

Інститут телекомунікацій і глобального інформаційного простору
Національної академії наук України

Кваліфікаційна наукова
праця на правах рукопису

Приходнюк Віталій Валерійович

УДК 004.82 + 004.91 + 005.94 + 528.065/.067

**ТЕХНОЛОГІЧНІ ЗАСОБИ ТРАНСДИСЦИПЛІНАРНОГО
ПРЕДСТАВЛЕННЯ ГЕОПРОСТОРОВОЇ ІНФОРМАЦІЇ**

05.13.06 – Інформаційні технології

05 Технічні науки

Подається на здобуття наукового ступеня кандидата технічних наук

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

Приходнюк В. В.

Науковий керівник
Стрижак Олександр Євгенійович,
доктор технічних наук,
старший науковий співробітник

Київ – 2017

АНОТАЦІЯ

Приходнюк В. В. Технологічні засоби трансдисциплінарного представлення геопросторової інформації. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 «Інформаційні технології». – Інститут телекомунікацій і глобального інформаційного простору Національної академії наук України, Київ, 2017.

Дисертаційну роботу присвячено вирішенню актуальної науково-технічної проблеми підвищення ефективності роботи користувачів з великими масивами слабо і неструктурованої інформації (зокрема, геопросторової), шляхом розробки та впровадження програмних засобів структуризації та трансдисциплінарного представлення такої інформації.

У дисертаційній роботі в першому розділі розглянуто існуючі програмні засоби, що можуть застосовуватись для обробки природномовних текстів. Перелік проаналізованих засобів складається із 35 систем та включає такі програмно-інформаційні рішення обробки знань, як: Palantir Gotham, Convera Excalibur, Oracle Endeca Information Discovery, General Architecture for Text Engineering (GATE), SAS Text Analytics. Розроблено список критеріїв відповідності, що включає в себе 30 критеріїв, об'єднаних у такі класифікаційні групи: загальні параметри, технічні параметри, можливості системи, когнітивні параметри, спеціалізовані параметри.

На основі розроблених критеріїв проведено оцінку і ранжування розглянутих програмних засобів. Показано, що в даний момент не існує програмних засобів, що повністю відповідали б поставленій задачі, однак найбільш оптимальним програмним засобом, що може використовуватись в якості основи для її розв'язку, є система ТОДОС (Трансдисциплінарні діалоги об'єктно-орієнтованих систем).

Також в першому розділі проведено оцінку і ранжування існуючих ГІС (зокрема, програмних продуктів компаній ESRI та Intergraph і ГІС з відкритим кодом GRASS) та дано загальну трактовку трансдисциплінарності.

В другому розділі розроблено метод структуризації природномовних (ПМ) текстів на основі їх рекурсивної редукції. Даний метод передбачає виконання

перетворення структуризації тексту, і являє собою багатокроковий процес. У рамках даного процесу для певного тексту визначається його первинна структура, яка послідовно перетворюється в тезаурус, таксономію і онтологію. Перший крок даного процесу – лексичний аналіз, в результаті якого текст представляється у вигляді послідовності речень, які складаються з лексем (слів або символів, таких, як розділові знаки). Також у ході аналізу формується множина синтаксичних зв'язків.

Отримана в результаті лексичного аналізу первинна структура тексту є вхідною інформацією для власне рекурсивної редукції, що проходить поетапно. В рамках відповідних етапів формуються множина об'єктів, зв'язків і функцій інтерпретації результуючої онтології. Редукція ПМ тексту виконується за допомогою спеціалізованого оператора редукції. Оператор редукції – комбінація кількох операторів, що представляють етапи аналізу ПМ тексту.

Кожна з складових оператора редукції є перетворенням, заданим відповідною базою правил виконання даного перетворення. Правило виконання перетворення має уніфіковану для всіх етапів структуру і складається з функції застосовності, що визначає, чи може правило бути застосоване до певного набору вхідної інформації, і функції перетворення, що задає перетворення вхідної інформації. Виконання функції застосовності полягає у перевірці певної умови. Умова складається з предикатів, серед яких є одномісні (предикати ідентифікації лексем, що задають умову, якій повинна відповідати певна лексема з вхідної множини) і двомісні (предикати ідентифікації зв'язків, що визначають, чи міститься між двома заданими лексемами зв'язок заданого типу).

Умова застосовності означає існування гомеоморфізму між орієнтованим графом, утвореним вхідною послідовністю лексем (а також синтаксичними зв'язками між ними) і певним еталонним орієнтованим графом, що являє собою обраний користувачем підграф первинного представлення певного тексту – як поточного, так і будь-якого іншого (наприклад, тезауруса ПдО).

Також в другому розділі розроблено модель інтерактивного документу, який представлений парою з онтології і натуральної системи виду {дія → результати}. Натуральна система забезпечує інтерактивну взаємодію з наявною в онтології

інформацією. На основі моделі інтерактивного документу розроблені моделі онтологічного ГІС-додатку і трансдисциплінарного представлення інформації. Онтологічний ГІС-додаток являє собою інтерактивний документ, натуральна система якого реалізує функціональне представлення афінного простору. Він призначений для представлення геопросторової інформації, що міститься в онтології, у вигляді ГІС. Трансдисциплінарне представлення інформації реалізується спеціалізованою натуральною системою, що використовує функції індексації і повнотекстового пошуку. Трансдисциплінарне представлення дозволяє динамічно формувати зв'язки між об'єктами, описаними в різних документах, що належать єдиному понятійному інформаційно-аналітичному середовищу, і забезпечувати таким чином оперативний доступ до великих обсягів інформації, важливої для виконання поставленої перед користувачем задачі.

В третьому розділі розроблено інформаційну та функціонально-компонентну модель системи трансдисциплінарного представлення геопросторової інформації. Інформаційну модель системи представлено сукупністю модулів власне системи трансдисциплінарного представлення геопросторової інформації, що інтегрується з певною підмножиною модулів системи ТОДОС. При цьому реалізується відображення інтеграції функцій окремих програмних модулів системи, що перетворює об'єднання множини функцій кожного з її програмних модулів в узагальнену (цільову) функцію – трансдисциплінарне представлення геопросторової інформації.

Функціонально-компонентна модель програмної системи трансдисциплінарної інтеграції геопросторової інформації включає в себе моделі: поведінки системи, структури системи, структури програмних сутностей, компонентів програмної системи а також включає в себе предикат цілісності системи. Дана модель включає в себе множину UML-діаграм варіантів використання, активності та взаємодії.

На основі вищенаведених моделей сформовано архітектуру програмної системи трансдисциплінарного представлення інформації. Вона показує склад, структуру та особливості функціонування даної системи.

В четвертому розділі представлено технологічні засоби трансдисциплінарного представлення геопросторової інформації. Зокрема, вказано особливості їх функціонування та взаємодії з користувачем.

В даному розділі описано приклади роботи розроблених інструментальних засобів, зокрема:

1) Представлення даних про сміттєзвалища, наданих у відкритий доступ Міністерством екології та природних ресурсів України, зі створенням онтологічного ГІС-додатку «Місця видалення відходів» на основі даних сайту «Інтерактивна карта мінприроди».

2) Представлення реєстру дорадників та дорадницьких організацій, а також наявної у відкритому доступі інформації про оренду агротехніки з допомогою онтологічного ГІС-додатку «Агро».

3) Представлення даних про житлово-комунальне господарство м. Біла Церква зі створенням ГІС-додатку «Біла Церква».

4) Обробка документів, присвячених Т. Г. Шевченку (книги обсягом до 500 сторінок) та їх трансдисциплінарне представлення у вигляді інтерактивних документів (зокрема, онтологічного ГІС-додатку).

Також в розділі показано результати обчислювальних експериментів, проведених для визначення ефективності роботи розроблених технологічних засобів. З допомогою обчислювальних експериментів показано, що інструментальні засоби можуть обробляти вхідні природномовні тексти з достатньо високою точністю (до 80%). При цьому інструментальні засоби мають високу швидкодію (близько 14 секунд на документ розміром в кілька сторінок), що дозволяє використовувати їх в процесах оперативного управління.

Розроблені інструментальні засоби створені в рамках дослідних робіт № 0110U002721, № 0112U007539, № 0113U004981, № 0116U000794, № 0117U000005 Інституту телекомунікацій і глобального інформаційного простору та дослідних робіт № 0114U002173, № 0114U002173, № 0115U002523, № 0117U004042 Національного центру «Мала академія наук України». Розроблені засоби впроваджені в наукову та прикладну діяльність: кафедри медичної інформатики

Національної медичної академії післядипломної освіти ім. П. Л. Шупика, Інституту модернізації змісту освіти МОН, ПРАТ «ЕСОММ», ТОВ «Експерт-Центр», Національного музею Тараса Шевченка, Одеського науково-дослідного інституту судових експертиз Мін'юсту України, Інституту водних проблем меліорації НААН України, Комітету з питань будівництва, містобудування і житлово-комунального господарства Верховної ради України.

Ключові слова: структуризація документів, обробка природної мови, трансдисциплінарність, геоінформаційна система, онтологія предметної області, таксономія, інформаційна технологія.

Список опублікованих праць за темою дисертації

Праці, в яких опубліковані основні наукові результати дисертації:

1. Величко В., Приходнюк В., Стрижак А., Марков К., Иванова К., Карастанев С. Построение таксономии документов для формирования иерархических слоев в геоинформационных системах. *Information Content and Processing*, 2015. С. 181–199.
2. Величко В. Ю., Попова М. А., Приходнюк В. В., Стрижак О. Є. ТОДОС – ІТ-платформа формування трансдисциплінарних інформаційних середовищ. *Системи озброєння і військова техніка*, 2017. №. 1(49). С. 10–19.
3. Приходнюк В. В., Стрижак О. Є. Онтологічна ГІС, як засіб впорядкування геопросторової інформації. *Наука і техніка Повітряних Сил Збройних Сил України*, 2017. №. 2(27). С. 167–174.
4. Приходнюк В. В., Стрижак О. Є., Лебідь О. Г. Онтологічне представлення функціональності систем. *Екологічна безпека та природокористування* : Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний університет будівництва і архітектури [та ін.], 2016. №. 3–4 (22). С. 5–23.
5. Приходнюк В. Таксономизация естественно-языковых текстов. *Information Models and Analyses*, 2016. №. 5. С. 270–284.
6. Приходнюк В. В., Стрижак О. Є. Множинні характеристики онтологічних систем. *Математичне моделювання в економіці* : Збірник наукових праць / редкол. : С. О. Довгий (голов. ред.) [та ін.]. К. : НАН України Інститут телекомунікацій і

глобального інформаційного простору, Інститут економіки та прогнозування, Інститут кібернетики ім. В. М. Глушкова, 2017. №. 8. С. 47–61.

7. Довгий С. О., Попова М. А., Приходнюк В. В., Стрижак О. Є., Яцухно В. А. Реєстр архівних документів, пов'язаних з життям, творчістю та вшануванням пам'яті Тараса Григоровича Шевченка: Монографія. К.: Центр розвитку особистості «УНІКУМ», 2017. 250 с. *(при грантовій підтримці Держаного фонду фундаментальних досліджень за конкурсним проектом Ф77/37489)*

8. Андрущенко Т. І., Довгий С. О., Глоба Л. С., Стрижак О. Є., Величко В. Ю., Гальченко С. А., Дем'яненко В. Б., Ільченко М. Ю., Лісовий О. В., Новогрудська Р. Л., Петренко М. Г., Приходнюк В. В., Слюсаренко О. О., Трофимчук О. М., Яценко Г. Ю. Засоби доступу до джерел знань та їх використання в навчальному процесі педагогічного навчального закладу: Методичні рекомендації. 2-ге вид., доповн. К.: Інститут обдарованої дитини, 2012. 192 с.

9. Андрущенко Т. І., Величко В. Ю., Гальченко С. А., Глоба Л. С., Гуляєв К. Д., Клімова Е. Я., Комова О. Б., Лісовий О. В., Попова М. А., Приходнюк В. В., Стрижак О. Є., Стус Д. М. Методики написання наукових робіт на основі онтологічного аналізу текстів: методичний посібник. К.: ТОВ «СІТІПРІНТ», 2013. 124 с.

10. Довгий С. О., Стрижак О. Є., Андрущенко Т. І., Гальченко С. А., Глоба Л. С., Гончар А. В., Копійка О. В., Кудляк В. М., Ляшук К. В., Попова М. А., Приходнюк В. В., Семенюк Г. Ф., Трофимчук О. М., Поліхун Н. І., Постова К. Г. Онтологічний кабінет дослідження життя та творчості Тараса Шевченка в середовищі науково-освітнього порталу KOVZAR.UA: Монографія. К.: Інститут обдарованої дитини, 2016. 175 с.

Праці, що засвідчують апробацію матеріалів дисертації:

11. Горборуков В. В., Приходнюк В. В., Стрижак О. Є., Франчук О. В. Система оцінювання інтелектуальних досягнень учнівської молоді. Онтологічний підхід. *Інтелектуальний аналіз інформації: Збірник праць XVI Міжнародної наукової конференції ім. Т. А.Таран.* К.: Просвіта, 2016. С. 36–42.

12. Величко В. Ю., Приходнюк В. В. Спосіб автоматизованого виділення відношень між термінами з природномовних текстів технічної тематики.

Knowledge – Dialogue – Solution : Збірник праць XX міжнародної конференції. К. : ITNEA, 2014. С. 27–28.

13. Величко В. Ю., Приходнюк В. В. Деякі способи виділення відношень між термінами в природномовному тексті. *Системний аналіз та інформаційні технології* : зб. наук. праць за матеріалами XV конференції (м. Київ, 27 - 31 травня 2013 р.). К. : НТУУ «КПІ», 2013. С. 406.

14. Величко В. Ю., Сирота С. В., Приходнюк В. В. Інструментарій автоматизованого виділення відношень з текстів технічної тематики. *Системний аналіз та інформаційні технології* : зб. наук. праць за матеріалами XVI конференції (м. Київ, 26 - 30 травня 2014 р.). К. : НТУУ «КПІ», 2014. С. 348.

15. Попова М. А., Приходнюк В. В. Формування тематичних ГІС на основі семантиколінгвістичної обробки документів. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Київ, Пуща-Водиця, 3 - 6 жовтня 2016 р.). К. ; Х. : НАН України, Ін-т телекомунікацій і глоб. інформ. простору [та ін.], 2016. С. 130–133.

16. Приходнюк В. В. Автоматизоване виділення інформації з текстів на основі правил, представлених у форматі λ -виразів. *Управління знаннями та конкурентна розвідка* : зб. наук. праць за матеріалами конференції (м. Харків, 19 - 21 квітня 2016 р.). Х. : Харківський національний університет радіоелектроніки, 2016. С. 48–49.

17. Приходнюк В. В. Автоматизоване формування електронних шарів геоінформаційних систем на основі структурованої і неструктурованої інформації. *Геоінформаційні технології в територіальному управлінні* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Одеса, 17 - 18 вересня 2015 р.). Одеса : Одеський регіональний інститут державного управління [та ін.], 2015. С. 73–76.

18. Приходнюк В. В. Онтолого-орієнтований інтерфейс відображення додатків в ГІС-середовищі. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Київ, Пуща-Водиця, 3 - 6 жовтня

2016 р.). К. ; Х : НАН України, Ін-т телекомунікацій і глоб. інформ. простору [та ін.], 2014. С. 127–129.

Prykhodniuk Vitalii. Technological means of transdisciplinary representation of geospatial information. – Manuscript.

Dissertation research for degree of PhDs. by specialty 05.13.06 – Information technology. Institute of Telecommunications and Global Information Space of the National Academy of Sciences of Ukraine, Kyiv, 2017.

Thesis of dissertation focuses on solving an actual scientific and technical problem of improving efficiency of processing large volumes of semi-structured and non-structured information (including geospatial) by creating applications for structuring and transdisciplinary representation of such information.

The first section of the thesis deals with existing tools, that can be used to process natural language texts. The list of analyzed tools consists of 35 systems and includes software for knowledge processing, such as: Palantir Gotham, Convera Excalibur, Oracle Endeca Information Discovery, General Architecture for Text Engineering (GATE), SAS Text Analytics. A list of compliance criteria has been developed, which includes 30 criteria, which are grouped into following classification groups: general parameters, technical parameters, system capabilities, cognitive parameters, specialized parameters.

On the basis of the developed criteria, the evaluation and ranking of the selected software was done. It is shown that at the current moment there are no software that fully complies to the task, but the most suitable tool that can be used as the basis for its solution is the TODOS system (Transdisciplinary dialogues of object-oriented systems).

Also, in the first section, the evaluation and ranking of existing GISs (ESRI and Intergraph software products and GRASS open source GIS in particular) were carried out and a general interpretation of transdisciplinarity was given.

In the second section a method for structuring natural-language (NL) texts based on their recursive reduction is developed. This method involves performing a structured text conversion, and is a multi-step process. Within this process, for a given text, its primary structure is determined, which is sequentially transformed into a thesaurus, taxonomy, and

ontology. The first step of this process is lexical analysis, in which the text is presented as a sequence of sentences consisting of tokens (words or symbols such as punctuation marks). Also, during the analysis, a set of syntactic relations is formed.

Obtained as a result of lexical analysis, the primary structure of the text is the input information for following recursive reduction, which proceeds in stages. Within the corresponding stages, a sets of objects, connections and interpretation functions are formed for the resulting ontology. Reduction of NL text is performed by a specialized reduction operator. The reduction operator is a combination of several operators that represent the stages of the NL text analysis.

Each of the components of the reduction operator is the transformation driven by the corresponding set of the rules. A transformation rule has a unified structure for all stages and consists of two functions: applicability function that determines whether a rule can be applied to a certain set of input information, and transformation function that specifies the transformation of the input information. Applying applicability function is equal to verifying a certain condition. The condition consists of predicates, among which are single (predicates of token identification that specify the condition, which must be met by a certain token) and double (predicates of link identification, which determine whether there is a relation between the two given tokens)

The condition of applicability means the existence of a homeomorphism between a directed graph formed by an input sequence of tokens (as well as syntactic relations between them) and a certain reference graph, which is a sub-graph, selected by the user from the primary representation of a certain text – current or any other (for example, some thesaurus).

Also in the second section a model of an interactive document is presented. Interactive document is represented by a pair: the ontology and the natural system. The natural system provides an interactive access to information from ontology. Based on the model of the interactive document, models of ontological GIS-applications and transdisciplinary information representation have been developed. An ontological GIS application is an interactive document, the natural system of which implements a functional representation of the affine space. It is used for representing of geospatial information, contained in the

ontology, in the form of GIS. The transdisciplinary representation of information is conducted by a specialized natural system that uses indexing and full text search functions. Transdisciplinary representation allows dynamically create relations between objects, contained in various documents belonging to a single conceptual information and analytical environment, and thus provide operational access to large volumes of information important for the solving of the current task.

In the third section an informational and functional-component model of the system of transdisciplinary representation of geospatial information is developed. The information model of the system is represented by a set of modules of the system of transdisciplinary representation of geospatial information, which is integrated with a certain subset of the TODOS modules. In this case, the integration reflection is realized, which converts sets of functions of individual program modules into a target function – transdisciplinary representation of geospatial information.

The functional-component model of the system includes models of: the behavior of the system, the structure of the system, the structure of software entities, components of the software system, and also includes the predicate of the integrity of the system. This model includes a set of UML diagrams: use, activity and interaction.

Based on described models, the architecture of the program system of the transdisciplinary representation of information is created. It shows structure and features of the system.

The fourth section presents tools of transdisciplinary representation of geospatial information. In particular, specifics of their functioning and interaction with users are described.

This section shows examples of work of the developed tools, in particular:

- 1) The data on landfills provided by the Ministry of Ecology and Natural Resources of Ukraine, with the creation of ontological GIS-application "Waste disposal sites" based on the data from the site "Interactive map of natural resources".

- 2) Processing of the register of advisers and advisory organizations, as well as available information on the lease of agricultural technics with the help of an ontological GIS-application "Agro".

3) Presentation of data on housing and communal services in the city of Bila Tserkva with the creation of the GIS-application.

4) Processing of documents related to Taras Shevchenko and their transdisciplinary presentation in the form of interactive documents (ontological GIS-application in particular).

Also, the section shows the results of computational experiments conducted to determine the effectiveness of the developed technological tools. With the help of computational experiments it is shown that tools can handle incoming NL texts with a sufficiently high accuracy (up to 80%). Also, tools have a high performance (about 14 seconds per document with a size of several pages), which allows them to be used in the processes of operational control.

The developed instrumental tools were created within the scope of research works № 0110U002721, № 0112U007539, № 0113U004981, № 0116U000794, № 017U000005 of the Institute of telecommunications and global information space and research works № 0114U002173, № 0114U002173, № 0115U002523, № 0117U004042 of the National center "Minor Academy of Sciences of Ukraine" . The developed tools were used by: the department of medical informatics of the National medical academy of postgraduate Education named after. P. L. Shupyk, Institute for the modernization of the content of education of the Ministry of education and science, private joint-stock company "Expert Center", LLC "ECOMM", National Taras Shevchenko museum, Odessa research institute of forensic expertise of the Ministry of justice of Ukraine, Institute of water problems of melioration National academy of agricultural sciences of Ukraine, Committee on issues of construction, urban development and housing and communal services of the Verkhovna Rada of Ukraine.

Keywords: document structuring, natural language processing, transdisciplinarity, geographic information system, domain ontology, taxonomy, information technology.

List of the candidate's publications

Articles in scientific and professional publications and foreign magazines:

1. Величко В., Приходнюк В., Стрижак А., Марков К., Иванова К., Карастанев С. Построение таксономии документов для формирования иерархических слоев в

геоинформационных системах. *Information Content and Processing*, 2015. С. 181–199.

2. Величко В. Ю., Попова М. А., Приходнюк В. В., Стрижак О. Є. ТОДОС – ІТ-платформа формування трансдисциплінарних інформаційних середовищ. *Системи озброєння і військова техніка*, 2017. №. 1(49). С. 10–19.

3. Приходнюк В. В., Стрижак О. Є. Онтологічна ГІС, як засіб впорядкування геопросторової інформації. *Наука і техніка Повітряних Сил Збройних Сил України*, 2017. №. 2(27). С. 167–174.

4. Приходнюк В. В., Стрижак О. Є., Лебідь О. Г. Онтологічне представлення функціональності систем. *Екологічна безпека та природокористування* : Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний університет будівництва і архітектури [та ін.], 2016. №. 3–4 (22). С. 5–23.

5. Приходнюк В. Таксономизация естественно-языковых текстов. *Information Models and Analyses*, 2016. №. 5. С. 270–284.

6. Приходнюк В. В., Стрижак О. Є. Множинні характеристики онтологічних систем. *Математичне моделювання в економіці* : Збірник наукових праць / редкол. : С. О. Довгий (голов. ред.) [та ін.]. К. : НАН України Інститут телекомунікацій і глобального інформаційного простору, Інститут економіки та прогнозування, Інститут кібернетики ім. В. М. Глушкова, 2017. №. 8. С. 47–61.

7. Довгий С. О., Попова М. А., Приходнюк В. В., Стрижак О. Є., Яцухно В. А. Реєстр архівних документів, пов'язаних з життям, творчістю та вшануванням пам'яті Тараса Григоровича Шевченка: Монографія. К.: Центр розвитку особистості «УНІКУМ», 2017. 250 с. (*with the grant support of the State Fund for Fundamental Research under a competitive project Ф77/37489*)

8. Андрущенко Т. І., Довгий С. О., Глоба Л. С., Стрижак О. Є., Величко В. Ю., Гальченко С. А., Дем'яненко В. Б., Ільченко М. Ю., Лісовий О. В., Новогрудська Р. Л., Петренко М. Г., Приходнюк В. В., Слюсаренко О. О., Трофимчук О. М., Яценко Г. Ю. Засоби доступу до джерел знань та їх використання в навчальному процесі педагогічного навчального закладу: Методичні рекомендації. 2-ге вид., доповн. К. : Інститут обдарованої дитини, 2012. 192 с.

9. Андрущенко Т. І., Величко В. Ю., Гальченко С. А., Глоба Л. С., Гуляєв К. Д., Клімова Е. Я., Комова О. Б., Лісовий О. В., Попова М. А., Приходнюк В. В., Стрижак О. Є., Стус Д. М. Методики написання наукових робіт на основі онтологічного аналізу текстів : методичний посібник. К. : ТОВ «СІТПРІНТ», 2013. 124 с.

10. Довгий С. О., Стрижак О. Є., Андрущенко Т. І., Гальченко С. А., Глоба Л. С., Гончар А. В., Копійка О. В., Кудляк В. М., Ляшук К. В., Попова М. А., Приходнюк В. В., Семенюк Г. Ф., Трофимчук О. М., Поліхун Н. І., Постова К. Г. Онтологічний кабінет дослідження життя та творчості Тараса Шевченка в середовищі науково-освітнього порталу KOBZAR.UA : Монографія. К. : Інститут обдарованої дитини, 2016. 175 с.

Published approbatory works:

11. Горборуков В. В., Приходнюк В. В., Стрижак О. Є., Франчук О. В. Система оцінювання інтелектуальних досягнень учнівської молоді. Онтологічний підхід. *Інтелектуальний аналіз інформації* : Збірник праць XVI Міжнародної наукової конференції ім. Т. А.Таран. К. : Просвіта, 2016. С. 36–42.

12. Величко В. Ю., Приходнюк В. В. Спосіб автоматизованого виділення відношень між термінами з природномовних текстів технічної тематики. *Knowledge – Dialogue – Solution* : Збірник праць XX міжнародної конференції. К. : ITNEA, 2014. С. 27–28.

13. Величко В. Ю., Приходнюк В. В. Деякі способи виділення відношень між термінами в природномовному тексті. *Системний аналіз та інформаційні технології* : зб. наук. праць за матеріалами XV конференції (м. Київ, 27 - 31 травня 2013 р.). К. : НТУУ «КПІ», 2013. С. 406.

14. Величко В. Ю., Сирота С. В., Приходнюк В. В. Інструментарій автоматизованого виділення відношень з текстів технічної тематики. *Системний аналіз та інформаційні технології* : зб. наук. праць за матеріалами XVI конференції (м. Київ, 26 - 30 травня 2014 р.). К. : НТУУ «КПІ», 2014. С. 348.

15. Попова М. А., Приходнюк В. В. Формування тематичних ГІС на основі семантиколінгвістичної обробки документів. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних*

ситуаціях : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Київ, Пуща-Водиця, 3 - 6 жовтня 2016 р.). К. ; Х. : НАН України, Ін-т телекомунікацій і глоб. інформ. простору [та ін.], 2016. С. 130–133.

16. Приходнюк В. В. Автоматизоване виділення інформації з текстів на основі правил, представлених у форматі λ -виразів. *Управління знаннями та конкурентна розвідка* : зб. наук. праць за матеріалами конференції (м. Харків, 19 - 21 квітня 2016 р.). Х. : Харківський національний університет радіоелектроніки, 2016. С. 48–49.

17. Приходнюк В. В. Автоматизоване формування електронних шарів геоінформаційних систем на основі структурованої і неструктурованої інформації. *Геоінформаційні технології в територіальному управлінні* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Одеса, 17 - 18 вересня 2015 р.). Одеса : Одеський регіональний інститут державного управління [та ін.], 2015. С. 73–76.

18. Приходнюк В. В. Онтолого-орієнтований інтерфейс відображення додатків в ГІС-середовищі. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Київ, Пуща-Водиця, 3 - 6 жовтня 2016 р.). К. ; Х. : НАН України, Ін-т телекомунікацій і глоб. інформ. простору [та ін.], 2014. С. 127–129.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ	18
ВСТУП	19
РОЗДІЛ 1. АНАЛІЗ ЗАСОБІВ ОБРОБКИ ПРИРОДНОМОВНИХ ТЕКСТІВ І ПОБУДОВИ ГІС-ДОДАТКІВ.....	26
1.1. Системи обробки природномовних текстів	26
1.1.1. Критерії оцінки систем обробки природномовних текстів.....	27
1.1.2. Огляд існуючих систем обробки природномовних текстів	37
1.2. Існуючі геоінформаційні системи	40
1.2.1. Огляд існуючих геоінформаційних систем	42
1.2.2. Порівняння існуючих геоінформаційних систем.....	53
1.3. Трансдисциплінарність інформаційних ресурсів. Загальні трактовки.	54
Висновки за розділом 1	56
РОЗДІЛ 2. МОДЕЛЬ ТРАНСДИСЦИПЛІНАРНОГО ПРЕДСТАВЛЕННЯ ГЕОПРОСТОРОВОЇ ІНФОРМАЦІЇ.....	58
2.1. Поняття онтологічного ГІС-додатку та його формування на основі природномовних документів	58
2.2. Структуризація тексту з допомогою методу рекурсивної редукції.....	62
2.2.1. Графове представлення первинної структури тексту.....	62
2.2.2. Метод рекурсивної редукції тексту	65
2.2.3. Множинність перетворення рекурсивної редукції	68
2.2.4. Структура оператора редукції та його складових.....	71
2.2.5. Формування правил на основі первинної структури тексту.....	80
2.3. Формування онтологічних ГІС-додатків і трансдисциплінарне представлення геопросторової інформації з їх допомогою	81
2.3.1. Трансдисциплінарне представлення інформації з допомогою інтерактивних документів	84
2.3.2. Трансдисциплінарне представлення геопросторової інформації з допомогою онтологічних ГІС-додатків.....	87
Висновки за розділом 2	94
РОЗДІЛ 3. АРХІТЕКТУРА СИСТЕМИ ТРАНСДИСЦИПЛІНАРНОГО ПРЕДСТАВЛЕННЯ ГЕОПРОСТОРОВОЇ ІНФОРМАЦІЇ.....	96
3.1. Інформаційна модель програмної системи трансдисциплінарного представлення геопросторової інформації.....	96
3.1.1. Інформаційна модель підсистеми структуризації тексту.....	97
3.1.2. Інформаційна модель підмножини модулів ТОДОС, що використовуються в задачі трансдисциплінарного представлення геопросторової інформації	101
3.2. Функціонально-компонентна модель програмної системи трансдисциплінарного представлення геопросторової інформації	104
3.2.1. Модель поведінки системи.....	105
3.2.2. Модель структури системи.....	110
3.2.3. Модель структури програмних сутностей.....	114
3.2.4. Схема компонентів програмної системи	115

3.3. Архітектура системи трансдисциплінарного представлення геопросторової інформації	119
Висновки за розділом 3	122
РОЗДІЛ 4. ІНСТРУМЕНТАЛЬНІ ЗАСОБИ ФОРМУВАННЯ ОНТОЛОГІЧНИХ ГІС-ДОДАТКІВ	124
4.1. Інструментальний засіб «Рекурсивний редуктор»	124
4.2. Інструментальний засіб ТОДОС-ГІС.....	133
4.2.1. Модуль формування інтерактивних документів.....	135
4.2.2. Модуль трансдисциплінарного представлення інформації	144
4.2.3. Модуль формування онтологічних ГІС-додатків	147
4.3. Емпірична перевірка засобів трансдисциплінарного представлення геопросторової інформації	154
Висновки за розділом 4	160
ВИСНОВКИ.....	162
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ	164
Додаток А. Акти впровадження.....	179
Додаток Б.Список опублікованих праць за темою дисертації	188
Додаток В. Відомості про апробацію результатів дисертації.....	191
Додаток Г.Результати аналізу існуючих програмних засобів	192
Додаток Д. Приклади вхідних файлів для формування ГІС-додатків	202
Додаток Е.Інструментальні засоби трансдисциплінарної обробки, що належать до середовища ТОДОС	216

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

CSV	Comma Separated Values
JSON	JavaScript Object Notation
XML	eXtensible Markup Language
ГІС	геоінформаційна система
ІАС	інформаційно-аналітична система
ОС	операційна система
ПдО	предметна область
ПЗ	програмне забезпечення
ПМ	природна мова
НС	натуральна система
ТОДОС Систем	Трансдисциплінарні Онтологічні Діалоги Об'єктно-орієнтованих

ВСТУП

Актуальність дослідження. Швидке зростання об'ємів тематичної інформації і потреба в якісній її обробці викликає необхідність створення ефективних методів і засобів отримання, аналізу та перетворення такої інформації в зручну для опрацювання форму. Аналіз досліджуваної інформації ускладнюється тим, що значна її частина міститься в слабко структурованих або неструктурованих документах (зокрема, природномовних (ПМ) текстах), що викликає необхідність створення відповідних методів і засобів її структуризації. Таким чином, актуальним є завдання ідентифікації описаних у ПМ текстах об'єктів, а також їх характеристик і зв'язків між ними.

Надалі такий процес дозволить сформувати топологічну структуру тексту у вигляді онтології. Її можна використовувати для створення інформаційного середовища корпоративних систем і для надання допомоги фахівцю у виборі та прийнятті рішень за заданою проблематикою. Структуризація інформації сприятиме високоякісному її представленню та надасть низку переваг. Зокрема, її структуризація дозволяє виявляти приховані закономірності, які не можна віднайти іншими методами. Створення та застосування можливостей складних та знання-орієнтованих інформаційних систем розглядалися у дослідженнях Глушкова В. М., Гладуна В. П., Палагіна О. В., Широкова В. А., Хорошевського В. Ф., Поспелова Д. А., Овдій О. М., Гаврилової Т. А., Загорулька Ю. А., Малишевського А. В., Андона П. І., Валькмана Ю. Р., Соловйової К. О., Стрижака О. Є., Яловця А. Л., Gruber T., Guarino N., Noy N., Corcho O., Gomez-Perez A., Gruninger M., Fernandez-Lopez M., Happel H., McGuinness D., Lee J. та інших.

Під час роботи з геопросторовою інформацією наявна інформація найефективніше представляється за допомогою геоінформаційних систем (ГІС), що характеризуються найбільш природним для людини поданням як просторової, так і атрибутивної інформації про розташовані в просторі об'єкти. Тому робота з масивами такої інформації вимагає формування ГІС-додатків у якості кінцевих результатів. Застосуванню можливостей ГІС для представлення просторово-розподіленої інформації в процесі розв'язання задач присвячено роботи Васюхіна

М. І., Довгого С. О., Красовського Г. Я., Лялька В. І., Мокіна В. Б., Палєхи Ю. М., Попова М. О., Стрижака О. Є., Трофимчука О. М, Aronoff S, Smith B., Stock O. та інших.

Також у наш час розвивається трансдисциплінарний підхід щодо інтегрованого використання інформаційних, а особливо геопросторових ресурсів. Такий підхід репрезентує розширення наукового світогляду, що полягає в розгляді того чи іншого явища поза рамками будь-якої наукової дисципліни та забезпечує представлення інформаційного середовища на основі використання множинної часткової впорядкованості гіпервідношень між множинами таксономічних і операціональних властивостей онтологічних моделей предметних областей. Дослідженню трансдисциплінарності приділяють увагу такі науковці, як Князева О. М., Лисак Л. В., Мокій М. С., Палагін О. В., Пригожин І. Р., Стрижак О. Є., Широков В. А., Haken H., Kahneman D., Köhler W., Morin E., Nicolescu B., Piage J., Pink D., Polak F., Rose J., Stengers I., Whole A. та інші. Проте в даний момент ще не створено засобів для єдиного представлення об'єктів та процесів у трансдисциплінарному вигляді.

Таким чином, дисертаційне дослідження має важливе наукове та практичне значення, оскільки спрямоване на вирішення актуальної науково-технічної проблеми, а саме підвищення ефективності формування ГІС на основі масивів слабо структурованої і неструктурованої інформації.

Зв'язок роботи з науковими програмами, темами, планами. Дисертацію виконано в межах наукових тем Інституту телекомунікацій і глобального інформаційного простору: «Розроблення програмно-апаратного комплексу безперервного контролю за станом здоров'я людини, що підключений до інформаційно-аналітичної телемедичної системи надання медичних послуг» 2013–2015 рр. (ДР № 0110U002721), «Науково-технологічні засади формування інформаційно-освітнього простору на основі використання комп'ютерних онтологічних описів мережних баз знань» 2013–2015 рр. (ДР № 0112U007539), «Створення інформаційних та алгоритмічних моделей і засобів трансдисциплінарної інтеграції мереж знань», 2014–2016 рр. (ДР № 0113U004981), «Створення програмно-інформаційних засобів інформаційно-аналітичного

забезпечення мережецентричних ситуаційних центрів», 2016–2020 рр. (ДР № 0116U000794) та «Створення методів та технологічних засад формування інтерактивних баз знань», 2017–2021 рр. (ДР № 017U000005). Також дисертаційне дослідження є частиною наукових тем Національного центру «Мала академія наук України»: «Розробка методів і засобів формування та використання мережних систем знань для створення електронних площадок віртуальних навчально-дослідницьких лабораторій», 2014–2016 рр. (ДР № 0114U002173), «Створення інформаційно-аналітичної системи та методичних засобів моніторингу та оцінювання навчальних досягнень учнівської молоді на основі онтологій рішення задачі вибору», 2014–2016 рр. (ДР № 0114U002173), «Створення інформаційно-аналітичної системи та методичних засобів моніторингу та оцінювання навчальних досягнень учнівської молоді на основі онтологій рішення задачі вибору», 2015–2017 рр. (ДР № 0115U002523), «Науково-технологічні засади створення та використання онтологічних засобів формування мережецентричного середовища віртуальних навчально-дослідницьких лабораторій», 2017–2019 рр. (ДР № 0117U004042).

Роль автора в науково-дослідних роботах та проектах полягає в розробці методів, засобів та відповідної інформаційної технології трансдисциплінарного представлення геопросторової інформації у формі інтерактивних документів та у вигляді онтологічних ГІС-додатків.

Об’єкт дослідження – інформаційні технології, що реалізують структурування природномовних документів та формування на їх основі онтологічних ГІС-додатків.

Предмет дослідження – моделі, методи та засоби трансдисциплінарного представлення геопросторової інформації та формування на їх основі онтологічних ГІС-додатків.

Метою дослідження є створення онтологічної моделі і засобів трансдисциплінарного представлення геопросторової інформації та формування на їх основі онтологічних ГІС-додатків. Реалізація поставленої мети передбачає виконання комплексу завдань:

1. Проаналізувати та узагальнити існуючі методи та засоби використання онтологій при формуванні тематичних ГІС-додатків.

2. На прикладі онтологічних моделей ГІС-додатків, які базуються на множинності класів просторових об'єктів, розробити концептуальні моделі трансдисциплінарного представлення геопросторової інформації, .

3. Розробити методи трансдисциплінарного представлення геопросторової інформації у процесі структуризації природномовних контекстів, які відображають семантичні властивості просторових об'єктів.

4. Дослідити системотехнічні принципи трансдисциплінарного представлення геопросторової інформації та формування і функціонування на її основі онтологічних ГІС-додатків, класи просторових об'єктів яких мають велику кількість міждисциплінарних зв'язків та характеризуються гіпермножинністю.

5. Створити мережеві програмно-інформаційні засоби трансдисциплінарного представлення геопросторової інформації на основі обробки природномовних текстів.

6. Створити інструментальні засоби формування трансдисциплінарних онтологічних додатків у середовищі ГІС.

7. Експериментально підтвердити ефективність роботи розроблених моделей і програмно-інформаційних засобів формування онтологічних ГІС-додатків на основі трансдисциплінарного представлення геопросторової інформації.

Методи та теоретичні засади дослідження. Для розв'язання поставлених задач було використано такі методи: *для розробки методів розв'язання задач* – теорія складних систем, системний аналіз, теорія множин, теорія графів, лямбда-числення; *для формалізації представлення знань* – алгебро-логічний та аксіоматичний методи, теорія категорій; *для програмної реалізації інформаційної технології* – шаблони проектування та об'єктно-орієнтований аналіз.

Наукова новизна дослідження визначається тим, що вперше:

- створено модель трансдисциплінарного представлення документів на основі множинної часткової впорядкованості гіпервідношень між множинами онтологій і натуральних систем;

- розроблено онтологічну модель інтерактивного документа, як технологічної основи формування тематичних шарів у середовищі онтологічної ГІС на основі множинної типізації;
- розроблено метод структуризації природномовних текстів на основі правил, представлених за допомогою безтипових лямбда-виразів;
- розроблено предикативне представлення рекурсивного редуктора, призначеного для формування множин бінарних відношень між термінами, які визначають контексти документів, що обробляються;
- удосконалено теорію трансдисциплінарних онтологічних систем на основі безтипових лямбда-виразів;
- удосконалено методи підтримки взаємодії користувачів з просторово розподіленими інформаційними ресурсами, які застосовують засоби онтологічної уніфікації представлення інформаційних одиниць предметних областей.

Практичне значення дисертаційного дослідження полягає в:

- реалізації програмних засобів формування тематичних бібліотек інтерактивних документів на основі процедур редукції природномовних текстів за допомогою правил, представлених у вигляді безтипових лямбда-виразів;
- реалізації процедури динамічного формування тематичних шарів онтологічних ГІС-додатків на основі онтологій інтерактивних документів;
- удосконаленні технологічних аспектів об'єктно-орієнтованого підходу до створення ГІС-додатків на основі формування тематичних шарів, що будуються з використанням онтологій трансдисциплінарних перетворень геопросторової інформації;
- удосконаленні мережових інструментів ІТ-ТОДОС, що можуть бути використані для трансдисциплінарного представлення геопросторової інформації у вигляді онтологічних ГІС-додатків.

Практичне значення одержаних результатів дослідження підтверджується актами про впровадження в наукову та прикладну діяльність: кафедри медичної інформатики Національної медичної академії післядипломної освіти ім. П. Л.

Шупика, Інституту модернізації змісту освіти МОН, ПРАТ «ЕСОММ», ТОВ «Експерт-Центр», Національного музею Тараса Шевченка, Одеського науково-дослідного інституту судових експертиз Мін'юсту України, Інституту водних проблем меліорації НААН України, Комітету з питань будівництва, містобудування і житлово-комунального господарства Верховної ради України.

Особистий внесок здобувача. Всі результати, які становлять основний зміст дисертації, автор отримав самостійно. Роботи [75–77, 81] виконувались без співавторів. У наукових працях, опублікованих у співавторстві, автору дисертації належать: узагальнена процедура формування таксономії документів [15, 16, 19, 31]; технологія реалізації онтологічних перетворень [79]; узагальнена структура функціонування рекурсивного редуктора [14]; опис аплікативних правил, що задають перетворення ідентифікації [17, 72]; опис предикатів ідентифікації географічної інформації, а також інтерфейсів керування редукцією текстів [80], процедура інтеграції онтологічних додатків у певні типи інформаційних систем [2, 3, 27, 32, 78].

Апробація результатів. Основні положення та результати дисертаційного дослідження було обговорено та викладено в доповідях на: XV і XVI Міжнародних конференціях «Системний аналіз та інформаційні технології» (м. Київ, 2013 р., 2014 р.); конференції «Управление знаниями и конкурентная разведка» (м. Харків, 2013 р., 2014 р., 2015 р., 2016 р., 2017 р.), XV і XVI Міжнародних науково-практичних конференціях «Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях» (м. Київ, Пуща-Водиця, 2016 р., 2017 р.), XX Міжнародній науковій конференції «Knowledge-Dialogue-Solution» (м. Київ, 2014р), XVI Міжнародній науковій конференції «Інтелектуальний аналіз інформації» ім. Т. А. Таран (м. Київ, 2016 р.), II міжнародній науково-практичній конференції «Геоінформаційні технології у територіальному управлінні» (м. Одеса, 2015 р.), XIV міжнародній науково-практичній конференції «Теоретичні та прикладні аспекти побудови програмних систем» (м. Київ, 2017 р.).

Публікації. За матеріалами дисертації надруковано 18 наукових праць, серед яких 6 статей у фахових наукових виданнях (із них: 1 – одноосібна, 2 – в наукових періодичних виданнях іноземних держав, 4 – у виданнях України, що включено до міжнародних наукометричних баз), 4 – монографії, 8 – в інших виданнях, збірниках і матеріалах наукових конференцій.

Структура роботи. Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел (150 позицій) та шести додатків. Загальний обсяг дисертації становить 267 сторінок, із них 148 сторінок основного тексту. Робота містить також 6 таблиць та 57 рисунків.

РОЗДІЛ 1. АНАЛІЗ ЗАСОБІВ ОБРОБКИ ПРИРОДНОМОВНИХ ТЕКСТІВ І ПОБУДОВИ ГІС-ДОДАТКІВ

1.1. Системи обробки природномовних текстів

Інформаційно-аналітичні системи (ІАС) у всіх сферах управління створюються з урахуванням необхідності задоволення потреб суспільства в предметно-орієнтованому розвитку із урахуванням перспектив та завдань створення інформаційного суспільства, де знання є двигуном усього людства [95].

У сучасних умовах використовуються засоби систем інформаційно-аналітичного забезпечення, що дозволять підвищити ефективність управління установою або організацією за рахунок об'єктивізації аналітичної діяльності та допоміжних пошукових та експертних інструментів.

Проблематика аналізу вихідної, або акумульованої інформації пов'язана, перш за все, з необхідністю адекватної обробки масивів інформації. Згідно з цим, під ІАС слід розуміти комплекс апаратних засобів, технологій, програмних рішень, інформаційних ресурсів, методик, які використовуються для забезпечення автоматизації аналітичних та експертних робіт з метою обґрунтування прийняття управлінських рішень та інших можливих застосувань. Згідно концепції інформаційно-аналітичного забезпечення ІАС проектується і функціонує з урахуванням наступних аспектів:

- витяг з багатьох джерел різнорідних даних, представлених у різних форматах і приведення їх до єдиного формату і єдиної структури;
- акумулювання інформації, створення інформаційних масивів даних, застосування технологій індексації і пошуку;
- організація надання користувачам необхідної інформації, яка потрібна для прийняття рішень, реалізації конкретних заходів чи програмних дій у сфері основної діяльності користувачів ІАС;
- інструменти оперативного та інтелектуального аналізу, підготовка планової та регулярної оцінки стану об'єктів управління та дослідження у вигляді документальних носіїв і екранних форм оглядів та звітів;

- подання інформації та результатів аналізу в упорядкованій формі для ефективного сприйняття користувачами.

Якщо інформаційно-аналітична система повинна бути інтегрована з ГІС [19], це накладає на неї додаткові вимоги:

- Підсистеми отримання інформації повинні мати можливість ідентифікувати, аналізувати і проводити верифікацію наявної у вхідних даних географічної інформації, зокрема, координат і адрес;
- Система повинна мати здатність перетворювати адреси в географічні координати (виконувати геокодування);
- Внутрішній формат представлення інформації системи повинен мати можливість зберігати географічну інформацію (можливо, в безтиповому вигляді – наприклад, як текстової строки з координатами, що пізніше будуть оброблені електронною картою);
- Одним з варіантів подання інформації користувачам повинна бути електронна карта, або має існувати механізм перетворення даних, що містяться в системі, в формат, що підтримується існуючими ГІС [19, 92] (наприклад, KML [126] або GeoJSON).

1.1.1. Критерії оцінки систем обробки природномовних текстів

При розгляді існуючих систем, які використовуються для обробки масивів інформації, що містять неструктуровані природномовні тексти [58, 59, 81], з метою побудови на їх основі інтегрованої з ГІС [19, 92] інформаційно-аналітичної системи [95] необхідно враховувати значну кількість параметрів, як основних так і другорядних. Усі ці параметри визначаються різними групами властивостей, які по суті можуть бути представлені у вигляді певних критеріїв [27]. Використання властивостей у якості критеріїв обумовлено тим, що фактично критерії є індексною ознакою певних властивостей, що використовуються функціоналом конкретної системи [78]. Основні критерії можна розбити на наступні категорії:

- Загальні параметри – базові відомості про систему.

- Технічні параметри – відображають технічну реалізацію системи і базові її можливості. У випадку невідповідності технічних параметрів заданим вимогам система мусить бути модифікована або розширена певним чином, що може призвести до значних втрат часу. При цьому в деяких випадках може виникнути необхідність в повній перебудові такої системи, що ставить під сумнів доцільність її використання в рамках поставленої задачі. Наприклад, в Україні не має змісту експлуатація системи, що не підтримує кириличні символи, а модифікація системи для забезпечення їх підтримки в загальному випадку вимагає переписування всіх або більшості програмних модулів.
- Можливості системи – задають спектр завдань, які може вирішувати система. Більш широкі можливості надають системі гнучкості і зменшують час, необхідний на її модифікацію чи інтеграцію у випадку виникнення нових задач. Необхідно концентруватись в першу чергу на можливостях, що використовуються в процесі аналізу ПМ текстів і формування електронних карт, але і загальні можливості керування, аналізу чи відображення даних є достатньо важливими, оскільки дозволяють формувати повноцінну ГІС.
- Когнітивні параметри – визначають можливості системи по роботі з інформацією. Такі можливості в сукупності з електронною картою і формують повноцінну ГІС.
- Спеціалізовані параметри – інші параметри, важливі для роботи з інформацією. Для вибору оптимальної системи (або систем) що можуть використовуватись в рамках реалізації поставленої задачі необхідно [27, 28, 89, 125]:
 - Сформувати критерії якості системи з допомогою числові оцінки значень її параметрів.
 - Оцінити важливість кожного з критеріїв для реалізації поставленої задачі.
 - Сформувати сумарну оцінку відповідності кожної з систем поставленій задачі.
 - Проранжувати системи у відповідності з їх сумарною оцінкою і вибрати систему з найвищою.

Виставлення числових оцінок таким параметрам дозволяє сформувати критерії відповідності системи поставленій задачі.

Пропонується визначивши важливість критеріїв по наступній лінгвістичній шкалі:

- Мала важливість
- Слабка важливість
- Середня важливість
- Значна важливість
- Абсолютна важливість

Значення параметрів будемо оцінювати по шкалі від нуля до десяти, де нуль – найгірше значення, десять – найкраще.

Параметри можна розділити на групи.

Загальні параметри – група основних параметрів системи. Серед загальних параметрів можна виділити клас, до якого належить система, і її доступність.

Клас системи є основним показником її можливостей. Можна виділити наступні класи:

Інформаційно-аналітичні платформи (значення 10) [78, 95] – є найбільш потужними системами, що об'єднують в собі широкий спектр засобів для роботи з даними, часто з клієнт-серверною архітектурою і можливістю обробки великих масивів даних. Основним недоліком таких систем є те, що вони, як правило, є або повністю закритими, або комерційними засобами, часто – корпоративного класу (з надзвичайно високими цінами). Крім того, такі системи як правило створюються за кордоном і можуть не підтримувати українську мову, а іноді – навіть кириличні кодування. Також інформаційно-аналітичні платформи часто є вузькоспеціалізованими, що може унеможливити їх використання для власне побудови ГІС, тому завжди необхідно оцінювати їх можливості і технічні параметри. Однак, завдяки складності і багатофункціональності інформаційно-аналітичних платформ, на їх основі можна легко формувати повноцінну багатофункціональну ГІС.

Однак, в складі інформаційно-аналітичних платформ часто може не бути підсистем аналізу ПМ текстів [14, 122], зокрема, спеціалізованих підсистем

ідентифікації географічної інформації. Тоді такі підсистеми необхідно створювати окремо.

Для їх реалізації можна скористатись спеціалізованими засобами аналізу природної мови, які можуть бути реалізовані як програми (значення 8) або фреймворки (значення 7). В загальному випадку програми є більш зручною альтернативою, але необхідно пам'ятати про необхідність вирішення спеціалізованої задачі обробки власне географічної інформації, що, в загальному випадку, вимагає модифікації цільового програмного засобу. Модифікація готової програми може бути значно ускладнена в порівнянні з фреймворком, або навіть неможлива, тому остаточне рішення завжди необхідно приймати на основі комплексної оцінки.

Як і у випадку інформаційно-аналітичних платформ, суттєвим мінусом для засобу обробки природної мови є відсутність підтримки української (або хоча б російської) мови. Особливою проблемою це є для фреймворків, які, як правило, розробляються закордонними організаціями і містять набори даних виключно для обробки англійської мови. Їх інтеграція в систему, що призначена для обробки української мови, може потребувати значних зусиль і часу.

У випадку відсутності готових засобів обробки природної мови виникає необхідність створення такого засобу. В такій ситуації виникає необхідність в програмі, здатній виконувати хоча морфологічний аналіз вхідного тексту [20, 53, 84] (морфологічному аналізаторі, значення 3). Справа в тому, що забезпечення достатньої якості такого аналізу вимагає величезної кількості лінгвістичної інформації, отримання якої є занадто довгим і трудомістким процесом, щоб виконувати його самостійно. У випадку відсутності навіть морфологічного аналізатора можна спробувати реалізувати базовий лексичний аналіз і стеммінг [20] (відділення основи слова від закінчення), але занадто низька якість роботи такого програмного продукту зробить реальне його використання неможливим.

У випадку відсутності доступної інформаційно-аналітичної платформи з потрібною для ГІС функціональністю необхідно формувати програмний комплекс

для роботи з даними з різноманітних допоміжних програмних елементів (значення 1), серед яких можна виділити:

- Пошукові машини [51] – дозволяють виконувати різноманітні допоміжні операції, такі, як вибірку даних. Деякі пошукові машини мають базові можливості по морфологічному аналізу слів, а також розвинені механізми зчитування ПМТ (зокрема, підпрограми конвертації файлів різного формату в текстовий), які можна застосувати в процесі ідентифікації географічної інформації. Важливо не використовувати комерційні пошукові машини, оскільки вони мають обмежені можливості по їх модифікації та інтеграції з ними;
- Аналітичне програмне забезпечення – може бути застосоване для обробки наявних даних, зокрема, їх валідації [89]. Деяке спеціалізоване програмне забезпечення може мати можливості по інтеграції з різними типами електронних карт, що дозволить спростити реалізацію цього процесу;
- Візуальні редактори, зокрема, редактори онтологій [97, 136, 147] – можуть служити для реалізації візуального користувацького інтерфейсу, зокрема, інтерфейсів для керування аналізом чи валідації/модифікації результатів аналізу. Редактори онтологій є корисними, оскільки онтологія є гнучким і зручним сховищем інформації, придатним для зберігання даних різної структури і призначення .

Інтеграція різних програмних елементів може бути складним завданням через невідповідність структур даних і протоколів обміну даними, особливо у випадку, якщо реалізація системи вимагає застосування великої кількості таких елементів.

Клас системи має високу важливість.

Доступність системи є другим важливим критерієм. Можна виділити:

- Системи з відкритим вихідним кодом.
- Безплатні системи з закритим вихідним кодом.
- Системи, що розповсюджуються на основі комерційної ліцензії.

Системи з відкритим кодом (значення 10) є найбільш зручними для використання, оскільки можуть бути легко модифікованими під конкретні цілі.

Також їх використання значно спрощує питання ліцензування системи, оскільки комерційне використання пропрієтарних систем може бути обмежене ліцензійною угодою.

Безплатні системи з закритим кодом (значення 5) не мають такої переваги, але все одно є достатньо зручною альтернативою комерційним ПЗ, оскільки дозволяють значно розширити функціональність системи, не збільшуючи її вартості.

Комерційне програмне забезпечення (значення 1) є небажаним для використання, оскільки ускладнює впровадження системи і збільшує її сумарну вартість.

Доступність системи має абсолютну важливість.

Технічні параметри:

Сумісність з операційними системами – параметр середньої важливості, що визначає спектр конфігурацій робочого місця оператора, що можуть використовуватись при роботі з системою. Якщо система побудована за клієнт-серверною архітектурою, сумісність клієнта має значно вище значення, ніж сервера, оскільки конфігурація сервера під потрібну ОС є достатньо простою задачею, що виконується адміністратором один раз при інсталяції системи, тоді як кількість робочих місць оператора може бути достатньо високою, а у випадку використання менш поширеної на ПК системи (такої, як Linux) може потребувати додаткового навчання персоналу. Системи за даним параметром можна розділити на:

- Веб-орієнтовані системи (значення 10) – мають веб-інтерфейс, доступ до якого можна отримувати за допомогою звичайного веб-браузера. Такі системи є найбільш гнучкими, оскільки можуть використовуватись незалежно від встановленої у клієнта операційної системи (а часто – навіть з мобільних пристроїв чи планшетів);
- Кросс-платформні (значення 9) системи підтримують кілька операційних систем (як правило, Linux і Windows). Часто список ОС є обмеженим і може не включати деяких (зокрема тих, що встановлюються на мобільні пристрої) ОС. Але в загальному випадку спектр ОС, що підтримуються є достатнім для організації повноцінного робочого місця оператора;

- Призначені для роботи з ОС Windows (значення 5) – такі системи позбавлені гнучкості, але, завдяки поширеності даної ОС вони все ще можуть використовуватись для побудови повноцінного робочого місця оператора.
- Не призначені для роботи з Windows – можуть використовуватись пристроями під керуванням інших ОС (зокрема, мобільними пристроями). Використання таких систем значно ускладнене через фізичні обмеження таких пристроїв, і в загальному випадку такі системи краще виключити з розгляду.

Підтримка кириличних символів є критично важливою для систем, призначених для роботи на території України. Вона може бути забезпечена з допомогою кодувань UTF-8 (значення 10) чи windows-1251 (значення 5). Кодування UTF-8 є значно зручнішим, оскільки дозволяє обробляти широкий спектр символів. При використанні windows-1251 часто може виникати ситуація, коли деякі символи вхідного тексту після закінчення обробки замінюються знаками питання. Особливо серйозною проблемою це може бути при обробці наукових (математичних, фізичних та ін.) текстів, що можуть містити формули, з включають грецькими, латинськими символи чи спеціальними знаками. Використання систем, що не підтримують кириличні символи, є небажаним (значення 0).

Даний критерій має абсолютну важливість.

