

РОЗРОБКА ІНТЕЛЕКТУАЛЬНИХ ІНФОРМАЦІЙНО-АНАЛІТИЧНИХ ВЕБПОРТАЛІВ НА ОСНОВІ СЕМАНТИЧНИХ ВІКІТЕХНОЛОГІЙ: ПРОБЛЕМИ ТА ПЕРСПЕКТИВИ

Проаналізовано основні функції сучасних вебпорталів та їхні відмінності від інших типів сайтів. Розглянуто тенденції розвитку інформаційно-аналітичних вебпорталів, що стосуються їхньої семантизації, та технології обробки знань, які можуть для цього застосовуватися. Проаналізовано типові проблеми, що виникають у процесі створення та розгортання таких порталів. На прикладі порталу, що базується на семантичних вікітехнологіях, досліджено практичні проблеми впровадження та масштабування засобів та сервісів обробки знань на порталі, запропоновано шляхи їх запобігання та розв'язання.

Ключові слова: інформаційно-аналітичний вебпортал, онтологія, вікітехнологія, сервіси

Вступ

Сьогодні інтелектуалізація інформаційних технологій (ІТ) є одним із базових напрямків їхнього розвитку, який має як широке теоретичне підґрунтя, так і значні практичні результати, пов'язані із розробкою інтелектуальних застосувань у різних сферах. Саме ці засоби семантизації широко застосовуються для вдосконалення та інтелектуалізації таких розповсюджених джерел інформації, як вебпортали. Потреба в інтелектуальних засобах аналізу інформації обумовлюється тим, що зростає обсяг даних, які потребують обробки, стає більше їх зовнішніх і внутрішніх джерел, вони стають складнішими за структурою та різноманітнішими за формами подання, використовують різні схеми індексації (реляційні, багатовимірні, noSQL тощо). Крім того, все більша кількість застосованих систем призначаються зараз для функціонування у відкритому середовищі (Web або його певних частинах).

Одним із найвідоміших проєктів з інтелектуалізації Web-середовища є Semantic Web, в рамках якого розроблено велику кількість стандартів та засобів подання й обробки інформації на рівні змісту, більшість з яких базується на онтологічному аналізі. Тому обробку знань в ІС усе частіше пов'язують зі створенням, використанням та обробкою онтологій, які підтримуються такими стандартами Semantic Web,

як OWL та RDF, а також різноманітними інструментальними засобами. Але, попри це, інтелектуалізація вебзастосувань великого обсягу та зі складною структурою, до яких належать вебпортали, залишається складною та комплексною проблемою, що потребує як розвитку моделей керування знаннями, так і методів їх практичного застосування.

Семантичні системи є підкласом інтелектуальних систем, в яких засоби подання знань явно відокремлені від засобів обробки. Вони можуть отримувати знання, потрібні для функціонування системи та задоволення інформаційних потреб користувачів, з будь-яких зовнішніх джерел та репозиторіїв [1], що відповідають умовам системи. В такому підході довільна зовнішня онтологія розглядається як "чорний ящик", про який невідомо нічого, крім його метаданих (які також можуть бути некоректними – застарілими, помилковими або описаними на основі іншої терміносистеми). Саме тому семантизація систем, орієнтованих на відкрите інформаційне середовище, дозволяє значно розширити функціонал таких застосунків та підвищити ефективність їх роботи, але водночас призводить до багатьох складнощів у процедурах аналізу та інтеграції отриманих відомостей.

Сьогодні користувачі Web мають доступ до величезної кількості вебсторінок і

сервісів, але з усього цього розмаїття більшість з них зазвичай використовують лише обмежений набір інформаційних ресурсів (ІР), що відповідають їхнім інтересам та потребам. Водночас кожен з користувачів використовує різні підмножини сторінок та сервісів (хоча можна визначати певні групи із подібними потребами, що використовують однакові або подібні набори). Тому доцільно виокремлювати такі набори потреб та створювати ІС, які забезпечують доступ до відповідних ресурсів та підтримують сервіси їх обробки. Саме до такого функціоналу ІС, що орієнтовані на функціонування у вебсередовищі, найближче подібне поняття вебпорталу.

Вебпортали дозволяють користувачам отримувати швидкий доступ саме до тих наборів вебсторінок, якими вони зазвичай користуються, і сервісів, орієнтованих на обробку саме таких сторінок. Уся інформація та сервіси, пропоновані на порталі, можуть розглядатися як спроба автоматизовано сформувати та зафіксувати саме таку підмножину інформації, яку обрав би сам користувач, якби мав достатній час і досвід для створення персонального набору вебсторінок [2]. Тому вебпортали доцільно розглядати як засіб персоніфікованої та задача-орієнтованої інтеграції відомостей та послуг, доступ до яких забезпечує вебсередовище. Саме такий підхід дозволяє ефективніше аналізувати потенційні напрямки розвитку вебпорталів, визначати існуючі проблеми та шляхи їх розв'язання на основі сучасних технологій обробки інформації та аналізу даних.

Серед усього розмаїття вебпорталів можна виділити підмножину інформаційно-аналітичних вебпорталів, які забезпечують ширший набір аналітичних сервісів та засобів здобуття знань з наявної сирової інформації. Вони передбачають можливість обробки контенту на рівні знань та використання елементів Data Mining, машинного навчання та логічного виведення. Такі інформаційно-аналітичні вебпортали потребують більш складних моделей контенту для інтеграції, порівняння, узгодження та візуалізації інформації, що надається користувачам.

Як і інші напрямки розвитку інформаційних систем, вебпортали можуть застосовувати засоби інтелектуалізації, менеджменту знань та машинного навчання для забезпечення користувачів більш розширеним функціоналом, однак вони мають враховувати специфіку роботи у відкритому середовищі та потребу в інтеграції різних моделей подання знань. Потрібно розуміти, що інтелектуальні вебпортали є лише досить обмеженою підмножиною всіх існуючих вебпорталів, і не всі з таких порталів можуть бути в майбутньому інтелектуалізовані (через обрані технології розробки, моделі контенту або особливості використання). Крім того, вибір інформаційних технологій, що припускають можливість інтелектуалізації, не забезпечує автоматичну обробку контенту вебпорталу на рівні знань, а лише надає відповідні інструментальні засоби для їх створення.

Основні властивості вебпорталів

Аналізуючи перспективи розвитку та можливості інтелектуалізації інформаційно-аналітичних вебпорталів, необхідно спочатку визначити, яку саме підмножину вебресурсів ми розглядаємо як вебпортали та які спільні властивості та функції відрізняють їх від інших типів сайтів. Наразі створено та розвивається багато різних типів порталів, і тому в літературі зустрічається багато різних визначень та наборів таких характеристик, що дозволяють називати певний вебсайт порталом.

Вебпортал – це сайт, організований як багаторівневе об'єднання різних ресурсів і сервісів, засобів навігації та інформаційних служб. Він може використовуватися як точка входу в мережу, що підтримує специфічні інформаційні потреби окремих користувачів або їх груп.

