

УДК 004.415

к.т.н., доцент Патракеєв І.М.,
ipatr@ukr.net, ORCID: 0000-0002-0448-8790,Денисюк Б.І.,
gis-knuba@ukr.net, ORCID: 0000-0003-1692-8551,

Київський національний університет будівництва і архітектури

DOI: 10.32347/2076-815x.2019.71.271-282

МОДЕЛЮВАННЯ БАЗ ЗНАНЬ В РЕДАКТОРІ ОНТОЛОГІЙ PROTÉGÉ

Найважливішою цінністю в сучасному інформаційному суспільстві є знання. Для їх отримання необхідні нові ефективні технології обробки інформації і перетворення її в потрібні знання. У багатьох випадках нові знання характеризуються неповнотою, неоднозначністю, невизначеністю і часто суперечливістю вихідної інформації, термінології або понять. У таких випадках здійснюється формалізований опис предметної області і формування бази знань для подальшого використання.

Стаття присвячена дослідженням в галузі збирання та подання знань у вигляді онтологій. Розглянуто види онтологій, відомі засоби розробки онтологій. Запропоновано способи побудови онтологій верхнього рівня і предметної області на основі апарату семіотичного моделювання. Онтології будуються у вигляді тезаурусів, що містять терміносистему і категорійно-понятійний апарат предметної області.

Для цілісного уявлення про складну систему або об'єкт, новому, ще не сталому понятті "система геоінформаційного моніторингу", а також для виявлення складу і основних компонентів і відносин між поняттями в сучасних умовах використовують програмні засоби онтологічного моделювання. Авторами розглянуто теоретичні аспекти розробки онтологій досліджуваної предметної сфери, проведено структурування і формалізацію знань щодо поняття "система геоінформаційного моніторингу" та визначено характеристики елементів онтології та їх значення. Розроблено класифікацію класів, підкласів, виявлено характеристики, які описують поняття "система геоінформаційного моніторингу" та розроблено структуру онтології предметної сфери у вигляді онтологічного графа.

На прикладі розробленої онтології показано, що геоінформаційний моніторинг, як метод пізнання, дозволяє проводити комплексні дослідження на основі яких здійснювати міждисциплінарне перенесення знань.

Ключові слова: система геоінформаційного моніторингу, онтологія предметної сфери, онтологічний граф, Protégé.

Вступ. На сьогоднішній день в області штучного інтелекту розроблено ряд засобів подання знань, одним з яких є *онтологія*. Онтологія – це детальна специфікація структури певної проблемної області, в нашому випадку такою проблемною областю є геоінформаційний моніторинг.

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, дозволяючи на основі загальної термінології пов'язувати інформацію, представлену у вигляді, необхідному для комп'ютерної обробки, з інформацією, представленою в зручній формі для сприйняття людиною.

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 так далі [1].

Вибір програмних засобів для онтологічного моделювання є важливим чинником, оскільки побудова онтологій є процесом складним 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 такі інтерфейси, які дозволяють виконувати концептуалізацію, реалізацію, перевірку узгодженості понять предметної сфери. За останні роки кількість інструментів моделювання онтологій різко зросло (сайт консорціуму *World Wide Web Consortium*, W3C, наприклад, налічує понад 50 інструментальних засобів моделювання онтологій). 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.

Найбільш відомим інструментарієм інженерії онтологій є Protégé, як відкритий редактор онтологій і фреймворк для побудови баз знань. Платформа Protégé підтримує два основних 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.

Метою статті є обґрунтувати визначення терміну "геоінформаційний моніторинг" та узагальнити класифікаційну схему видів моніторингових систем з визначенням в ній геоінформаційного моніторингу, розглянути побудову узагальненого графа онтології терміну "система геоінформаційного моніторингу" (СГМ) з застосуванням сучасних програмних засобів, на прикладі вільного відкритого редактора побудови онтологій та баз знань Protégé.

Основна частина. Моніторинг як спостереження є інструментом пізнання і може розглядатися з узагальнених філософських позицій. В даний час переважає концепція технологічного розгляду моніторингу. Це обумовлено технологічною диференціацією моніторингу за видами, об'єктами, цільовим призначенням та часом моніторингових спостережень і досліджень.

