process. Today there is a huge number of services, and their number is constantly growing, and the process of services discovery is very complex and multifaceted. It should take into account the structural, syntactic, but, first of all, the semantic suitability of services, to ensure their comparative analysis according to the maximum possible number of characteristics that are significant for matching. It should be step-by-

step process and it has been designed in such a way that each step reduces the set of candidates, leaving services with a greater degree of compliance. This requires algorithms that return quantitative estimates for each step, each type of matching, and algorithms that effectively generalize these estimates to define the final values of the proximity of services and the request. These researches are devoted to problems of using the descriptive logics formalisms for web services matching by their contexts, which, as a rule, contain information about the services purposes, the area of use, business functions, etc. That is, it is information representing the semantics of the service, but in text form, that is not convenient for automated processing. There are many studies that try to solve this problem by applying standard text analysis methods to contextual service descriptions. This study proposes an ontological approach to matching web services by context. It is determined the extension of the previously proposed the top-level service DL ontology. It also involves the use of a special ontology of the general textual service description, a fragment of the taxonomy of which is presented in the paper. This ontology should cover all important semantic aspects of contextual descriptions. It have not only promote to determining the matching a service and a request, but it also have to allow semantically categorizing the available services: by subject area, implemented functions, etc.

Key words: web service, discovery task, matching process, matching types, matching degree defining methods, semantical matching, contextual matching, ontological approaches, top level service ontology, web services categorization and classification, a taxonomy of general textual description of web services, matching of web services by context.

Вступ

Процес співставлення полягає у попарному порівнянні характеристик об'єктів з метою встановлення їх відповідності. У ракурсі вирішення задач веб-сервісів, головні вимоги до процесу співставлення можна визначити наступним чином:

- співставлення не повинно спиратися лише на порівняння описів сервісів за ключовими словами, до уваги повинні прийматися семантики оголошень сервісів та запиту;
- процес має бути автоматизованим;
- процес співставлення має бути точним. Якщо повертається результат з певним ступенем відповідності, то необхідне чітке визначення самого ступеня відповідності.
- процес має бути ефективним та швидким. Алгоритм має відбирати невеличку кількість сервісів, найбільш відповідних запиту.

Визначення та загальна архітектура процесу співставлення

Веб-сервіси є формально визначеними об'єктами з певним набором характеристик, щойменше, множиною вхідних та вихідних параметрів, протоколом взаємодії, що використовується, мовами представлення, типом функціональної моделі тощо. Всі ці характеристики є спільними для всіх сервісів та дозволяють класифікувати їх певним чином на верхньому рівні, не зважаючи на детальний опис.

Це дозволяє звзвити множину сервісів на початковому етапі для подальшого виконання їх вартісного складного аналізу та співставлення з запитом. Приклад такої класифікації показано на рис. 1.

Для забезпечення постійного, своєчасного процесу співставлення, що гарантує адекватний результат, його інформаційна модель має містити:

- реєстр або база даних (БД) оголошень веб сервісів, їх формальні визначення, онтології сервісів;
- глобальна онтологія понять, що використовуються у загальних описах сервісів;
- загальна онтологія сервісу;
- онтології доменів прикладної області (ПО) бізнес-задачі. В процесі співставлення можуть використовуватися додаткові інформаційні ресурси (БД пар термінів з відстанями між ними, ієрархії сутностей, використані в оголошеннях сервісів, які представлені у реєстрі оголошень, внутрішня термінологія), як, наприклад, це реалізовано у Larks [1].

Загальна класифікація підходів співставлення сервісів

Умовно, можна виділити:

- синтаксичне співставлення;
- семантичне співставлення;
- контекстне співставлення.

Зрозуміло, що для досягнення ефективного результату, процес співставлення має являти собою погоджене поєднання вказаних підходів.

Синтаксичне співставлення – це суто технічний процес, що сам по собі може призводити до не коректних висновків щодо результатів співставлення: синтаксично схожі сервіси можуть реалізовувати зовсім різні задачі, та навпаки, сервіси, що вирішують певну проблему, можуть мати різний синтаксис.

Найбільш цікавим є *семантичне співставлення* сервісів. Так як, перш за все, метою пошуку сервісу є виконання конкретної задачі і лише семантичний опис сервісу визначає його зміст, цілі, функції тощо.

Семантичне співставлення веб-сервісів. Підходи семантичного співставлення можна класифікувати як ті, що базуються:

- на онтології домену;
- на використанні загальної онтології;
- на синтаксисі та семантиці.

Підходи на основі онтології домену розглядають онтології як структури знань для визначення атрибутів сервісів, їх властивостей та відношень між ними. База знань (БЗ) онтології використовується як для визначення запиту, так й для представлення сервісу. Концепти, що не були визначені у первинному представленні, можуть бути додані до опису, наприклад, як семантичні анотації, та використовуватись у процесі співставлення. Для їх визначення використовуються знання з онтології домену. Це дозволяє отримати більш точний результат – знайти семантичну відповідність сервісів та запиту, що не містять точно співпадаючих атрибутів, на де-якому абстрактному рівні. Окрім цього, такі підходи можуть забезпечити категоризацію сервісів на основі їх семантики та дозволяють досягти функціонального рівня категоризації на основі онтології, а також надають можливість вдосконалити семантичний опис запиту, що, в свою чергу, забезпечує більш якісний результат пошуку сервісів.

