

УДК 004.94:[004.822+007.5

DOI: 10.31673/2412-9070.2024.060447

І. М. ПАТРАКЕЄВ, канд. техн. наук, доцент;

ORCID: 0000-0002-0448-8790

Р. К. НАФЄЄВ, канд. фіз.-мат. наук, доцент,

ORCID: 0000-0003-2721-9718

Державний університет інформаційно-комунікаційних технологій, Київ

ОНТОЛОГІЧНИЙ ПІДХІД ДО МОДЕЛЮВАННЯ СИСТЕМНОГО ГЕОІНФОРМАЦІЙНОГО АНАЛІЗУ

Одним з основних завдань концептуального моделювання є отримання з великого масиву різноманітної інформації даних, які необхідні для вирішення конкретних завдань або проблем. Важливим елементом сучасних інформаційних технологій інженерії знань є онтологія. Онтологія описує факти та взаємозв'язки у певній предметній області. Це один із важливих напрямків у галузі штучного інтелекту, який дозволяє систематизувати й організувати інформацію для ефективного її використання, забезпечити аналіз семантики інформації про предметну сферу задля її ефективного використання та подання. У статті розглянуто онтологічний підхід до отримання предметно-орієнтованих знань із текстів оригінальною мовою, який передбачає використання формально-логічних методів для структурування та аналізу інформації. Такий підхід дозволяє створити концептуальну модель, ідентифікувати сутності, визначити зв'язки між ними, а також встановити правила, які описують предметну область. Дослідження виконане на прикладі дефініції «системний геоінформаційний аналіз». Використання системного онтологічного підходу сприяє формалізації предметної сфери, уточненню понять та визначенню загальної термінологічної бази. Показано, що програмне середовище Protégé є ефективним інструментом для структуризації та аналізу предметної області, термінів і взаємозв'язків між ними. Проведено концептуальний аналіз терміну "системний геоінформаційний аналіз", який надає характеристики його властивостям, показує класи основних понять і відносин між класами для побудови онтології досліджуваної предметної сфери. Показано, що наведений інструментарій дозволяє розглядати терміни і поняття предметної сфери у розвитку, який подається ланцюжком квантів знань про досліджувані поняття або об'єкт. У роботі показано важливість онтологічного підходу для розвитку та вдосконалення методів системного геоінформаційного аналізу, що в свою чергу, сприятиме ефективнішому прийняттю рішень у різних сферах, де просторовий аналіз є критично важливим.

Ключові слова: онтологія, інженерія знань, складно-структурована предметна сфера, моделювання знань, системний геоінформаційний аналіз, системний аналіз.

Вступ

Сьогодні геоінформаційні системи (ГІС) відіграють важливу роль у різних сферах, від екологічного моніторингу до міського планування і стратегічного управління ресурсами. Проте, із зростанням складності даних, виникає потреба в нових методах їх аналізу та інтеграції. Використання онтологічного підходу дозволяє структурувати знання і підвищити точність та ефективність аналізу геоінформаційних даних. У сучасних умовах існує потреба у системному підході до обробки геоінформаційних даних, що включає формування нових методів і моделей для об'єднання різних типів даних, їхньої обробки та аналізу. Системний геоінформаційний аналіз (СГА) являє собою підхід, який дозволяє розглядати просторові дані як комплексну систему з урахуванням численних факторів і взаємозв'язків.

Однією з основних проблем у дослідженні нових методів є відсутність чіткого понятійного апарату, що ускладнює сприйняття і застосування концепції. Введення онтології в аналіз ГІС

дозволяє формалізувати та уніфікувати ключові поняття та терміни, необхідні для СГА, а також забезпечити логічну структуру для подальшого використання. У цьому контексті онтологічний підхід є ефективним засобом для створення формальної моделі предметної області, що дозволяє врахувати взаємозв'язки між різними елементами геоінформаційних систем і структуровано представляти знання. Це сприяє кращому розумінню і обробці просторових даних, особливо в умовах постійного збільшення їхнього обсягу та різноманітності.

Основне призначення онтологій – інтеграція інформації. Онтології пов'язують два важливі аспекти: по-перше, вони визначають формальну семантику інформації, дозволяючи обробку цієї інформації комп'ютером і, по-друге, визначають семантику реального світу, що дозволяє на основі загальної термінології пов'язувати інформацію, представлену у вигляді, необхідному для комп'ютерної обробки, з інформацією, яка представлена у зручній формі для сприйняття людиною.

