

УДК 004.942 + 528.88.04

© В.Ю. Вишняков, аспірант

Інститут телекомунікацій і глобального інформаційного простору НАН України, м. Київ
ЦПОСІ та КНП, м. Дунаївці

ТЕМАТИЧНА ОНТОЛОГІЯ ПРОЦЕСІВ ЗАСТОСУВАННЯ ДАНИХ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ДЛЯ МОНІТОРИНГУ ЗЕМЛІ З КОСМОСУ

У статті розглянуто алгоритм побудови онтології відбору даних дистанційного зондування Землі за тематичним призначенням та навіпаки. Сформовано вимоги до розроблення онтології. Наведено фрагменти глосарію.

Ключові слова: ДЗЗ, онтологія, бази даних

Вступ

На сьогоднішній день актуальним завданням є створення автоматизованого визначення можливості застосування даних дистанційного зондування Землі (ДЗЗ) та їх відповідної необхідності для виконання певних тематичних завдань фахівцями різноманітних областей, таких як геологія, лісові ресурси, охорона навколишнього середовища. Згідно з «Класифікатором тематичних задач оцінки природних ресурсів та навколишнього середовища, що вирішуються з використанням матеріалів дистанційного зондування Землі», розробленим «Вост-СибНИИГГиМС» і ІТЦ «СканЭкс» і затвердженим Департаментом науки і інформаційних систем МПР Росії 23.03.2001р., визначено 209 завдань тематичного оброблення, у тому числі 88 завдань з геології, 47 - по поверхневих водах, 25 - по лісових ресурсах, 49 - з охорони навколишнього середовища [1].

На сьогоднішній день дані ДЗЗ з космосу надають понад 49 штучних супутників Землі (ШСЗ) [2], кожен з яких має відповідну знімальну апаратуру з певними технічними властивостями. Процес роботи з кожним ШСЗ може розглядатися як певний інформаційний процес, пов'язаний з пошуком, зберіганням, передачею, обробкою і подальшим використанням інформації [3].

Для формалізації опису цих завдань, їх структурування та подання в машинній формі було використано онтології. Відомо, що онтологія, як правило, описує ієрархію концептів предметної області й істотні властивості кожного концепту за допомогою механізму "атрибут - значення". Зв'язки між концептами можуть бути описані за допомогою додаткових логічних тверджень [4]. У даному випадку онтологія особливо проявляється в уявленні та управлінні знаннями, моделюванні об'єктів і процесів тематичного оброблення, проектуванні баз даних ДЗЗ, інформаційній інтеграції отриманих знань [5].

Розділ 3. Науково-технологічна безпека та інтелектуальні ресурси

Для опису логіки процесу запропоновано використання онтологічного підходу, а саме онтології предметної області. Комп'ютерна онтологія предметної області (ПдО) – це:

- ієрархічна структура скінченної множини понять, що описують задану предметну область;
- структура є онтограф, вершинами якого є поняття, а дугами – семантичні відношення між ними;
- поняття і відношення інтерпретуються відповідно до загальнозначущих функцій інтерпретації, взятих з електронних джерел знань заданої ПдО;
- визначення понять і відношень виконується на основі аксіом і обмежень (правил) їх області дії;
- існує засіб формального опису онтографу;
- функції інтерпретації та аксіоми описані в нотації формальної теорії.

Для побудови онтології автоматизованого визначення можливості застосування даних ДЗЗ та використання їх для виконання тематичних завдань визначимося з поняттями. Під онтологією розуміється кортеж множин:

$$O = \langle X, R, F \rangle \quad (1)$$

де X, R, F - кінцеві множини відповідно:

X - множина понять, які використовуються для опису тематичного завдання $TЗ_i$ та відповідних даних ДЗЗ ($ДД_k$);

$$X = \{TЗ_i, ДД_j\} \quad (2)$$

R – кінцева множина відношень між поняттями X ;

F - множина функції інтерпретації, яка задається на поняттях із множини X та на відношеннях із множини R .

Таким чином, узагальнену процедуру інтеграції можливо представити у наступному вигляді:

$$TЗ_i \times ДД_j = r_k | r_k \neq 0 \quad (3)$$

$$X \times R \rightarrow x_{i,j} \times r_k \rightarrow f \in F \quad (4)$$

У базовій онтології приймемо, що множина F тотожна множині аксіом A , що представляють істинні висловлювання про відповідні поняття X .

$$O = \langle X, R, F, A(D, G) \rangle, \quad (5)$$

де A - скінченна множина аксіом, яка складається з множини визначень D_i і множини обмежень G_i для поняття X_i .