Програмна і процедурна розширюваність впливають одна з одною і є важливим критерієм для забезпечення інтеграції системи. Даний критерій має різну важливість для різних типів систем: він критично важливий для програмних засобів аналізу ПМТ, тоді як для допоміжних програмних елементів може бути просто бажаним (наприклад, для візуальних редакторів) або навіть не потрібним взагалі (для пошукових машин, що не інтегруються в процес аналізу вхідних текстів). Система може або підтримувати (значення 10), або не підтримувати (значення 0) розширюваність. В загальному критерій має значну важливість.

Мова вхідних даних. Програмний засіб може або підтримувати (значення 10) або не підтримувати (значення 0) обробку певної мови. Для України критично важливою є підтримка української мови (абсолютна важливість). Підтримка російської мови також є достатньо важливою (середня важливість), оскільки

документи російською достатньо поширені на території України, а також, завдяки схожості двох мов, програмні засоби, призначені для аналізу російської мови, достатньо просто модифікувати для аналізу української. Підтримка англійської та інших мов має малу важливість, оскільки через несхожість їх з українською і російською інтеграція можливостей по їх обробці значно ускладнена і вимагає занадто великих зусиль.

Наступною групою параметрів є «Можливості системи», до яких належить:

Можливість аналізу ПМ текстів [13, 16, 23, 57] є критично важливою для ключового компоненту результуючої системи – інформаційно-аналітичної платформи. Якщо доступної платформи з такою функціональністю немає, то необхідно використовувати спеціалізоване програмне забезпечення для забезпечення такої можливості (в такому випадку ключовим елементом системи стає дане ПЗ). Даний критерій має абсолютну важливість.

Можливість інтеграції з ГІС [19, 70, 92] з допомогою експорту в файли підтримуваного ГІС формату або інших засобів, що надаються ними (такими, як ArcGIS API [100]), є зручною функцією, що дозволяє використовувати результати аналізу ПМТ не для роботи з інтегрованою в систему електронною картою, а і з електронними картами, що надаються відповідними ГІС. В деяких випадках можна використати в якості інформаційно-аналітичної платформи саму ГІС, а систему побудувати з допомогою додавання до неї підсистеми аналізу ПМТ. У випадку відсутності такої можливості даний механізм може бути реалізованим окремо в якості додаткового конвертера, але це вимагає вивчення форматів даних, що підтримується цільовою ГІС. Даний параметр має значну важливість.

Можливість інтеграції з системами оптичного розпізнавання документів (OCR, Optical Character Recognition) є зручною функцією, оскільки дозволяє розширити спектр джерел інформації, що можуть використовуватись системою (наприклад, за рахунок розпізнавання не-електронних договорів). Однак OCR-програми дають на вихід звичайний текст, який може бути наданий системі в ручному режимі, що знижує необхідність в автоматичній інтеграції з ними. Також часто розпізнавання документів є частиною переведення їх в електронну форму, що

включає в себе ручне їх сканування і обробку. В такому випадку переваги від автоматичної інтеграції з OCR зменшуються майже до нуля, тому даний параметр має слабку важливість.

Можливість таксономізації документів [81, 93] – як вже було сказано, онтології являються зручним і гнучким сховищем інформації, і складність реалізації системи можна значно зменшити, якщо вихідною інформацією для процесу аналізу тексту є таксономія, що представляє онтологію предметної області, яку описує текст. Це дозволяє використовувати в якості інтерфейсу користувача для перегляду, валідації і модифікації результатів один з наявних в даний момент безплатних редакторів онтологій, а також інші програмні засоби, що працюють з онтологіями. Критерій має середню важливість.

Можливість категоризації може використовуватись для формування власне шарів ГІС [92], тобто групування не зв'язаних між собою об'єктів, виділених з вхідного тексту, в категорії, які можуть використовуватись в якості шарів. Крім того, категоризація є корисною операцією, що може застосовуватись до вхідних даних в рамках виконання аналітичних операцій.

Можливість автоматичної інтеграції [78] – системи з іншими системами показує її гнучкість і значно спрощує побудову результуючої системи. Система з такою можливістю може використовуватись в якості центрального сховища даних і керуючого механізму, який координує роботу інших механізмів. Однак як правило можливості автоматичної інтеграції є обмеженими і для повноцінної роботи різних механізмів необхідне узгодження форматів їх даних, що потребує реалізації додаткових підпрограм. Критерій має середню важливість.

Можливість автоматизації відображення таксономій [13] – система з такою можливістю є зручним механізмом для побудови користувацьких інтерфейсів, призначених для перегляду даних (без можливості редагування), який доповнює електронну карту. Автоматизація процесу забезпечує високу швидкість інтеграції з мінімальними зусиллями. Критерій має значну важливість.

Можливість створення інтерактивних документів [57, 73] доповнює можливість автоматизації відображення онтологій, розширюючи її функцією

редагування даних. Таким чином можна створювати широкий спектр інтерфейсів різного призначення, зокрема, реалізовувати електронні форми. Одним з варіантів використання такої функції є також реалізація інтерфейсу редагування і валідації результатів аналізу тексту, хоча такий інтерфейс може страждати через недостатню функціональність. Критерій має середню важливість.

Можливість знаходження латентної інформації [49] є зручною аналітичною функцією, що дозволяє збільшити повноту обробки тексту завдяки знаходженню неявно вказаних фактів, зокрема, зв'язків між об'єктами, описаними в тексті. Критерій має середню важливість.

Підтримка форматів і протоколів Semantic Web [1, 90, 133, 147] розширює функціональність системи за рахунок надання можливості інтеграції з іншими інструментами, що підтримують цей формат. Критерій має середню важливість.

Можливість опрацювання великих даних (Big Data) [31, 90] важлива в зв'язку з тим, що карти часто формуються великою кількістю елементів, а в деяких випадках такі карти повинні мати можливість врахування можливості частотої зміни стану (зокрема, положення) елементів на ній. Однак в зв'язку з тим, що джерелом даних для карти є ПМ тексти, це накладає відчутні обмеження на швидкість зміни даних, якими вона оперує, тому наявність такої можливості не є критичною. Критерій має середню важливість.

Спеціалізовані параметри:

Глибина бінарних відношень таксономії [90] – даний критерій може бути застосований тільки до систем, що оперують таксономіями. Деякі системи накладають відношення на глибину вкладеності категорій, що в термінах таксономії означає, що є гранична відстань між листовою і кореневою вершинами, після досягнення якої система перестає обробляти таку онтологію коректно. Це обмеження може викликати деякі проблеми при роботі (особливо якщо граничне значення є достатньо малим), і тому систем з обмеженнями на глибину вкладеності категорій варто уникати (хоча такі системи все ще є кращою альтернативою, ніж системи, що не оперують таксономіями). Системи без обмежень на вкладеність

мають оцінку 10, з обмеженнями – 5, тоді як системи, що не оперують таксономіями, мають оцінку 0. Критерій має слабку важливість.

Відношення порядку, що підтримується [42, 90] – критерій, що вказує на особливості упорядкування об'єктів у внутрішньому сховищі системи. Деякі системи використовують лінійний порядок для організації об'єктів (наприклад, пошукові машини можуть розташовувати лексеми, отримані з текстів, в лексикографічному порядку). При роботі з отриманою з текстів інформацією часто можуть виникати об'єкти, що є непорівнюваними, між собою, в той час, як вони є порівнюваними з деякими іншими об'єктами. Для зберігання такої інформації необхідна спеціалізована структура сховища даних а також інтерфейси для її редагування і перегляду. Системи, що підтримують тільки лінійний порядок, мають оцінку 1, тоді як системи, що підтримують частковий порядок – 10. Критерій має слабку важливість.

Мережецентризм [90] означає можливість роботи в рамках мережі, взаємодія між вузлами якої функціонує на основі їх багатозв'язності як об'єктів. Системи, що збудовані на основі концепції мережецентризму, мають оцінку 10. Критерій має слабку важливість.

Емерджентність [90] – критерій показує, що система складається з багатьох підсистем, що при взаємодії між собою починають демонструвати нові властивості, яких не було в окремих компонентів системи. Системи, що проявляють емерджентність, мають оцінку 10. Критерій має слабку важливість.

1.1.2. Огляд існуючих систем обробки природномовних текстів

Для вибору використовувались наступні системи:

- Apache Lucene/Solr, пошукова платформа [99];
- Convera Excalibur, веб-орієнтована інформаційно-аналітична платформа;
- FAST Search, інформаційно-аналітична платформа;
- Galaktika-ZOOM, інформаційно-аналітична система;
- KNIME, система інтелектуального аналізу даних [129];

- NeurOK Semantic Suite, програмні продукти для пошуку та аналізу в текстових масивах інформації;
- Palantir Gotham, інформаційно-аналітична платформа [140];
- RapidMiner, безкоштовне OpenSource середовище для прогнозної аналітики [142];
- RCO (Russian Context Optimizer), комплекс програм для синтаксичного і морфологічного аналізу текстів [33–35];
- SAS Text Analytics, інформаційно-аналітична система ;
- Hummingbird, Enterprise Information Management Systems, (EIMS), корпоративне рішення до інформації та ресурсам підприємства;
- ABBYY Compeno, система додаткових функцій повнотекстового пошуку і морфологічного аналізу текстів [98];
- Affix Grammars Over a Finite Lattice (AGFL), формальна граматики [139];
- Artificial Intelligence Information Retrieval Engine (AIRE), лінгвопроцесор;
- Autonomy, інформаційно-аналітична платформа;
- DictaScope Tokenizer, лексичний аналізатор [37];
- Oracle Endeca Information Discovery, інформаційно-аналітична система бізнес аналітики [111];
- Exalead CloudView, інформаційно-аналітична платформа [113];
- General Architecture for Text Engineering (GATE), архітектура обробки природної мови [101, 106, 107, 133];
- IBM Watson Analytics, суперкомп'ютер, платформа для прогнозної аналітики, обробки і візуалізації даних [116];
- LexicoSyntactic Pattern Language (LSPL), мова для формального опису конструкцій природної мови [7];
- Link Grammar Parser, Link Grammar Parser for Russian, синтаксичний аналізатор [131];
- OntoEdit, система для роботи з онтологіями [147];
- Protégé, універсальний редактор онтологій [97, 109, 115];
- rymorphy2, морфологічний аналізатор [53];

- Russian Morphological Dictionary, програма для синтаксичного і морфологічного аналізу [84];
- Sphinx, система повнотекстового пошуку [146];
- Statistical Package for the Social Sciences (SPSS Statistics), комп'ютерна програма для статистичної обробки даних;
- Stemka, модуль морфологічного аналізу [20];
- ИКОП, інтелектуальні комп'ютерні навчальні програми [61, 62];
- ИСИДА-Т, інтелектуальна система видобування даних та їх аналізу [13, 46];
- Конспект (TextTermin), система лінгвістичного аналізу текстів [23, 24];
- ЛоТА, система аналізу текстових документів [13, 54];
- ТОДОС, онтолого-керована інформаційно-аналітична платформа [14, 19, 88];
- Томіта-парсер, інструмент для отримання структурованих даних з ПМ текстів [117].

Таблиця 1.1. Результати ранжування програмних засобів.

#	Програмний засіб	Значення	#	Програмний засіб	Значення	#	Програмний засіб	Значення
1	ТОДОС	0.819	13	FAST Search	0.341	25	Stemka	0.269
2	Palantir Gotham	0.591	14	NeurOK Semantic Suite	0.34	26	ABBYY Compreno	0.269
3	Convera Excalibur	0.574	15	OntoEdit	0.328	27	SAS Text Analytics	0.257
4	IBM Watson Analytics	0.422	16	Apache Lucene/Solr	0.325	28	pymorphy2	0.244
5	Томіта-парсер	0.395	17	LSPL	0.313	29	Конспект (TextTermin)	0.216
6	GATE	0.378	18	ИСИДА-Т	0.312	30	Exalead CloudView	0.197
7	AGFL	0.37	19	Link Grammar Parser	0.302	31	Oracle Endeca	0.189
8	Protege	0.368	20	ИКОП	0.301	32	Russian Morphological Dictionary	0.121
9	RCO	0.368	21	Sphinx	0.289	33	DictaScope Tokenizer	0.119
10	AIRE	0.347	22	Galaktika-ZOOM	0.288	34	Autonomy	0.116
11	KNIME	0.346	23	Hummingbird	0.288	35	SPSS Statistics	0.094
12	RapidMiner	0.346	24	ЛоТА	0.275			

Отримані результати показують, що для виконання задачі формування ГІС-додатків на основі ПМ текстів є оптимальним рішенням є використання системи ТОДОС: вона задовольняє основним параметрам, таким, як можливість інтеграції з ГІС, підтримка української мови (зокрема, з допомогою кодування UTF-8), відкритість і простота розширення. Інтегрований в ТОДОС аналізатор природної мови Texttermin надає можливості по лексичному і синтаксичному аналізу текстів а також містить допоміжні функції, такі, як визначення узгодженості слів за родом, відмінком та ін.

Таблиця критеріїв для програмних засобів наведена в Додатку Г. Результати ранжування показано в Таблиці 1.1.

Менш оптимальними рішеннями є комерційні інформаційно-аналітичні системи Convera Excalibur і Palantir Gotham, використання яких майже неможливе через закритість і високу ціну. Четверте місце займає суперкомп'ютер компанії IBM – Watson Analytics, який має надзвичайно широкі інформаційно-аналітичні можливості, однак його використання ускладнене відсутністю підтримки української мови. П'яте місце займає безплатний програмний засіб з закритим програмним кодом Tomita-парсер російської компанії Yandex, який показує достатньо хороші результати в обробці української і російської мов, але не має ніяких засобів для модифікації чи розширення своєї функціональності.

1.2. Існуючі геоінформаційні системи

ГІС – це інформаційна система обробки даних, яка має засоби накопичення, збереження, відновлення, пошуку і видачі даних та надає можливості маніпулювання і обробки просторової (просторово-розподіленої, просторово-координованої) інформації [12, 86, 89, 91]. Відмінною рисою ГІС є наявність в їх складі специфічних методів аналізу просторових даних, що в сукупності із засобами введення, збереження, маніпулювання і представлення просторово-координованої інформації і складають основу технології географічних інформаційних систем. ГІС можна розглядати як систему технологічних засобів, програмного забезпечення і

процедур, призначену для збору просторових даних, їх аналізу, моделювання і відображення.

Призначення ГІС – підтримка рішень по управлінню просторово-розподіленими об'єктами. Одна з основних функцій ГІС – створення і використання електронних карт, атласів і іншої картографічної продукції.

З погляду інформатики будь-яку інформаційну систему можна представити як чотирьохкомпонентну модель [91], що включає:

- апаратне забезпечення (весь комплекс технічних засобів – процесори, периферія та ін.);
- програмне забезпечення (методи і засоби, що забезпечують функції зберігання, аналізу і надання даних);
- дані (якісні і кількісні характеристики досліджуваного об'єкту або явища);
- персонал (технічні фахівці, що розробляють і підтримують систему, що створюють і управляють даними, і безпосередні користувачі).

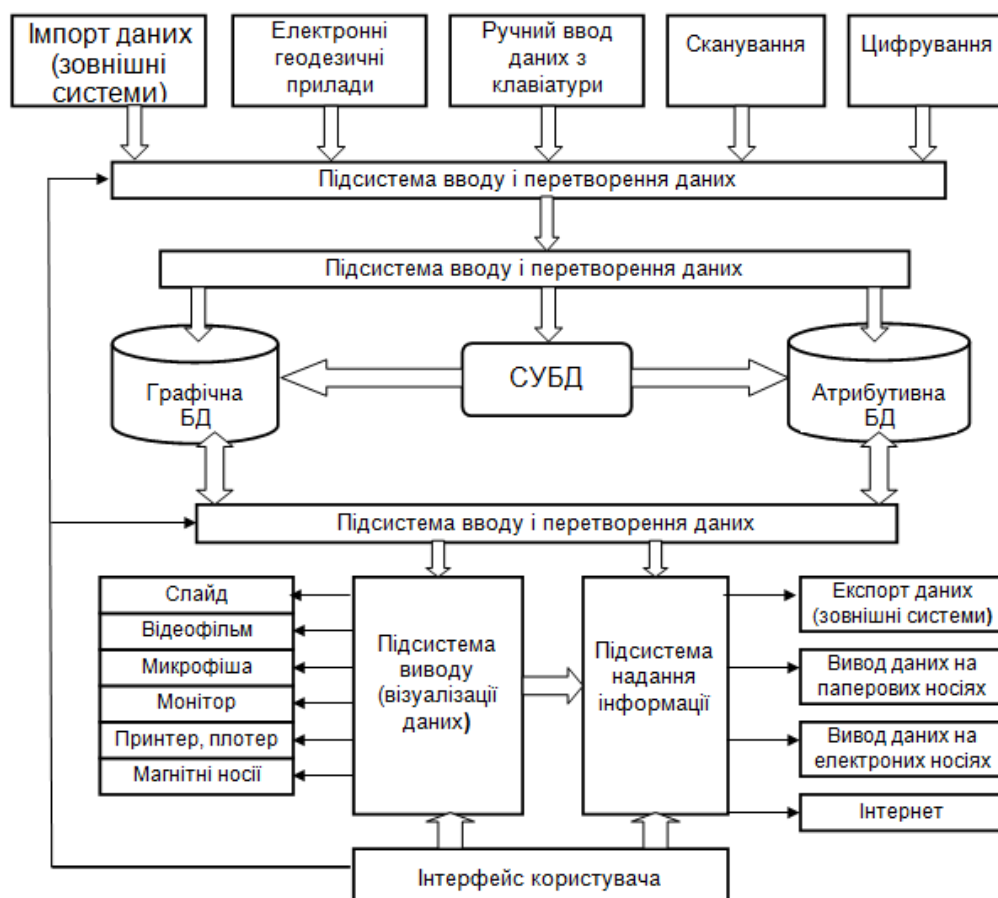


Рис. 1.1. Структура типової ГІС

Більш стійкою є модель ГІС, в основу якої покладений функціональний принцип. Основні компоненти такої системи наведені на Рис. 1.1. Кожен з компонентів виконує певні функції, і відсутність хоч би одного з них свідчить про неповноцінність ГІС.

1.2.1. Огляд існуючих геоінформаційних систем

Програмні ГІС – пакети фірми ESRI [112]. Одним з перших продуктів компанії є система ArcInfo, що призначена для аналізу і обробки великих масивів даних із контролем топологічних взаємозв'язків. Структура ArcInfo складається з двох незалежно встановлюваних програмних пакетів ArcInfo Desktop («настільна» версія) і ArcInfo Workstation (версія для робочих станцій), що належать до професійних ГІС з розвиненими аналітичними можливостями [92].

Пакет ArcInfo Desktop складається з базового модуля і системи розширень. Він містить повну версію додатку ArcToolbox, близько двох сотень операторів якого забезпечують великої кількості векторних і растрових форматів, а також деяких САПР-форматів.

У пакеті є можливість модифікації призначеного для користувача інтерфейсу з використанням засобів настройки. Використовуються внутрішні мови програмування (макриси) – ARC Macro Language (AML) і Visual Basic for Application (VBA), і мови високого рівня через COM і OLE (C++, Delphi, Visual Basic та ін.).

Пакет ArcInfo Workstation є класичною професійною ГІС з повною підтримкою всіх функцій системи ArcInfo. Запускається на персональних комп'ютерах під управлінням Windows і Unix. Пакет має блокову структуру, проте і базовий модуль і модулі розширення відрізняються від настільного варіанта пакета – ArcInfo Desktop.

Базовий модуль, що є повнофункціональною ГІС, містить компоненти:

STARTER KIT – інтерфейсний модуль, який забезпечує управління роботою системи і виклик інших підсистем;

ARCTOOLS – графічний інтерфейс, призначений для користувача;

ARCPlot – підсистему візуалізації і графічного виводу на зовнішні пристрої;

ARCEDIT – підсистему введення і редагування просторових даних;

DATA CONVERSION – підсистему перетворення систем координат і картографічних проекцій.

Додатковими модулями пакета ArcInfo Workstation є:

ArcCOGO – здійснює підтримку координатної геометрії (набір засобів і функцій для роботи з геодезичними даними), її інтеграцію в ArcInfo;

ArcNetwork – забезпечує підтримку моделювання і аналізу просторових (географічних) мереж;

ArcGRID – надає можливості растрового моделювання і перетворює ArcInfo в інтегровану векторно-растрову ГІС;

ArcTIN – забезпечує подання поверхонь у тривимірному просторі у вигляді триангуляційної мережі або нерегулярної матриці точок;

ArcStorm – менеджер просторових даних, що забезпечує цілісність баз даних в розрахованому на багато користувачів режимі роботи;

ArcExpress – підвищує швидкість візуалізації зображень на дисплеї і оперативність роботи з наборами даних на робочих станціях;

ArcScan – забезпечує введення картографічних даних зі сканерів;

ArcPress – забезпечує швидкий високоякісний друк карт і зображень на растрових пристроях виведення.

Призначений для користувача інтерфейс ГІС ArcInfo ґрунтується на командах, що вводяться з клавіатури або за допомогою програм, написаних на мові SML (AML для робочих станцій). Це забезпечує велику гнучкість при рішенні різноманітних задач формування числових карт і проведення географічного аналізу, але потребує досить високої кваліфікації користувача. ARC/INFO часто використовувалася в парі з ArcView GIS, остання в цьому випадку виконує функції візуалізації даних, що готуються за допомогою ARC/INFO.

Пакет ArcView входить до складу сімейства ArcGIS й інтегрує можливості і пакетів ArcView GIS 3.x, і низку інших пакетів компанії ESRI. Пакет належить до ГІС настільного типу і працює під операційними системами Windows.

ArcView має простий призначений для користувача інтерфейс, із можливістю його зміни. Цей продукт відзначається тим, що його функціональність можна нарощувати за рахунок окремих модулів. Також в його склад входить мова програмування Avenue, використовуючи яку можна писати додатки для своїх завдань. ArcView GIS може використовуватися в різних сферах діяльності для візуалізації, запиту і аналізу будь-якої просторової інформації. За допомогою ArcView можна завантажувати будь-які дані, пов'язані з географічним місцеположенням і графічно відображати у вигляді карт, графіків і таблиць. Крім того можливе редагування даних, зміна способів їх відображення, приєднання додаткових даних, створення деяких власних даних, виконання запитів з метою отримання відповідей на специфічні питання, а також аналізувати як статистичну, так і просторову інформацію. ArcView GIS об'єднує векторні, растрові, табличні дані в єдину аналітичну систему.

У 2000 р. ESRI об'єднала пакет для створення даних ArcInfo і пакет для візуалізації ArcView GIS в єдиний пакет із принципово новим інтерфейсом і розширеними можливостями – ArcGIS. Включення функціональності ArcView і ArcInfo в цьому пакеті здійснюється за принципом ліцензування, тобто можна говорити про ArcGIS із функціональністю ArcView, ArcGIS із функціональністю ArcInfo (розширеної) і так далі, залежно від наявності потрібної ліцензії.

Пакет ArcGIS, залежно від типу ліцензії, може бути представлений у вигляді одного з 3 типів продуктів (у порядку зростання функціональності):

ArcView – базові можливості картографування, створення звітів і картографічного аналізу;

ArcEditor – функціональність більша, ніж у ArcView, підтримує можливість побудови баз геопросторових даних;

ArcInfo – максимальна функціональність, найбільш могутній додаток ArcGIS Desktop, що включає ArcInfo Workstation і безліч додаткових інструментів.

До складу ArcGIS входить багато інтегрованих програмних продуктів, призначених як для розробки і експлуатації ГІС різного рівня складності, так і для геоінформаційного забезпечення завдань, пов'язаних з використанням просторової

інформації, включаючи польову зйомку і роботу в комп'ютерних мережах, у тому числі і в Інтернет.

Основними компонентами ArcGIS є:

- настільні ГІС (ArcGIS Desktop), у тому числі ГІС- пакети ArcInfo, ArcEditor і ArcView з набором додаткових модулів;
- серверне програмне ГІС-забезпечення (Server GIS), до складу якого входять пакети ArcIMS, ArcSDE і ArcGIS Server;
- мобільні ГІС (Mobile GIS), представлені пакетом ArcPad, і ГІС-інструменти, або вбудовані інструментальні ГІС (Embedded GIS), представлені пакетом ArcGIS Engine;
- програми-переглядачі (Viewers, Web-viewers), такі, як ArcReader і ArcExplorer.

Настільні ГІС сімейства ArcGIS (ArcGIS Desktop) є різномасштабними програмними ГІС-пакетами, функціональні і аналітичні можливості яких можуть бути істотно збільшені за рахунок великої колекції додаткових модулів, які називаються розширеннями і поставляються окремо.

Для перегляду і друку карт, підготовлених з використанням настільних інструментальних ГІС, використовується пакет ArcReader.

Серверне програмне ГІС-забезпечення (Server GIS) використовується для створення і управління серверними ГІС-додатками, які поширюють функції ГІС і просторово-розподілену інформацію всередині і між організаціями по комп'ютерних мережах, у тому числі і глобальній мережі Інтернет.

Пакет ArcGIS Server – це сервер додатків, що містить загальну бібліотеку програмних ГІС-об'єктів, які використовуються для створення серверних додатків для функціонування як у локальних (рівня підприємства), так і в глобальній (Інтернет) комп'ютерних мережах. Пакет ArcGIS Server є зручною платформою для створення *корпоративних ГІС*, під якими розуміють сукупність додатків з центральним управлінням, які оперують з єдиною базою картографічних даних, що мають розвинену ГІС-функціональність і одночасно підтримують велику кількість користувачів.

Пакет ArcIMS (Arc Internet Map Server) є картографічним Інтернет-сервером, призначеним для публікації карт, даних і метаданих у глобальній мережі з використанням стандартних відкритих Інтернет-протоколів, забезпечує створення ГІС-порталів.

Пакет ArcSDE (Arc Spatial Database Engine) є серверним програмним забезпеченням, що надає доступ до баз географічної інформації через комерційні реляційні системи керування базами даних (СКБД), у тому числі Oracle8i, Oracle9i, MS SQL Server, Informix і DB2.

Web-переглядачі представлені пакетом ArcExplorer, який реалізує функції запитів, вибірки і відображення даних через Інтернет (але може використовуватися і для роботи з локальними наборами даних).

Пакет ArcGIS Engine є бібліотекою взаємопов'язаних компонентів настільних інструментальних ГІС, призначеною для розробників прикладного програмного забезпечення. Цей пакет дозволяє створювати нові картографічні і ГІС-додатки та розширювати можливості вже існуючих додатків (наприклад, Microsoft Excel) за рахунок «вбудовування» в них ГІС-функцій. Додатки, створені за допомогою цього пакета, можуть відповідати діапазону від простих картографічних переглядачів до настільних ГІС, що за функціональністю відповідають пакетам ArcGIS Desktop.

Пакет ArcPad є програмним ГІС-забезпеченням, що запускається на кишенькових персональних комп'ютерах (КПК) під управлінням операційної системи Windows CE. Пакет забезпечує доступ до існуючих цифрових баз просторових даних, а разом з GPS-приймачем – їх поповнення безпосередньо в польових умовах.

ArcGIS Engine пропонує ГІС-компоненти, що вбудовуються і можуть використовуватися при створенні додатків для надання ГІС-логіки будь-якому числу користувачів у межах організації. ArcGIS Engine – це інфраструктура для надання підмножин можливостей ГІС, що відповідають спеціалізованим вимогам. ArcGIS Engine являє собою бібліотеку ГІС-компонентів, які розробники можуть вмонтувати у свої користувацькі додатки. Використовуючи ArcGIS Engine, розробники можуть оснастити функціями ArcGIS такі додатки, як Microsoft Word і Excel, а також

набудовані додатки для надання спеціалізованих ГІС-рішень багатьом користувачам. ArcGIS Engine працює в настільних середовищах Windows, UNIX і Linux, підтримує широкий спектр засобів розробки додатків, таких як Visual Basic 6, Microsoft Visual Studio .NET, багато середовищ розробки на Java, включаючи Eclipse і JBuilder. До складу ArcGIS Engine входить пакет розробника для створення користувацьких додатків. Він встановлюється на комп'ютері з обраними розробником мовою програмування і середовищем розробки. Для вбудовування ГІС-функцій у додатки ArcGIS Engine додає в середовище розробки елементи управління, інструменти, панелі інструментів і бібліотеки об'єктів. Наприклад, програміст може створити додаток, що містить створену в ArcMap карту, деякі інструменти роботи з картою з ArcGIS Engine і інші користувацькі функції.

При установці ліцензії запуску ArcGIS Engine Runtime мається можливість наростити доступну в створюваному додатку функціональність шляхом додавання можливостей чотирьох додаткових модулів. Функції, підтримувані цими опціями, подібні з пропонованими в додаткових модулях до продуктів ArcGIS Desktop, а відповідні ліцензії видаються на кожне робоче місце з Engine Runtime.

Spatial. Модуль Spatial Option додає в середовище ArcGIS Engine Runtime функції розвитку обробки растрів. Додаткові можливості доступні через бібліотеку об'єктів Spatial Analyst.

3D. Модуль 3D Option додає в стандартне середовище ArcGIS Engine Runtime засоби 3D-аналізу і візуалізації. Додаткові можливості включають елементи управління й інструменти Scene і Globe, а також набір бібліотек об'єктів 3D для Scene і Globe.

Geodatabase Update. Модуль Update Option додає можливість записувати й оновлювати базу геопросторових даних з використанням додатків ArcGIS Engine. Він корисний при створенні користувацьких додатків із засобами ГІС-редагування. Ці додаткові можливості доступні через бібліотеку об'єктів корпоративної бази геопросторових даних.

Розробники створюють додатки на основі ArcGIS Engine в обраному інтегрованому середовищі розробки (IDE), такому як:

Microsoft Visual Studio чи Delphi™ для розробників на Windows;

ECLIPSE®, Sun™ ONE Studio™ чи Borland® Jbuilder™ для розробників на Java.

Програмні ГІС – пакети фірми Intergraph. Програмний пакет MGE (Modular GIS Environment) є багатогалузевим модульним середовищем ГІС, розробленим фірмою Intergraph [123]. На сьогодні розроблено більше 60 програмних модулів, що надають засоби створення і супроводу ГІС-проектів, аналізу геопросторової інформації, виведення і відображення електронного картографічного матеріалу. На основі модулів MGE розробляються прикладні ГІС будь-якого ступеню складності, розраховані на одне чи декілька взаємозалежних робочих місць.

Серед усієї сукупності компонентів MGE три модулі є базовими, оскільки містять основні функції створення і супроводу ГІС-проекту. Це – MGE Basic Nucleus (програмне ядро системи), MGE Basic Administrator (базовий адміністратор) і MGE Base Mapper (базовий картограф). Наявність цих продуктів у системі необхідна практично при будь-якій конфігурації робочих місць. Набір додаткових модулів MGE у системі залежить від конкретного розв'язуваного завдання. Взаємодія системи з базою даних забезпечує модуль RIS – Relational Interface System, що входить у стандартне постачання базових продуктів MGE. MGE може працювати з такими відомими промисловими СУБД, як SYBASE, ORACLE, INFORMIX, INGRES, MS SQL, DB2 і т.п.

Програмою основою системи MGE є модуль MGE Basic Nucleus (MGNUC), що забезпечує взаємодію модулів системи між собою і з базою даних. На рівні адміністрування проекту MGNUC відповідає за створення нового проекту і його зв'язок з базою даних, здійснює налаштування проекту на необхідні додатки MGE, створення архівних копій проекту, контроль за виконанням функцій. Цей модуль дозволяє налагодити єдину систему координат для карт проекту.

Для інших програмних модулів MGNUC надає широкий набір функцій для формування запитів до бази даних, виведення результатів запитів, редагування атрибутки, перегляду і відображення семантичних і графічних даних. Функція формування запитів дозволяє здійснювати пошук записів бази даних за значенням

одного або декількох атрибутів об'єкта, а також безпосередньою вказівкою графічного зображення об'єкта на карті. При формуванні запитів можуть використовуватися створені раніше SQL-запити, виділені області і групи об'єктів, а також просторові критерії запитів.

Як базовий графічний пакет у системі MGE використовується графічний редактор корпорації Bentley Systems – MicroStation різних версій. Усі графічні додатки MGE є програмними надбудовами над MicroStation, що являє собою могутній графічний редактор і має розвинуті засоби створення власних програмних додатків. Організаційною одиницею збереження даних є проект.

Основним форматом даних є векторні графічні файли формату DGN. До складу графічних об'єктів входять точкові, лінійні і полігональні об'єкти, а також їх різні комбінації. Використовуються різні графічні блоки, мультилінії, сплайнові лінії і поверхні.

Залежно від розв'язуваних завдань користувачу може бути наданий різний набір базових і прикладних модулів. Для забезпечення основних типів завдань (створення і підтримка карт графічних баз даних, аналізу і подання даних) розроблений стандартний пакет модулів GIS Office. Для роботи цього пакета необхідний комп'ютер з оперативною пам'яттю 32 Мб і 300 Мб дискової пам'яті.

У комплекті з базовими модулями MGE Basic Nucleus, MGE Basic Administrator і MGE Base Mapper у складі GIS Office поставляються:

MGE Analyst – модуль просторового аналізу, що забезпечує побудову і виконання складних запитів, генерацію й аналіз топологічних відношень, побудову буферних зон, агрегування просторових контурів, побудову тематичних карт, відображення топологічно структурованих геопросторових даних і генерацію текстових звітів;

I/RAS C – модуль обробки напівтонових, кольорових і кольорових індексованих аерокосмічних зображень і растрових карт, який дозволяє проводити геометричну корекцію зображень, геоприв'язування, спектральну обробку й аналіз; зшивати/виризувати растри; поліпшувати якість зображень; виконувати монтаж

фотопланів; проводити екранну векторизацію; маніпулювати гібридними растрово-векторними зображеннями і друкувати їх;

MGE Map Finisher – модуль для створення картографічної продукції найвищої якості на основі інформації, що міститься в базі даних ГІС. У ньому автоматизовані процеси створення оформлення, розміщення легенд, виведення на друк;

MGE Grid Generation – інструмент для генерації картографічних сіток і зарамкового оформлення у вигляді векторного зображення;

MRF Clean Tool Kit – комплект із трьох додатків для перевірки й автоматичного коригування топології векторних карт та ін.

Сімейство пакетів GeoMedia. Сімейство продуктів GeoMedia забезпечує високу технологічність усього циклу створення засобів АМ/FM/GIS (Автоматизованого картографування/Управління устаткуванням/ Географічних інформаційних систем), дозволяє в короткий термін створити систему кінцевого користувача і забезпечити її швидке впровадження. Однією із принципових властивостей продуктів сімейства GeoMedia є можливість використовуватися як універсальний ГІС-клієнт.

GeoMedia розрахована на ГІС-аналітиків і ГІС-глядачів, які синтезують нові дані, знання і рішення на основі вже наявних даних. Ці дані зможуть бути використані для комплексного географічного аналізу і прийняття обґрунтованих управлінських рішень. Використовуючи сервери даних, GeoMedia дозволяє підключатися до джерел, розташованих у різних географічних точках, і одночасно аналізувати дані різних типів і форматів.

Поточний список форматів, з якими працює продукт, містить ГІС-формати корпорації Intergraph, формати ArcGIS компанії ESRI, Oracle SDO, Microsoft, Access і деяких інших.

Продукт має різні функціональні можливості і дозволяє: здійснювати запити до баз даних; проводити просторовий аналіз; створювати тематичні карти; створювати і компонувати легенди.

Можливе створення власних додатків на основі GeoMedia використовуючи стандартні засоби програмування у тому числі: Visual Basic і Visual C++; PowerBuilder, Delphi, Excel, FoxPro; Microsoft Access.

Сімейство програмних продуктів фірми Bentley Systems – MicroStation [135] складається з програмних продуктів, об'єднаних єдиним концептуальним підходом і інтегрованих з базовим програмним продуктом – MicroStation. MicroStation є багатофункціональним, високопродуктивним пакетом автоматизованого проектування (CAD, САПР), який забезпечує функції креслення, візуалізації, аналізу, моделювання і керування базами даних. Пакет особливо корисний для груп розробників і користувачів, що працюють у мережному режимі. Застосовуваний у MicroStation багато-користувальницький підхід забезпечує ефективну спільну роботу групи проектувальників незалежно від типу комп'ютерних платформ. Внутрішня архітектура пакета дозволяє використовувати макроси і створювати нові додатки. MicroStation це комплексна система автоматизованого проектування на базі тривимірної графічної системи, що включає:

- обробку дослідницьких даних;
- створення ГІС (MicroStation GeoGraphics);
- інженерне проектування (InRoads);
- проектування схемних рішень (MicroStation Schematics);
- технологічне проектування (PlantSpace DesignSeries);
- архітектурно-будівельне проектування (MicroStation TriForma).

Структура системи – модульна. Внутрішній формат графічних даних – DGN, DWG. Внутрішній формат баз даних – ASCII.

Основним форматом даних є векторний графічний формат DGN. До складу графічних об'єктів входять точкові, лінійні і полігональні об'єкти, а також їх різні комбінації. Використовуються різні графічні блоки, мультилінії, сплайнові лінії і поверхні. У структурі файлу DGN жорстко задані 64 рівні, які ідентифіковані номерами. До кожного рівня прив'язаний набір графічних змінних – стиль, колір, заповнення, товщина.

GRASS GIS (Geographic Resources Analysis Support System) – геоінформаційна система з розвиненими функціями управління просторово-координованими даними. GRASS є модульною, універсальною, багатофункціональною геоінформаційною системою.

В поточній версії GRASS реалізовано 2D/3D топологічне ядро для управління графічною векторною інформацією з функціями мережевого топологічного аналізу даних. Модернізований менеджер відображення графічних даних і додані нові інструменти візуалізації тривимірної графічної інформації. Управління атрибутивними даними виконується за допомогою системи управління базами даних.

Мова програмування системи: ANSI-C відповідно до C-API, підтримується C++ інтерфейс. Система розповсюджується з відкритими початковими кодами. Оскільки архітектура системи побудована за модульним принципом можливо написання власних модулів для вирішення прикладних завдань і включення їх в систему GRASS. На інтерфейс системи накладає певний відбиток початкова орієнтація GRASS на Unix-системи. Охарактеризувати застосоване рішення можна як поєднання командного і віконного інтерфейсів. Причому загальна концепція інтерфейсу вгадується у версіях під різні платформи. Крім стандартного графічного інтерфейсу користувача можливо застосування різних оболонок GUI, наприклад широко відома оболонка QGIS для ядра GRASS.

Існує також JAVA версія системи GRASS – JAVAGRASS, яка забезпечує повну багатоплатформність системи.

Архітектура системи характерна Unix-додатками і множинними модулями, зв'язаних між собою, де кожен модуль виконує певні функції. Система може бути доповнена модулями для вирішення прикладних завдань користувача. У стандартний склад GRASS входить значна кількість модулів для вирішення різних завдань.

Модулі операції над растровими даними дозволяють вирішувати достатньо широкий круг завдань.

Сьогодні GRASS містить понад 350 додатків і інструментів і дозволяє: відображати карти і різні зображення на моніторі і папері; маніпулювати растром, вектором, і даними сайтів; обробляти мультиспектральні дані зображення; створювати, управляти, і запам'ятовувати просторові дані; просторовий аналіз; візуалізацію даних; створення імітаційних моделей. У систему інтегровані модулі,

що забезпечують широкі можливості просторового моделювання і 3D візуалізації. В цілому система GRASS виглядає гідною альтернативою комерційним ГІС, забезпечуючи широку функціональність.

Leaflet [130] – не являється ГІС в повній мірі і представляє собою бібліотеку з відкритим вихідним кодом, призначену для відображення карт на веб-сайтах. Бібліотека підтримує більшість мобільних і стаціонарних платформ. Leaflet є одною з найбільш популярних відкритих бібліотек, підтримує Web Map Service (WMS), GeoJSON, векторні і растрові шари і може, зокрема, працювати серверними програмним забезпеченням ArcGIS.

Також можна виділити наступні менш популярні ГІС:

- AutoDesk CAD Overlay (компанія AutoDesk);
- Geo (компанія Blue Marble Geographies);
- ERDAS (компанія Leica GeoSystems);
- Mapinfo Professional (компанія Mapinfo);
- ELT (компанія Paragon Imaging);
- Geomatica (компанія PCI geomatics).

1.2.2. Порівняння існуючих геоінформаційних систем

Для порівняння існуючих ГІС були вибрані параметри, показані в Таблиця 1.1. Вибрано п'ять основних параметрів, кожен з яких може мати значення від оди ниці до п'яти, і оцінюється за шкалою важливості (абсолютна, значна, середня, слабка, мала важливість).

Основним критерієм є наявність веб-орієнтованих засобів представлення геопросторової інформації. Значну важливість мають вид ліцензії (який показує, чи можливе безплатне використання вибраного програмного засобу) і можливість програмної інтеграції (що дозволяє інтегрувати вибраний програмний засіб в розроблену систему). Всі інші критерії є другорядними.

Критерії порівняння існуючих ГІС:

- Є(5) / нема (1) настільні (Desktop) засоби – мала важливість;
- Є(5) / нема (1) серверні засоби – середня важливість;

- Є(5) / нема (1) Web-орієнтовані засоби – абсолютна важливість;
- Є(5) / нема (1) програмна інтеграція – значна важливість;
- Є(5) / нема (1) підтримка PostGIS – середня важливість;
- Ліцензія: відкрите ПЗ (5), комерційна + некомерційна (4), комерційна (1) – значна важливість.

Результати порівняння ГІС представлені в Таблиця 1.2. Як видно з таблиці, найкращою програмною системою для виконання поставленої задачі є лінійка продуктів ArcGIS компанії ESRI, яка має в собі основні види програмного забезпечення, такі, як сервер, настільна програма і засоби формування веб-додатків. Також достатньо ефективною в рамках поставленої задачі є бібліотека Leaflet, яка, не зважаючи на значно меншу функціональність, виконує основне завдання – формування веб-орієнтованих ГІС-додатків.

Таблиця 1.2. Результати порівняння існуючих ГІС

	Елементи	Значення	Настільні (Desktop) засоби	Серверні засоби	Web-орієнтовані засоби	Програмна інтеграція	Підтримка PostGIS	Ліцензія
1	ArcGIS	0.946	так	так	так	так	так	комерційна і некомерційна
2	Leaflet	0.813	ні	ні	так	так	так	відкрите ПЗ
3	GRASS	0.505	так	ні	ні	ні	так	відкрите ПЗ
4	Mapinfo Professional	0.389	так	ні	ні	ні	так	комерційна
5	AutoDesk CAD Overlay	0.313	так	ні	ні	ні	ні	комерційна
6	ELT	0.313	так	ні	ні	ні	ні	комерційна
7	ERDAS	0.313	так	ні	ні	ні	ні	комерційна
8	Geo	0.313	так	ні	ні	ні	ні	комерційна
9	Microstation GeoGraphics	0.313	так	ні	ні	ні	ні	комерційна
10	PcI geomatics	0.313	так	ні	ні	ні	ні	комерційна

1.3. Трансдисциплінарність інформаційних ресурсів. Загальні трактовки.

Забезпечення трансдисциплінарності досліджень – спосіб розширення наукового світогляду, який полягає в розгляді того чи іншого явища поза рамками якоїсь однієї наукової дисципліни. Розробка механізмів забезпечення трансдисциплінарності забезпечує функціонування загального інформаційно-аналітичного робочого середовища, і є пріоритетним напрямком для задач

інформаційної підтримки наукових та інших досліджень, особливо у великих комплексних проектах або у випадках, коли ресурси розподілені географічно внаслідок особливостей проблем, що вирішуються [38, 39, 43, 66, 90, 94, 128, 137, 149]

Доцільно навести п'ять основних трактовок трансдисциплінарності (по Дж. Клейн) [127, 128]:

Перший підхід пов'язаний з вивченням складної єдності світу і відповідним йому «складним» мисленням, розширенням устояного образу науки.

Другий підхід визначає трансдисциплінарність виходячи з можливостей трансгресивного переходу за межі дисциплінарного знання в соціологічних і культурологічних дослідженнях.

Третій підхід розкриває трансдисциплінарність як нову форму організації наукових досліджень, що передбачає об'єднання міждисциплінарних ресурсів в єдиних методологічних і теоретичних рамках.

Четвертий підхід – так званий «транс-сектор трансдисциплінарності», що фокусується на виявленні механізмів ефективної взаємодії між академічною спільнотою, сферою промисловості і бізнес-сектором. У такому контексті трансдисциплінарність може бути розглянута як ресурс синергії перспектив експертного (академічного) та повсякденного (непрофесійного) типів знання у вирішенні проблем.

П'ятий підхід представлений теоретичними концепціями «другого типу виробництва знання» та «постнормальної науки».

Найбільш загальною є трансдисциплінарність, яка ґрунтується на зусиллях формального взаємозв'язку розумінь окремих дисциплін, при якому забезпечується формування логічних мета-рамок, за допомогою яких знання, що викладені в цих дисциплінах, можуть бути інтегровані на більш високому рівні абстракції, ніж це відбувається в міждисциплінарності. Даний вид трансдисциплінарності часто використовується при роботі різних експертних систем та експертних груп.

Кожен підхід має свої недоліки і переваги, що можуть впливати на вирішення конкретних задач. Однак доцільність використання трансдисциплінарності

безперечна, про що свідчить текст «Всесвітньої Декларації про Вищу освіту для XXI століття: підходи та практичні заходи», яка була прийнята учасниками Міжнародної конференції у жовтні 1998 р. у Парижі, Штаб-квартирі ЮНЕСКО, зокрема, її статті 5 і 6 [150].



Рис. 1.2. Міждисциплінарний характер онтологій в наукових дослідженнях

Для ефективного використання трансдисциплінарності в науці, техніці та освіті її слід представити у вигляді наукового підходу. Тобто трансдисциплінарність має стати самостійним науковим напрямом, що володіє своїм предметом дослідження і концепцією, підходом і мовою, одиницями вимірювання та моделями дійсності, методом аналізу інформації і методикою аналізу ризику прийнятих рішень, що дозволить їй здійснити загальнонаукову класифікацію та систематизацію дисциплінарних знань, зокрема, у формі онтологій [90] (Рис. 1.2.).

Висновки за розділом 1

В першому розділі проведено аналіз та узагальнення стану проблеми трансдисциплінарного представлення геопросторової інформації, що міститься у

природномовних текстах, що дозволили сформулювати взаємопов'язані задачі дисертаційного дослідження:

- розробити концептуальні моделі трансдисциплінарного представлення геопросторової інформації, на прикладі онтологічних моделей ГІС-додатків, які базуються на множинності класів просторових об'єктів.;
- розробити методи трансдисциплінарного представлення геопросторової інформації у процесі структуризації природномовних контекстів, які відображають семантичні властивості просторових об'єктів;
- дослідити системотехнічні принципи трансдисциплінарного представлення геопросторової інформації та формування і функціонування на її основі онтологічних ГІС-додатків, класи просторових об'єктів яких мають велику кількість міждисциплінарних зв'язків та характеризуються гіпермножинністю;
- створити мережеві програмно-інформаційні засоби трансдисциплінарного представлення геопросторової інформації на основі обробки природномовних текстів;
- створити інструментальні засоби формування трансдисциплінарних онтологічних додатків в середовищі ГІС.
- експериментально підтвердити ефективність роботи розроблених моделей і програмно-інформаційних засобів формування онтологічних ГІС на основі трансдисциплінарного представлення геопросторової інформації.

Було показано, що найкраще поставленій задачі відповідає програмна система ТОДОС, і висунуто гіпотезу про необхідність її розширення з допомогою модулів, призначених для автоматизованого аналізу текстів і трансдисциплінарного представлення отриманої в результаті аналізу інформації (зокрема, з допомогою ГІС-додатків), що дозволить підвищити якість та ефективність діяльності експерта за рахунок швидкого доступу та більш ефективного представлення великих масивів інформації.

Основні результати розділу опубліковано в [19, 31, 32].

РОЗДІЛ 2. МОДЕЛЬ ТРАНСДИСЦИПЛІНАРНОГО ПРЕДСТАВЛЕННЯ ГЕОПРОСТОРОВОЇ ІНФОРМАЦІЇ

2.1. Поняття онтологічного ГІС-додатку та його формування на основі природномовних документів

ГІС являє собою найбільш природній і зручний спосіб представлення геопросторової інформації [71, 90, 92]. Однак, побудова ГІС може бути достатньо складним процесом, якщо наявні геопросторові дані представлені в документах, що містять слабо структуровану або навіть неструктуровану інформацію. Обробка таких документів вручну може стати надзвичайно трудомістким процесом, а обробка великих масивів таких документів – майже неможливим [2, 31, 32].

Перед початком роботи з слабо структурованими або неструктурованими документами необхідно провести їх структуризацію. В ході даного процесу дані представляються в зручній для автоматичної обробки формі, що легко може бути зчитана стандартними засобами ГІС а також відображена зручним чином кінцевому користувачу. Це, зокрема, може надавати можливість знаходити у вхідних даних приховану інформацію [4, 11, 21, 25, 74, 85, 88, 120].

Найскладнішим є виконання структуризації ПМ текстів, оскільки даний процес вимагає достатньо повного формального опису підмножини мови, до якого вони належать. Кожен з текстів описує певну предметну область (ПдО) або її частину. При цьому в тексті використовуються належні до ПдО терміни, що формують її термінополе. Структуризація тексту полягає у виділенні з нього цього термінополя, зокрема, ідентифікація концептів (понять) відповідної ПдО, а також їх атрибутів і взаємозв'язків.

Сформоване термінополе може бути представлене з допомогою онтології, що являє собою впорядковану трійку [55, 59, 60, 68, 71, 90]:

$$O = \langle X, R, F \rangle \quad (2.1)$$

де X – множина концептів предметної області,

R – множина відношень між концептами X ,

F – множина функцій інтерпретації X та/або R .

Таким чином структуризацію певного ПМ тексту T^T можна представити як певне перетворення (перетворення структуризації):

$$F_{str} : T^T \rightarrow O \quad (2.2)$$

В реальності, однак, не завжди можна виділити з тексту всю необхідну інформацію, тому можуть мати місце граничні випадки виразу (2.1) при виконанні умов $X = \emptyset$, $R \neq \emptyset$ або $F = \emptyset$. Всі можливі комбінації даних умов дають різні варіанти онтологічних конструкцій, починаючи від простого словника до формальної структури концептуальної бази знань. Зокрема, згідно [90], можна виділити:

- 1) $X = \emptyset, R = \emptyset, F = \emptyset$ – неструктурований текст T^T .
- 2) $X \neq \emptyset, R = \emptyset, F \neq \emptyset$ – глосарій O^1 ;
- 3) $X \neq \emptyset, R \neq \emptyset, F = \emptyset$ – таксономія O^2 ;
- 4) $X \neq \emptyset, (R \neq \emptyset | R = R_t \cup R^+), F = \emptyset$ – тезаурус;
- 5) $X \neq \emptyset, R = \emptyset, card(F) = 1$ – проста онтологія O^* ;
- 6) $X \neq \emptyset, R \neq \emptyset, card(F) > 1$ – активна онтологія O .

Така схема класифікації онтологій за функціональною ознакою узгоджується з описом, наведеним в [4, 8, 13, 22, 28–30, 36, 41, 48, 58, 63–65, 67–69, 119].

Тематична онтологія [90] – це така онтологія, в якій, крім забезпечення виконання умов $R \neq \emptyset$ і $F \neq \emptyset$, до функцій інтерпретації додаються аксіоми, визначення та обмеження за тематикою даної ПдО. При цьому опис всіх компонент представлений деякою формальною мовою, яка може інтерпретуватися деякою процедурою (алгоритмом). Схема формальної моделі тематичної онтології описується наступним чином:

$$O_t = \langle X, R, F, A, D, R_s \rangle \quad (2.3)$$

де X – множина концептів (понять) заданої ПдО;

R – кінцева множина семантично значущих відношень між концептами ПдО.

F – кінцева множина функцій інтерпретації, заданих на концептах і/або відношеннях.

A – скінченна множина аксіом, які використовуються для запису завжди істинних висловлювань (визначень і обмежень) в термінах тематики ПдО;

D – множина додаткових визначень понять в термінах тематики ПдО;

R_s – множина обмежень, що визначають область дії понятійних структур визначеної тематики ПдО.

Тематична онтологія є формальним представленням концептуальних знань про предметну область і може бути представлена певною інформаційною системою. Процес побудови такої інформаційної системи можна представити як композицію певних висловлювань, суджень, тверджень, термінів-понять і відношеннями між ними, а його результат – основою для побудови складової частини наукової теорії – онтологічної бази знань у заданій предметній області, описаній в декларативній формі [21, 25, 44, 87]. Така схема класифікації онтологій за функціональною ознакою узгоджується з описом: «Онтологія або концептуальна модель предметної області складається з ієрархії понять предметної області, зв'язків між ними і законів, які діють в рамках цієї моделі».

Однак структура ПдО, що описується текстом, в загальному випадку обмежується концептами ($X \neq \emptyset$), зв'язками між ними ($R \neq \emptyset$). Тобто на основі текстів можна побудувати глосарій O^1 і таксономію O^2 . Побудована таксономія матиме вигляд (2.4).

$$O^2 = \langle X, R \rangle \quad (2.4)$$

де X – множина концептів заданої ПдО;

R – кінцева множина семантично значущих відношень між концептами ПдО.

Однак згідно [90], всі непусті добутки підмножин X і R утворюють певну множину дій $F_t \subset F$. Тому множину функцій інтерпретації завжди можна сформулювати при виконанні умов $X \neq \emptyset$ і $R \neq \emptyset$, що виконується для виразу (2.4). Таким чином, будь-якому виразу виду (2.4) завжди відповідатиме певний вираз виду (2.1).

Фактично перетворення структуризації тексту (2.2) є багатоетапним процесом, кожен етап якого потребує використання спеціалізованих моделей і процедур.

Загальну схему виконання такого перетворення можна представити виразом виду (2.5).

$$T^T \rightarrow T_{sn} \rightarrow O^1 \rightarrow O^2 \rightarrow O \quad (2.5)$$

де T^T – ПМ текст;

T_{sn} – первинна структура тексту;

O^1 – глосарій;

O^2 – таксономія (2.4);

O – онтологія (2.1).

При цьому перший етап, лексичний аналіз $T^T \rightarrow T_{sn}$, потребує використання значної кількості лінгвістичної інформації, і тому виконується окремою системою – лексичним аналізатором [13, 14, 16–19, 23, 47, 58].

Отримання структуризованого представлення тексту у вигляді онтології (2.1) дозволяє виконувати його автоматичну обробку наявної в ньому інформації з допомогою тих чи інших засобів, однак для роботи експертів таке представлення не підходить. Оскільки діяльність експертів у корпоративному інформаційному середовищі може бути представлена у вигляді системи {дія $\rightarrow$ результати}, то здійснювати представлення наявної в структуризованому документі інформації експерту доцільно з допомогою системи аналогічної структури. Систему такого типу можна визначити як натуральну – SN [52, 90].

Натуральна система (НС), використана для відображення певного документу, забезпечує інтерактивну взаємодію з його змістом. Такий документ можна визначити як інтерактивний і представити з допомогою пари (2.6).

$$\langle O, SN \rangle \quad (2.6)$$

де O – онтологія, що представляє собою структуризоване представлення деякого документу;

SN – НС, побудована на основі O .

Частковим випадком інтерактивного документу є *онтологічний ГИС-додаток*, який являє собою інтерактивний документ, НС якого побудована над афінним простором [80]. Афінний простір виступає математичною основою для

відображення геопросторової інформації, і тому така НС отримує можливість природним чином представляти експерту виділену з тексту в процесі структуризації інформацію такого типу. Сформований на основі такої НС ГІС-додаток може бути представлений з допомогою виразу (2.7) [57, 70, 71, 73, 80].

$$\langle X_g, R_g, A^{op}, S, O \rangle \quad (2.7)$$

де X_g – множина географічних сутностей, над якими виконуються аналітичні операції для вирішення задачі;

R_g – множина відношень між географічними об'єктами, які визначають тип виконуваних операцій;

A^{op} – множина аналітичних операцій над географічними об'єктами, виконуваних в процесі вирішення задачі;

S – множина станів задачі, які візуалізуються на карті в процесі її вирішення;

O – онтологічний опис географічних об'єктів, процесів і задач заданої предметної області.

2.2. Структуризація тексту з допомогою методу рекурсивної редукції

Процес структуризації тексту (2.5) можна умовно розбити на два основних етапи: лексичний аналіз $T^T \rightarrow T_{sn}$, що формує первинну структуру тексту, і формування онтології $T_{sn} \rightarrow O$, що дозволяє виділити з первинної структури необхідну інформацію і представити її у вигляді онтології. Для виконання другого етапу запропоновано метод рекурсивної редукції, що передбачає послідовне перетворення первинної структури з допомогою певного набору динамічно заданих користувачем правил.

2.2.1. Графове представлення первинної структури тексту

Первинна структура тексту T_{sn} є результатом роботи лексичного аналізатора і вхідними даними для виконання процесу рекурсивної редукції. Первинна структура містить структуризоване представлення лексем (слів або символів), а також

синтаксичні зв'язки між ними. Дана структура, по суті, являє собою орієнтований граф, а лексеми – вершини даного графу.

Будь-який ПМ текст T^T представляється множиною лексем L , на якій визначене відношення передування $\prec$. Дане відношення перетворює L в лінійно впорядковану множину. Також текст T^T можна представити як послідовність речень S , на яких також визначене відношення передування:

$$T^T = \{S_1 \prec S_2 \prec \dots \prec S_{n_s}\} \quad (2.8)$$

де n_s – загальна кількість речень в тексті.