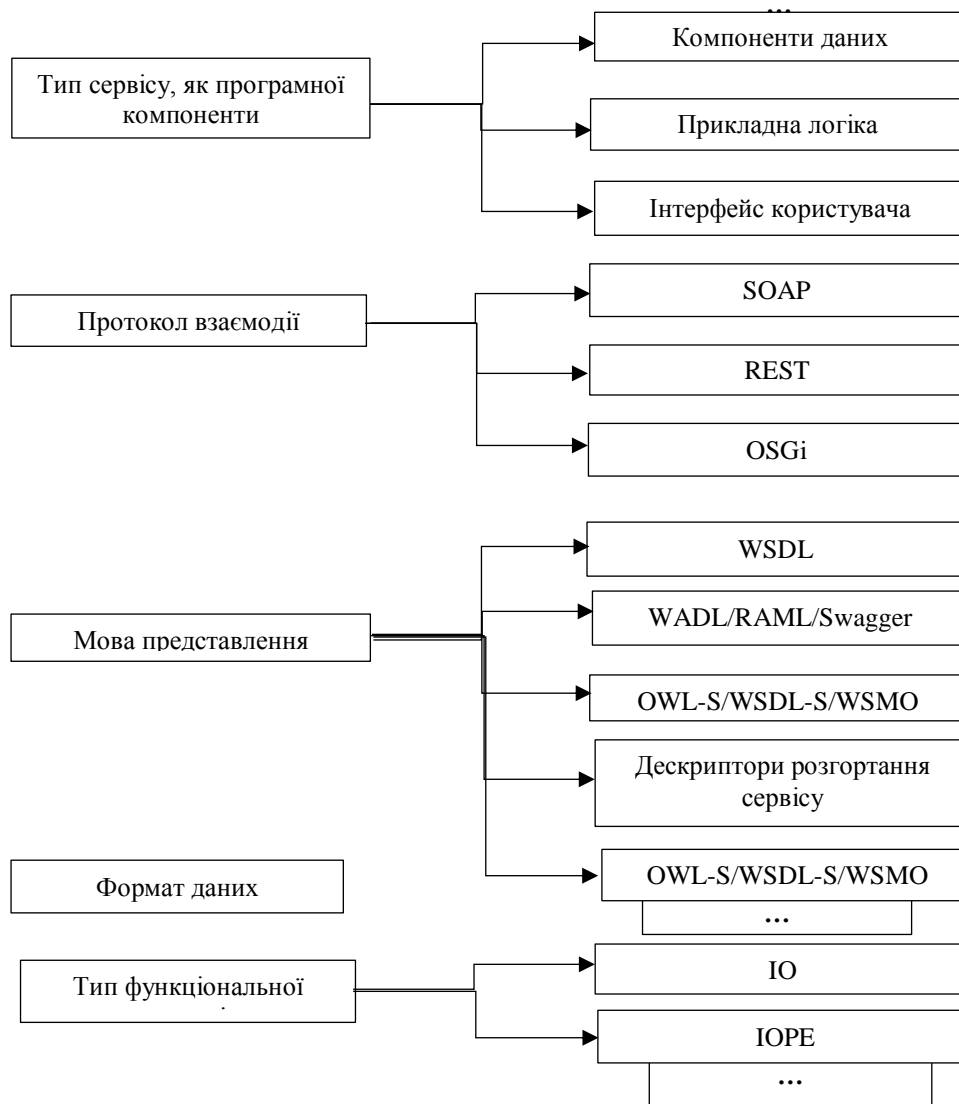


Рис. 1. Класифікація сервісів за загальними характеристиками

Слід зазначити, що існує чимало методів співставлення на основі онтології доменів, які пропонують різноманітні алгоритми для визначення відповідності сервісів та запиту, використовують різні моделі представлення сервісів, засоби обчислення семантичної відстані, порівнюють різні характеристики тощо. Їх перевагою є спрямованість на автоматизовану ідентифікацію, інтеграцію та виявлення сервісів, та використання семантик, що дозволяє покращити релевантність їх виявлення. Але, більшість з цих підходів базуються на конкретних мовах представлення сервісів, як, наприклад, OWL-S, WSMO чи WSDL-S, що вимагає від кінцевого користувача певного рівня знань для їх використання. Окрім цього, запитувач сервісу може не володіти у повній мірі знаннями, що містяться в онтології домену, не знати всіх умов, що пов'язані з запитом, і запит, таким чином, може виявитися не достатньо коректно визначеним. Це може призвести до того, що у процесі виявлення буде втрачено багато відповідних сервісів. Інше обмеження існуючих підходів полягає у тому, що вони базуються на припущенні, що провайдери сервісів та запитувач сервісу використовують одну й ту саму

онтологію домена для опису сервісу та запиту, але у реальному житті, як правило, це не так і подолати цю різницю між онтологіями неможливо без техніки відображення онтологій.

Підходи на основі загальної онтології дозволяють користувачам доповнювати сервіси семантичною інформацією, не виконуючи семантичного анотування. Таким чином, на відміну від традиційного підходу, вони не потребують значних витрат на управління онтологією та семантичне анотування, та є незалежними від онтології домену. Мова йде про визначення семантичних тегів на частини повідомлень веб сервісів. Ці підходи базуються на неявних семантиках веб сервісів, а для формування результату співставлення використовують загальну базу знань та семантичні ризонери. Для семантизації сервісу, в даному випадку, будується онтологія сервісу. Це може бути спеціальна самовизначена онтологія або загальновикористовувана онтологія. Використання загальної онтології спрощує застосування методів даної групи та забезпечує можливість їх повторного використання. Таким, добре відомим, словником загального використання є, наприклад, WordNet. Його перевагою є те, що терміни та концепти в ньому мають точну семантику. WordNet може використовуватися як онтологія сервісу, а його концепти (семантичні слова) – для анотування бібліотеки сервісів, сервісу та його інтерфейсу. Насьогодні існує чимало прикладів використання WordNet для семантичного анотування сервісів [2–6]. Здебільшого це пояснюється незалежністю WordNet від конкретного домену, що є його головною перевагою.

Підходи на основі синтаксису та семантики. До цієї групи відносяться процеси виявлення сервісів, що базуються на співставленні функціональних та не функціональних вимог сервісу та запита ([17]) та встановлення якості сервісів на основі спеціальних метрик. Це можуть бути суто синтаксичні підходи, суто семантичні підходи, або методи, які комбінують як синтаксичні, так і семантичні алгоритми. При цьому в процесі співставлення вони можуть використовувати різні онтології: онтологію для інтеграції стандартів якості та встановлення властивостей сервісу, онтологію для моделювання відношень між цими стандартами та уподобаннями користувачів для ранжування сервісів за функціональними характеристиками, онтологію, що дозволяє встановити відношення між моделями якості та функціональністю веб-сервісів тощо.

Процес виявлення веб-сервісів, у даному випадку, може комбінувати різні типи пошуку: за контекстом, за функціональністю та за метриками якості. Пошук може враховувати виявлення синонімів. При встановленні відповідності підходи даної групи можуть використовувати як специфікації сервісів, так і їх сигнатури, комбінувати логічне та не логічне співставлення.

Контекстні підходи. Бізнес-задачі дуже залежать від ситуаційного контексту, в якому вони виконуються, але технології семантичних веб сервісів зазвичай не заохочують до представлення таких ситуацій у домені. Більше того, опис складного поняття конкретної ситуації з урахуванням всіх її аспектів, є досить вартісною задачею та може ніколи не досягти достатньої семантичної виразності. Жодна ситуація у реальному світі не може повністю співпадати з іншою, але має відповідати кінцевому набору семантично визначених описів параметрів, щоб забезпечити адаптивність до контексту. Дана група методів гарантує найточніше автоматизоване виявлення на основі вичерпного семантичного опису запиту та сервісу, та з урахуванням ситуаційного контексту, або контексту користувача [7].

