

УДК 02:(031)

Штерн І. В.

МОДЕЛІ РЕПРЕЗЕНТАЦІЇ ЗНАНЬ І СТРУКТУРУВАННЯ ІНФОРМАЦІЙНОГО ПРОСТОРУ В СУЧАСНИХ ГУМАНІТАРНИХ ДИСЦИПЛІНАХ ЯК ПРОТОТИПНІ КОНФІГУРАЦІЇ ДЛЯ ПРОЕКТУВАННЯ ГУМАНІТАРНИХ БАЗ ЗНАНЬ

Сучасне гуманітарне знання відзначається, в певному розумінні, аномальною структурою та непередбачуваною динамікою. Це спонукає до пошуку нових підходів у проектуванні інфраструктури великомасштабних баз знань (БЗ) з гуманітарних дисциплін, призначених для ефективного впровадження у глобальні світові інформаційні мережі (електронні бібліотеки, віртуальні університети, системи дистанційного навчання). У статті обґрунтовано гіпотезу про те, що одним з таких універсальних засобів репрезентації гуманітарних царин в БЗ є глобальна енциклопедична організація знаннєвого простору. Визначено та систематизовано нестандартні енциклопедичні моделі, притаманні саме гуманітарній сфері (на відміну від дисциплін, що вивчають "фізичний світ"). Вони розглядаються як можливі прототипні конфігурації для адекватної репрезентації базових знань та відтворення інфраструктури гуманітарних предметних галузей при проектуванні великомасштабних повнотекстових БЗ.

Дослідження фінансується Державним фондом фундаментальних досліджень Міністерства освіти та науки України (Проект 7.4/79). Автор користується нагодою висловити найщирішу подяку цій установі за підтримку даного наукового напрямку.

Вступ. Проблеми організації знань у контексті сучасних комунікативно-інформаційних технологій: гуманітарні дисципліни

Стрімкий розвиток потужних когнітивно-орієнтованих інформаційних технологій на рубежі ХХ—ХХІ ст. стає принципово важливою передумовою та водночас надійним інструментальним підґрунтям для формування нових соціокультурних парадигм у розбудові цілісного знаннєвого простору, у становленні нових сфер та способів використання знань. Йдеться, звичайно, насамперед, про глобальні телекомунікаційні мережі типу Internet, які реалізують концепцію "відкритого знання" та забезпечують нові види інформаційного сервісу, що суттєво виходять за межі традиційного інформаційного пошуку — електронні бібліотеки, віртуальні університети, системи дистанційного навчання тощо.

Нові парадигми створення інформаційного простору та споживання знань потребують нових підходів до комплектування і подання знання. Це стосується, перш за все, змінення його "операційного" статусу — із сукупності статичних документальних блоків, систематизованих за прийнятими у бібліотечній практиці класифікаційними критері-

ями, знання має перетворитися на активну інформаційну структуру. Отже, на передній план виступає проблема конструктивної організації всіх шарів та типів знання, що мають бути представленими в інформаційному просторі. Центральне місце серед них, на наш погляд, посідатиме наукове знання, зокрема гуманітарні дисципліни, що відзначаються досить неординарною структурою та непрогнозованою динамікою і, на відміну від дисциплін, які вивчають "фізичні світи", погано піддаються традиційній жорсткій формалізації. Ці характерні риси гуманітарної царини є принципово важливими для проектування великомасштабних повнотекстових баз знань, оскільки саме вони визначатимуть "м'які" концепції та технології для розбудови дисциплінарної інфраструктури, вибору репрезентативних форматів та побудови пошукових стратегій [1].

Отже, мета даного дослідження — по-перше, з'ясувати суттєві для моделювання концептуальні та структурні властивості гуманітарних знань, що кардинально відрізняють їх від знань "про фізичний світ"; по-друге, показати доцільність енциклопедичного підходу до розв'язання концептуальних та структурних колізій під час репрезентації знань, і, по-третє, реконструювати та систематизувати типові енциклопедичні моделі, що можна викорис-

тати як прототипні конфігурації для адекватного відтворення інфраструктури базових знань у гуманітарних предметних галузях під час проектування великомасштабних повнотекстових БЗ (див., наприклад, [2]).