Кожне речення S_i , в свою чергу, представлене деякою підмножиною лексем:

$$L_{S_i} = \{l_{ij}, j = \overline{1..n_i}\} \quad (2.9)$$

де n_i – кількість лексем в i -му реченні.

Очевидно, що виконується умова:

$$\forall l_1 \in S_1, \forall l_2 \in S_2, S_1 \prec S_2 \Rightarrow l_1 \prec l_2 \quad (2.10)$$

де S_1, S_2 – довільні речення тексту;

l_1, l_2 – лексеми.

Кожна лексема, в свою чергу, має ряд ознак:

$$l_{ij} = \langle l_{ij}^T, P_{ij} \rangle \quad (2.11)$$

де l_{ij}^T – текстове представлення лексеми l_{ij} ;

P_{ij} – ознаки лексеми l_{ij} .

Лексема може бути зв'язана з іншими лексемами з допомогою синтаксичних зв'язків $r_{sn} \in R_{sn}$:

$$r_{sn} = \langle l^1, l^2, k \rangle \quad (2.12)$$

де l^1, l^2 – лексеми, між якими існує зв'язок;

k – тип зв'язку.

Таким чином орієнтований граф, що представляє первинну структуру ПМ тексту, має вигляд (2.13).

$$T_{sn} = \langle L, R_{sn} \rangle \quad (2.13)$$

Основною проблемою структури (2.13), і, зокрема, (2.11), є неефективність роботи з текстовим представленням l_{ij}^T лексеми, що є надлишковим і вимагає побудови спеціалізованих функцій, визначених на множині текстових представлень слів. Такі функції є громіздкими і неефективними, а при програмній реалізації – часто залежать від особливостей реалізації обробки текстових змінних в заданій мові програмування.

Оскільки множина текстових представлень лексем, очевидно, є зліченною, можливо побудувати перетворення виду (2.14) [51].

$$V: L^T \rightarrow \mathbb{N} \quad (2.14)$$

де L^T – множина текстових представлень лексем.

Нехай текст написано в певному алфавіті W , кількість символів в якому $n_W = \text{card}(W)$. Даний алфавіт можна розглядати як систему числення з основою n_W . Відповідно, кожній букві $w \in W$ можна поставити у відповідність певне число $i_w \in \mathbb{N}$, що являє собою індекс даної букви в алфавіті. Будь-яке слово вхідного тексту являє собою послідовність (2.15).

$$l^T = \{w_1, w_2 \dots w_{n_l}\} \quad (2.15)$$

де n_l – довжина слова $n_l > 0$;

w_i – букви алфавіту W .

Якщо розглядати букви w_i як розряди числа у відповідній системі числення, то таке число можна перевести в десяткову систему числення з допомогою формули (2.16).

$$V(l^T) = i_{w_1} \times (n_W)^{n_l} + i_{w_2} \times (n_W)^{n_l-1} + \dots + i_{w_{n_l}} \times (n_W)^0 \quad (2.16)$$

Де i_{w_j} – індекс букви w_j в алфавіті W ;

n_W – кількість символів в алфавіті W .

З допомогою функції V можна замінити всі l^T на відповідні їм $l^V = V(l^T)$. В результаті такої операції можна отримати більш ефективне представлення (2.17) множини лексем.

$$\langle l^V, P \rangle \in L^V \quad (2.17)$$

де l^V – кодове представлення лексеми l ;

P – граматичні характеристики лексеми;

L^V – множина кодових представлень лексем.

В подальшому L^V можна розглядати в якості множини лексем L .

2.2.2. Метод рекурсивної редукції тексту

Згідно (2.5), процес структуризації тексту, що виконується методом рекурсивної редукції, має структуру (2.18) і перетворює первинну структуру тексту (2.13) в онтологію (2.1).

$$T_{sn} \rightarrow O^1 \rightarrow O^2 \rightarrow O \quad (2.18)$$

Метод рекурсивної редукції полягає в рекурсивному виконанні процесу редукції вхідного ПМ тексту, яке полягає в застосуванні до нього спеціалізованого оператора (2.19).

$$F_{rd} : T_{sn} \rightarrow O \quad (2.19)$$

Оператор редукції (2.19) описується в термінах λ -теорії.

λ -теорія (λ -числення) – формальна теорія, розроблена для формалізації і аналізу поняття обчислюваності [5]. Основними поняттями λ -теорії є аплікація (застосування функції до аргументу) і абстракція. Абстракція означає, що якщо $t(x)$ – вираз який, можливо, містить в собі вільну змінну x , то запис $\lambda x.t(x)$ – це функція f , яка ставить у відповідність значенню a значення $t(a)$, інакше кажучи, має місце рівність (2.20).

$$(\lambda x.t(x))a = t(a) \quad (2.20)$$

де x – певна змінна;

t – вираз, що містить, можливо, входження змінної x ;

a – аргумент, що визначає значення x .

Центральним елементом λ -теорії є поняття терму [5, 40]. Множина Λ , λ -термів визначається індуктивно, як показано виразом (2.21):

$$\begin{aligned} x &\in \Lambda \\ M &\in \Lambda \Rightarrow (\lambda x M) \in \Lambda \\ M, N &\in \Lambda \Rightarrow (MN) \in \Lambda \end{aligned} \quad (2.21)$$

де x – довільна змінна;

M, N – довільні терми.

Крім того в λ -теорії діє так зване спостереження Штейнфінкеля (2.22), що дозволяє зводити функції багатьох аргументів до функції одного аргументу [5].

$$\lambda x_1 \dots x_n. M = \lambda x_1 (\lambda x_2 (\dots (\lambda x_n (M)) \dots)) \quad (2.22)$$

де x_i – довільні змінні;

M – довільний терм.

Важливе значення має поняття β -редукції, яке можна визначити як відношення (2.23) між двома термами:

$$\beta = \{((\lambda x. M) N, M[x := N]) \mid M, N \in \Lambda\} \quad (2.23)$$

де x – довільна змінна;

M, N – довільні терми;

$M[x := N]$ – вираз, отриманий підстановкою N замість змінної x в виразі M .

Оператор редукції (2.19) являє собою комбінацію чотирьох операторів (2.24). при цьому три з операторів виконують кроки перетворення, задані виразом (2.18), а один виконує допоміжну функцію.

$$F_{rd} = F_{l*} \circ F_x \circ F_{smr} \circ F_{ct} \quad (2.24)$$

де F_{l*} – оператор агрегації (2.25), що виконує допоміжну функцію, перетворюючи множину лексем L в множину конструктів L^* . Конструкти є особливою формою лексем, і об'єднують в собі послідовності слів або символів, зокрема, словосполучення. Особливістю конструктів є те, що з точки зору подальшої обробки вони можуть розглядатись як лексеми (тобто як одне слово чи символ).

Таким чином множина $L \cup L^*$ може використовуватись у тих же випадках, що і множина L , зокрема, в якості вхідних даних для оператора

$$F_{l*} : L \cup L^* \rightarrow L^* \quad (2.25)$$

F_X – оператор (2.26) ідентифікації концептів X .

$$F_X : L \cup L^* \rightarrow L \cup L^* \cup X \quad (2.26)$$

F_{smr} – оператор (2.27) ідентифікації онтологічних зв'язків [50] R , що поділяються на зв'язки між концептами R_{sem} і допоміжні зв'язки між концептом і його контекстами R_{sem}^* .

$$F_{smr} : \langle L \cup L^* \cup X, R_{sn} \rangle \rightarrow \langle L \cup L^* \cup X, R_{sn} \cup R_{sem}^* \cup R_{sem} \rangle \quad (2.27)$$

F_{ct} – оператор (2.28) ідентифікації контекстів, які в подальшому виступають в якості атрибутів концептів A_X .

$$F_{ct} : \langle X, R_{sem}^* \rangle \rightarrow A_X \quad (2.28)$$

Сумарне перетворення F_{rd} має вигляд (2.29), де $R = R_{sem}^* \cup R_{sem}$.

$$F_{rd} : \langle L \cup L^* \cup X, A_X, R_{sn} \cup R \rangle \rightarrow \langle L \cup L^* \cup X, A_X, R_{sn} \cup R \rangle \quad (2.29)$$

На першому кроці $L^* = \emptyset$; $X = \emptyset$; $A_X = \emptyset$; $L_{sem}^* = \emptyset$; $L_{sem} = \emptyset$ і тому перетворення (2.29) вироджене в перетворення (2.30). Однак з кожним кроком дані множини поповнюються новими елементами.

$$F_{rd} : \langle L, R_{sn} \rangle \rightarrow \langle L \cup L^* \cup X, A_X, R_{sn} \cup R_{sem}^* \cup R_{sem} \rangle \quad (2.30)$$

Для повного аналізу тексту необхідно використовувати оператор (2.24) рекурсивно, для чого в λ -теорії використовується оператор нерухомої точки [5]. Для цього необхідно побудувати допоміжну функцію F' :

$$F' = \lambda f x. \begin{cases} f F_{rd} x, & F_{rd} x \neq x \\ x & , F_{rd} x = x \end{cases} \quad (2.31)$$

Застосуємо до неї оператор нерухомої точки Y (2.33). Терм YF' є нерухомою точкою для перетворення, що задається функцією F' , тобто для нього виконується

співвідношення (2.34). З допомогою відношення (2.34) можна здійснювати рекурсивне виконання функції (2.31), доки не буде досягнута наявна в ній умова завершення (2.35), тобто доки множина поточних даних (2.32) не стане нерухомою точкою для перетворення F_{rd} .

$$\langle L \cup L^* \cup X, A_X, R_{sn} \cup R_{sem}^* \cup R_{sem} \rangle \quad (2.32)$$

$$Y = \lambda f. (\lambda x. f(xx)) \lambda x. f(xx) \quad (2.33)$$

$$F'(YF') = YF' \quad (2.34)$$

$$F_{rd}x = x \quad (2.35)$$

Умова (2.35) може бути інтерпретована як «застосування F_{rd} більше не ідентифікує нової інформації».

2.2.3. Множинність перетворення рекурсивної редукції

Конструкти $\tilde{L}^*$, що виділяються в ході перетворення (2.25), формуються з допомогою спеціалізованої операції. Як правило, концепт, що описується у вхідному тексті, представлений в ньому певним словосполученням, що в свою чергу представлене певною підмножиною лексем $\tilde{L} \subset L$. Над цією множиною виконується спеціалізоване перетворення (2.36).

$$\tilde{L} \rightarrow \langle V^+(\tilde{l}_1^V, \tilde{l}_2^V \dots \tilde{l}_n^V), P_{M(\tilde{L}^*)}, L^{im} \rangle \quad (2.36)$$

де n – кількість лексем в $\tilde{L}$;

l^V – кодове представлення лексеми l ;

V^+ – операція об'єднання кодових представлень лексем (2.37);

$M(L)$ – операція знаходження головної лексеми множини L ;

L^{im} – утворююча множина ($L^{im} = \tilde{L}$).

Операція V^+ задана рекурсивно виразом (2.37).

$$\begin{aligned} V^+(l^V) &= l^V \\ V^+(l_1^V, l_2^V \dots l_n^V) &= l^V \times [\log_{n_W}(V^+(l_2^V, l_3^V \dots l_n^V)) + 1] + V^+(l_2^V, l_3^V \dots l_n^V) \end{aligned} \quad (2.37)$$

Де $\lfloor \cdot \rfloor$ – операція знаходження цілої частини числа;

n_W – кількість символів в алфавіті W , що використовується для текстового представлення лексем.

Перетворення V^+ є аналогом застосування перетворення V до вихідних текстових представлень лексем (2.38).

$$V^+(l_1^V, l_2^V \dots l_n^V) = V(V^{-1}(l_1^V) + V^{-1}(l_2^V) + \dots + V^{-1}(l_n^V)) \quad (2.38)$$

де V – операція формування кодового представлення лексеми;

V^{-1} – обернена до V операція, тобто $V \circ V^{-1} = V^{-1} \circ V = \lambda_{X.X}$;

$+$ – операція об'єднання (конкатенація) текстових представлень.

Конструкти можуть створюватись також на основі комбінації конструктів і лексем ($\tilde{L} \subset L \cup L^*$). При цьому основною відмінністю є формування утворюючої множини L^{im} з допомогою виразу (2.39).

$$L^{im} = \{l \mid l \in L \cap \tilde{L}\} \cup \bigcup_{l^* \in \tilde{L}} L_l^{im} \quad (2.39)$$

де L – множина лексем вхідного тексту;

$\tilde{L}$ – множина лексем і конструктів, з якої був сформований конструкт;

l – лексеми, що належать $\tilde{L}$;

l^* – конструкти, що належать $\tilde{L}$;

L_l^{im} – утворююча множина конструкта l^* .

Поняття утворюючої множини також можна розширити на множину L , прийнявши (2.40).

$$l \in L \Rightarrow L_l^{im} = \emptyset \quad (2.40)$$

На множині лексем L визначене відношення строгого лінійного порядку $\prec$.

Дане відношення можна розширити і на множину L^* :

$$\forall l_1 \in L_{l_1^*}^{im}, \forall l_2 \in L_{l_2^*}^{im}, l_1 \prec l_2 \Rightarrow l_1^* \prec l_2^* \quad (2.41)$$

де l_1, l_2 – лексеми;

$L_{l_1^*}^{im}, L_{l_2^*}^{im}$ – утворюючі множини конструктів l_1^*, l_2^* .

Але на множині L^* відношення $\prec$ в загальному випадку не буде відношенням лінійного порядку, оскільки можливе виконання умови (2.42).

$$\exists l_1 \in L_{l_1^*}^*, \exists l_2, l_3 \in L_{l_2^*}^*, l_2 \prec l_1 \prec l_3 \quad (2.42)$$

де l_1, l_2, l_3 – лексеми;

$L_{l_1^*}^{im}, L_{l_2^*}^{im}$ – утворюючі множини конструктів l_1^*, l_2^* .

Відношення (2.41) можна тривіальним чином розширити на $L \cup L^*$.

Також на множині $L \cup L^*$ визначене відношення R :

$$l_1 \in L \cup L^*, l_2 \in L^*, l_1 \in L_{l_2^*}^{im} \Rightarrow l_1 R l_2 \quad (2.43)$$

Відношення R , а також відношення слідування $\prec$ являються відношеннями строгого часткового порядку на $L \cup L^*$.

Тепер розглянемо детальніше комбінацію перетворень (2.25) і (2.26), на першому кроці редукції, коли вона має вигляд (2.44).

$$F_{l^*} \circ F_x : L \rightarrow L \cup L^* \cup X \quad (2.44)$$

Дане перетворення ставить у відповідність кожній лексемі $l \in L$ саму лексему, а також певний набір конструктів або концептів, які сформовані на основі даної лексеми. Якщо взяти в якості вихідної множини множину конструктів L^* , то таку відповідність можна представити у вигляді гіпервідношення $G_{im} : L^* \rightarrow L$:

$$l \in L, Y = \{l^* \mid l \in L_{l^*}^{im}\} \Rightarrow Y G_{im} l \quad (2.45)$$

де l – лексема;

l^* – довільний конструкт;

$L_{l^*}^{im}$ – утворююча множина конструкту l^* .

Дане гіпервідношення можна розширити на множину концептів X з допомогою умови (2.46).

$$l \in L, Y = \{x \mid \exists l^*, x(l^*) = x \wedge l \in L_{l^*}^{im}\} \Rightarrow Y G_{im} l \quad (2.46)$$

де l – лексема;

x – довільний концепт;

l^* – довільний конструкт;

$x(l^*)$ – операція формування концепту на основі конструкту l^* ;

$L_{l^*}^{im}$ – утворююча множина конструкту l^* .

З допомогою формул (2.40), (2.45) і (2.46) дане відношення можна тривіальним чином розширити до $G_{im} : L \cup L^* \cup X \rightarrow L \cup L^* \cup X$.

Гіпервідношення G_{im} показує множинність відношення рекурсивної редукції і може бути використане в якості додаткового механізму формування онтологічних зв'язків R між концептами ПдО.

2.2.4. Структура оператора редукції та його складових

Як вже було сказано вище, оператор (2.24) являє собою композицію операторів. Кожен з них виконує один крок перетворення (2.5). Повний цикл даного перетворення структурує частину інформації, що міститься у вхідному тексті, після чого перетворення рекурсивно викликається заново, доки не буде виділена вся інформація. Однак кожна зі складових оператора редукції також може бути розбита на складові.

В загальному випадку оператор виконання перетворення F задається базою правил G_R виконання даного перетворення. Правило $g \in G_R$ має уніфіковану для всіх етапів структуру:

$$g = \langle f_{ap}^g, f_{tr}^g \rangle \quad (2.47)$$

де f_{ap}^g – функція застосовності, що визначає, чи може правило бути застосоване до певного набору вхідної інформації;

f_{tr}^g – функція перетворення, що задає перетворення вхідної інформації.

Задане правилом g перетворення $F_g : X \rightarrow Y$ має вигляд (2.48).

$$F_g(x) = \begin{cases} f_{tr}^g(x), f_{ap}^g(x) \\ x, \neg f_{ap}^g(x) \end{cases} \quad (2.48)$$

Кожна функція застосовності являє собою лямбда-терм виду [15, 17, 76]:

$$f_{ap} = (\lambda x_1, x_2 \dots x_{n_g}. t_{ap}(x)) a_1, a_2 \dots a_{n_g} = t_{ap}(a_1, a_2 \dots a_{n_g}) \quad (2.49)$$

де запис λx вказує, що дана конструкція являється λ -термом;

x_i – змінна, що приймає значення на множині $L \cup L^*$;

a_i – аргумент функції, що задає значення x_i ;

n_g – кількість аргументів, що повинні бути подані на вхід функції перетворення;

t_{ap} – умова застосовності, вираз, що містить n_g змінних.

В загальному випадку умова застосовності t_{ap} в загальному випадку означає існування гомеоморфізму між орієнтованим графом, утвореним вхідною послідовністю лексем (а також синтаксичними зв'язками між ними) і певним еталонним орієнтованим графом G_{ap} , що являє собою вибраний користувачем підграф первинного представлення T_{sn}^e певного тексту. В якості T_{sn}^e може виступати первинне представлення як поточного тексту T_{sn} , так і будь-якого іншого тексту (наприклад, тезауруса ПдО). Умова має структуру (2.50) і складається з *предикатів ідентифікації* [16, 80, 81]. Такі предикати дозволяють ідентифікувати контексти певної лексеми, і на основі цього робити висновок про необхідність або відсутність необхідності виконання перетворення. Кожен з предикатів задає певну умову, і умовою застосовності правила є виконання всіх умов, заданих кожним з предикатів. Кількість предикатів у виразі задає число n_g .

$$t_{ap} = c_{p_1}(x_1) \& \dots c_{p_n}(x_{n_g}) \& r_{k_{11}}(x_1, x_1) \& \dots r_{k_{n_g n_g}}(x_{n_g}, x_{n_g}) \quad (2.50)$$

Одномісні предикати, присутні в виразі – це предикати ідентифікації лексем. Такий предикат задає умову, якій повинна відповідати певна лексема (або конструкт) з вхідної множини. Предикат має структуру (2.51).

$$c_p(l) = \begin{cases} 1, p = 0 \vee p = l^T \vee p \in P_l \\ 0, p \neq 0 \wedge p \neq l^T \wedge p \notin P_l \end{cases} \quad (2.51)$$

Робота такого предикату залежить від *шаблонного параметра* p . В залежності від типу даного параметра предикат може бути:

- 1) Стандартним предикатом ідентифікації. В такого предиката p – це морфологічна характеристика лексеми. Такий предикат визначає, чи має вхідна лексема задану характеристику ($p \in P_l$).
- 2) Предикатом ідентифікації ключових слів. В таких предикатів p – це текстове представлення необхідної лексеми, і це значення порівнюється зі значенням вхідної ($p = l^T$). Даний предикат завжди має значення 0 для конструктів.
- 3) Нульовий предикат ($p = 0$). Такий предикат завжди має значення 1, незалежно від поданої на вхід лексеми.

Двомісний предикат – це предикат ідентифікації зв'язків. Такий предикат визначає, чи міститься між двома заданими лексемами зв'язок заданого типу. Предикат має вигляд (2.52).

$$r_k(l_1, l_2) = \begin{cases} 1, k = 0 \vee \langle l_1, l_2, k \rangle \in R_{sn} \\ 0, k \neq 0 \wedge \langle l_1, l_2, k \rangle \notin R_{sn} \end{cases} \quad (2.52)$$

Як і предикат ідентифікації лексем, даний предикат має нульову модифікацію, яка має значення 1 незалежно від вхідних даних. При цьому для коректного використання умови (2.50) повинна виконуватись задана структурою лексичного аналізатора умова:

$$k_{ij} = 0, i = j \quad (2.53)$$

База правил задає перетворення F , що має вигляд:

$$F_G(L) = \bigcup_{\tilde{L} \in P(L)} \bigcup_{g \in G} F_g^*(\tilde{L}) \quad (2.54)$$

де $P(L)$ – множина всіх підмножин L ;

F_g^* – модифікована функція (2.48), доповнена додатковими умовами.

Додаткових умов накладається дві:

Умова порядку означає, що всі елементи вхідної підмножини повинні бути лінійно впорядковані певним відношенням строгого порядку G , і має вигляд:

$$f_{ord}(x) = \begin{cases} 1, \forall x_1, x_2 \in x, x_1 \prec x_2 \vee x_2 \prec x_1 \\ 0, \exists x_1, x_2 \in x, x_1 \not\prec x_2 \wedge x_2 \not\prec x_1 \end{cases} \quad (2.55)$$

Відношення порядку, що можуть використовуватись в якості G :

- 1) Відношення слідування $\prec$
- 2) Транзитивне замикання відношення R (2.43)
- 3) Транзитивне замикання відношення, заданого зв'язками R_{sem}

Умова узгодженості f_{ap+}^g визначає, чи є даний елемент $\tilde{L} \in P(L)$ придатним для обробки правилом g :

$$f_{ap+}^g(x) = \begin{cases} 1, card(x) = n_g \\ 0, card(x) \neq n_g \end{cases} \quad (2.56)$$

З урахуванням даних умов перетворення (2.48) перетворюється на F_g^* :

$$F_g^*(x) = \begin{cases} f_{tr}^g(x) & , f_{ord}(x) \wedge f_{ap+}^g(x) \wedge f_{ap}^g(x) \\ x & , \neg f_{ord}(x) \vee \neg f_{ap+}^g(x) \vee \neg f_{ap}^g(x) \end{cases} \quad (2.57)$$

Розглянемо особливості виконання різних кроків редукції в рамках структури (2.57).

Перетворення агрегації – перетворення, що формує структури (конструкти) зі зв'язаних між собою лексем, які в подальшому стають кандидатами в концепти або контексти концептів. Особливістю роботи перетворення є те, що воно застосовується до речення в цілому, тобто до множини лексем L , асоційованої з певним реченням S . Перетворення формує множину конструктів L^* , яка в подальшому також асоціюється з реченням S , і використовується з разом з початковою множиною лексем. Всі подальші перетворення використовують в якості вхідних даних множину $L \cup L^*$, на якій зберігається відношення слідування $\prec$, але яке при цьому перетворюється у відношення часткового порядку.

Правила даного перетворення відрізняються від стандартної структури способом застосування і способом формування функції перетворення. На даному етапі використовується модифікована функція F_g^* , що має вигляд (2.58).

$$F_g^*(x) = \begin{cases} f_{tr}^g(x) & , f_{ord}(x) \wedge f_{ap+}^g(x) \wedge f_{ap}^g(x) \\ \emptyset & , \neg f_{ord}(x) \vee \neg f_{ap+}^g(x) \vee \neg f_{ap}^g(x) \end{cases} \quad (2.58)$$

Функція (2.58), на відміну від функції (2.57), повертає пусту множину у випадку своєї незастосовності. Завдяки цьому в результуючу множину не попадають елементи вхідної множини L .

Функція f_{tr}^g являє собою функцію агрегації вигляду (2.59).

$$f_g^{tr} : \{l_1 \dots l_{n_g}\} \rightarrow \{l^* = \langle V^+(\tilde{l}_1^V, \tilde{l}_2^V \dots \tilde{l}_{n_g}^V), P_{l_{n_r}}, L_{l^*}^* \rangle\} \quad (2.59)$$

де l^* – результуючий конструкт;

$L_{l^*}^*$ – множина лексем чи конструктів, з яких сформований конструкт;

V^+ – операція об'єднання числових представлень $\tilde{l}_i^V$ лексем вхідної множини;

n_r – певний індекс ($1 \leq n_r \leq n_g$), що визначає, яке слово є головним в виділеному словосполученні (наприклад, в словосполученні з прикметника і іменника головним буде іменник). Індекс n_r визначає, від якої лексеми конструкт l^* наслідує свої морфологічні характеристики P_{l^*} . У випадку, якщо конструкт сформований не словами (наприклад, дата), даний параметр не має змісту і його можна вважати нулем.

На основі даного індексу будується операція M знаходження головного слова словосполучення:

$$M(L) = l_{n_r} \quad (2.60)$$

Власне перетворення агрегації виконується рекурсивним застосуванням функції (2.58) за схемою (2.61).

$$F_{ag}(\tilde{L}) = \begin{cases} \tilde{L} & , F_G(L \cup \tilde{L}) = \tilde{L} \\ F_{ag}(F_G(L \cup \tilde{L})), F_G(L \cup \tilde{L}) \neq \tilde{L} \end{cases} \quad (2.61)$$

Перетворення ідентифікації концептів – перетворення, що формує множину концептів X на основі множини $L \cup L^*$. Дане перетворення виконується після перетворення агрегації і використовує сформовані з допомогою даного перетворення конструкти L^* .

В загальному дане перетворення має стандартну структуру. Умова застосовності часто має тривіальну структуру:

$$t_{ap} = c_{p_1}(x_1) \quad (2.62)$$

Де c_{p_1} ідентифікує іменники. Таким чином словосполучення, головним словом для яких є іменник, і для яких були сформовані конструкти на етапі агрегації, зразу попадають під дію даного правила і з них формуються концепти.

Функція перетворення f_{tr}^g тривіальна:

$$f_{tr}^g(L^*) = x(l_1^*) \in X \quad (2.63)$$

де $x(l)$ – функція формування концепту x з лексеми l .

При цьому з лексем L формуються ті концепти, що описані одним словом, а з конструктів L^* – кількома.

Перетворення ідентифікації зв'язків – перетворення, що формує множину семантичних зв'язків між концептами R_{sem} , семантичних зв'язків між концептами і лексемами/конструктами R_{sem}^* на основі множин $L \cup L^*$ і X .

Перетворення має стандартну структуру, при цьому умова застосовності як правило має вигляд:

$$t_{ap} = c_{p_1^*}(x_1) \& c_{p_1}(x_2) \& \dots \& c_{p_{n_g-2}}(x_{n_g-1}) \& \\ \& c_{p_2^*}(x_{n_g}) \& r_{k_{12}}(x_1, x_2) \& \dots \& r_{k_{n_g n_g}}(x_{n-1}, x_{n_g}) \quad (2.64)$$

Де p_1^* , p_2^* – умови, що визначають тип лексем чи конструктів, між якими (або відповідними їм концептами) буде встановлений зв'язок;

$p_1 \dots p_{n_g-2}$ – умови, що визначають ключові слова, характерні для типу зв'язку, що ідентифікується правилом.

Функція перетворення має вигляд (2.65).

$$f_{tr}^g(L^*) = \begin{cases} \langle x(l_1^*), x(l_{n_g}^*), k_{sem}^g \rangle \in R_{sem}, x(l_1^*) \in X \wedge x(l_{n_g}^*) \in X \\ \langle x(l_1^*), l_{n_g}^*, k_{sem}^g \rangle \in R_{sem}^* & , x(l_1^*) \in X \wedge x(l_{n_g}^*) \notin X \\ \langle x(l_{n_g}^*), l_1^*, k_{sem}^g \rangle \in R_{sem}^* & , x(l_1^*) \notin X \wedge x(l_{n_g}^*) \in X \\ 0 & , x(l_1^*) \notin X \wedge x(l_{n_g}^*) \notin X \end{cases} \quad (2.65)$$

де $l_1^* \dots l_{n_g}^*$ – конструкти вхідної множини;

$x(l)$ – функція формування концепту x з лексеми l ;

k_{sem}^g – тип семантичного відношення, що ідентифікується поточним правилом.

Дане перетворення визначає, який саме зв'язок ідентифікується в його результаті. Якщо на етапі ідентифікації концептів були ідентифіковані і створені концепти як на основі l_1^* , так і на основі $l_{n_g}^*$, то сформований зв'язок буде віднесений до множини R_{sem} . Якщо тільки для одного з конструктів концепт не був створений, то такий зв'язок попаде в множину R_{sem}^* і буде використовуватись в для ідентифікації контекстів зв'язаного з ним концепту. Якщо ж концепт не був створений на основі жодного з концептів, то зв'язок не створюється. Скоріше за все, така ситуація означає, що відсутнє правило агрегації, яке об'єднало б дані конструкти в один.

Перетворення ідентифікації контекстів призначене для аналізу контекстів концептів, заданих відношеннями $R_{sem}^* \cup R_{sem}$, і формування на їх основі атрибутів, що в подальшому задають функції інтерпретації концептів.

Перетворення ідентифікації контекстів використовує модифікований предикат ідентифікації зв'язків (2.52), що призначений для ідентифікації семантичних зв'язків між концептами. Предикат має наступний вигляд:

$$r_k^*(x_1, x_2) = \begin{cases} 1, \langle x_1, x_2, k \rangle \in R_{sem}^* \cup R_{sem} \\ 0, \langle x_1, x_2, k \rangle \notin R_{sem}^* \cup R_{sem} \end{cases} \quad (2.66)$$

де x_1, x_2 – певні концепти, $x_1, x_2 \in X$.

Функція застосовності не використовує предикати ідентифікації лексем (2.51), і має наступний вигляд:

$$t_{ap} = r_{k_{12}}^*(x_1, x_2) \quad (2.67)$$

Функція перетворення для даного випадку також тривіальна і має вигляд:

$$f_{tr}^g(L^*) = a_{l_2}^*(x(l_1^*)) \in A_X \quad (2.68)$$

де $a_l(x)$ – функція формування атрибуту з лексеми або конструкту l для концепту x .

Перетворення попередньої структуризації F_{ps} не є частиною оператора редукції, але здійснюється аналогічно до його частин. Дане перетворення може доповнювати процес синтаксичного аналізу.

Попередня структуризація тексту часто буває необхідною при роботі з слабо структурованими документами. Особливо важливо це при обробці таблиць, що містять заголовки, але при цьому значення комірок таблиць містять природну мову, яку необхідно аналізувати аналогічно до ПМ тексту для отримання необхідної інформації. Також як правило концепти містяться в заголовках розділів тексту, причому наявний в заголовку концепт може бути назвою певної категорії, до якої належать задані в даному розділі концепти.

Все, що необхідно зробити при попередній структуризації – це ідентифікувати в тексті іменовані області (такі, як розділи чи колонки таблиці). Для цього потрібно:

- 1) Визначити лексеми тексту, що формують ім'я іменованої області.
- 2) Визначити лексеми тексту, що формують знайдену іменовану область.

Ім'я області – це, як правило, певний блок тексту (часто речення), що виділене розміткою особливим чином, що, в свою чергу, може бути здійснено двома способами:

- 1) Явно – з допомогою певного символу-роздільника, як, наприклад, кома в CSV (Comma Separated Values) файлах.
- 2) Неявно – з допомогою налаштувань відображення тексту, наприклад, з допомогою виділення іншим шрифтом.

При зчитуванні документів, що підтримують неявне виділення, необхідно виконувати додаткову обробку з метою отримання метаданих відображення лексеми. В такому випадку множина ознак P_l лексеми l складатиметься з двох частин – морфологічних ознак P_l^{mph} і метаданих відображення P_l^{dis} .

$$P_l = P_l^{mph} \cup P_l^{dis} \quad (2.69)$$

Функція перетворення попередньої структуризації має стандартний вигляд (2.48) – (2.52). Однак, вона використовує спеціалізовану структуру умови застосовності (2.50), що має форму (2.70).

$$\begin{aligned} t_{ap} = & c_d(x_1) \& c_t(x_2) \& \dots \& c_t(x_{n_g-1}) \& \\ & \& c_d(x_{n_g}) \& r_0(x_1, x_2) \& \dots \& r_0(x_{n-1}, x_{n_g}) \end{aligned} \quad (2.70)$$

Умова використовує тільки два шаблонних параметра: d являє собою умову-обмежувач, а t – умову співставлення:

- 1) Для явного способу задавання c_d ідентифікує символ-роздільник, тоді як c_t не накладає ніяких умов ($c_t = c_0$).
- 2) Для неявного способу задавання c_d ідентифікує будь-які лексеми, що, згідно стилів поточного документу, не відносяться до заголовків, тоді як c_t навпаки, ідентифікує будь-які лексеми, що до них відносяться.

На етапі попередньої структуризації не використовуються синтаксичні зв'язки між лексемами, тому умови на них також не накладаються ($\forall i, \forall j, k_{ij} = 0$).

При знаходженні лексем, що належать іменованому блоку, при явному способі задавання правило аналогічне, тоді як при неявному задаванні міняються місцями умови d і t .

Функція перетворення попередньої структуризації виконує перетворення, що має вигляд (2.71) або (2.72), де p_n і p_n^* – властивість належності лексеми до іменованої області і належність лексеми до назви іменованої області відповідно.

$$f_{tr} : \langle l^T, P \rangle \rightarrow \langle l^T, P \cup \{p_n\} \rangle \quad (2.71)$$

$$f_{tr} : \langle l^T, P \rangle \rightarrow \langle l^T, P \cup \{p_n^*\} \rangle \quad (2.72)$$

Умови p_n і p_n^* при необхідності можна використовувати на наступних етапах, задавши в шаблонах правил відповідні параметри. Зокрема таким чином можна включити заголовки тексту у вихідну онтологію в якості об'єктів.

2.2.5. Формування правил на основі первинної структури тексту

Оскільки принцип роботи кожного з предикатів, що входять до умови застосовності (2.50) залежить виключно від асоційованого з ним шаблонного параметра p або k , то будь-яка функція застосовності ідентифікується набором таких параметрів. Функція перетворення f_{tr} , з іншої сторони, залежить виключно від кроку, що виконується.

В результаті правило виконання перетворення g можна представити в значно більш компактній формі – у вигляді шаблону правила $\tilde{g}$, що представляється виразом виду (2.73).

$$\tilde{g} = \langle \{p_i\}, \{k_{ij}\} \rangle \quad (2.73)$$

Тривіальним чином можна побудувати перетворення інтерпретації шаблонів:

$$F_{int} : \tilde{G} \rightarrow G \quad (2.74)$$

Представлення (2.73) є значно зручнішим для користувача, оскільки його легше представляти в текстовій формі. Однак процес формування правил перетворення можна спростити навіть більше.

Вираз (2.73), по суті накладає певну умову на топологію графової структури, сформованої вхідною множиною і зв'язками між її елементами. Оскільки як конструкти, так і концепти в рамках виконуваних над ними операцій можуть розглядатись як аналог лексем, то і сформована ними графова структура може розглядатись як аналог структури, сформованої лексемами, а саме – первинної структури тексту (2.13). Важливим наслідком даного факту є те, що для створення шаблонів виду (2.73) легко розробити процедуру автоматизованого створення, оскільки така процедура зводиться до простої в реалізації функції вибору підграфу

$G_{ap} \subset T_{sn}^e$. Важливо зазначити, що в якості T_{sn}^e може використовуватись і первинна структура тексту, що обробляється – тобто формування шаблонів може виступати в ролі додаткового кроку після синтаксичного аналізу і перед агрегацією (2.25).

Іншим важливим наслідком існування гомоморфізму між первинною структурою тексту і структурою інформації, що з нього виділяється, є можливість побудови спеціалізованої НС NS' , призначеної для роботи не з об'єктами онтології, а лексемами первинної структури [80]. Побудований на її основі інтерактивний документ $\langle T_{sn}, NS' \rangle$ може здійснюватися для формування G_{ap} .

Процедуру формування шаблонів можна спростити ще більше, формуючи спеціалізовані еталонні тексти, що містять виключно приклади вживання деяких словесних конструкцій. В такому випадку можна виключити крок вибору підграфу, прийнявши $G_{ap} = T_{sn}^e$.

2.3. Формування онтологічних ГІС-додатків і трансдисциплінарне представлення геопросторової інформації з їх допомогою

Як вже було сказано, експертів у корпоративному інформаційному середовищі може бути представлена у вигляді системи $\{\text{дія} \rightarrow \text{результати}\}$ і являти собою аналог натуральної системи SN . На основі такої системи можна будувати інтерактивні документи виду (2.6). Для формування інтерактивного документу можна побудувати спеціалізоване перетворення виду (2.75).

$$\check{G}: O \Rightarrow SN \quad (2.75)$$

Розглянемо натуральні системи SN , що характеризуються набором з n «дій» $\check{x}^1 \dots \check{x}^n$ і одним «результатом» $\check{y}$ [52], що зв'язані залежністю:

$$\check{y} = \check{f}(\check{x}^1 \dots \check{x}^n) \quad (2.76)$$

При формуванні натуральних система на основі онтологій залежність $\check{f}$ в загальному випадку, має наступний вигляд:

$$\check{f}(\check{x}^1 \dots \check{x}^n) = D(Q_n(Q_{n-1}(\dots Q_1(Q_0(X), \check{x}^1) \dots, \check{x}^{n-1}), \check{x}^n)) \quad (2.77)$$

Де X – множина об'єктів початкової онтології O ;

Q_i – допоміжні функції обробки;

D – функція відображення, що дозволяє відображувати користувачу результати обробки об'єктів.

Можна визначити базовий набір допоміжних функцій наступним чином:

Функція ієрархічної фільтрації (2.78).

$$Q_h(X, x^*) = \{\bar{x} \in X \mid \bar{x} R x^* \vee (\exists \tilde{x}, \tilde{x} \bar{R} x^* \wedge \bar{x} \in Q_h(\bar{x}))\} \quad (2.78)$$

де X – множина об'єктів, що належать певній онтології;

x^* – об'єкт, відносно якого здійснюється фільтрація;

$\bar{R}$ – певне відношення між об'єктами.

Така функція дозволяє фільтрувати об'єкти, що зв'язані певним відношенням $\bar{R}$. Також умова перевіряється рекурсивно, що еквівалентно використанню транзитивного замикання $\bar{R}$.

Відношення $\bar{R}$ повинно базуватись на множині зв'язків між об'єктами R :

$$\exists k, \langle x_1, x_2, k \rangle \in R \Rightarrow x_1 \bar{R} x_2$$

$$\exists k, \langle x_1, x_2, k \rangle \in R \Rightarrow x_2 \bar{R} x_1$$

$$\exists k, \langle x_1, x_2, k \rangle \in R \vee \langle x_2, x_1, k \rangle \in R \Rightarrow x_1 \bar{R} x_2$$

де x_1, x_2 – певні об'єкти;

R – множина зв'язків між об'єктами;

k – тип відношення між об'єктами.

При необхідності тип відношення k можна зафіксувати, що приведе до формування відношення $\bar{R}_k$, заданого виразами (2.79) – (2.81).

$$\langle x_1, x_2, k \rangle \in R \Rightarrow x_1 \bar{R}_k x_2 \quad (2.79)$$

$$\langle x_1, x_2, k \rangle \in R \Rightarrow x_2 \bar{R}_k x_1 \quad (2.80)$$

$$\langle x_1, x_2, k \rangle \in R \vee \langle x_2, x_1, k \rangle \in R \Rightarrow x_1 \bar{R}_k x_2 \quad (2.81)$$

Функція атрибутивної фільтрації має вигляд:

$$Q_a(X, A) = \{\bar{x} \in X \mid A \cap A_{\bar{x}} = A\} \quad (2.82)$$

де X – множина об'єктів, що належать певній онтології;

A – множина атрибутів, відносно яких здійснюється фільтрація;

A_x – множина атрибутів об'єкту x .

Функція Q_a дозволяє виділити множину об'єктів, що володіють певним атрибутом або певним значенням певного атрибуту.

Функція контекстної зв'язки дозволяє встановлювати зв'язки між об'єктами, описаними в документах, що належать до певного інформаційного середовища. Дана функція є комбінацією двох функцій.

Функція індексації має вигляд:

$$Q_I(C) = \bigcup_{T \in C} \bigcup_{l \in L_T} \{ \langle V(l), V(T) \rangle \} \quad (2.83)$$

де C – вхідна множина документів;

T – певний документ;

L_T – набір лексем, що формує текстове представлення T ;

$V(l), V(T)$ – ідентифікатори лексеми l і документу T відповідно.

Результати функції індексації використовуються функцією пошуку:

$$Q_S(I, l) = \{ T \mid \langle V(l), V(T) \rangle \in I \} \quad (2.84)$$

де I – індекс, що являє собою результат роботи Q_I ;

$V(l), V(T)$ – ідентифікатори лексеми l і документу T відповідно.

З допомогою функції пошуку можна формувати зв'язки між об'єктами, що належать різним онтологіям (або між об'єктом онтології і неструктурованими інформаційним ресурсом). При цьому необхідно враховувати дві особливості:

- 1) Контекст об'єкту як правило складається з великої кількості атрибутів, які при побудові індексу можна розглядати як єдиний текст.
- 2) Як назва, так і контекст об'єкту в загальному випадку складається з багатьох лексем.

З врахуванням даних особливостей функція контекстної зв'язки матиме вигляд:

$$Q_c(x) = \bigcup_{l \in L_x} Q_S(Q_I(C), l) \quad (2.85)$$

де C – множина документів, що представляє інформаційне середовище, в рамках якого здійснюється зв'язка;

x – об’єкт, з яким здійснюється зв’язка;

L_x – текстове представлення контексту x .

l – певна лексема.

Основним недоліком функції (2.85) є те, що в її результаті формується неупорядкована множина документів, яка може бути достатньо великою за розміром, і тому незручною для обробки експертом. Вирішити цю проблему можна двома способами:

- 1) Задати на множині результатів відношення порядку, в якості якого може виступити відношення релевантності R_{rel} :

$$T_1 R_{rel} T_2 \Rightarrow card(L_{T_1} \cap L_x) > card(L_{T_2} \cap L_x) \quad (2.86)$$

де L_{T_1}, L_{T_2} – текстові представлення документів T_1, T_2 ;

L_x – контекст об’єкту x , з яким була здійснена контекстна зв’язка.

- 2) Виключити з результату документи, що мають недостатню релевантність, що можна застосуванням в функції (2.85) операції перетину замість операції об’єднання.

2.3.1. Трансдисциплінарне представлення інформації з допомогою інтерактивних документів

Трансдисциплінарне представлення множин онтологій базується на перетворенні між двома гіпермножинами:

$$f^{ct} : R \rightarrow F \quad (2.87)$$

Де R, F – множини зв’язків і функцій інтерпретації певної онтології O .

Для множини онтологій можна побудувати гіпермножини:

$$\mathfrak{R} = \bigcup_i R_i, \mathcal{F} = \bigcup_i F_i \quad (2.88)$$

Де i – індекс, що визначає певну онтологію $O_i = \langle X_i, R_i, F_i \rangle$.

На цих гіпермножинах можна побудувати зворотне до f^{ct} перетворення:

$$f^{tt} : \mathcal{F} \rightarrow \mathfrak{R} \quad (2.89)$$

Тобто певний зв'язок між об'єктами різних за тематикою онтологій може бути представлений не пустою множиною інтерпретуючих функцій з даних онтологій.

Дане перетворення можна розширити і на множину неструктурованих текстів, застосувавши до кожного з текстів оператор редукції (2.24).

Розглянемо процедуру трансдисциплінарного представлення певної множини онтологій засобами інтерактивних документів. Дане представлення базується на функції (2.85). Для його здійснення необхідно виконати трансдисциплінарне перетворення з допомогою функції:

$$Q_{\Pi}(C) = \bigcup_{x \in X_C} \{Q_C(x)\} \quad (2.90)$$

Де C – множина онтологій;

X_C – множина об'єктів, що належать об'єднанню онтологій $\bigcup_{O \in C} O$.

Застосування функції $Q_C(x)$ формує множину об'єктів з різних онтологій, що представляє певне гіпервідношення між відповідними об'єктами. З допомогою сформованих таким чином гіпервідношень можна побудувати трансдисциплінарне представлення O' множини онтологій $C = \{< X_i, R_i, F_i >\}$:

$$C \xrightarrow{Q_{\Pi}} < \bigcup_i X_i, \bigcup_i R_i \cup Q_{\Pi}(C), \bigcup_i F_i > \quad (2.91)$$

Перетворення (2.91) задає найбільш повне представлення наявної в C інформації, що не завжди зручно. Часто необхідно виконати представлення одної вибраної онтології O . В такому випадку (2.90) необхідно змінити:

$$Q_{TO}(O, C) = \bigcup_{x \in X_O} \{Q_C(x)\} \quad (2.92)$$

Де C – множина онтологій;

X_O – множина об'єктів, що належать онтології O .

Для побудови систем, що використовують трансдисциплінарне представлення інформації, важливими є такі твердження:

Твердження 2.1. На основі однієї онтології O , що належить множині документів C , можна сформулювати довільну кількість НС.

Доведення. Нехай існує певна НС виду (2.76), побудована на основі онтології $O = \langle X, R, F \rangle$ з допомогою перетворення (2.75), і множина документів C . Оскільки НС визначається своєю функцією виду (2.77), можна сформувати довільну кількість натуральних систем, міняючи склад і кількість функцій $Q_0 \dots Q_n$. Зокрема:

- 1) Якщо зафіксувати довільним чином об'єкт $\tilde{x} \in X$ і застосувати до онтології функцію ієрархічної фільтрації, то можна взяти $Q_0(X) = Q_h(X, \tilde{x})$ і отримати систему, що відображує певний клас об'єктів;
- 2) Якщо вибрати в якості Q_0 функцію, що виконує трансдисциплінарне представлення (2.91), то отримаємо систему, що відображує трансдисциплінарне представлення онтології O'

Послідовно застосовуючи функцію Q_h з різними параметрами і функцію, що виконує трансдисциплінарне представлення, можна формувати довільну кількість початкових перетворень онтології O , використання яких в формулі (2.77) і дозволить сформувати довільну кількість натуральних систем.

Наслідком Твердження 2.1 є те, що на основі одної онтології можна побудувати довільну кількість інтерактивних документів.

Натуральні системи виду (2.76) є найпростішим варіантом натуральних систем. Однак така система може бути неефективною, зокрема, у наступних випадках:

- Якщо інформація може бути представлена кількома способами (наприклад, у вигляді таблиці та електронної карти), то, як правило, зарання не відомо, який зі способів є більш зручним для експерта;
- Якщо інформація є різною за змістом і структурою, то може виникнути необхідність представляти різні її підмножини різними способами;

Тому часто є зміст формувати на основі одної онтології кілька натуральних систем, що діятимуть незалежно і прийматимуть один і той же набір вхідних «дій» від користувача, і надаючи йому різні «результати», з яких він зможе вибрати той, що відповідає поставленій задачі.

Твердження 2.2. Комбінація незалежних натуральних систем SN_i , що приймають на вхід один і той же набір «дій», є натуральною системою.

Доведення. Згідно до наведеного в [52], визначення натуральної системи, якщо хоча б один елемент в складі системи при виконанні дії даватиме такий же результат, як і при роботі незалежно від інших елементів системи. Якщо розглядати кожну з SN_i як елемент певної метасистеми S , то завдяки незалежності SN_i їх робота в складі системи не буде відрізнятись від роботи за межами системи. Тому S також буде натуральною системою.

Така система S буде розширенням (2.76) і матиме вигляд (2.93).

$$\begin{aligned}\bar{y}^1 &= \bar{f}^1(\bar{x}^1 \dots \bar{x}^n) \\ &\vdots \\ \bar{y}^m &= \bar{f}^m(\bar{x}^1 \dots \bar{x}^n)\end{aligned}\tag{2.93}$$

2.3.2. Трансдисциплінарне представлення геопросторової інформації з допомогою онтологічних ГІС-додатків

Онтологічний ГІС-додаток виду (2.7) також формується на основі НС, що може бути представлена виразом (2.76). В найпростішому випадку таку НС можна розглядати як звичайну НС зі спеціальною функцією відображення – функцією відображення в якості маркера електронної карти D_m . Однак основною особливістю даної функції є те, що відображувати в якості маркера є можна не всі концепти, а тільки ті, що являються елементами шару (який, наприклад, може бути представлений певною категорією концептів), і атрибути яких містять географічну інформацію. Тому перед застосуванням D_m необхідно провести фільтрацію:

$$\bar{T}_{lr} = Q_a(Q_h(X, x_{lr}), A_{geo})\tag{2.94}$$

де $\bar{T}_{lr}$ – таксономія, сформована підмножиною концептів онтології, що можуть бути відображені в якості маркерів певного шару;

x_{lr} – об'єкт, що визначає клас об'єктів, які формують шар ГІС;

A_{geo} – множина атрибутів, що можуть містити географічну інформацію.

В результаті такої фільтрації можна отримати набір таксономій (як правило, таких, що не пересікаються) $\tilde{T}_{lr}$. Даний набір може бути відображений з допомогою D_m :

$$D_m \left(\bigcup_{x_{lr} \in X_{lr}} \tilde{T}_{lr} \right) \quad (2.95)$$

де X_{lr} – множина класів концептів, що представляють собою шари ГІС;

x_{lr} – концепт, що визначає клас концептів, які формують шар ГІС;

$\tilde{T}_{lr}$ – таксономічне представлення шару, заданого x_{lr} ;

D_m – функція відображення об'єкту в якості маркера електронної карти.

Сформована на основі такої процедури НС і буде формувати онтологічний ГІС-додаток виду (2.7). Привести вираз (2.7) до стандартної структури інтерактивного документу можна наступним чином:

Множину географічних сутностей X_g можна сформувати, об'єднавши таксономічні представлення шарів ГІС (2.94), як і у виразі (2.95):

$$X_g = \bigcup_{x_{lr} \in X_{lr}} \tilde{T}_{lr} \quad (2.96)$$

На основі сформованої таким чином множини X_g можна сформувати і множину $R_g \subset R$ з допомогою виразу (2.97). Як правило, така множина відповідатиме об'єднанню множин зв'язків таксономічних представлень $\tilde{T}_{lr}$ шарів ГІС, однак можливі і спеціалізовані зв'язки між об'єктами різних таксономічних представлень.

$$R_g = \{ \langle x_1, x_2 \rangle \in R \mid x_1 \in X_g \wedge x_2 \in X_g \} \quad (2.97)$$

Множина станів задачі S представляється послідовністю «результатів» $\tilde{y}$ роботи натуральної системи, що були сформовані у відповідь на надані користувачем «дії»:

$$S = \{ \tilde{y} \} \quad (2.98)$$

Множина аналітичних операцій над географічними об'єктами A^{op} по суті являє собою підмножину множини допоміжних функцій Q , що формують залежність (2.77):

$$A^{op} \subset Q \quad (2.99)$$

Дану множину можна розбити на дві підмножини: множину операцій аналізу просторового відношення об'єктів A^{rel} і множину вимірювальних операцій A^{ms} .

Таблиця 2.1. Операції аналізу просторових відношень об'єктів

Бінарні відношення	Відношення включення	Співпадає
		Містить в собі
		Міститься в
	Відношення перетину	Перетинається з
		Межує з
		Відділений від
	Відношення суміжності	Сусідство
Інші відношення	віддаленості	

Операції аналізу просторового відношення об'єктів дозволяють виділити множину об'єктів, таких, що зв'язана з ними геопросторова інформація відповідає певному критерію. Такі операції можуть використовуватись для фільтрації об'єктів, в комбінації з функціями (2.78) – (2.84). Класифікація таких операцій представлена в Таблиця 2.1.

Більшість таких відношень є бінарними. При заданні користувачем геопросторового об'єкту-еталону (точки або полігону), необхідно виділити серед об'єктів онтології такі, що містять геопросторову інформацію, яка, в свою чергу, зв'язана заданим бінарним відношенням з еталоном:

$$Q^r(X, g) = \{x \in X \mid A_{geo}(x), g > r\} \quad (2.100)$$

де r – певне бінарне геопросторове відношення;

X – множина об'єктів таксономії;

g – просторовий об'єкт, що виступає в якості еталону;

$A_{geo}(x)$ – функція вибірки зв'язаної з об'єктом x геопросторової інформації.

Відношення віддаленості може також використовуватись для фільтрації, якщо задати певне значення максимальної віддаленості d_{max} . Така функція є надзвичайно корисною в якості альтернативи відношенню «співпадає» для випадків, коли і еталон, і множина геопросторової інформації, для якої здійснюється фільтрація, є точками (наприклад, при обробці події «клік мишкою по маркеру»). Дану функцію можна представити наступним чином:

$$Q^d(X, g) = \{x \in X \mid d(A_{geo}(x), g) < d_{max}\} \quad (2.101)$$

де d – функція віддаленості;

X – множина об'єктів таксономії;

g – просторовий об'єкт, що виступає в якості еталону;

$A_{geo}(x)$ – функція вибірки зв'язаної з об'єктом x геопросторової інформації;

d_{max} – значення максимально допустимої віддаленості.

Вимірювальні операції A^{ms} принципово відрізняються від операцій аналізу просторових відношень об'єктів. Їх результатом є певне числове значення, для представлення якого користувачу потрібна спеціальна функція відображення D_n . Класифікація таких операцій подана в Таблиця 2.2.

Основною відмінністю обчислювальних операцій є те, що вони приймають в якості аргументу не еталон для фільтрації, а конкретний об'єкт, над зв'язаною з яким геопросторовою інформацією необхідно виконати обчислення. Однак при необхідності на їх основі можна побудувати функції фільтрації, зокрема, аналогічні до (2.101).

Накладемо на усі елементи множини онтологій C наступну умови – у кожного з них є непуста підмножина об'єктів, що містять географічну інформацію:

$$\forall O \in C, Q_a(X_O, A_{geo}) \neq \emptyset \quad (2.102)$$

Де X_O – множина об'єктів онтології O ;

A_{geo} – множина географічних атрибутів;

Q_a – функція атрибутивної фільтрації (2.82)

Довільну ГІС можна представити у наступному вигляді [71]:

$$\langle X_g, R_g, A^{op}, T^s \rangle \quad (2.103)$$

де X_g – множина географічних сутностей, над якими виконуються аналітичні операції для вирішення задачі;

R_g – множина відношень між географічними об'єктами, які визначають тип виконуваних операцій;

A^{op} – множина аналітичних операцій над географічними об'єктами, виконуваних в процесі вирішення задачі;

T^s – множина станів задачі, які візуалізуються на карті в процесі її вирішення;

Таблиця 2.2. Вимірювальні операції

Визначення координат точки на карті	визначення координат точки перетину двох прямих		
	визначення координат центроїдів і центрів		
Вимірювання відстаней	довжина прямої/кривої між двома точками		
	найкоротшої відстані за найкоротшою прямою з врахуванням/без врахування сферичності земної поверхні		
	найкоротшої відстані від заданої точки до полігону та між полігонами		
	функціональної відстані за заданим маршрутом з використанням точок повороту/з врахуванням нерівностей рельєфу («по землі»)		
	відстані та витрат по мережі/по поверхні		
Вимірювання полігонів	периметрів		
	площі		
	форм	просторова цілісність	
		співвідношення великої й малої осей	
		відношення периметра до площі	
		розвиненість меж	
Вимірювання об'ємів			

Твердження 2.3. Кожну ГІС можна представити у вигляді певної натуральної системи.

Дане твердження впливає з визначення натуральної системи [52], оскільки ГІС може бути представлена як система з n впливів і одним «результатом» (2.76), в

якості якого виступає представлення користувачу геопросторової інформації в потрібній формі.

Твердження 2.4. Сукупність усіх ГІС можна представити у вигляді гіпермножини натуральних систем.

Функціональні характеристики ГІС можуть бути проінтерпретовані на основі афінного простору. Афінним простором над полем K називається трійка:

$$\langle A_n, V_n, + \rangle \quad (2.104)$$

де A_n – множина, елементи якої називаються точками;

V_n – векторний простір над полем K ;

$+$ – бінарна операція $A_n \times V_n \rightarrow A_n$, що задовольняє певним аксіомам [6].

Множину точок A_n називають n -вимірним афінним простором над K , а векторний простір V_n – напрямним для A_n [6, 26].

Такий простір можна легко побудувати в рамках певного онтологічного ГІС-додатку. Для цього необхідно використати функцію вибірки зв'язаної з об'єктом геопросторової інформації $A_{geo}(x)$:

$$A_{geo}^x = \{A_{geo}(x) \mid x \in X\} \quad (2.105)$$

Сформована таким чином множина точок $A_n = A_{geo}^x$ і дозволяє сформувати афінний простір над полем $\mathbb{R}^2$ або $\mathbb{R}^3$ (в залежності від структури інформації в вихідному онтологічному ГІС-додатку і вимог задачі, що розв'язується). Дана операція здійснюється спеціалізованим перетворенням:

$$\tilde{G}_A : \langle X, R, F \rangle \rightarrow \langle A_{geo}^x, V_n, + \rangle \quad (2.106)$$

Де X, R, F – множини об'єктів, зв'язків та функцій інтерпретації онтології;

A_{geo}^x – сформована на основі X з допомогою (2.105) множина географічних точок;

$V_n, +$ – векторний простір над полем $\mathbb{R}^2$ або $\mathbb{R}^3$ і певна бінарна операція.