Загальними вимогами до розробки онтології є:

- 1) ясність - онтологія повинна ефективно передавати зміст запроваджених концептів;
- 2) узгодженість - всі поняття і визначення онтології повинні бути логічно не суперечливі;
- 3) розширюваність - онтологія повинна допускати можливість розширювати словник термінів без необхідності ревізії існуючих понять;
- 4) структурованість - простота для розуміння і пошуку понять, що досягається застосуванням системологічного підходу до аналізу предметної області [6].

Процес побудови онтології складається з ряду етапів. Основні етапи побудови онтології включають: визначення цілей створення і використання онтології, переліку питань, на які дозволяє дати відповідь онтологія; розгляд варіантів повторного використання онтології; побудова списку понять, термінів предметної області та їх властивостей; визначення класів об'єктів предметної області та їх ієрархії; визначення властивостей класів - слотів; завдання обмежень і аксіом, побудова діаграм неієрархічних (функціональних) відносин; введення екземплярів об'єктів онтології.

Метою побудови необхідної онтології були формальний опис і класифікація тематичних завдань за даними ДЗЗ та зв'язків між ними для оцінки функціональності застосування відповідних даних ДЗЗ та планування подальшого їх тематичного оброблення. Тобто система мала являти собою опис процедур та компонентів, які дозволяють вирішити завдання користувача. Таким чином, для базової онтології було здійснено обмеження функціональності шляхом представлення інформації про залежність між окремими компонентами по запиту користувача, включаючи тематичні завдання та відповідні дані ДЗЗ з урахуванням усіх множин їх технічних властивостей.

Для заданої предметної області було створено глосарій, який включає терміни тематичних завдань та їх областей застосування. Всього описано 209 тематичних завдань та визначено таксономії тематичних класів з даними 58 штучних супутників ДЗЗ. Фрагмент глосарію наведено в таблиці 1.

Таблиця 1

Термін	Опис терміну
Бази даних ДЗЗ	Файл (або кілька файлів) даних, які зберігаються в ПК або на носіях даних і містять інформацію з штучних супутників Землі, необхідну для виконання тематичного завдання. Умовно діляться на бібліотеки по кожному штучному супутнику.
Бібліотека тематичних завдань, що вирішуються із застосуванням даних ДЗЗ	База даних типових тематичних завдань, яка містить в собі функціональні рішення та інформацію про вхідні дані, необхідні для реалізації цих функцій.
Бібліотеки скануючих пристроїв	Є частиною Баз даних ДЗЗ. Містить інформацію про властивості скануючих пристроїв та рівнів їх оброблення.

З метою спрощення представлення даних було обрано дані ДЗЗ, які на сьогоднішній день приймаються на території України. На основі глосарію було побудовано класифікаційну

Розділ 3. Науково-технологічна безпека та інтелектуальні ресурси

схему з використанням таксономічних відносин "підклас-надкласа" і "клас-екземпляр".
Фрагмент представлено в таблиці 2.

Таблиця 2

Клас	відносіна	Об'єкти								
Супутники ДЗЗ	клас-екземпляр	Terra	Aqua	NOAA	Січ-2					
Terra	клас-екземпляр	MODIS								
Aqua	клас-екземпляр	MODIS								
NOAA	клас-екземпляр	AVHRR								
Січ-2	клас-екземпляр	БСП	СІЧ							
роздільна здатність см	значення	780	3 800	25 000	50 000	100 000	110 000			
780	підклас-надкласа	БСП								
7800	підклас-надкласа	СІЧ								
25000	підклас-надкласа	MODIS								
50000	підклас-надкласа	MODIS								
100000	підклас-надкласа	MODIS								
110000	підклас-надкласа	AVHRR								
спектр	клас-екземпляр	фіолетовий	синій	голубий	зелений	оранжевий	червоний	БІЧ*	СІЧ*	ДІЧ*
фіолетовий	підклас-надкласа	MODIS								
синій	підклас-надкласа	MODIS								
голубий	підклас-надкласа	MODIS								
зелений	підклас-надкласа	MODIS	БСП							
оранжевий	підклас-надкласа	MODIS								
червоний	підклас-надкласа	MODIS	БСП	AVHRR						
БІЧ*	підклас-надкласа	MODIS	AVHRR	БСП						
СІЧ*	підклас-надкласа	MODIS	СІЧ							
ДІЧ*	підклас-надкласа	MODIS	AVHRR							

* БІЧ - ближній інфрачервоний; СІЧ - середній інфрачервоний; ДІЧ - дальній інфрачервоний.