Застосування геоінформатики в багатьох прикладних сферах позначається виникненням нових понять, які зв'язані з ознакою "геоінформаційний", зокрема геоінформаційні системи, геоінформаційне картографування, геоінформаційний моніторинг тощо. Цю ознаку найчастіше розглядають як чисто технологічну, що вказує на нове інформаційно-технологічне середовище здійснення традиційної діяльності в певній предметній сфері. Але, як правило, це лише один аспект впровадження нової технології. Нерідко воно супроводжується зміною не тільки методів і засобів, а й розширенням змісту та зміною парадигми діяльності. Виявити та зафіксувати подібні зміни можна на основі онтологічного підходу до концептуалізації певної предметної сфери до та після впровадження нової технології.

Геоінформаційний моніторинг — це технологія та автоматизована система планування й проведення моніторингу на основі інтегрування даних з різних джерел, моделювання, оцінювання та прогнозування стану об'єктів моніторингу в середовищі геоінформаційних систем із застосуванням баз геопросторових даних і баз знань [2].

Застосування програмних засобів побудови онтологій надає можливість автоматизувати процеси роботи зі знаннями, включаючи етапи придбання, вилучення, структурування, подання та практичного застосування знань фахівцями різної професійної діяльності. Також повинно забезпечуватися можливість порівняння і аналізу змісту фрагментів, повноти, несуперечності і узагальнення знань, розміщених у інформаційному середовищі.

Модель онтології предметної області Ξ може бути надано кортежем [2]:

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

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

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

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

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

$\Lambda: \Theta \times \Psi$ — кінцева множина функцій інтерпретації, які задано на концептах і/або стосунках. 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 .

Онтологія 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, в тому числі, судженнями, ствердженнями, поняттями і відношеннями між ними. Останні є фундаментом, основою для побудови складової частини наукової теорії – інженерії знань в даній предметній сфері.

Стосовно предметної сфери "система геоінформаційного моніторингу" в якості декомпозиції переважним є визначення складу, властивостей, інформаційних ресурсів, переліку моделей та методів аналізу інформації, методів моделювання навколишнього середовища та ін.

Розглянемо більш детально підхід до моделювання знань системи геоінформаційного моніторингу 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 між ними;
- графічне (візуальне) проектування онтографа і складання формалізованого опису онтології предметної області.

Побудова множини Θ вважається найбільш важливим кроком в 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 засновано на результатах етапу попереднього аналізу предметної сфери. Тобто, вимагається встановити для кожного елемента $\theta_i \in \Theta$ 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 (чи вершинами) онтографа є родові поняття, які не мають надкласу, а нижчий рівень є конкретними поняттями. Для створення онтології предметної сфери "система геоінформаційного моніторингу" було застосовано інструментальний засіб моделювання онтології Protégé 4.2 beta [6].

Розробка онтології для складної предметної області, як правило, є трудомістким і ітеративним процесом, до того ж аж до кінця 1990-х років

відзначався недолік опрацьованих методологій для створення онтологій. До теперішнього часу здобули популярність кошти 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 призначена для проектування, пристосування і імпорту / експорту моделей знань в форматах RDF, DAML + OIL, Flogis 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 і нарощувати бази знань. Система WebDeso призначена для створення онтологій предметної області. У табл.1 наведено порівняльну таблицю коштів управління онтологіями.