Точки, що належать множині географічної інформації A_{geo}^x , будуть пов'язані з векторами, що належать V_n , наступним чином:

- 1) Кожній впорядкованій парі точок A, B можна поставити у відповідність вектор $\overrightarrow{AB} = v \in V_n$, що можна позначити як $v = B - A$.
- 2) Для даної точки A і вектора $v \in V_n$ $\exists! B \in A_n : B - A = v$.
- 3) Для довільних трьох точок A, B, C виконується в V_n тотожність трикутника:

$$(B - A) + (A - C) + (C - B) = 0 \quad (2.107)$$

Можна виділити властивості:

- 1) Двом однаковим точкам відповідає нульовий вектор.
- 2) Вектор, що визначається парою точок B, A протилежний вектору, що визначається парою точок A, B .
- 3) Якщо $B - A = B - A'$ то $A = A'$.
- 4) Якщо виберемо якусь початкову точку O , то $\forall v \in V_n$ відповідає тільки одна точка $B \in A_n$, а $\forall B \in A_n$ відповідає тільки один вектор $v = B - O$.

Розглянемо площину, як афінний простір A_2 , і розглянемо всі взаємно однозначні відображення $A_2 \rightarrow A_2$, що зберігають відстані між точками і кути між векторами. Класифікація основних таких відображень показана в Таблиця 2.3

Кожному з основних видів відображень відповідає певна функція, що виконує таке відображення над певним афінним простором $\check{A}_n$, утвореним певною підмножиною об'єктів онтології $\check{X} \subset X$. Такі функції мають структуру (2.108).

$$Q_t(O) = \check{G}_A^{-1}(t(\check{G}_A(O))) \quad (2.108)$$

Де $\check{G}_A, \check{G}_A^{-1}$ – перетворення (2.106) та обернене до нього;

t – певне афінне відображення $A_n \rightarrow A_n$.

Функції Q_t, Q_{rt}, Q_{in} та інші, задані спеціалізованими видами відображень, можуть використовуватись в комбінації зі стандартними функціями інтерактивного документу для побудови онтологічних ГІС-додатків. Онтологічні ГІС-додатки, в

свою чергу, забезпечують трансдисциплінарну інтеграцію геопросторової інформації засобами інтерактивних документів та відображень афінного простору.

Таблиця 2.3. Класифікація основних відображень афінного простору

Група	Позначення	Примітка	Ф-я інт. документу
Група всіх невідроджених лінійних відображень $V_n \rightarrow V_n$	$GL_n(R)$, загальна лінійна група	Задає всі взаємно однозначні перетворення вект. прост. V_n	
Група лінійного відображення $V_n \rightarrow V_n$, що мають матриці з визначником 1.	$SL_n(R)$, спеціальна лінійна група	Взаємно однозначні перетворення V_n , що зберігають довжини і об'єми.	
Група трансляцій	$T_n(R)$	Паралельні переноси простору A_n	Q_t
Ортогональна група O_n	$O_n(R)$	Взаємно однозначні перетворення дійсного простору $V_n \rightarrow V_n$, що зберігають кути і відстані.	
Група поворотів	$SO_n(R)$, спеціальна ортогональна група.	Зберігає кути, відстані, а також орієнтацію систем векторів.	Q_{rt}
Множина інверсій		Не утворює групу, оскільки добуток двох інверсій матиме матрицю з визначником 1, а отже, зображує деякий поворот.	Q_{in}

Висновки за розділом 2

На основі аналізу первинної структури тексту, що формується в ході його лексичного аналізу, визначено можливість і запропоновано метод структуризації природномовних текстів на основі процедури рекурсивної редукції тексту з допомогою правил, представлених у формі лямбда-виразів. Показано можливість динамічного формування таких правил користувачами без спеціальної підготовки на основі первинної структури вже проаналізованих текстів. Розроблений метод дозволяє формувати структуризоване представлення текстів, і надавати таким чином можливість обробки наявної в тексті інформації автоматичними і автоматизованими системами.

Розроблено онтологічну модель інтерактивного документу, призначену для відображення експерту результатів структуризації тексту найбільш оптимальним чином для виконання поставленої задачі. Розроблено модель онтологічного ГІС-

додатку як варіант інтерактивного документу, натуральна система якого визначена над афінним простором і завдяки цьому придатна для відображення геопросторової інформації найбільш природним чином. Моделі інтерактивного документу та онтологічного ГІС-додатку забезпечують високий рівень репрезентативності наявної в текстових документах інформації (зокрема, геопросторової) за рахунок використання структуризованого представлення тексту.

Розроблено модель трансдисциплінарного представлення інформації як функції інтерактивного документу, що надає можливість отримання оперативного доступу до великих масивів тематичної інформації, а в комбінації з можливостями онтологічних ГІС-додатків – вирішує задачу трансдисциплінарного представлення геопросторової інформації.

Результати досліджень, наведені в другому розділі, опубліковано в роботах [14, 16, 19, 29, 31, 32, 72, 75, 78–81].

.

РОЗДІЛ 3. АРХІТЕКТУРА СИСТЕМИ ТРАНСДИСЦИПЛІНАРНОГО ПРЕДСТАВЛЕННЯ ГЕОПРОСТОРОВОЇ ІНФОРМАЦІЇ

Програмна система трансдисциплінарного представлення геопросторової інформації є системою, призначеною для обробки великих масивів слабо і неструктурованих документів, що можуть бути представлені у вигляді природномовних текстів, написаних українською мовою [19, 75, 81]. Система виконує виділення з них інформації за заданими користувачем правилами, її представлення у вигляді онтологічної структури [78, 79], а в подальшому – трансдисциплінарне представлення. Для документів, що містять геопросторову інформацію, також створюється представлення у вигляді онтологічного ГІС-додатку [72, 77, 80].

Система трансдисциплінарного представлення геопросторової інформації призначена для реалізації множини компонентів інформаційної технології структуризації текстів [19, 75, 81], а також розширення множини компонентів інформаційної технології представлення інформації, що реалізується системою ТОДОС.

З точки зору програмної інженерії програмна система розглядається у вигляді набору описів, представлених у вигляді математичних моделей, формалізмів і технік моделювання [103, 104].

Структура математичних моделей ПС такого роду включає в себе наступні моделі [103, 104]:

- 1) інформаційна модель;
- 2) функціонально-компонентна модель.

3.1. Інформаційна модель програмної системи трансдисциплінарного представлення геопросторової інформації

Інформаційна модель використовується для представлення і опису потоків інформації, структур даних а також програмних модулів в програмній системі.

Узагальнена інформаційна модель програмної системи трансдисциплінарного представлення геопросторової інформації Π_R має вигляд (3.1). Вона

представляється деякою скінченною сукупністю програмних модулів Π_{R_i} , що інтегруються в інформаційно-аналітичну систему ТОДОС Π_T :

$$\Pi_R = \sum_{i=1}^n \Pi_{R_i} \cup \Pi_T \quad (3.1)$$

При цьому реалізується відображення G_{Π_R} інтеграції функцій окремих програмних модулів системи, що має вигляд (3.2). Дане відображення перетворює об'єднання множини функцій S_{R_i} кожного з її програмних модулів Π_{R_i} в узагальнену (цільову) функцію F_R – трансдисциплінарне представлення геопросторової інформації.

$$G_{\Pi_R} : \bigcup_{i=1}^n S_{R_i} \rightarrow F_R \quad (3.2)$$

Система ТОДОС, в свою чергу, також має велику кількість багатфункціональних модулів різного призначення. В роботі системи трансдисциплінарного представлення геопросторової інформації використовується тільки сукупність $\Pi_T^R \subset \Pi_T$ модулів, що релевантні задачі формування шарів ГС. Таким чином формули (3.1) і (3.2) перетворюються в (3.3) і (3.4).

$$\Pi_R = \sum_{i=1}^n \Pi_{R_i} \cup \sum_{i=1}^m \Pi_{T_i}^R \quad (3.3)$$

$$G_{\Pi_R} : \bigcup_{i=1}^n S_{R_i} \cup \bigcup_{i=1}^m S_{T_i}^R \rightarrow F_R \quad (3.4)$$

3.1.1. Інформаційна модель підсистеми структуризації тексту

Сукупність модулів системи трансдисциплінарного представлення геопросторової інформації, що не входять до складу базової структури ТОДОС, можна розглядати як окрему підсистему – підсистему структуризації

тексту $\Pi_{TX} = \sum_{i=1}^n \Pi_{R_i}$. Дана сукупність має вигляд (3.5).

$$\Pi_{TX} = \{\Pi_{ПА}, \Pi_{ЛА}, \Pi_{ШШ}, \Pi_{ЗШШ}, \Pi_{П}, \Pi_{ВГР}, \Pi_{ФТД}, \Pi_{РШ}\} \quad (3.5)$$

Розглянемо детальніше функції різних модулів підсистеми виділення географічної інформації з тексту.

$\Pi_{ПА}$ – модуль виконання попереднього аналізу вхідного документу. Даний модуль необхідний для приведення документу до текстового формату. Множина функцій даного модуля має вигляд (3.6).

$$S_{ПА} = \langle S_1^{ПА}, S_2^{ПА}, S_3^{ПА}, S_4^{ПА}, S_5^{ПА} \rangle \quad (3.6)$$

До його функцій відносяться:

- 1) $S_1^{ПА}$ – зчитування вхідного файлу одного з підтримуваних форматів (текст, документи Word, електронні таблиці Excel, файли HTML та ін.);
 - 2) $S_2^{ПА}$ – отримання текстової інформації, що знаходиться в файлі;
 - 3) $S_3^{ПА}$ – нормалізація кодування отриманої текстової інформації;
 - 4) $S_4^{ПА}$ – виділення і збереження оригінальної розмітки вхідного документу;
 - 5) $S_5^{ПА}$ – збереження вкладеної в документ мультимедійної інформації (картинок, відео, аудіо та ін.), якщо тип вхідного документу підтримує це;
- Вихідними даними для даного модулю є текстовий файл.

$\Pi_{ЛА}$ – модуль виконання лексичного аналізу (3.7).

$$S_{ЛА} = \langle S_1^{ЛА}, S_2^{ЛА}, S_3^{ЛА}, S_4^{ЛА}, S_5^{ЛА}, S_6^{ЛА}, S_7^{ЛА}, S_8^{ЛА} \rangle \quad (3.7)$$

Функції модулю:

- 1) $S_1^{ЛА}$ – зчитування вхідного текстового файлу;
- 2) $S_2^{ЛА}$ – розбиття вхідного тексту на речення, а речень – на лексеми (слова і символи);
- 3) $S_3^{ЛА}$ – нормалізація розбиття у відповідності з правилами мови (наприклад, видалення розривів речень, викликаних крапками в складі скорочень і чисел);
- 4) $S_4^{ЛА}$ – нормалізація слів;

- 5) S_5^{LA} – визначення частини мови для слів і деяких символів (таких, як тире, що замінює певне слово);
- 6) S_6^{LA} – визначення додаткових характеристик лексем;
- 7) S_7^{LA} – визначення синтаксичних зв'язків між лексемами;
- 8) S_8^{LA} – формування структур даних первинного представлення тексту.

Π_{III} – модуль інтерпретації шаблонів (3.8).

$$S_{III} = \langle S_1^{III}, S_2^{III}, S_3^{III}, S_4^{III} \rangle \quad (3.8)$$

Функції модулю:

- 1) S_1^{III} – зчитування файлів шаблонів;
- 2) S_2^{III} – формування внутрішнього представлення для предикатів, з яких формуються шаблони;
- 3) S_3^{III} – формування внутрішнього представлення для шаблонів правил;
- 4) S_4^{III} – формування внутрішнього представлення для правил ідентифікації інформації;

Π_{3III} – модуль застосування шаблонів (3.9).

$$S_{3III} = \langle S_1^{3III}, S_2^{3III}, S_3^{3III}, S_4^{3III} \rangle \quad (3.9)$$

- 1) S_1^{3III} – вибірка підпоследовностей лексем з вхідного тексту;
- 2) S_2^{3III} – вибірка шаблонів-кандидатів з наявної множини шаблонів;
- 3) S_3^{3III} – застосування шаблонів-кандидатів до вибраних підпоследовностей лексем;
- 4) S_4^{3III} – вибір шаблону, що найкраще відповідає вхідній підпоследовності;

Π_{II} – модуль інтерпретації інформації.

$$S_{II} = \langle S_1^{II}, S_2^{II}, S_3^{II}, S_4^{II} \rangle \quad (3.10)$$

- 1) S_1^{II} – вибір застосовної функції інтерпретації інформації;
- 2) S_2^{II} – встановлення відповідності між елементами вхідної підпоследовності, що містить інформацію, і аргументами вибраної функції інтерпретації;
- 3) S_3^{II} – застосування функції інтерпретації;
- 4) S_4^{II} – визначення належності отриманої інформації до географічної.

Π_{BGI} – модуль валідації географічної інформації (3.11).

$$S_{BGI} = \langle S_1^{BGI}, S_2^{BGI}, S_3^{BGI}, S_4^{BGI} \rangle \quad (3.11)$$

Функції модуля:

- 1) S_1^{BGI} – визначення коректності всіх частин географічної інформації (наприклад, в географічних координатах значення градусів не повинно перевищувати 180);
- 2) S_2^{BGI} – нормалізація координат, при необхідності – переведення їх в систему координат WGS-84 [138, 145];
- 3) S_3^{BGI} – геокодування текстових адрес з допомогою основної або резервних систем геокодування [134, 148];
- 4) S_4^{BGI} – визначення належності результату дозволених географічній області;

$\Pi_{\Phi TD}$ – модуль формування онтологій.

$$S_{\Phi TD} = \langle S_1^{\Phi TD}, S_2^{\Phi TD}, S_3^{\Phi TD}, S_4^{\Phi TD} \rangle \quad (3.12)$$

Функції модуля:

- 1) $S_1^{\Phi TD}$ – нормалізація строкового представлення об'єктів, описаних в тексті;
- 2) $S_2^{\Phi TD}$ – формування ієрархії об'єктів згідно з наданою текстом інформацією;

- 3) $S_3^{\Phi TD}$ – формування атрибутів об’єктів, зокрема тих, що містять географічну інформацію;
- 4) $S_4^{\Phi TD}$ – запис результуючої таксономії у файл формату XML [1, 121, 124].

$\Pi_{PШ}$ – модуль редактора шаблонів. Даний модуль являє собою користувацький інтерфейс системи. Він містить в собі набір візуальних редакторів для різноманітних структур даних, що використовуються системою Множина функцій даного модуля має вигляд (3.13).

$$S_{PШ} = \langle S_1^{PШ}, S_2^{PШ}, S_3^{PШ}, S_4^{PШ}, S_5^{PШ}, S_6^{PШ}, S_7^{PШ}, S_8^{PШ} \rangle \quad (3.13)$$

Функції модуля:

- 1) $S_1^{PШ}$ – створення і редагування предикатів ідентифікації інформації;
- 2) $S_2^{PШ}$ – створення і редагування правил ідентифікації інформації;
- 3) $S_3^{PШ}$ – встановлення відповідності між правилами ідентифікації інформації і функціями інтерпретації;
- 4) $S_4^{PШ}$ – перегляд результатів лексичного аналізу тексту;
- 5) $S_5^{PШ}$ – автоматизоване створення правил і предикатів на основі результатів лексичного аналізу;
- 6) $S_6^{PШ}$ – перегляд результатів роботи системи;
- 7) $S_7^{PШ}$ – виправлення і уточнення результатів роботи системи;
- 8) $S_8^{PШ}$ – формування файлів шаблонів.

3.1.2. Інформаційна модель підмножини модулів ТОДОС, що використовуються в задачі трансдисциплінарного представлення геопросторової інформації

Більшість модулів ТОДОС тим чи іншим чином використовується в задачі трансдисциплінарного представлення геопросторової інформації. В складі ТОДОС

присутній багатофункціональний редактор з можливістю збереження інформації в різних форматах, і розширюваний набір конвертерів для перетворення таксономій в ті формати даних, що не підтримуються редактором. Також ТОДОС має розширювану програмну підсистему відображення онтологій, на основі якої можна побудувати модулі трансдисциплінарного представлення і формування онтологічних ГІС-додатків.

В загальному підмножину Π_T^R модулів ТОДОС, що використовуються в процесі формування ГІС, можна представити структурою (3.14).

$$\Pi_T^R = \{\Pi_{ГД}, \Pi_{ТБ}, \Pi_{ТК}, \Pi_{СХ}, \Pi_{РТ}, \Pi_{ТДП}, \Pi_{ДК}\} \quad (3.14)$$

Деякі модулі ТОДОС створені для виконання тільки одної певної функції, зокрема:

- 1) $\Pi_{ГД}$ – модуль, призначений для відображення онтології у вигляді онтологічного ГІС-додатку;
- 2) $\Pi_{ТБ}$ – модуль представлення списку об'єктів онтології у вигляді таблиці;
- 3) $\Pi_{ТК}$ – модуль відображення онтології у вигляді онтографу;

Інші модулі являються багатофункціональними підсистемами:

$\Pi_{СХ}$ – сховище таксономій, що являє собою серверний комплекс, призначений для зберігання онтологій і їх метаданих. Множина функцій модуля має вигляд (3.15)

$$S_{СХ} = \langle S_1^{СХ}, S_2^{СХ}, S_3^{СХ}, S_4^{СХ}, S_5^{СХ} \rangle \quad (3.15)$$

Функції модуля:

- 1) $S_1^{СХ}$ – зберігання онтологій, відкритих для публічного доступу;
- 2) $S_2^{СХ}$ – зберігання особистих онтологій користувача;
- 3) $S_3^{СХ}$ – контроль доступу до онтологій;
- 4) $S_4^{СХ}$ – забезпечення програмного інтерфейсу зчитування і запису онтологій;

5) S_5^{CX} – зберігання метаданих онтологій.

Π_{PT} – редактор таксономій загального призначення (3.16).

$$S_{PT} = \langle S_1^{PT}, S_2^{PT}, S_3^{PT}, S_4^{PT}, S_5^{PT}, S_6^{PT} \rangle \quad (3.16)$$

Функції модуля:

- 1) S_1^{PT} – зчитування і запис онтологій в форматі XML;
- 2) S_2^{PT} – додавання і видалення об'єктів;
- 3) S_3^{PT} – додавання і видалення зв'язків між об'єктами;
- 4) S_4^{PT} – редагування атрибутів об'єктів;
- 5) S_5^{PT} – автоматичне розміщення об'єктів в робочій області;
- 6) S_6^{PT} – фільтрація множини об'єктів, що відображаються.

Π_{TDP} – модуль трансдисциплінарного представлення.

$$S_{TDP} = \langle S_1^{TDP}, S_2^{TDP}, S_3^{TDP} \rangle \quad (3.17)$$

Функції модуля:

- 1) S_1^{TDP} – контекстна зв'язка об'єктів різних онтологій.
- 2) S_2^{TDP} – трансдисциплінарне представлення інформації з допомогою онтологічного куба.
- 3) S_3^{TDP} – керування процесом індексації інформаційних ресурсів.

Π_{DK} – блок допоміжних конвертерів. Даний блок містить програмний інтерфейс для створення підпрограм-конвертерів, а також базовий набір таких підпрограм. Множина функцій даного модуля має вигляд (3.18), де S_1^{DK} – забезпечення програмного інтерфейсу для підпрограм-конвертерів; S_{xiyi}^{DK} –

підпрограма-конвертер з формату x_i в формат y_i ; n_{DK} – загальна кількість наявних підпрограм-конвертерів.

$$S_{DK} = \langle S_1^{DK} \{ S_{x_i y_i}^{DK} | i \in [1; n_{DK}] \} \rangle \quad (3.18)$$

Основними наявними в даний момент конвертерами є:

- з CSV в XML;
- з XML в CSV;
- з HTML в текстовий файл;

А також спеціалізовані конвертери для забезпечення інтеграції з важливими програмами:

- з HTML, що генерується Microsoft Office, в текст;
- з XML в CSV форматі, що може бути зчитаний ArcGIS for Desktop.

3.2. Функціонально-компонентна модель програмної системи трансдисциплінарного представлення геопросторової інформації

Функціонально-компонентна модель використовується для представлення взаємодій, відношень і залежностей програмних модулів, а також для детального опису компонентів системи. Узагальнено дану модель для програмного комплексу трансдисциплінарного представлення геопросторової інформації можна представити структурою (3.19).

$$S_R = \langle M_D, M_S, M_P, M_C, P_0(M_D, M_S) \rangle \quad (3.19)$$

Елементи, що входять в дану модель:

- 1) M_D – модель, що задає поведінку системи;
- 2) M_S – модель, що задає структуру системи;
- 3) M_P – модель, що задає структуру програмних сутностей;
- 4) M_C – модель (схема) компонентів програмної системи;
- 5) $P_0(M_D, M_S)$ – предикат цілісності системи.

3.2.1. Модель поведінки системи

Модель поведінки системи має структуру (3.20).

$$M_D = \langle d_{use}, d_{act}, d_{seq} \rangle \quad (3.20)$$

Вона включає в себе:

- 1) d_{use} – множину UML-діаграм варіантів використання системи трансдисциплінарного представлення геопросторової інформації.
- 2) d_{act} – множину UML-діаграм активності системи трансдисциплінарного представлення геопросторової інформації.
- 3) d_{seq} – множину UML-діаграм взаємодії системи трансдисциплінарного представлення геопросторової інформації.

UML-діаграма варіантів використання [102, 132, 141] програмної системи трансдисциплінарного представлення геопросторової інформації представлена на Рис. 3.1.

Дана діаграма показує основі дії, що можуть здійснюватися в рамках системи, і користувачів, які можуть їх здійснювати.

Базовим користувачем системи є звичайний користувач, для якого доступне тільки представлення онтологій – а саме стандартні способи представлення ТОДОС (таблиця, онтограф) і спеціалізовані способи, що надаються модулями програмної системи трансдисциплінарного представлення геопросторової інформації (онтологічний ПС-додаток і трансдисциплінарний куб). В рамках процесу трансдисциплінарного представлення користувач також може здійснювати контекстну зв'язку інформаційних ресурсів.

Адміністратор ТОДОС виконує дві спеціалізовані функції – формування публічної бібліотеки онтологій (зокрема, збереження в ній створених експертом ПдО онтологій) і адміністрування підсистеми трансдисциплінарного представлення, що полягає у виборі множини інформаційних ресурсів, що можуть включатись в таке представлення.

Експерт ПдО виконує основну функцію – формування тематичних онтологій, що в подальшому можуть бути опубліковані адміністратором в публічній бібліотеці.

Формування може бути виконане вручну (з допомогою редактора онтологій), а може здійснюватися автоматично – на основі процесу структуризації тексту, який також здійснюється експертом ПдО.

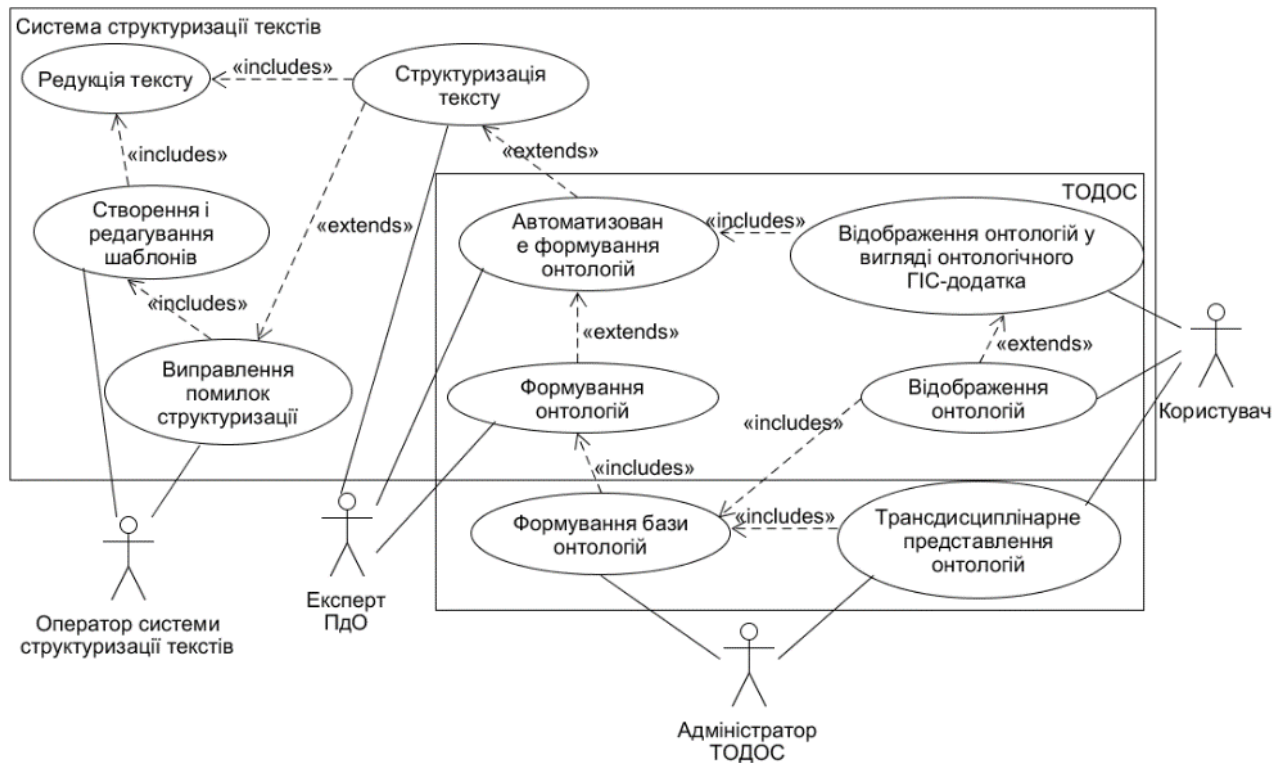


Рис. 3.1 Способи використання системи

Процес структуризації включає в себе редукцію тексту, що здійснюється під керуванням оператора системи структуризації. Для коректного виконання структуризації необхідно мати базу правил виконання редукції, що створюється на основі результатів редукції з пустою базою правил оператором системи структуризації. В подальшому, після завершення експертом ПдО структуризації, може бути перевірена якість виконання даного процесу, і у випадку виявлення помилок результати структуризації направляються оператору, який виправляє їх в рамках відповідного процесу.

UML-діаграма активності [102, 132, 141] програмної системи трансдисциплінарного представлення інформації показана на Рис. 3.2.

Діаграма описує алгоритм роботи з системою трансдисциплінарного представлення геопросторової інформації. Вхідною точкою для роботи з нею є веб-інтерфейс ТОДОС, що надає інтерфейси для здійснення всіх дій, а також керує

правами доступу до них. ТОДОС може працювати в режимі перегляду, або в режимі створення (який також забезпечує функцію редагування).

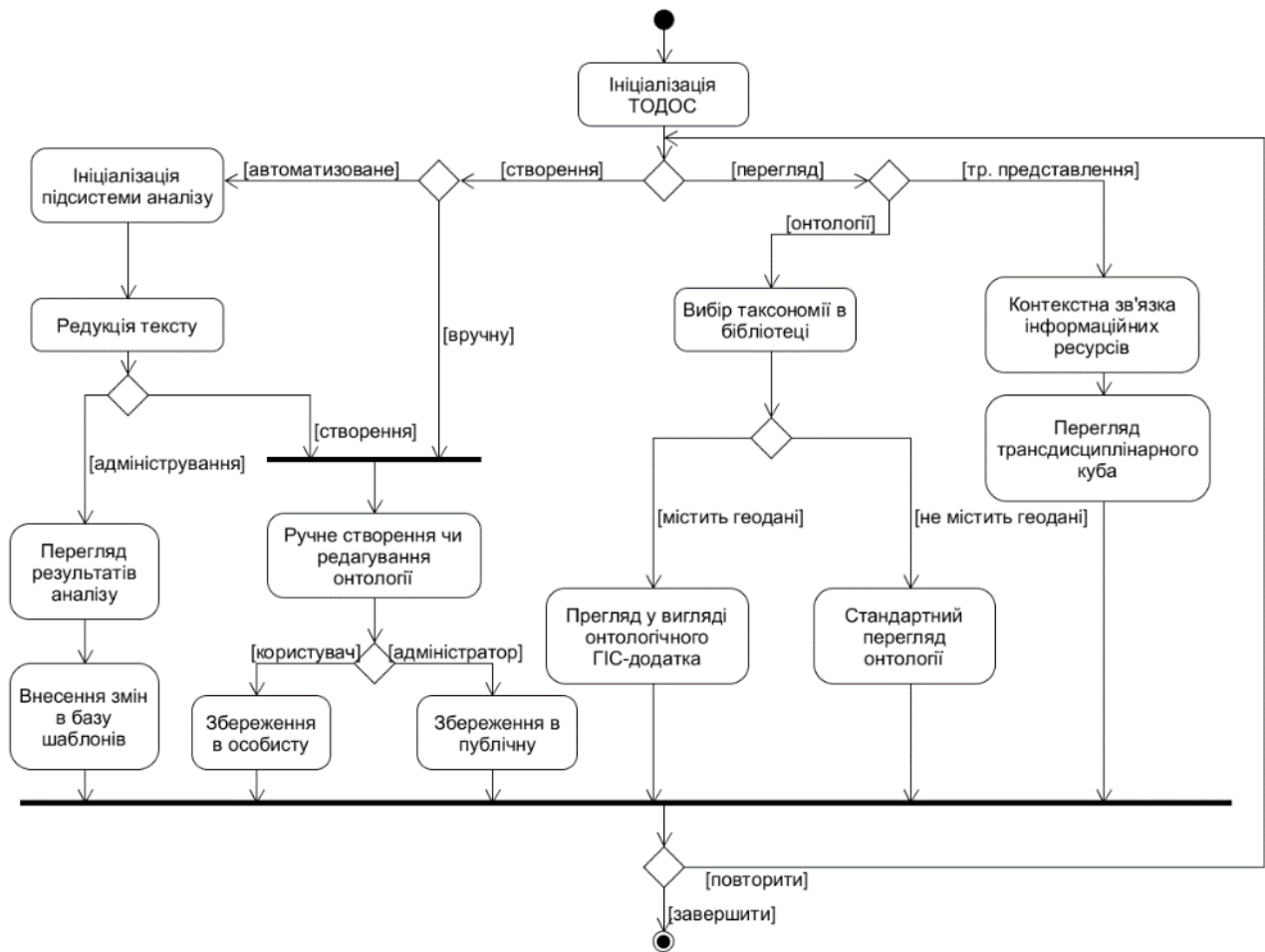


Рис. 3.2. Діаграма активності системи

При роботі в режимі створення можливі наступні варіанти:

- 1) Створення онтології вручну – стандартний режим, доступний для всіх користувачів. Після закінчення створення онтології користувач може:
 - зберегти результат своєї роботи локально. Результат доступний виключно користувачу;
 - зберегти результат своєї роботи в особисту бібліотеку. Користувач повинен бути авторизованим в системі. Результат доступний в веб-орієнтованому середовищі користувачу і людям, для яких він явно відкріє публічний доступ;
 - зберегти результат своєї роботи в публічну бібліотеку. Користувач повинен мати права адміністратора ТОДОС. Результат доступний всім (в тому числі і незареєстрованим користувачам) через публічну бібліотеку онтологій.

2) Автоматизоване створення онтологій – для виконання даного процесу необхідно надати системі вхідний текст, з який необхідно структурувати. Даний режим використовує підсистему аналізу текстів, тому користувач повинен мати доступ до її інтерфейсу. Залежно від вибору користувача можливі варіанти:

- власне створення онтології – в результаті редукції відбувається формування попереднього варіанту онтології, що завантажується в редактор таксономій. Подальші дії користувача такі ж, як і у випадку ручного створення;
- адміністрування модулю аналізу текстів. Замість редактора онтологій результати редукції завантажуються в спеціальний адміністративний модуль, що дозволяє керувати базою правил. Необхідний адміністративний доступ до модулю структуризації текстів.

При переході в режим перегляду ТОДОС надає користувачу доступ до публічної (а при наявності достатніх прав доступу – і до особистої) бібліотеки, в якій можна вибирати онтологію для перегляду. Перегляд можливий з допомогою кількох модулів відображення, окрема:

- Модуль відображення таксономії;
- Модуль табличного відображення;
- Модуль об'єктного відображення;
- Відображення у вигляді онтологічного ГІС-додатку (необхідні геопросторові дані).

Особливим режимом перегляду є перегляд трансдисциплінарного представлення. В даному режимі користувач послідовним застосуванням функції контекстної зв'язки формує таке представлення, яке в подальшому може бути відображене в якості трансдисциплінарного кубу.

UML-діаграма взаємодії програмної системи трансдисциплінарного представлення геопросторової інформації показана на Рис. 3.3.

Дана діаграма представляє собою основний процес системи трансдисциплінарного представлення геопросторової інформації – власне аналіз документів з формуванням в результаті онтології.

обробки тексту (хоча редактор таксономій все ще доступний завдяки його веб-орієнтованості). Після завершення обробки з її результатів формується таксономія, і користувачу надається можливість редагувати її з допомогою редактора таксономій.

Після закінчення редагування створеної автоматично або закінчення створення вручну таксономії вона зберігається в бібліотеку. Для цього виконується асинхронний запит до серверної інфраструктури ТОДОС. Серверна інфраструктура приймає таксономію, перевіряє її цілісність і у випадку успішної перевірки зберігає її в своїй базі даних. Після закінчення користувачу відправляється повідомлення про успішність або неуспішність збереження, яке приймається і відображується редактором таксономій.

Після закінчення обробки весь комплекс даних про оброблений файл (включаючи сам файл, його текстове представлення і результати аналізу) зберігаються системою аналізу текстів. На основі цих даних може бути проведене виправлення помилок обробки.

Виправлення помилок обробки відбувається з допомогою інтерфейсу редагування шаблонів, перехід на який стає можливим після завершення обробки з керуючого інтерфейсу. Інтерфейс редагування шаблонів також зв'язаний з своєю власною серверною інфраструктурою, що керує базою шаблонів. Після завершення роботи з редактором шаблонів він відправляє запит на цю інфраструктуру, передаючи через нього список внесених змін. Як і у випадку збереження таксономії, запит виконується асинхронно.

3.2.2. Модель структури системи

Модель структури системи має вигляд (3.21).

$$M_S = \langle d_{class}, b_{stat}, req \rangle \quad (3.21)$$

Вона включає в себе:

- 1) d_{class} – множину діаграм класів системи формування електронних шарів ГІС;
- 2) b_{stat} – множину статичних блок-схем системи формування електронних шарів ГІС;

- RuleReader зчитує файли шаблонів і виконує їх інтерпретацію, формуючи внутрішнє представлення зчитаних шаблонів;
- InputFileReader – виконує зчитування вхідного XML-файлу з результатами лексичного аналізу, або зчитування вхідного текстового файлу і його лексичний аналіз з допомогою відповідного модуля (залежно від налаштувань запуску).

Дані класи створюються класом Worker, який запускає підпрограми, що реалізуються ними, отримує результати роботи цих підпрограм і передає їх контролеру ідентифікації інформації.

Підпрограми обробки виконуються класами, що реалізують заданий класом AbstractResultProcessor інтерфейс (а також наслідують його методи). До них відносяться:

- 1) XMLHolderTransducer – створюється і зберігається контролером ідентифікації інформації. Клас містить набір об'єктів типу XMLNodeTransducer з різними конфігураціями, що виконують інтерпретацію того чи іншого блоку ідентифікованої інформації. Клас XMLHolderTransducer визначає, який саме з об'єктів буде використовуватись для обробки, на основі інформації про вхідний блок інформації і правило, що його ідентифікувало.
- 2) XMLNodeTransducer – клас, що виконує інтерпретацію ідентифікованої інформації тим чи іншим чином, в залежності від конфігурації.
- 3) ResultProcessor – клас, що перетворює інтерпретовану інформацію, приводячи його до вигляду, придатного для запису до вихідного файлу.

Аналогічна діаграма для редактора шаблонів показана на Рис. 3.5.

Структура класів редактора шаблонів в деякій мірі схожа на структуру аналізатора текстів, за виключенням того, що замість ієрархії контролерів вона містить ієрархію візуальних елементів.

Структура класів даних аналогічна: є два класи, RuleFileIO для зчитування файлів правил і SentencesFileIO для зчитування результатів аналізу тексту. Обидва класи наслідуються від FileIOController, який реалізує методи для відкриття і запису (зокрема, виклик системних форм вибору файлу).

Контролер у редактора шаблонів один, глобальний. Він викликає класи даних і запускає зчитування файлів, як вибраних користувачем, так і заданих налаштуваннями запуску. також контролер надає інтерфейс редагування результуючої структури даних, що в подальшому використовується візуальними елементами.

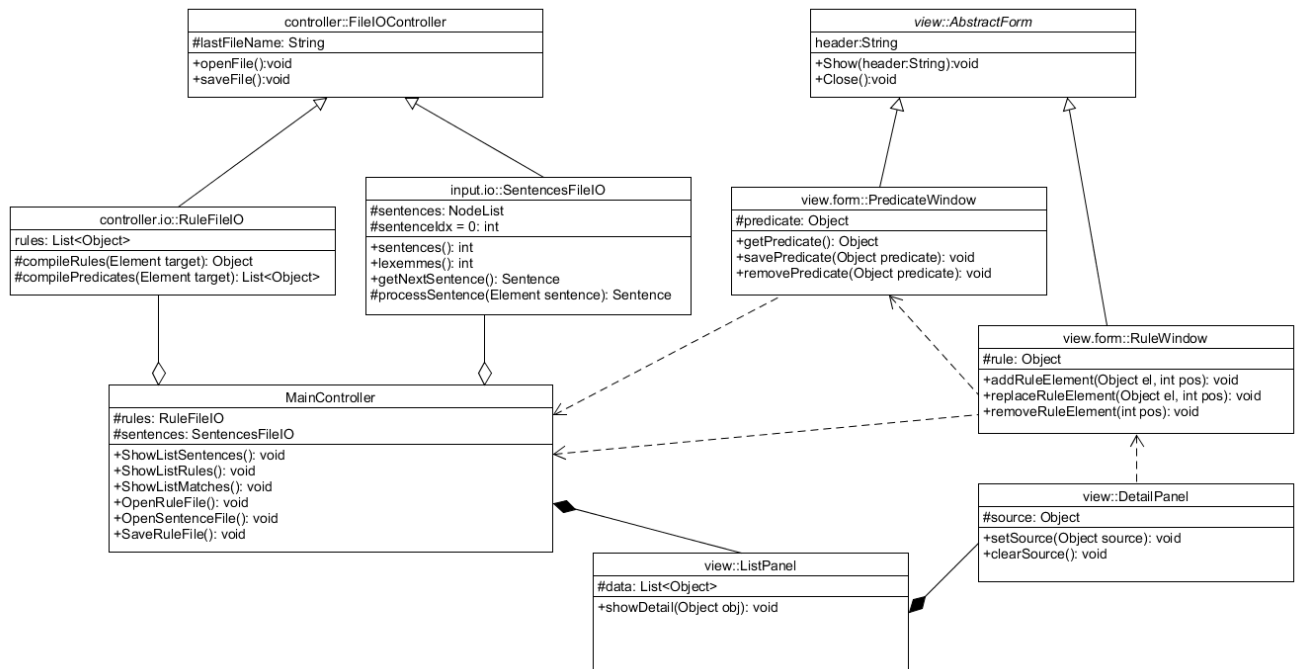


Рис. 3.5. Діаграма класів редактора шаблонів

Ієрархія візуальних елементів наступна:

- 1) Елементом найвищого рівня є панель ListPanel, що відображає список заданих елементів і дозволяє вибрати їх для детального перегляду;
- 2) Панель DetailPanel знаходиться на панелі ListPanel і приймає від неї інформацію про вибраний елемент. Отримана інформація в подальшому відображається на цій панелі.
- 3) Вікно RuleWindow призначене для створення або редагування правил і містить список елементів правила.
- 4) Вікно PredicateWindow призначене для створення чи редагування окремих елементів правила. Воно викликається в процесі редагування правила з допомогою RuleWindow.

Елементи-вікна наслідуються від абстрактного класу AbstractForm, що надає базові функції відображення форм, зокрема, задає правила позиціонування вікна.

Панель ListPanel може працювати в трьох режимах, що задаються контролером: режимі перегляду речень, режимі перегляду правил і режимі перегляду ідентифікованої інформації. DetailPanel може працювати в режимі перегляду речення (вмикається коли ListPanel в режимі перегляду речень або інформації) чи в режимі перегляду правил (вмикається коли ListPanel в режимі перегляді правил). В першому режимі відображаються наявні в реченні лексеми, в другому – наявні в правилі предикати.

Вікна PredicateWindow і RuleWindow використовують надані контролером MainController функції, вносячи зміни в загальну структуру бази шаблонів.

3.2.3. Модель структури програмних сутностей

Модель структури програмних сутностей розглядається як файли вихідного коду, бібліотеки, виконувані файли і відношення між ними. Дана модель описується з допомогою UML Package Structure і показана на Рис. 3.6.

Фізично, файли підсистеми аналізу текстів і ТОДОС розміщуються окремо. Кожна з систем має два блоки сутностей:

- веб-інтерфейс, що безпосередньо відображується користувачу і надає йому керуючі елементи для запуску виконання тих чи інших дій;
- серверна інфраструктура, що приймає запити від веб-інтерфейсу і виконує задані запитами дії;

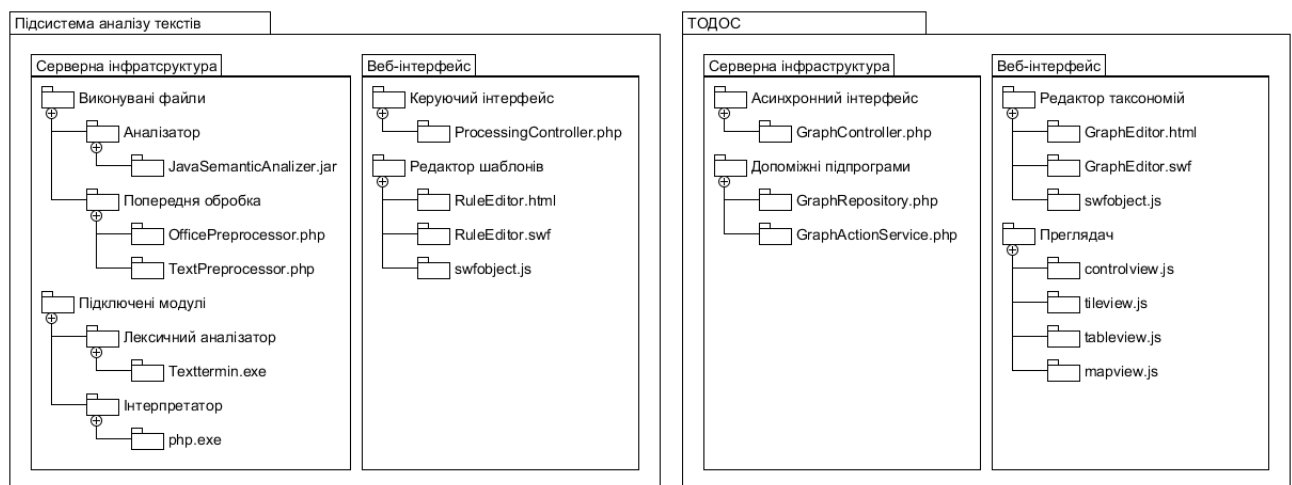


Рис. 3.6. Структура програмних сутностей системи формування шарів ГІС

Для підсистеми аналізу текстів веб-інтерфейс складається з двох частин – керуючого інтерфейсу (який також за приймає запити від користувача в синхронному режимі), і редактора шаблонів.

Серверна інфраструктура підсистеми аналізу текстів представлена власне аналізатором текстів, підключеним до нього модулем лексичного аналізу, а також блоком підпрограм попереднього аналізу і додатковим консольним інтерпретатором PHP для їх виконання.

Для ТОДОС веб-інтерфейс доступний в двох варіантах: редактор таксономій, що включає в себе модулі комунікації з серверною інфраструктурою, і переглядач онтологій, призначений для автономної роботи в веб-браузері користувача. Переглядач онтологій, зокрема, реалізує перегляд у вигляді електронної карти.

Серверна інфраструктура ТОДОС призначена для роботи в асинхронному режимі, і включає в себе асинхронну точку доступу і набір допоміжних підпрограм, що реалізують низькорівневі операції зберігання онтологій в базу та інші допоміжні функції, такі, як контроль доступу.

3.2.4. Схема компонентів програмної системи

Модель (схема) компонентів програмної системи реалізується з допомогою UML-діаграми компонентів і використовується для візуалізації високорівневої структури системи і поведінки служб (сервісів), що надаються і споживаються цими елементами (компонентами) через інтерфейси. Схема компонентів системи трансдисциплінарного представлення інформації показана на Рис. 3.7.

В контексті взаємодії всі компоненти системи поділені на дві частини: веб-орієнтований інтерфейс і серверна інфраструктура. В даному контексті не має значення належність чи неналежність того чи іншого компонента базовій структурі ТОДОС, оскільки весь програмний комплекс будується як інтегрована система.

Однак, як компоненти ТОДОС, так і компоненти підсистеми аналізу текстів функціонують за схожою схемою:

- 1) Існує сховище (онтологій або шаблонів відповідно), що використовується для зберігання даних, якими оперує підсистема.

- 2) Існує точка доступу (асинхронна для компонентів ТОДОС, синхронна для компонентів підсистеми аналізу текстів).
- 3) Існує набір інтерфейсних модулів, що взаємодіють з точкою доступу.

Основні відмінності між схемами роботи підсистем:

- В підсистеми аналізу текстів є окремий серверний модуль-аналізатор для виконання обробки, тоді як в модулях ТОДОС вся обробка відбувається в інтерфейсних компонентах;
- Точка доступу підсистеми аналізу текстів взаємодіє не тільки з модулями даної підсистеми, а також з модулями ТОДОС: вона зберігає результати у вигляді таксономії в базі онтологій, також викликає редактор таксономій для редагування цих даних.

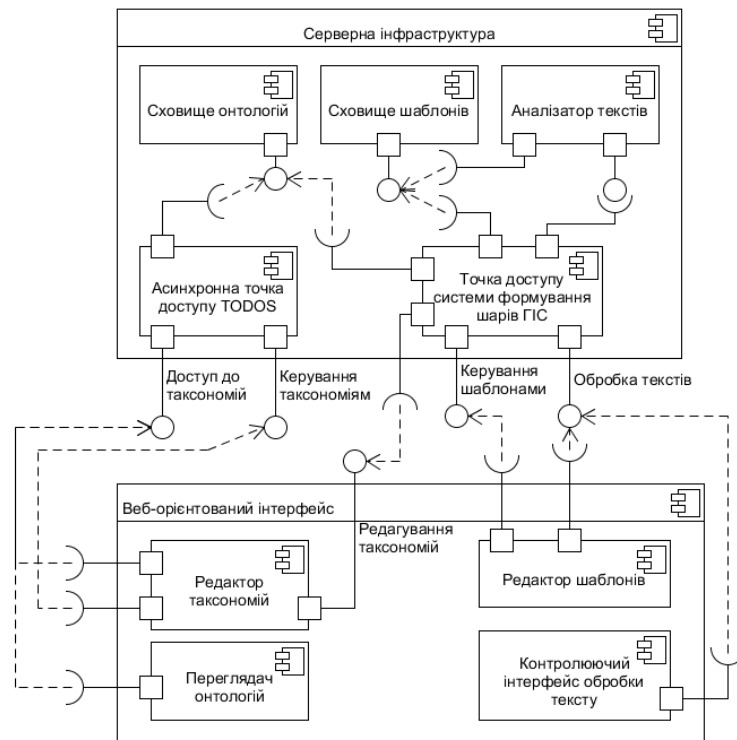


Рис. 3.7. Схема компонентів системи

Два з наведених в схемі компонентів – аналізатор текстів і редактор шаблонів – є складними підсистемами з внутрішньою структурою компонентів. Крім того дані підсистеми є відокремленими від інших, і виконані в формі окремих виконуваних модулів, що можуть взаємодіяти з іншими модулями тільки з допомогою спеціалізованих інтерфейсів, таких, як HTTP-точка доступу чи командна консоль операційної системи. Розглянемо їх компонентні структури детальніше.

Компонентна структура редактора правил показана на Рис. 3.8.

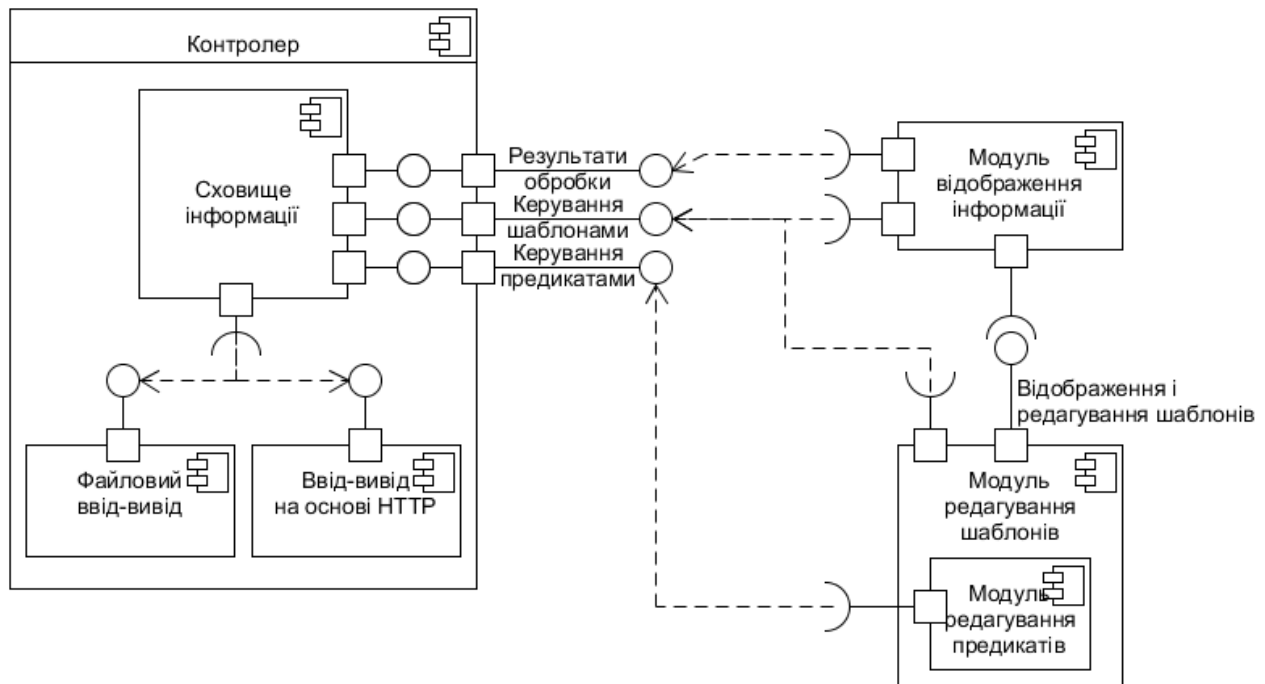


Рис. 3.8. Компонентна структура редактора шаблонів

Основний компонент редактора – сховище інформації. Воно є складовою частиною контролера. До складу підсистеми контролера входять також модулі вводу/виводу, які реалізують дану функцію двома різними способами – з допомогою бібліотек доступу до файлової системи і бібліотек здійснення HTTP-запитів. Який саме контролер вводу/виводу буде використовуватись, залежить від налаштувань запуску і дій користувача в ході роботи з редактором. Можливе комбіноване використання – наприклад, попереднє зчитування через веб-запити з подальшим збереженням з допомогою файлової системи.

Підсистема контролера інкапсулює сховище даних і надає іншим компонентам інтерфейс доступу до нього. Основні функції, доступ до яких надається, зв'язані з типами інформації, що зберігається сховищем. таких типів є три:

- 1) Предикати ідентифікації інформації;
- 2) Шаблони ідентифікації інформації;
- 3) Тестова інформація про результати обробки тексту.

Для предикатів і шаблонів надається інтерфейс, що дозволяє зчитування і модифікацію даних, тоді як інтерфейс доступу до результатів обробки дозволяє тільки зчитування.

Основний інтерфейсний модуль (модуль відображення інформації) являє собою набір панелей, що відображають об'єкти або списки об'єктів, що подаються на вхід. В залежності від типу списку, що поданий на вхід, даний інтерфейс може працювати в режимі відображення речень, в режимі відображення правил або режимі відображення ідентифікованої інформації.

Модуль відображення інформації використовує надані контролером інтерфейси, що дозволяють зчитування шаблонів і результатів аналізу тексту. Також в даному модулі є керуючі елементи, що дозволяють видаляти правила. Підпрограми що запускаються даними елементами, використовують інтерфейс контролера, що дозволяє керування правилами. Також модуль відображення інформації має можливість запускати інший модуль, модуль редагування інформації, з заданими параметрами.

Модуль редагування інформації запускається модулем відображення інформації, в при запуску отримує в якості параметру правило, що має бути відредаговане, або набір лексем вхідного тексту, з яких необхідно автоматично згенерувати новий шаблон.

При запуску модуль створює набір субмодулів (модулів редагування предикатів). Кожен з таких субмодулів виконує завдання редагування або створення одного з елементів шаблону. Після закінчення редагування або створення шаблону кожен з субмодулів виконує підпрограму, що перевіряє наявність поточного предиката в сховищі інформації, і у випадку його відсутності – створює новий. Дана підпрограма використовує можливості контролера по зчитуванню і запису предикатів. Після закінчення всіх підпрограм субмодулів модуль редагування шаблонів запускає свою підпрограму, яка генерує нове правило і додає його в сховище інформації, або замінює ним існуюче правило.

3.3. Архітектура системи трансдисциплінарного представлення геопросторової інформації

В процесі розробки і експлуатації програмної системи трансдисциплінарного представлення геопросторової інформації може виникати ряд проблем, зв'язаних з масштабуванням, швидкодією і надійністю. Для вирішення вищевказаних проблем в програмній інженерії використовується трирівнева архітектура [105, 108, 110, 114]. Трирівнева архітектура складається з рівня представлення (Presentation tier), рівня логіки (Logic tier), рівня даних (Data tier) і передбачає наявність клієнтської програми, сервера додатків і сервера бази даних. Однак в реальній ситуації, при обмеженості наявних обчислювальних ресурсів, розгортання такої архітектури може бути зв'язане зі значними труднощами, що можуть виражатись в недостатній потужності обчислювальних машин або наявності в модулів системи специфічних системних вимог, що можуть конфліктувати між собою.

Архітектура програмної системи трансдисциплінарного представлення геопросторової інформації базується на описаних в попередніх розділах моделей і являє собою набір незалежних програмних засобів. Деякі компоненти програмної системи створюються на базі веб-орієнтованих API системи ТОДОС і можуть функціонувати виключно в її складі, однак більшість компонентів реалізовані у вигляді окремих модулів, що можуть виконуватись і в складі ТОДОС, і на окремих, виділених серверах, а в особливих випадках – навіть локально на робочій машині користувача.

Базова і найбільш проста архітектура, що може використовуватись в рамках задачі трансдисциплінарного представлення геопросторової інформації, представлена на Рис. 3.9. Така архітектура містить виключно модулі підсистеми структуризації текстів, і тому не може використовуватись для представлення інформації. Однак, дана архітектура є дуже важливою, оскільки може виконуватись користувачем локально, і дозволяє виконувати найбільш ресурсоємкі процеси, не витрачаючи ресурсів серверу. Це природним чином забезпечує паралельну обробку масивів інформаційних ресурсів, оскільки надає можливість обробляти різні інформаційні ресурси на різних машинах різними операторами.



Рис. 3.9. Базова архітектура підсистеми структуризації текстів

Основних модулів в підсистемі структуризації текстів є три – редактор шаблонів, що реалізує собою інтерфейс користувача, лексичний аналізатор, що виконує попередній аналіз тексту, і модуль виконання обробки, що власне виконує виділення інформації.

Однак для повноцінного виконання задачі трансдисциплінарного представлення геопросторової інформації необхідно виконувати обробку документів в середовищі ТОДОС. Загальна архітектура системи в цьому середовищі матиме вигляд, показаний на Рис. 3.10.