Екологічна безпека та природокористування

Фрагмент представлення класів тематичних завдань та їх зв'язок з об'єктами представлено в таблиці 3.

Таблиця 3

Клас тематичного завдання	відносини	Об'єкти			
592 Виявлення змін стану лісів, що відбуваються в результаті негативних впливів (у т.ч. змін, що відбулися в результаті впливу шкідливих організмів, лісових пожеж, вітровалів, використання лісів і т.п.), а також причин, що призвели до таких змін	роздільна здатність см	5 000	20 000	500	1 500
592 Виявлення змін стану лісів, що відбуваються в результаті негативних впливів (у т.ч. змін, що відбулися в результаті впливу шкідливих організмів, лісових пожеж, вітровалів, використання лісів і т.п.), а також причин, що призвели до таких змін	довжини хвиль	фіолетовий	синій	голубий	зелений
611 Оцінка пожежонебезпеки лісів	роздільна здатність см	100 000	250 000		
611 Оцінка пожежонебезпеки лісів	довжини хвиль	дальній інфрачервоний			
612 Визначення координат ділянок загорання та оперативне спостереження за лісовими пожежами	роздільна здатність см	100 000	250 000		
612 Визначення координат ділянок загорання та оперативне спостереження за лісовими пожежами	довжини хвиль	дальній інфрачервоний			
613 Прогнозування поширення лісових пожеж	роздільна здатність см	100 000	250 000		
613 Прогнозування поширення лісових пожеж	довжини хвиль	фіолетовий	синій	голубий	зелений

Після визначення функціональних відносин у онтологію було додано примірники об'єктів, які необхідні для виконання певного тематичного завдання, наприклад \$ 8, \$11 на рис.3.