Аналіз літературних даних та постановка проблеми дослідження

Найчастіше онтологія деякого складного поняття (системи, об'єкта) має ієрархічну структуру, для побудови якої послідовно застосовується метод декомпозиції, коли система, об'єкт або поняття i -го рівня розбивається на складові частини, елементи (компоненти) рівня $i + 1$ і т. д. [1]. Вибір програмних засобів для онтологічного моделювання є важливим чинником, оскільки побудова онтологій процес складний і займає досить багато часу. Щоб полегшити його, в середині 90-х років почали створюватися перші програмні середовища для процесу розробки онтологій. Такі програмні засоби забезпечують інтерфейси, які дозволяють виконувати концептуалізацію, реалізацію, перевірку узгодженості понять предметної сфери. За останні роки кількість інструментів моделювання онтологій різко зросла (сайт консорціуму World Wide Web Consortium, W3C, наприклад, налічує понад 50 інструментальних засобів моделювання онтологій) [4]. Інженерію онтологій можна визначити як сукупність дій, що стосуються процесу розробки онтологій; життєвого циклу онтологій; методів побудови онтологій; набору інструментів і мов для їх побудови і підтримки.

Проблема визначення термінів і формулювання нового поняття, як "системний геоінформаційний аналіз" (СГА), пов'язана з кількома важливими аспектами:

- складність інтеграції різних дисциплін: СГА об'єднує елементи з кількох наукових сфер, таких як картографія, інформаційні технології, системний аналіз. СГА вимагає міждисциплінарного підходу, оскільки кожна з дисциплін має власні методології та понятійний апарат;

- відмінності між класичним і геоінформаційним підходами: класичний системний аналіз зазвичай орієнтований на абстрагування та моделювання процесів у загальному вигляді, тоді як геоінформаційний аналіз враховує геопросторовий контекст і специфіку даних. Це створює додаткові труднощі при адаптації існуючих методів системного аналізу до геоінформаційної сфери;

- онтологічні проблеми: визначення термінів для побудови онтології СГА вимагає чітких, логічно впорядкованих понять і взаємозв'язків між ними, оскільки СГА є новим і комплексним поняттям; створення відповідної онтології потребує додаткових досліджень, аби уникнути розбіжностей та забезпечити однозначність у його трактуванні.

Найбільш відомим інструментарієм інженерії онтологій є програма Protégé, як відкритий редактор онтологій і фреймворк для побудови баз знань. Платформа Protégé підтримує два основні способи моделювання онтологій за допомогою редакторів Protégé-Frames та Protégé-OWL. Онтологічне моделювання надає змогу з'ясувати, які концепції відповідають заданим визначенням та перевірити чи взаємно узгоджуються поняття та визначення в онтології.

Онтологія є формальним поданням концептуальних знань про предметну сферу та за своєю значністю співвідноситься з базою знань інтелектуальної інформаційної системи, а її побудова являє собою специфічну форму людського мислення. Мислення в процесі пізнання оперує судженнями, ствердженнями, поняттями та відношеннями між ними. Останні — це фундамент, заснований для побудови складової частки наукової теорії – онтологічної бази знань у визначеній предметній сфері. Процес проектування та розробки онтології предметної сфери

визначаються як онтологічний інжиніринг. Основи онтологічного інжинірингу викладені в [4]. Існують приклади розробки онтології в медицині [5], природокористуванні [6], техніці [7, 8, 9] та інших сферах.

Мета дослідження

Мета даної роботи – обґрунтувати визначення терміну "Системний геоінформаційний аналіз" (СГА) та представити науково обґрунтоване, логічно впорядковане тлумачення, яке відображає різноманітність і специфіку змісту цього поняття. В основі підходу – порівняльний аналіз класичного та геоінформаційного методів пізнання, що дозволяє виокремити особливості СГА. Крім того, в роботі розглядається побудова графа онтології для терміну СГА з використанням сучасних інформаційних технологій на прикладі використання редактора онтологій і баз знань Protégé.