1. Загальна структура та динаміка гуманітарної царини. Основні опозиції гуманітарного та ф'зичного світів на концептуальному та структурному рівнях

Для розроблення "м'якої" парадигми репрезентації гуманітарних знань в інформаційному просторі суттєвими є такі чотири чинники, що характеризують концептуальні та структурні властивості гуманітарного знання:

- неоднорідна концептуальна структура гуманітарних знань;
- нестабільна дисциплінарна матриця гуманітарної царини;
- деформалізована структура одиниць знань та нерегламентований стиль їх тлумачення;
- герменевтичний принцип організації цілісного простору знань.

Коротко розглянемо ці чинники.

Чинник перший. Концептуальна (проблемно-тематична) неоднорідність, "розмитість" знанневого простору в гуманітарних дисциплінах зумовлюється багатомасштабністю, багаторівневістю, багатаспектністю явищ, що вивчаються, розмаїттям концепцій, теорій та метатеорій, поліфонією точок зору на одні й ті ж самі проблеми та явища тощо. Подібні спостереження накопичені у штучному інтелекті, наукознавстві, інформатиці, документах істині, термінографії тощо. Зокрема, проблема аномальності гуманітарних знань у порівнянні з іншими науками актуалізується в тому напрямі штучного інтелекту, який називають "штучний інтелект у гуманітарних науках" ([3—5] та ін.). У програмній праці В. К. Фінна [3] з цього приводу наголошується, що "гуманітарні науки є класичним прикладом тієї галузі людського знання, яку важко формалізувати". Це пояснюється, на думку автора, неточністю, метафоричністю та образністю мови гуманітарних наук, наявністю великої кількості ідей, що не мають точних визначень, відсутністю контролю за ступенем правдоподібності висновків тощо. Важливою передумовою формалізованого опису гуманітарних знань автор справедливо вважає створення основ теорії гуманітарних понять, а також розв'язання проблеми розгалуження понять, що мають визначення та не мають таких.

Чинник другий. Особливістю дисциплінарної організації гуманітарної царини є принципова нестабільність дисциплінарної матриці та неусталеність міждисциплінарних кордонів, що взагалі перешкоджає будь-якому строгому впорядкуванню та сегментації гуманітарного простору. В сучасній

науці надзвичайна лабільність гуманітарної царини спричиняється, принаймні, двома обставинами.

По-перше, з'явилося багато нових способів утворення гуманітарних дисциплін, напрямів, предметних галузей. Так, якщо більш традиційні гуманітарні дисципліни формувалися навколо життєво важливих ментальних феноменів (наприклад, психологія — навколо сфери психічного, соціологія — навколо сфери людських колективів, етнографія — навколо сфери культурної спільноти), то гуманітарні дисципліни наступних генерацій створювалися навколо професійної діяльності (наприклад, маркетинг, управління), нових методологій аналізу об'єктів або нових теоретичних конструктів (наприклад, біхевіористська психологія, що вивчає поведінку через співвідношення стимулів та реакцій) тощо. По-друге, домінуючою тенденцією сучасного розвитку гуманітарної царини є утворення гібридних, інтердисциплінарних галузей знань. За приклад може слугувати комплекс гібридних лінгвістичних дисциплін та напрямів, що утворилися у лінгвістичному просторі протягом другої половини ХХ ст.: генеративна лінгвістика, комп'ютерна лінгвістика, комунікативна лінгвістика, герменевтична лінгвістика, лінгвістика тексту, когнітивна лінгвістика, теорія мовленнєвих актів, лінгвістична філософія, логічний аналіз мови тощо (див. [6]). Причиною створення такого широкого спектра дисциплін стала стрімка експансія в традиційний лінгвістичний простір концепцій та методів інших дисциплін. Подібні процеси є виявом характерного для ХХ ст. загальнонаукового руху до переділу кордонів та сфер компетенції наук на тлі глобальних парадигмальних зрушень, що є особливо актуальними для гуманітарної царини у цілому. Отже, завдяки дисциплінарним трансформаціям гуманітарна царина має непрогнозований характер і відкрита для поповнення новими дисциплінами.

