



ПРОГРАМНО-ТЕХНІЧНІ КОМПЛЕКСИ

УДК 004.318

О.В. ПАЛАГІН

Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна,
e-mail: *palagin_a@ukr.net*.

М.Г. ПЕТРЕНКО

Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна,
e-mail: *petrng@ukr.net*.

К.С. МАЛАХОВ

Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна,
e-mail: *k.malakhov@incyb.kiev.ua*.

ПРОБЛЕМИ І РОЛЬ ОНТОЛОГІЧНОЇ ІНЖЕНЕРІЇ У ФОРМУВАННІ ІНДУСТРІЇ ЗНАНЬ З ПОЗИЦІЙ ДОСЛІДНОГО ПРОЄКТУВАННЯ¹

Анотація. Розглянуто моделі та механізми трансдисциплінарних наукових досліджень для створення кластерів конвергенції дисциплін і наукових теорій, формального представлення знань і формування індустрії знань за допомогою єдиного інструментарію онтологічної інженерії. Приділено увагу можливостям дослідного проєктування для створення нових знань і технологій. Проаналізовано концептуальні терміни онтологічної інженерії та показано важливість і ефективність використання онтологічних знань та механізмів для розв'язання задач користувачів. Детально розглянуто предметну онтологію, її онтологічний граф та їхні суттєві відмінності і переваги над owl-онтологіями. Наведено приклад застосування онтології задач для проєктування обчислювальних пристроїв на програмовних мікросхемах.

Ключові слова: індустрія знань, онтологічна інженерія, трансдисциплінарні наукові дослідження, кластери конвергенції, дослідне проєктування, нові знання та технології, онтологія предметної галузі, онтологічний граф, онтологія задач.

ПОСТАНОВКА ПРОБЛЕМИ

Постійне ускладнення проблем підтримки сталого цілеспрямованого розвитку людства зумовлює невідпинне розширення обсягу знань, зокрема, про їхнє використання у поточній діяльності людини. Спрямованість цього процесу на сучасному трансдисциплінарному етапі розвитку науки зводиться до поетапного збільшення знань. Щодо процесу видобування знань, то його можна представити такою послідовністю фрагментів життєвого циклу знань:

- когнітивний процес (пізнання);
- видобування нових знань;
- перетворення загальних знань на систему споживчих сервісів (сервісних знань);
- апробація знань в процесі їхнього використання у духовній (навчання, концепції, теорії) та матеріальній (матеріальні об'єкти) діяльності;
- розроблення перспективних концепцій розвитку знань та їхнє застосування [1].

¹ Дослідження виконано за підтримки Національного фонду досліджень України (FundRef ID: 10.13039/100018227, грант № 2021.01/0136 «Розробка хмарної платформи пацієнт-центричної телереабілітації онкологічних хворих на основі математичного моделювання») на базі Інституту кібернетики ім. В.М. Глушкова НАН України.

Рівень розвитку інтелектуальних інформаційних технологій значною мірою впливає на ефективність процесів, що відбуваються в економічній, науково-технічній, освітній та інших сферах діяльності людини. Процеси глобальної інформатизації світового співтовариства орієнтовані насамперед на побудову знаннеорієнтованого суспільства і мають дедалі яскравіше виражений трансдисциплінарний характер. Безсумнівним лідером при цьому є технології інженерії знань, зокрема, її новий напрям — онтологічна інженерія. Ці технології реалізують процеси керування знаннями (Knowledge Management) і успіхи в цьому напрямі багато в чому визначаються рівнем інтелектуалізації та ефективності комп'ютерних систем [2, 3].

Основними різновидами знань вважають: наукові, техніко-технологічні, інноваційні та компетенційні. При цьому загальне означення поняття (або концепту) «знання» зазнало безліч інтерпретацій. Найпростішим, але конструктивним означенням є таке: знання — це інструмент розв'язання задач у предметній галузі (Knowledge Domain, KD).

Розвиток повноцінної індустрії знань як головного фактора еволюційного процесу зумовлює чітке і вичерпне означення терміна «знання» як понятійного об'єкта відповідно до сьогодишнього уявлення про прагматичну картину світу і водночас як ринкової категорії, а також формування ефективної моделі «попит–пропозиція» на ринку знань як товару та на її основі — продуктивної стратегії розвитку промисловості знань [2].

У життєвому циклі знань важливою є мотивація споживачів та простота задоволення їхніх потреб. Отже, перспективним вважають створення паралельно із загальною індустрією знань інформаційних інтерактивних систем, головним призначенням яких є наближення користувача до джерела знань. Знаннеорієнтовані інтелектуальні інформаційні інтерактивні системи посідають серед них особливе місце. Основні їхні особливості: підтримка повного когнітивного циклу, семантичне оброблення інформації, активні діалогові функції, що забезпечують перетворення пасивних знань на активні. Вони є незамінними в сучасних системах підтримки прийняття рішень, прогнозуванні, наукових дослідженнях тощо.

Водночас через єдність і системну складність світу як об'єкта наукових досліджень доцільно разом із процесом диференціації розглядати процес інтеграції наукових дисциплін та відповідних технологій. Адже саме інтеграція є фундаментальною парадигмою еволюційної теорії академіка В.І. Вернадського. Процес інтеграції, що почався нещодавно здавалося б на стихійній (інтуїтивній) основі, сьогодні стає усвідомленим та очевидним. А головне, згідно з еволюційною теорією він спрямований на створення «Єдиного загального знання», сутність якого однозначно апелює до побудови «Наукової картини світу» на основі трансдисциплінарної концепції розвитку науки.

Початковий етап процесу інтеграції можна назвати етапом кластеризації, наприклад, створення кластера конвергенції NBIC (нано-біо-інформаційно-когнітивні технології). У низці успішних системних проєктів на основі кластера NBIC вже доведено його перспективність. Навіть поверхневий аналіз явища кластеризації показує, що воно базується на глибокій взаємодії методів, засобів та можливостей компонентів кластера, синтез яких породжує синергетичні ефекти за рахунок інтеграції функціональних властивостей цих компонентів та відкриває широкі перспективи для створення невідомих раніше наукових теорій, зразків нової техніки та технологій. З процесом інтеграції, що вже розпочався, пов'язано безліч проблем, розв'язання яких дасть змогу визначити нові

перспективні шляхи еволюції знань і науки в цілому відповідно до ноосферної теорії В.І. Вернадського. При цьому однією із задач є аналіз структури знань на понятійному рівні. Загальне знання та кожна його дисципліна можуть бути представлені сумарною сукупністю понять-термінів, що становлять онтологічну основу формального чи неформального опису знань. Кожен термін має своє загальновизначене означення, яке описується за допомогою термінів (чи понять) нижнього рівня. Створено технологічний багаторівневий онтологічний опис як загального знання, так і окремих його дисциплін, розділів, теорій тощо, який дає змогу формально представляти знання за допомогою єдиного інструментарію онтологічної інженерії, що відкриває широкі перспективи розвитку когнітивних технологій та їхнього продуктивного використання [2].