Можна розглядати портал як певний шлюз для доступу до ресурсів мережі, але зараз майже будь-яка вебсторінка відповідає цьому критерію. У Веб можна знайти велику кількість визначень, наприклад, [3] визначає вебпортал як спеціально розроблений вебсайт, який уніфікує надання інформації з різних джерел на одній платформі.

мі, допомагає в навігації та персоналізації сповіщень із розширеними функціями, такими як керування завданнями, співпраця, бізнес-аналітика тощо.

Основні функції вебпорталу:

– оновлення контенту сайту: портал має містити набір інструментів, які покращують пошук інформації, підвищують її якість;

– покращення взаємодії з користувачами сайту: портал надає додаткові можливості спілкування власників сайту з користувачами, для вдосконалення їхніх взаємин з клієнтами, надаючи високоякісну та відфільтровану інформацію на зручній і комплексній платформі;

– орієнтація на домен: вебпортал пов'язаний з певною предметною областю і має забезпечувати різноманітні потреби користувачів, які шукають відповідну інформацію та послуги;

Розробка та дизайн вебпорталу є складним процесом, оскільки його функціональні можливості досить важко заздалегідь визначити, формалізувати та реалізувати. В цілому процес розробки вебпорталу містить усі кроки, необхідні для розробки будь якого вебсайту. Однак різниця полягає в складності реалізації кожного з цих кроків та в наявності додаткових етапів, що стосуються визначення його функцій, кодування, тестування та інтеграції. Отож, вебпортали відрізняються від інших типів вебсайтів ще й тими аспектами, що останні в основному зосереджені на залученні трафіка, тоді як вебпортали призначені для його скорочення за рахунок ефективнішого та персоналізованого аналізу. Тому мови програмування та фреймворки, які в основному використовуються у розробці вебпорталів, мають певну специфіку та потребують значно ширшого функціоналу, а поширеність їхнього застосування різниться порівняно з використанням для створення інших вебсайтів. У [4] наводять набір найбільш поширених фреймворків для створення вебпорталів: AngularJS, Laravel, React.JS, NodeJS, Ruby on Rails.

Вебпортал забезпечує легкість навігації та підвищення продуктивності в обробці інформації. Він забезпечує універ-

сальний вхід і за потреби забезпечує інтеграцію з іншими програмами та системами.

Хоча можна знайти багато різноманітних характеристик вебпорталів, більшість із них певним чином визначаються саме орієнтацією на деяку обмежену аудиторію користувачів та можливістю надати інтегрований набір найбільш корисних послуг, які підтримують більшість потрібних функцій та не потребують використання зовнішніх сервісів та джерел інформації.

Портали можна розглядати як системи управління знаннями, які дозволяють створювати, обмінюватися та повторно використовувати знання. Вони містять величезну кількість інформації, в тому числі і мультимедійної, що реалізується за рахунок зберігання контенту не в базах даних, а в сховищах і вітринах даних.

Вебпортал є окремим випадком вебсайту. Портал, зазвичай, значною мірою орієнтований на кожного окремого користувача, тоді як вебсайт може не містити елементів персоналізації. Контент вебпорталу є динамічним і часто змінюється. Вебпортал містить елементи захисту від входу в систему і може підтримувати інформаційну взаємодію, що є специфічною для користувача, а його інтерфейс може бути загальним або приватним. Контент порталу збирається з різних джерел і може бути унікальним для користувача на основі його персональних налаштувань. Портали спрямовані на забезпечення різних типів сервісів, які можуть знадобитися користувачам для отримання та обробки його контенту. Тому вибір сервісів, які підтримуються порталами, великою мірою залежить від тематики порталу та потреб його користувачів.

У [5] вебпортал визначають як вебсайт, який спеціально розроблений так, щоб бути єдиною точкою доступу до інформації. Його також можна вважати бібліотекою персоналізованого та категоризованого контенту. Вебпортали допомагають у пошуку та навігації, забезпечують персоналізацію та інтеграції сповіщень, а також часто надають сервіси, що стосуються керування завданнями, співпраці, бізнес-аналітики та інтеграції з іншими застосунками. Досить часто вирізняють відмінності вебпорта-

лів від інших типів сайтів (таблиця 1), що дозволяють визначити, які саме додаткові функції та обмеження потрібні для розробки таких порталів та які переваги надають ці додаткові ускладнення.

Таблиця 1. Відмінності вебпорталів від інших типів сайтів

Вебсайт	Вебпортал
Доступ через унікальну адресу URL	Також використовують для доступу пароль та логін
Всі бачать весь контент	Тільки члени спільноти користувачів, що мають паролі, отримують доступ до всього контенту
Всі користувачі отримують однаковий контент	Контент відрізняється для різних груп користувачів
Сайтом керує єдина команда адміністраторів	Для різних рівнів роботи з контентом застосовують різні команди адміністраторів
Контент є досить стабільним	Динамічний контент часто змінюється, що потребує спеціальних інструментів
Розрахований на анонімних користувачів та публічний доступ	Дозволяє персоніфікувати контент та сервіси, що доступні користувачам, підтримує різні рівні доступу до інформації з різних джерел

Зазвичай, вебпортали забезпечують доступ із різних платформ та видів пристроїв, таких як персональні комп'ютери та смартфони.

Вебпортал можна розглядати як сайт, що надає інформацію та сервіси для використання цієї інформації.

Тенденції розвитку порталів пов'язані із вдосконаленням усіх цих типів сервісів та зі створенням інформаційної інфраструктури порталу, яка забезпечує більш ефективно виконання цих сервісів.

Класифікація вебпорталів

Існує багато типів класифікації порталів, найбільш поширеними з яких є:

- за цільовою аудиторією;
- за типом контенту;
- за призначенням;
- за технічними ознаками.

За цільовою аудиторією порталів: публічні і корпоративні. Публічні портали орієнтовані на всіх користувачів, а корпоративні – тільки на певну підмножину, наприклад, співробітників і партнерів організації. Крім цього, портали також поділяють за іншими параметрами на інтернаціональні та регіональні. Інтернаціональні портали орієнтовані на всіх користувачів, незалежно від місця їх проживання і мовної приналежності, тоді як регіональні орієнтовані на користувачів із певної місцевості або мовної групи. Приклади горизонтальних порталів — Yahoo! (англомовний), Ukr.net.

За типом контенту вебпортали поділяють на горизонтальні і вертикальні: перші націлені на великі спільноти користувачів, другі ж більш специфічні щодо вмісту та об'єктів. Вертикальні портали містять контент, що стосується однієї теми, а горизонтальні портали охоплюють одночасно багато різних тем, але надають однаковий набір сервісів для обробки інформації з різних напрямків. Вертикальний портал – це точка доступу до певної галузі, ринкової ніші чи сфери інтересів, що стосуються обмеженої тематики, торкаючись лише однієї теми або одного типу тем, таких як страхування, освіта або охорона здоров'я. Наприклад, вертикальний портал www.contractor.com містить інформацію та сервіси для будівельної галузі.