Результати дослідження

Важливо зазначити, що на відміну від традиційного аналізу даних, СГА має справу з багатовимірними просторовими даними, які можуть мати як статичні, так і динамічні характеристики. Це включає інформацію про розташування, форму, площу, відстань і просторові взаємовідносини між об'єктами. З цією метою використовуються спеціалізовані програмні засоби, які дозволяють виконувати аналіз з урахуванням геометричної природи даних. Однак в даний час термін «системний геоінформаційний аналіз» є узагальненням багатьох видів аналізу і використовується для позначення не тільки аналізу в галузі геоінформатики, але й в галузі управління, кадастру, транспорту, екології, містобудування та інших сферах. Наприклад, у земельному кадастрі геоінформаційний аналіз реалізується для визначення меж земельних ділянок і оцінки їх вартості, тоді як у транспортній сфері — для оптимізації маршрутів та аналізу потоків транспорту.

Використовуючи системний підхід, геоінформаційний аналіз можна розглядати як складну систему, яка має певний склад та структуру термінів, визначень та відношень між ними.

Спочатку термін "онтологія" визначав науку про буття — розділ філософії, нині "онтологія" широко використовується у штучному інтелекті. Онтологію розуміють як систему понять (концептів, сутностей, класів), відносин між ними і правил операцій над ними у певній предметній сфері або, інакше кажучи, онтологія є специфікацією концептуалізації предметної сфери. Модель онтології предметної області Ξ може бути подано кортежем [8]:

$$\Xi = \langle \Theta, \Psi, \Lambda \rangle,$$

де $\Theta = \{ \theta_1, \theta_2, \dots, \theta_i, \dots, \theta_n \}$, $i = 1, n$, $n = \text{Card } \Theta$ – множина понять (концептів), яка є основою для побудови онтології предметної області;

$$\Psi = \{ \psi_1, \psi_2, \dots, \psi_k, \dots, \psi_m \}, \quad \Psi : \theta_1 \times \theta_2 \times \dots \times \theta_n, \quad k = 1, m,$$

де $m = \text{Card } \Psi$ – множина семантично значних відношень між концептами предметної області.

Відношення визначають тип взаємодії між поняттями. У загальному випадку відношення діляться на загальнозначущі і конкретні відношення заданої предметної області;

$\Lambda : \Theta \overset{\sim}{\sim} \Psi$ – кінцева множина функцій інтерпретації, які задано на концептах і/або відношеннях. Приватним випадком завдання множини функцій інтерпретації Λ є глосарій, який складено для множини понять Θ . Визначення поняття Θ_i , в загальному випадку, включає підмножину понять $\{\theta_{i-1}\}$, через які визначається Θ_i ; відношення Ψ_i , яке зв'язує Θ_i з $\{\theta_{i-1}\}$; та множину атрибутів (ознак), властивих Θ_i .

Онтологію предметної сфери за своєю значущістю можна порівняти з базою знань інтелектуальної інформаційної системи, а побудова онтології предметної сфери є специфічною формою мислення. Мислення в процесі пізнання оперує судженнями, ствердженнями, поняттями і відношеннями між ними. Останні є фундаментом, основою для побудови складової частини наукової теорії – інженерії знань у даній предметній сфері.

Стосовно предметної сфери СГА в якості декомпозиції переважним є визначення складу, властивостей, перелік моделей та методів аналізу просторової інформації, методів моделювання просторових об'єктів та ін.

Розглянемо більш детально підхід до моделювання знань СГА на основі методики, яка включає три етапи проектування [5]:

- попередній аналіз предметної сфери;
- побудова онтографа (онтологічного графа) предметної сфери. Під онтографом розуміється дводольний граф, вершинами якого є поняття, а дугами — відношення між ними;
- графічне (візуальне) проектування онтографа і складання формалізованого опису онтології предметної області.

Побудова множини Θ вважається найбільш важливим кроком у розробці онтології предметної сфери. Наступним кроком є впорядкування списку понять за деяким типом відношення "вище-нижче" на основі професійних знань розробника предметної сфери і, можливо, слід повторити деякі фрагменти процесу аналізу, які виконано на попередньому етапі.