Нині відбувається активізація наукових досліджень як на стику різних предметних дисциплін (міждисциплінарні дослідження), так і в кластерах конвергенції (трансдисциплінарні дослідження). Для підтримки цих досліджень важливими чинниками є побудова знаннеорієнтованих інформаційних систем, удосконалення процесів дослідного проєктування (Research-related Design, RD), розроблення методів та інструментальних засобів онтологічного аналізу природномовних об'єктів для видобування з них знань, прикладних аспектів застосування онтологій, метаонтологій, систем інтеграції знань у трансдисциплінарних кластерах конвергенції [3].

На рис. 1 наведено інноваційну модель видобування і використання знань, яка відрізняється від традиційної моделі розвитку нових знань тим, що:

- керування процесом наукових досліджень доповнюється керуванням процесом трансформації знань, який має певну кількість етапів залежно від KD;
- базується на трансдисциплінарній парадигмі розвитку науки, створенні трансдисциплінарних кластерів конвергенції, які є основою перетворення «знання → інновація» в галузі природничих наук і відповідних технологій.

Також на рис. 1 показано деталізацію верхнього рівня вказаної інноваційної моделі, нижній рівень якої сконцентровано на моделях і методах «Knowledge Engineering», зокрема і «Ontology Engineering» [2].

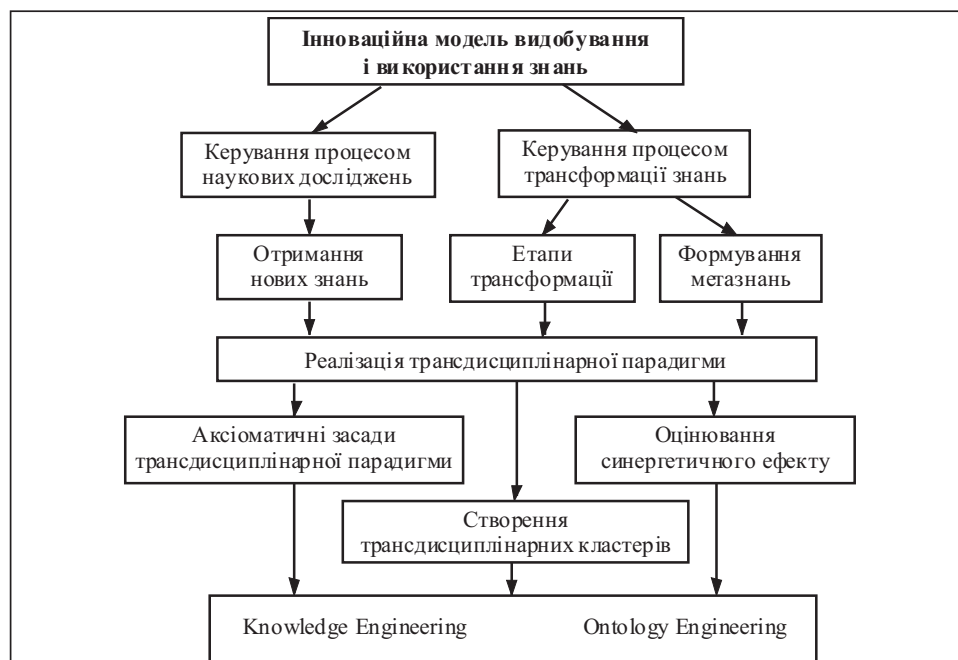


Рис. 1. Часткова інноваційна модель видобування і використання знань

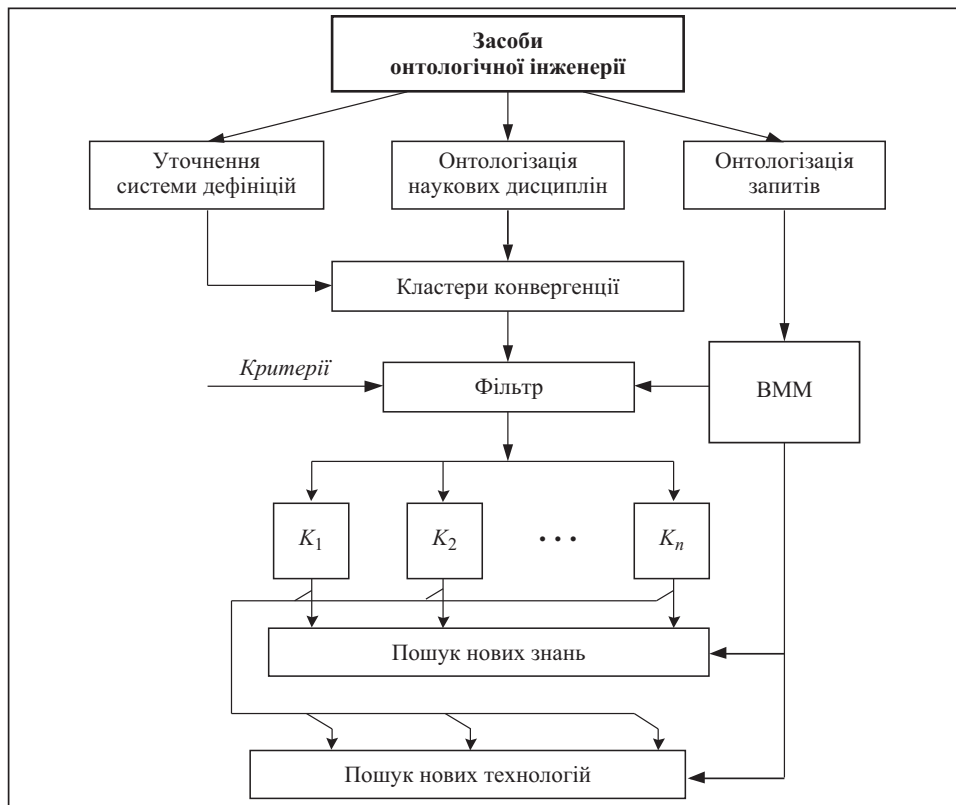


Рис. 2. Блок-схема продукування нових знань з використанням засобів штучного інтелекту

Розглянемо детальніше блок «Створення трансдисциплінарних кластерів» (див. рис. 1). При цьому зазначимо, що важливими є задача обмеження кількості рівнів онтологічних описів кластерів конвергенції та ступінь їхнього представництва у технології Knowledge Engineering. Як зазначено в [2], головним обмеженням слід вважати обсяг накопичених знань у цій KD. Таким чином, кількість рівнів онтологічного опису може бути однією з об'єктивних характеристик загальних знань та окремих дисциплін. Задача побудови багаторівневих онтологічних описів KD зазвичай трудомістка і вимагає залучення засобів штучного інтелекту, зокрема Великих Мовних Моделей (ВММ), що базуються на аналізі та генерації природномовних текстів за допомогою методів онтологічної інженерії. Блок-схему засобів онтологічної інженерії для формування нових знань і нових технологій наведено на рис. 2.

Використання ВММ забезпечує розширення і поглиблення взаємодії всіх рівнів представлення (насамперед онтологічного) дисциплін і за цей рахунок пошук потенційного синергетичного ефекту як результату цієї взаємодії.