Розуміння контексту є ключовим компонентом у будь-якій розповсюдженій системі та забезпечує інтелект системи, дозволяючи обчислювальним пристроям приймати відповідні та своєчасні рішення від імені користувача. Такі системи мають адаптуватися до змін і варіацій контексту користувача [8]. Але поняття контексту є занадто широким, оскільки контекст може бути будь-якою інформацією, що стосується користувача або програмної системи [9]. Його дуже складно моделювати, тим більше єдиним способом для різних застосунків. Його використання в автоматизованих алгоритмах також не є тривіальним. Це обумовлює складність масового використання контекстно залежних підходів та цілу низку проблем, що пов'язані з їх застосуванням для вирішення задач веб сервісів.

Типи співставлень та класифікація методів встановлення відповідності

Більшість задач веб-сервісів базуються, перш за все, на визначенні ефективних підходів для виявлення потрібних сервісів відповідно до особливостей їх представлення. Множина та особливості елементів, що використовуються для визначення сервісу та запиту, модель представлення сервісу й запиту та специфіка самої бізнес-задачі обумовлюють доцільність різних типів співставлення сервісів:

– *співставлення описів сервісів та запиту природньою мовою (контекстів).* Синтаксис WSDL дозволяє створювати текстові описи сервісів, їх типів та операцій, що сформовані у групи за тегами <Documentation>. Таким чином, у даному випадку задача зводиться до визначення ступеня відповідності текстових описів сервісів, тобто до традиційної задачі інформаційного пошуку;

– *структурне співставлення сервісів та запиту (співставлення профілів).* Структурне співставлення сервісів полягає у порівнянні структур їх функціональних моделей представлення. Існують різні типи функціональної моделі сервісу. Найбільш розповсюдженою є IOPE модель, де функціональні можливості сервісу визначаються його вхідними та вихідними параметрами, перед-умовами та ефектами. Наочно, IO (входи, виходи) надають синтаксичну інформацію про веб-сервіси, тоді як PE (передумови та ефекти) – відображають їх семантику. Методи, які використовуються у співставленні виходів-входів відрізняються від

тих, що використовується для передумов та ефектів. Семантики, які відображаються входами-виходами й передумовами та ефектами також різні. Насьогодні досягнуто значних результатів у співставленні веб-сервісів за входами-виходами, але відсутній ефективний підхід щодо співставлення передумов та ефектів, тому винесемо PE співставлення до окремої категорії;

- співставлення сервісів та запиту за передумовами та ефектами (співставлення умов та обмежень);
- співставлення операцій сервісів. Визначення відповідності операцій сервісів є послідовним кроком у вирішенні загальної задачі виявлення сервісу. Ступінь відповідності операцій визначається на основі того, в якому ступені подібні їх вхідні та вихідні повідомлення на попередніх кроках встановлення відповідності. Загальна оцінка ступеня відповідності двох сервісів визначається шляхом встановлення попарної відповідності їх операцій, що максимізує загальну суму співпадань для окремих пар;

- семантичне структурне співставлення за ідентифікаторами операцій (співставлення сигнатур). Обрані імена типів, операцій та сервісів зазвичай відображають семантики властивостей сервісу. Даний тип співставлення є процесом встановлення семантичної відповідності ідентифікаторів типів даних, операцій та сервісів [10].

Відповідно до перелічених типів матчингу можна виділити наступні групи методів.

Методи встановлення структурної відповідності. Насьогодні більшість веб сервісів мають WSDL [11] представлення профілю. Тому, більшість підходів структурного співставлення базуються на співставленні вхідних та вихідних WSDL специфікацій, що включає порівняння набору операцій, які пропонуються сервісами, що, у свою чергу, базується на порівнянні структур вхідних та вихідних повідомлень операцій, основанийому на порівнянні типів даних об'єктів, які передаються цими повідомленнями. Результатом такого співставлення є оцінки ступеня відповідності всіх типів даних джерела та цілі. У роботі [12] пропонується підхід, де вказані оцінки формують матрицю, яка в подальшому може бути використана для визначення загальної оцінки ступеня відповідності сервісів як за вхідними та вихідними параметрами, так і в цілому. Слід зазначити, що по-перше, типи даних веб-сервісів у WSDL є XML елементами і вони можуть бути досить складними структурами. По-друге, WSDL специфікації сервісів не включають їх семантику і таке співставлення є суто технічним.

Інша підгрупа підходів структурного співставлення – методи на основі DAML-S онтологій. Мова DAML-S [13] побудована на основі логік та підтримує специфікацію семантичної інформації в RDF форматі. DAML-S забезпечує представлення сервісу, яке охоплює поведінкові специфікації операцій. Профіль сервісу DAML-S визначається IOPE моделлю. DAML-S робить можливим виявлення сервісів через співставлення специфікацій. Відомі різні підходи до реалізації алгоритмів семантичного співставлення на основі DAML-S онтологій [14]. Спільним для них є те, що оголошення сервісу та запиту звертаються до DAML концептів та пов'язаних семантик. Використання DAML надає можливість процесу співставлення здійснювати логічне виведення на ієрархії категорій, що призводить до розпізнавання семантичних відповідностей, незважаючи на синтаксичні відмінності та відмінності у моделюванні абстракцій у оголошеннях сервісів. DAML також забезпечує точність співставлення: відповідність має місце лише, якщо вона виводиться з DAML онтологій, які використовуються у реєстрі. Більше того, семантики DAML-S описів дозволяють визначити функції ранжування для ступенів відповідності. Щонайменше, DAML-S дозволяє не лише порівняти два сервіси, а й формально довести, що знайдений та цільовий сервіси не конфліктують. Загальна архітектура механізму співставлення на основі DAML-S може бути подібна, показаний на рис. 2.