Таблиця 1

Характеристика засобів управління онтологіями

Засіб управління онтологією	Розробник	Модель подання знань	Метод моделювання понять та відношень	Засіб реалізації
1	2	3	4	5
Ontoligua	Лабораторія систем знань Університету Стенфорда	Логіка першого порядку	Таксономія	Клієнт—HTML інтерфейс
OntoEdit	Компанія Ontoprise GmbH	Фреймова модель знань	Складна таксономія та ієрархія	Java_програма
OilEd	Проект "On_To_Knowledge_Project", Університет Манчестера	Дескрипційна логіка (description logics)	Складна таксономія та ієрархія	Java програма
Proterger	Лабораторія медичної інформатики Університету Стенфорда	Фреймова модель знань	Складна таксономія	Java програма
Web-Deso	"Система інтеграції знань", Інститут інформатики та автоматизації, Спб.	Об'єктно-орієнтовані мережі	Складна таксономія, ієрархія, асоціативні відношення	Клієнт – HTML інтерфейс, Java script

У даній роботі розглянуті основні аспекти створення онтологій верхнього рівня предметної області в середовищі Protégé 4.2 beta, призначеної для загального доступу і користування на прикладі поняття геоінформаційний моніторинг.

На рис.1 показано діалогове вікно програми, в якому подано основні класи та підкласи термінів, які складають поняття СГМ. 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, для необхідно визначити семантичні відношення між класами або підкласами термінів, які визначають поняття СГМ. На рис.3 показано 22 семантичних відношення між класами (підкласами) онтологічного графа.

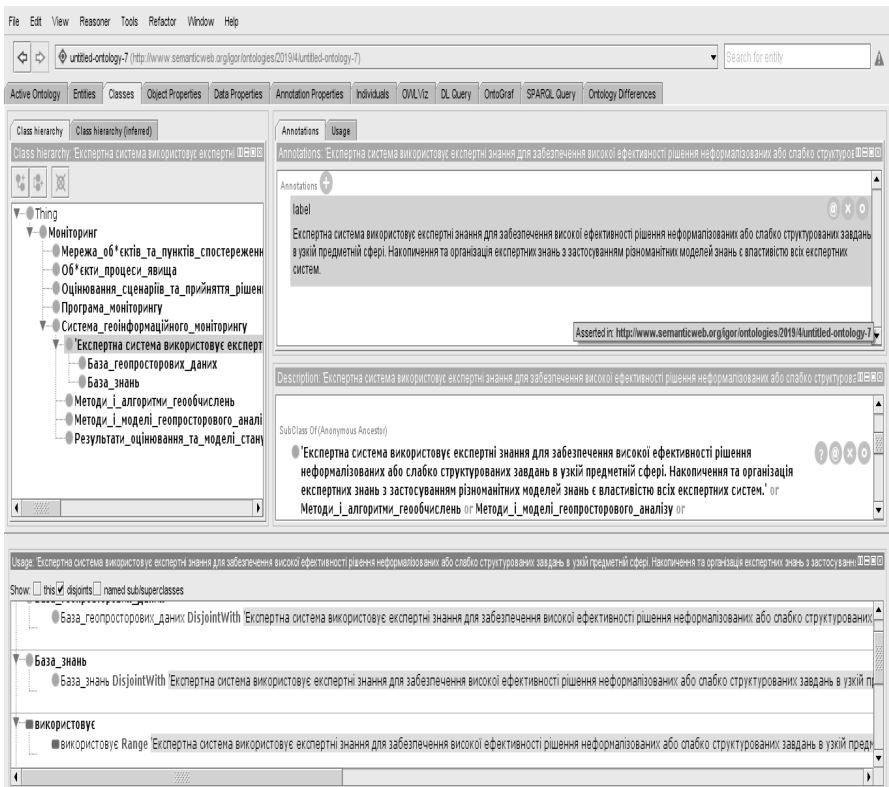


Рис. 1. Діалогове вікно програми Protégé 4.2 beta: визначено основні класи та підкласи системи геоінформаційного моніторингу

Розглянемо докладніше об'єкти СГМ. Під об'єктами СГМ в даній онтології розуміються експертна система, база геопросторових даних і база знань (рис. 4). Наявність таких об'єктів в дозволяє розглядати СГМ як систему з інтелектуальними можливостями.