Метою цієї статті є аналіз моделей і засобів онтологічної інженерії для розв'язання задач трансдисциплінарної кластеризації, формування нових знань і нових технологій та розбудови конкретних сервісів користувачів. (Автори цієї статті, а саме Олександр Палагін та Кирило Малахов входять до складу Міжвідомчої робочої групи з питань розбудови телемедицини в Україні [4] (наказ МОЗ України від 3 листопада 2023 р., № 1906).)

ДОСЛІДНЕ ПРОЄКТУВАННЯ НОВИХ ЗАСОБІВ І ТЕХНОЛОГІЙ ФОРМУВАННЯ НОВИХ ЗНАНЬ

Як різновид наукових досліджень RD характеризується тим, що основні його етапи пов'язані з процесом опису передбачуваного образу об'єкта нової техніки (Object of New Hardware, ONH), що проєктується [5, 6]. Сам процес проєктування будується як низка інтерактивних процедур дедалі глибшого занурення в опис ONH і формування проміжних варіантів технічного рішення. При цьому перетворення опису об'єкта проєктування залежить як від знань про KD загалом, так і від знань і досвіду проєктувальника.

Процес RD охоплює такі етапи: збір матеріалу, що представляє KD; формування мети проєктування; аналіз матеріалу та онтологічний опис KD; виявлення суперечностей і формулювання проблемної ситуації; постановка задачі; уточнення проблемної ситуації; виявлення аналогів-прототипів, формування сукупності технічних ознак та вигляду ONH; виконання ескізно-технічного етапу проєктування і підготовки матеріалів до патентування.

У [7] детально розглянуто методологічні засади побудови систем RD, які містять моделі, технологію та інструментальні засоби когнітивних процесів оброблення інформації, зокрема, проєктування ONH, автоматизованої побудови онтологічних баз знань KD. Останні є базовими компонентами інтелектуальних технологій і систем під час проведення складних наукових досліджень міждисциплінарного і трансдисциплінарного характеру. Технологія системної інтеграції онтологій KD дає змогу дослідити взаємодію предметних знань як на межі предметних дисциплін, так і в кластерах конвергенції, що відкриває широкі можливості для отримання нових знань та розроблення наукових теорій.

ОЗНАЧЕННЯ І СУТНІСТЬ ОСНОВНИХ ТЕРМІНІВ ОНТОЛОГІЧНОЇ ІНЖЕНЕРІЇ

Наукова та практична складові дисципліни «Комп'ютерні науки» мають широке розгалуження. У всіх цих напрямках так чи інакше використовують концептуальну термінологію, а саме: Knowledge Engineering, Knowledge Processing, Knowledge-based System, Knowledge Base, Knowledge Graph, Ontology Engineering, Ontology, Ontology-related System, Owl-ontology, Other Ontologies і Ontology Graph.

У царині інтелектуальних інформаційних систем та інженерії знань концепт вважають фундаментальною складовою наукового пізнання. Він слугує когнітивною основою, яка фіксує та містить у собі суттєві ознаки об'єктів і явищ у межах об'єктивної реальності. В контексті онтології концепт можна розглядати як окреме подання, що відрізняється від звичайних тверджень тим, що воно зосереджене на внутрішній природі об'єкта. Такі формулювання відповідають історичній перспективі в логіці, що часто називаються реальними визначеннями об'єктів. Концепт стає смисловим значенням терміна і визначає його суть. Семантичне значення терміна тісно пов'язане з концептом, який його представляє, та відіграє основну роль у визначенні його суті. У межах концентрації знань і когнітивного розвитку поняття є концентрованою формою знань, що являє собою результат пізнання на певному етапі. Водночас воно слугує і точкою відліку, і засобом для подальших досліджень [8, 9].

У багатьох публікаціях не існує єдиної узгодженості щодо семантичного значення і практичних застосувань наведених термінів. Для інформативнішого подальшого викладення на рис. 3 показано структурно-семантичну схему концептів, які представляють зацікавленість. Також, на думку авторів, невинновдано забуті поняття з інших видів онтологій, окрім owl-онтологій, та їхніх практичних застосувань.

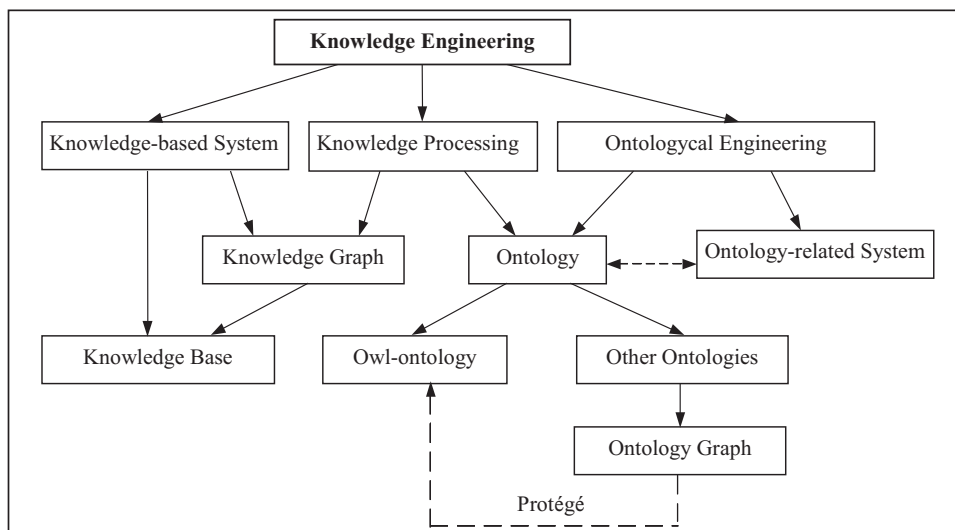


Рис. 3. Структурно-семантична схема концептуальних термінів

Насамперед зацікавленість стосується термінів Knowledge Graph, Ontology і Ontology Graph, тлумачення яких мають найбільші розбіжності, а щодо останнього терміна, то в англомовних публікаціях про нього взагалі не йдеться. У запропонованій роботі було проаналізовано такі авторитетні джерела, як [10–20] та багато інших. Крім того, інші змістовні інформаційні джерела розглянуті далі.

Визначальним концептом у [10–18] вважається термін «Граф знань». Він слугує базовим концептом для інших знаннеорієнтованих структур (див. рис. 3). У публікаціях пострадянського періоду його синонімом був термін «Семантична мережа».

У дослідженнях графів знань передбачено поєднання методів із різних дослідницьких галузей для максимізації знань, які можуть бути вилучені з різних джерел за допомогою графової абстракції даних [20]. Хоча прогрес у цих окремих дисциплінах, безсумнівно, продовжуватиметься і матиме подальший вплив, особливо цікаві теми для дослідження графів знань виникають на перетині дисциплін. Це корелює з методологічними засадами міждисциплінарних і трансдисциплінарних наукових досліджень і побудовою кластерів конвергенції останніх [19].

Наведемо означення терміна «знання» та пов'язаних з ним похідних термінів.