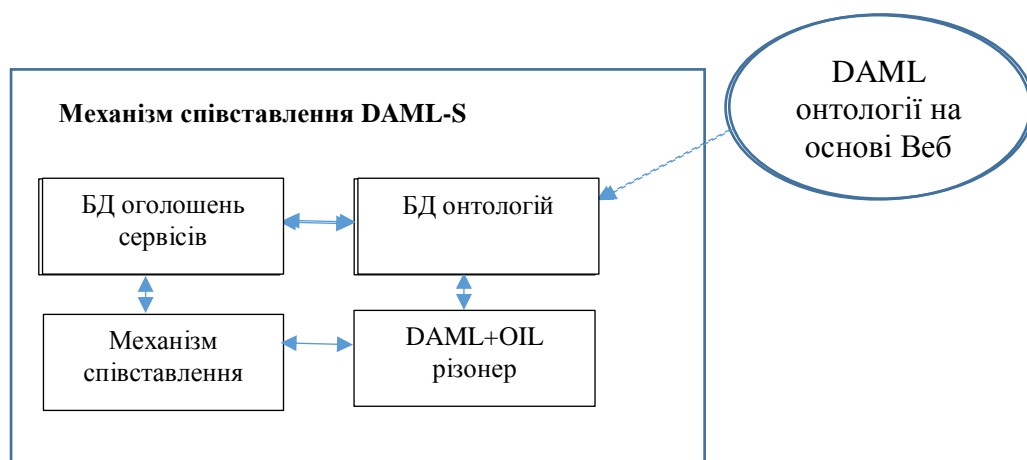


Рис. 2. Архітектура механізму співставлення DAML-S

Аналогічна архітектура може бути застосована й у інших онтологічних підходах. У роботі [15] наводиться алгоритм PE співставлення, на основі формалізму ДЛ SHOIN+(D). Його сутність полягає у тому, що

якщо два сервіси представлені РЕ моделлю на базі одного ациклічного Tbox: $S_r=(pre_r, effect_r)$ та $S_a=(pre_a, effect_a)$, де S_r – сервіс-запит, а S_a – наявний сервіс, то вважають, що S_a відповідає S_r , якщо:

- передумови оголошеного сервісу є простішими ніж передумови запиту, або щонайменше, їх так само легко задовільнити;
- ефекти оголошеного сервісу не менші ніж ефекти запиту.

Формально, для сервісу та запиту, що представлені IOPE моделлю, задачу встановлення структурної відповідності можна визначити наступним чином. Нехай сервіс $S = \{s/ s = (CI; I; O; A)\}$ з репозиторію \mathcal{R} , де CI – передумови сервісу, I та O – кінцеві множини, відповідно, вхідних та вихідних параметрів, A – ефекти сервісу, а запит Q визначається як бажаний сервіс $Q = (CI'; I'; A'; O')$, де CI' – передумови, I' , O' – кінцеві множини, відповідно, вхідних та вихідних параметрів, A' – ефекти сервісу, то сервіс S відповідає запиту Q , якщо [16] $I \subseteq I'$, $O' \subseteq O$, та передумови запиту CI' ширші ніж передумови сервісу CI , тобто з виконання передумов запиту випливає виконання запиту CI' ширші ніж передумови сервісу, а для ефектів навпаки – з виконання ефектів сервісу випливає виконання ефектів запиту.

Методи встановлення відповідності текстових описів сервісів. Головним фокусом попередніх досліджень було вирішення задач виявлення та композиції сервісів, що представлені із застосуванням ДЛ-онтологій [17–20]. Але запропоновані підходи взагалі не враховували контексту сервісу, а це є досить важливим аспектом у встановленні відповідності сервісу цілі. І навіть звичайне WSDL визначення профілю веб сервісу містить тег, який дозволяє задати контекст сервісу з важливою семантичною інформацією. Її аналіз дозволив би запобігти, у подальшому, потраплянню у множину кандидатів на композицію сервісів, які вирішують зовсім інші бізнес-задачі.

У свою чергу, використання контекстного опису веб-сервісів при вирішенні їх задач, вимагає ефективних автоматизованих методів аналізу та порівняння текстової інформації. Найбільш поширеними методами вирішення задач визначення подібності текстових описів є методи, що базуються на моделях векторного простору. Іншими словами, здійснюється пошук сервісів, найбільш подібних вхідному опису відповідних векторів. Відповідні підвектори з документів та запиту співставляються, і визначаються оцінки відповідності. При цьому для отримання більш адекватних оцінок часто використовують методи визначення вагів, а при пошуку та співставленні – різні лексичні словники (наприклад, словники синонімів). Відомі розширення методів векторного аналізу [21] на базі використання лексичної бази даних WordNet [22–26], як основи векторного простору. Найвища загальна оцінка вказує на більшу ступінь подібності між вихідною та цільовою специфікацією. Такий тип співставлення подібний співставленню сигнатур при пошуці компонент програмного забезпечення [27] та співставленню з документом шаблоном при інформаційному пошуку.

Недоліками таких методів визначення подібності текстової інформації [28] є, по-перше, їх складність, а – по-друге, вони фактично визначають лише лексичну схожість тексту до цільового, але одна й та сама сутність може бути по-різному виражена природньою мовою.

Останнім часом активно з'являються онтологічні методи семантизації даних, у тому числі текстових. Одним з таких підходів є використання таксономій концептів для обчислення семантичних відстаней між поняттями. Але таксономія визначає лише відношення узагальнення/деталізації на основі визначення концептів через ролі та набори атрибутів. Цього може бути не достатньо, для більш глибокого розуміння семантики доцільне визначення додаткових асоціативних зв'язків між концептами. У роботі [4] для цього використовуються так звана зважена асоціативна мережа – семантична мережа з концептами у вузлах та спрямованими зваженими ребрами між ними. У цій семантичній мережі допускається три типи бінарних зважених відношень: 1) узагальнення, 2) деталізація, 3) позитивна асоціація серед концептів [29] – найбільш загальне відношення між концептами. Такі асоціативні зв'язки можуть виявлятися, наприклад, на основі онтології WordNet.

Встановлення семантичної відповідності ідентифікаторів типів даних, операцій та сервісів. Даний процес, як правило, зводиться до встановлення відповідності ідентифікаторів та є подібним до процесу структурного співставлення WSDL описів. Результатом може, наприклад, бути матриця оцінок відповідності [30] для всіх попарних комбінацій типів даних вихідного та цільового сервісів. Результатом подальшого співставлення операцій може бути визначення ступеня відповідності двох операцій на базі семантичної відстані між назвами операцій, та наборів їх параметрів у термінах ідентифікаторів, які вони містять. Загальний показник відповідності обчислюється шляхом співставлення назв сервісів та попарної відповідності їх операцій, що максимізує загальну суму співпадань окремих пар.