Побудова множини Ψ також заснована на результатах етапу попереднього аналізу предметної сфери. Тобто, потрібно встановити для кожного елементу $\theta_i \in \Theta$ семантичне відношення Ψ_k з елементом $\theta_j \in \Theta$, $\theta_i \Psi_k \theta_j$, $i, j = 1, n$, $i \neq j$, $k = 1, m$.

Іншими словами, необхідно побудувати відношення, які зв'язують елементи онтографа. Вершиною (або вершинами) онтографа є родові поняття, які не мають надкласу, а нижчий рівень є конкретними поняттями. Для створення онтології визначення СГА було застосовано інструментальний засіб моделювання онтології Protégé 5 [6].

Розробка онтології для складної предметної області, як правило, є трудомістким та ітеративним процесом, тому до кінця 1990-х років відзначався недолік опрацьованих методологій для створення онтологій [4]. До теперішнього часу здобули популярність засоби створення онтологій, такі, як Ontoligua, OntoEdit, OilEd, Proterger, WebDeso. Середовище розробки Ontoligua призначено для колективного використання системи базових знань для побудови онтологій. Середовище розробки OntoEdit призначено для проектування, пристосування та імпорту/експорту моделей знань у форматах RDF, DAML + OIL, Flogic для/з прикладних систем. Редактор онтологій OilEd більшою мірою призначений для перевірки розроблених онтологій на узгодженість. Система Protégé є бібліотекою, що надає доступ для перегляду баз знань і дозволяє редагувати та нарощувати бази знань. Система WebDeso призначена для створення онтологій предметної області. У таблиці наведено порівняльну таблицю засобів управління онтологіями.

Характеристика засобів управління онтологіями

Засіб управління технологією	Розробник	Модель подання знань	Метод моделювання понять та відношень	Засіб реалізації
1	2	3	4	5
<i>Ontoligua</i>	Лабораторія систем знань Університету Стенфорда	Логіка першого порядку	Таксономія	Клієнт—HTML інтерфейса
<i>OntoEdit</i>	Компанія Ontoprise GmbH	Фреймова модель знань	Складна таксономія та ієрархія	Java програма
<i>OilEd</i>	Проект "On_To_Knowledge_Project", Університет Манчестера	Дескрипційна логіка (description logics)	Складна таксономія та ієрархія	Java програма

<i>Proterger</i>	Лабораторія медичної інформатики Університету Стенфорда	Фреймова модель знань	Складна таксономія	Java програма
<i>Web-Deso</i>	"Система інтеграції знань", Інститут інформатики та автоматизації	Об'єктно-орієнтовані мережі	Аксномія, ієрархія, асоціативні відношення	Клієнт – HTML інтерфейс, Java script

Результати дослідження

На рис.1 показано діалогове вікно програми, в якому подано основні класи та підкласи термінів, що складають поняття СГА. Всього налічується 12 класів та підкласів, які складають узагальнений онтологічний граф (рис. 2). Визначення елементів множини Ψ засновано на результатах етапу попереднього аналізу предметної сфери. Тобто, необхідно визначити семантичні відношення між класами або підкласами термінів, які визначають поняття СГА. На рис. 2 показано 22 семантичних відношення між класами (підкласами) онтологічного графа.