Тенденції розвитку вебпорталів

Основні напрямки розвитку вебпорталів можна поділити на кілька напрямків (хоча на практиці вони досить часто використовуються разом та у різних комбінаціях): персоніфікацію, вдосконалення сервісів та інтелектуалізацію.

Тенденції розробки інформаційно-аналітичних вебпорталів теж стосуються цих трьох напрямків, але передбачають більшу

увагу до визначення набору сервісів, що надає портал, та додаткові вимоги до подання контенту, які підтримують виконання цих сервісів.

Багато підходів та визначень підкреслюють важливу роль персоніфікації інформації на вебпорталах та можливість застосування знань із зовнішніх джерел та досвіду роботи самого порталу. Така персоніфікація передбачає індивідуальні налаштування та надання індивідуалізованого набору послуг. У найбільш розвинутих порталах – можливість враховувати досвід взаємодії з користувачем та навчатися на цьому досвіді. Але цей параметр більш важливий для вертикальних порталів, де основні відмінності між користувачами стосуються саме вибору підмножини контенту та засобів його обробки. Для горизонтальних порталів це менш важливо і багато в чому залежить від спрямованості порталу. Наприклад, для інформаційних та енциклопедичних порталів персоніфікація зазвичай стосується визначення повноважень, створення персональної сторінки та збереження історії запитів.

Тому питання персоніфікації пов'язане із захистом інформації вебпорталу, а саме – з визначенням набору функцій, що дозволяються певному користувачеві, як щодо перегляду контенту, так і щодо можливостей його редагування.

Сучасні інформаційно-пошукові системи (зокрема, Google) та сервіси прагнуть накопичувати й використовувати досвід взаємодії з конкретним користувачем. Перспективні напрямки розвитку вебпорталів мають також підтримувати ці можливості та забезпечувати доступ як до зовнішніх пошукових сервісів, так і надавати розширений функціонал пошуку в контенті порталу з урахуванням індивідуальних потреб користувача та досвіду його взаємодії з порталом.

Для того, щоб формалізувати інформаційні потреби користувачів та описати структуру й властивості потрібних йому інформаційних об'єктів (ІО), на порталі можуть використовувати такі джерела знань про учасників процесу пошуку:

– *персональний профіль користувача*, у якому містяться відомості про індиві-

дуальні особливості та переваги користувача у споживанні інформації: дані, отримані при реєстрації користувача; результати його анкетування; заявлений користувачем рівень компетентності в різних Про, історію його попередніх запитів, модель Про його інтересів тощо;

– *знання щодо предметної області пошуку* (наприклад, онтологія відповідної Про, тезаурус задачі, посилання на зовнішні БЗ);

– *досвід взаємодії користувача з порталом*: історії запитів цього та інших користувачів, чимось подібних до нього;

– *зовнішні джерела знань*, обрані користувачем і доступні через Web (наприклад, Вікіпедію, словники, Web-сайти).

Значною мірою можливості семантизації порталу визначаються: загальною концепцією розробки порталу, його цілями та призначенням; обраним програмним забезпеченням; моделлю та структурою бази знань порталу; виразною здатністю реалізованих пошукових сервісів; засобами та методами структурування контенту порталу, можливостями експорту зовнішніх баз знань.

Але слід зазначити, що вдалий вибір моделі подання контенту, набору сервісів та їх програмної реалізації не є автоматичною запорукою створення дійсно ефективного семантичного порталу. Необхідно враховувати, що вебпортал функціонує у відкритому інформаційному середовищі, що безперервно змінюється, і тому розробка порталу потребує постійного оновлення та інтеграції як моделі знань, що використовуються на порталі, так і самого контенту, щоб забезпечити його актуальність та несуперечність.

Набір сервісів, що надається вебпорталом, значною мірою визначає його функціонал та сферу застосування. Тому тенденції розвитку вебпорталів мають враховувати розширення набору таких сервісів та вдосконалення якості їхньої роботи. Саме сервіси вебпорталів є одним із важливих критеріїв їх порівняння. Важливо відрізняти, які сервіси дійсно підтримуються на порталі, а які лише декларуються або надаються з обмеженим функціоналом.

Багато дослідників пропонують різноманітні методи, що дозволяють ранжувати вебпортали на основі їхніх умов та політики надання послуг. Наприклад, в [6] розглядаються методи, які застосовують машинне навчання та нейронну мережу обробки природної мови, для порівняння сайтів. Для цього аналізуються умови надання сервісів та угоди, що визначають набір правил, які повинен виконати користувач для того, щоб використовувати послугу, що надається вебсайтом або пошуковою системою.

Напрямки семантизації вебпорталів

Основні тенденції розвитку порталів стосуються їхньої семантизації та можливості повторного використання інформації. Обидві ці тенденції безпосередньо пов'язані із проектом Semantic Web та широко застосовують стандарти й мови подання знань, створених в рамках цього проекту. Технології Semantic Web (SW) [7] і Linked Data [8] є багатообіцяючим підходом для вирішення проблем семантичної сумісності в середовищі створення розподіленого контенту. На основі комп'ютерно «зрозумілих» зв'язаних великих даних можна створювати інтелектуальні програми для дослідників цифрових гуманітарних наук (DH) і громадськості.

Наприклад, у [9] автори розглядають набір семантичних порталів, що підтримує поєднання портальних технологій з засобами та стандартами обробки розподілених знань та забезпечує створення національної онтологічної інфраструктури культурної спадщини та семантичних порталів “Sampo” (таких як BookSampo, BiographySampo, AcademySampo) на основі пов'язаних даних (Linked Data), технологій Semantic Web та принципів відкритої науки FAIR [10]. Модель Sampo містить три основні компоненти: 1. бізнес-модель для узгодження, агрегування та публікації різноманітного, розподіленого контенту на основі спільної інфраструктури онтології; 2. дизайн інтерфейсу, де дані можна повторно використовувати та отримувати доступ незалежно від різних точок програми, тоді

як дані знаходяться в одній кінцевій точці SPARQL; 3. двоетапну модель для доступу та аналізу даних, де фокус інтересу спочатку фільтрується за допомогою фасетного семантичного пошуку, а потім візуалізується та аналізується за допомогою інструментів порталу.