Елементи процесу співставлення, для визначення яких доцільне використання засобів ДЛ. Виходячи з визначених типів співставлення та видів семантик [9], можна дійти висновку, що процес співставлення має охоплювати наступні елементи сервісів:

- 1) загальний контекстний опис сервісу;
- 2) вхідні параметри сервісу;
- 3) вихідні параметри сервісу;
- 4) перед-умови сервісу;
- 5) пост-умови сервісу;
- 6) ефекти сервісу;

- 7) об'єкти, що виробляються ефектами сервісів;
- 8) змінні внутрішніх операцій сервісів;
- 9) результати виконання операцій сервісів;
- 10) умови виконання операцій сервісів;
- 11) повідомлення сервісу.

ДЛ можуть використовуватися як для визначення формального семантичного опису сервісів та цілі, так і для виведення результатів співставлення. Будь-яка формалізація семантики сервісу за допомогою ДЛ дозволяє використовувати їх апарат міркувань на різних етапах процесу співставлення. Так, співставлення та сервісів за входами/виходами є стандартною задачею виводу ДЛ.

Використання апарату ДЛ для вирішення задачі співставлення веб сервісів за контекстом

Співставлення загальних контекстних описів сервісів має бути одним з початкових етапів у вирішенні задачі виявлення, мета якого – відсікти сервіси, які семантично не підходять запиту: мають інші цілі, працюють в іншій прикладній галузі з іншими даними тощо. Таксономія для вирішення такої задачі має бути в де-чому схожа на БД, що використовуються для порівняння текстів, але враховуючи можливість формалізації опису, бажано, щоб вона сприяла не лише визначенню близькості цільового та наявних описів, але, перш за все, дозволяла категоризувати наявні сервіси: за ПО, типами функцій, що реалізуються, тощо, тобто охоплювала всі важливі семантичні аспекти контекстних описів. Можлива структура таксономії показана на рис. 4.

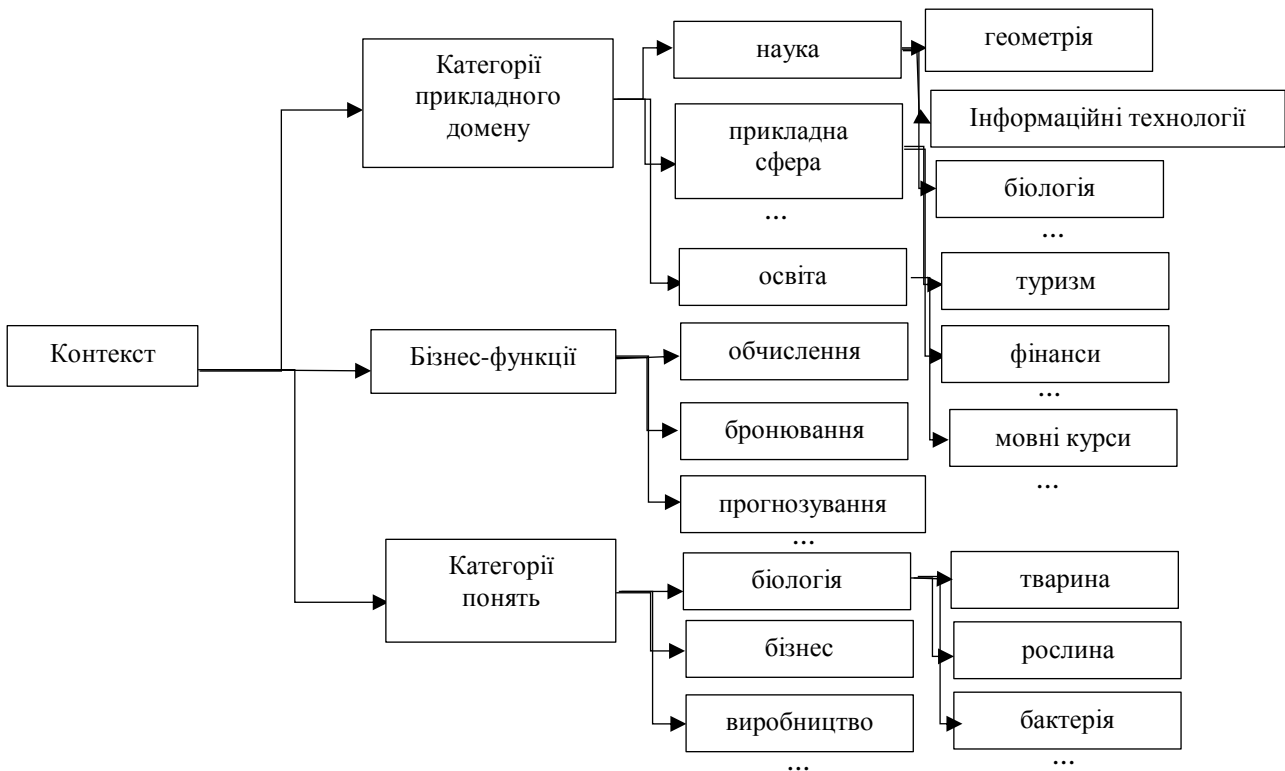


Рис. 4. Архітектура таксономії загального текстового опису сервісу

На базі таксономії понять, що характеризують контекст сервісу, доцільне визначення асоціативних зв'язків, що, наприклад, дозволять визначити можливі відношення між бізнес-функціями та прикладними доменами. Дуже важливо визначити асоціативні лексичні відносини на категоріях понять, перш за все, відношення синонімії. Відповідна до наведеного фрагменту архітектури онтологія дозволяє визначити основні аспекти опису сервісу та запиту. Далі наводиться фрагмент відповідної ДЛ онтології *DocumentationOntology*:

TBox

```

ServiceContext
BusinessDomenCategory ⊆ ServiceContext
...
BFunctionType ⊆ ServiceContext
MainEntitiesTypes ⊆ ServiceContext
ClassicScience ⊆ BusinessDomenCategory
    
```

AppliedSphere \sqsubseteq *BusinessDomenCategory*

Geometry \sqsubseteq *ClassicScience*

IT \sqsubseteq *ClassicScience*

Algebra \sqsubseteq *ClassicScience*

...

Turism \sqsubseteq *AppliedSphere*

Trade \sqsubseteq *AppliedSphere*

Finance \sqsubseteq *AppliedSphere*

...

LanguageCourses \sqsubseteq *Education*

...

Calculation \sqsubseteq *BFunctionType*

...

Mathematic \sqsubseteq *MainEntitiesTypes*

Zoology \sqsubseteq *MainEntitiesTypes*

...

Average \sqsubseteq *Mathematic*

Square \sqsubseteq *Mathematic*

...