Розділ 3. Науково-технологічна безпека та інтелектуальні ресурси

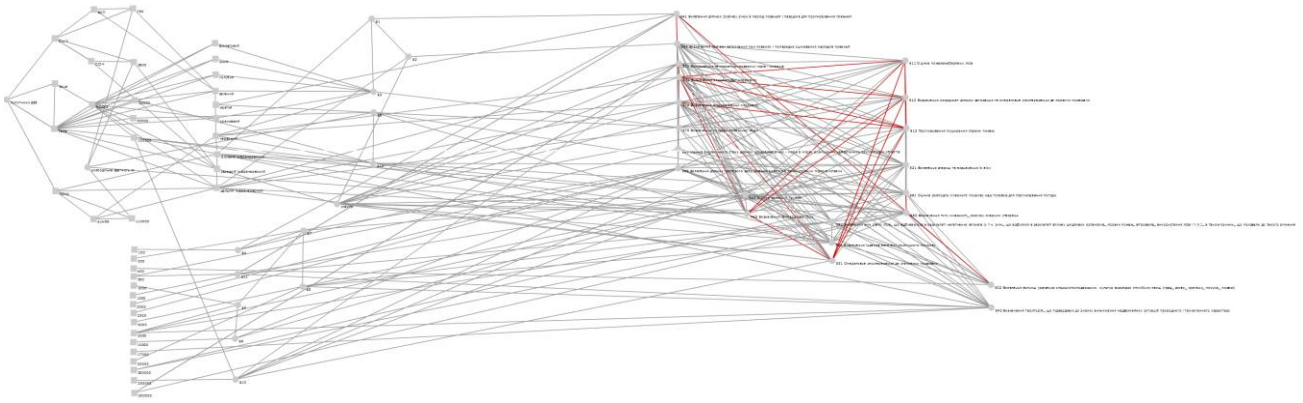


Рис. 1 - Таксономія тематичних зв'язків за даними ДЗЗ

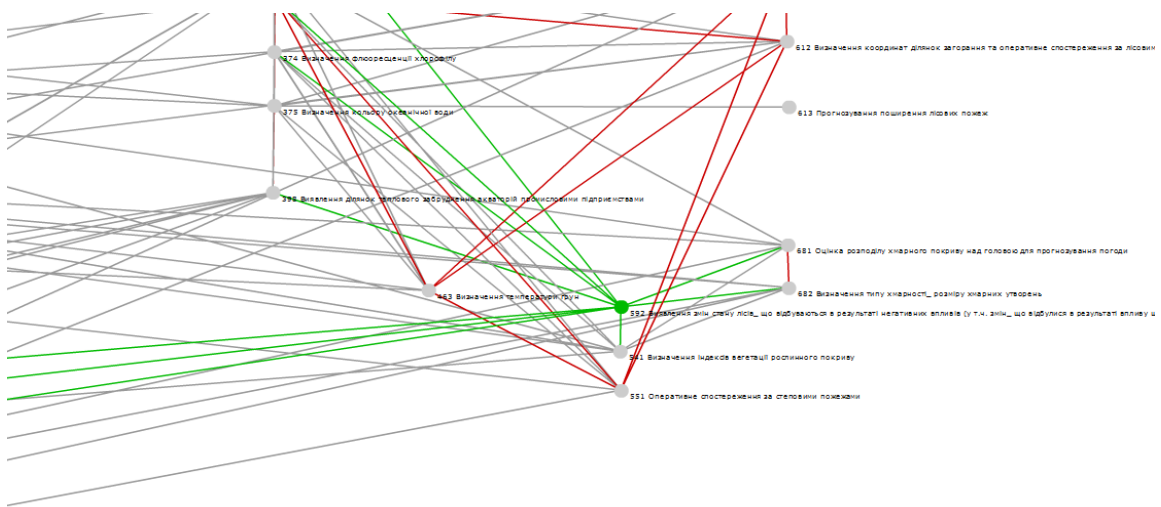


Рис. 2 - Фрагмент таксономії тематичного класу «Виявлення змін стану лісів, що відбуваються в результаті негативних впливів...»

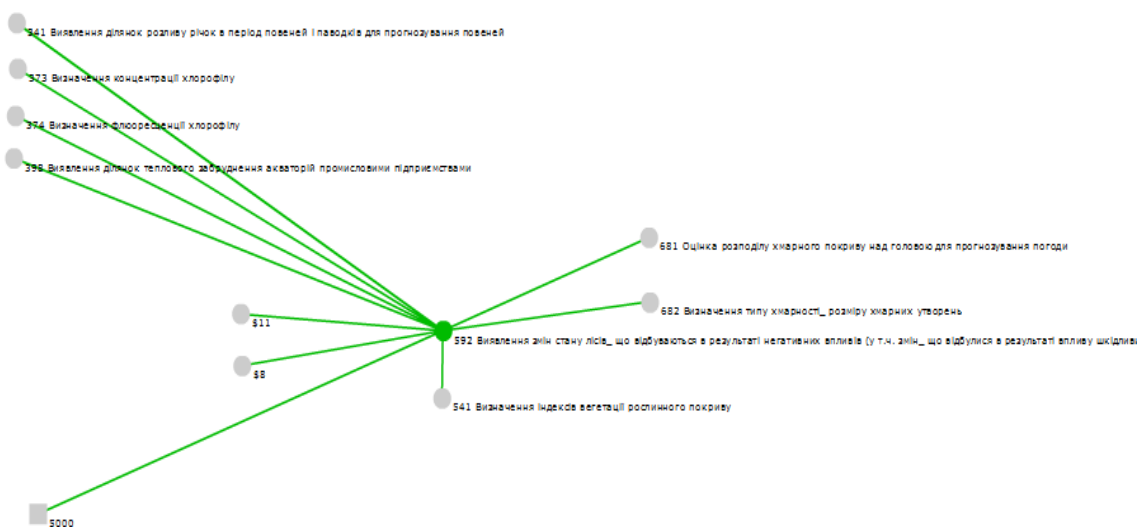


Рис. 3 - Субмножини класу «Виявлення змін стану лісів, що відбуваються в результаті негативних впливів...»