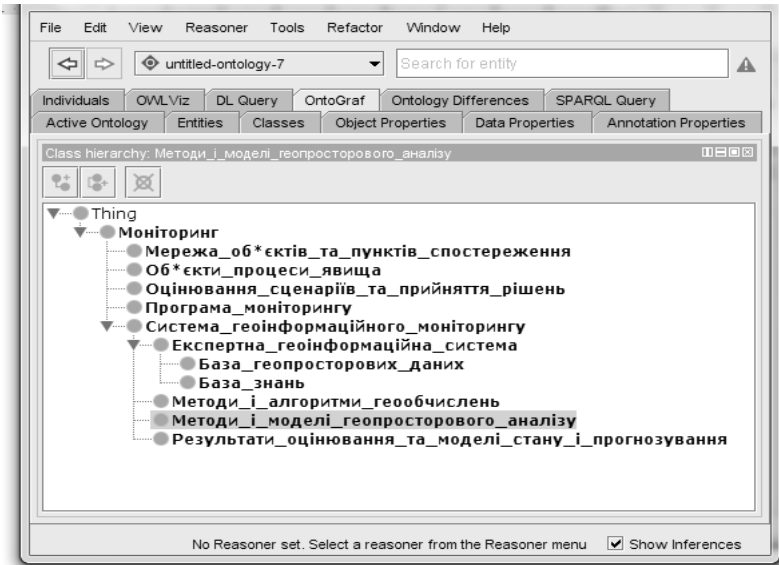


Рис.2. Основні класи та підкласи термінів, які складають поняття системи геоінформаційного моніторингу

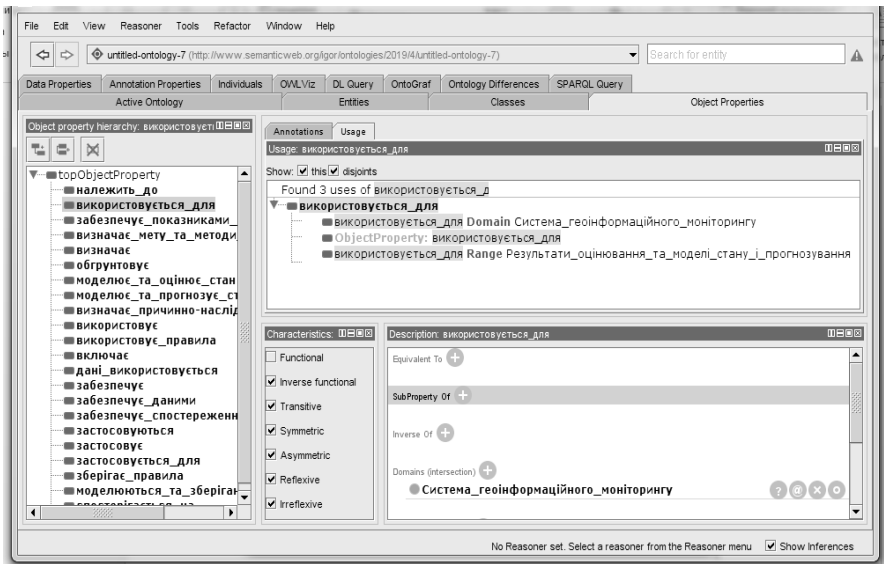


Рис.3. Основні семантичні відношення між класами та підкласами термінів, які складають поняття системи геоінформаційного моніторингу

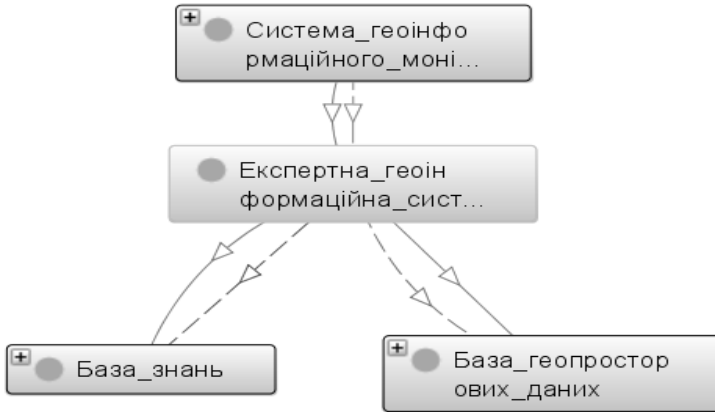


Рис.4. Фрагмент взаємозв'язків між класами в онтології, наявність яких дозволяє розглядати систему геоінформаційного моніторингу, як систему з інтелектуальними можливостями.