Основні відмінності даної архітектури від зображеної на Рис. 3.9. наступні:

- наявна підсистема попередньої структуризації, що дозволяє розширити множину форматів файлів, з що можуть оброблятись. Зокрема, з її допомогою можуть бути зчитані документи Microsoft Office. Дана система спрощує і пришвидшує процес аналізу документів, оскільки позбавляє оператора необхідності конвертувати їх вручну;
- використовується редактор онтологій ТОДОС, що дозволяє виконувати перегляд результатів структуризації і при необхідності – вносити в них виправлення;

- використовується веб-орієнтований інтерфейс представлення інформації, що містить API, на основі яких реалізуються підсистеми трансдисциплінарного представлення інформації та онтологічних ГІС-додатків; Для підсистеми трансдисциплінарного представлення інформації доступний модуль перегляду у вигляді трансдисциплінарного куба, для підсистеми онтологічних ГІС-додатків – модулі інтеграції з основними ГІС-бібліотеками, зокрема, ArcGIS API for JavaScript [100], Leaflet.js [130], Google Maps [118];
- доступна серверна інфраструктура ТОДОС, зокрема, бібліотека онтологій, що використовується для зберігання результатів структуризації, а модифікація даної бібліотеки використовується для зберігання правил структуризації;
- доступний керуючий інтерфейс ТОДОС, в який інтегрується керуючий інтерфейс підсистеми структуризації текстів. Засобами ТОДОС реалізується контроль доступу до інтерфейсів;

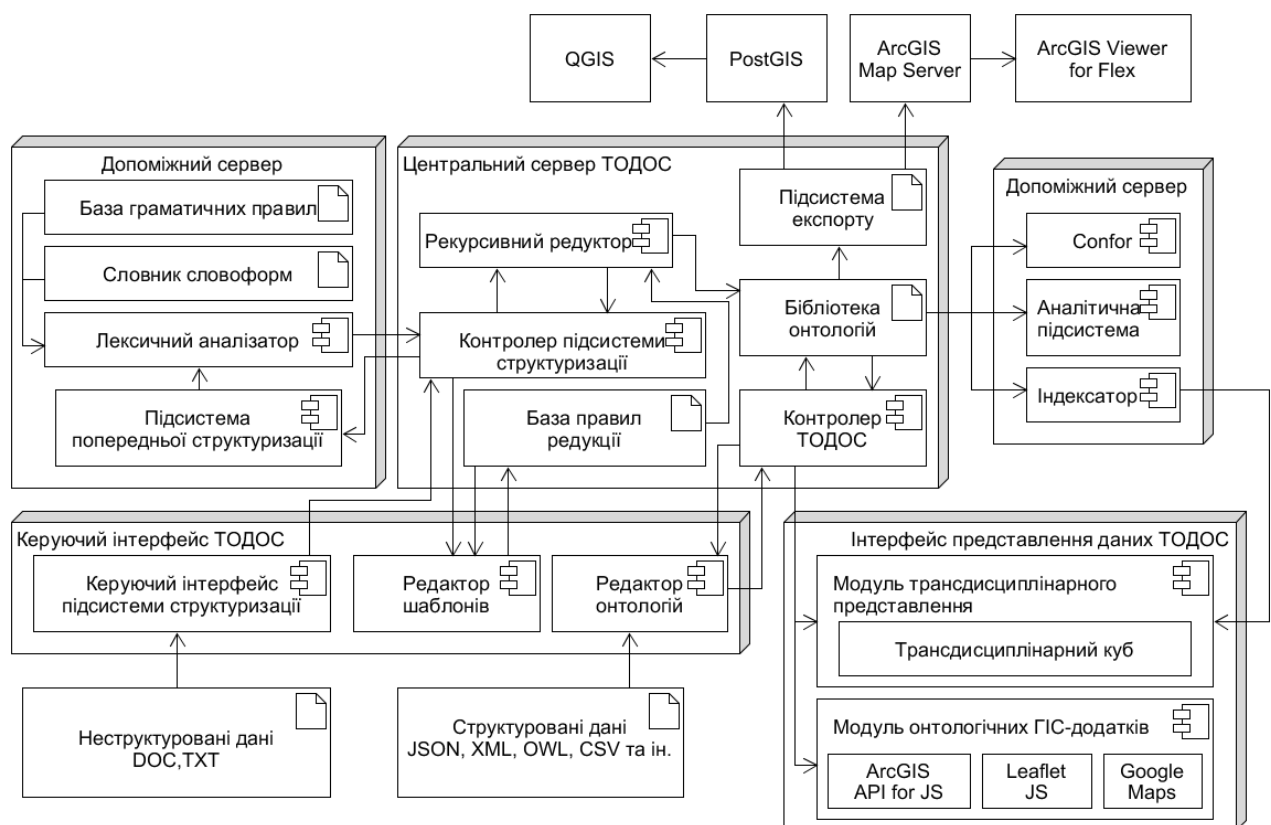


Рис. 3.10. Архітектура системи трансдисциплінарного представлення геопросторової інформації в середовищі ТОДОС

- доступна підсистема експорту онтологій, що може бути використана для формування ГІС-додатків з допомогою різноманітних ГІС [92], зокрема, QGIS і продуктів компанії ESRI [112].
- доступні аналітичні підсистеми ТОДОС, що дозволяють виконувати над ідентифікованими даними різноманітні операції, серед яких прогнозування, багатокритеріальна оптимізація та ін.
- можливе зчитування широкого спектру документів, що містять структуровану інформацію, зокрема з файлів CSV, OWL, KML [126] та ін.

Необхідно зазначити, що вказана структура системи відповідає трирівневій архітектурі. При цьому функції рівень представлення виконують інтерфейси ТОДОС, функції рівня даних – бібліотека онтологій і база правил редукції, а всі інші модулі виконують функції рівня логіки.

Архітектура, представлена на Рис. 3.10, задає також показує можливий варіант міжсерверної взаємодії. Як видно з рисунку, деякі підсистеми, зокрема, лексичний аналізатор, аналітичні модулі та індексатор, можуть розміщуватись на допоміжних серверах. дані підсистеми є найбільш ресурсоємкими, і тому винесення їх на окремі сервери дозволяє значно зменшити навантаження на основний сервер, і таким чином збільшити надійність його роботи.

Висновки за розділом 3

Розроблено інформаційну модель програмної системи трансдисциплінарного представлення геопросторової інформації, що розділяється на модель підсистеми структуризації тексту і модель інтегрованої в ТОДОС підсистеми представлення інформації.

Розроблено функціонально-компонентну модель програмної системи трансдисциплінарного представлення геопросторової інформації, що включає в себе модель поведінки системи, модель структури системи, модель структури програмних сутностей і схему компонентів програмної системи.

На основі інформаційної та функціонально-компонентної моделей програмної системи трансдисциплінарного представлення геопросторової інформації

розроблено архітектуру даної системи, що визначає структуру, склад і особливості функціонування розроблених програмних засобів.

Результати досліджень, наведені в третьому розділі, опубліковано в роботах [14–17, 19, 72, 76–78, 81].

РОЗДІЛ 4. ІНСТРУМЕНТАЛЬНІ ЗАСОБИ ФОРМУВАННЯ ОНТОЛОГІЧНИХ ГІС-ДОДАТКІВ

Процеси трансдисциплінарного представлення геопросторової інформації реалізуються на основі наступних мережових інструментальних засобів:

- «Рекурсивний редуктор» – інструментальний засіб, призначений для структуризації текстів і формування їх онтологічного представлення;
- ТОДОС-ГІС – інструментальний засіб, призначений відображення користувачу ідентифікованої в тексті інформації, що, зокрема, виконує її трансдисциплінарне представлення і представлення у виді онтологічного ГІС-додатку.

Дані інструментальні засоби є практичними реалізаціями наведених в попередніх розділах моделей і методів.

4.1. Інструментальний засіб «Рекурсивний редуктор»

Інструментальний засіб «Рекурсивний редуктор» [80] складається з модулю виконання рекурсивної редукції та редактора шаблонів.

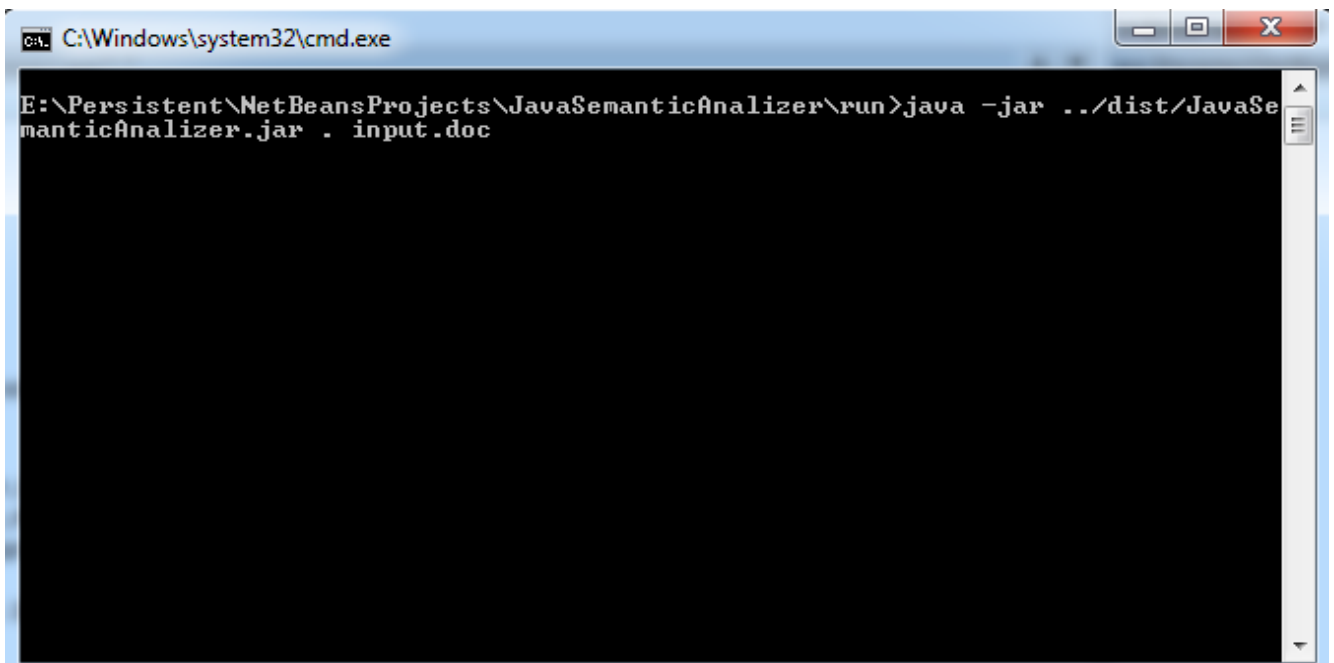


Рис. 4.1. Вікно запуску модуля редукції

Модуль рекурсивної редукції являється консольною програмою, що може бути запусненою з командної строки з допомогою Java Runtime Environment. Вікно

запуску матиме вигляд, показаний на Рис. 4.1. Програма приймає два аргументи: робочу директорію (на прикладі це поточна директорія ”.”), і назву вхідного файлу (input.doc).

Допустимі типи вхідних файлів:

- 1) TXT – завжди доступний.
- 2) HTML – доступний при наявності в поточній конфігурації модулів попередньої структуризації ТОДОС. Такий файл можна згенерувати, зберігши в якості HTML-файлу документ Microsoft Word (Рис. 4.2).
- 3) DOC/DOCX – необхідний встановлений на комп’ютері Microsoft Word.

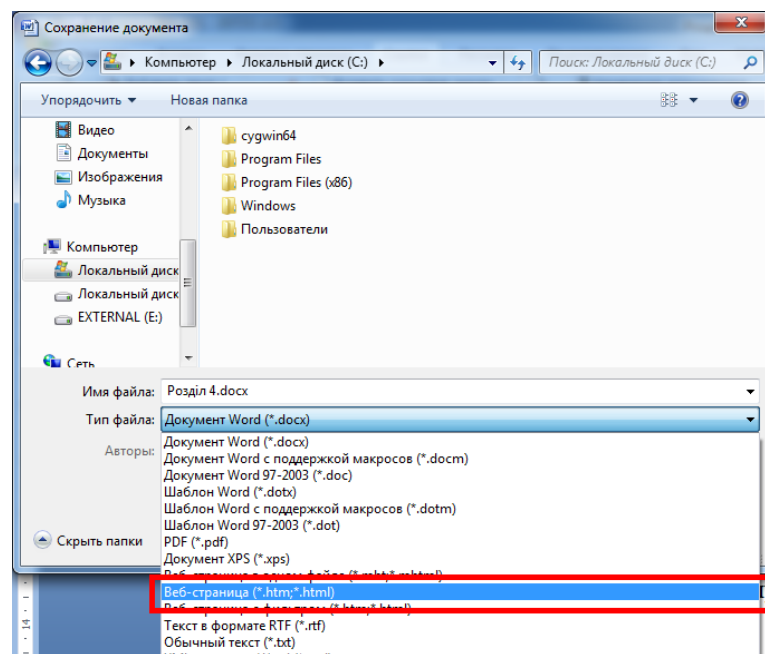


Рис. 4.2. Збереження документу Word в форматі HTML

Після закінчення роботи редуктора в папці, що була вказана як робоча, буде створено файл graph.xml, що містить результати роботи у вигляді онтології ТОДОС. Дана онтологія може бути зчитана засобами ТОДОС.

Також буде створений файл agsuccess.csv, що містить додаткову інформацію про роботу редуктора, яка може бути зчитана редактором шаблонів.

Редактор шаблонів являє собою реалізований з допомогою Adobe Flash модуль, що може бути запущений будь-яким веб-браузером з підтримкою відповідного розширення. Загальний вигляд основного вікна показано на Рис. 4.3.

В основного вікна є наступні робочі області:

- 1) Зверху розташована контрольна панель, з кнопками збереження бази правил (тільки при інтеграції з ТОДОС), а також для її локального збереження і зчитування.
- 2) Під контрольною панеллю розташований перемикач вкладок, що дозволяє переключатись між списками речень, виділеної інформації і правил.
- 3) Зліва розташований список поточних елементів, заданий перемикачем вкладок.
- 4) В центрі розташована робоча область, що відображає елемент, вибраний в списку зліва.
- 5) Знизу розташована панель керування відображенням, що відображає зв'язані з поточним елементом і дозволяє показувати їх в робочій області.

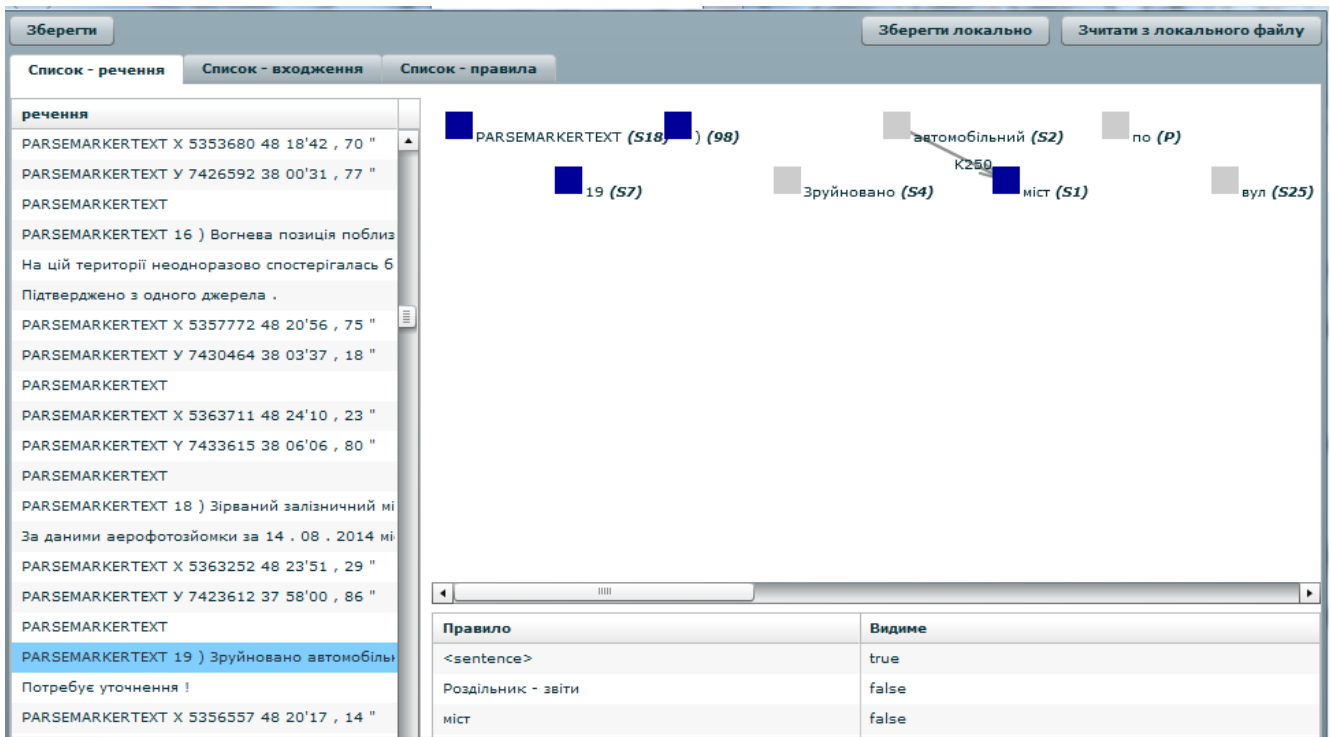


Рис. 4.3. Головне вікно редактора шаблонів

Режими перегляду працюють наступним чином:

Режим перегляду речень (Рис. 4.4) відображує в списку кожне з речень тексту. При виборі речення в робочій області відображаються лексеми, що належать даному реченню. Зв'язаними з реченням об'єктами є правила, що були застосовані до наявних в реченні підмножин лексем.

Режим перегляду виділеної інформації (Рис. 4.5) призначений для відображення всіх речень, що містять певний концепт або конструкт, а також правил, з допомогою яких він був виділений.

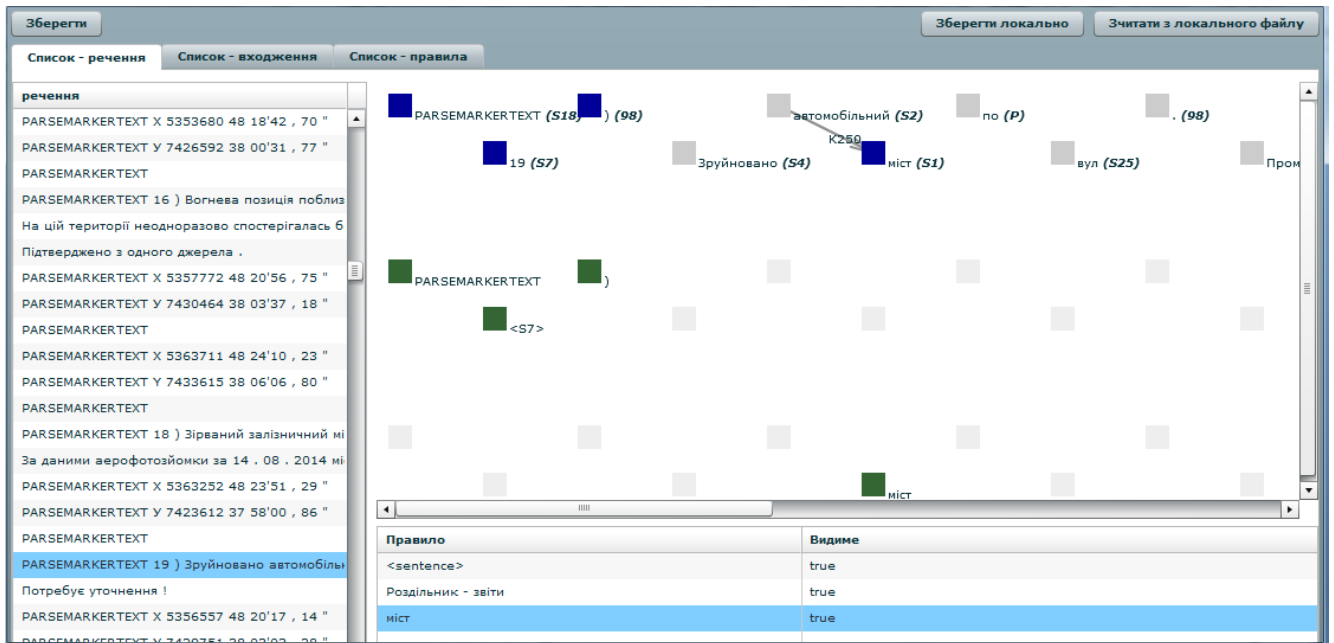


Рис. 4.4. Режим перегляду речень

В даному режимі всі виділені елементи даних згруповані за їх текстовим представленням. Текстові представлення відображаються в списку справа. При виборі елемента в цьому списку всі входження будуть відображені в робочій області у вигляді списку лексем, причому ті лексеми, що увійшли в склад потрібного елемента даних, виділені іншим кольором. Також входження відображаються в списку знизу, і можуть бути приховані кліком мишки на відповідний елемент цього списку.

Режимі перегляду виділеної інформації відрізняється від інших тим, що він показує не тільки сам лексеми, але і предикати, що ідентифікували дані лексеми. Відображується дана інформація через роздільник «—>» – зліва від нього показана назва вихідної лексеми, а справа – текстове представлення шаблону відповідного предикату.

Режим перегляду правил (Рис. 4.6) відображує список шаблонів правил, наявний в системі. Зв'язаними з правилами об'єктами є елементи даних, що були виділені даним правилом.

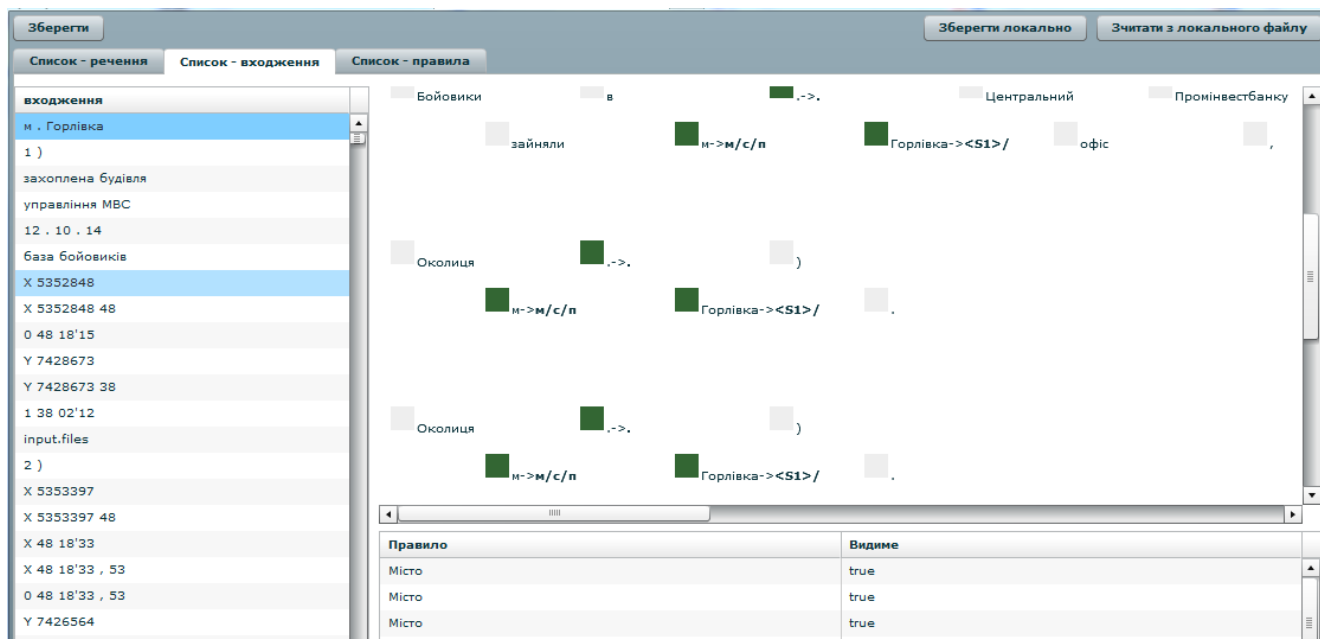


Рис. 4.5. Режим перегляду виділеної інформації

Для будь-якого з інтерфейсів доступні наступні комбінації клавіш:

- 1) Shift+клік викликає редагування правила.
- 2) Ctrl+клік видаляє правило.

Комбінації працюють тільки тоді, якщо елемент, по якому здійснюється клік, є правилом. Клік по реченню або ідентифікованому елементі даних не виконає ніяких дій.

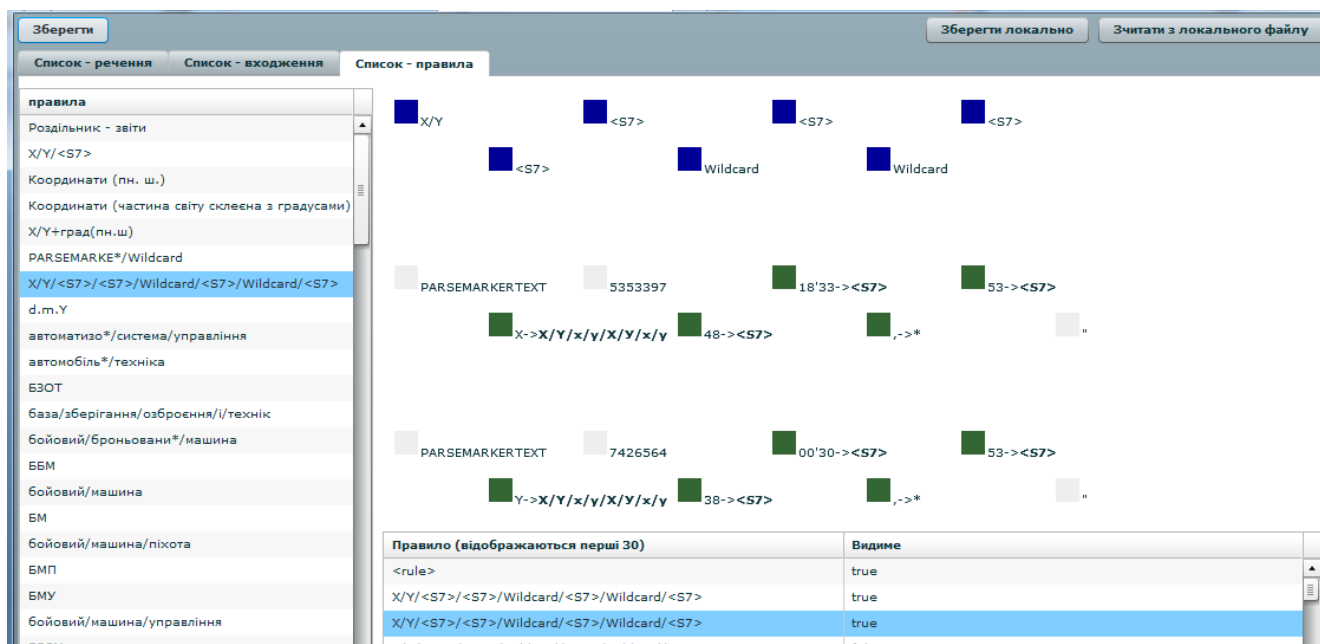


Рис. 4.6. Режим перегляду правил

Крім того, в режимі перегляду речень доступна додаткова функція – вибір лексем для автоматичної генерації нового шаблону. Вона працює наступним чином:

- 1) Необхідно натиснути і утримувати клавішу Shift.
- 2) Необхідно клікнути в робочій області на всі лексеми поточного речення, що повинні увійти до нового шаблону, що помітить їх як виділені.
- 3) Потрібно відпустити Shift.

Дана функція працює виключно при кліку на елементи вихідного речення, що розміщується зверху (помічене в нижньому списку як <sentence>). Її використання з допомогою наведених вище дій відкрис вікно створення шаблону (Рис. 4.7).

№	Входження	Предикат	Варіанти співставлення	Пропуск	Порівняння	Заміна	Інтерпретація	Del
1	зруйнувати	<S4>	#<S4>	<S4>	0	По-замовч.	Немає	По-замовчуванню Del
2	автомобільний	<S2>	#<S2>	<S2>	0	По-замовч.	Немає	По-замовчуванню Del
3	міст	<S1>	#<S1>	<S1>	0	По-замовч.	Немає	По-замовчуванню Del
4	по	<P>	#<P>	<P>	0	По-замовч.	Немає	По-замовчуванню Del
5	вул	<S25>	#<S25>	<S25>	0	По-замовч.	Немає	По-замовчуванню Del
6	.	<98>	#<98>	<98>	0	По-замовч.	Немає	По-замовчуванню Del
7	промисловий	<S2>	#<S2>	<S2>	0	По-замовч.	Немає	По-замовчуванню Del

Клас елемента даних:

Якість правил:

☐ Створити категорію ☐ Створити вершину

Рис. 4.7. Вікно створення шаблону

Основним елементом даного вікна є власне список шаблонів предикатів, що містить наступні елементи:

- 1) «Входження» – кнопки швидкого створення предикату. Предикат може бути створений на основі текстового представлення певної лексеми (нормалізованого), або на основі його граматичного класу (частини мови). Натиск на одну з кнопок, що відображують відповідну інформацію, сформує такий предикаті автоматично згенерує його назву.
- 2) «Предикат» – назва предикату або вибір існуючого. Кожен із створених предикатів зберігається в базі і при необхідності може бути використаний

повторно. Доступний комбінований елемент управління, що дозволяє або вибрати з випадуючого списку існуючий предикат, або ввести ім'я нового.

- 3) «Варіанти співставлення» – шаблон поточного предиката. Дане поле доступне тільки при створенні нового предиката, якщо вибрати в полі «Предикат» існуючий предикат, то поле «Варіанти співставлення» стане недоступним для редагування і відображатиме шаблон вибраного існуючого предиката.
- 4) Пропуск – набуває значень від нуля до п'яти. В залежності від даного значення на основі даного шаблону можна згенерувати кілька правил. Якщо значення дорівнює нулю, то генерується тільки стандартне правило. Якщо значення дорівнює одиниці, то генерується стандартне правило, а також правило, в якому після поточного предиката стоїть нульовий предикат, ідентифікована яким лексема відкидається функцією перетворення. Якщо значення дорівнює двійці, генерується три правила – стандартне, з одним нульовим предикатом і з двома нульовими предикатами, і т.д.
- 5) Порівняння – може набувати значень «пропуск» і «закінчення». Якщо поле має значення «пропуск», то на основі даного шаблону можна згенерувати два правила: звичайне і таке, в якому поточний предикат відсутній. Якщо поле має значення «Закінчення», то також можна згенерувати два правила – звичайне, і таке, в якому поточний предикат є останнім.
- 6) «Заміна» – комбінований керуючий елемент, що дозволяє визначити, на що буде замінене значення ідентифікованої предикатом лексеми в ході її інтерпретації. Основними значеннями є «відсутня», «вказане значення» (перетворює керуючий елемент на поле вводу, в якому і вводиться потрібне значення), «пусте значення» (означає, що лексема відкидається функцією інтерпретації). Також доступні інші значення, задані наявними в системі функціями інтерпретації.
- 7) Інтерпретація – показує, яка саме буде інтерпретована дана лексема функцією інтерпретації. Прикладами значень є «Адреса – місто», «Адреса – вулиця», «Дата – рік». Список варіантів інтерпретації задається наявними в системі функціями інтерпретації.

- 8) Кнопка «Del» – дозволяє видалити поточний шаблон предиката з шаблону правила.

Крім того, існує так званий блок метаданих, що також керує інтерпретацією виділеної правилом множини лексем. Доступні керуючі елементи:

- 1) «Клас елемента даних» – означає, що множина лексем повинна бути проінтерпретована як текстове значення певного атрибуту концепту.
- 2) Якість правила – метаінформація правила, що означає, наскільки точним є дане правило, і наскільки висока імовірність, що воно виділятиме некоректні елементи даних. Якість правила знижується, коли в нього явним чином включають нульові предикати. Правило, що повністю складається з нульових предикатів, має найнижчу точність і буде виділяти будь-які послідовності лексем заданої довжини.
- 3) «Створити вершину» – означає, що дана підмножина лексем задає назву концепту. У вихідній таксономії буде створена вершина з таким іменем.
- 4) «Створити категорію» – означає, що поточна множина лексем задає назву класу концептів. В результуючій таксономії буде створена вершина з таким іменем, і між нею і останнім ідентифікованим концептом буде створене відношення «клас-підклас».
- 5) Назва і опис правила – текстові поля в нижній і правій частині відповідно. Дозволяють задати назву правила, що буде відображатись в списку, а також опис правила, що можна буде переглядати при його редагуванні. Якщо не вказати цю інформацію, то назва буде створена автоматично, а опис буде пустим.

Якщо викликати редагування правила з допомогою комбінації Shift+клік, то відкриється вікно редагування правила (Рис. 4.8).

Основні відмінності від вікна створення:

- 1) Відсутній блок «Входження», оскільки вихідне речення, з якого було згенероване правило, не зберігається.
- 2) Всі поля «Предикат» містять вибраний існуючий шаблон предикату (той, який був створений при збереженні шаблону правила, або заданий при

формуванні), і тому всі поля «Варіанти співставлення» не доступні для редагування.

№	Предикат	Варіанти співставлення	Пропуск	Порівняння	Заміна	Інтерпретація	Del
1	PARSEMARKERTEXT	PARSEMARKERTEXT	0	По-замовч.	Пусте значення (= видалити)	По-замовчуванню	Del
2	<S7>	<S7>	0	По-замовч.	Немає	По-замовчуванню	Del
3	)	)	0	По-замовч.	Немає	По-замовчуванню	Del

Клас елемента даних:

Якість правила:

☐ Створити категорію ☒ Створити вершину

Додати строку на початок Додати строку в кінець

Ok Відмінити Роздільник - звіт

Рис. 4.8. Вікно редагування правила

Для керування всіма наявними модулями в режимі інтеграції з ТОДОС використовується керуючий інтерфейс (Рис. 4.9). Інтерфейс складається з двох блоків: блоку створення так званої «обробки» («processing»), і блоку списку існуючих обробок. Обробка – це блок інформації, що включає в себе певний вихідний текст, результуючої таксономії а також всіх проміжних файлів, що були створені в процесі аналізу.

Для створення обробки необхідно вибрати файл і, при необхідності, вказати її ім'я (якщо не вказати – ім'я буде вибране на основі імені файлу). Типи файлів, доступних для зчитування, залежать від складу поточної інсталяції ТОДОС. В найгіршій можливій конфігурації, при інсталяції ТОДОС під Linux і відсутності допоміжних серверів, можливе завантаження локальної інсталяції лексичного аналізатора (Texttermin), яка дозволяє проводити лексичний аналіз текстів на локальній машині, після чого сформованих XML-файл лексичного аналізу можна завантажити в систему.

Дата	Ім'я	Стан	Граф	Дії
03:27 17-05-2017	gorlivka	Завершено	gorlivka	Від'єднати Редуктор Уточнення правил Видалити Обробити заново

Рис. 4.9. Керуючий інтерфейс редуктора

Доступні керуючі елементи:

- 1) Керування лексичним аналізом – колонка «стан» показує стан виконання аналізу: в черзі, оброблено, помилка. Запуск аналізу можливий крайньою правою кнопкою «Запустити обробку»/ «Обробити заново».
- 2) Кнопка «Редуктор» запускає редукцію тексту, зберігає в базі ТОДОС сформовану таксономію і відкриває її. Якщо обробка вже проводилась, то існуюча таксономія замінюється, і всі внесені в неї зміни будуть втрачені.
- 3) Кнопка «Уточнення правил» запускає редактор шаблонів.
- 4) Колонка «Граф» дозволяє працювати напряму зі сформованою в результаті редукції таксономією. З допомогою кнопки, на якій зображено ім'я файл таксономії, її можна переглядати (на відміну від кнопки «Редуктор», таксономія не буде змінена»). Кнопка «Від'єднати» розриває зв'язок між обробкою і таксономією. Після наступного виконання редукції буде сформована нова таксономія, і подальші запуски редукції замінюватимуть її, тоді як початкова таксономія залишиться незмінною.

4.2. Інструментальний засіб ТОДОС-ГІС

Інструментальний засіб ТОДОС-ГІС [31, 32, 69, 80] призначений для обробки, управління та відображення різноманітної структурованої інформації, зокрема, структуризованих з допомогою рекурсивної редукції природномовних документів.

Базовим поняттям ТОДОС-ГІС є онтологічний кабінет – відокремлене інформаційне середовище, в якому користувач може розміщувати свої інформаційні

ресурси, працювати з ними, а при необхідності – створювати їх трансдисциплінарне представлення. Онтологічний кабінет визначає обмеження, що накладаються на користувача. Зокрема, робота з інтерфейсами виконання редукції можлива тільки в рамках онтологічного кабінету з відповідними дозволами.

Для отримання доступу до онтологічного кабінету необхідно пройти:

- 1) Процедуру реєстрації – надання персональних даних адміністратору для отримання облікового запису.
- 2) Процедуру авторизації – введення логіну та паролю для входу до кабінету під персональним обліковим записом в середовищі ТОДОС (Рис. 4.10).

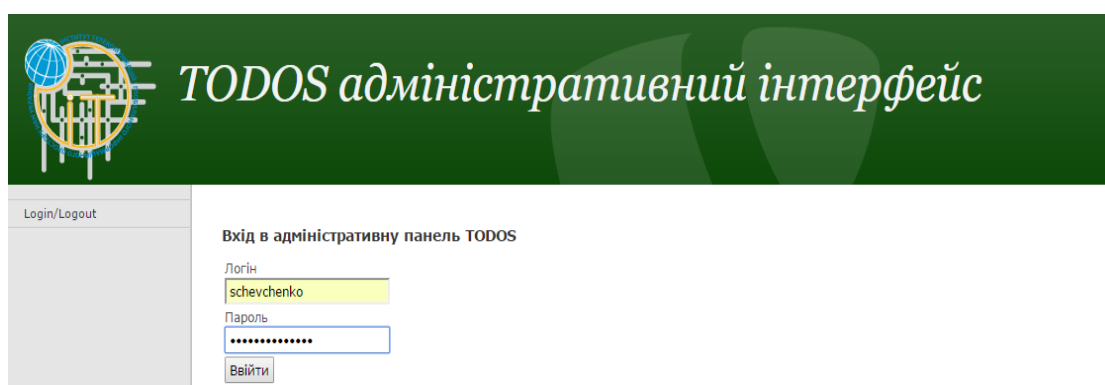


Рис. 4.10. Авторизація для доступу до онтологічного кабінету

Керування документами в особистому кабінеті здійснюється з допомогою інструментальних засобів, наданих ТОДОС, зокрема, підсистеми «TODOS Graph editor», що описана в Додатку Е.

Сам засіб ТОДОС-ГІС складається з трьох субмодулів:

- модулю формування інтерактивних документів;
- модулю трансдисциплінарного представлення інформації;
- модулю формування онтологічних ГІС-додатків.

Дані модулі призначені для представлення різними способами інформації, що може бути присутня в атрибутах об'єктів онтології. Таким чином користувачу надається можливість вибирати представлення, що найбільше відповідає поставленій задачі.

4.2.1. Модуль формування інтерактивних документів

Даний модуль може бути доступним з допомогою веб-орієнтованої точки входу (наприклад, ontology4.inhost.com.ua для серверу TODOS4) або після натиску кнопки «Об’єктна онтологія», яка відкриває точку входу поточного сервера в новому вікні браузера. Загальний вигляд модулю формування інтерактивних документів показано на Рис. 4.11.

Базовий режим відображення модулю – відображення у вигляді галереї зображень, що дозволяє досліджувати зв’язки між об’єктами певної онтології. Саме цей режим перегляду і зображено на Рис. 4.11.

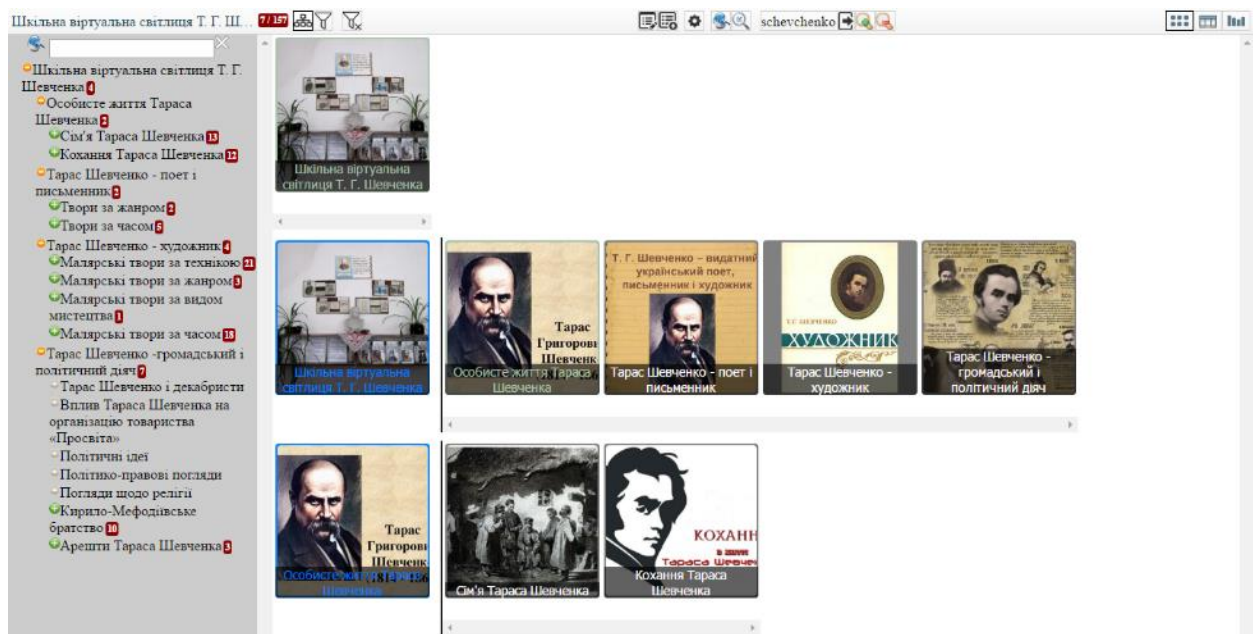


Рис. 4.11. Загальний вигляд переглядача онтологій в режимі галереї

Інтерфейс модулю складається з трьох частин: ліва – навігатор онтології, що відповідає таксономії онтології, верхня – функціональне меню, центральна – основна форма відображення об’єктів онтології.

Загальний вигляд навігатора показано на Рис. 4.12. Він містить дерево, що відповідає структурі онтології, а також область пошуку по назвам об’єктів.

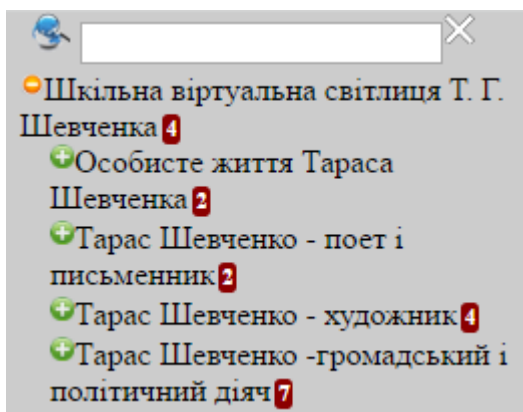


Рис. 4.12. «Меню» онтології, що відповідає структурі поточної онтології

В області пошуку здійснюється пошук за назвою об'єкта онтології (Рис. 4.13). В результаті пошуку об'єкти онтології, в назвах яких знайдено пошуковий запит, виділяються кольором, їх батьківські об'єкти розкриваються, а всі інші – приховуються. Якщо скоротити пошуковий запит до одного символу або натиснути на кнопку очистки справа від поля вводу пошукового запиту, пошук припиниться, всі об'єкти відображаються в початковому стані.

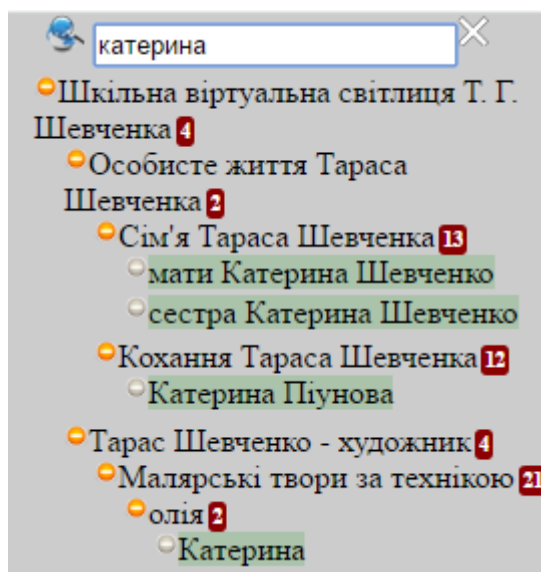


Рис. 4.13. Пошук за назвою об'єкта в навігаторі онтології

Якщо об'єкт онтології має залежні від нього (пов'язані певним відношенням), зліва від його імені розміщується перемикач «+ закрито / - розкрито», що дозволяє показувати/приховувати перелік залежних об'єктів, а справа знаходиться індикатор їх кількості.

Натиснення лівою кнопкою миші на назві об'єкту онтології робить його головним об'єктом в робочій області екрану (Рис. 4.14).

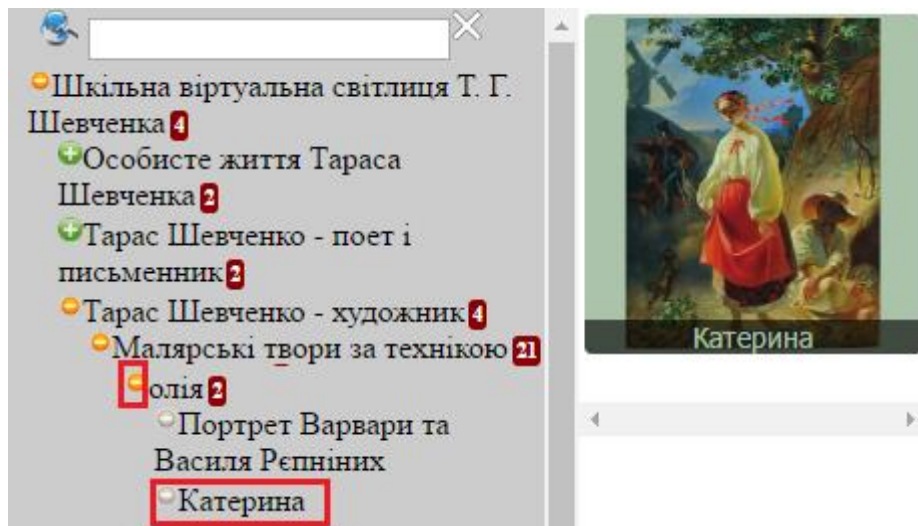


Рис. 4.14. Вибір вершини з допомогою навігатора онтології

Верхнє функціональне меню містить основні керуючі елементи, що використовуються в роботі з онтологією (Рис. 4.15).



Рис. 4.15. Верхнє функціональне меню переглядача онтологій ТОДОС

Поряд із назвою поточної онтології розміщений індикатор кількості видимих на даний момент об'єктів і загальну кількість об'єктів онтології.

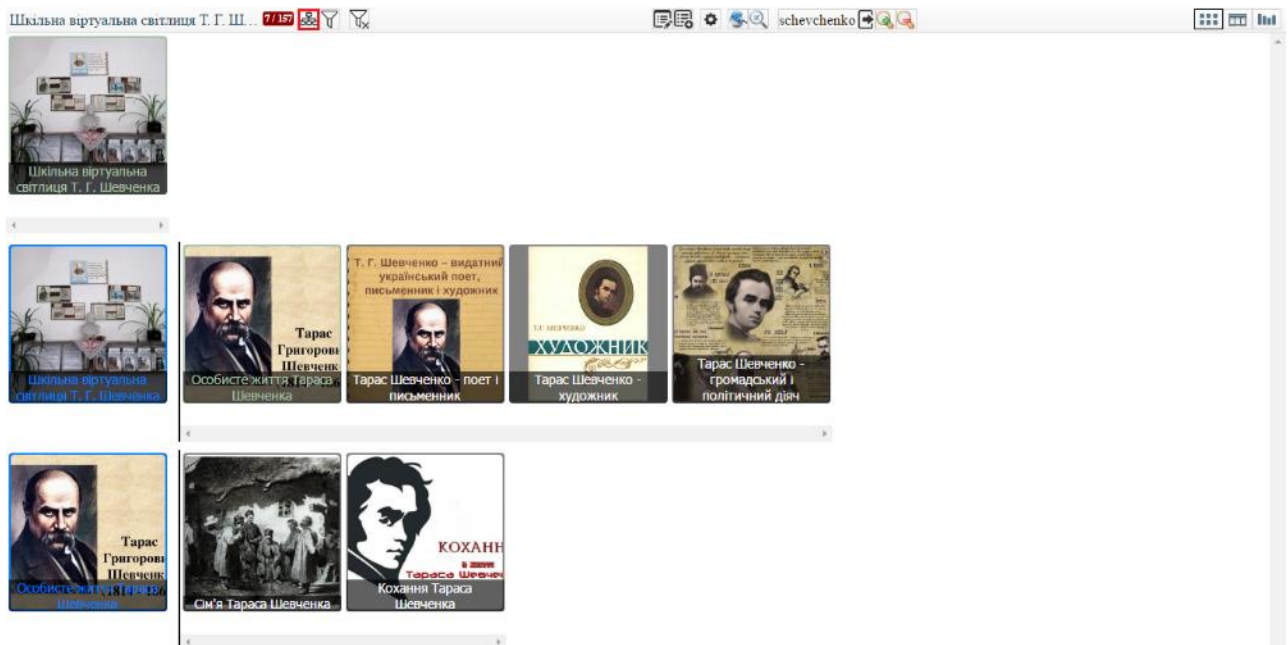


Рис. 4.16. Кнопка ієрархічного фільтра

Кнопка «Ієрархічний фільтр» (Рис. 4.16) дозволяє показати або приховати навігатор онтології. Для його показу необхідно натиснути на кнопку в момент, коли

фільтр не відображується. Повторне натискання приховає фільтр і розширить робочу область, як і показано на Рис. 4.16.

Кнопка «Фільтр атрибутів» (Рис. 4.17) дозволяє показати або приховати інтерфейс фільтрації об'єктів за значеннями їх атрибутів. Загальний вигляд фільтру показано на Рис. 4.17 і Рис. 4.18.

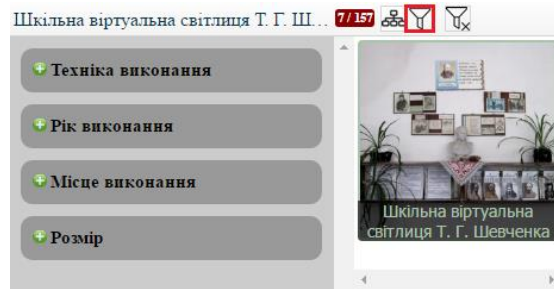


Рис. 4.17. Фільтра атрибутів і кнопка керування його видимістю

Фільтр по атрибутах дозволяє виключити з відображення об'єкти, значення атрибутів яких не відповідають заданим параметрам. Даний фільтр доступний тільки для онтологій, об'єкти яких мають іменовані (з встановленим класом) властивості.

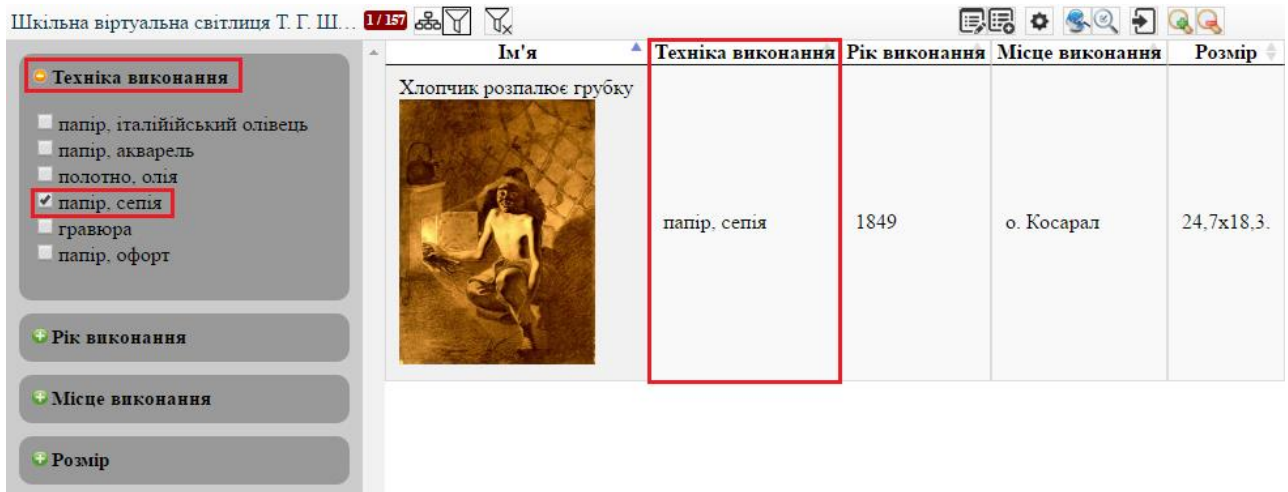


Рис. 4.18. Відбір об'єктів онтології за атрибутом «Техніка виконання» зі значенням «папір, сепія»

Основним варіантом фільтрації є фільтрація за точним значенням (Рис. 4.18). При використанні даного варіанту в фільтрі відображуються всі можливі значення певного атрибуту, і користувачем виділяються значення, які повинні мати об'єкти, що відображуються. Даний режим може працювати з будь-якими даними, однак якість роботи в деяких ситуаціях може бути незадовільною. Так, необхідно виключати з фільтрації поля з текстовими описами, оскільки в загальному випадку

вони унікальні для кожного з об'єктів (фільтрація по ним не має змісту, і замість неї краще використовувати внутрішній пошук).

Для числових значень пошук по точному значенню теж є незручним, тому використовується пошук по числовому інтервалу (Рис. 4.19). Деякі інші типи даних (зокрема, географічні координати і посилання на зображення) не відображаються в фільтрі.

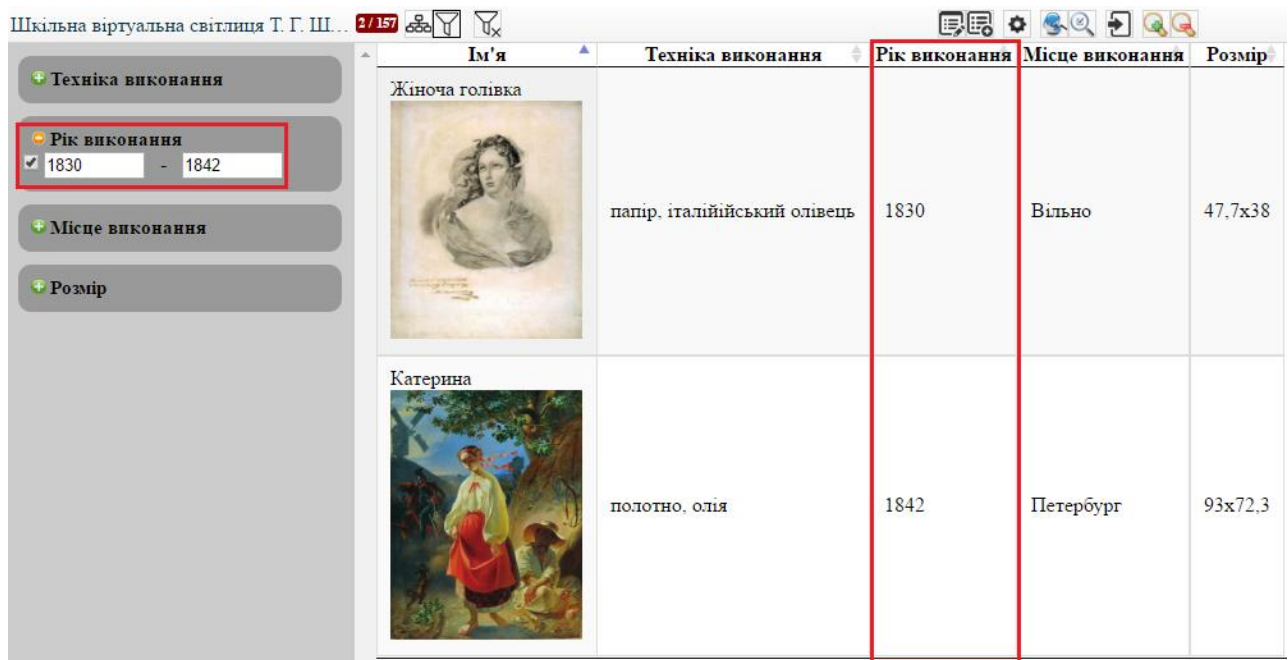


Рис. 4.19. Відбір об'єктів онтології з атрибутом «Рік виконання» в числовому інтервалі від 1830 до 1842

Кнопка «Зняти фільтр» відмінює дію фільтрів і повертає початковий режим відображення об'єктів онтології.

Кнопка «Бібліотека онтології» та «Бібліотека додатків» відкривають перелік доступних для перегляду таксономій з публічної бібліотеки, і список додаткових сервісів (зокрема, керуючого інтерфейсу редуктора).

Кнопка «Налаштування» відкриває форму зміни налаштувань роботи з інтерфейсом переглядача онтологій ТОДОС (Рис. 4.20). Доступні наступні налаштування:

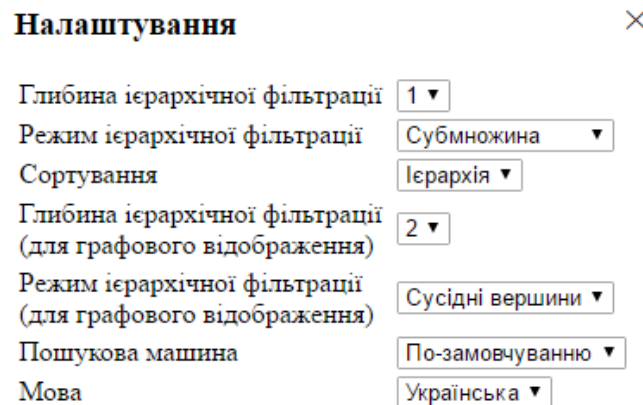
«Глибина/режим ієрархічної фільтрації» визначає, які об'єкти онтології чи вершини графа будуть показані після виконання ієрархічної фільтрації (субмножина – відображатимуться дочірні об'єкти/вершини, супермножина – відображатимуться

батьківські об'єкти/вершини, оточення – відображатимуться всі пов'язані з поточним об'єктом/вершиною). Дане налаштування визначається окремо для об'єктного і для ієрархічного вигляду.

«Сортування» визначає спосіб відображення об'єктів/вершин: ієрархія – за положенням в таксономії, назва – в лексикографічному порядку (за алфавітом).