Наведений фрагмент таксономії – лише демонстраційний приклад загальної структури понять для опису контексту. Головною метою такої онтології є визначення максимально повної узагальненої системи понять, що надають можливість структурованого семантичного опису контексту будь-якого веб сервісу, не залежно від сфери його застосування чи бізнес-призначення. Та лише за умови адекватності та досить великого ступеня повноти як системи концептів, так й відповідної бази знань, така онтологія може сприяти ефективному вирішенню задач веб-сервісів. Тому, її розробка є самостійною, не тривіальною задачею, яка вимагає залучення експертів.

ДЛ опис веб сервісу повинен будуватися як на базі онтології сервісу верхнього рівня, так і на базі прикладних онтологій, що саме й визначають семантику вхідних і вихідних параметрів, передумов, ефектів тощо. У роботі [17] вже запропонована онтологія сервісу вищого рівня для IOPE моделі **ServiceOntology**:

TBox

Service, *InputParameter*, *OutputParameter*, *PreCondition*, *PostCondition*, *Parameter*,

Name, *Value*, *Type*

Service \sqsubseteq *has.InputParameter*;

Service \sqsubseteq *has.OutputParameter*

Service \sqsubseteq *has.PreCondition*;

PreCondition \sqsubseteq *Condition*

PostCondition \sqsubseteq *Condition*;

Condition \sqsubseteq *has.Value.Boolean*

Service \sqsubseteq *has.PostCondition*;

InputParameter \sqsubseteq *Parameter*

OutputParameter \sqsubseteq *Parameter*;

Parameter \sqsubseteq *has.Name*

Parameter \sqsubseteq *has.Value*;

Type = {*String*, *Numeric*, *Boolean*}

I_s \sqsubseteq *InputParameter*

O_s \sqsubseteq *OutputParameter*;

CI_s \sqsubseteq *PreCondition*;

CO_s \sqsubseteq *PostCondition*;

Name \sqsubseteq *hasType*.{*String*}; *Value* \sqsubseteq *hasType.Type*

Запропонована онтологія фактично визначає базові сутності сервісу. Класи онтології дозволяють визначити основні елементи IOPE моделі. Але, дана онтологія не забезпечує можливості формалізації загального контекстного опису сервісу. Для цього необхідно не лише розширити онтологію сервісу верхнього рівня, але й визначити зв'язки з прикладними онтологіями та загальною онтологією понять, що дозволило б категоризувати сервіс певним чином та здійснювати матчинг сервісів не лише за входами та виходами, а й за контекстними описами, які, зазвичай, надають саме семантичну інформацію. Відповідне розширення TBox онтології **ServiceOntology**:

Documentation, *DocItem*

Service \sqsubseteq *has.DocItem*;

$Service \sqsubseteq has.Documentation$
 $DocItem \sqsubseteq Documentation;$
 $DocItem \sqsubseteq has.ItemName;$
 $DocItem \sqsubseteq has.ItemValue;$
 $DocItem \sqsubseteq BusinessDomenCategory / *зв'язок з онтологією DocumentationOntology$
 $DocItem \sqsubseteq BFunctionType / *зв'язок з онтологією DocumentationOntology$
 $DocItem \sqsubseteq MainEntitiesTypes / *зв'язок з онтологією DocumentationOntology$
 $ItemName \sqsubseteq hasType.\{String\};$
 $ItemValue \sqsubseteq hasType.Type$

Визначення задачі встановлення відповідності за контекстом. Нехай S – сервіс, а Q – запит. Визначимо задачу відповідності як:

$S \sqsubseteq Service, Q \sqsubseteq Service$
 $S \sqsubseteq has.DocItem$
 $Q \sqsubseteq has.DocItem$
 $DocItemS \sqsubseteq DocItem, DocItemQ \sqsubseteq DocItem$

Сервіс S відповідає запиту Q за контекстом, якщо його опис містить описи всіх семантик, що визначені у контексті запиту, тобто: $DocItemQ \sqsubseteq DocItemS$.

Для вирішення проблеми із застосуванням апарату ДЛ, кожний сервіс, як і запит, розглядається як специфічна онтологія зі зв'язками з множиною інших онтологій. І рішення щодо включення концептів приймається резонером на основі аналізу всієї сукупності пов'язаних концептів і різних типів зв'язків між ними (включаючи асоціативні зв'язки).

Як і у випадку співставлення за входами/виходами, можна встановити різні ступені відповідності за контекстом. Як загальні показники відповідності можна визначити:

- **Exact.** Точна відповідність має місце за умови еквівалентності концептів $DocItemQ \equiv DocItemS$.
- **Plug In.** $DocItemQ \sqsubseteq DocItemS$ – умова, що є достатньою для прийняття рішення про відповідність сервісу запиту: сервіс у контексті містить всі концепти, що запитуються.
- **Subsumes.** $DocItemS \sqsubseteq DocItemQ$, в даному випадку сервіс згідно контексту не повністю задовільняє запит, бо його опис не є достатньо повним.
- **Fail** – невдача, не знайдено будь-якого *subsumption* між $DocItemS$ та $DocItemQ$.
- **Intersection.** $DocItemIntersection \equiv DocItemQ \sqcap DocItemS$. Враховуючи специфіку екземплярів $DocItem$ та самого опису контексту, уваги заслуговують всі сервіси, контекстний опис яких має не пустий перетин з описом у запиті. При чому відсоткові оцінки цього перетину, наприклад, відношення кардинальності множини $DocItemIntersection$ до кардинальності множини $DocItemQ$ можуть точніше визначати ступінь відповідності. Ймовірно ця оцінка має бути більш комплексною та враховувати категорії елементів, що потрапили до перетину, або обчислювати близькість концептів $DocItemS$ та $DocItemQ$. Це окрема складна задача, яка потребує додаткових досліджень та вироблення спеціальних підходів до формування таких оцінок, оцінювання їх складності та доцільності застосування.

Висновки

Зважаючи на величезну кількість існуючих сервісів та динамічність їх створення, зрозуміло, що задача знаходження потрібного сервісу чи сервісів для вирішення поставленої бізнес-задачі є досить складною та вимагає ефективних підходів до виявлення сервісів-кандидатів, встановлення ступеня їх відповідності, фільтрації тощо. Процес співставлення сервісів має бути покроковим й таким, що на кожному кроці множина кандидатів суттєво звужується. Контекстне співставлення є необхідним кроком загального процесу та дозволяє відсікти сервіси, які можуть мати схожий набір параметрів, але реалізовувати різні функції.