Знання. Інформація, яка може бути виражена у вигляді набору фактів і відома агенту або програмі. Знання можна відрізнити від інформації або даних за їхнім втіленням в агенті, наприклад, агент може отримувати інформацію, яка збільшує його знання [11, с. 718]. Знання — це інформація про процеси розв'язання, логікового виведення, про закономірності, в результаті застосування якої до даних породжується нова інформація чи знання [21].

База знань. Набір знань, зазвичай релевантних до певної КД, які були формалізовані у відповідній схемі для підтримки процесів міркувань. Часто використовують формалізми на основі правил, але існують й інші методи представлення знань. Бази знань відрізняються від баз даних тим, що

- не тільки зберігають дані, але й полегшують модифікацію, перегляд та інші форми внутрішньої маніпуляції зі знаннями;
- здатні обробляти знання, які є неповними, несуперечливими та невизначеними;
- можуть використовувати як імперативні, так і декларативні форми знань [11, с. 718–719].

У комп'ютерних науках, інформатиці та системній інженерії онтологічна інженерія — це галузь, яка вивчає методи та методології побудови онтологій, що містить в собі подання, формальне визначення понять, властивостей та зв'язків між поняттями, даними та сутностями певної КД, що представляє зацікавленість. У ширшому розумінні ця сфера також включає побудову знань про КД за допомогою формальних представлень онтологій, зокрема, таких, як OWL/RDF. Онтологічна інженерія є одним з напрямів використання прикладної онтології і застосування в філософському означенні онтології. Основні ідеї та цілі онтологічної інженерії також є центральними в концептуальному моделюванні [15].

Онтологія. Опис певних понять та взаємовідношень між ними для визначення суджень, достатніх для того, щоб комп'ютер міг їх сприймати та міркувати про них. Таким чином, комп'ютерна онтологія описує основні будівельні блоки знань, які відображають те, що комп'ютер може сприймати і про що він може міркувати. Це своєрідна модель і, як така, дуже корисна для розуміння того, що агенти або навчальні програми можуть знати і про що вони можуть спілкуватися. Онтології зазвичай складаються для певного «домена», наприклад, для енергетики, медичної діагностики або навігації в офісі, і вони є більш формальними ніж доменні знання [11, стор. 836].

Наведемо загальновизнані означення онтологій [22–25]:

- експліцитна специфікація концептуалізації (формально онтологія складається з термінів, організованих у таксономію, їхніх визначень та атрибутів, а також пов'язаних з ними аксіом і правилами виведення);
- база знань спеціального типу, яка може «читатися» і розумітися, відчуватися від розробника і/або фізично розподілятися користувачами;
- формалізоване представлення основних понять КД та зв'язків між ними. Основні задачі, які розв'язують із застосуванням онтологій:
 - створення та використання баз знань;
 - організація ефективного пошуку в базах даних, інформаційних каталогах, базах знань;
 - побудова систем, які реалізують механізми міркувань (системи керування, експертні системи, інтелектуальні роботи та ін.);
 - організація пошуку за сенсом у текстовій інформації;
 - семантичний пошук в Інтернеті;
 - побудова та використання баз знань для створення інформаційних систем різного призначення;
 - забезпечення загальної термінології для сукупності спеціалістів і застосунків, які сумісно використовуються;
 - багатократне застосування баз знань та інформаційних масивів, які надають відомості про технічні системи на різних стадіях їхнього життєвого циклу;
 - побудова знаннеорієнтованих інформаційних систем з онтолого-керівною архітектурою;
 - інформаційна та інструментальна підтримки міждисциплінарних та трансдисциплінарних наукових досліджень.

Розглянемо основні типи онтологій залежно від їхнього застосування.

ОНТОЛОГІЯ ПРЕДМЕТНОЇ ГАЛУЗІ ТА ЇЇ ОНТОЛОГІЧНИЙ ГРАФ

Як зазначається в [10, с. 6–7], «... граф знань представляється як направлений, з поміченими ребрами, причому в загальному випадку кількість вершин, ребер і міток може бути незліченною». І далі відзначається, що «багато прак-

тичних реалізацій накладають обмеження на зв'язки в графах знань, при цьому визначаючи схему чи онтологію» (мається на увазі owl-онтологію).

На відміну від owl-онтологій, в багатьох роботах (переважно пострадянського періоду) уведено поняття (C, R, F, A) -онтології KD, яка своєю чергою описує концептуальну (термінологічну) складову онтологічної бази знань певної KD [5, 19, 24, 25]. Таку онтологію можна описати так:

$$O^{KD} = \{G^O, F, A\}, \quad (1)$$

де G^O — онтологічний граф (означення буде розглянуто далі); F — функції інтерпретації термінів і семантичних відношень між ними; A — аксіоми, які в загальному випадку визначають обмеження та область дії термінологічних структур D .

Граф G^O є невід'ємною частиною онтології (1).

Онтологічним графом G^O KD називається орієнтований ациклічний граф вигляду

$$G^O = (C, R) = \{(C_0, C_j, C_n), (R(C_0, C_j), R(C_j, C_n))\}, \quad (2)$$

де C_0 — коренева вершина, яка не містить вхідних ребер, належить нульовому рівню ієрархії в графі, може бути зв'язана з вершинами C_j у довільному порядку; C_j^i — вершини j -го рівня, можуть бути зв'язані тільки з вершинами більш високого рівня ($j+$), зокрема, з прикінцевими вершинами; C_n^s — кінцеві вершини або листя. Нижній індекс у позначеннях вершин — номер рівня у графі G^O , а верхній індекс — порядковий номер вершини у відповідному рівні.

Узагальнене відображення графу G^O показано на рис. 4.

Ребра R графу G^O направлені від вершини C_0 до вершин $C_j^1, \dots, C_j^i$ і до вершин-листя C_n^s . Інакше кажучи, ребра $\{R\} = \{R_0\} \rightarrow \{R_j\} \rightarrow \{R_n\}$ є семантичними відношеннями (переважно таксономічними) та специфічними для конкретної галузі семантичними відношеннями. Відповідно до цих відношень визначається структура графу G^O і його розбиття на рівні. З рис. 4 випливає, що із збільшенням рівня вершин-термінів графу $(0, 1, 2, \dots, n-1, n)$ їхній обсяг зменшується, а вміст збільшується.