«Пошукова машина» визначає зовнішнє джерело здійснення пошуку (електронні бібліотеки, архіви, оцифровані фонди та ін.).

«Мова» визначає мову інтерфейсу переглядача онтологій ТОДОС.



Налаштування	
Глибина ієрархічної фільтрації	1 ▼
Режим ієрархічної фільтрації	Субмножина ▼
Сортування	Ієрархія ▼
Глибина ієрархічної фільтрації (для графового відображення)	2 ▼
Режим ієрархічної фільтрації (для графового відображення)	Сусідні вершини ▼
Пошукова машина	По-замовчуванню ▼
Мова	Українська ▼

Рис. 4.20. Вікно налаштувань інтерфейсу переглядача онтологій

Кнопка «Вхід»/«Вихід» дозволяє увійти або вийти з онтологічного кабінету.

Кнопки «Збільшити» та «Зменшити» дозволяють змінити масштаб відображення об'єктів онтології.

У правій частині верхнього меню знаходяться кнопки зміни способу відображення об'єктів в основній частині інтерфейсу переглядача онтологій ТОДОС. Доступні наступні режими відображення:

Об'єктне відображення виводить об'єкти у вигляді рівнів. На верхньому рівні користувачем обирається об'єкт онтології, і на основі семантичних відношень пов'язаних з ним об'єктів (батьківських, дочірніх або суміжних – залежно від налаштувань) формується наступний рівень. За замовченням відображається об'єкт онтології, що відповідає темі таксономії, і сформований її дочірніми об'єктами рівень.

Кожна вершина містить три керуючих елемента – «Фільтрувати оточення», «Деталі», «Деталі (повноекранний режим)».

Перший елемент (Рис. 4.21) має функції, аналогічні до функцій «Ієрархічного фільтру».

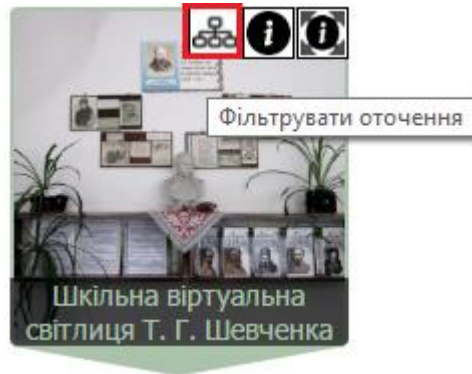


Рис. 4.21. Керуючий елемент «Фільтрувати оточення» об'єкта онтології

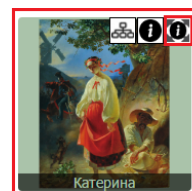
Інші два відображають інформацію про обраний об'єкт – у скороченому вигляді (Рис. 4.22) та в повноекранному режимі (Рис. 4.23).



Рис. 4.22. Керуючий елемент «Деталі» та відомості про об'єкт онтології в скороченому вигляді

При цьому в скороченому режимі дані відображаються також в зменшеній формі: показані менші розміри картинок, а також довгі тексти обрізуються до довжини в 300 символів, якщо початкова їх довжина більша. В цілому скорочений режим призначений для швидкого перегляду об'єктів, зокрема, при пошуку потрібної інформації.

Катерина



Опис: «КАТЕРИНА» (полотно, олія, 93х72,3) — картина Шевченка, намальована 1842 у Петербурзі. На полотні внизу праворуч два написи, один чорною, а нижче — майже на краю — червоною фарбою (з певними втратами фарбового шару): «1842. // Т: Шевченко». Зберігається у НМТШ (№ ж—100). На фототипії картини, вперше репрод. у журн. «Киевская старина» (1896. № 2. С. 4), видно лише фрагменти букв, що виконано чорною фарбою унизу праворуч біля намальованого листка. Б. Грінченко як укладач Каталогу Музею Тарновського у розд. «VIII. Его картины масляными красками» першою за номером подає картину «К.» (С. 193. № 427), в описі якої вказує, що підпис і дата, виконані червоною фарбою, — авторські. Цієї думки дотримувалися дослідники І. Дорофієнко та В. Яшук. Про експоновану на Ювілейній Шевч. виставці 1939 у Києві картину «К.» у каталозі 1941 сказано: «Шевченко Тарас Григорович. Катерина. 93,5х72. Полотно, олія. Внизу червоною фарбою: "1842. Т. Шевченко"». Була в Чернігівському Обласному музеї (кол. збірка В. Тарновського). Галерея Картин Т. Г. Шевченка. Харків» (С. 27. № 138). З цього опису випливає, що в

Рис. 4.23. Керуючий елемент «Деталі (повноекранний режим)» та відомості про об'єкт онтології в повноекранному режимі

В інформаційній картці об'єкта онтології, що відкривається керуючим елементом «Деталі (повноекранний режим)» також є можливість пошуку в зовнішніх та внутрішніх джерелах інформації (кнопки «Зовнішній пошук», «Внутрішній пошук» у верхньому правому куті екрана).

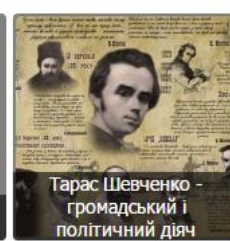
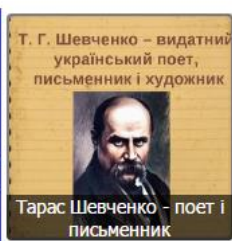
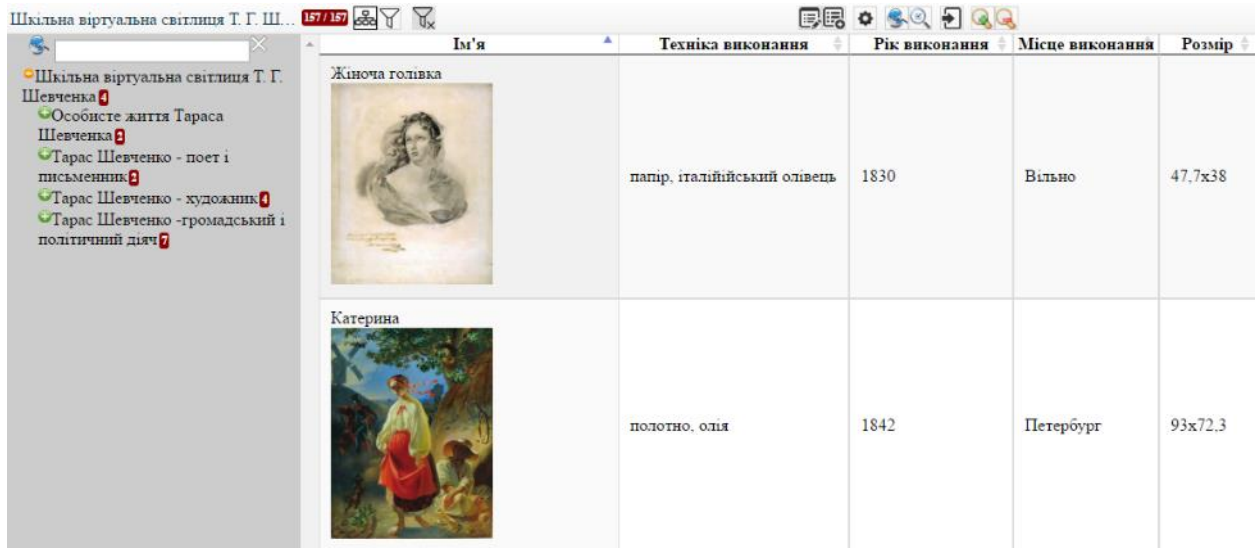


Рис. 4.24. Об'єктне відображення дочірніх елементів

Якщо поточний об'єкт має пов'язані (згідно налаштувань за замовченням) об'єкти, при наведенні курсору нижня частина об'єкта перетвориться на стрілку, то при натисненні лівої кнопки миші на такому об'єкті утвориться додатковий рівень нижче поточного із пов'язаних з ним елементів (Рис. 4.24).

На початку кожного рівня (крім вищого) є іконка копії обраного на вищому рівні об'єкта – об'єкт-класу. Дана копія є пасивною і не містить керуючих елементів.

Табличне відображення об'єктів онтології є таблицею іменованих атрибутів об'єктів та їх значень. Верхній рядок (з іменами атрибутів) є фіксованим і зберігає своє положення при прокрутці таблиці з іменами об'єктів, їх зображеннями та значеннями атрибутів (Рис. 4.25).



Ім'я	Техніка виконання	Рік виконання	Місце виконання	Розмір
Жіноча голівка 	папір, італійський олівець	1830	Вільно	47,7x38
Катерина 	полотно, олія	1842	Петербург	93x72,3

Рис. 4.25. Табличне відображення елементів онтології

Даний режим зручний для перегляду багатьох об'єктів онтології одночасно, зокрема, для їх порівняння або в ситуації, коли необхідно знайти певні значення, але немає можливості задати запит до пошукової системи або встановити фільтр.

Режим табличного перегляду доступний тільки для онтологій, об'єкти яких мають іменовані (з встановленим класом) властивості.

Графове відображення представляє таксономію у вигляді графа (Рис. 4.26). Даний режим дублює роботу підсистеми відображення редактора таксономій, і основною відмінністю при цьому є те, що керуючі елементи керування фільтрацією розташовані в налаштуваннях.



Рис. 4.26. Графове відображення таксономії області дослідження

4.2.2. Модуль трансдисциплінарного представлення інформації

Модуль трансдисциплінарного представлення інформації призначений для контекстної зв'язки об'єктів різних онтологій і представлення їх у вигляді трансдисциплінарного куба.

Пошуковий запит: Катерина			Ok
Граф	Вершина	Керування	
Культурна спадщина Т. Г. Шевченка	Сліпа з дочкою	Деталі	Таксономія
		Об'єктна онтологія	
Культурна спадщина Т. Г. Шевченка	Портрет Толстого Федора Петровича	Деталі	Таксономія
		Об'єктна онтологія	
Культурна спадщина Т. Г. Шевченка	Портрет невідомої в бузковій сукні	Деталі	Таксономія
		Об'єктна онтологія	
Портал Шевченка	Олійні роботи	Деталі	Таксономія
		Об'єктна онтологія	
http://kobzar.ua/uploads/comp/src/doc/13.doc		Деталі	Перехід
http://kobzar.ua/map/world		Деталі	Перехід
http://kobzar.ua/item/5329		Деталі	Перехід
http://kobzar.ua/item/show/5329		Деталі	Перехід

Рис. 4.27. «Внутрішній пошук»: результати пошуку

Трансдисциплінарне представлення відбувається з допомогою кнопки «Внутрішній пошук», яка відкриває зображений на Рис. 4.27. інтерфейс контекстної зв'язки. Даний інтерфейс призначений для відображення списку інформаційних ресурсів, що контекстно зв'язані з поточним згідно вказаного пошукового запиту. Запит можна ввести вручну у відповідне поле, або вибрати з контексту поточного об'єкту перед натиском кнопки «Внутрішній пошук» (як показано на Рис. 4.28). До інформаційних джерел належать структуризовані документи, створені вручну онтології та інші веб-ресурси, що були задані адміністратором.

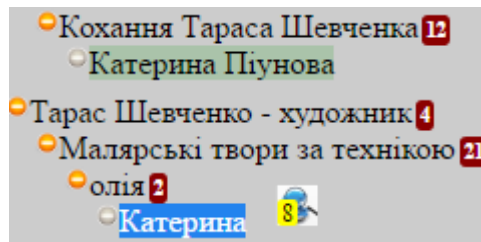


Рис. 4.28. Формування запиту для контекстної зв'язки

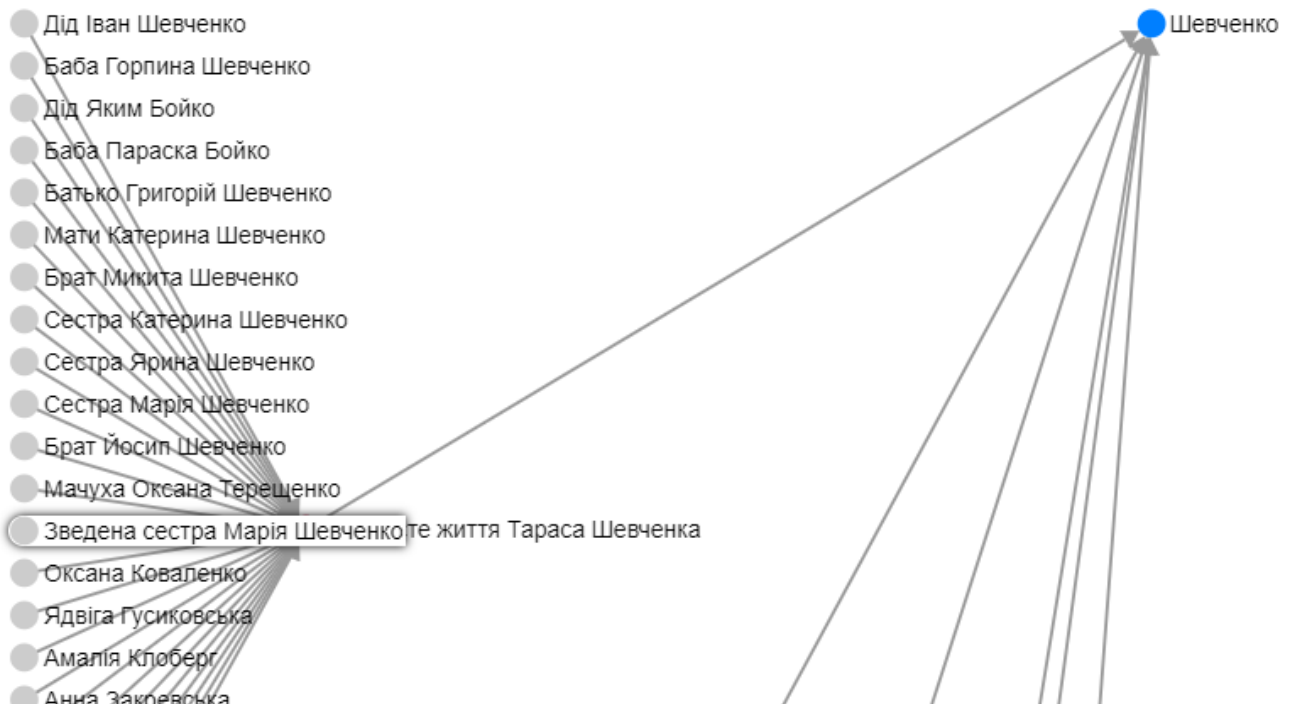


Рис. 4.29. Фрагмент трансдисциплінарного представлення,
присвяченого Т. Г. Шевченку

У вікні пошуку (Рис. 4.27.) відображується список знайдених документів, і базова інформація про них. Біля кожного з документів є кнопка «Деталі», що дозволяє переглядати основну інформацію про знайдений документ без переходу на

нього, зокрема – знайдені входження пошукового запиту, а для таксономій – також базову інформацію про вершину, контекст якої містить пошуковий запит. Також доступний перехід, на графове або об'єктне представлення для таксономій, і на сам ресурс – для не таксономізованих веб-ресурсів.

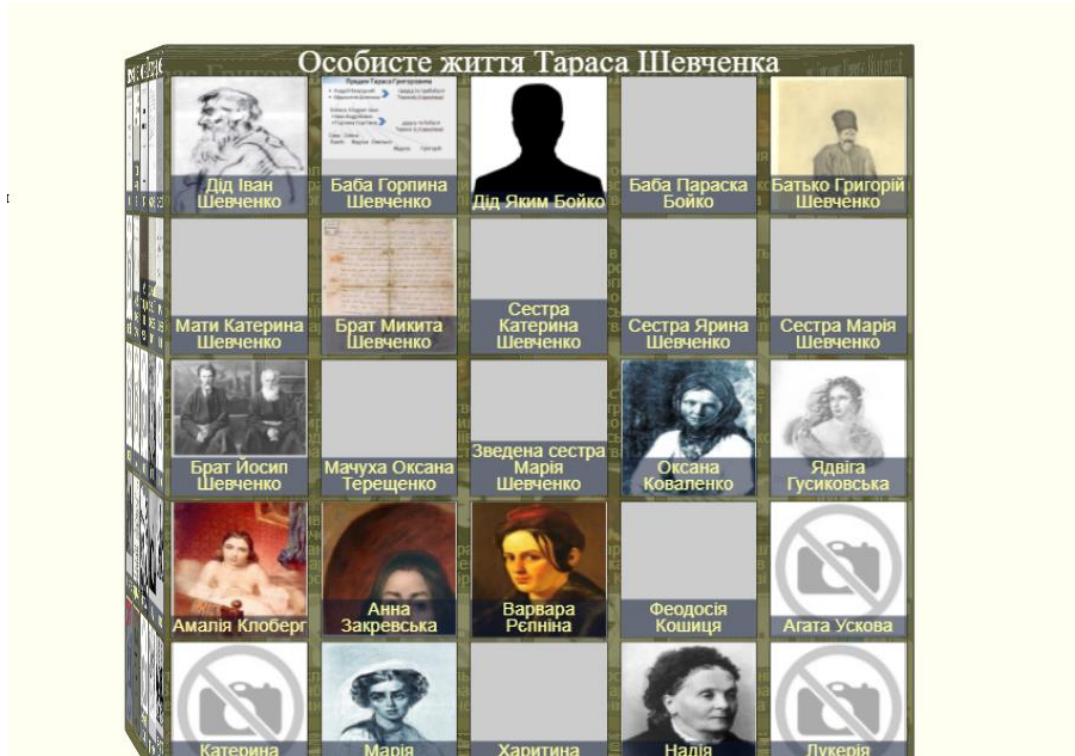


Рис. 4.30. Трансдисциплінарний куб, присвячений Т. Г. Шевченку

З допомогою контекстної зв'язки користувач може створювати трансдисциплінарне представлення інформації. Для цього необхідно:

- 1) Створити пусту онтологію з одною кореневою вершиною.
- 2) Створити для кореневої вершини від одної до чотирьох дочірніх вершин, що представляють пошукові запити.
- 3) Виконати кожен з пошукових запитів з допомогою функції «Внутрішній пошук»
- 4) Вибрати з наданого списку інформаційних ресурсів ті, що релевантними до поставленої користувачу задачі.
- 5) Здійснити «Перехід» для кожного з вибраних інформаційних ресурсів.
- 6) Якщо ресурс представляє собою вершину певного онтографу, її можна скопіювати в новий граф з допомогою функції «Копіювати»/«Вставити», доступної через контекстне меню вершини. Якщо ресурс представляє

неструктурований текст, необхідно вручну створити в новому графі вершин і перенести в неї всю необхідну інформацію.

Таким чином можна побудувати нову онтологію, що і буде трансдисциплінарним представленням. Приклад такого представлення показано на Рис. 4.29.

В подальшому таке представлення може бути відображене у вигляді трансдисциплінарного кубу, приклад якого показано на Рис. 4.30. Кожен з пошукових запитів представляється гранню такого куба і відображує певну ПдО або її частину.

4.2.3. Модуль формування онтологічних ГІС-додатків

В рамках ТОДОС-ГІС існує можливість представлення геопросторової інформації у вигляді ГІС з допомогою широкого спектра допоміжних систем або бібліотек. ГІС-додатки можуть формуватися одним із двох способів: з допомогою підсистеми експорту в існуючі ГІС-системи або у вигляді інтегрованого в ТОДОС-ГІС онтологічного ГІС-додатку.

Server:
http://server1.inhost.com.ua:6080/arcgis/rest/services/Test/JSONTest/FeatureServer

Update ○		Create ○		Clear ○	View ●
Fields:					
Ontology link field:					
ontology					
Compare field:		Name field:			
label		label			
		Coordinate field:			
		coordx:coordy			
Graph:					
342					
Auth:					
Login:					
user					
Password:					
pass					

Submit

Рис. 4.31. Інтерфейс експорту онтології в ГІС

Як було показано вище, більшість онтології можна представити у вигляді таблиці. В подальшому дану таблицю можна використовувати як джерело даних для

певної ГІС, для чого необхідно виконати експорт у її сховище даних. Даний процес виконується з допомогою інтерфейсу, показаного на Рис. 4.31.

Даний інтерфейс призначений для експорту онтології в ArcGIS, але може бути модифікований для роботи з іншими системами, зокрема такими, що побудовані на основі СУБД PostgreSQL [96] з розширенням PostGIS [144].

Інтерфейс, зображений на Рис. 4.31, може працювати в чотирьох режимах: створення об'єктів (Create), оновлення інформації в існуючих об'єктах (Update), видалення об'єктів (Clear) і перегляд наявних в ГІС об'єктів (View). Для роботи необхідно ввести посилання, за яким знаходиться точка доступу до ArcGIS Map Server, а у випадку, якщо публічний доступ до даного серверу закритий – також користувацькі логін і пароль для авторизації.

При виконанні створення або оновлення інформації в об'єктах додатково необхідно вказувати внутрішній ідентифікатор онтології (Graph), назву поля в ГІС, що використовується для зберігання імені об'єкту (Compare field/Name field), і назву поля, в яке необхідно вписати посилання на об'єкт онтології. При створення також потрібно вказувати поле онтології, в якому знаходиться географічна інформація у формі координат (coordinate field).

Прикладом ГІС-додатку, створеного з допомогою підпрограми експорту в ГІС, є карта встановлених тарифів на послуги ліцензіатів з централізованого опалення, централізованого постачання гарячої води та послуги на теплову енергію з 1 травня (квітня) 2015 р. постановами Національної комісії, що здійснює державне регулювання у сфері комунальних послуг та Національної комісії, що здійснює державне регулювання у сферах енергетики та комунальних послуг. Вихідними даними для формування даного ГІС-додатку є інтегрована порівняльна таблиця тарифів, фрагмент якої показано на Рис. 4.32. Вихідні дані для побудови даної таблиці приведені в Додатку Д (Таблиця 1).

На Рис. 4.32. в стовпчику «НАЗВА» наведено перелік ліцензіатів, на житлово-комунальні послуги яких встановлено тарифи. В наступних стовпчиках представлена структура тарифу, найменування якого визначається за заголовком стовпчика. Число, виділене жирним, є значенням встановленого тарифу, інші числа

– складові частини тарифної ставки (витрати на теплову енергію, витрати на оплату праці та внески на соціальні заходи, послуги банку, решта, податок на додану вартість та інше).

Відсотковий показник в першому рядку кожної комірки визначає відношення значення тарифної ставки даного ліцензіата до середньозваженого показника тарифу всіх ліцензіатів, наведених в таблиці по даному тарифу. Якщо цей показник перевищує 100% (за 100% приймається значення середньозваженого показника по певному тарифу), то він відображається червоним кольором, що означає перевищення тарифної ставки по відношенню до середнього, в іншому випадку – зеленим.

№	НАЗВА	ТАРИФ З 1 ТРАВНЯ 2015 НА ПОСЛУГИ З ЦЕНТРАЛІЗОВАНОГО ОПАЛЕННЯ ДЛЯ НАСЕЛЕННЯ З БУДИНКОВИМИ ТА КВАРТИРНИМИ ПРИЛАДАМИ ОБЛІКУ ТЕПЛОВОЇ ЕНЕРГІЇ	ТАРИФ З 1 ТРАВНЯ 2015 НА ПОСЛУГИ З ЦЕНТРАЛІЗОВАНОГО ОПАЛЕННЯ ДЛЯ НАСЕЛЕННЯ БЕЗ БУДИНКОВИХ ТА КВАРТИРНИХ ПРИЛАДІВ ОБЛІКУ ТЕПЛОВОЇ ЕНЕРГІЇ	ТАРИФ З 1 ТРАВНЯ 2015 НА ПОСЛУГИ З ЦЕНТРАЛІЗОВАНОГО ПОСТАЧАННЯ ГАРЯЧОЇ ВОДИ ДЛЯ НАСЕЛЕННЯ З РУШНИКОСУШИЛЬНИКАМИ	ТАРИФ З 1 ТРАВНЯ 2015 НА ПОСЛУГИ З ЦЕНТРАЛІЗОВАНОГО ПОСТАЧАННЯ ГАРЯЧОЇ ВОДИ ДЛЯ НАСЕЛЕННЯ БЕЗ РУШНИКОСУШИЛЬНИКІВ
1	Відокремлений підрозділ «Ладихинська ТЕС» ПАТ «ДТЕК Західенерго»	344,79 (Вот) 51% 279,14 (Вит) 80.96% 5,31 (Вит) 1.54% (Вит) (Вит) 2,87 (Пос) 0.83% 0 (Реш) 57,47 (Под) 16.67%	8,86 (Вот) 49% 7,17 (Вит) 80.93% 0,14 (Вит) 1.58% (Вит) (Вит) 0,07 (Пос) 0.79% 0 (Реш) 1,48 (Под) 16.7%	20,53 (Вот) 50% 14,99 (Вит) 73.02% 5,31 (Вит) 25.88% (Вит) (Вит) 2,87 (Пос) 13.98% -6,07 (Реш) 3,42 (Под) 16.66%	18,92 (Вот) 50% 13,74 (Вит) 72.62% 0,28 (Вит) 1.48% 0,28 (Вит) 3.22% 0,98 (Вит) 5.18% 0,16 (Пос) 0.85% 0 (Реш) 3,15 (Под) 16.65%
2	ДП «Теплокомуненерго Маяк» ВАТ «Маяк»	699,54 (Вот) 104% 574,64 (Вит) 82.15% 3,07 (Вит) 0.44% (Вит) (Вит) 5,25 (Пос) 0.75% 0 (Реш) 116,59 (Под) 16.67%	16,85 (Вот) 93% 13,84 (Вит) 82.14% 0,07 (Вит) 0.42% (Вит) (Вит) 0,13 (Пос) 0.77% 0 (Реш) 2,81 (Под) 16.68%	42,12 (Вот) 103% 30,83 (Вит) 73.2% 0,18 (Вит) 0.43% 0,53 (Вит) 1.28% 3,24 (Вит) 7.69% 0,32 (Пос) 0.78% 0 (Реш) 7,02 (Под) 16.67%	38,94 (Вот) 103% 28,26 (Вит) 72.57% 0,17 (Вит) 0.44% 0,49 (Вит) 1.28% 3,24 (Вит) 8.32% 0,29 (Пос) 0.74% 0 (Реш) 6,49 (Под) 16.67%

Рис. 4.32. Фрагмент інтегрованої порівняльної таблиці тарифів

Сформований на основі наявної в інтегрованій таблиці інформації про тарифи ГІС-додаток представлено на Рис. 4.33. На ньому візуально представлено географічне розташування ліцензіатів надання житлово-комунальних послуг за тарифами на централізоване опалення з будинковими та квартирними приладами обліку теплової енергії в межах адміністративно-територіальних одиниць та кліматичних зон, та наведено їх градацію за 4 шкалами:

– За значенням тарифної ставки (0 - 971,34 грн/Гкал);

- За відношенням витрат на теплову енергію до тарифної ставки (78,36% - 82,61%);
- За відношенням витрат на оплату праці до тарифної ставки (0% - 3,8%);
- За відношенням витрат на послуги банку до тарифної ставки (0% - 1,74%).



Рис. 4.33. ГІС-додаток, сформований на основі інтегрованої таблиці тарифів

Даний ГІС-додаток також надає можливість проведення аналітики з допомогою інтегрованих засобів ArcGIS. Наприклад, з допомогою іконки «Діаграма» на панелі інструментів карти можна виконувати статистичний аналіз та візуалізацію його результатів для обраних на карті ліцензіатів.

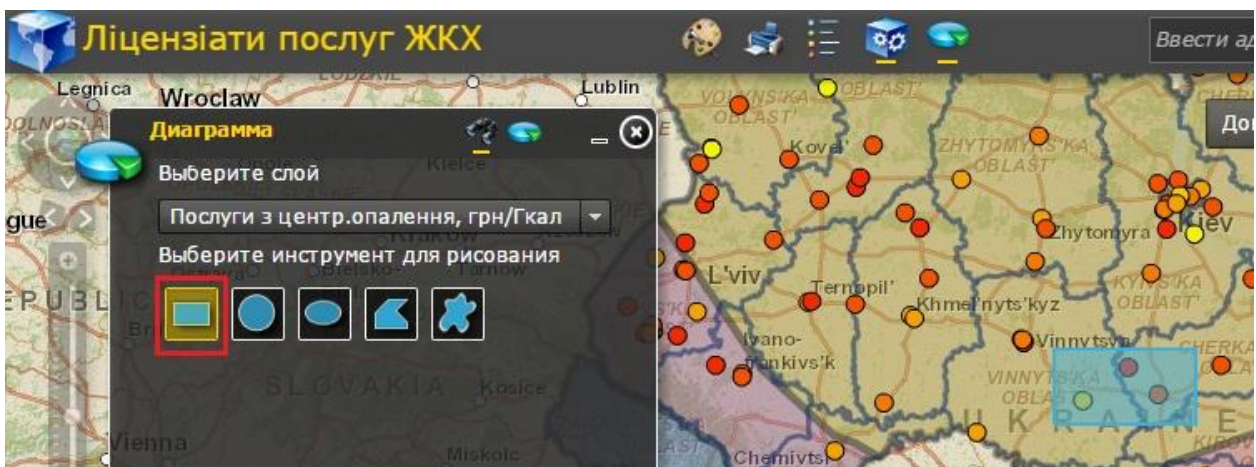


Рис. 4.34. Вибір об'єктів для побудови діаграми

Вибір ліцензіатів здійснюється шляхом окреслення навколо обраної множини об'єктів геометричних фігур, зразки яких наведені у вікні інструменту «Диаграмма» (Рис. 4.34.).

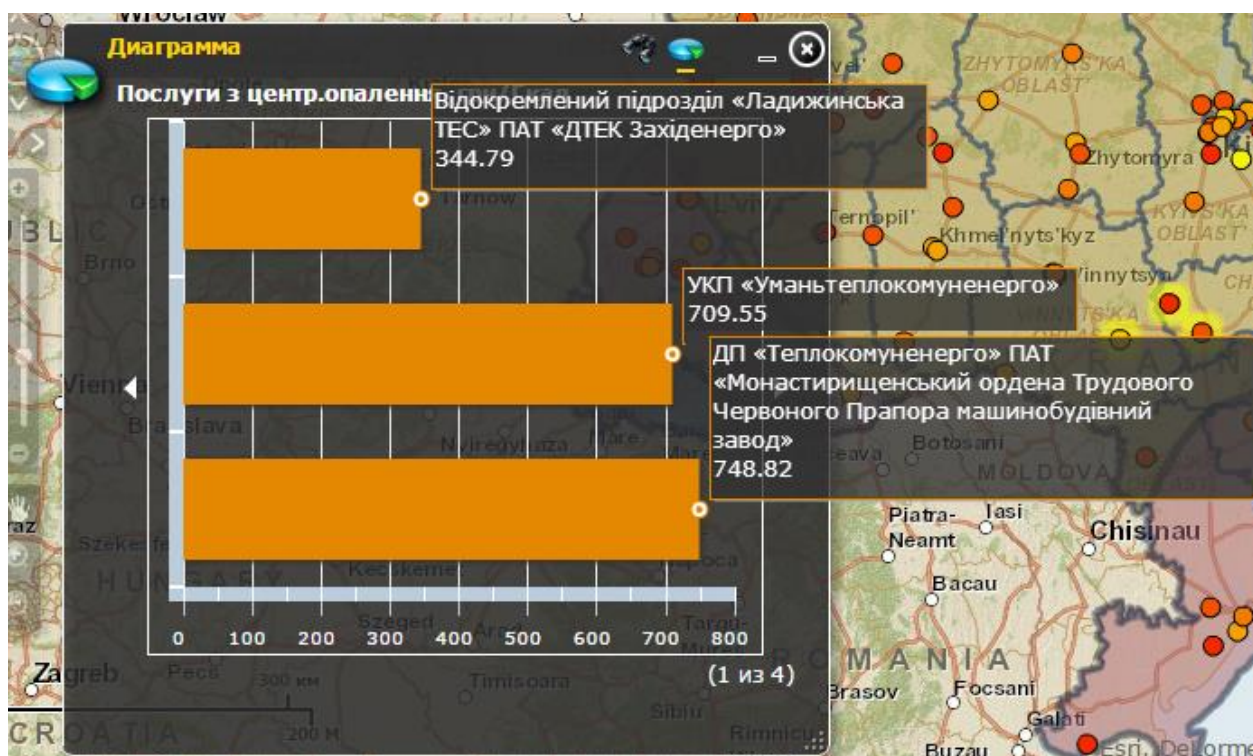


Рис. 4.35. Загальний вигляд діаграми

Результат порівняння обраних ліцензіатів відображається у вигляді діаграми, наведення курсору миші на стовпчики якої відкриває інформаційне вікно із зазначенням назви ліцензіату та значенням порівнюваного параметру. Обрані ліцензіати підсвічені на карті жовтим кольором (Рис. 4.35).

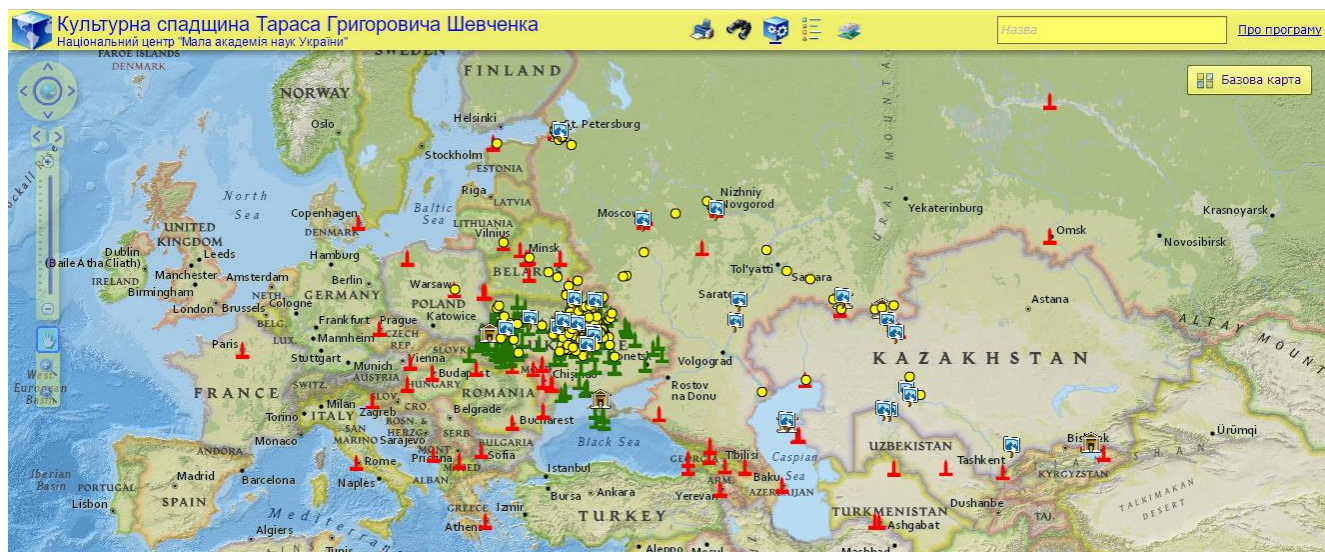


Рис. 4.36. ГІС-додаток «Культурна спадщина Тараса Григоровича Шевченка»

Іншим прикладом такого ГІС-додатку є додаток «Культурна спадщина Тараса Григоровича Шевченка» [31, 32] (Рис. 4.36).



Рис. 4.37. Вхідні дані для ГІС-додатку «Культурна спадщина Т. Г. Шевченка»

Даний додаток був сформований в процесі обробки значної кількості інформаційних ресурсів (зокрема, книг і статей) присвячених життю і творчості Т.Г. Шевченка [32]. Результати обробки були представлені у вигляді онтологій (одна з яких, зокрема, представлена на Рис. 4.11), які в подальшому були експортовані в ArcGIS. Деякі інформаційні ресурси, що були оброблені в процесі створення ГІС-додатку, показані на Рис. 4.37.

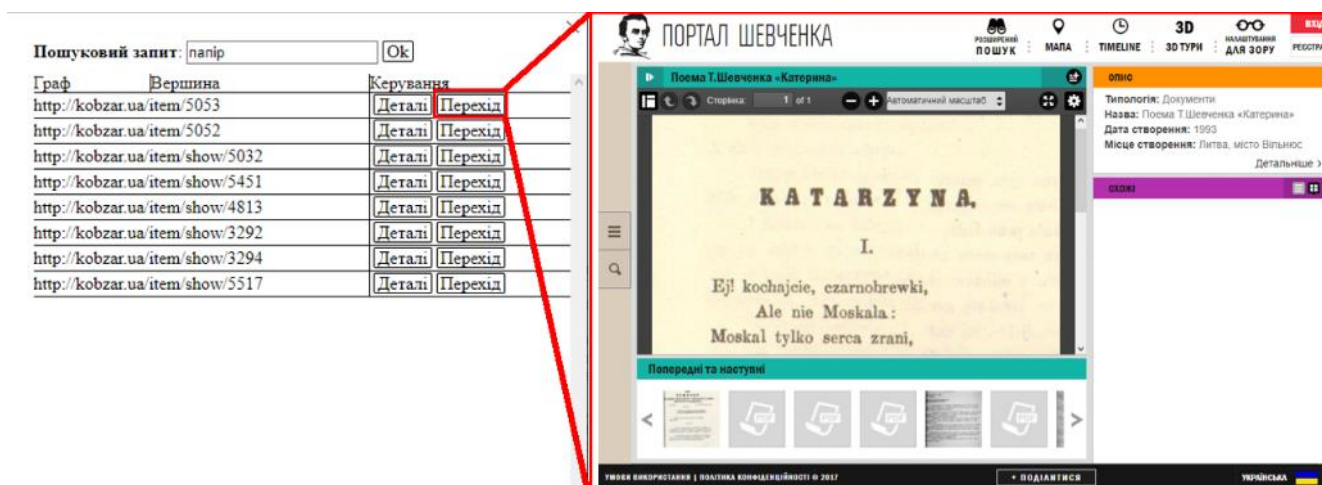


Рис. 4.38. Трансдисциплінарне представлення інформації, присвяченої Т. Г. Шевченку, в середовищі Порталу Шевченка

Особливістю даного ГІС-додатку є те, що він знаходиться в середовищі Порталу Шевченка [32], і тому наявна в ньому інформація може бути трансдисциплінарно представлена в даному середовищі, як, зокрема, показано на Рис. 4.38.

Експорт даних в ГІС часто є незручним, оскільки може призводити до втрат даних (зокрема, множинних зв'язків між об'єктами), і в більшості випадків потребує втручання адміністратора для первинної настройки серверів. Для спрощення процедури формування ГІС-додатків призначений модуль онтологічних ГІС-додатків, який дозволяє формувати такі додатки на основі онтології – без використання допоміжних механізмів. Модуль інтегрований у веб-інтерфейс ТОДОС і стає доступним автоматично при перегляді онтологій, що містять геопросторові дані. Приклади сформованих з допомогою даного модуля ГІС-додатків:

ГІС-додаток «Місця видалення відходів» сформований в ході обробки даних про сміттєзвалища, наданих у відкритий доступ Міністерством екології та природних ресурсів України. Дані були отримані з набору таблиць виду, показаного в Додатку Д (Таблиця 2.). Загальний вигляд онтологічного ГІС-додатку «Місця видалення відходів» показано на Рис. 4.39.

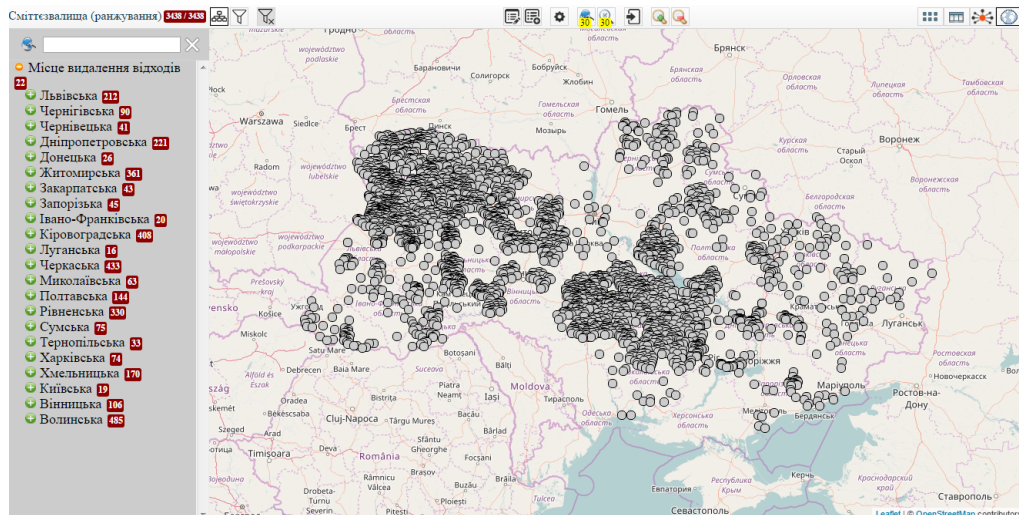


Рис. 4.39. Онтологічний ГІС-додаток «Місця видалення відходів»

ГІС-додаток «Агро» був сформований в ході обробки реєстру сільськогосподарських дорадників та експертів-дорадників [82], сільськогосподарських дорадчих служб [83], а також наявної у відкритому доступі

інформації про оренду агротехніки [56]. Загальний вигляд ГІС-додатку «Агро» показано на Рис. 4.39.

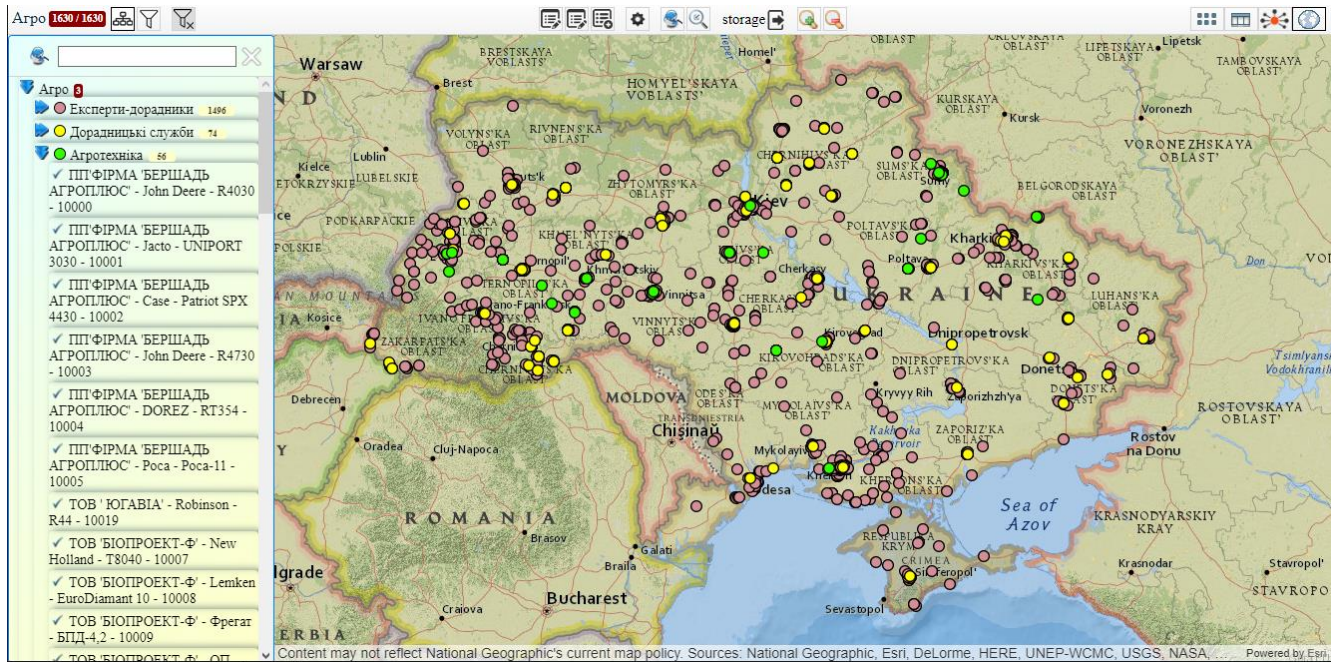


Рис. 4.40. Онтологічний ГІС-додаток «Агро»

Більшість наведених вище ГІС-додатків розміщені у відкритому доступі, хоча деякі з них (зокрема, «Агро») є частиною закритого користувацького середовища і потребують авторизації для перегляду.

4.3. Емпірична перевірка засобів трансдисциплінарного представлення геопросторової інформації

Найкращими вхідними даними для засобів трансдисциплінарного представлення геопросторової інформації (зокрема, засобу «Рекурсивний редуктор») є великі масиви однотипних текстів. Це пов'язано з тим, що для його роботи необхідний формальний опис підмножини мови, що аналізується, а найпростіший спосіб отримати такий опис – аналізувати існуючі тексти. Тому вхідну множину текстів T необхідно розділити на три підмножини [14, 81]:

- 1) Навчальну вибірку T_l .
- 2) Тестову вибірку T_t .
- 3) Робочу вибірку T_w .

Такі, що $T_l \cup T_t \cup T_w = T$.

Всі тексти, що належать T_l необхідно проаналізувати підсистемою структуризації. Аналіз необхідно проводити послідовно. Кожен з текстів буде відображений з допомогою редактора шаблонів, який показуватиме існуючі лексеми і їх інтерпретації, сформовані лексичним аналізатором. На даному етапі необхідно ідентифікувати помилки лексичного аналізу, серед яких основними є:

- 1) Поява не ідентифікованих лексем, викликана неповнотою бази знань аналізатора.
- 2) Наявність у вхідному тексті помилок, що не дозволяють аналізатору правильно ідентифікувати лексему.
- 3) Неправильна ідентифікація лексем, викликана омонімією (особливо для низькоякісних текстів, що містять фрагменти, що належать різним мовам).

Оскільки в загальному випадку модифікація вхідних текстів чи бази знань лексичного аналізатора є занадто трудомісткою справою, необхідно для кожної підмножини створювати спеціалізовану базу правил, що враховує дані некоректності і містить спеціально сформовані предикати для ідентифікації всіх можливих (правильних і неправильних) варіантів.

Після закінчення аналізу текстів з вихідної вибірки необхідно аналогічним чином аналізувати тексти з T_t , але не модифікувати базу правил, а визначати якість обробки.

Якщо отримана якість є достатньою, то можна запускати аналіз текстів з T_w . Інакше необхідно вибрати нову тестову вибірку $T_t^* \subset T_w$, і сформувані нові множини $T_l^* = T_l \cup T_t$ і $T_w^* = T_w \setminus T_t^*$. Даний процес необхідно повторювати, доки не буде отримане задовільне значення якості.

Досліджувати якість роботи інструментарію можна з допомогою F-міри [75, 143].

Для оцінки якості обчислюються наступні три параметра:

- TP (True Positive) – кількість блоків інформації, що правильно ідентифіковані системою;

- FP (False Positive) – кількість послідовностей лексем, що не містять релевантної інформації, але були ідентифіковані системою;
- FN (False Negative) – кількість блоків інформації, що не були ідентифіковані системою;

Дані значення можна отримати з допомогою порівняння результатів роботи інструментальних засобів і результатів обробки вхідного тексту експертом ПдО.

Оцінка якості роботи алгоритму здійснюється за трьома параметрами:

- Точність (Precision) являє собою відношення коректно ідентифікованих об'єктів до кількості всіх ідентифікованих системою об'єктів;
- Повнота (Recall) являє собою відношення кількості правильно виділених ідентифікованих системою об'єктів до кількості всіх об'єктів в тексті.

F-міра являє собою інтегральний показник точності і повноти, і обчислюється як їх середнє гармонічне.

Зазначені вище параметри обчислюються за формулами (4.1)–(4.3).

$$Precision = \frac{TP}{TP + FP} \quad (4.1)$$

$$Recall = \frac{TP}{TP + FN} \quad (4.2)$$

$$F = (\beta^2 + 1) \frac{Precision \times Recall}{\beta^2 Precision + Recall} \quad (4.3)$$

Оскільки важко оцінити відносну важливість точності і повноти в процесі виділення текстів, є зміст використовувати збалансовану F-міру (4.4):

$$F = \frac{Precision \times Recall}{Precision + Recall} \quad (4.4)$$

Для одержання кількісних та якісних результатів роботи засобів трансдисциплінарного представлення геопросторової інформації були проведені експериментальні дослідження на настільному комп'ютері з наступними характеристиками:

- двоядерний процесор Intel Core i3-2100T, 2,5 ГГц;
- 4 ГБ оперативної пам'яті DDR3;
- графічний прискорювач NVIDIA GeForce GT 640M

Для тестування інструментальних засобів було довільним чином вибрано 10 текстів, кожний з яких мав обсяг від 1 до 26 сторінок. Тексти містили географічну інформацію у вигляді координат і адрес, а також назви об'єктів, до яких відносилась дана географічна інформація. Результати обчислення якості обробки даних текстів приведені в Таблиця 4.1.

Таблиця 4.1. Результати обчислення якості обробки текстів

	Точність	Повнота	F-міра
Імена	0,72	0,85	0,78
Географічна інформація	0,92	0,74	0,82
Загальна оцінка	0,82	0,78	0,80

Як видно з таблиці, якість виділення імен і географічної інформації кардинально відрізняються.

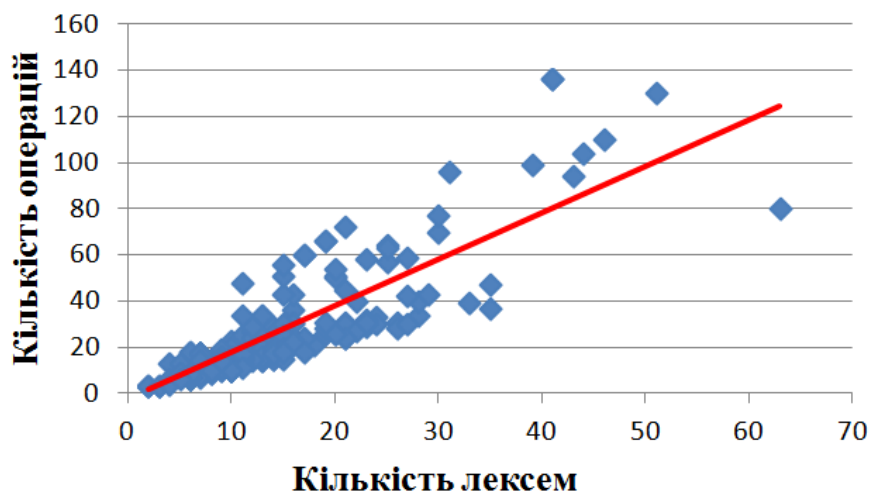


Рис. 4.41. Графік залежності швидкості обробки речення від його довжини

Географічна інформація після ідентифікації проходить процедуру валідації, що дозволяє добитись надзвичайно високої точності. Однак, в ході цієї процедури деякі виділені елементи даних відкидаються, що значно знижує повноту.

Для імен ситуація прямо протилежна – для них не існує ефективних алгоритмів валідації, тому точність їх ідентифікації відносно низька. Але завдяки тому, що не відбувається відкидання елементів даних, підвищується повнота.

В цілому ефективність виділення імен дещо нижча, оскільки в способах написання імен є значно більша варіативність.

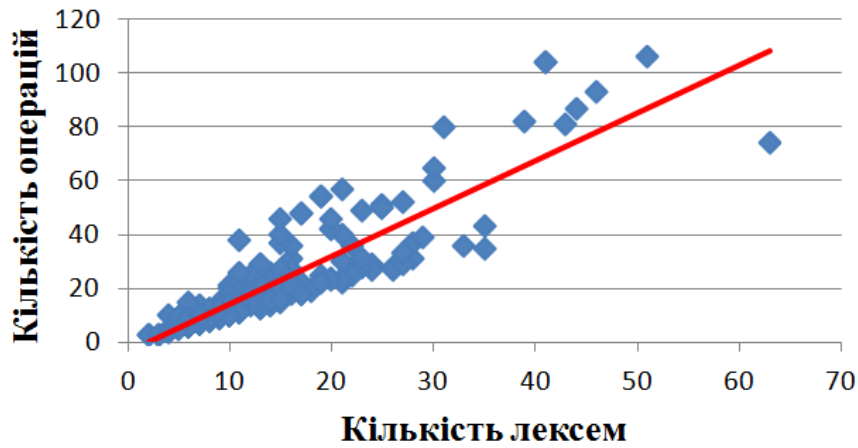


Рис. 4.42. Швидкість обробки речень з допомогою урізаної бази правил

Швидкодія роботи інструментальних засобів на пряму залежить від кількості виклику операцій застосування предикату до вхідної лексеми. Оскільки кожне правило працює в рамках одного речення, то залежність швидкодії від кількості речень є лінійною. Залежність швидкості від довжини речення показана на графіку, показаному на Рис. 4.41.

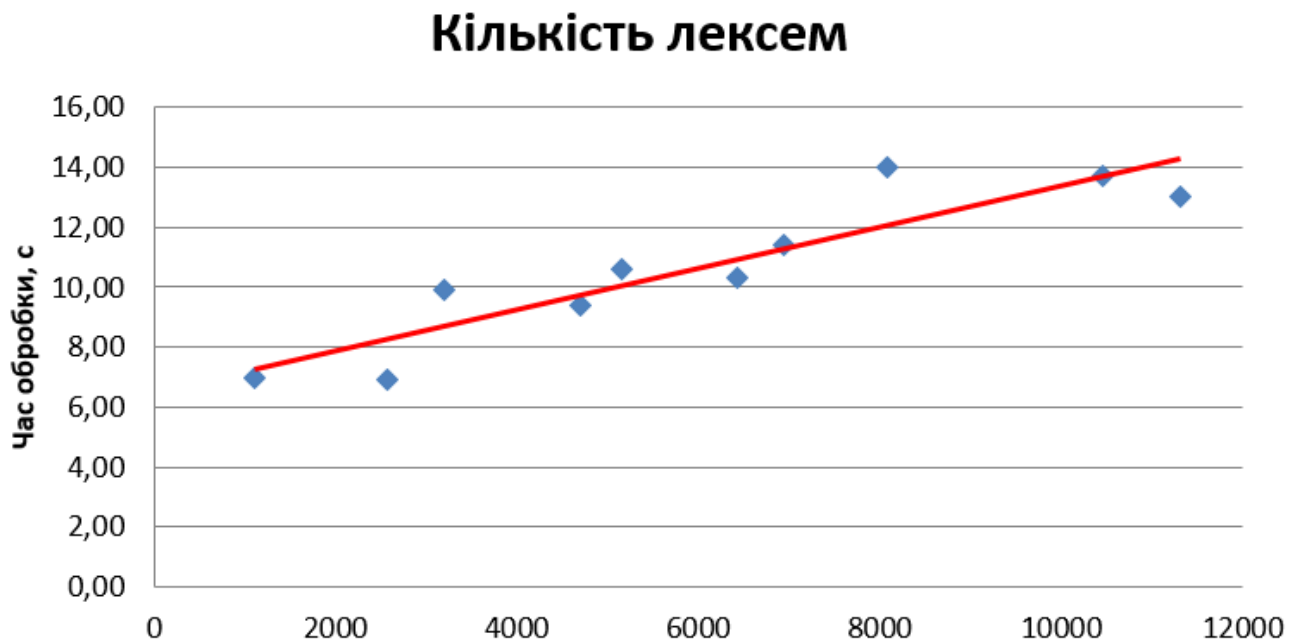


Рис. 4.43. Швидкість обробки тексту в залежності від кількості лексем

Як видно, залежність кількості викликів операцій (і, відповідно, швидкодії) від довжини речення близька до лінійної.

Також швидкодія залежить від розмірів бази правил. Наприклад, якщо відкинути близько половини правил, швидкодія інструментальних засобів дещо підвищиться (Рис. 4.42.).

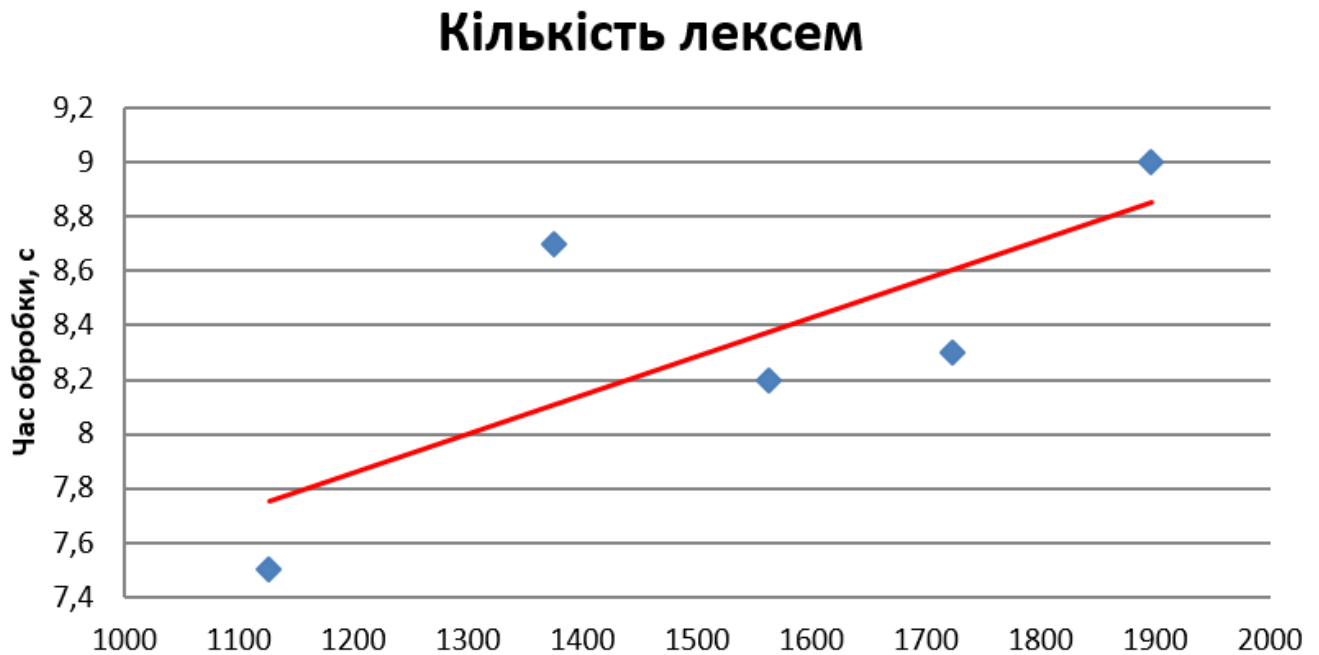


Рис. 4.44. Швидкість обробки вхідних даних для ГІС-додатку «Ліцензіати послуг ЖКХ»

Як можна побачити, основним чинником, що впливає на швидкість роботи інструментальних засобів, є розміри вхідного тексту. Залежність швидкості роботи алгоритму від довжини тексту (в лексемах) показана на графіку, показаному на Рис. 4.43.