Нижче наведено фрагмент запису онтології на мові OWL 2 в Манчестерському синтаксисі, що описує ієрархію, показану на рис.4.

```

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"@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"@en
    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 x в узкій предметній сфері"@en
    label "KnowledgeBase"@en
  SubclassOf: ExpertGeosystem
  disjointwith: GeospatialDatabase, ExpertGeosystem

```

Розроблена базова онтологія, що описує основні компоненти СГМ і відносини між цими компонентами, дає в сукупності цілісне подання про таке поняття як "система геоінформаційного моніторингу".

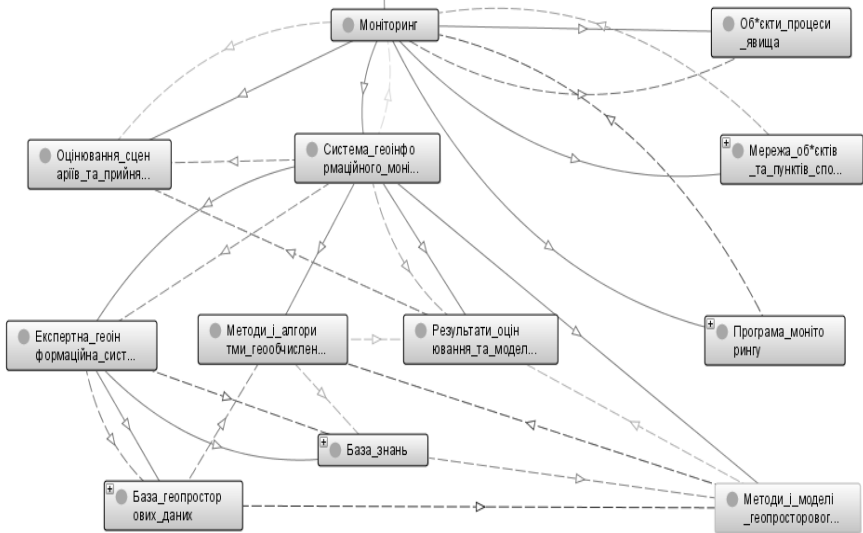


Рис. 5. Узагальнена структура поняття системи геоінформаційного моніторингу, яка побудована з застосуванням засобів моделювання онтологій Protégé 4.2 beta.

Висновки. В роботі запропоновано підхід до створення онтологій з використанням існуючих програмних засобів. Дан короткий огляд програмних засобів для створення онтологій і показано їх обмеження і напрямки розвитку. Проведено концептуальний аналіз терміна "система геоінформаційного моніторингу", який показує класи основних понять і відносин для побудови онтології досліджуваній предметній сфері. Показано, що розглянутий інструментарій дозволяє розглядати терміни і поняття предметної сфери в розвитку, який подається ланцюжком квантів знань розробників онтології предметної сфери в умовах колективної роботи.

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 в статті показано нові можливості і важливі переваги від впровадження розглянутих методів і засобів, для побудови термінів та

визначень, їх взаємовідношень в галузі знань – система геоінформаційного моніторингу.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Лященко А.А. Онтологія та особливості компонентів геоінформаційного моніторингу за технологією баз геопросторових даних / А.А. Лященко, І.М. Патракеєв // 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 / I.M. Patrakeyev // 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.- Str. 22-28

3. Тузовский А.Ф. Системы управления знаниями (методы и технологии) / А.Ф. Тузовский, С.В. Чириков, В.З. Ямпольский // под общ. ред. В.З. Ямпольского. – Томск: Изд-во НТЛ, 2005. – 260 с.

4. Наталья Ф. Ной. Разработка онтологии 101: руководство по созданию Вашей первой онтологии / Н.Ф. Ной, и Д.Л. МакГиннесс. – Стэнфордский Университет, Стэнфорд, Калифорния, 94305.