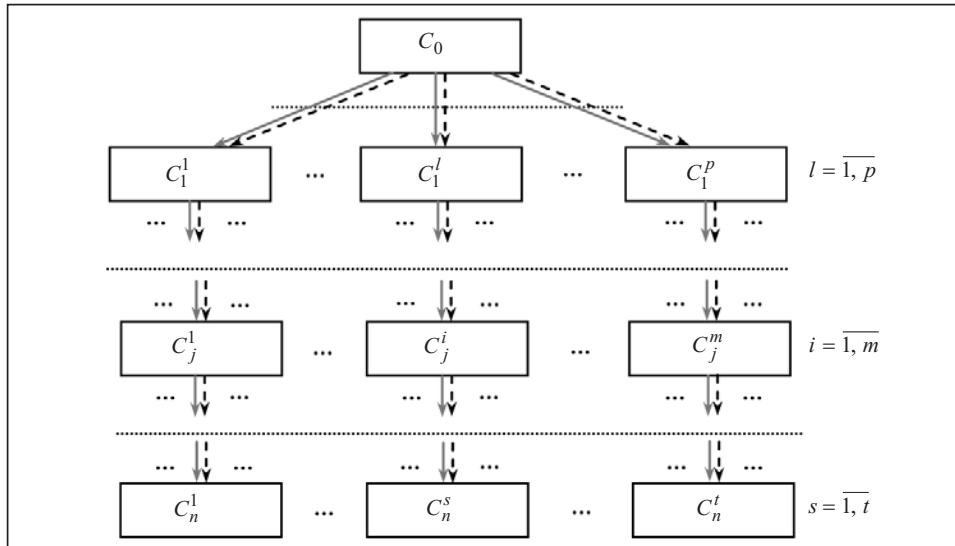


Рис. 4. Узагальнене відображення графу G^O

Ребра та їхня спрямованість (див. рис. 4) позначені суцільною і штриховою лініями, що означає «зв'язок актуальний» для відповідного ребра у першому випадку і «зв'язок неактуальний» (не існує) — у другому випадку. Актуальність (і відповідно неактуальність) ребер залежить від інженера зі знань, що розробляє граф G^O .

Іншою важливою задачею для графу G^O є його застосування як структурованого семантичного подання класів KD для швидкого, наглядного ознайомлення з призначенням і обсягом KD або її певної частини [26].

Якщо KD містить від декількох десятків термінів, то інженер зі знань має автоматизувати процес побудови графу G^O . Відомі інструментальні засоби, що реалізують цей процес, зокрема, для української мови — це Конспект [21, 27] і Поліедр [28].

Як аналог графу G^O (якщо як множина C описуються концепти категоріального рівня) зазначимо онтологію DOLCE [29].

Зазвичай для розв'язання певної множини прикладних задач інженер зі знань не розробляє максимально повної онтології (зокрема, і граф G^O), а до неї включає тільки ті вершини-концепти, які безпосередньо використовує для розв'язання вказаних прикладних задач. Відповідно і кількість ребер R буде меншою за можливу максимальну кількість.

Як зазначалось раніше, онтологія KD складає її концептуальну структуру (ту її частину, що представляє зацікавленість), а для отримання онтологічної бази знань KD потрібно до O^{KD} додати відповідну базу даних. Одним із аналогів онтологічної бази знань для англійської мови є WordNet.

ОНТОЛОГІЯ ЗАДАЧ ПРЕДМЕТНОЇ ГАЛУЗІ

Із предметно-проблемних знань у загальному випадку виокремлюють поняття і процеси KD, поняття і класи задач проблемного простору (Problem Domain, PD) предметної галузі та методи розв'язання останніх, а також алгоритми, що реалізують відповідні методи. Вказані типи знань за темпоральною ознакою поділяють на статичні та динамічні, а за онтологічною схемою — на такі онтології: об'єктів KD; процесів KD; застосунків; класу задач PD; методів розв'язання задач.

У [5, 19] (C, R, F, A) -онтологія KD розглядається як трикомпонентна. Зокрема, вона складається з онтологій об'єктів, процесів і задач. Таке розбиття (C, R, F, A) -онтології KD має своє практичне значення. Для розв'язання конкретного типового набору задач у PD розробляється тільки онтологія задач (Task Ontology, TO) цього PD, інші дві компоненти розробляються тільки один раз на початку розроблення онтології KD. На рис. 5 показано узагальнену схему онтології задач.

Онтологія об'єктів певної KD з добре розвинутим теоретичним базисом обов'язково містить таке поняття, як «Теоретичний базис». Концептом нижнього рівня цього поняття є «Метод». За допомогою методів, що описують цей концепт, виконується постановка і розв'язання задач у відповідному PD. Концепт «Метод» є абстрактним поняттям, яке на нижньому рівні класифікується (мається на увазі онтологічна класифікація) за ознакою «Клас задач». Алгоритми, що реалізують відповідні методи для розв'язання певного класу задач, не розглядаються ні в онтології об'єктів, ні в онтології процесів. Описи вказаних алгоритмів у прийнятому представленні для описів онтологій об'єктів і процесів зберігаються в блоці «Формувач методів розв'язання задач» (див. рис. 5), який належить складу TO.

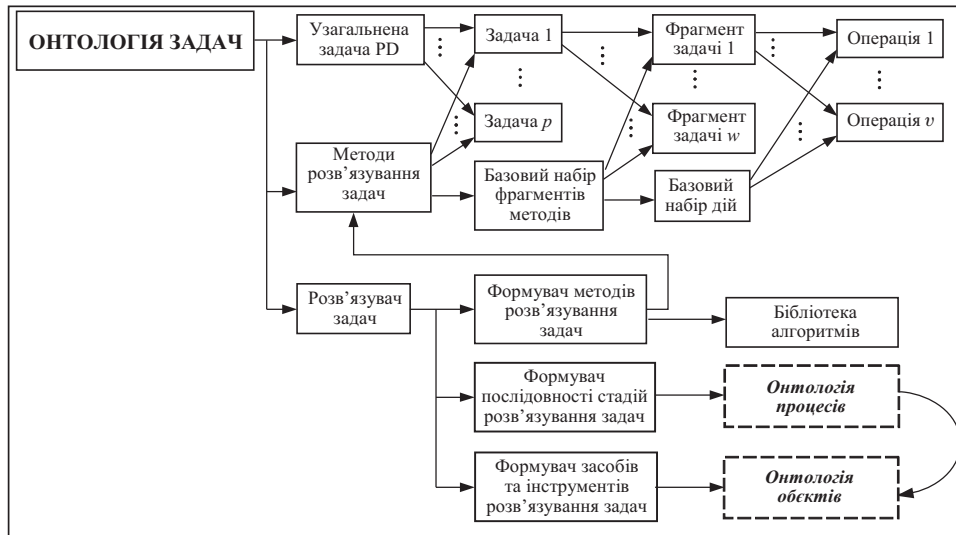


Рис. 5. Узагальнена схема онтології задач PD

Зазначимо, що базовими онтологічними задачами є: класифікація понять, побудова таксономій з відношеннями, які належать їм, інтеграція онтологічних структур, відображення і модифікація онтологій, навігація (інформаційний пошук і запити), і насамкінець, розв'язання прикладних задач KD на рівні сервісної архітектури тощо [5].

Для побудови онтології задач певного PD у загальному вигляді зазвичай задається така інформація.

1. Предметна галузь з областю інтерпретації KD.

2. Онтологія KD.