Чинник третій. Принципова розбіжність у природі об'єктів фізичного світу та об'єктів гуманітарної царини, у методах їх вивчення, і, відповідно, у самих когнітивних процедурах, за допомогою яких утворюються знання, обумовлює розбіжність і у ментальному статусі типових одиниць знань, а саме: якщо типовими базисними одиницями знань про фізичний світ є поняття, то для гуманітарних знань такими одиницями будуть структури дещо іншої ментальної природи, які ми назовемо понятійними образами, або концептуальними гештальтами. Тлумачення понять мають переважно дефініційну форму; вони будуються за принципом "портретування" відповідної реалії або, умовно кажучи, "мовної візуалізації" її глибинних компонентів (див. [7] та ін.). На відміну від понять, тлумаченнями концептуальних гештальтів стають здебільшого не дефініції, а м'які дескрипції, або

топіки, тобто тексти з нерегламентованою та слабо регламентованою дискурсивною структурою. Дескриптивна форма репрезентації гештальтів дозволяє поєднувати знання різних рівнів завершеності, зіставляти різноманітні контексти, в які водночас включаються одні і ті ж самі гуманітарні реалії і з яких послідовно складається цілісний образ тлумачуваної гуманітарної реалії. Саме за цих причин у практиці гуманітарних дисциплін переважає контекстуальний спосіб введення і тлумачення базисних гуманітарних реалій, коли реалії спочатку взагалі не дається визначення, але її смисл нарощується поступово, протягом всього викладу матеріалу (див., наприклад, культурологічні праці В. фон Гумбольдта та М. Бахтіна).

Чинник четвертий. Відмінність між репрезентацією знань про фізичний світ та репрезентацією гуманітарних знань виявляється також у способах організації цілісного знанневого простору.

Простір знань про фізичний світ на формальному рівні, тобто на рівні структурної презентації одиниць знань та логіко-понятійних відношень між ними, може бути представлений переважно дискретною та адитивною структурою, що відбиває принцип послідовного нарощення знань у відповідній предметній царині; при подальшій формалізації така структура може вкладатися в ієрархічні схеми та жорсткі класифікаційні фасети. Натомість, цілісний знанневий простір у гуманітарних науках має суто континуальний характер та організований за герменевтичним принципом. Це означає, що сукупне знання, представлене у цілісному знанневому просторі, виводиться із знань, представлених у певних знанневих одиницях та блоках знань, водночас з тим, що знанневі одиниці та блоки знань визначаються через той же самий цілісний знанневий контекст. З формального погляду знанневий простір гуманітарних дисциплін нагадує мультиграф, насичений циклічними структурами, і може бути репрезентований у комп'ютерному середовищі у вигляді гіпертексту.

I 2. *Енциклопедичний принцип організації знанневого простору та сучасні типові енциклопедичні моделі репрезентації гуманітарних знань*

Очевидно, що зазначені в попередньому розділі концептуальні та структурні властивості гуманітарних дисциплін вимагають специфічних форм їх репрезентації. Саме енциклопедичний принцип репрезентації знань [8], в певному розумінні, стає ключем до розв'язання структурних аномалій та колізій гуманітарної царини.

Це пояснюється тим, що, по-перше, знанневий простір в енциклопедії має "природну" гіпертекстову структуру (подібно до Біблії). Він створюється завдяки крос-референціям, тобто логіко-понятій-

ним зв'язкам між енциклопедичними статтями, які оперезують тлумачувані поняття у такий спосіб, що, вирушаючи у навігацію крізь цей простір від певного поняття за маршрутами, визначеними крос-референціями, неодмінно повертаєшся до вихідного поняття, подібно до бумеранга. По-друге, виявляється, що енциклопедичні структури релевантні не тільки для опису усталених реалій, але й для відображення динамічних процесів в описуваній предметній галузі, що, як вже зазначалося, є принципово важливою рисою сучасного гуманітарного дискурсу. По-третє, енциклопедичний опис передбачає презентацію узагальненої, резюмуючої інформації про предметну царину, яка охоплює найрізноманітніші концептуальні пласти — ситуаційні, фактажні, подійні, інструментальні тощо — в м'якій дескриптивній формі. Саме такі і деякі інші особливості енциклопедичної організації знанневого простору (див. [8], [9]) корелюють з розмитою інфраструктурою гуманітарних дисциплін.

Таким чином, енциклопедична організація може бути використана як глобальна прототипна конфігурація для відтворення інфраструктури відповідної гуманітарної дисципліни при проектуванні баз знань, що є носіями базових або узагальнюючих гуманітарних знань.