У [11] наводяться базові вимоги до технологій створення семантичних порталів. Бажані функції семантичного порталу разом із вимогами до інструментів для їх створення, які поділяються на дві групи:

– *вимоги, яким повинні відповідати інструменти для реалізації інтерфейсів користувача семантичного порталу :*

– швидке створення застосунків для конкретних випадків використання;

– можливість запитувати дані безпосередньо з кінцевої точки SPARQL [12];

– дослідження ієрархічних даних за допомогою онтологій;

– підтримка масштабування для обробки великої кількості ІО за обраним сценарієм, використовуючи обмежену кількість ресурсів;

– інтуїтивно зрозумілих і зручних інтерфейсів користувача;

– можливість візуалізацій, які автоматично адаптуються до доступних ресурсів;

– *вимоги, які стосуються структури та функціональних можливостей інтерфейсів користувача семантичних порталів:*

– наявність кількох точок зору для доступу до графів знань або онтологій, що відображають структуру бази знань вебпорталу;

– можливість перегляду, пошуку та дослідження структури бази знань вебпорталу;

– підтримка різних парадигм пошуку в контенті вебпорталу (наприклад, шукати екземпляри певного класу з подібними значеннями властивостей);

– наявність різних форм подання результатів пошуку (як-от, на географічних картах, на часовій шкалі або у вигляді графічної діаграми);

– можливість перегляду та аналізу структури ІО порталу та його конкретних екземплярів;

- користувачі можуть досліджувати посиланнями між екземплярами Ю ;
- користувачі можуть експортувати результати пошуку у різні формати відповідно до їхніх потреб.

Технології Semantic Web для керування знаннями вебпорталів

Проект Semantic Web розвивається вже багато років та залишається на сьогодні джерелом найбільш ефективних технологій керування знаннями у вебсередовищі. Основна ідея цього проекту пов’язана з прагненням перетворити Web на глобальну базу знань та з розробкою стандартів, мов і програмних засобів для семантичного пошуку інформації в цих ресурсах. Саме ці засоби дозволяють розширювати функціонал вебпорталів та інтегрувати їх з іншими ресурсами.

Semantic Web – це потужний напрям для підвищення ефективності розподіленого й сумісного доступу до інформації та її використання прикладними програмами. Це досить ефективна спроба інтеграції застосовних Web-технологій з науковими розробками у сфері подання знань (knowledge representation, KR), що є підгалуззю штучного інтелекту. В рамках проекту Semantic Web створюються засоби для керування великою кількістю інформації та інформаційних послуг. Урахування семантики інформаційних ресурсів дає змогу зробити інформацію більш корисною. Але це потребує створення нових засобів подання даних, які покращують здатність до збирання й спільного використання знань, а також нових конструкцій програмування та інструментів, що уможливають обробку цієї інформації в прикладних програмах.

Використання елементів Semantic Web дозволяє підтримувати важливі для інформаційно-аналітичних порталів функції:

- індексація та пошук інформації;
- розробка та підтримка метаданих;
- розробка та підтримка методів анодування;
- подання контенту у вигляді інтероперабельної бази даних;
- організація машинного здобуття даних;

- виявлення й надання вебсервісів;
- інтеграція та аналіз контенту;
- дослідження в галузі інтелектуальних програмних агентів;
- розвиток підтримки інтерфейсів різних пристроїв і всепроникні обчислення.

Основними компонентами Semantic Web, які безпосередньо стосуються розвитку порталних технологій, є онтології та вебсервіси. Для того, щоб використовувати їх у застосовних програмах, у рамках проекту Semantic Web розроблено такі відкриті стандарти, що дають змогу формалізувати семантику IP і програмних засобів їх пошуку та обробки: RDF – мова опису метаданих; OWL – мова подання онтологій; SPARQL – мова запитів до RDF та OWL; SWRL – мова визначення правил. Semantic Web продовжує активно розвиватися, про що свідчать розробка й удосконалення мов і стандартів роботи з розподіленими знаннями.

У контексті Semantic Web вебпортал розглядається як вебсайт, що надає інформацію та забезпечує можливості обміну для певної спільноти на основі технологій Semantic Web [13]. У такому порталі вирізняють три рівні (рис.1): 1. доступ до інформації для користувачів; 2. засоби обробки інформації на порталі; 3. базові технології.

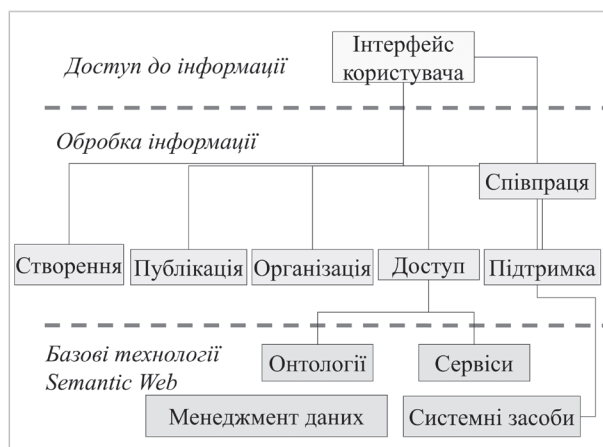


Рис.1. Рівні вебпорталу Semantic Web

Елементи, що використовуються на кожному рівні, можуть значно різнитися залежно від реалізації. Наприклад, можуть використовуватися різні типи онтологій, що можуть мати різний обсяг та виразність,

але вони мають застосовуватися замість неформалізованих природномовних описів предметної області вебпорталу. Залежно від типу онтології можуть бути реалізовані різні види логічного виведення (перевірка наявності елементів класу, виявлення симетричних або транзитивних відношень між ІО тощо), але портал має підтримувати сервіси для здобуття та аналізу поданих в онтології знань.

Фундаментальна семантична проблема публікації, інтеграції, аналізу та використання на основі порталів полягає в тому, як зробити різноманітну інформацію семантично сумісною, щоб можна було шукати потрібні користувачам відомості, з'єднати їх між собою та представляти узгодженим способом із різних наборів даних. Проблема пов'язана зі способом створення контенту порталів: дані збираються, зберігаються та публікуються різними суб'єктами, які використовують власні стандарти та найкращі практики, що можуть бути несумісними між собою. Важливо розуміти, що використання для цього методів створення інтелектуальних інформаційно-аналітичних систем та впровадження вже існуючих семантичних вебтехнологій, має враховувати такі важливі аспекти функціонування вебпорталів, як динамічність контенту, можливість масштабування як сервісів, так і моделей подання знань.

У семантизації порталів можна виділити кілька досить тісно пов'язаних аспектів: 1. семантичний пошук; 2. імпорт знань із зовнішніх інформаційних джерел; 3. експорт інформації з порталу в сучасних форматах подання знань (таких як RDF та OWL); 4. створення метаданих, що описують контент порталу на семантичному рівні.

Наразі швидко зростання обсягу інформації, доступ до якої забезпечують портали, а також ускладнення її структури зумовлюють потребу в розвитку різноманітних сервісів для знаходження необхідних користувачам відомостей [14]. Ці сервіси мають позбавити користувачів рутинного аналізу інформації, яка не стосується їхніх інтересів, та виокремлювати відомості щодо ІО, потрібних цим користувачам.