2.1. Онтологія об'єктів KD, формальна модель якої описується четвіркою:

$$O^O = \langle C, R^O, F^O, A^O \rangle, \quad (3)$$

де

$$C = \{c_1, c_2, \dots, c_i, \dots, c_I\} \quad \forall i = \overline{1, I},$$

$$F^O = C \times R^O = \{f_h^O\} = \{c_i\} \times \{r_j^O\} \quad \forall h = \overline{1, H},$$

$$A^O = \{D_j^O\} \quad \forall j = \overline{1, J}.$$

2.2. Онтологія процесів KD, формальна модель якої описується трійкою:

$$O^P = \langle P, R^P, F^P \rangle, \quad (4)$$

де P — процес, $P = \{sp_1, sp_2, \dots, sp_l, \dots, sp_L\}$, sp_l — підпроцеси (subprocesses) $P \quad \forall l = \overline{1, L}$,

$$R^P = \{r_1^P, r_2^P, \dots, r_k^P, \dots, r_K^P\} \quad \forall k = \overline{1, K},$$

$$F^P = P \times R^P = \{f_g^P\} = \{sp_l\} \times \{r_k^P\} \quad \forall g = \overline{1, G}.$$

3. У PD задано типовий набір задач $OT_n^{PD} \in OT^{PD}$, де OT^{PD} — множина задач, які розв'язуються в проблемному просторі; $\forall n = \overline{1, N}$ — множина типових наборів задач.

4. У PD задана множина методів $\{M_{n_q}^{PD}\}$, за допомогою яких розв'язуються задачі $\{T_{n_p}\}$, де q — множина методів для n -го типового набору задач; p — множина задач у n -му типовому наборі.

4.1. Якщо $q < p$, то для деяких (або всіх) задач $\{T_{n_p}\}$ існує не більше одного методу їхнього розв'язання.

4.2. Якщо $q = p$, то для кожної задачі $\{T_{n_p}\}$ існує тільки один метод її розв'язання $\{M_{n_q}^{PD}\}$.

4.3. Якщо $q > p$, то для деяких (або всіх) задач $\{T_{n_p}\}$ існує більше одного методу їхнього розв'язання.

Потрібно для типового набору задач $\{T_n\}$ сформувати підмножини задач $\{T_{n_p}\}$, фрагментів задач w і операцій v , а також конкретизувати функції «Розв'язувача задач» (див. рис. 5) для заданого типового набору. Для цього слід виконати такі дії:

- розробити «Формувач методів розв'язання задач» (для множини методів $\{M_{n_q}^{PD}\}$ потрібно сформувати підмножини базових наборів фрагментів методів і дій);

- сформувати послідовності процесів, задіяних в алгоритмах відповідних методів розв'язання задач, і об'єктів, які задіяні в розв'язанні задач, з можливими значеннями, обмеженнями та умовами;

- вказати в явному вигляді, які конкретні об'єкти із онтології об'єктів задіяні у відповідних процесах (алгоритмах) розв'язання задач.

А також потрібно розробити узагальнений алгоритм розв'язання заданого типового набору задач і розробити алгоритмічно-програмний комплекс, який, окрім реалізації узагальненого алгоритму, зв'язував би в єдину інтелектуальну інформаційну технологію всі компоненти «Онтології задач» і забезпечував інтерфейс з користувачами.

ОНТОЛОГІЯ ЗАДАЧ КД «ПРОЄКТУВАННЯ ОБЧИСЛЮВАЛЬНИХ ПРИСТРОЇВ І СИСТЕМ НА ПЛІС»

Специфіка заданої КД «Проектування обчислювальних пристроїв і систем на ПЛІС» полягає в тому, що в ній практично розв'язуються два набори типових задач: у першому — задачі проектування, а у другому — перепроєктування певного обчислювального пристрою (Computing Device, CD) або системи на програмних логікових інтегральних схемах (ПЛІС) відповідно до технічних завдань (ТЗ) на проектування. Перший набір орієнтований на проектування нового CD чи системи, другий набір — на перепроєктування вже наявного CD чи системи в елементний базис ПЛІС. Крім того, є різні способи описування проекту [30–32], для яких відомі методики та детальні алгоритми їхньої реалізації.

Постановка задачі для другого набору задач згідно з описаною раніше схемою може бути сформульована таким чином: перепроєктувати вже наявний CD, наприклад апаратний лінгвістичний процесор (hardware linguistic processor, HLP), що являє собою класичний мікропрограмний процесор [5, 33]. Таке перепроєктування, виконане ручним способом, описане там же.

Вхідними даними для перепроєктування є:

- тип CD чи системи і його (її) призначення;
- специфікація елементної бази, на якій реалізовано CD чи система;
- схема електрична принципова CD чи системи;
- інтерфейс із «зовнішнім середовищем»;
- технічні вимоги та критерії, які має задовольняти CD чи система, що перепроєктується.

Крім того, зазвичай іще задається тип системи автоматизації проектування, якою буде виконуватися перепроєктування, і діапазон доступних кристалів програмних логікових інтегральних схем.

У результаті розв'язання задачі отримуємо:

- готовий проєкт у кристалі, який містить готовий виріб чи файл конфігурації кристала;
- характеристику проєкту;
- технічну документацію;
- звіт про виконання проєкту.

Методом розв'язання для другого типового набору задач переважно є метод модульного проєктування із структурною декомпозицією CD чи системи.

Як впливає з контексту «Онтології задач», наведений раніше узагальнений алгоритм складається із двох частин: проблемної (зокрема, і предметної) та прикладної. У першій частині виконується конкретизація блоків, зображених на рис. 5. Практичний приклад реалізації другої частини описано в [6].

Розглянемо детальніше методичні засади проєктування лінгвістичного процесора.

Оброблення мовленнєвої або текстової інформації забезпечується лінгвістичним процесором чи то на рівні «мовної» свідомості людини, чи то в комп'ютерній системі. В останній він є основним компонентом, що реалізує розпізнавання і розуміння вхідної природномовної інформації, видобування з неї первинних знань з їхнім подальшим формально-логічним представленням. Отриманої інформаційної структури вже достатньо для реалізації (знаннеорієнтованих) процедур для розв'язання прикладних задач, прийняття рішень тощо.

Лінгвістичний процесор (апаратний чи програмний) інтерпретує текстову інформацію (деякий природномовний документ, стаття, монографія або лінгвістичний корпус текстів (ЛКТ)) відповідно до етапів лінгвістичного аналізу: графематичного, морфологічного, синтаксичного та семантичного (точніше, поверхнево-семантичного). Результатом роботи лінгвістичного процесора є інформаційна структура, призначена для проведення глибинно-семантичного аналізу в підсистемі екстралінгвістичного опрацювання, завданням якої є видобування та формування структури понять, реалізація їхньої прагматичної інтерпретації в термінах прикладного завдання або відповідна реакція, притаманна людині.

На рис. 6 наведено функціональну схему HLP, що містить відповідні підсистеми реалізації етапів лінгвістичного аналізу та мовно-онтологічну картину світу, використання якої є однією з основних відмінностей HLP від класичного лінгвістичного процесора, а проєктування мовно-онтологічної картини світу розглянуто в [19]. У процесі виконання загального алгоритму лінгвістичного аналізу текст поетапно перетворюється на графематичну, морфологічну, синтаксичну та семантичну структури, кожна з яких має свою модель подання.