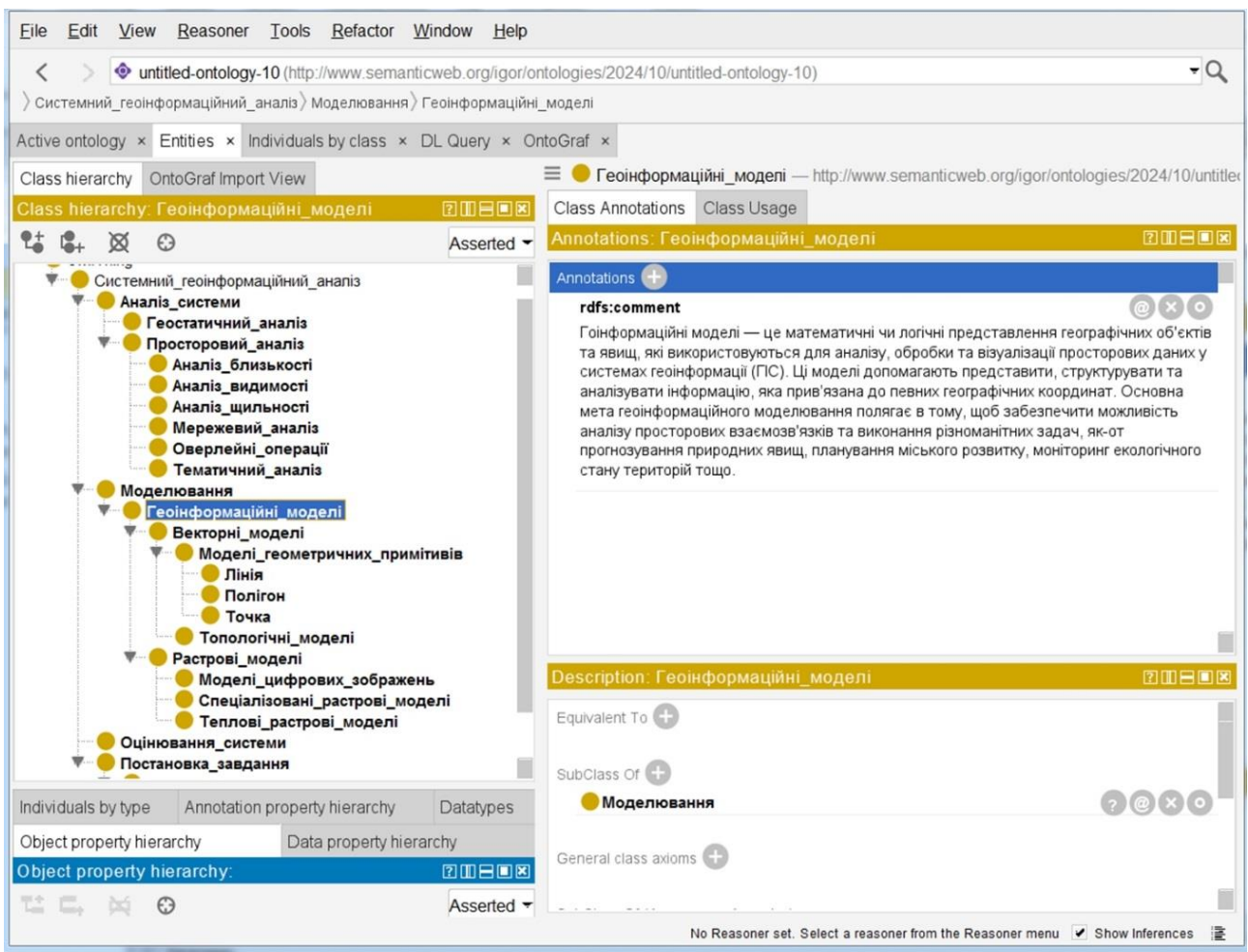


Рис. 1. Діалогове вікно програми Protégé 5.6.4: основні класи та підкласи

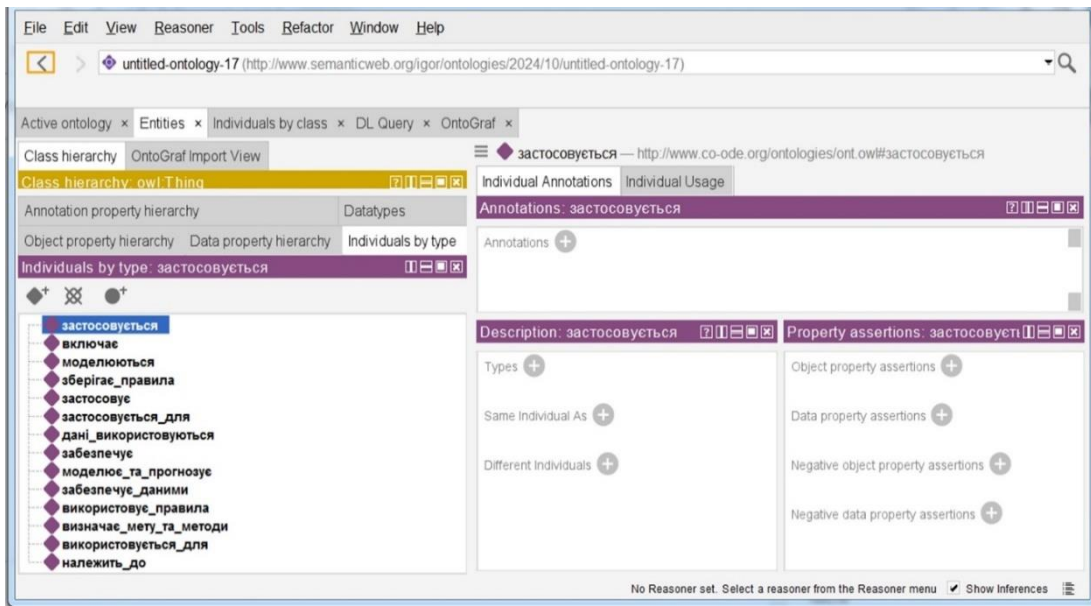


Рис. 2. Основні семантичні відношення між класами та підкласами термінів, які складають поняття СГА

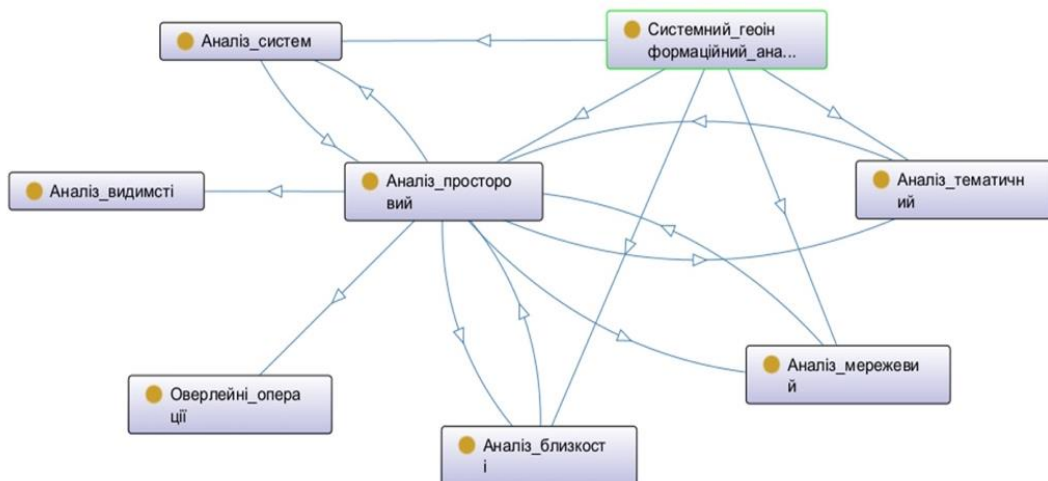


Рис. 3. Фрагмент взаємозв'язків між класами в онтології СГА

Нижче наведено фрагмент запису онтології на мові OWL 2 в Манчестерському синтаксисі [3], що описує ієрархію, показану на рис. 3.

```

class: ExpertGeosystem
  Annotations:
    label `Експертна геоінформаційна система`@ua
    label `Експертна система використовує експертні знання для забезпечення високої ефективності рішення неформалізованих або слабо структурованих завдань в узкій предметній сфері. Накопичення та організація експертних знань з застосуванням різноманітних моделей знань є властивістю всіх експертних систем`
    label `ExpertGeosystem`@en
class: GeospatialDatabase
  Annotations:
    label `База геопросторових даних`@ua
    label `База геопросторових даних забезпечує зберігання та обробку запитів до даних про просторові об'єкти, які подано у вигляді абстракцій: точки, лінії, полігони тощо`
    label `GeospatialDatabase`@en
  SubClassOf: ExpertGeosystem
class: KnowledgeBase
  Annotations:
    label `База знань`@ua
    label `База знань вміщує правила виведення та інформацію про експертні знання людини, ґрунтуючись на її досвіді та знаннях в узкій предметній сфері`
    label `KnowledgeBase`@en
  SubClassOf: ExpertGeosystem
  DisjointWith: GeospatialDatabase, ExpertGeosystem
    
```

Розроблена базова онтологія, яка описує основні компоненти та відношення між цими компонентами, що надає в сукупності цілісне подання про таке поняття як СГА. Узагальнена структура СГА подана у вигляді Онтографа, який налічує 26 термінів, на яких існує 25 відношень подано на рис. 4.