Використання порталів має забезпечувати більш ефективний пошук порівняно зі спробами шукати інформацію безпосередньо за місцем її збереження – у базах даних та інших джерелах структурованої інформації (наприклад, XML-документах і сховищах) або на тих вебсайтах, що застосовуються як джерела контенту порталу. Для кожного типу інформаційних джерел і кожного типу ІО, які шукає користувач, доцільно мати свій набір пошукових компонентів з уніфікованим інтерфейсом взаємодії.

Якщо в процесі пошуку інформації використовують зовнішні знання щодо предметної області, яка цікавить користувача, та засоби керування знаннями, то йдеться про *семантичний* або *інтелектуальний* пошук.

Семантичний пошук – це один з підкласів інформаційного пошуку, в якому релевантність знайденої інформації до запиту користувача оцінюється за близькістю змісту для ефективнішого задоволення інформаційних потреб користувача. Такий пошук іноді розглядають як інтелектуальну надбудову над традиційним інформаційним пошуком. Тож можна вважати, що сервіси семантичного пошуку на порталі розширюють його функціонал та доповнюють традиційні пошукові сервіси (за ключовими словами, категоріями тощо). Наприклад, у порталах на основі семантичних вікітехнологій підтримуються рідні види пошуку на основі семантичних властивостей сторінок та їхніх значень.

Для семантичного пошуку на вебпорталі, що виконується у відкритому середовищі, характерним є те, що потрібні для пошуку знання також динамічно здобуваються з цього відкритого середовища, а не закладаються в систему в процесі її створення. Те, які саме знання використовуються для цього, залежить як від специфіки та типу порталу, так і від концепції, обраної його розробниками.

Результати пошуку можуть автоматично вбудовуватися до відповідних сторінок порталу. Зокрема, портали на основі семантичних вікітехнологій можуть містити знайдені дані для автоматичного оновлення контенту (рис.2).

Статті автора у друкованій версії	Статті автора на порталі
Цим автором підготовлені такі статті друкованої «Великої української енциклопедії»:	Цим автором підготовлені такі статті для електронної версії «Великої української енциклопедії»:
Аве Марія	Авгури
Авраам	Аве Марія
Адат	Авраам
Алі ібн Абу Таліб	Адат
Аллаг	Алі ібн Абу Таліб

Рис. 2. Результати виконання вбудованого запиту для генерації контенту

Слід враховувати, що семантичний пошук на вебпорталі зазвичай забезпечує знаходження екземплярів типових ІО, характерних для даного порталу, а також специфічних сукупностей таких об'єктів (наприклад, портал транспортних послуг дозволяє знаходити узгоджену сукупність переміщень між обраними пунктами, а портал освітньої установи дозволяє знаходити навчальний курс з обраним набором компетенцій).

Під час створення таких систем варто враховувати, з яких саме зовнішніх баз знань можуть здобуватися відомості щодо структури таких ІО, та регулярно перевіряти, чи не змінюється контент цих баз.

БЗ вебпорталу можуть змінювати структуру й доступність незалежно від розробників ІПС. Тому результати пошуку у Web є функцією ще й від вмісту зовнішніх БЗ.

Підтримка парадигми FAIR у контенті вебпорталів

Одним із перспективних напрямків вдосконалення вебпорталів є можливість забезпечення повторного використання їхнього контенту на основі парадигми відкритих даних. Це потребує розв'язання різноманітних проблем, що стосуються збереження великих масивів даних, дотримання прав власників цих даних, забезпечення безпечного інформаційного і правового середовища для користувачів тощо. Ці проблеми можуть бути досить ефективно вирішені для обробки результатів наукових досліджень. Але технології, що підтримують пошук та повторне використання специфічних наукових ресурсів, виявляються при-

датними і для вдосконалення порталних технологій. Наприклад, FAIR дані (FAIR_data) – це дані, які відповідають принципам знаходжуваності, доступності, інтеоперабельності та повторного використання [15]. Якщо вебпортал надає дані та сервіси, що відповідають цій парадигмі, то це гарантує користувачам більший функціонал для аналітичної обробки та практичного застосування інформації.

Такі дані мають наступні властивості:

- *Findable*: щоб використовувати дані, їх необхідно спочатку знайти там, де вони зберігаються. Тому дані та їх метаописи мають бути доступними як для людей, так і для комп'ютерних програм.
- *Accessible*: Коли користувач знаходить потрібні дані, він/вона повинен знати, як отримати доступ до них (можливо, включаючи аутентифікацію та авторизацію). І дані, і їхні метаописи мають бути зрозумілими людині та придатними для автоматизованої обробки.
- *Interoperable*: дані мають бути придатними для інтеграції з іншими даними та для обробки іншими застосунками.
- *Reusable*: дані та метадані повинні бути описані з відповідною виразністю, щоб їх можна було відтворювати та/або комбінувати в різних налаштуваннях.

Згідно FAIR, сервіси пошуку, здобуття і представлення даних реалізують не користувачі, а інформаційна система. Водночас мова йде не тільки про власне дані і метадані, а і про алгоритми й інструменти керування ними.

Керівні принципи FAIR не потребують будь-якої стандартизації чи конкретної технології їхньої підтримки. Тому вони досить незалежні від конкретних програмних рішень, що можуть застосовуватися для їхньої реалізації.

Практична реалізація нових тенденцій у створенні вебпорталів

Сьогодні один із перспективних напрямків створення семантичних вебпорталів з відкритими даними базується на

вікітехнологіях та їх семантичному розширенні. Прикладом реалізації такого підходу є портал Великої української енциклопедії (vue.gov.ua). Зараз цей портал відповідає багатьом вимогам, пов'язаним з тенденціями розвитку порталів, але потребує подальшої роботи та підтримання структури та контенту у відповідності до змін відкритого середовища та змін в інформаційних ресурсах, котрі є джерелами його контенту.

Семантизовані Wiki-ресурси, такі як Semantic MediaWiki, що дозволяють створювати семантичні дані та базуються на використанні стандартів Semantic Web, дають потужне рішення для спільного редагування даних та їхніх метаописів, створення різних довільних наборів властивостей у шаблонах цих метаописів, з одночасним поданням їх як в машинно-оброблюваній формі, так і формі, придатній для розуміння людиною, що в результаті дає можливість оперувати цими даними, автоматизовано керувати, здійснювати аналіз, публікувати.

Використання вікітехнології на основі MediaWiki [16] та його семантичного розширення Semantic MediaWiki [17], яке історично себе показало досить потужним інструментом, має широку спільноту використання, відкритий код та відкриті принципи розробки і підтримки, постійно розвивається, забезпечує відповідність створеного інформаційного ресурсу вимогам FAIR [18] та надає достатній функціонал для роботи з контентом [19].