Доцільність використання ДЛ є обґрунтованою на всіх етапах вирішення задачі виявлення. Це інструмент, що забезпечує як формальне визначення сервісу, так і механізми міркувань, які дозволяють виконувати логічне виведення. Кожний сервіс представляється ДЛ онтологією на основі онтології сервісу верхнього рівня. Онтологія конкретного сервісу базується на базі знань інтегрованої ДЛ-онтології домену. Онтології самі по собі не описують можливості сервісів, але вони забезпечують додаткові концепти, що дозволяють правильно зрозуміти та сприйняти специфікацію можливостей сервісу. Існуючі резонери ДЛ дозволяють здійснювати автоматичне співставлення. Це стандартні задачі виводу ДЛ. В даній роботі запропонований підхід співставлення веб-сервісів за контекстними описами на основі апарату ДЛ. Створення та використання при пошуку спеціальної онтології найбільш використовуваних типів сутностей не лише сприяє підвищенню ефективності вирішення задачі співставлення, але й дозволяє категоризувати/класифікувати сервіси певним чином: за областю застосування, реалізованою функцією тощо. Використання такої онтології при анотуванні або формальному визначенні сервісів може суттєво спростити задачу виявлення та

співставлення сервісів. А при використанні підходів формального визначення та співставлення веб-сервісів на базі ДЛ – онтологій, можливе автоматичне використання ДЛ представлення такої онтології для вирішення задач із застосуванням резонерів ДЛ.

Однак, ДЛ є досить складним інструментом для самостійної формалізації описів сервісів їх розробниками, але, враховуючі, що описи профілів семантичних веб-сервісів є WSDL документами, які доповнені семантичними анотаціями, тобто це звичайні XML, можливий, та є доцільним, автоматичний розбір цього XML-визначення та побудова онтології конкретного сервісу на основі онтології сервісу верхнього рівня.

Література

1. Sycara K., Widoff S., Klusch M. and Lu J. LARKS: Dynamic Matchmaking Among Heterogeneous Software Agents in Cyberspace. *Autonomous Agents and Multi-Agent Systems*. 2002. Vol. 5. P. 173–203.
2. Karimpour R., Taghiyareh F. Conceptual discovery of web services using WordNet. Service Computing conference, Singapore, Dec. 2009.
3. Ganapathy G., Surianarayanan C. An approach to identify candidate services for semantic web service discovery. IEEE international conference on service-oriented computing and applications, Perth, WA, Dec. 2010.
4. Peng Y. Two levels semantic web service discovery. Seventh international conference on fuzzy systems and knowledge discovery. Yantai, Shandong. Aug. 2010.
5. Becker J., Mueller O., Woditsch M. An ontology-based natural language service discovery engine-design and experimental evaluation. Proceeding of the European conference on information systems. Pretoria, South Africa. 2010.
6. Levy N., Losavio F., Matteo A., Ramdane-Cherif A., Hadj Salem H. Quality standards for ontology web service discovery. International conference with peer: XXXVII CLEI. Quito, Ecuador. October 2011.
7. Xiao H., Zou Y., Ng L., Nigul L. An approach for context-aware service discovery and recommendation. IEEE international conference on web services. Miami, FL. July 2010.
8. TalebiFard P., Leung C.M. A data fusion approach to context-aware service delivery in heterogeneous. The 8th international conference on mobile web information systems. 2011.
9. Rong W, Liu K. A survey of context aware web service discovery: from user's perspective. SOSE '10 proceedings of the 2010 fifth IEEE international symposium on service oriented system engineering. Washington, DC, USA. 2010.
10. Zaremski A. M. and Wing J. M. Signature Matching: a Tool for Using Software Libraries. *ACM Transactions on Software Engineering and Methodology*. 1995. Vol. 4, N 2. P. 146–170.
11. Web Services Description Language (WSDL) (WSDL) <http://www.w3.org/TR/wsdl>
12. <http://dit.unitn.it/~p2p/RelatedWork/Matching/icsoc2004.pdf>
13. The DARPA Agent Markup Language Homepage. <http://www.daml.org/>
14. https://www.researchgate.net/publication/221466239_Semantic_Matching_of_Web_Services_Capabilities
15. Hai Wang, Zengzhi Li. A Semantic Matchmaking Method of Web Services Based On SHOIN+(D). /Institute of Computer System Structure and Networks School of Electronics & Information Engineering, Xi'an Jiaotong University, Xi'an Shaanxi 710049, PR China hwang@mailst.xjtu.edu.cn, lzz@mail.xjtu.edu.cn.
16. Kona S., Bansal A. Gopal Gupta Department of Computer Science. The University of Texas at Dallas Richardson, TX 75083, Thomas D. Hite Metallec Corp. 2400 Dallas Parkway Plano, TX 75093 Automatic Composition of Semantic Web Services.
17. Захарова О. Визначення та вирішення задачі виявлення веб-сервісів за допомогою апарату дескриптивних логік. *Проблеми програмування*. 2017. N 4. P. 66–78.
18. https://www.researchgate.net/publication/221466239_Semantic_Matching_of_Web_Services_Capabilities
19. <http://dspace.nbu.gov.ua/bitstream/handle/123456789/144567/07-Zakharova.pdf?sequence=1>
20. Захарова О. Семантична анотація веб-сервісів. *Проблеми програмування*. 2015. N 4. P. 49–57.
21. Salton G., Wong A. and Yang C.S. A vector-space model for information retrieval. *Journal of the American Society for Information Science*. 1975. Vol. 18. P. 13–620.
22. Voorhees E. Using WordNet for Text Retrieval. C. Fellbaum (ed.), WordNet: An Electronic Lexical Database 1998, The MIT Press, Cambridge, MA. 1999, 285–303.
23. Miller G.A., Beckwith R., Felbaum C., Gross D. and Miller K. Introduction to WordNet: An On-line Lexical Database. *International Journal of Lexicography*. 1990. Vol. 3, N 4. P. 235–244.
24. Miller G.A. Nouns in WordNet: A Lexical Inheritance System. *International Journal of Lexicography*. 1990. Vol. 3, N 4. P. 245–264.
25. Mandala R., Takenobu T. and Hozumi T. The Use of WordNet in Information Retrieval. Proceedings of the COLING/ACL Workshop on Usage of WordNet in Natural Language Processing Systems. Montreal, 1998. P. 31–37.
26. Richardson R. and Smeaton A.F. Using WordNet in a knowledge-based approach to information retrieval. Dublin City University School of Computer Applications Working Paper CA-0395.
27. Putilo J. and Atlee J.M. Module Reuse by Interface Adaptation. *Software Practice and Experience*., 1991. Vol. 21, N 6. P. 539–556.
28. Faloutsos C., and Oard D.W. A survey of Information Retrieval and Filtering Methods, University of Maryland. Technical Report CS-TR-3514, August 1995.
29. Fankhauser P., Kracker M., and Neuhold E.J. Semantic vs. structural resemblance of classes. Special issue: Semantic issues in multidatabase systems. *ACM SIGMOD RECORD*. 1991. Vol. 20, N 4. P. 59–63.
30. Wang Y. and Stroulia E. Flexible Interface Matching for Web-Service Discovery. Proceedings of 4th International Conference on Web Information Systems Engineering, December. 10th - 12th, 2003 (to appear).