Від глобальних прототипних конфігурацій перейдемо до розгляду локальних енциклопедичних моделей, типових для сучасних гуманітарних дисциплін. Такі моделі були реконструйовані та систематизовані нами раніше [5], тому у даному разі обмежимося лише короткою репрезентативною характеристикою деяких з них та прикладами. До них належать:

1. Послідовно розгортуваний монотематичний дискурс з блочною структурою знань, побудований на герменевтичних принципах тлумачення; знанневий простір у цьому разі організовано у вигляді зіцплення взаємопов'язаних топіків, де поєднуються деревоподібна і континуальна схеми логічного впорядкування блоків, що забезпечують композиційну цілісність викладу. Така енциклопедична модель є класичною, а її прикладом може слугувати енциклопедичний словник з античної філософії О. Ф. Лосева (докладніше див. [5]).

2. Ретроспективно-етимологічний дискурс, де тлумачення реалій побудовано на етимологічних реконструкціях, в яких простежуються ретроспективні глибинні зв'язки між реаліями.

Тлумачення у даному разі представляє собою відтворення концептів-архетипів та відображення їх еволюції. Таку енциклопедичну модель можна знайти, зокрема, у фундаментальній праці Е. Бенвеніста "Словник індоєвропейських соціальних термінів" (див. докладніше у [5]), де об'єктами тлумачення є культурні концепти — соціальні та

культурні реалії індоевропейської епохи: " родина, друг, враг, чужий " тощо.

3. Циклічний дискурс, де об'єктами тлумачення стають світові проблеми людства, актуальні для певних часових відрізків, представлений у довіднику ЮНЕСКО "Енциклопедія світових проблем та людського потенціалу" (докладно див. [5]). Дискурс складається з замкнених циклів — зачарованих кіл проблем. Кожний цикл подає певне проблемне явище, яке описується у вигляді причинно-наслідкових ланцюжків, що створюють своєрідну семантичну петлю проблем. Наприклад: "відчуженість — молодіжні банди — контроль над місцевістю з боку кримінальних елементів — психологічний стрес, спровокований життям у міському середовищі — нахабна поведінка — розпад родини — відчуженість".

4. Поліпарадигмальний та політематичний дискурс, що складається з різноманітних взаємодіючих між собою дисциплінарних та тематичних зон. Такого роду енциклопедичні моделі необхідні для репрезентації знань у гібридних гуманітарних дисциплінах, типових для сучасної гуманітарної царини. Подібну модель було розроблено нами для створення есеїстичної енциклопедії з сучасних гібридних лінгвістичних дисциплін [6], утворення яких призвело до формування нової дисциплінарної матриці в лінгвістиці і виникнення великого блоку цілком нових реалій та категорій. Маючи на меті не просто зафіксувати нові гібридні лінгвістичні реалії, а й відтворити їх динаміку у контексті сучасних парадигмальних зрушень та нової дисциплінарної матриці, ми розробили спеціалізовану багаторівневу методологію, що складається з п'яти рівнів репрезентації знань:

- презентації лексикона;
- тлумачення реалій;
- фіксації логіко-понятійних зв'язків між реаліями;
- парадигмальної та дисциплінарної атрибуції реалій;
- презентації узагальнюючої когнітивно-тематичної карти.

Принципово важливими концептуальними та структурними особливостями розробленої методології є:

- розгортання енциклопедичного опису навколо "парадигмально-дисциплінарних меридіанів". Для відтворення парадигмальної динаміки реалій, спричиненої інтенсивним розвитком міждисциплінарних зв'язків у гуманітарній царині, з одного боку, та, з другого — кібернетичною трансформацією науки, у кожній енциклопедичній статті передбачено парадигмальну та дисциплінарну атрибуції тлумачуваних реалій;

- організація цілісного енциклопедичного простору як сукупності зв'язаних між собою концентричних структур, де концентрами є головні, "титульні" реалії певних тематичних груп, а периферійні реалії-сателіти утворюють глосарій до тлумачення титульних реалій;

- комбінація протилежних засобів енциклопедичного "портретування" (тлумачення) реалій: з одного боку, використання топіків, тобто розгорнутих, деталізованих текстів з нерегламентованою структурою для тлумачення центральних титульних реалій; з другого — дефініційних форм для тлумачення периферійних реалій-сателітів;

- на рівні презентації узагальнюючої когнітивно-тематичної карти тлумачуваних реалій об'єднуються у великі тематичні розділи, як-то "мова як знакова система"; "мова та смисл"; "мова як механізм втілення думки"; "мова як засіб спілкування"; "мова та когнітивні процеси"; "логічний аналіз мови"; "мова та структура знань" тощо;