Насамкінець у результаті роботи HLP буде сформовано таблицю шаблонів речень тексту, у якій зберігатиметься інформаційна структура для всього тексту T_k загалом і яка буде вхідною інформацією для підсистеми екстралінгвістичного оброблення тексту T_k . В останній виконується формально-логічне представлення (переведення) речень і тексту загалом у відповідну формальну теорію першого порядку, наприклад у логіку предикатів першого порядку.

Іншим прикладом застосування «Онтології задач» є розглянуте в [6] створення та оформлення винаходу для пристрою «Hardware morphology processor».

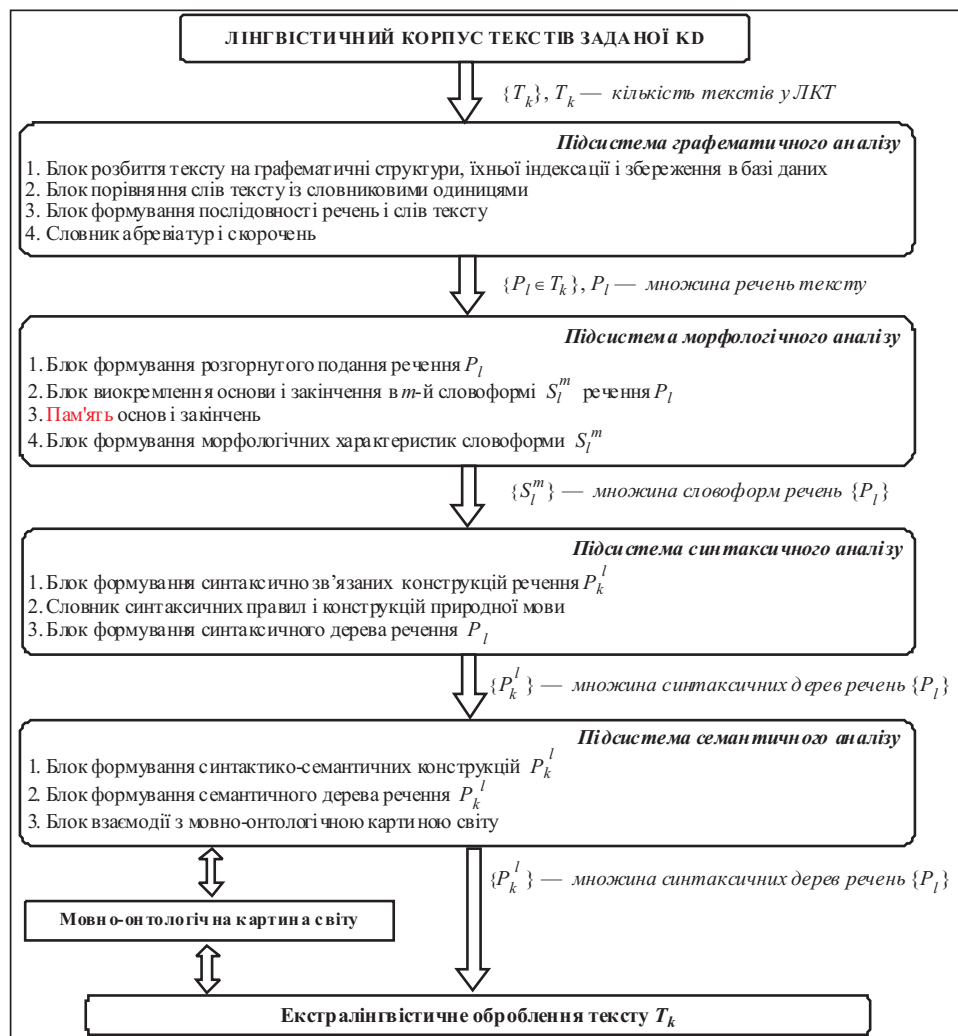


Рис. 6. Функціональна схема апаратного лінгвістичного процесора

ВИСНОВКИ

У статті розглянуто методологічні засади формування індустрії знань та її вплив на цілеспрямовані процеси поступового розвитку людства, що на сучасному трансдисциплінарному етапі науки зводиться до поступового при множування знань. Приділено увагу можливостям дослідного проектування для створення нових знань і технологій.

Виконано аналіз сутнісних засад та взаємозв'язку концептуальних термінів у комп'ютерній науці. Підкреслено особливу роль понять онтологічної інженерії в становленні системології трансдисциплінарних наукових досліджень, зокрема, у формуванні кластерів конвергенції. Насамкінець розглянуто модель онтології та її складові, а саме онтологію задач. Остання дає змогу значно спростити процес розроблення та розв'язання типового набору задач у предметній галузі. Показано важливість і ефективність застосування онтологічних знань і механізмів у розв'язанні задач користувачів.

СПИСОК ЛІТЕРАТУРИ

1. Palagin A.V. Transdisciplinarity problems and the role of informatics. *Cybernetics and Systems Analysis*. 2013. Vol. 49, N 5. P. 643–651. <https://doi.org/10.1007/s10559-013-9551-y>.