5. Палагин А.В. Системно-онтологический анализ предметной области / А.В. Палагин, Н.Г. Петренко // УСиМ. – 2009. – № 4. – С. 3–14.

6. Protege 4.2 beta [Электронный ресурс]. – Режим доступа: <http://translate.yandex.ru/translate?srv=yasearch&ur1=http%3A%2F%2Fprotege.stanford.edu%2F&lang=enru&ui=ru> (дата обращения: 15.10.2014).

к.т.н., доцент Патракеєв І.М., Денисюк Б.І.
Київський національний університет будівництва та архітектури

МОДЕЛИРОВАНИЕ БАЗ ЗНАНИЙ В РЕДАКТОРЕ ОНТОЛОГИЙ PROTÉGÉ

Для целостного представления о сложной системе или объекте, новом, еще не установленном понятии "система геоинформационного мониторинга", а также для выявления состава и основных компонентов и отношений между понятиями в современных условиях используют программные средства онтологического моделирования. Авторами рассмотрены теоретические аспекты разработки онтологии исследуемой предметной области, проведено структурирование и формализация знаний относительно понятия "система

геоинформационного мониторинга" и определены характеристики элементов онтологии и их значения.

Сложившаяся классификация классов, подклассов, обнаруженные характеристики, описывающие понятие "система геоинформационного мониторинга" и разработана понятийная структура онтологии предметной области в виде онтологического графа.

На примере разработанной онтологии показано, что геоинформационный мониторинг, как метод познания, позволяет проводить комплексные исследования, на основе которых осуществлять междисциплинарный перенос знаний.

Ключевые слова: система геоинформационного мониторинга, онтология предметной области, онтологический граф, Protégé.

Ph.D., associate Professor Patrakeev I., Denysiuk B.,
Kyiv National University of Construction and Architecture

MODELING KNOWLEDGE BASES IN PROTÉGÉ ONTOLOGY EDITOR

For a holistic view of a complex system or object, a new, not yet established concept of a "geoinformation monitoring system", as well as to identify the composition and main components and relationships between concepts in modern conditions, use ontological modeling software. The authors considered the theoretical aspects of the development of the ontology of the studied subject area, carried out the structuring and formalization of knowledge regarding the concept of "geographic information monitoring system" and determined the characteristics of the ontology elements and their significance.

The existing classification of classes, subclasses, characteristics found, describing the concept of "geoinformation monitoring system" and the conceptual structure of the ontology of the subject area in the form of an ontological graph.

Using the developed ontology as an example, it is shown that geoinformation monitoring, as a method of cognition, allows for comprehensive studies on the basis of which to carry out an interdisciplinary transfer of knowledge.

Keywords: geographic information monitoring system, domain ontology, ontological graph, Protégé.

REFERENCES

1. Lyashenko, A.A., & Patrakeev, I.M. (2015) Ontolohiya ta osoblyvosti komponentiv heoinformatsynoho monitorynhu za tekhnolohiyeyu baz

heoprostorovykh danykh [Ontology and features of geoinformation monitoring components based on geospatial data base technology]. Suchasni dosyahnennya heodezychnoyi nauky ta vyrobnytstva – *Modern achievements of geodesic science and production*, 1 (29), 174-177 [in Ukrainian]

2. Patrakeyev I.M. (2015) Logic-ontological approach to geoinformation monitoring modeling. Zbior raportov 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

3. Tuzovsky, A.F., Chirikov, S.V., Yampolsky, V.Z. (2005). *Knowledge management systems (methods and technologies)* / under total. ed. V.Z. Yampolsky. - Tomsk: Publishing house NTL.

4. 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.

5. Palagin, A.V., & Petrenko, N.G. (2009). Sistemno-ontologicheskiiy analiz predmetnoy oblasti [System-ontological analysis of the subject area] // UsiM, № 4. 3–14.

6. Protege 4.2 beta [Electronic resource]. – Access mode: <http://translate.yandex.ru/translate?srv=yasearch&url=http%3A%2F%2Fprotege.stanford.edu%2F&lang=enru&ui=ru>(appeal date: 10/15/2014).