- універсальна схема енциклопедичного "портретування" реалії містить такі інформаційні блоки: 1) визначення; 2) детальну, розширену інформацію про сутність та структуру реалії; 3) характеристику компонентів або об'єктів, які входять до сфери реалії, або на які розповсюджується його вплив; 4) додаткові стандартні та нестандартні властивості реалії; 5) інформацію про входження тлумачуваної реалії до різноманітних контекстів (наприклад, застосування); 6) модифікації, версії; 7) перспективи розвитку; 8) відомості про генезис та еволюцію реалії; 9) демонстрацію способу, яким отримано знання про реалії; 10) дискусії щодо останнього (наприклад, щодо трактування даної реалії у різних школах, напрямках тощо); 11) ілюстрації; 12) приклади тощо.

3. Висновки

Отже, в даному дослідженні відображено ті особливості організації гуманітарного знання, що є важливими для його репрезентації у сучасному інформаційному просторі. З'ясовано суттєві для моделювання концептуальні та структурні властивості гуманітарних знань, що кардинально відрізняються від знань "про фізичний світ" та показано доцільність енциклопедичного підходу до розв'язання концептуальних і структурних колізій під час репрезентації знань. Нарешті, систематизовано типові енциклопедичні моделі, що можна використати як прототипні конфігурації для адекватного відтворення інфраструктури базових знань у гуманітарних предметних галузях під час проектування великомасштабних повнотекстових БЗ.

1. Штерн И. Б. Проблемы систематизации знаний в контексте современных когнитивных ориентации и коммуникативно-информационных технологий // Искусственный интеллект. Спец. выпуск: Материалы Восьмой Межд. конф. KDS-99 "ЗНАНИЯ — ДИАЛОГ — РЕШЕНИЕ" (Крым, 1999 г. Донецк 1999,— С. 92—97.
2. Lena! D. B. СУС: A Large-Scale Investment in Knowledge Infrastructure // Communications of the ACM. 1995.— Vol. 38.—№ 11.
3. Финн В. К. Гуманитарное знание и искусственный интеллект // Тр. Национальной конференции по искусственному интеллекту КИИ-94.— Рыбинск, 1994.— С. 1—4.
4. Постелов Д. А. Что привносит искусственный интеллект в гуманитарные науки? // V Национальная конференция по искусственному интеллекту КИИ-96.— Казань, 1996.
5. Штерн И. Б. Представление гуманитарных знаний vs представление знаний о физическом мире: единицы, связи и структурные модели введения реалій // Новости искусственного интеллекта.— М., 1997.— Вып. 3.— С. 57—68.
6. Штерн И. Б. Вибрані топіки та лексикон сучасної лінгвістики. Енциклопедичний словник.— Київ: АРТ-ЕК, 1998.— С. 358.
7. Штерн И. Б. Канонические знания в модели исследователя: энциклопедия как информационная и как креативная среда // Теория и системы управления. Известия РАН.— М.: Наука, МАИК, 1996 — № 3 — С. 153—159.
8. Штерн И. Б. Энциклопедия в зеркале когнитивной парадигмы // Научно-техническая информация.— М.: ВИНТИ РАН И МИННАУКИ РФ, 1998 — Сер. 2. Информационные процессы и системы.— № 1.— С. 1—11.
9. Штерн И. Б. Энциклопедические модели в искусственном интеллекте // Теория и системы управления. Известия РАН — М.: Наука, МАИК, 1999.— № 1 — С. 6—59.

Stern I. B.

THE PROTOTYPICAL MODELS OF KNOWLEDGE REPRESENTATION AND INFRASTRUCTURE OF MODERN HUMANITIES FOR DESIGNING THE LARGE SCALE KNOWLEDGE BASES OF THE HUMANITIES

The structure and dynamics of modern humanities is principally anomalistic and in a sense paradoxical. That is why a very special approach to designing the large scale humanities knowledge bases (KB) is required to provide thier successful introduction into the global worldwide information networks (electronic libraries, virtual universities, distant education etc.). The most appropriate universal tool for presentation of humanities areas in the KB of the kind is suggested as the global encyclopedic organization of knowledge space. The non-standard encyclopedic patterns inherent exclusively in humanities in opposite to ones inherent in the "physical world" are discovered and systematized. They are designed as the prototypical configurations for the adequate representation of the infrastructure of the basic knowledge of humanities in the large scale full-text KB.