Але потрібно зауважити, що ускладнення структури portalу та розширення його функціоналу може викликати появу багатьох додаткових проблем: обробка контенту на рівні семантики є значно складнішою за традиційну, а орієнтація на відкрите середовище вимагає враховування різноманітних форм подання інформації, знаходження шляхів розв'язання неоднозначностей, виявлення помилок та визначення процедур інтеграції як даних, так і їхніх моделей [20], а також більш якісно обирати інформаційні джерела, які забезпечують якісні дані та застосовують релевантні формати їх представлення.

Як бачимо, сучасні тенденції потребують від розробників portalу значно вищої кваліфікації та більшого часу на генерацію

контенту, адже додавання нової інформації має супроводжуватися її семантичною розміткою, структуризацією та узгодженням з уже наявними даними. Але створений таким чином контент є значно зручнішим для користувачів, тому що забезпечує автоматизоване виконання різноманітних сервісів, пов'язаних із аналітичною обробкою та узагальненням даних. Крім того, такі портали дозволяють генерувати персоніфіковані бази знань за запитам користувачів, і саме це дозволяє користувачам ефективніше реалізувати свої інформаційні потреби у взаємодії з іншими інтелектуальними застосунками.

Критерії оцінки можливостей розвитку вебпорталів

Як показує наведений вище аналіз перспективних напрямків розвитку вебпорталів, вдосконалення вже існуючих порталів базується на наявності:

- відкритого програмного забезпечення, що підтримує масштабованість контенту та має достатню виразність для подання інформації зі складною структурою;
- моделі представлення знань portalу, що сумісна із технологіями менеджменту знань у відкритому середовищі та підтримує імпорт та експорт знань із зовнішніх інформаційних ресурсів;
- семантичного структурування контенту та створення мета даних, що характеризують цей контент;
- засобів перевірки узгодженості контенту та його автоматизованої генерації на основі семантичної розмітки;
- набору сервісів для пошуку та навігації на порталі, для аналізу контенту та його автоматизованої генерації за запитам.

Важливо розрізнити потенційні технологічні можливості та їхню практичну реалізацію. Наприклад, використання семантичних розширень вікітехнологій забезпечує можливість семантичної розмітки контенту та явного визначення змісту посилань між окремими сторінками. Але ці визначення потрібно створювати (і цю роботу мають виконувати експерти предметної області, що розуміють зміст таких зв'язків, а

не оператори). Саме наявність повноцінної семантичної розмітки є необхідною умовою ефективного виконання різноманітних пошукових запитів.

Важливо розрізняти потенційні технологічні можливості розвитку вебпорталів та їхню практичну реалізацію.

Розглянемо це на прикладі порталу Великої української енциклопедії (vue.gov.ua). У Таблиці 2 наведено технологічні можливості семантичного вікісередовища, в якому реалізовано цей портал, та ступінь їхнього практичного використання.

Таблиця 2. Реалізація тенденцій розвитку вебпорталів

Напрямок розвитку	Практична реалізація
Персоніфікація	Тільки для розробників контенту
Захист персональних даних	
– можливість перегляду	+ (простори імен сторінок)
– можливість редагування	+ (тільки для зареєстрованих користувачів, але без градації повноважень)
Експорт знань у відкриті формати	Потрібно писати та виконувати запити сторінок, де такі запити вбудовано
Імпорт знань із зовнішніх джерел	Зараз не підтримується (було у попередніх версіях)
Розширений набір сервісів навігації	
– спеціальні сторінки	Для зареєстрованих користувачів
– інтерфейс порталу	обмежено
Сервіси пошуку	
– спеціальні сторінки	+
– інтерфейс порталу	частково
Семантична розмітка контенту	

– типових ІО	частково
– довільних ІО та посилань між сторінками	Менше 1%
Семантичний пошук	
– типових ІО	частково
– довільних ІО	Менше 1%
Підтримка парадигми FAIR	Зараз не використовується
Інтелектуальні сервіси аналізу контенту	Тільки у тестових версіях

Тож бачимо, що такий портал потенційно має великі можливості для вдосконалення, але наразі значна частина таких можливостей не реалізована практично (або реалізована лише частково). Слід відзначити, що семантичне розширення вікітехнологій на основі плагіну Semantic MediaWiki підтримує більшість тенденцій розвитку вебпорталів, але сама технічна можливість не забезпечує автоматично відповідний функціонал. Наявність семантично розміченого контенту є необхідною умовою розвитку порталу та можливості створення набору сервісів, що задовольняють потреби користувачів.

Проблеми у створенні семантичних вебпорталів

У створенні та розвитку семантичних вебпорталів виникає ціла низка проблем, специфічних саме для систем обробки та аналізу знань, які оперують з великими обсягами інформації та з гетерогенними моделями ІО. Їх умовно можна поділити на такі категорії: 1. проблеми моделювання та подання семантики; 2. проблеми масштабування; 3. організаційні проблеми.

Значна частина проблем подання семантичного контенту пов'язана з необхідністю обробки даних, які є неточними, розпливчастими, невизначеними, непослідовними, неповними тощо і тому не можуть бути віднесені до класичних даних. Цей термін близький до брудних даних, але для систем обробки знань вони доповнюються семантично некоректними даними різних типів. Дані вважаються брудними, якщо

програма, що працює коректно, не в змозі отримати результат їхньої обробки, або отримувє неправильний результат. Водночас потрібно аналізувати два різні аспекти – з якої причини дані стали брудними та що можна зробити, щоб вони стали придатними для аналізу. Наприклад, джерелами брудних даних можуть бути помилки введення або оновлення даних, помилки передавання даних або некоректно обрана форма подання даних. Семантично некоректні дані можна виявити тільки на етапі їх інтерпретації, якщо їхні значення не відповідають обмеженням та правилам предметної області (наприклад, у загальному випадку вік людини може бути довільно позитивним числом, але вік працівника не може бути менше певного значення). Для семантичних даних виникає також необхідність перевірки змістовної узгодженості даних щодо ІО, між якими існують певні семантичні зв'язки (як-от, рік народження дитини має бути менше року народження її батьків).

Причини появи на порталі такої інформації можуть бути різними. В деяких випадках інформація може перетворюватися у класичні дані автоматично або автоматизовано, в інших потребує участі експерта предметної області або інженера зі знань (наприклад, для інтеграції різних онтологічних моделей однієї області). В деяких випадках дані не можуть бути перетворені на класичні через відсутність, неповноту або надлишковість інформації. Класифікація таких даних для семантичного вебпорталу на основі Wiki-технології наводиться у [19].

Участь великої кількості спеціалістів у створенні вебпорталів викликає розбіжності у розумінні правил подання та структурування даних. Тому виникає не тільки необхідність додаткової верифікації контенту, а й створення детальної методології семантизації та узгодження з моделлю предметної області. Для цього потрібно створювати формалізовані та масштабовані рішення для опрацювання різноманітних типів нечіткості, неповноти та семантичної некоректності контенту. Однак наявність таких рішень та методологій є не достатньою, а лише необхідною умовою побудови семантичного вебпорталу.