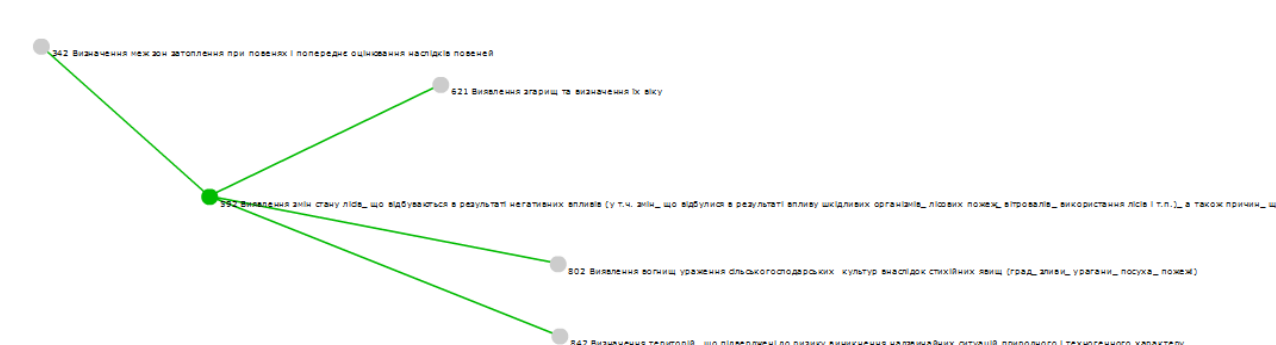


Рис. 4 - Супермножини класу «Виявлення змін стану лісів, що відбуваються в результаті негативних впливів...»

Аналіз сформованих примірників об'єктів визначає необхідність додавання в алгоритм автоматизованого визначення можливості застосування даних дистанційного зондування Землі (ДЗЗ) та їх відповідної необхідності для виконання певних тематичних завдань додаткових даних ДЗЗ, методів їх оброблення, або визначає можливість застосування інших тематичних завдань для їх вирішення. Таким чином, загальний онтологічний вираз приймає більш інформативний вигляд:

$$O = \langle X, R, F, \$, A(D, G) \rangle, \quad (6)$$

Висновки

1. Розроблено базову онтологію автоматизованого визначення можливості застосування даних ДЗЗ та їх відповідної необхідності для виконання певних тематичних завдань фахівцями різноманітних областей, а також відносини, семантично значущі для цієї предметної області.

2. На наступному етапі дана онтологія може бути доповнена екземплярами об'єктів, які необхідно або можливо додатково застосовувати в процесах тематичного оброблення.

3. Онтологія дозволила чітко визначити зв'язки різноманітних примірників об'єктів як ознак різних інформаційних процесів.

4. Послідовна побудова онтології з використанням результатів попередніх онтологій дозволяє збільшувати інформативність кожної наступної, шляхом врахування нових примірників об'єктів.

Список використаної літератури

1. Классификатор тематических задач оценки природных ресурсов и окружающей среды, решаемых с использованием материалов дистанционного зондирования Земли. Редакция 7. – Иркутск: ООО «Байкальский центр», 2008. - 80 с.
2. Спутники ДЗЗ. Режим доступа: http://ru.wikipedia.org/wiki/Категория:Спутники_ДЗЗ.
3. Інформаційні процеси. Режим доступу: http://uk.wikipedia.org/wiki/Інформаційні_процеси.

4. Палагин А.В., Романов В.А., Марков К., Величко В.Ю. Базовая онтология распределенной виртуальной лаборатории проектирования сенсорных систем "International Science and Computing", International Book Series N 15, V.3/2009, P.19-23
5. Гладун В. П. Процессы формирования новых знаний / В. П. Гладун. – София: СД "Педагог 6", 1994. – 397 с.
6. Соловьева Е.А. Естественная классификация: системологические основы. ХТУРЕ Харьков. – 1999.– 222 с.

Стаття надійшла до редакції 23.01.13 українською мовою

© В.Ю. Вишняков

Тематическая онтология процессов применения данных дистанционного зондирования для мониторинга Земли из космоса

В статье рассмотрен алгоритм построения онтологии отбора данных дистанционного зондирования Земли согласно тематическим задачам и наоборот. Сформированы требования по разработке онтологии. Приведены фрагменты глоссария.

© V.U. Vishniakov

Thematic ontology of procedures of using remote sensing data for monitoring the Earth from space

In this article describes an algorithm for constructing ontology selection of remote sensing data according to thematic objectives and vice versa. Formed requirements for the development of the ontology. Presented by glossary fragments.