References

1. Sycara K., Widoff S., Klusch M. and Lu J. LARKS: Dynamic Matchmaking Among Heterogeneous Software Agents in Cyberspace. *Autonomous Agents and Multi-Agent Systems*. 2002. V. 5. P. 173–203.
2. Karimpour R., Taghiyareh F. Conceptual discovery of web services using WordNet. Service Computing conference, Singapore, Dec. 2009.
3. Ganapathy G., Surianarayanan C. An approach to identify candidate services for semantic web service discovery. IEEE international conference on service-oriented computing and applications, Perth, WA, Dec. 2010.
4. Peng Y. Two levels semantic web service discovery. Seventh international conference on fuzzy systems and knowledge discovery. Yantai, Shandong. Aug. 2010.

5. Becker J., Mueller O., Woditsch M. An ontology-based natural language service discovery engine-design and experimental evaluation. Proceeding of the European conference on information systems. Pretoria, South Africa. 2010.
6. Levy N., Losavio F., Matteo A., Ramdane-Cherif A., Hadj Salem H. Quality standards for ontology web service discovery. International conference with peer: XXXVII CLEI. Quito, Ecuador. October 2011.
7. Xiao H., Zou Y., Ng L., Nigul L. An approach for context-aware service discovery and recommendation. IEEE international conference on web services. Miami, FL. July 2010.
8. TalebiFard P., Leung C.M. A data fusion approach to context-aware service delivery in heterogeneous. The 8th international conference on mobile web information systems. 2011.
9. Rong W, Liu K. A survey of context aware web service discovery: from user's perspective. SOSE "10 proceedings of the 2010 fifth IEEE international symposium on service oriented system engineering. Washington, DC, USA. 2010.
10. Zaremski A. M. and Wing J. M. Signature Matching: a Tool for Using Software Libraries. *ACM Transactions on Software Engineering and Methodology*. 1995. Vol. 4, No. 2. P. 146–170.
11. Web Services Description Language (WSDL) (WSDL) <http://www.w3.org/TR/wsdl>
12. <http://dit.unitn.it/~p2p/RelatedWork/Matching/icsoc2004.pdf>
13. The DARPA Agent Markup Language Homepage. <http://www.daml.org/>
14. https://www.researchgate.net/publication/221466239_Semantic_Matching_of_Web_Services_Capabilities
15. Hai Wang, Zengzhi Li. A Semantic Matchmaking Method of Web Services Based On SHOIN+(D). /Institute of Computer System Structure and Networks School of Electronics & Information Engineering, Xi'an Jiaotong University, Xi'an Shaanxi 710049, PR China hwang@mailst.xjtu.edu.cn, lzz@mail.xjtu.edu.cn.
16. Kona S., Bansal A. Gopal Gupta Department of Computer Science. The University of Texas at Dallas Richardson, TX 75083, Thomas D. Hite Metallec Corp. 2400 Dallas Parkway Plano, TX 75093 Automatic Composition of Semantic Web Services.
17. Zakharova O. Defining and resolving Web-services discovery problems using description logics formalism. *Problems in programming*. 2017. N 4. P. 66–78.
18. https://www.researchgate.net/publication/221466239_Semantic_Matching_of_Web_Services_Capabilities
19. <http://dspace.nbu.gov.ua/bitstream/handle/123456789/144567/07-Zakharova.pdf?sequence=1>
20. Zakharova O. Semantic annotation of web services. *Problems in programming*. 2015. N 4. P. 49–57.
21. Salton G., Wong A. and Yang C.S. A vector-space model for information retrieval. *Journal of the American Society for Information Science*. 1975. Vol. 18. P. 13–620.
22. Voorhees E. Using WordNet for Text Retrieval. C. Fellbaum (ed.), *WordNet: An Electronic Lexical Database 1998*, The MIT Press, Cambridge, MA. 1999, 285–303.
23. Miller G.A., Beckwith R., Felbaum C., Gross D. and Miller K. Introduction to WordNet: An On-line Lexical Database. *International Journal of Lexicography*. 1990. Vol. 3, N.4. P. 235–244.
24. Miller G.A. Nouns in WordNet: A Lexical Inheritance System. *International Journal of Lexicography*. 1990. Vol. 3, N 4. P. 245–264.
25. Mandala R., Takenobu T. and Hozumi T. The Use of WordNet in Information Retrieval. Proceedings of the COLING/ACL Workshop on Usage of WordNet in Natural Language Processing Systems. Montreal, 1998. P. 31–37.
26. Richardson R. and Smeaton A.F. Using WordNet in a knowledge-based approach to information retrieval. Dublin City University School of Computer Applications Working Paper CA-0395.
27. Puri J. and Atlee J.M. Module Reuse by Interface Adaptation. *Software Practice and Experience*. 1991. Vol. 21, N 6. P. 539–556.
28. Faloutsos C., and Oard D.W. A survey of Information Retrieval and Filtering Methods, University of Maryland. Technical Report CS-TR-3514, August 1995.
29. Fankhauser P., Kracker M., and Neuhold E.J. Semantic vs. structural resemblance of classes. Special issue: Semantic issues in multidatabase systems. *ACM SIGMOD RECORD*. 1991. Vol. 20, N 4. P. 59–63.
30. Wang Y. and Stroulia E. Flexible Interface Matching for Web-Service Discovery. Proceedings of 4th International Conference on Web Information Systems Engineering, December. 10th - 12th, 2003 (to appear).

Одержано 02.03.2020

Про автора:

Захарова Ольга Вікторівна,

кандидат технічних наук,

старший науковий співробітник.

Кількість наукових публікацій в українських виданнях – 30.

<http://orcid.org/0000-0002-9579-2973>.**Місце роботи автора:**

Інститут програмних систем НАН України,

проспект Академіка Глушкова, 40.

Тел.: +38(044)526 5139.

E-mail: ozakharova68@gmail.com.