Як показує практика розробки семантичних вебпорталів, проблеми обробки знань значно різняться на етапі створення й розгортання та на етапі масового поповнення контенту. Якщо в першому випадку більшість проблем пов'язані із невідповідністю моделі об'єктам реального світу та потребують внесення принципів змін (наприклад, припускати кілька значень певного параметру, зміни структури), то у другому випадку такі моделі потребують уточнення та розширення, щоб коректно обробляти більш специфічні випадки, які не враховували початкові варіанти.

Проблеми масштабування виникають у процесі ускладнення бази знань порталу, коли інформація про взаємодію великої кількості ІО різних типів має бути інтегрованою та узгодженою. Як правило, існуючі порталні технології не підтримують такі функції безпосередньо, але допускають можливість застосування зовнішніх сервісів, призначених для менеджменту знань. Інша проблема масштабування пов'язана з тим, що зі зростанням кількості складних ІО кількість зв'язків між ними може зростати нелінійно, й це може призвести до збільшення часу аналізу інформації та виконання пошукових запитів. У таких випадках потрібно знаходити певний компроміс між прийнятним часом обробки та необхідною виразністю моделі. Крім того, деякі складні для виконання аналітичні запити доцільно виносити на окремі сторінки, а не пропонувати усім користувачам.

Організаційні проблеми пов'язані з тим, що на початку розробки семантичного вебпорталу основні роботи, як правило, виконує команда кваліфікованих спеціалістів з менеджменту знань. Але після формування основи структури бази знань порталу обсяг робіт для таких спеціалістів зменшується, і це може призвести до спроб зменшення такої команди через поєднання функцій або припинення розвитку певних напрямків семантизації. Об'єднуючи дані з різних джерел, можна отримувати нову інформацію за допомогою різноманітних засобів аналізу даних. Але якість нової інформації залежить не лише від алгоритмів аналізу, а й від якості даних. Внаслідок цього через деякий час структура бази знань не оновлюється та пе-

рестає відповідати потребам доданого контенту, нові працівники порталу з недостатньою кваліфікацією некоректно вносять семантичну розмітку (або не вносять її взагалі, тому що це не впливає явно на формальні характеристики порталу). Через це семантичні запити та аналітичні сервіси працюють некоректно (вони обробляють некоректно задані дані або не мають потрібних відомостей). І тому портал, що формально вважається семантичним, втрачає значну частку своїх переваг як аналітична інформаційна система, але надалі продовжує витрачати додаткові обчислювальні потужності на підтримку семантичної складової.

Практичні рекомендації зі створення семантичних вебпорталів

На основі наведеного вище аналізу можна сформулювати наступні практичні рекомендації зі створення семантичних вебпорталів, спрямованих на запобігання та розв'язання проблем обробки знань:

1) модель бази знань вебпорталу, створену на основі початкового набору контенту, потрібно постійно вдосконалювати та узгоджувати з усіма особливими випадками інформаційних ІО, які могли не зустрічатися у початковому наборі даних;

2) моделі ІО бажано не змінювати, а доповнювати додатковими властивостями, а типи даних, що використовуються для опису їх значень, змінювати на ширші, а не на більш специфічні (наприклад, замінювати “Число” на “Текст”, але не навпаки) – це має забезпечити узгодженість моделі з попередньо оприлюдненим контентом;

3) необхідно постійно генерувати семантичну розмітку контенту, яка є основою функціонування усіх семантичних сервісів: потрібно явно визначати семантику всіх (або хоча б більшість) зв'язків між елементами контенту;

4) набір семантичних відношень між елементами контенту має постійно поповнюватися, щоб коректно відображати все розмаїття таких зв'язків, але разом із зростанням кількості таких відношень доцільно розробити систему їх класифікації (наприклад, таксономію або тезаурус), надавати

чіткі формальні описи та намагатися запобігти дублюванню;

5) виконання складних аналітичних та пошукових запитів доцільно виносити в окремі сторінки або пункти меню, щоб виконувати їх тільки тоді, коли це потрібно користувачам. Це дозволяє економити обчислювальні ресурси порталу;

6) постійно проводити інструктування персоналу, який готує контент, щоб забезпечити коректність створюваної семантичної розмітки, та чітко розрізняти функції зі створення моделі такої розмітки та її застосування (для складних порталів це підтримується за рахунок детальнішого визначення повноважень різних груп розробників);

7) здійснювати регулярні операції для оновлення структури бази знань, створення її резервних копій та контролю змістовної цілісності (як з використанням внутрішніх можливостей порталу, так і за допомогою зовнішніх систем менеджменту знань), для запобігання появи значних семантичних некоректностей даних.

Висновки

Основна мета даної роботи полягає в здійсненні наукових досліджень та експериментальних перевірок отриманих результатів у створенні інтелектуальних інформаційних систем, які використовують Semantic Web, онтологічний аналіз та вікітехнології для забезпечення можливості аналізу та використання для практичних цілей великих обсягів і потоків інформації різних типів. У роботі аналізуються вебпортали, які є одним з перспективних засобів інтеграції та аналізу великих обсягів інформації, переробки масивів гетерогенних «сирих» даних у форму, що забезпечує їх уніфіковане подання, здобуття з них корисних знань та повторне використання. Розглядаються можливості, які надає семантизація вебпорталів. Розвиток таких систем потребує поєднання портальних технологій з можливостями засобів аналізу та обробки знань у відкритому середовищі, і таке поєднання призводить до появи специфічних проблем.

Наступним кроком розвитку вебпорталів можуть стати портали випадкового виявлення знань, де інструменти базують-

ся на техніках штучного інтелекту і здатні автоматично знаходити випадкові цікаві явища та дослідницькі питання в даних і вирішувати проблеми користувачів за допомогою пояснень.