Дана залежність отримана для тестового масиву текстів, що використовувався для обчислення точності роботи інструментальних засобів.

Швидкодія роботи алгоритму обчислювалась не тільки для тестової вибірки, але і в рамках процесів обробки вхідних даних для ГІС-додатків «Ліцензіати послуг ЖКХ» (Рис. 4.44.) і «Культурна спадщина Тараса Григоровича Шевченка» (Рис. 4.45). При цьому з масивів вхідних даних були вибрані вибірки з п'яти документів різного розміру.

В цілому розроблені засоби мають достатньо високу швидкодію, і придатні для використання в процесах підтримки прийняття рішень, що вимагають оперативного

аналізу даних. Інструментальні засоби також придатні для структуризації великих обсягів даних, таких, як книги.

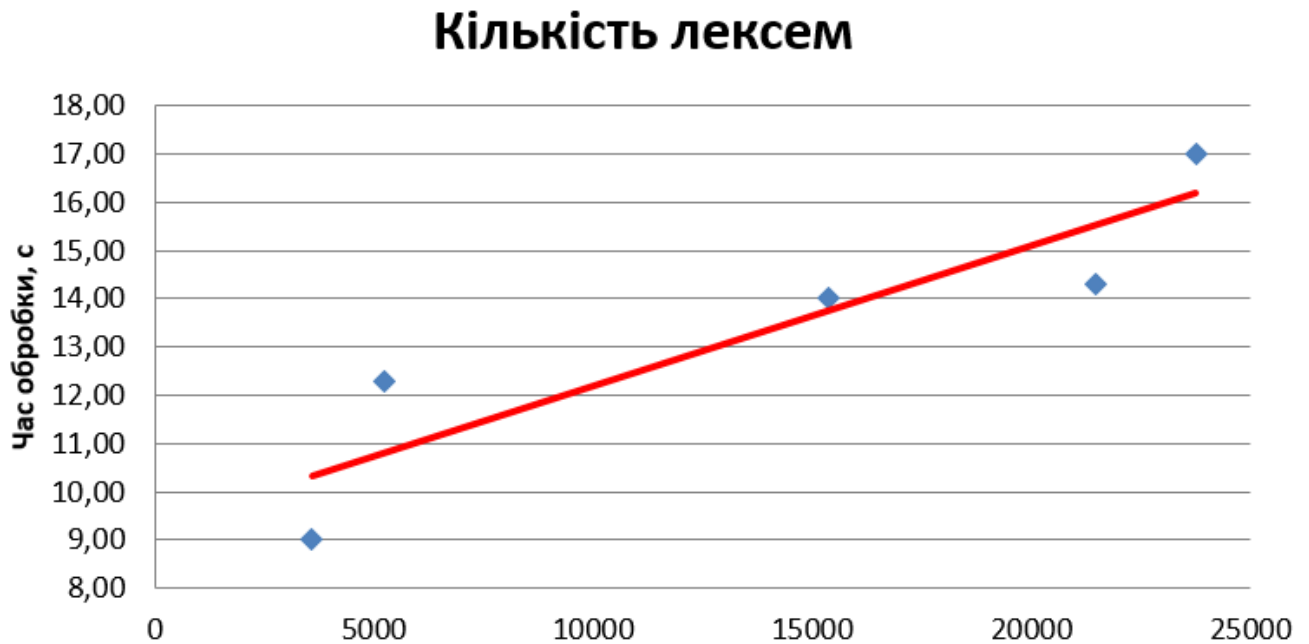


Рис. 4.45. Швидкість обробки вхідних даних для ГІС-додатку «Культурна спадщина Тараса Григоровича Шевченка»

Висновки за розділом 4

Розроблено інструментальний засіб «Рекурсивний редуктор», що дозволяє аналізувати ПМ тексти і формувати їх структуризоване представлення в автоматичному або автоматизованому режимі.

Розроблено інструментальний засіб «ТОДОС-ГІС», що дозволяє виконувати обробку і відображення структурованої інформації, зокрема, отриманої а результаті структуризації природномовних текстів. В склад даного інструментального засобу входять модулі, що надають можливість відображення геопросторової інформації у вигляді ГІС-додатку, і дозволяють формувати трансдисциплінарне представлення інформації, а в сукупності вирішують задачу трансдисциплінарного представлення геопросторової інформації.

Проведено ряд обчислювальних експериментів, в ході яких побудовано онтологічні ГІС-додатки на основі наявних у відкритому доступі або наданих організаціями масивів документів. В ході експериментів показано високу

швидкодію і достатньо високу точність розроблених інструментальних засобів, що дозволяє використовувати їх в процесах оперативного управління.

Основні результати розділу опубліковано в [2, 3, 17, 27, 31, 32, 72, 76–80].

ВИСНОВКИ

В процесі виконання дисертаційного дослідження розроблено необхідні концептуальні та системотехнічні засади трансдисциплінарного представлення геопросторової інформації у вигляді онтологічних ГІС-додатків. Основні результати роботи є такі:

1. Проведено аналіз існуючих ГІС та програмних засобів, призначених для роботи з онтологіями, який показав, що існує потреба в створенні програмного комплексу, що міг би виконувати структурування ПМ текстів і трансдисциплінарне представлення отриманої таким чином геопросторової інформації у формі інтерактивних документів та у вигляді онтологічних ГІС-додатків.
2. Розроблено метод структурування документів шляхом рекурсивного застосування процедури редукції тексту на основі правил, що формуються користувачем.
3. Розроблено моделі трансдисциплінарного представлення інформації та онтологічного інтерактивного документу, на основі яких сформовано модель онтологічного ГІС-додатку і програмний модуль, призначений для формування таких додатків.
4. На основі розроблених моделей і методів реалізовано, введено в експлуатацію та інтегровано в систему «ТОДОС» програмний комплекс «Рекурсивний редуктор» і модуль «ТОДОС-ГІС», що дозволяють виконувати структурування великих масивів слабо- і неструктурованих документів, а також трансдисциплінарне представлення наявної в них геопросторової інформації у вигляді онтологічного ГІС-додатку.
5. Виконано реалізацію тематичної бібліотеки інтерактивних документів і її розширення – тематичну бібліотеку онтологічних ГІС-додатків, що дозволяють розміщувати структуровані документи в веб-орієнтованому середовищі.
6. Експериментально підтверджено ефективність роботи запропонованих моделей та засобів шляхом обробки з їх допомогою великих масивів текстової і табличної

інформації і автоматизованого формування на їх основі бібліотек онтологічних ГІС-додатків.

Результати досліджень можуть бути використані для підвищення ефективності обробки великих інформаційних текстових масивів експертами, а також в процесах управління для оперативного аналізу потоків даних при реалізації різноманітних інфраструктурних проектів великої складності.

Результати впровадження підтверджено відповідними актами.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Андон Ф. И., Гришанова И. Ю., Резниченко В. А. Semantic Web как новая модель информационного пространства интернет. *Проблемы программирования.*, 2008. №. 2–4. С. 417–430.
2. Андрущенко Т. І., Величко В. Ю., Гальченко С. А., Глоба Л. С., Гуляєв К. Д., Клімова Е. Я., Комова О. Б., Лісовий О. В., Попова М. А., Приходнюк В. В., Стрижак О. Є., Стус Д. М. Методики написання наукових робіт на основі онтологічного аналізу текстів : методичний посібник. К. : ТОВ «СІТІПРІНТ», 2013. 124 с.
3. Андрущенко Т. І., Довгий С. О., Глоба Л. С., Стрижак О. Є., Величко В. Ю., Гальченко С. А., Дем'яненко В. Б., Ільченко М. Ю., Лісовий О. В., Новогрудська Р. Л., Петренко М. Г., Приходнюк В. В., Слюсаренко О. О., Трофимчук О. М., Яценко Г. Ю. Засоби доступу до джерел знань та їх використання в навчальному процесі педагогічного навчального закладу : Методичні рекомендації. 2-ге вид., доповн. К. : Інститут обдарованої дитини, 2012. 192 с.
4. Артемьева И. Л., Рештаненко Н. В. Интеллектуальная система, основанная на многоуровневой онтологии химии. *Программные продукты и системы*, 2008. №. 1. С. 84–87.
5. Барендрегт Х. Ламбда-исчисление. Его синтаксис и семантика: Пер. с англ. М. : Мир, 1985. 606 с.
6. Беклемишев Д. В. Курс аналитической геометрии и линейной алгебры. М. : Высшая школа, 1998. 320 с.
7. Большакова Е. И., Васильева Н. Э. Формализация лексико-синтаксической информации для распознавания регулярных конструкций естественного языка. *Программные продукты и системы*, 2008. №. 4. С. 103–106.
8. Боровикова О. И., Загоруйко Ю. А. Организация порталов знаний на основе онтологий. *Диалог '2002* : Труды международного семинара. Протвино, 2002. С. 76–82.
9. Буч Г. Объектно-ориентированное проектирование с примерами применения: Пер. с англ. М. : Конкорд, 1992. 519 с.

10. Вайсфельд М. Объектно-ориентированное мышление. СПб.: Питер, 2004. 304 с.
11. Валькман Ю. Р., Гриценко В. И., Рыхальский А. Ю. Модельно-параметрическое пространство: теория и применение. К. : Наукова думка, 2012. 192 с.
12. Варламов А. А., Гальченко С. А. Земельный кадастр. Т6. Географические и земельные информационные системы. М. : Колос, 2006. 400 с.
13. Величко В. Ю., Малахов К. С., Семенов В. В., Стрижак А. Е. Комплексные инструментальные средства инженерии онтологий. *Information Models and Analyses*, 2014. №. 3. С. 336–361.
14. Величко В. Ю., Попова М. А., Приходнюк В. В., Стрижак О. Є. ТОДОС – ІТ-платформа формування трансдисциплінарних інформаційних середовищ. *Системи озброєння і військова техніка*, 2017. №. 1(49). С. 10–19.
15. Величко В. Ю., Приходнюк В. В. Деякі способи виділення відношень між термінами в природномовному тексті. *Системний аналіз та інформаційні технології* : зб. наук. праць за матеріалами XV конференції (м. Київ, 27 - 31 травня 2013 р.). К. : НТУУ «КПІ», 2013. С. 406.
16. Величко В. Ю., Приходнюк В. В. Спосіб автоматизованого виділення відношень між термінами з природномовних текстів технічної тематики. *Knowledge – Dialogue – Solution* : Збірник праць XX міжнародної конференції. К. : ІТНЕА, 2014. С. 27–28.
17. Величко В. Ю., Сирота С. В., Приходнюк В. В. Інструментарій автоматизованого виділення відношень з текстів технічної тематики. *Системний аналіз та інформаційні технології* : зб. наук. праць за матеріалами XVI конференції (м. Київ, 26 - 30 травня 2014 р.). К. : НТУУ «КПІ», 2014. С. 348.
18. Величко В., Волошин П., Свитла С. Автоматизированное создание тезауруса терминов предметной области для локальных поисковых систем. *Knowledge – Dialogue – Solution* : International Book Series “INFORMATION SCIENCE & COMPUTING”. Sofia, Bulgaria : ІТНЕА, 2009. №. 15. С. 24–31.
19. Величко В., Приходнюк В., Стрижак А., Марков К., Иванова К., Карастанев С. Построение таксономии документов для формирования иерархических слоев в геоинформационных системах. *Information Content and Processing*, 2015. С. 181–

199.

20. Вероятностный морфологический анализатор русского и украинского языков Stemka. URL: <http://www.keva.ru/stemka/stemka.html> (дата звернення: 12.10.2016).
21. Гаврилова Т. А., Хорошевский В. Ф. Базы знаний интеллектуальных систем. СПб.: Питер, 2001. 384 с.
22. Гладун А. Я., Рогушина Ю. В. Онтологии в корпоративных сетях. URL: <http://www.management.com.ua/ims/ims115.html> (дата звернення: 10.06.2017).
23. Гладун В. П., Величко В. Ю. Конспектирование естественных языковых текстов. *Knowledge – Dialogue – Solution : Proceedings of the XI-th International Conference*, 2005. С. 344–347.
24. Гладун В. П. Партнерство с компьютером. Киев : «Port-Royal», 2000. 128 с.
25. Гладун В. П. Процессы формирования новых знаний. София : СД «Педагог 6», 1994. 192 с.
26. Голод П. І. Симетрія та методи теорії груп у фізиці (дискретні симетрії). К. : Києво-Могилянська академія, 2005. 215 с.
27. Горборуков В. В., Приходнюк В. В., Стрижак О. Є., Франчук О. В. Система оцінювання інтелектуальних досягнень учнівської молоді. Онтологічний підхід. *Інтелектуальний аналіз інформації* : Збірник праць XVI Міжнародної наукової конференції ім. Т. А.Таран. К. : Просвіта, 2016. С. 36–42.
28. Горборуков В. В., Стрижак О. Є., Франчук О. В. Використання онтологій у системах підтримки прийняття рішень. *Математичне моделювання в економіці* : Збірник наукових праць / редкол. : С. О. Довгий (голов. ред.) [та ін.]. К. : НАН України, Інститут телекомунікацій і глобального інформаційного простору, Інститут економіки та прогнозування, Інститут кібернетики ім. В. М. Глушкова, 2013. С. 33–39.
29. Гуржій А. М., Стрижак О. Є. Онтологічні інструменти управління мережевими інформаційними ресурсами та їх використання в освітній та науковій діяльності. *Наукові записки Малої академії наук України* : Збірник наукових праць. Серія : Педагогічні науки. К. : Національний центр «Мала академія наук України»,

2013. С. 427–434.

30. Добров Б. В., Иванов В. В., Лукашевич Н. В., Соловьев В. Д. Онтологии и тезаурусы: модели, инструменты, приложения: учебное пособие. М. : Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2009. 173 с.
31. Довгий С. О., Попова М. А., Приходнюк В. В., Стрижак О. Є., Яцухно В. А. Реєстр архівних документів, пов'язаних з життям, творчістю та вшануванням пам'яті Тараса Григоровича Шевченка: Монографія. К.: Центр розвитку особистості «УНІКУМ», 2017. 250 с.
32. Довгий С. О., Стрижак О. Є., Андрущенко Т. І., Гальченко С. А., Глоба Л. С., Гончар А. В., Копійка О. В., Кудляк В. М., Ляшук К. В., Попова М. А., Приходнюк В. В., Семенюк Г. Ф., Трофимчук О. М., Поліхун Н. І., Постова К. Г. Онтологічний кабінет дослідження життя та творчості Тараса Шевченка в середовищі науково-освітнього порталу KOBZAR.UA: Монографія. К.: Інститут обдарованої дитини, 2016. 175 с.
33. Ермаков А. Е. Метаязык описания грамматики в синтаксических анализаторах естественного языка на основе объектно-ориентированного языка программирования. *Программная инженерия*, 2015. №. 12. С. 12–21.
34. Ермаков А. Е., Плешко В. В. Семантическая интерпретация в системах компьютерного анализа текста. *Информационные технологии*, 2009. №. 6. С. 2–7.
35. Ермаков А. Е. Язык семантических трансформаций для компьютерной интерпретации текста. *Информационные технологии*, 2017. №. 23. С. 403–412.
36. Загоруйко Ю. А. Автоматизация сбора онтологической информации об интернет-ресурсах для портала научных знаний. *Известия Томского политехнического университета*. Управление, вычислительная техника и информатика, 2008. С. 114–119.
37. Извлечение объектов - ООО Диктум. URL: <http://dictum.ru/ru/named-entities-extraction/blog> (дата звернення: 21.11.2015).
38. Киященко Л. П., Гребенщикова Е. Г. Современная философия науки:

трансдисциплинарные аспекты. М.: МГМСУ, 2011. 172 с.

39. Киященко Л. П., Моисеев В. И. Философия трансдисциплинарности. М. : ИФРАН, 2009. .
40. Клещев А. С., Артемьева И. Л. Необогатенные системы логических соотношений. *Научно-техническая информация*, 2000. №. 7–8. С. 18–28.
41. Клещев А. С., Шалфеева Е. А. Классификация свойств онтологий. Онтологии и их классификации. Владивосток : ИАПУ ДВО РАН, 2005. 19 с.
42. Клини С. К. Введение в метаматерику. М. : Иностранная литература, 1957. 526 с.
43. Князева Е. Н. Трансдисциплинарные стратегии исследований. *Вестник ТГПУ*, 2011. №. 10. С. 193–201.
44. Колмогоров А. Н., Драгалев А. Г. Математическая логика. М.: УРСС, 2005. 240 с.
45. Константайл Л., Локвуд Л. Разработка программного обеспечения. СПб.: Питер, 2004. 592 с.
46. Кормалев Д. А., Куршев Е. П. Развитие языка правил извлечения информации в системе ИСИДА-Т. *Программные системы: теория и приложения* : Труды международной конференции. М. : Физматлит, 2006. С. 365–377.
47. Коршунова С. О. Роль тезаурусного моделирования в организации терминополья «ТЕХТ-ТЕКСТ». *Вестник Иркутского государственного лингвистического университета*, 2009. №. 1. С. 116–123.
48. Котурова М. П. Стилистика научной речи. М. : Академия, 2010. 240 с.
49. Ланде Д. В. Поиск знаний в Internet. Профессиональная работа. М. : Издательский дом «Вильямс», 2005. 272 с.
50. Лосский Н. О. Логика. Рипол Классик, 1928. 167 с.
51. Мазур А. А., Надутенко М. В., Остапова И. В., Петрук В. С., Пустовойт С. В., Широков В. А. Формування національної термінологічної системи в галузі зварювання на основі віртуальних лексикографічних лабораторій. *Вісник Національної Академії наук України*. К. : НАН України, 2014. С. 75–83.
52. Малишевский А. В. Качественные модели в теории сложных систем. М.: Наука.

Физматлит, 1998. 528 с.

53. Морфологический анализатор pymorphy2. URL: <https://pymorphy2.readthedocs.io/en/latest/> (дата звернення: 05.04.2016).
54. Невзорова О. А., Федунев Б. Е. Система анализа технических текстов «ЛоТА»: основные концепции и проектные решения. *Изв. РАН. Теория и системы управления*, 2001. №. 3. С. 138–149.
55. Олецкий О. В. Застосування графу «онтологія–документ» до задачі інтелектуального аналізу поведінки відвідувачів веб-ресурсів. *Наукові записки НаУКМА, Комп'ютерні науки*, 2011. №. 125. С. 90–92.
56. Онлайн-сервіс пошуку та планування техніки і роботи на полях. URL: <https://agrocounrty.com> (дата звернення: 30.11.2017).
57. Палагин А. В., Величко В. Ю., Стрижак А. Е., Попова М. А. Инструменты поддержки процессов аналитической деятельности эксперта при тематическом исследовании информационных ресурсов и источников. *Information Technologies and Knowledge*, 2010. №. 4. С. 329 – 347.
58. Палагин А. В., Крытый С. Л., Величко В. Ю., Петренко Н. Г. К анализу естественно-языковых объектов. «*Intelligent Processing*» *International Book Series “INFORMATION SCIENCE & COMPUTING”*: Supplement to the International Journal “INFORMATION TECHNOLOGIES & KNOWLEDGE”. Sofia, Bulgaria : ITNEA, 2009. №. 3. С. 36–43.
59. Палагин А. В., Крытый С. Л., Петренко Н. Г. Знание-ориентированные информационные системы с обработкой естественно-языковых объектов: основы методологии и архитектурно-структурная организация. *УСiМ*, 2009. №. 3. С. 42–55.
60. Палагин А. В., Крытый С. Л., Петренко Н. Г. Онтологические методы и средства обработки предметных знаний: монография. Луганск: изд-во ВНУ им. В. Даля, 2012. 324 с.
61. Палагин А. В., Петренко Н. Г., Величко В. Ю., Малахов К. С. Развитие формальных моделей, алгоритмов, процедур, разработки и функционирования программной системы “Инструментальный комплекс онтологического

- назначения”. *Інтелектуальні інформаційні технології*, 2014. №. 2–3. С. 221–232.
62. Палагин А. В., Петренко Н. Г., Величко В. Ю., Малахов К. С., Тихонов Ю. Л. К вопросу разработки инструментального комплекса онтологического назначения. *Моделі та засоби систем баз даних і знань*, 2012. №. 2–3. С. 289–298.
63. Палагин А. В., Петренко Н. Г. К вопросу системно-онтологической интеграции знаний предметной области. *Математические машины и системы*, 2007. №. 3,4. С. 63–75.
64. Палагин А. В., Петренко Н. Г. К проектированию онтологоуправляемой информационной системы с обработкой естественно-языковых объектов. *Математические машины и системы*, 2008. №. 2. С. 14–23.
65. Палагин А. В., Риппа С. П., Саченко А. А. Концептуализация и проблематика онтологий. «Искусственный интеллект», №3, 2008. 374 с.
66. Палагин А. В. Трансдисциплинарность, информатика и развитие современной цивилизации. *Вісник Національної Академії наук України*. К. : НАН України, 2014. №. 7. С. 25–33.
67. Палагин А. В., Яковлев Ю. С. Системная интеграция средств компьютерной техники. Винница : УНІВЕРСУМ, 2005. 680 с.
68. Палагин А. В., Петренко Н. Г. Системно-онтологический анализ предметной области. *УСuM*, 2009. №. 4. С. 3–14.
69. Попова М. А. Методика формування та використання комп’ютерних онтологій в галузі екологічної освіти : [монографія]. К. : «СІТІПРІНТ», 2013. 200 с.
70. Попова М. А. Модель онтологического интерфейса агрегации информационных ресурсов и средств ГИС. *Information Technologies and Knowledge*, 2013. №. 7. С. 362–370.
71. Попова М. А. Онтологія взаємодії в середовищі геоінформаційної системи: дис. канд. техн. наук. К., 2014. 250 с.
72. Попова М. А., Приходнюк В. В. Формування тематичних ГІС на основі семантиколінгвістичної обробки документів. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в*

- надзвичайних ситуаціях* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Київ, Пуща-Водиця, 3 - 6 жовтня 2016 р.). К. ; Х. : НАН України, Ін-т телекомунікацій і глоб. інформ. простору [та ін.], 2016. С. 130–133.
73. Попова М. А., Стрижак О. Є. Онтологічний інтерфейс як засіб представлення інформаційних ресурсів в ГІС-середовищі. *Ученые записки Таврического национального университета имени В. И. Вернадского, Серия : География*, 2013. №. 26 (65). С. 127–135.
 74. Поспелов Д. А. Ситуационное управление: теория и практика. М. : Наука, 1986. 288 с.
 75. Приходнюк В. В. Автоматизоване виділення інформації з текстів на основі правил, представлених у форматі λ -виразів. *Управління знаннями та конкурентна розвідка* : зб. наук. праць за матеріалами конференції (м. Харків, 19 - 21 квітня 2016 р.). Х. : Харківський національний університет радіоелектроніки, 2016. С. 48–49.
 76. Приходнюк В. В. Автоматизоване формування електронних шарів геоінформаційних систем на основі структурованої і неструктурованої інформації. *Геоінформаційні технології в територіальному управлінні* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Одеса, 17 - 18 вересня 2015 р.). Одеса : Одеський регіональний інститут державного управління [та ін.], 2015. С. 73–76.
 77. Приходнюк В. В. Онтолого-орієнтований інтерфейс відображення додатків в ГІС-середовищі. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Київ, Пуща-Водиця, 3 - 6 жовтня 2016 р.). К. ; Х : НАН України, Ін-т телекомунікацій і глоб. інформ. простору [та ін.], 2014. С. 127–129.
 78. Приходнюк В. В., Стрижак О. Є., Лебідь О. Г. Онтологічне представлення функціональності систем. *Екологічна безпека та природокористування* : Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний

університет будівництва і архітектури [та ін.], 2016. №. 3–4 (22). С. 5–23.

79. Приходнюк В. В., Стрижак О. Є. Множинні характеристики онтологічних систем. *Математичне моделювання в економіці* : Збірник наукових праць / редкол. : С. О. Довгий (голов. ред.) [та ін.]. К. : НАН України Інститут телекомунікацій і глобального інформаційного простору, Інститут економіки та прогнозування, Інститут кібернетики ім. В. М. Глушкова, 2017. №. 8. С. 47–61.
80. Приходнюк В. В., Стрижак О. Є. Онтологічна ГІС, як засіб впорядкування геопросторової інформації. *Наука і техніка Повітряних Сил Збройних Сил України*, 2017. №. 2(27). С. 167–174.
81. Приходнюк В. Таксономизация естественно-языковых текстов. *Information Models and Analyses*, 2016. №. 5. С. 270–284.
82. Реєстр сільськогосподарських дорадників і експертів-дорадників. URL: <http://minagro.gov.ua/node/3511> (дата звернення: 30.11.2017).
83. Реєстр сільськогосподарських дорадчих служб. URL: <http://minagro.gov.ua/node/3512> (дата звернення: 30.11.2017).
84. Русская и английская морфология для Windows и Linux. URL: <http://www.solarix.ru/grammatical-dictionary-api.shtml> (дата звернення: 10.10.2016).
85. Рыженко Л. И. Когнитивный инжиниринг. Омск : СибАДИ, 2012. 172 с.
86. Світличний О. О., Злотницький С. В. Основи геоінформатики: Навчальний посібник. Суми: Університетська книга, 2006. 295 с.
87. Стрижак А. Е. Инвариантные задачи онтологических систем. *INFORMATION TECHNOLOGIES & KNOWLEDGE*, 2014. №. 8. С. 356–360.
88. Стрижак А. Е. Онтологические аспекты трансдисциплинарной интеграции информационных ресурсов. *Открытые информационные и компьютерные интегрированные технологии*, 2014. №. 65. С. 211–223.
89. Стрижак О. Є., Горборуков В. В., Франчук О. В., Попова М. А. Онтологія задачі вибору та її застосування при аналізі лімнологічних систем. *Екологічна безпека та природокористування* : Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Київ : Міністерство освіти і науки

України, Київський національний університет будівництва і архітектури, НАН України Інститут телекомунікацій і глобального інформаційного простору., 2014. С. 172–183.

90. Стрижак О. Є. Трансдисциплінарна інтеграція інформаційних ресурсів: дис. д-ра техн. наук. К., 2014. 470 с.
91. Тикунов В. И. Геоинформатика. Академический учебник. М. : Академия, 2006. 540 с.
92. Цветков В. Я. Геоинформационные системы и технологии. М. : Финансы и статистика, 1998. 286 с.
93. Шаталкин А. И. Таксономия. Основания, принципы и правила. М. : Товарищество научных изданий КМК, 2012. 600 с.
94. Широков В. А. Язык. Информация. Система. Palamarium Academic Publishing, 2017. 270 с.
95. Яловець А. Л. Представление и обработка знаний с точки зрения математического моделирования. Пролемы и решения. К. : Наукова думка, 2011. 359 с.
96. Ahmed I., Fayyaz A., Shahzad A. PostgreSQL Developer's Guide. Packt Publishing, 2015. 468 с.
97. Alani H., Kim S., Millard D., Weal M. J., Hall W., Lewis P., Shadbolt N. Using Protege for automatic ontology instantiation. *7th International Protégé Conference*, 2004. .
98. Anisimovich K. V., Druzhkin K. Y., Zuev K. A., Minlos F. R., Petrova M. A., Selegei V. P. Syntactic and semantic parser based on ABBYY Compreno linguistic technologies. *ДИАЛОГ-2012 : материалы XVIII международной конференции* (Бекасово, 30 мая-03 июня 2012 г.). Российский государственный гуманитарный университет, 2012. С. 91–103.
99. Apache Lucene - Apache Lucene Core. URL: <https://lucene.apache.org/core/> (дата звернення: 17.12.2016).
100. ArcGIS API for JavaScript. URL: <https://developers.arcgis.com/javascript/> (дата звернення: 25.06.2017).
101. Bontcheva V., Tablan D., Maynard, Cunningham H. Evolving GATE to Meet New

- Challenges in Language Engineering. *Natural Language Engineering*, 2004. №. 10 (3/4). С. 349–373.
102. Booch G., Rumbaugh J., Jacobson I. The Unified Modeling Language User Guide. Reading, MA : Addison-Wesley, 1999. 482 с.
 103. Broy M. Mathematical Methods in System and Software Engineering. *NATO ASI SERIES F COMPUTER AND SYSTEMS SCIENCES*. Berlin : Springer, 1997. №. 158. С. 271–312.
 104. Broy M. Mathematical System Models as a Basis of Software Engineering. Berlin : Springer, 1995. 292 с.
 105. Chadwick J., Snyder T., Panda H. Programming ASP.NET MVC 4. O'Reilly Media, 2012. 492 с.
 106. Cunningham H. GATE, a General Architecture for Text Engineering. *Computers and the Humanities*, 2002. №. 36. С. 223–254.
 107. Cunningham H., Maynard D., Bontcheva K., Tablan V. Developing Language Processing Components with GATE Version 8. University of Sheffield Department of Computer Science, 2014. .
 108. Deployment Patterns (Microsoft Enterprise Architecture, Patterns, and Practices). URL: <http://msdn.microsoft.com/en-%0Aus/library/ff646997.aspx> (дата звернення: 07.12.2016).
 109. Drummond N., Horridge M., Knublauch H. Protégé-OWL Tutorial. *8th International Protégé Conference*, 2005. .
 110. Eckerson W. Three Tier Client/Server Architectures: Achieving Scalability, Performance, and Efficiency in Client Server Applications. *Open Information Systems*, 1995. №. 3. С. 46–50.
 111. Endeca Information Discovery - Overview. URL: <http://www.oracle.com/technetwork/middleware/endeca/overview/index.html> (дата звернення: 15.11.2015).
 112. Esri: GIS Mapping Software, Spatial Data Analytics & Location Platform. URL: <https://www.esri.com/en-us/home> (дата звернення: 08.12.2016).
 113. EXALEAD™ Data in Business Solution - Dassault Systèmes®. URL:

- <https://www.3ds.com/products-services/exalead/> (дата звернення: 21.10.2016).
114. Fowler M. Patterns of Enterprise Application Architecture. Addison Wesley, 2003. 560 с.
 115. García-Castro R., Gómez-Pérez A. Interoperability of Protégé using RDF (S) as interchange language. *9th International Protégé Conference*, 2006. .
 116. Get business insights in minutes | IBM Watson Analytics. URL: <https://www.ibm.com/watson-analytics> (дата звернення: 04.11.2016).
 117. GitHub: Yandex Tomita-parser. URL: <https://github.com/yandex/tomita-parser> (дата звернення: 17.12.2016).
 118. Google Maps APIs | Google Developers. URL: <https://developers.google.com/maps/?hl=ru> (дата звернення: 25.06.2017).
 119. Gruber T. R. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 1993. №. 5. С. 199 – 220.
 120. Guarino N. Understanding, building and using ontologies. *Human-Computer Studies*, 1997. С. 293–310.
 121. Horst H. J. Completeness, decidability and complexity of entailment for RDF Schema and a semantic extension involving the OWL vocabulary. *J. of Web Semantics*, 2005. №. 3. С. 79–115.
 122. Ian W., Elbe F. Data Mining: Practical Machine Learning Tools and Techniques. San Francisco : Morgan Kaufmann publishers, 2005. 525 с.
 123. Intergraph Corporation | Process, Power and Marine | Hexagon Safety & Infrastructure. URL: <http://www.intergraph.com> (дата звернення: 08.12.2016).
 124. Introduction to XML. URL: https://www.w3schools.com/xml/xml_what_is.asp (дата звернення: 10.05.2017).
 125. ISO/IEC 15504-1:2004. Information technology – Process assessment. Part 1: Concepts and vocabulary.
 126. Keyhole Markup Language. URL: <https://developers.google.com/kml/> (дата звернення: 12.07.2017).
 127. Klein J. T. Interdisciplinarity: History, Theory, and Practice. Wayne State University Press, 1990. .

128. Klein J. T. Transdisciplinarity: Joint Problem Solving Among Science, Technology, and Society: An Effective Way for Managing Complexity. Birkhäuser, 2001. .
129. KNIME Analytics Platform | KNIME. URL: <https://www.knime.com/knime-analytics-platform> (дата звернення: 15.12.2016).
130. Leaflet – an open-source JavaScript library for interactive maps. URL: <http://leafletjs.com> (дата звернення: 25.06.2017).
131. Link Grammar. URL: <http://www.link.cs.cmu.edu/link/> (дата звернення: 20.10.2016).
132. Marshall C. Enterprise Modelling with UML. Reading, MA : Addison-Wesley, 2000. 288 с.
133. Maynard D., Bontcheva K., Augenstein I. Natural Language Processing for the Semantic Web. Morgan and Claypool, 2016. 194 с.
134. Melo F., Martins B. Automated Geocoding of Textual Documents: A Survey of Current Approaches. *Transactions in GIS*, 2017. .
135. MicroStation - ПО САПР для разработки в 3D архитектурных и инженерных проектов. URL: <https://www.bentley.com/ru/products/brands/microstation> (дата звернення: 08.12.2016).
136. Munzner T. Interactive visualization of large graphs and networks: Ph. D. thesis. Department of Computer Science, Stanford University, 2000. 175 с.
137. Nicolescu B. Transdisciplinarity - Theory and Practice. Hampton Press, Cresskill, NJ, USA, 2008. 320 с.
138. NIMA Department of Defense World Geodetic System 1984. *Technical Report 8350.2, Third Edition*, 2000. .
139. NLnet; AGFL; Description. URL: <https://nlnet.nl/project/agfl/description.html> (дата звернення: 12.10.2015).
140. Palantir Gotham | Palantir. URL: <https://www.palantir.com/palantir-gotham/> (дата звернення: 13.11.2016).
141. Quatrani T. Visual Modeling with Rational Rose and UML. Reading, MA : Addison-Wesley, 1998. 222 с.
142. RapidMiner: Data Science Platform. URL: <https://rapidminer.com/> (дата звернення:

12.12.2016).

143. Rijsbergen C. J. van Information retrieval. *Information Retrieval: Searching in the 21st Century*. London, GB; Boston, MA : Butterworth, 1979. 208 с.
144. Spatial and Geographic objects for PostgreSQL. URL: <http://postgis.net> (дата звернення: 12.05.2017).
145. Spatial reference for EPSG:4326. URL: <https://epsg.io/4326> (дата звернення: 20.08.2017).
146. Sphinx | Open Source Search Engine. URL: <http://sphinxsearch.com> (дата звернення: 02.10.2016).
147. Sure Y., Erdmann M., Angele J., Staab S., Studer R., Wenke D. OntoEdit: Collaborative Ontology Development for the Semantic Web. *The Semantic Web*. Berlin : Springer, 2002. №. 2342. С. 221–235.
148. The Google Geocoding API. URL: <https://developers.google.com/maps/documentation/geocoding/intro?hl=ru> (дата звернення: 22.07.2017).
149. Volckmann R. Transdisciplinarity: Basarab Nicolescu Talks with Russ Volckmann. *Integral review*, 2007. №. 4. С. 73–90.
150. World Declaration on Higher Education for the Twenty-First Century: Vision and Action. URL: http://www.unesco.org/education/educprog/wche/declaration_eng.htm (дата звернення: 02.12.2016).

ДОДАТКИ

Додаток А. Акти впровадження

Затверджую

“ ” березень 2017 р.
Генеральний директор ТОВ «ЕСОММ Со»

М. М. Мазуренко

АКТ

впровадження науково-технічної продукції в процес розробки ГІС-проектів

1. **Найменування пропозиції для впровадження:** Науково-технологічні та методичні засоби ІТ-технології ТОДОС, щодо трансдисциплінарного представлення геопросторової інформації, для створення у середовищі ГІС-систем та ГЕОПОРТАЛІВ онтологічних інформаційних додатків. Створення ГІС-орієнтованих, онтологічних інформаційно-аналітичних систем та середовищ щодо моделювання компонентів інформаційної системи управління інфраструктурними проектами.
2. **Ким запропоновано, адреса:** відділ онтологічних систем та прикладної алгебраїчної комбінаторики Інституту телекомунікацій і глобального інформаційного простору НАН України; адреса: Чоколівський б-р, 13, 03186, м. Київ.
3. **Виконавець:** ПРИХОДНЮК Віталій Валерійович.
4. **Джерело інформації:**
 - Построение таксономии документов для формирования иерархических слоев в геоинформационных системах [Текст]: / [Приходнюк В.В. та ін.] // International Journal "Information Content and Processing", Volume 2, Number 2, 2015 – с. 181-199.
 - Онтологічне представлення функціональності систем / В.В. Приходнюк, О.Є. Стрижак, О.Г. Лебідь // Екологічна безпека та природокористування: Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний університет будівництва і архітектури, НАН України Інститут телекомунікацій і глобального інформаційного простору. – Київ, 2016. – Вип. № 3–4 (22), стр. 5-23
5. **Портал мережних онтологічних інструментів:** <http://editor4.inhost.com.ua>
6. **Реєстр онтологій:** <http://ontology4.inhost.com.ua/>
7. **Терміни впровадження:** 2015 – 2017 рр.
8. **Ефективність впровадження:** Підвищення якості й ефективності інформаційно-аналітичного процесу на основі використання онтологічних засобів семантичного аналізу природно-мовних текстів, які репрезентують геопросторову інформацію. Підвищення якості й ефективності прийняття рішень в процесі управління інфраструктурними проектами на основі засобів адаптивної генерації тематичних категорій. Генерація онтологічних категорій меню WEB-семантичного інтерфейсу онтологічних ГІС-додатків. Створення знансво-орієнтовних систем у середовищі геопорталів.
9. **Зауваження, пропозиції:** Розповсюдження отриманих позитивних результатів впровадження шляхом застосування технологічних засобів трансдисциплінарного представлення геопросторової інформації з метою створення інтегрованих онтологічних ГІС-додатків та інформаційно-аналітичних середовищ для підвищення якості та ефективності управління інфраструктурними проектами з використанням ГІС-середовищ.

Відповідальний за впровадження:
Начальник відділу ТОВ «ЕСОММ Со»



Мальцев С. В.

Затверджую

“ ” травень 2017 р.
Генеральний директор
Національного музею Тараса Шевченка

Д. В. Стус



АКТ

впровадження науково-технічної продукції

1. **Найменування пропозиції для впровадження:** Науково-технологічні та методичні засоби ІТ-технології ТОДОС, щодо створення єдиного історико-культурного та інформаційно-освітнього простору на основі використання засобів трансдисциплінарного представлення геопросторової інформації, для створення у середовищі ГІС-систем онтологічних інформаційних додатків історико-культурного та освітнього профілю. Створення історико-культурних та освітніх ГІС-орієнтованих онтологічних інформаційно-аналітичних систем та середовищ щодо репрезентації у мережевому середовищі життєдіяльності Т.Г. Шевченко.
2. **Ким запропоновано, адреса:** відділ онтологічних систем та прикладної алгебраїчної комбінаторики Інституту телекомунікацій і глобального інформаційного простору НАН України; адреса: Чоколівський б-р, 13, 03186, м. Київ.
3. **Виконавець:** ПРИХОДНЮК Віталій Валерійович.
4. **Джерело інформації:**
 - Построение таксономии документов для формирования иерархических слоев в геоинформационных системах [Текст]: / [Приходнюк В.В. та ін.] // International Journal "Information Content and Processing", Volume 2, Number 2, 2015 – с. 181-199.
 - Онтологічне представлення функціональності систем / В.В. Приходнюк, О.Є. Стрижак, О.Г. Лебідь // Екологічна безпека та природокористування: Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний університет будівництва і архітектури, НАН України Інститут телекомунікацій і глобального інформаційного простору. – Київ, 2016. – Вип. № 3–4 (22), стр. 5-23
5. **Портал мережних онтологічних інструментів:** <http://editor4.inhost.com.ua>
6. **Реєстр онтологій:** <http://ontology4.inhost.com.ua/>
7. **Терміни впровадження:** 2015 – 2017 рр.
8. **Ефективність впровадження:** Підвищення якості й ефективності інформаційно-аналітичного процесу при використанні мережних інформаційних ресурсів історико-культурного та освітнього профілів, на основі використання онтологічних засобів семантичного аналізу природно-мовних текстів, які репрезентують геопросторову інформацію в історико-культурних текстах, особливо таких, які репрезентують життєдіяльність Т.Г. Шевченко. Підвищення якості й ефективності управління історико-культурними мережевими ресурсами різного тематичного напрямку на основі засобів адаптивної генерації тематичних категорій. Генерація онтологічних категорій меню WEB-семантичного інтерфейсу онтологічних ГІС-додатків.
9. **Зауваження, пропозиції:** Розповсюдження отриманих позитивних результатів впровадження щодо створення єдиного інформаційно-освітнього простору шляхом застосування технологічних засобів трансдисциплінарного представлення геопросторової інформації для створення історико-культурних та освітніх трансдисциплінарних онтологій і мережних баз знань з метою надання освітньо-інформаційних послуг на основі онтологічних, консультативних мережних площадок та використанні у навчальному процесі ЗОНЗах, ВНЗах України та ефективності управління інфраструктурними проектами в галузі культури та освіти.



ЗАТВЕРДЖУЮ

Ректор НМАПО імені П. Л. Шупика
академік НАМН України

професор _____ Ю. В. Вороненко

06.11. 2017 р.

АКТ

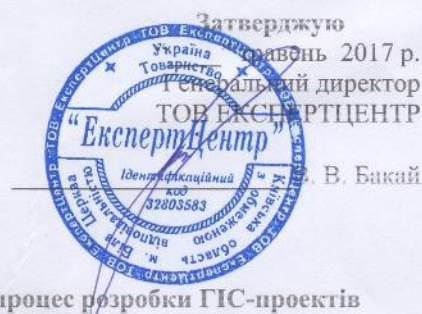
впровадження науково-технічної продукції

1. **Найменування пропозиції для впровадження:** науково-технологічні та методичні засоби ІТ-технології ТОДОС, щодо трансдисциплінарного представлення геопросторової інформації, для створення у середовищі ГІС-систем онтологічних інформаційних додатків медичного профілю. Створення ГІС-орієнтованих, онтологічних інформаційно-аналітичних систем та середовищ щодо моделювання компонентів інформаційної системи управління інфраструктурними проектами в галузі охорони здоров'я населення.
2. **Ким запропоновано, адреса:** відділ онтологічних систем та прикладної алгебраїчної комбінаторики Інституту телекомунікацій і глобального інформаційного простору НАН України; адреса: Чоколівський б-р, 13, м. Київ, 03186.
3. **Виконавець:** Приходнюк Віталій Валерійович.
4. **Джерело інформації:** 1) Построение таксономии документов для формирования иерархических слоев в геоинформационных системах [Текст] / В. В. Приходнюк, А. Е. Стрижак, В. Ю. Величко [и др.] // International Journal "Information Content and Processing". - 2015. - Vol. 2, № 2. - Р. 181-199.
2) Онтологічне представлення функціональності систем / В. В. Приходнюк, О. Є. Стрижак, О. Г. Лебідь // Екологічна безпека та природокористування: Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний університет будівництва і архітектури, НАН України Інститут телекомунікацій і глобального інформаційного простору. – Київ, 2016. – Вип. № 3–4 (22), С. 5-23.
5. **Портал мережних онтологічних інструментів:** <http://editor4.inhost.com.ua>
6. **Реєстр онтологій:** <http://ontology4.inhost.com.ua>
7. **Терміни впровадження:** 2015 – 2017 рр.
8. **Ефективність впровадження:** підвищення якості й ефективності інформаційно-аналітичного процесу на основі використання онтологічних засобів семантичного аналізу природно-мовних текстів, які репрезентують геопросторову інформацію. Підвищення якості й ефективності прийняття рішень в процесі управління інфраструктурними проектами на основі засобів адаптивної генерації тематичних категорій. Генерація онтологічних категорій меню WEB-семантичного інтерфейсу онтологічних ГІС-додатків. Створення знансво-орієнтовних систем у мережевому середовищі.
9. **Зауваження, пропозиції:** розповсюдження отриманих позитивних результатів впровадження шляхом застосування технологічних засобів трансдисциплінарного представлення геопросторової інформації для створення трансдисциплінарних медичних та фармацевтичних онтологій і мережних баз знань з метою надання інформаційно-консультаційних послуг лікарям і широкім колам населення на основі онтологічних, консультативних мережних площадок, а також використанні у навчальному процесі ВНЗах України медичного, фармацевтичного профілю, закладах медичної післядипломної освіти та в управлінні інфраструктурними проектами в галузі охорони здоров'я населення.

Відповідальний за впровадження:

Завідувач кафедри медичної інформатики
д.мед.н. професор

О. П. Мінцер



АКТ

впровадження науково-технічної продукції в процес розробки ГІС-проектів

1. **Найменування пропозиції для впровадження:** Науково-технологічні та методичні засоби ІТ-технології ТОДОС, щодо трандисциплінарного представлення геопросторової інформації, для створення у середовищі ГІС-систем та ГЕОПОРТАЛІВ онтологічних інформаційних додатків. Створення ГІС-орієнтованих, онтологічних інформаційно-аналітичних систем та середовищ щодо моделювання компонентів інформаційної системи управління інфраструктурними проектами.
2. **Ким запропоновано, адреса:** відділ онтологічних систем та прикладної алгебраїчної комбінаторики Інституту телекомунікацій і глобального інформаційного простору НАН України; адреса: Чоколівський б-р, 13, 03186, м. Київ.
3. **Виконавець:** ПРИХОДНЮК Віталій Валерійович.
4. **Джерело інформації:**
 - Построение таксономии документов для формирования иерархических слоев в геоинформационных системах [Текст]: / [Приходнюк В.В. та ін.] // International Journal "Information Content and Processing", Volume 2, Number 2, 2015 – с. 181-199.
 - Онтологічне представлення функціональності систем / В.В. Приходнюк, О.Є. Стрижак, О.Г. Лебідь // Екологічна безпека та природокористування: Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний університет будівництва і архітектури, НАН України Інститут телекомунікацій і глобального інформаційного простору. – Київ, 2016. – Вип. № 3–4 (22), стр. 5-23
5. **Портал мережних онтологічних інструментів:** <http://editor4.inhost.com.ua>
6. **Реєстр онтологій:** <http://ontology4.inhost.com.ua/>
7. **Терміни впровадження:** 2015 – 2017 рр.
8. **Ефективність впровадження:** Підвищення якості й ефективності інформаційно-аналітичного процесу на основі використання онтологічних засобів семантичного аналізу природно-мовних текстів, які репрезентують геопросторову інформацію. Підвищення якості й ефективності прийняття рішень в процесі управління інфраструктурними проектами на основі засобів адаптивної генерації тематичних категорій. Генерація онтологічних категорій меню WEB-семантичного інтерфейсу онтологічних ГІС-додатків. Створення знансво-орієнтованих систем у середовищі геопорталів.
9. **Зауваження, пропозиції:** Розповсюдження отриманих позитивних результатів впровадження шляхом застосування технологічних засобів трандисциплінарного представлення геопросторової інформації з метою створення інтегрованих онтологічних ГІС-додатків та інформаційно-аналітичних середовищ для підвищення якості та ефективності управління інфраструктурними проектами з використанням ГІС-середовищ.

Відповідальний за впровадження:

В.В. Приходнюк *Н.Н. Мавшин*

Затверджую

“___” вересень 2017 р.
Директор Одеського науково-дослідного інституту
судових експертиз Міністерства юстиції України

А.І. Ріпенко

АКТ

впровадження науково-технічної продукції

1. **Найменування пропозиції для впровадження:** Науково-технологічні та методичні засоби ІТ-технології ТОДОС, щодо трансдисциплінарного представлення геопросторової інформації, для створення у середовищі ГІС-систем та ГЕОПОРТАЛІВ онтологічних інформаційних додатків. Створення ГІС-орієнтованих, онтологічних інформаційно-аналітичних систем та середовищ для забезпечення процесів експертного аналізу та управління інфраструктурними проектами.
2. **Ким запропоновано, адреса:** відділ онтологічних систем та прикладної алгебраїчної комбінаторики Інституту телекомунікацій і глобального інформаційного простору НАН України; адреса: Чоколівський б-р, 13, 03186, м. Київ.
3. **Виконавець:** ПРИХОДНЮК Віталій Валерійович.
4. **Джерело інформації:**
 - Построение таксономии документов для формирования иерархических слоев в геоинформационных системах [Текст]: / [Приходнюк В.В. та ін.] // International Journal "Information Content and Processing", Volume 2, Number 2, 2015 – с. 181-199.
 - Онтологічне представлення функціональності систем / В.В. Приходнюк, О.Є. Стрижак, О.Г. Лебідь // Екологічна безпека та природокористування: Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний університет будівництва і архітектури, НАН України Інститут телекомунікацій і глобального інформаційного простору. – Київ, 2016. – Вип. № 3–4 (22), стр. 5-23
5. **Портал мережних онтологічних інструментів:** <http://editor4.inhost.com.ua>
6. **Реєстр онтологій:** <http://ontology4.inhost.com.ua/>
7. **Терміни впровадження:** 2017 р.
8. **Ефективність впровадження:** Підвищення якості й ефективності інформаційно-аналітичного процесу на основі використання онтологічних засобів семантичного аналізу природно-мовних текстів, які репрезентують геопросторову інформацію. Підвищення якості й ефективності прийняття рішень в процесі проведення експертного аналізу на основі засобів адаптивної генерації тематичних категорій. Генерація онтологічних категорій меню WEB-семантичного інтерфейсу онтологічних ГІС-додатків для забезпечення управління територіальним розвитком, експертній діяльності та судових експертиз..
9. **Зауваження, пропозиції:** Розповсюдження отриманих позитивних результатів впровадження шляхом застосування технологічних засобів трансдисциплінарного представлення геопросторової інформації з метою створення інтегрованих онтологічних ГІС-додатків та інформаційно-аналітичних середовищ для підвищення якості та ефективності управління інфраструктурними проектами з використанням ГІС-середовищ.

Відповідальний за впровадження:
Заступник директора з експертної роботи

А.А. Колосюк

АКТ
впровадження науково-технічної продукції в процес розробки ГІС-проектів

1. **Найменування пропозиції для впровадження:** Науково-технологічні та методичні засоби ІТ-технології ТОДОС, щодо трансдисциплінарного представлення геопросторової інформації, для створення у середовищі ГІС-систем та ГЕОПОРТАЛІВ онтологічних інформаційних додатків. Створення ГІС-орієнтованих, онтологічних інформаційно-аналітичних систем та середовищ щодо моделювання компонентів інформаційної системи управління інфраструктурними проектами у аграрному секторі.
2. **Ким запропоновано, адреса:** відділ онтологічних систем та прикладної алгебраїчної комбінаторики Інституту телекомунікацій і глобального інформаційного простору НАН України; адреса: Чоколівський б-р, 13, 03186, м. Київ.
3. **Виконавець:** ПРИХОДНЮК Віталій Валерійович.
4. **Джерело інформації:**
 - Построение таксономии документов для формирования иерархических слоев в геоинформационных системах [Текст]: / [Приходнюк В.В. та ін.] // International Journal "Information Content and Processing", Volume 2, Number 2, 2015 – с. 181-199.
 - Онтологічне представлення функціональності систем / В.В. Приходнюк, О.Є. Стрижак, О.Г. Лебідь // Екологічна безпека та природокористування: Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний університет будівництва і архітектури, НАН України Інститут телекомунікацій і глобального інформаційного простору. – Київ, 2016. – Вип. № 3–4 (22), стр. 5-23
5. **Портал мережних онтологічних інструментів:** <http://editor4.inhost.com.ua>
6. **Реєстр онтологій:** <http://ontology4.inhost.com.ua/>
7. **Терміни впровадження:** 2017 рік.
8. **Ефективність впровадження:** Підвищення якості й ефективності інформаційно-аналітичного процесу на основі використання онтологічних засобів семантичного аналізу природно-мовних текстів, які репрезентують геопросторову інформацію. Підвищення якості й ефективності прийняття рішень в процесі управління інфраструктурними проектами на основі засобів адаптивної генерації аграрних тематичних категорій. Генерація онтологічних категорій меню WEB-семантичного інтерфейсу онтологічних ГІС-додатків. Створення знансво-орієнтованих систем у середовищі геопорталів.
9. **Зауваження, пропозиції:** Розповсюдження отриманих позитивних результатів впровадження шляхом застосування технологічних засобів трансдисциплінарного представлення геопросторової інформації з метою створення інтегрованих онтологічних ГІС-додатків та інформаційно-аналітичних середовищ для підвищення якості та ефективності управління інфраструктурними проектами в агросекторі з використанням ГІС-середовищ.

Завідувач відділення агресурсів і
 використання меліорованих земель
 Інституту водних проблем і меліорації НААН
 доктор сільськогосподарських наук, професор,
 чл.-кор. НААН

 Ю.О. Тараріко

Підпис Ю.О. Тараріко засвідчую
 Завідувач відділу наукових кадрів, інфраструктури
 та правового забезпечення

 С.В. Єгорова





ВЕРХОВНА РАДА УКРАЇНИ

Комітет з питань будівництва,
містобудування і житлово-комунального господарства

01008, м. Київ-8, вул. М. Грушевського, 5, тел.: 255-46-15, факс: 255-41-11

№ 04-12/11-476

"27" квітня 2015 р.

Інститут телекомунікацій і глобального
інформаційного простору НАНУ

На виконання доручень Голови Верховної Ради України В.Б.Гройсмана від 07.04.2015 р. та 10.04.2015 р. створено Робочу групу з розгляду питання щодо обґрунтованості підвищення цін/тарифів на енергоносії та комунальні послуги.

Зважаючи на великий обсяг первинної інформації, наданої Національною комісією, що здійснює державне регулювання у сферах енергетики та комунальних послуг та короткі строки її обробки Комітет звертається з проханням надати консультації щодо можливості науково-технічної допомоги в обробці та проведенні змістовного аналізу документів, які надійшли до робочої групи на основі ІТ-технології ТОДОС, яка розроблена фахівцями Вашого інституту.

Голова Комітету

С.І. Скуратовський



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ДЕРЖАВНА НАУКОВА УСТАНОВА «ІНСТИТУТ МОДЕРНІЗАЦІЇ ЗМІСТУ ОСВІТИ»

вул. Митрополита Василя Липківського, 36, м. Київ, 03035, тел./факс: (044) 248-25-14

16. 11. 2017 № 21.1/10 - 2551

На № _____ від _____

Довідка про впровадження науково-технічної продукції

1. **Найменування пропозиції для впровадження:** Науково-технологічні та методичні засоби ІТ-технології ТОДОС щодо створення єдиного інформаційно-освітнього простору на основі використання засобів трансдисциплінарного представлення геопросторової інформації для створення у середовищі ГІС-систем онтологічних інформаційних додатків освітнього профілю. Створення освітніх ГІС-орієнтованих онтологічних інформаційно-аналітичних систем та середовищ.
2. **Ким запропоновано, адреса:** відділ онтологічних систем та прикладної алгебраїчної комбінаторики Інституту телекомунікацій і глобального інформаційного простору НАН України; адреса: Чоколівський б-р, 13, 03186, м. Київ.
3. **Виконавець:** ПРИХОДНЮК Віталій Валерійович.
4. **Джерело інформації:**
 - Построение таксономии документов для формирования иерархических слоев в геоинформационных системах [Текст]: / [Приходнюк В.В. та ін.] // International Journal "Information Content and Processing", Volume 2, Number 2, 2015 – с. 181-199.
 - Онтологічне представлення функціональності систем / В.В. Приходнюк, О.Є. Стрижак, О.Г. Лебідь // Екологічна безпека та природокористування: Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний університет будівництва і архітектури, НАН України Інститут телекомунікацій і глобального інформаційного простору. – Київ, 2016. – Вип. № 3–4 (22), стор. 5-23
5. **Портал мережних онтологічних інструментів:** <http://editor4.inhost.com.ua>
6. **Реєстр онтологій:** <http://ontology4.inhost.com.ua/>

7. **Терміни впровадження:** 2015 – 2017 рр.
8. **Ефективність впровадження:** Підвищення якості та ефективності навчального та інформаційно-аналітичного процесів на основі використання онтологічних засобів семантичного аналізу природно-мовних текстів, які репрезентують геопросторову інформацію. Підвищення якості та ефективності управління освітніми процесами різного тематичного напрямку на основі засобів адаптивної генерації тематичних категорій. Генерація онтологічних категорій меню WEB-семантичного інтерфейсу онтологічних ГІС-додатків. Створення навчальних знанево-орієнтовних систем у мережевому середовищі.
9. **Зауваження, пропозиції:** Розповсюдження отриманих позитивних результатів впровадження щодо створення єдиного інформаційно-освітнього простору шляхом застосування технологічних засобів трансдисциплінарного представлення геопросторової інформації для створення освітніх трансдисциплінарних онтологій та мережових баз знань навчального призначення з метою надання освітніх послуг на основі онтологічних консультативних мережових площадок, використання у навчальному процесі загальноосвітніх навчальних закладів, вищих навчальних закладів, закладів післядипломної педагогічної освіти та підвищення ефективності управління інфраструктурними проектами в галузі освіти.