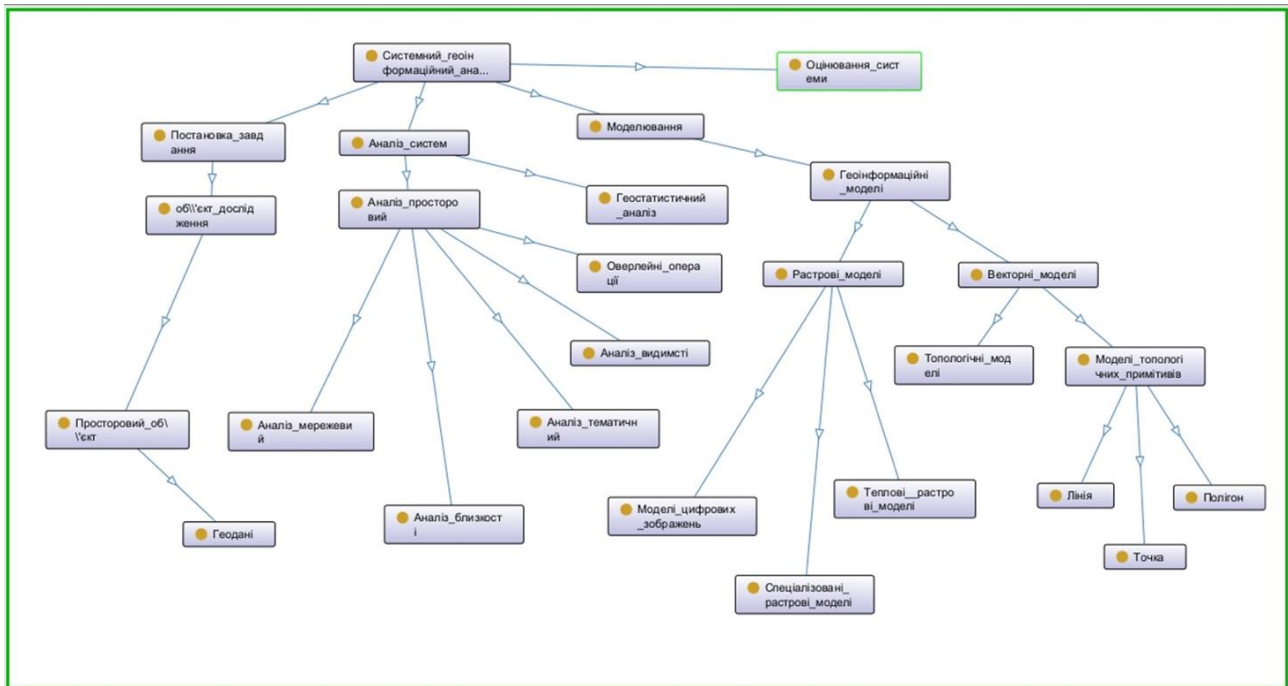


Рис. 4. Узагальнена структура поняття СГА, побудована із застосуванням засобів моделювання онтологій Protégé 5.6.4

Висновки

У роботі запропоновано підхід до створення та аналізу нових термінів та визначень на основі застосування онтологій. Зроблено стислий огляд програмних засобів для створення онтологій та показано їх обмеження і напрямки розвитку.

Проведено концептуальний аналіз терміну "системний геоінформаційний аналіз", який надає характеристики його властивостям, показує класи основних понять і відносин між класами для побудови онтології досліджуваної предметної сфери. Показано, що наведений інструментарій дозволяє розглядати терміни і поняття предметної сфери у розвитку, який подається ланцюжком квантів знань про досліджувані поняття або об'єкт.

Переваги запропонованого підходу до побудови тезауруса складного поняття:

- можливість автоматичного добудовування мережі асоціативних зв'язків, цілісної, що розвивається системи термінів і понять для побудови тезауруса геоінформаційного моніторингу;
- підвищення якості та ефективності побудови тезауруса терміна "система геоінформаційного моніторингу" за допомогою прийняття узгоджених рішень;
- можливість швидкої навігації по семантичній мережі термінів, відносин, асоціацій підвищують ступінь очікуваності в отримуваних результатах;

Таким чином, у роботі показано важливість онтологічного підходу для розвитку та вдосконалення методів системного геоінформаційного аналізу, що в свою чергу, сприятиме ефективнішому прийняттю рішень у різних сферах, де просторовий аналіз є критично важливим.