Література

1. *Гладун А. Я., Рогущина Ю. В.* Репозитории онтологии как средство повторного использования знаний для распознавания информационных объектов. *Онтология проектирования*, 1 (7), 2013, с.35-50.
2. *Strauss H.* Web portals: the future of information access and distribution. In *Transforming Serials*, 2021, pp. 27-35.
3. What is the Difference between Website and Web Portal. www.geeksforgeeks.org/difference-website-web-portal/ [Accessed 4/09/2023].
4. Difference between Website and Web Portal. forum.huawei.com/enterprise/en/Difference-between-Website-and-Web-Portal [Accessed 4/09/2023].
5. Web Portal. <https://www.techopedia.com/definition/17352/web-portal> [Accessed 4/09/2023].
6. *Bhat M. U. D., Singh S.* Automatic Method to Rate Websites Based on Terms of Services. *Journal Of Creative Writing*, 4(2), 2020.
7. *Berners-Lee T., Hendler, J., Lassila, O.* The semantic web. *Scientific american*, 284(5), 2001, pp.34-43.
8. *Bizer C., Heath T., Berners-Lee, T.* Linked data: The story so far. In *Semantic services, interoperability and web applications: emerging concepts*, 2011, pp. 205-227.
9. *Hyvönen E.* “Sampo” model and semantic portals for digital humanities on the Semantic Web. In *Proceedings of the Digital Humanities in the Nordic Countries 5th Conference (DHN 2020)*, 2020, pp.374-378.
10. FAIR_data. https://en.wikipedia.org/wiki/FAIR_data. [Accessed 4/09/2023].
11. *Ikkala E., Hyvönen E., Rantala H., Koho M.* Sampo-UI: A full stack JavaScript framework for developing semantic portal user interfaces. *Semantic Web*, 13(1), 2022, pp.69-84. <https://content.iospress.com/articles/semantic-web/sw210428>.
12. *Koho M., Heino E., Hyvönen E.* SPARQL Faceter-Client-side Faceted Search Based on SPARQL. In *LIME/SemDev@ ESWC*. 2016. <https://www.seco.tkk.fi/publications/2016/koho-et-al-sparql-faceter.pdf>.
13. *Lausen H., Ding Y., Stollberg M., Fensel D., Lara Hernández R, Han S.-K.* Semantic web portals: State-of-the-art survey, *Journal of Knowledge Management* 9: (5), 2005, pp.40–49. doi:10.1108/13673270510622447.
14. *Gladun A., Rogushina J., Schreurs J., Salem A. B.* Ontology-based knowledge recognition in service-oriented virtual research environments. In *7th International Conference on Information Technology*, 2015, pp. 148-155.
15. The FAIR data principles. Available from: <https://www.force11.org/group/fairgroup/fair-principles> [Accessed 4/09/2023].
16. MediaWiki. Available from: <https://www.mediawiki.org/wiki/MediaWiki/>.
17. *Krötzsch M., Vrandečić D., Völkel M.* Semantic mediawiki. *International semantic web conference*, 2006, pp. 935-942. https://link.springer.com/content/pdf/10.1007/11926078_68.pdf.
18. *Rogushina J.* Use of ontologies and Semantic Mediawiki for representation and retrieval of scientific data in the FAIR paradigm *Proc.of International scientific conference “Information technologies and management in higher education and sciences”*, 2023, pp.69-74. <https://doi.org/10.30525/978-9934-26-277-7-200>.
19. *Рогущина Ю. В., Гришанова І. Ю.* Проблеми масштабування семантичних інформаційних ресурсів зі складною структурою. *Проблеми програмування*, 2022, № 2-3, с.171-182.
20. *Рогущина Ю. В.* Нечіткі дані у семантичних wiki-ресурсах: моделі, джерела та методи обробки. *Проблеми програмування*, 2023, №2, с. 67-83. URL: pp.isoftware.kiev.ua/index.php/ojs1/article/viewFile/569/620.

References

1. Gladun A., Rogushina J. (2013). Ontology repositories as a means of reusing knowledge for recognizing information objects. *Design Ontology*, 1 (7), 35-50 (in Russian).
2. Strauss H. (2021) Web portals: the future of information access and distribution. In *Transforming Serials*, 7-35.
3. What is the Difference between Website and Web Portal. www.geeksforgeeks.org/difference-website-web-portal/ [Accessed 4/09/2023].

4. Difference between Website and Web Portal. forum.huawei.com/enterprise/en/Difference-between-Website-and-Web-Portal [Accessed 4/09/2023].
5. Web Portal. <https://www.techopedia.com/definition/17352/web-portal> [Accessed 4/09/2023].
6. Bhat M. U. D., Singh S. (2020) Automatic Method to Rate Websites Based on Terms of Services. Journal Of Creative Writing , 4(2).
7. Berners-Lee T., Hendler, J., Lassila, O. (2001) The semantic web. Scientific american, 284(5), 34-43.
8. Bizer C., Heath T., Berners-Lee, T. Linked data: The story so far. In Semantic services, interoperability and web applications: emerging concepts, 2011, pp. 205-227.
9. Hyvönen E. (2020) “Sampo” model and semantic portals for digital humanities on the Semantic Web. In Proceedings of the Digital Humanities in the Nordic Countries 5th Conference (DHN 2020), 374-378.
10. FAIR_data. https://en.wikipedia.org/wiki/FAIR_data. [Accessed 4/09/2023].
11. Ikkala E., Hyvönen E., Rantala H., Koho M. (2022) Sampo-UI: A full stack JavaScript framework for developing semantic portal user interfaces. Semantic Web, 13(1), 69-84. <https://content.iospress.com/articles/semantic-web/sw210428>.
12. Koho M., Heino E., Hyvönen E. (2016) SPARQL Faceter-Client-side Faceted Search Based on SPARQL. In LIME/SemDev@ ESWC. . <https://www.seco.tkk.fi/publications/2016/koho-et-al-sparql-faceter.pdf>.
13. Lausen H. , Ding Y. , Stollberg M. , Fensel D., Lara Hernández R, Han S.-K. (2005) Semantic web portals: State-of-the-art survey, Journal of Knowledge Management 9: (5), 40–49. doi:10.1108/13673270510622447.
14. Gladun A., Rogushina J., Schreurs J., Salem A. B. (2015) Ontology-based knowledge recognition in service-oriented virtual research environments. In 7th International Conference on Information Technology, 148-155.
15. The FAIR data principles. Available from: <https://www.force11.org/group/fairgroup/fair-principles> [Accessed 4/09/2023].
16. MediaWiki. Available from: <https://www.mediawiki.org/wiki/MediaWiki/> [Accessed 4/09/2023].
17. Krötzsch M., Vrandečić D., Völkel M. (2006) Semantic mediawiki. International semantic web conference, 935-942. https://link.springer.com/content/pdf/10.1007/11926078_68.pdf.
18. Rogushina J. (2023) Use of ontologies and Semantic Mediawiki for representation and retrieval of scientific data in the FAIR paradigm Proc.of International scientific conference “Information technologies and management in higher education and sciences”, 69-74. <https://doi.org/10.30525/978-9934-26-277-7-200>.
19. Rogushina J., Gryshanova I. (2022) Problems of scaling semantic information resources with a complex structure. Problems of programming, , 2-3, 171-182.(in Ukrainian).
20. Rogushina J., (2023). Fuzzy data in semantic Wiki-resources: models, sources and processing methods. Problems in programming, (2), 67-83 .(in Ukrainian)..

Одержано: 26.08.2023

Про авторів:

Рогущина Юлія Віталіївна,
Канд.фіз.-мат.наук, с.н.с.
Публікації в українських виданнях – 220,
Публікації в іноземних журналах – 40.
Індекс Хірша: Scopus – 5,
Google Scholar – 21.
ORCID <http://orcid.org/0000-0001-7958-2557>.

Місце роботи авторів:

Інститут програмних систем
НАН України, 03181, Київ-187,
проспект Академіка Глушкова, 40,
e-mail: ladamandraka2010@gmail.com,
066 550 1999.