2. Палагін О.В. Системний підхід до формування індустрії знань. «Глушковські читання»: *Матеріали XI Міжнародної конф.* Київ, 2022. С. 51–55. URL: https://drive.google.com/file/d/1QwFk8O2gF7tFI8t1XYDC0vyUlyxqrPtv/view?usp=share_link.
3. Palagin A.V., Petrenko M.N. Methodological foundations for development, formation and IT-support of transdisciplinary research. *Journal of Automation and Information Sciences*. 2018. Vol. 50, Iss. 10. P. 1–17. <https://doi.org/10.1615/JAutomatInfScien.v50.i10.10>.
4. Malakhov K.S. Insight into the digital health system of Ukraine (eHealth): Trends, definitions, standards, and legislative revisions. *Int. J. Telerehab*. 2023. Vol. 15, N 2. <https://doi.org/10.5195/ijt.2023.6599>.
5. Palagin O., Petrenko M., Kryvyi S., Boyko M., Malakhov K. Ontology-driven processing of transdisciplinary domain knowledge. Iowa State University Digital Press, 2023. 189 p. <https://doi.org/10.31274/isudp.2023.140>.
6. Палагин А.В., Петренко Н.Г., Малахов К.С. Информационная технология и инструментальные средства поддержки процессов исследовательского проектирования Smart-систем. *Управляющие системы и машины*. 2018. № 2. С. 19–30. <https://doi.org/10.15407/usim.2018.02.019>.
7. Palagin A.V. Functionally oriented approach in research-related design. *Cybernetics and Systems Analysis*. 2017. Vol. 53, N 6. P. 986–992. <https://doi.org/10.1007/s10559-017-0001-0>.
8. Malakhov K., Petrenko M., Cohn E. Developing an ontology-based system for semantic processing of scientific digital libraries. *South African Computer Journal*. 2023. Vol. 35, N 1. P. 19–36. <https://doi.org/10.18489/sacj.v35i1.1219>.
9. Кононюк А.Е. Общая теория понятий. Киев: Освіта України, 2014. 512 с. URL: https://ecat.ust.edu.ua/ft/Theory_Of_Terms_1.pdf.
10. Hogan A. et al. Knowledge graphs. Cham: Springer, 2021. 237 p. <https://doi.org/10.1007/978-3-031-01918-0>.
11. A dictionary of computer science. 7th ed. Butterfield A., Ekembe Ngondi G., Kerr A. (Eds.). Oxford University Press, 2016. 1211 p. <https://doi.org/10.1093/acref/9780199688975.001.0001>.
12. Encyclopedia of complexity and systems science. Meyers R.A. (Ed.). New York, NY: Springer, 2009. 10453 p. <https://doi.org/10.1007/978-0-387-30440-3>.
13. What is a knowledge graph? URL: <https://www.ontotext.com/knowledgehub/fundamentals/what-is-a-knowledge-graph> (Last accessed 12/21/2023).
14. Khan A. Knowledge graphs querying. *ACM SIGMOD Record*. 2023. Vol. 52, N 2. P. 18–29. <https://doi.org/10.1145/3615952.3615956>.
15. Bergman M.K. A common-sense view of knowledge graphs. Adaptive information. Adaptive innovation. Adaptive infrastructure blog, July 2019. URL: <http://www.mkbergman.com/2244/a-common-sense-view-of-knowledge-graphs/>.
16. Bonatti P.A., Decker S., Polleres A., Presutti V. Knowledge graphs: New directions for knowledge representation on the semantic web (Dagstuhl Seminar 18371). *Dagstuhl Reports*. 2018. Vol. 8, Iss. 9. P. 29–111. <https://doi.org/10.4230/DagRep.8.9.29>.
17. Ehrlinger L., Wöß W. Towards a definition of knowledge graphs. Martin M., Cuquet M., Folmer E. (Eds.). *Joint Proc. of the Posters and Demos Track of the 12th International Conference on Semantic Systems — Semantic and the 1st International Workshop on Semantic Change and Evolving Semantics (SuCCESS) Co-Located with the 12th International Conference on Semantic Systems (Semantic)* (Leipzig, Germany, Sept. 12–15, 2016). CEUR Workshop Proceedings. 2016. Vol. 1695. URL: <http://ceur-ws.org/Vol-1695/paper4.pdf>.
18. Noy N., Gao Y., Jain A., Narayanan A., Patterson A., Taylor J. Industry-scale knowledge graphs: lessons and challenges. *Communications of the ACM*. 2019. Vol. 62, Iss. 8. P. 36–43. <https://doi.org/10.1145/3331166>.
19. Палагин А.В., Крывий С.Л., Петренко Н.Г. Онтологические методы и средства обработки предметных знаний. Луганск: Изд-во ВЛУ им. В. Даля, 2012. 324 с.

20. Hogan A. Knowledge graphs: Research directions. Manna M., Pieris A. (Eds.). Reasoning Web. Declarative Artificial Intelligence. Reasoning Web 2020. *Lecture Notes in Computer Science*. 2020. Vol. 12258. P. 223–253. Cham: Springer, 2020. https://doi.org/10.1007/978-3-030-60067-9_8.
21. Гладун В.П. Процессы формирования новых знаний. София: СД «Педагог 6», 1994. 192 с.
22. Guarino N. Formal ontology in information systems. *Proc. of FOIS'98* (June 6–8, 1998, Trento, Italy). Amsterdam: IOS Press IOS Press, 1998. P. 3–15.
23. Gómez-Pérez A., Fernández-López M., Corcho O. Ontological engineering. 1st ed. London: Springer, 2004. 404 p. <https://doi.org/10.1007/b97353>.
24. Гаврилова Т.А. Онтологический подход к управлению знаниями при разработке корпоративных информационных систем. *Новости искусственного интеллекта*. 2003. № 2. С. 24–30.
25. Гаврилова Т.А. Базы знаний интеллектуальных систем. СПб.: Питер, 2001. 384 с.
26. Horridge M., Brandt S. A practical guide to building OWL ontologies using Protégé 4 and CO-ODE tools. Ed. 1.3. The University of Manchester, March 24, 2011. URL: https://mariajulianadascalu.com/wp-content/uploads/2014/02/owl-cs-manchester-ac-uk_-eowltutorialp4_v1_3.pdf.
27. Гладун В.П. Партнерство с компьютером. Киев: Port-Royal, 2000. 128 с.
28. Інструкція користувача когнітивної ІТ-Платформи «ПОЛІЕДР». Київ: Національна академія наук України, 2020. 71 с. URL: https://storage.ulif.org.ua/storage/instructions/polyhedron_instruction.pdf (дата звернення: 01.03.2023).
29. Borgo S., Ferrario R., Gangemi A., Guarino N., Masolo C., Porello D., Sanfilippo E.M., Vieu L. DOLCE: A descriptive ontology for linguistic and cognitive engineering. *Applied Ontology*. 2022. 17(1) P. 45–69. <https://doi.org/10.3233/AO-210259>.
30. Грушвицкий Р.И., Мурсаев А.Х., Угрюмов Е.П. Проектирование систем на микросхемах программируемой логики. СПб.: БХВ. Петербург, 2002. 608 с.
31. Petrenko M., Cohn E., Shchurov O., Malakhov K. Ontology-driven computer systems: Elementary senses in domain knowledge processing. *South African Computer Journal*. 2023. Vol. 35, N 2. P. 127–144. <https://doi.org/10.18489/sacj.v35i2.17445>.
32. Petrenko M., Sofiyuk A. On one approach to the transfer of an information structures interpreter to PLD-implementation. *Upravlyayushchie Sistemy i Mashiny*. 2003. Vol. 188(6). P. 48–57. URL: <https://cdn.e-rehab.pp.ua/u/usim-2003-petrenko-sofiyuk.pdf>.
33. Kurgaev A.F., Petrenko M.G. Processor structure design. *Cybernetics and Systems Analysis*. 1995. Vol. 31, N 4. P. 618–625. <https://doi.org/10.1007/BF02366417>.

O. Palagin, M. Petrenko, K. Malakhov

CHALLENGES AND ROLE OF ONTOLOGY ENGINEERING IN CREATING THE KNOWLEDGE INDUSTRY: A RESEARCH-RELATED DESIGN PERSPECTIVE

Abstract. The article discusses models and mechanisms of transdisciplinary research in the perspective of creating clusters of disciplinary convergence and scientific theories, formal representation of knowledge, and the formation of the knowledge industry using a unified ontology engineering devkit. Special attention is focused on the possibilities of research-related design for the creation of new knowledge and technologies. A conceptual analysis of ontology engineering is conducted, indicating the importance and effectiveness of applying ontology knowledge and mechanisms in solving user tasks. The domain ontology, its ontology graph, and their essential differences and advantages from owl ontologies are examined in depth. An example of the application of task ontology to the design of programmable microchip computing devices is proposed.

Keywords: knowledge industry, ontology engineering, transdisciplinary research, convergence clusters, research-related design, new knowledge and technologies, domain ontology, ontology graph, task ontology.

Надійшла до редакції 16.01.2024