Список літератури

1. Патракеєв І.М. Лященко А.А. Онтологія та особливості компонентів геоінформаційного моніторингу за технологією баз геопросторових даних / Збірник наукових праць «Сучасні досягнення геодезичної науки та виробництва». – Львів: Видавництво Львівської політехніки. – 2015. – I (29). С.174-177.

2. Patrakeyev I.M. Логіко-онтологічний підхід до моделювання геоінформаційного моніторингу / Zbior raportow naukowych. «Inzynieria i technologia. Wspolczesna nauka. Nowy wyglad» (30.01.2015-31.01.2015) — Warszawa: Wydawca: Sp. z. o. o. «Diamond trading tour», 2015.- p.p. 22-28
3. Development of ontology 101: a guide to creating your first ontology / Natalya F. Noy (Natalya F. Noy) and Debor L. McGuinness, Stanford University, Stanford, California 94305
4. Asunci on Gomez-Perez, Mariano Fernandez-Lopez, Oscar Corcho. Ontological Engineering: With Examples from the Areas of Knowledge Management, E-commerce and the Semantic Web. Springer, 2004
5. Мінцер О.П., Денисенко С. В., Стрижак О. Є. Використання принципів медичної онтології для побудови сценарних моделей післядипломної освіти лікарів і провізорів // Медична інформатика та інженерія. - РОД.- № 2. - С. 18-23.
6. Моделювання і прогнозування стану довкілля: підручник / В.І.Лаврик, Б.М. Боголюбов, Л.Б. Полетаєва та ін. –К.:ВЦ "Академія", 2010. – 400с.
7. Patrakeyev I.M. Logic-ontological approach to geoinformation monitoring modeling. Zbior raportow naukowych.«Inzynieria i technologia. Wspolczesna nauka. Nowy wyglad» (30.01.2015-31.01.2015) — Warszawa: Wydawca: Sp.z.o.o. «Diamond trading tour».- Str. 22-28
8. Патракеєв І. М. Онтологічне дослідження міського середовища / Управління розвитком складних систем : зб. наук. праць // Київ. нац. ун-т буд-ва і архітектури ; гол. ред. Лізунов П. П. – Київ : КНУБА, 2015. – № 23. – С. 159-168.
9. Лященко А.А., Патракеєв І.М. Онтологія та особливості компонентів геоінформаційного моніторингу за технологією баз геопросторових даних / Збірник наукових праць «Сучасні досягнення геодезичної науки та виробництва». – Львів: Видавництво Львівської політехніки. – 2015. – I (29). С.174-177.

I. Patrakeev, R. Nafieiev

ONTOLOGICAL APPROACH TO SYSTEM GEO-INFORMATION ANALYSIS

One of the main tasks of conceptual modeling is to obtain from a large array of diverse information the data necessary to solve specific tasks or problems. Ontologies are an important element of modern information technologies of knowledge engineering. An ontology describes facts and relationships in a particular subject area. This is one of the important directions in the field of artificial intelligence, which allows to systematize and organize information for its effective use, to provide an analysis of the semantics of information about the subject area for its effective use and presentation.

The article examines the application of the principle of the ontological approach to obtaining subject-oriented knowledge from a set of natural language texts, their formal-logical presentation and application processing using the example of the definition "system geoinformation analysis". The application of the systemic ontological approach allows you to formally describe the subject area, clarify the concept of the subject area, and define a common terminological base. It is shown that the Protégé software environment is an effective means of structuring and analyzing the subject area, terms and relationships between them. A conceptual analysis of the term "systemic geo-informational analysis" was carried out, which gives characteristics to its properties, shows the classes of basic concepts and relations between classes for the construction of the ontology of the studied subject area. It is shown that the given toolkit allows considering the terms and concepts of the subject area in development, which is provided by a chain of quanta of knowledge about the studied concept or object. The paper shows the importance of the ontological approach for the development and improvement of methods of systematic geoinformation analysis, which in turn will contribute to more effective decision-making in various areas where spatial analysis is critically important.

Keywords: ontology, knowledge engineering, complexly structured subject area, knowledge modeling, system geoinformation analysis, system analysis.