В. о. директора

Ю. І. Завалевський

Додаток Б. Список опублікованих праць за темою дисертації

Праці, в яких опубліковані основні наукові результати дисертації:

1. Величко В., Приходнюк В., Стрижак А., Марков К., Иванова К., Карастанев С. Построение таксономии документов для формирования иерархических слоев в геоинформационных системах. *Information Content and Processing*, 2015. С. 181–199.
2. Величко В. Ю., Попова М. А., Приходнюк В. В., Стрижак О. Є. ТОДОС – ІТ-платформа формування трансдисциплінарних інформаційних середовищ. *Системи озброєння і військова техніка*, 2017. №. 1(49). С. 10–19.
3. Приходнюк В. В., Стрижак О. Є. Онтологічна ГІС, як засіб впорядкування геопросторової інформації. *Наука і техніка Повітряних Сил Збройних Сил України*, 2017. №. 2(27). С. 167–174.
4. Приходнюк В. В., Стрижак О. Є., Лебідь О. Г. Онтологічне представлення функціональності систем. *Екологічна безпека та природокористування* : Збірник наукових праць / редкол. : О. С. Волошкіна, О. М. Трофимчук (голов. ред.) [та ін.]. Міністерство освіти і науки України, Київський національний університет будівництва і архітектури [та ін.], 2016. №. 3–4 (22). С. 5–23.
5. Приходнюк В. Таксономизация естественно-языковых текстов. *Information Models and Analyses*, 2016. №. 5. С. 270–284.
6. Приходнюк В. В., Стрижак О. Є. Множинні характеристики онтологічних систем. *Математичне моделювання в економіці* : Збірник наукових праць / редкол. : С. О. Довгий (голов. ред.) [та ін.]. К. : НАН України Інститут телекомунікацій і глобального інформаційного простору, Інститут економіки та прогнозування, Інститут кібернетики ім. В. М. Глушкова, 2017. №. 8. С. 47–61.
7. Довгий С. О., Попова М. А., Приходнюк В. В., Стрижак О. Є., Яцухно В. А. Реєстр архівних документів, пов'язаних з життям, творчістю та вшануванням пам'яті Тараса Григоровича Шевченка: Монографія. К.: Центр розвитку особистості «УНІКУМ», 2017. 250 с. (*при грантовій підтримці Держаного фонду фундаментальних досліджень за конкурсним проектом Ф77/37489*)
8. Андрущенко Т. І., Довгий С. О., Глоба Л. С., Стрижак О. Є., Величко В. Ю., Гальченко С. А., Дем'яненко В. Б., Ільченко М. Ю., Лісовий О. В., Новогрудська

Р. Л., Петренко М. Г., Приходнюк В. В., Слюсаренко О. О., Трофимчук О. М., Яценко Г. Ю. Засоби доступу до джерел знань та їх використання в навчальному процесі педагогічного навчального закладу: Методичні рекомендації. 2-ге вид., доповн. К.: Інститут обдарованої дитини, 2012. 192 с.

9. Андрущенко Т. І., Величко В. Ю., Гальченко С. А., Глоба Л. С., Гуляєв К. Д., Клімова Е. Я., Комова О. Б., Лісовий О. В., Попова М. А., Приходнюк В. В., Стрижак О. Є., Стус Д. М. Методики написання наукових робіт на основі онтологічного аналізу текстів: методичний посібник. К.: ТОВ «СІТІПРІНТ», 2013. 124 с.

10. Довгий С. О., Стрижак О. Є., Андрущенко Т. І., Гальченко С. А., Глоба Л. С., Гончар А. В., Копійка О. В., Кудляк В. М., Ляшук К. В., Попова М. А., Приходнюк В. В., Семенюк Г. Ф., Трофимчук О. М., Поліхун Н. І., Постова К. Г. Онтологічний кабінет дослідження життя та творчості Тараса Шевченка в середовищі науково-освітнього порталу KOBZAR.UA: Монографія. К.: Інститут обдарованої дитини, 2016. 175 с.

Праці, що засвідчують апробацію матеріалів дисертації:

11. Горборуков В. В., Приходнюк В. В., Стрижак О. Є., Франчук О. В. Система оцінювання інтелектуальних досягнень учнівської молоді. Онтологічний підхід. *Інтелектуальний аналіз інформації*: Збірник праць XVI Міжнародної наукової конференції ім. Т. А.Таран. К.: Просвіта, 2016. С. 36–42.

12. Величко В. Ю., Приходнюк В. В. Спосіб автоматизованого виділення відношень між термінами з природномовних текстів технічної тематики. *Knowledge – Dialogue – Solution*: Збірник праць XX міжнародної конференції. К.: ІТНЕА, 2014. С. 27–28.

13. Величко В. Ю., Приходнюк В. В. Деякі способи виділення відношень між термінами в природномовному тексті. *Системний аналіз та інформаційні технології*: зб. наук. праць за матеріалами XV конференції (м. Київ, 27 - 31 травня 2013 р.). К.: НТУУ «КПІ», 2013. С. 406.

14. Величко В. Ю., Сирота С. В., Приходнюк В. В. Інструментарій автоматизованого виділення відношень з текстів технічної тематики. *Системний аналіз та інформаційні технології*: зб. наук. праць за матеріалами XVI конференції

(м. Київ, 26 - 30 травня 2014 р.). К. : НТУУ «КПІ», 2014. С. 348.

15. Попова М. А., Приходнюк В. В. Формування тематичних ГІС на основі семантиколінгвістичної обробки документів. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Київ, Пуща-Водиця, 3 - 6 жовтня 2016 р.). К. ; Х. : НАН України, Ін-т телекомунікацій і глоб. інформ. простору [та ін.], 2016. С. 130–133.

16. Приходнюк В. В. Автоматизоване виділення інформації з текстів на основі правил, представлених у форматі λ -виразів. *Управління знаннями та конкурентна розвідка* : зб. наук. праць за матеріалами конференції (м. Харків, 19 - 21 квітня 2016 р.). Х. : Харківський національний університет радіоелектроніки, 2016. С. 48–49.

17. Приходнюк В. В. Автоматизоване формування електронних шарів геоінформаційних систем на основі структурованої і неструктурованої інформації. *Геоінформаційні технології в територіальному управлінні* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Одеса, 17 - 18 вересня 2015 р.). Одеса : Одеський регіональний інститут державного управління [та ін.], 2015. С. 73–76.

18. Приходнюк В. В. Онтолого-орієнтований інтерфейс відображення додатків в ГІС-середовищі. *Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях* : зб. наук. праць за матеріалами XIII міжнар. наук.-практ. конф. (м. Київ, Пуща-Водиця, 3 - 6 жовтня 2016 р.). К. ; Х. : НАН України, Ін-т телекомунікацій і глоб. інформ. простору [та ін.], 2014. С. 127–129.

Додаток В. Відомості про апробацію результатів дисертації

№ п/п	Назви конференції, конгресу, симпозіуму, семінару, школи	Місце проведення	Дата проведення	Форма участі
1	Системний аналіз та інформаційні технології	м. Київ	27-31.05.2013	Заочна
2	Системний аналіз та інформаційні технології	м. Київ	26-30.05.2014	Заочна
3	Управление знаниями и конкурентная разведка	м. Харків	22-24.04.2013	Очна
4	Управление знаниями и конкурентная разведка	м. Харків	14-16.04.2014	Очна
5	Управление знаниями и конкурентная разведка	м. Харків	20-22.04.2015	Очна
6	Управление знаниями и конкурентная разведка	м. Харків	19-21..04.2016	Очна
7	Управление знаниями и конкурентная разведка	м. Харків	25-27.04.2017	Очна
8	Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях	м. Пуща-Водиця	03-06.10.2016	Очна
9	Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях	м. Пуща-Водиця	03-04.06.2017	Очна
10	Knowledge-Dialogue-Solution	м. Київ	08-10.09.2014	Заочна
11	Інтелектуальний аналіз інформації	м. Київ	18-20.05.2016	Очна
12	Геоінформаційні технології у територіальному управлінні	м. Одеса	17-18.09.2015	Очна
13	Теоретичні та прикладні аспекти побудови програмних систем	м. Київ	04-08.12.2017	Очна

Додаток Г. Результати аналізу існуючих програмних засобів

Таблиця 1. Порівняльна характеристика існуючих програмних засобів

Програмний засіб	Значення	Клас системи	Доступність	Сумісність з ОС	Підтримка кирилических символів
ТОДОС	0.819	Інформаційно-аналітична платформа	Відкритий код	Веб-орієнтованість	UTF-8
Palantir Gotham	0.591	Інформаційно-аналітична платформа	Комерційна ліцензія	Кросс-платформність	
Convera Excalibur	0.574	Інформаційно-аналітична платформа	Комерційна ліцензія	Веб-орієнтованість	UTF-8
IBM Watson Analytics	0.422	Інформаційно-аналітична платформа	Безплатна ліцензія	Веб-орієнтованість	
Томіта-парсер (Компанія Yandex)	0.395	Система обробки природної мови	Безплатна ліцензія	Windows	UTF-8
General Architecture for Text Engineering (GATE)	0.378	Фреймворк (обробка природної мови)	Відкритий код	Кросс-платформність	UTF-8
Affix Grammars Over a Finite Lattice (AGFL)	0.37	Фреймворк (обробка природної мови)	Відкритий код	Windows	UTF-8
Protege	0.368	Редактор онтологій	Відкритий код	Кросс-платформність	UTF-8
RCO (Russian Context Optimizer)	0.368	Інформаційно-аналітична платформа	Комерційна ліцензія	Windows	UTF-8
Artificial Intelligence Information Retrieval Engine (AIRE)	0.347	Фреймворк (обробка природної мови)	Відкритий код	Windows	UTF-8
KNIME	0.346	Інформаційно-аналітична платформа	Відкритий код	Windows	UTF-8
RapidMiner	0.346	Інформаційно-аналітична платформа	Відкритий код	Windows	UTF-8
FAST Search	0.341	Інформаційно-аналітична платформа	Комерційна ліцензія	Windows	UTF-8
NeurOK Semantic Suite	0.34	Інформаційно-аналітична платформа	Комерційна ліцензія	Windows	UTF-8
OntoEdit	0.328	Редактор онтологій	Безплатна ліцензія	Кросс-платформність	UTF-8
Apache Lucene/Solr	0.325	Пошукова машина	Відкритий код	Веб-орієнтованість	UTF-8
LexicoSyntactic Pattern Language (LSPL)	0.313	Лексичний аналізатор	Відкритий код	Кросс-платформність	UTF-8

Продовження таблиці 1.

ИСИДА-Т	0.312	Система обробки природної мови	Комерційна ліцензія	Кросс-платформність	UTF-8
Link Grammar Parser, Link Grammar Parser for Russian	0.302	Фреймворк (обробка природної мови)	Відкритий код	Веб-орієнтованість	
ИКОП	0.301	Система обробки природної мови	Комерційна ліцензія	Windows	UTF-8
Sphinx	0.289	Пошукова машина	Відкритий код	Веб-орієнтованість	UTF-8
Galaktika-ZOOM	0.288	Інформаційно-аналітична платформа	Комерційна ліцензія	Windows	UTF-8
Hummingbird, Enterprise Information Management Systems, (EIMS)	0.288	Інформаційно-аналітична платформа	Комерційна ліцензія	Windows	UTF-8
ЛоТА	0.275	Система обробки природної мови	Комерційна ліцензія	Windows	UTF-8
Stemka	0.269	Лексичний аналізатор	Безплатна ліцензія	Кросс-платформність	UTF-8
ABBYY Compreno	0.269	Лексичний аналізатор	Комерційна ліцензія	Windows	UTF-8
SAS Text Analytics	0.257	Інформаційно-аналітична платформа	Комерційна ліцензія	Кросс-платформність	UTF-8
py morphology2	0.244	Лексичний аналізатор	Комерційна ліцензія	Кросс-платформність	UTF-8
Конспект (TextTermin)	0.216	Система обробки природної мови	Комерційна ліцензія	Windows	
Exalead CloudView	0.197	Пошукова машина	Комерційна ліцензія	Веб-орієнтованість	
Oracle Endeca Information Discovery	0.189	Інформаційно-аналітична платформа	Комерційна ліцензія	Windows	windows-1251
Russian Morphological Dictionary	0.121	Лексичний аналізатор	Комерційна ліцензія	Windows	
DictaScope Tokenizer	0.119	Лексичний аналізатор	Безплатна ліцензія	Windows	
Autonomy	0.116	Інформаційно-аналітична платформа	Комерційна ліцензія	Windows	
Statistical Package for the Social Sciences (SPSS Statistics)	0.094	Аналітична система	Комерційна ліцензія	Кросс-платформність	

Таблиця 2. Порівняльна характеристика існуючих програмних засобів (продовження)

Програмний засіб	Програмна і процедурна розширюваність	Мова вхідних даних: Англійська	Мова вхідних даних: Російська	Мова вхідних даних: Інші	Когнітивні властивості: Вибір	Можливості: Автоматична інтеграція
ТОДОС	Так		Низька якість		Так	Так
Palantir Gotham	Так	Так	Так	Так	Так	Так
Convera Excalibur	Так	Так	Так	Так	Так	Так
IBM Watson Analytics	Так	Так	Так		Так	Так
Томіта-парсер (Компанія Yandex)		Так	Так	Так		
General Architecture for Text Engineering (GATE)		Так	Так	Так		
Affix Grammars Over a Finite Lattice (AGFL)		Так	Так	Так		
Protege		Так	Так	Так		
RCO (Russian Context Optimizer)	Так	Так	Так	Так		
Artificial Intelligence Information Retrieval Engine (AIRE)			Так			
KNIME	Так	Так	Так	Так		
RapidMiner	Так	Так	Так	Так		
FAST Search	Так	Так	Так	Так		
NeurOK Semantic Suite	Так	Так	Так	Так	Так	
OntoEdit		Так	Так			
Apache Lucene/Solr	Так	Так	Так	Так	Так	Так
LexicoSyntactic Pattern Language (LSPL)		Так				
ИСИДА-Т		Так	Так			

Продовження таблиці 2.

Link Grammar Parser, Link Grammar Parser for Russian		Так	Так			
ИКОП			Так			
Sphinx	Так	Так	Так		Так	
Galaktika-ZOOM	Так	Так	Так	Так		
Hummingbird, Enterprise Information Management Systems, (EIMS)	Так	Так	Так	Так		
ЛоТА			Так			
Stemka		Так	Так			
ABBYY Compreno		Так	Так	Так		
SAS Text Analytics	Так	Так	Так	Так		
рyморphy2		Так	Так			
Конспект (TextTermin)						
Exalead CloudView					Так	Так
Oracle Endeca Information Discovery		Так	Так			
Russian Morphological Dictionary			Так			
DictaScope Tokenizer						
Autonomy						
Statistical Package for the Social Sciences (SPSS Statistics)						

Таблиця 3. Порівняльна характеристика існуючих програмних засобів (продовження)

Програмний засіб	Відношення порядку, що підтримується	Когнітивні властивості: Структуризація	Когнітивні властивості: Аналіз	Когнітивні властивості: Синтез	Можливості: Інтеграція з ГІС	Можливості: Інтеграція з системами розпізнавання документів
ТОДОС	Часткового порядку	Так	Так	Так	Так	
Palantir Gotham		Так	Так	Так	Так	
Convera Excalibur	Часткового порядку	Так	Так	Так	Так	Так
IBM Watson Analytics		Так	Так	Так		Так
Томіта-парсер (Компанія Yandex)		Так	Так		Так	
General Architecture for Text Engineering (GATE)		Так	Так			
Affix Grammars Over a Finite Lattice (AGFL)		Так	Так			
Protege	Часткового порядку					
RCO (Russian Context Optimizer)			Так		Так	
Artificial Intelligence Information Retrieval Engine (AIRE)		Так	Так			
KNIME						
RapidMiner						
FAST Search					Так	
NeurOK Semantic Suite			Так		Так	
OntoEdit	Часткового порядку					
Apache Lucene/Solr	Лінійного порядку					

Продовження таблиці 3.

LexicoSyntactic Pattern Language (LSPL)		Так	Так			
ИСИДА-Т		Так	Так			
Link Grammar Parser, Link Grammar Parser for Russian						
ИКОП		Так	Так			
Sphinx						
Galaktika-ZOOM					Так	
Hummingbird, Enterprise Information Management Systems, (EIMS)					Так	
ЛоТА		Так	Так			
Stemka						
ABBYY Compreno						Так
SAS Text Analytics						
py morphology2						
Конспект (TextTermin)	Часткового порядку	Так	Так			
Exalead CloudView	Лінійного порядку					
Oracle Endeca Information Discovery						
Russian Morphological Dictionary						
DictaScope Tokenizer						
Autonomy						
Statistical Package for the Social Sciences (SPSS Statistics)						

Таблиця 4. Порівняльна характеристика існуючих програмних засобів (продовження)

Програмний засіб	Можливості: Категоризація	Можливості: Автоматизація відображення таксономій	Можливості: Опрацювання великих даних (Big Data)	Глибина бінарних відношень таксономії	Можливості: Аналіз ПМТ	Можливості: Таксономізація
ТОДОС	Так	Так	Можлива	Без обмежень	Так	Так
Palantir Gotham	Так		Так	Без обмежень	Так	Так
Convera Excalibur	Так	Так	Так	Без обмежень		
IBM Watson Analytics	Так		Так			
Томіта-парсер (Компанія Yandex)						Можлива реалізація
General Architecture for Text Engineering (GATE)					Так	Можлива реалізація
Affix Grammars Over a Finite Lattice (AGFL)					Так	Можлива реалізація
Protege	Так	Так		Без обмежень		
RCO (Russian Context Optimizer)					Так	
Artificial Intelligence Information Retrieval Engine (AIRE)					Так	Можлива реалізація
KNIME					Так	
RapidMiner					Так	
FAST Search					Так	
NeurOK Semantic Suite						
OntoEdit		Так		Без обмежень		
Apache Lucene/Solr						
LexicoSyntactic Pattern Language (LSPL)					Так	
ИСИДА-Т					Так	

Продовження таблиці 4.

Link Grammar Parser, Link Grammar Parser for Russian					Так	
ИКОП		Так				Так
Sphinx						
Galaktika-ZOOM						
Hummingbird, Enterprise Information Management Systems, (EIMS)						
ЛоТА		Так				
Stemka						
ABBYY Compreno						
SAS Text Analytics						
py morphology2						
Конспект (TextTermin)				Без обмежень		Так
Exalead CloudView	Так		Так	Обмежена архітектурою		
Oracle Endeca Information Discovery						
Russian Morphological Dictionary						
DictaScope Tokenizer						
Autonomy						
Statistical Package for the Social Sciences (SPSS Statistics)						

Таблиця 5. Порівняльна характеристика існуючих програмних засобів (продовження)

Програмний засіб	Можливості: Знаходження латентної інформації	Мова вхідних даних: Українська	Можливості: Підтримка форматів і протоколів Semantic Web	Можливості: Створення інтерактивних документів	Мережецентризм	Емерджентність
ТОДОС	Так	Так		Так	Так	Так
Palantir Gotham	Так					
Convera Excalibur						
IBM Watson Analytics						
Томіта-парсер (Компанія Yandex)		Так				
General Architecture for Text Engineering (GATE)						
Affix Grammars Over a Finite Lattice (AGFL)						
Protege		Так				
RCO (Russian Context Optimizer)						
Artificial Intelligence Information Retrieval Engine (AIRE)						
KNIME						
RapidMiner						
FAST Search						
NeurOK Semantic Suite						
OntoEdit		Так				
Apache Lucene/Solr						
LexicoSyntactic Pattern Language (LSPL)						
ИСИДА-Т						

Продовження таблиці 5.

Link Grammar Parser, Link Grammar Parser for Russian		Так				
ИКОП						
Sphinx						
Galaktika-ZOOM						
Hummingbird, Enterprise Information Management Systems, (EIMS)						
ЛюТА						
Stemka		Так				
ABBYY Compreno		Так				
SAS Text Analytics						
py morphology2		Так				
Конспект (TextTermin)						
Exalead CloudView						
Oracle Endeca Information Discovery						
Russian Morphological Dictionary						
DictaScope Tokenizer						
Autonomy						
Statistical Package for the Social Sciences (SPSS Statistics)						

Додаток Д. Приклади вхідних файлів для формування ГІС-додатків

Таблиця 1. Приклад вхідної таблиці для створення ГІС-додатку «Ліцензіати послуг ЖКХ»

№ з/п	Найменування ліцензіатів	Тарифи на послуги з централізованого опалення, грн/м?	у тому числі за складовими:						
			Витрати на оплату праці та внески на соціальні заходи	Витрати на проведення періодичної перевірки, обслуговування і ремонту квартирних засобів обліку, в тому числі їх демонтажу, транспортування та монтажу після перевірки	Витрати на придбання води на послугу з централізованого постачання гарячої води	Послуги банку	Решта	Податок на додану вартість (ПДВ)	
1	2	3	4	5	6	7	8	9	10
Сумарні та середньозважені показники		17,19	13,97	0,12	0	0	0,22	0,01	2,86
	Вінницька								
1	Відокремлений підрозділ "Ладизинська ТЕС" ПАТ "ДТЕК Західенерго"	8,86	7,17	0,14			0,07	0	1,48
2	ДП "Теплокомуненерго Маяк" ВАТ "Маяк"	16,85	13,84	0,07			0,13	0	2,81
3	КП "Вінницяоблтеплоенерго"	20,42	16,62	0,22			0,17	0	3,4
4	КП ВМР "Вінницяміськтеплоенерго" одноставковий еквівалент двоставкового тарифу	17,05	13,83	0,12			0,24	0,02	2,84
4.1.	у т.ч. умовно-змінна частина двоставкового тарифу	13,21	10,82				0,19	0	2,2
4.2.	у т.ч. умовно-постійна частина двоставкового тарифу, грн/Гкал/год на місяць	25 439,08	19 940,44	784,09	0	0	360,39	114,32	4 239,85
5	Могилів-Подільське міське КП "Теплоенергетик"	15,86	12,9	0,12			0,19	0	2,64
	Волинська	0	0	0			0	0	0
1	Володимир-Волинське ПТМ "Володимир-	18,1	14,73	0,07			0,28	0	3,02

	Волинськтеплокомун енерго"								
2	ДКП "Луцьктепло"	15,54	12,7 7	0,04			0,14	0	2,59
3	КП "Нововолинськтеплог омуненерго" житлово- комунального об'єднання Нововолинської міської ради	22,96	18,7 7	0,08			0,29	0	3,83
4	ПТМ "Ковельтепло"	16,23	13,3	0,09			0,14	0	2,7
5	ТзОВ "Західна Теплоенергетична Група"	0	0	0			0	0	0
	Дніпропетровська	0	0	0			0	0	0
1	ДП "Криворізька теплоцентраль"	19,4	15,8 1	0,08			0,27	0	3,23
2	Комунальне підприємство "Тернівське житлово- комунальне підприємство"	11,58	9,49	0,17			0	0	1,93
3	КП "Дніпродзержинськте пломережа"	19,06	15,4 7	0,11			0,3	0	3,18
4	КП "Жовтоводськтеплом ережа"	0	0	0			0	0	0
5	КП "Коменергосервіс" Дніпропетровської міської ради	12,51	10,1	0,12			0,21	0	2,09
6	КП "Марганецьтепломер ежа"	20,07	16,2 7	0,12			0,33	0	3,34
7	КП "Новомосковськтепло енерго"	17,7	14,4 9	0,11			0,15	0	2,95
8	КП "Павлоградтеплоенер го"	18,6	15,0 6	0,12			0,31	0,0 1	3,1
9	КП "Теплоенерго" Дніпропетровської міської ради	20,12	16,3 5	0,15			0,25	0,0 1	3,35
10	КП "ТПТЕ "Теплотранс" Дніпропетровської міської ради	25,58	20,7 3	0,16			0,43	0	4,26
11	КПТМ "Криворіжтепломер ежа"	21,77	17,6 9	0,14			0,31	0	3,63
12	МКП "Дніпропетровські міські теплові мережі"	18,98	15,3 3	0,17			0,32	0	3,16

13	Нікопольське КП «Нікопольтеплоенерг о»	20,33	16,5 4	0,07			0,33	0	3,39
14	Орджонікідзевському міському КП "Орджонікідзетеплоє нерго"	19,18	15,6 6	0,11			0,21	0	3,2
15	ПАТ "ДТЕК Дніпроенерго" ВП "Криворізька ТЕС"	10,48	8,54	0,1			0,09	0	1,75
16	ПАТ "ДТЕК Дніпроенерго" ВП "Придніпровська ТЕС"	10,51	8,24	0,2			0,18	0,1 5	1,75
17	Першотравенське міське житлово- комунальне підприємство	11,79	9,7	0,13			0	0	1,97
18	ТОВ "Теплосервіс"	24,74	20,2 6	0,12			0,24	0	4,12
	Донецька	0	0	0			0	0	0
1	КВП «Краматорська тепломережа» Краматорської міської ради	20,48	16,7	0,14			0,23	0	3,41
2	ККП "Донецькіськтеплом ережа"	19,07	15,4 3	0,14			0,32	0	3,18
3	ККП "Маріупольтепломер ежа"	15,76	12,8 6	0,13			0,13	0	2,63
4	КП "Вуглик" Горлівської міської ради	23,61	19,4	0,1			0,16	0,0 1	3,93
5	КП "Красноармійськтепл омережа"	21,11	17,2 9	0,12			0,18	0	3,52
6	КП "Макіївтепломережа"	21,17	17,2 7	0,14			0,23	0	3,53
7	КП "Тепломережа" (м. Донецьк)	21,62	17,5 9	0,16			0,27	0	3,6
8	ОКП "Донецьктеплокомун енерго"	20,53	16,5 5	0,2			0,34	0,0 1	3,42
9	ПАТ "Донбасенерго"	0	0	0			0	0	0
10	ПАТ "ДТЕК Донецькобленерго" ВП "Миронівська ТЕС"	14,97	12,3 4	0,14			0	0	2,49
11	ПАТ "Часівоярський вогнетривкий комбінат"	25,57	20,8 5	0,25			0,21	0	4,26
12	ПрАТ "Горлівськтепломере жа"	19,38	15,8 9	0,12			0,12	0,0 1	3,23
13	ТДВ "Шахта "Білозерська"	16,5	13,5	0,11			0,14	0	2,75
14	ТОВ "Артемівськ- Енергія"	18,29	14,9 4	0,1			0,2	0,0 1	3,05

15	ТОВ "Водотеплокомунікація" (м. Вугледар)	15,89	12,95	0,14			0,15	0	2,65
16	ТОВ "ДТЕК "Добропіллявугілля"	17	13,87	0,1			0,2	0	2,83
17	ТОВ "ДТЕК Східенерго" ВП "Курахівська ТЕС"	16,07	13,21	0,19			0	0	2,68
18	ТОВ "Краматорськтеплоенерго"	27,9	22,75	0,15			0,35	0	4,65
	Житомирська	0	0	0			0	0	0
1	КП "Бердичівтеплоенерго"	17,66	14,43	0,09			0,2	0	2,94
2	КП "Житомиртеплокомуненерго"	14,91	12,22	0,11			0,09	0	2,48
3	КП "Озерне" Новоїтвинської селищної ради	20,13	16,28	0,16			0,34	0	3,35
4	КП Новоград-Волинської міської ради «Новоград-Волинськтеплокомуненерго»	14,73	12,05	0,1			0,12	0	2,46
5	КП теплозабезпечення	18,36	14,9	0,09			0,31	0	3,06
	Запорізька	0	0	0			0	0	0
1	Концерн "Міські теплові мережі"	14,53	11,93	0,1			0,06	0,01	2,42
2	КП "Дніпрорудненські теплові мережі" одноставковий еквівалент двоставкового тарифу	16,04	12,96	0,13			0,27	0,01	2,67
2.1.	у т.ч. умовно-змінна частина двоставкового тарифу	12,39	10,12				0,21	0	2,07
2.2.	у т.ч. умовно-постійна частина двоставкового тарифу, грн/Гкал/год на місяць	23 536,01	18 335,26	842,4	0	0	392,27	43,42	3 922,67
3	КП "Токмак теплоенергія" Токмацької міської ради	17,44	14,14	0,13			0,26	0	2,91
4	ПАТ "Бердянське підприємство теплових мереж"	21,58	17,53	0,15			0,31	0	3,6
5	ПАТ "Мотор Січ" (м. Запоріжжя)	0	0	0			0	0	0
6	ТОВ "Мелітопольські теплові мережі" одноставковий	15,1	12,25	0,14			0,19	0	2,52

	еквівалент двоставкового тарифу								
6. 1.	у т.ч. умовно-змінна частина двоставкового тарифу	12,12	9,95				0,16	0	2,02
6. 2.	у т.ч. умовно-постійна частина двоставкового тарифу, грн/Гкал/год на місяць	19 815,22	15 301,91	926,56	0	0	253,47	30,74	3 302,54
	Івано-Франківська	0	0	0			0	0	0
1	ДМП "Івано-Франківськтеплокому енерго" одноставковий еквівалент двоставкового тарифу	15,7	12,81	0,07			0,2	0,01	2,62
1. 1.	у т.ч. умовно-змінна частина двоставкового тарифу	12,28	10,08				0,15	0	2,05
1. 2.	у т.ч. умовно-постійна частина двоставкового тарифу, грн/Гкал/год на місяць	23 748,51	18 970,15	487,51	0	0	296,86	35,91	3 958,08
2	КП "Водотеплосервіс" Калуської міської ради	18,7	15,24	0,08			0,27	0	3,12
3	ТОВ "Станіславська теплоенергетична компанія" одноставковий еквівалент двоставкового тарифу	18,52	15,11	0,09			0,23	0,01	3,09
3. 1.	у т.ч. умовно-змінна частина двоставкового тарифу	14,52	11,92				0,18	0	2,42
3. 2.	у т.ч. умовно-постійна частина двоставкового тарифу, грн/Гкал/год на місяць	26 250,41	20 921,85	595,64	0	0	328,13	29,72	4 375,07
	Київська	0	0	0			0	0	0
1	Вишгородське РКП "Вишгородтепломережа"	22,05	17,88	0,15			0,34	0	3,67
2	КП "Боярське ГВУЖКГ Боярської міської ради Києво-Святошинського району Київської області"	19,15	15,54	0,17			0,25	0	3,19
3	КП "Броваритепловодое нергія"	14,85	12,16	0,1			0,12	0	2,48
4	КП "Васильківтепломережа"	21,62	17,75	0,08			0,18	0	3,6

5	КП "Вишнівськтеплоенерго" Вишневої міської ради Кірово-Святошинського району Київської області	16,96	13,8	0			0,21	0,12	2,83
6	КП "Кірово-Святошинська тепломережа" Київської обласної ради	21,12	17,32	0,15			0,14	0	3,52
7	КП "Управління житлово-комунального господарства" (м. Славутич)	16,86	13,75	0,08			0,21	0	2,81
8	КП БМР "Білоцерківтепломережа"	16,87	13,77	0,15			0,14	0	2,81
9	КПТМ "Бориспільтепломережа"	17,93	14,65	0,1			0,19	0	2,99
10	ПАТ "Енергія"	0	0	0			0	0	0
11	ПКПП "Теплокомунсервіс" (м. Буча)	19,68	16,1	0,06			0,25	0	3,28
	Кіровоградська	0	0	0			0	0	0
1	ДП "Кіровоградтепло" ТОВ "Центр науково-технічних інновацій Української нафтогазової академії"	18,3	14,8	0,15			0,31	0	3,05
2	КП "Теплоенергетик"	19,92	16,23	0,11			0,26	0	3,32
3	КП "Теплокомуненерго" Олександрійської міської ради	21,03	17,14	0,14			0,25	0	3,5
4	СП-ТОВ "Світловодськпобут"	18,07	14,87	0,08			0,08	0,04	3,01
5	ТОВ "ДОЛИНСЬКІ ОБ'ЄДНАНІ МЕРЕЖІ"	15,38	12,62	0,09			0,1	0	2,56
	Луганська	0	0	0			0	0	0
1	АМКП "Теплокомуненерго"	0	0	0			0	0	0
2	ДП "Сєвєродонецька ТЕЦ"	23,15	18,83	0,1			0,35	0	3,86
3	КП "Лисичанськтепломережа"	0	0	0			0	0	0
4	КП "Первомайськтеплокомуненерго" Первомайської міської ради	0	0	0			0	0	0

5	КП "Сєверодонецьктепло комуненерго"	17,08	13,8 3	0,11			0,3	0	2,85
6	КП "СТП "Ровенькитеплокомун енерго"	0	0	0			0	0	0
7	КСТП "Рубіжнєтеплокомуне нерго" Рубіжанської міської ради	19,74	16,0 7	0,08			0,3	0	3,29
8	КТП "Алчевськтеплокомун енерго"	19,77	15,9 8	0,16			0,33	0,0 1	3,29
9	Луганське МКП "Теплокомуненерго"	18,63	15,1	0,16			0,26	0	3,11
10	ТОВ "ДТЕК Ровенькиантрацит"	0	0	0			0	0	0
	Львівська	0	0	0			0	0	0
1	ДКП "Стебниктеплокомуне нерго"	16,97	13,8 6	0,14			0,14	0	2,83
2	КП "Дрогобичтеплоенерг о" ДМР	19,72	16,2 1	0,06			0,16	0	3,29
3	КП "Бориславтеплоенерг о"	17,38	14,1 4	0,14			0,2	0	2,9
4	КП "Бродитеплоенерго"	21,43	17,5	0,17			0,2	0	3,57
5	КП "Жовкватеплоенерго"	20,83	16,8 6	0,19			0,31	0	3,47
6	КП "Стрийтеплоенерго"	18,5	15,2	0,07			0,15	0	3,08
7	КП "Трускавецьтепло"	15,35	12,4 6	0,07			0,23	0,0 4	2,56
8	КП "Червоноградтеплоко муненерго"	25,08	20,4 1	0,08			0,42	0	4,18
9	КП Сокальської міської ради "Сокальтеплокомунен ерго"	20,76	16,8 7	0,22			0,2	0	3,46
10	ЛКМП "Львівтеплоенерго" одноставковий еквівалент двоставкового тарифу	15,53	12,7 2	0,07			0,16	0	2,59
10 .1.	у т.ч. умовно-змінна частина двоставкового тарифу	12,52	10,3 1				0,13	0	2,09
10 .2.	у т.ч. умовно-постійна частина двоставкового тарифу, грн/Гкал/год на місяць	21 927,14	17 556, 23	487,1 1	0	0	219, 27	10, 01	3 654, 52
11	ЛКП "Залізничнєтеплоене рго"	15,08	12,3 2	0,08			0,16	0	2,51

11	у т.ч. умовно-змінна частина								
.1.	двоставкового тарифу	11,66	9,59				0,13	0	1,94
11	у т.ч. умовно-постійна частина								
.2.	двоставкового тарифу, грн/Гкал/год на місяць	25 631,07	20 468,40	595,5	0	0	277,67	17,65	4 271,84
12	ПТМ "Самбіртеплокомуненерго"	19,03	15,27	0,31			0,29	0	3,17
13	ТзОВ "Енергія-Новий Розділ"	16,18	13,16	0,12			0,2	0	2,7
14	ТзОВ НВП "Енергія-Новояворівськ"	14,62	11,92	0,09			0,17	0	2,44
	м. Київ	0	0	0			0	0	0
1	ПАТ "Київенерго"	16,14	13,04	0,15			0,22	0,03	2,69
2	ТОВ "ЄВРО-РЕКОНСТРУКЦІЯ"	16,91	13,59	0,18			0,24	0,09	2,82
3	ТОВ "Теплопостачсервіс" (м. Київ)	12,8	10,38	0,13			0,16	0	2,13
	Миколаївська	0	0	0			0	0	0
1	КП "Теплопостачання та водо-каналізаційне господарство"	4,93	3,98	0,08			0,03	0,02	0,82
2	КП Первомайської міської ради "Тепло"	16,91	13,83	0,12			0,14	0	2,82
3	ОКП "Миколаївоблтеплоенерго"	17,07	13,81	0,19			0,21	0,01	2,84
4	ПАТ "Миколаївська ТЕЦ"	15,26	12,32	0,22			0,17	0	2,54
	Одеська	0	0	0			0	0	0
1	КВЕП "Котовськтеплокомуненерго"	20,39	16,55	0,1			0,34	0	3,4
2	КП «Теплові мережі Ізмаїлтеплокомуненерго»	24,22	19,68	0,15			0,34	0	4,04
3	КП "Білгород-Дністровськтеплоенерго"	19,33	15,64	0,16			0,31	0	3,22
3.	у т.ч. умовно-змінна частина								
1.	двоставкового тарифу	14,72	12,03				0,24	0	2,45
3.	у т.ч. умовно-постійна частина								
2.	двоставкового тарифу, грн/Гкал/год на місяць	26 876,28	21 043,52	921,12	0	0	432,26	0	4 479,38
4	КП "Іллічівськтеплоенерго"	14,58	11,88	0,1			0,17	0	2,43
5	КП "Теплопостачання міста Одеси"	16,77	13,61	0,1			0,27	0	2,79

6	КПТМ "Южтеплокомуненерг о"	14,89	12,2 4	0,07			0,1	0	2,48
7	ТОВ "Теплодаренерго"	17,14	14,1 6	0			0	0,1 2	2,86
	Полтавська	0	0	0			0	0	0
1	КВП "Комсомольськтепло енерго"	17,2	14,1 7	0,08			0,07	0,0 1	2,87
2	КП "Теплоенерго" (м. Кременчук)	16,29	13,4	0,07			0,11	0	2,72
3	КПТГ "Гадячтеплоенерго"	16,11	13,2 5	0,11			0,07	0	2,68
4	ОКВПТГ "Лубнитеплоенерго"	20,9	17,1 8	0,05			0,17	0,0 1	3,48
5	ОКВПТГ "Миргородтеплоенер го"	19,27	15,7 8	0,08			0,16	0,0 3	3,21
6	ПАТ "Полтаваобленерго" (Кременчуцька ТЕЦ)	14,75	12,0 2	0,12			0,15	0	2,46
7	ПОКВПТГ "Полтаватеплоенерго "	14,86	12,1 4	0,14			0,1	0	2,48
	Рівненська	0	0	0			0	0	0
1	КП "Дубнокомуненергія" Дубенської міської ради	18,2	14,8 6	0,07			0,23	0	3,03
2	КП "Здолбунівкомуненер го" Здолбунівської міської ради	21,21	17,3 5	0,15			0,17	0	3,54
3	КП "Костопількомуненер гія"	18,3	14,9 8	0,12			0,15	0	3,05
4	Кузнецовське міське комунальне підприємство	3,5	2,86	0,03			0,03	0	0,58
5	ТОВ "Рівнетеплоенерго"	18,23	14,9 3	0,11			0,14	0,0 1	3,04
	Сумська	0	0	0			0	0	0
1	ДП "Конотопський авіаремонтний завод "АВІАКОН"	18,74	15,3	0,12			0,2	0	3,12
2	КП "Глухівський тепловий район"	18,72	15,3 3	0,11			0,16	0	3,12
3	КП "Ромникомунтепло"Р МР"	14,52	11,9 2	0,08			0,1	0	2,42
4	КП "Теплогарант" (м. Конотоп)	20,12	16,4 4	0,11			0,22	0	3,35
5	КП "Шосткинський казенний завод "Імпульс"	17,97	14,6 6	0,15			0,16	0	2,99
6	КП Білопільської міської ради	17,24	14,1 6	0,06			0,14	0	2,87

	"Теплосервіс Білопілья"								
7	ПАТ "Сумське машинобудівне НВО ім. М.В.Фрунзе"	17,4	14,2 2	0,13			0,13	0,0 2	2,9
8	ТОВ "Брок-Енергія"	19,23	15,5 6	0,3			0,16	0	3,21
9	ТОВ "Сумитеплоенерго"	21,39	17,4 9	0,11			0,22	0	3,56
10	ТОВ "ТЕПЛОДОПОСТАЧ "	17,55	14,2 9	0,08			0,25	0	2,92
11	ТОВ "Шосткінське підприємство "Харківенергоремонт"	18,4	15,0 7	0,11			0,15	0	3,07
	Тернопільська	0	0	0			0	0	0
1	КПТМ "Тернопільміськтепло комуненерго"	14,88	12,1 8	0,09			0,12	0	2,48
2	КПТМ Тернопільської обласної ради "Тернопільтеплокому ненерго"	21,57	16,9 8	0,82			0,18	0	3,6
	Харківська	0	0	0			0	0	0
1	Борівське КП ТМ	21,94	17,9 2	0,13			0,23	0	3,66
2	Вовчанське підприємство теплових мереж	18,51	15,1 1	0,24			0,07	0	3,08
3	Ізюмське КП ТМ	19,86	16,2	0,11			0,23	0	3,31
4	КП "Тепловодосервіс" Лозівської районної ради	26,95	22,1 9	0,13			0,13	0	4,49
5	КП "Теплоенерго" Лозівської міської ради Харківської області	21,38	17,6 5	0			0	0,1 6	3,56
6	КП "Харківські теплові мережі"	16,42	13,3 1	0,09			0,27	0,0 1	2,74
7	КП "Чугуївтепло"	18,51	15,0 2	0,09			0,31	0	3,08
8	КП БРР "Балаклійські теплові мережі"	20,46	16,7 3	0,08			0,24	0	3,41
9	КП ТМ Харківського району Харківської районної державної адміністрації	26,06	21,2	0,09			0,43	0	4,34
10	Красноградське ПТМ	19,29	15,6 6	0,09			0,32	0	3,22
11	ПКП "Тепломережі"	15,97	13,0 4	0,07			0,19	0	2,66
12	ПрАТ «Теплоенергетичний центр Роганського промвузла»	13,89	11,1 6	0,19			0,23	0	2,32
13	ТОВ "Котельні лікарняного	26,29	21,1 5	0,42			0,34	0	4,38

	комплексу" (м. Харків)								
14	Харківське ОКП "Дирекція розвитку інфраструктури території"	23,73	19,49	0,11			0,18	0	3,95
	Херсонська	0	0	0			0	0	0
1	КПТМ "Каховтеплокомунерго"	22,77	18,66	0,12			0,19	0	3,79
2	МКП "Херсонтеплоенерго"	14,81	12,01	0,14			0,19	0,01	2,47
3	ПАТ "Херсонська теплоелектроцентраль"	22,65	18,47	0,21			0,19	0	3,77
4	ПП "Херсонтеплогенерація"	16,45	13,25	0,19			0,27	0	2,74
	Хмельницька	0	0	0			0	0	0
1	Волочиське комунальне підприємство теплових мереж "Тепловик"	20,07	16,26	0,25			0,21	0	3,34
2	КП "Міськтепловоденергія"	21,34	17,44	0,16			0,18	0	3,56
3	КП "Південно-західні тепломережі"	12,25	9,96	0,09			0,15	0	2,04
4	КП "Славутське житлово-комунальне об'єднання"	19,14	15,64	0,09			0,22	0	3,19
5	КП по експлуатації теплового господарства "Тепловик" Старокостянтинівської міської ради	22,73	18,61	0,05			0,28	0	3,79
6	МКП "Хмельницьктеплокомуненерго"	13,72	11,17	0,1			0,16	0	2,29
7	ТОВ "Шепетівка Енергоінвест"	20,62	16,9	0,11			0,17	0	3,44
	Черкаська	0	0	0			0	0	0
1	Ватутінське комунальне підприємство теплових мереж	16,69	13,54	0,13			0,21	0,02	2,78
2	ДП "Теплокомуненерго" ПАТ "Монастирищенський орден Трудового Червоного Прапора машинобудівний завод"	19,25	15,81	0,09			0,14	0	3,21
3	Канівське КПТМ	21,14	17,27	0,17			0,18	0	3,52

4	КПТМ "Черкаситеплокомуне нерго" Черкаської міської ради	18,07	14,8	0,04			0,21	0	3,01
5	ПАТ "Черкаське хімволокно"	15,25	12,5 4	0,04			0,12	0,0 1	2,54
6	ТОВ "Сміла Енергоінвест"	21,16	17,2 7	0,14			0,22	0	3,53
7	ТОВ "Смілаенергопромтра нс"	17,49	14,1 1	0,17			0,29	0	2,92
8	УКП "Уманьтеплокомунен ерго"	18,72	15,3 1	0,1			0,19	0	3,12
	Чернівецька	0	0	0			0	0	0
1	МКП "Чернівцітеплокомун енерго"	13,73	11,2 1	0,11			0,11	0	2,29
	Чернігівська	0	0	0			0	0	0
1	КП "Прилуkiteпловодоп остачання" Прилуцької міської ради Чернігівської області	17,17	14,1 1	0,06			0,14	0	2,86
2	ПАТ "Облтеплокомуненер го"	18,52	15,0 8	0,1			0,22	0,0 3	3,09
3	ТОВ "НіжинТеплоМережі"	22,95	18,7 7	0,17			0,19	0	3,83
4	ТОВ фірма "ТехНова" (Чернігівська ТЕЦ)	17,27	14,0 6	0,12			0,22	0	2,88
	Діяльність в декількох областях	0	0	0			0	0	0
1	ДП "НАЕК "Енергоатом"	2,62	2,12	0,04			0,02	0	0,44
2	ПАТ "Центренерго"	7,02	5,65	0,14			0,07	0	1,17
3	ТзОВ «Дзержинське управління регіонального будівництва»	15,06	12,0 6	0,29			0,21	0	2,51

Таблиця 2. Вхідна таблиця для створення ГІС-додатку «Місця видалення відходів»

Код організованого МВВ	0.1.	MVV_07107_00025
Реєстраційний номер МВВ, дата реєстрації	A.1.	№113 від 17.01.2015
Дата останньої ревізії МВВ	A.2.	42355
Найменування МВВ, код	A.3.	тверді побутові відходи, 7720.3.1
Місце знаходження МВВ (населений пункт, район), код КОАТУУ	A.4.	С.Нова лішня – 0,6 км на південь 0710700000
Повне найменування та адреса власника МВВ, код за ЄДРПОУ	A.5.	Виробниче управління комунального господарства ЖКО НМР 03339331
Режим функціонування МВВ (діюче, закрите, законсервоване)	A.6.	діюче
Категорія екологічної безпеки МВВ	A.7.	A малонебезпечні
Характеристика МВВ (обсяг, площа, наземне, підземне)	A.8.	296,27 тн – обсяг, 7,76 га – площа, Підземне
Характеристика відходів (найменування, група, клас небезпеки)	A.9.	ТПВ, 0008 IV клас небезпеки
Географічна широта	B.1.1.	50.699098
Географічна довгота	B.1.2.	24.205384
Район:	B.2.1.	Іваничівський
КОАТУУ:	B.2.2.	721100000
Зона:	B.2.3.	03
Квартал:	B.2.4.	000
Номер ділянки	B.2.5.	0721100000:03:000:0209
Тип власності	B.2.6.	державна
цільове призначення	B.2.7.	Для будівництва ТПВ
Площа ділянки	B.2.8.	7,7611
Чи є МВВ перевантаженим ?	B.4.1.	ні
Чи відповідає МВВ нормам екобезпеки ?	B.4.2.	так
Чи є МВВ закритим ?	B.4.3.	ні
Чи МВВ підлягає рекультивзації?	B.4.4.	ні
Чи МВВ рекультивовано у звітному періоді?	B.4.5.	ні
Чи МВВ підлягає санації?	B.4.6.	ні
Чи є МВВ санованим у звітному періоді?	B.4.7.	ні
Середньодобовий обсяг відходів, що поступають на МВВ (м3 / тонн)	D.1.	298,9 м3 / 56,2 т
Середньодобовий обсяг вторинної сировини, що вивозиться з МВВ (тонн)	D.2.	Макулатура – 0,49 тн Полімери- 0,07 тн Метали – 0,438 тн
Номенклатура вторинної сировини, що відбирається на МВВ Назва – річний обсяг відбору (тонн)	D.3.	Макулатура – 125 тн Полімери- 18 тн Метали – 110 тн
Перелік населених пунктів, звідки поступають побутові відходи на МВВ (код КОАТУУ:- назва – орієнтовна кількість мешканців)	D.4.	с.Гряди – 0721181200 – 780 осіб с.Ст.Лішня -0721187700-420 осіб с.Нова Лішня- 0721187705 -260 осіб,с.Литовеж- 0721182401 -1500 осіб.с.мт.Благодатне- 0710700000- 4850 осіб
Перелік юридичних осіб та ФОП, які здійснюють завезення ТПВ на МВВ	D.5.	Завод «Промлит», КП "Теплокомуненерго», Ливарний завод, Хлібозавод, Завод дезодорованих олій,

		ФОП БадзюньВ.М., ФОП Лисюк А.В., Тзов ВМП, Тзов Імперія жирів,
Перелік юридичних осіб та ФОП, які здійснюють сортування ТПВ та вилучення ресурсоцінних компонентівРежим функціонування МВВ (діюче, закрите, законсервоване)	D.6.	ВУКГ

Додаток Е. Інструментальні засоби трансдисциплінарної обробки, що належать до середовища ТОДОС

1. Загальна інформація

1.1. Призначення системи ТОДОС

Концептуальна модель предметної області, або онтологія, складається з онтографу (ієрархії понять) предметної області, зв'язків між ними і законів, які діють в рамках цієї моделі.

Система «Трансдисциплінарні Онтологічні Діалоги Об'єктно-орієнтованих Систем» (ТОДОС) призначена для реалізації технології онтологічної інтеграції розподілених інформаційних ресурсів шляхом побудови онтологічного графа, вершинами якого є поняття і процеси предметних областей (концепти). ТОДОС забезпечує динамічне декларування імен понять і процесів, формування тематичних термінополів з підключенням нових властивостей концептів.

Система ТОДОС передбачає:

- створення єдиного інформаційного простору для забезпечення ефективної навчально-дослідницької діяльності всіх груп користувачів;
- аналіз і створення баз знань на основі мережових інформаційних масивів;
- поширення результатів персоналізованих запитів між користувачами;
- створення спеціалізованих територіально-розподілених інформаційних систем;
- маршрутизацію й класифікацію вхідної електронної пошти великих структур;
- забезпечення учителів, науковців, учнів необмеженою контекстно значущою інформацією з Інтернету без безпосередньої роботи в мережі (інформація із заданих вузлів Інтернету доставляється в локальну мережу закладу й безпосередньо на робочі місця користувачів);

Представлення процесу обробки контенту мережових інформаційних ресурсів засобами системи ТОДОС відображається у вигляді онтологічного графу (онтограф).

Декларативність реалізується поданням концептів онтології у вигляді упорядкованої множини дводольних онтографів. Ребра онтографа визначають властивості між кожною парою концептів. Множина, що інтерпретує функції, що визначають семантику кожної онтології, може постійно динамічно доозначатися множиною нових відносин і спеціальними групами міждисциплінарних аксіом. Зазначені групи аксіом окрім нових властивостей охоплюють ще й обмеження на концепти різних тематичних онтологій, що включаються в процес інтеграції.

Представлення інформації у вигляді онтографу дозволяє вивчати не лише окремий термін (поняття), але й отримувати всі його семантичні зв'язки з іншими поняттями, тим самим осмислюючи його роль у даній системі знань чи в ході розв'язання задачі.

Однак мережевий онтограф може виступати не лише засобом організації знань. Розширюючи його традиційні функції, онтограф можна перетворити на середовище, в якому забезпечується активна робота зі знаннями, а також оригінальним чином вирішуються навчальні завдання.

У процесі виділення ієрархій між термінами-поняттями виконується локалізація таксономічних невизначеностей на рівні входу й виходу певних вершин онтографу, що визначають неможливість безпосередньо визначити певне відношення між об'єктами чи процесами. Розгляд цих вершин, як таксономічних невизначеностей, забезпечує синхронізацію смислових сутностей, що, в свою чергу, дозволяє оптимізувати умови переходу й дозначити класифікацію об'єктів у онтологічній моделі.

1.2. Вимоги до рівня підготовки користувача

Для роботи з системою ТОДОС користувачу необхідні знання базового набору операцій при роботі з операційною системою Windows: вміння запускати програми, працювати з вікнами та стандартними елементами вікон (меню, закладки, кнопки, випадаючими списками, полями внесення даних тощо); основами роботи із мережею Internet: вміння переходити до web-сторінок, користуватись стандартними функціями (завантажити файл, зчитати, видалити тощо), володіти найпростішими навичками роботи з пошуковими web-системами а також вміння

працювати з текстовим редактором MS Word та володіння основами роботи з табличним процесором MS Excel.

1.3. Функціональний склад системи ТОДОС

Функціонально систему ТОДОС складають наступні підсистеми:

КОНСПЕКТ (TextTermin) – побудова термінологічних дерев на основі аналізу природно-мовного тексту;

ЕДИТОР (Graph Editor) – формування онтологічних моделей;

КОНСПЕКТ (TextTermin)

Програмний модуль «Конспект (TextTermin)» виконує наступні функції:

- лінгвістичний аналіз тексту до рівня поверхневого синтактико-семантичного аналізу;
- виділення термінів предметної області з релевантних текстів;
- виділення і стисле конспектування фрагментів природно-мовних текстів, що належать до заданої тематики, яка задається ключовим словом або словосполученням;
- генерацію за результатами семантичного аналізу заданого числа вторинних ключів, використання яких в циклічному режимі дозволяє поглибити розкриття тематики у сформованих конспектах;
- використання стислих тематичних конспектів для вибору з множини текстових документів таких, які найбільшою мірою релевантні заданій тематиці.

Результатом роботи Конспекту є виділення з тексту термінів та зв'язків між ними, відображення їх у вигляді дерева, підрахунок частоти вживання термінів в тексті та їх контекстне оточення. Конспект також надає можливість побудови вихідних таблиць для Web-додатку Graph Editor.

ЕДИТОР (Graph Editor)

Web-додаток Graph Editor призначений для створення, редагування, перегляду та аналізу мереж понять і формування закономірностей, представлених у вигляді набору значень ознак, якими описуються початкові поняття предметної області.

Таким чином, опис предметної області є наочним і таким, що легко інтерпретується користувачем.

Виділення закономірностей відбувається методом індуктивного формування понять на основі пірамідальної мережі.

Пірамідальною мережею називається ациклічний орієнтований онтограф, вершинами якого виступають терміни (поняття) предметної області, а ребрами (дугами) – семантичні зв'язки між термінами.

Представлення інформації у вигляді онтографу дозволяє вивчати не лише окремий термін (поняття), але й отримувати всі його семантичні зв'язки з іншими поняттями, тим самим осмислюючи його роль у даній системі знань чи в ході розв'язання задачі. Алгоритм побудови онтографу у системі ТОДОС.

Підготовчий етап

Текст, для якого проводиться лінгвістичний аналіз, зберігається у форматі *.txt. В залежності від розміру, великий (понад 1 МБ), або різномірний за структурою контенту текст може бути розбитий на окремі файли згідно структури документу (наприклад, за змістом).

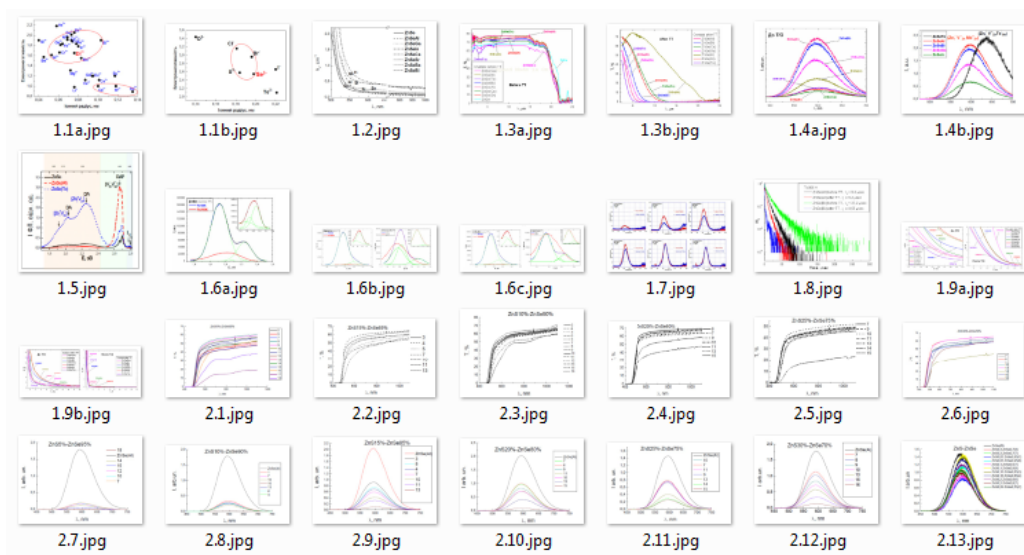


Рис. 1.1. Збереження рисунків і таблиць в окремих файлах

Рисунки й таблиці з тексту, що аналізується, потрібно зберегти в окремі файли формату *.jpg (*.JPEG, *.gif, *.png) – рис. 1.1– та завантажити їх на сервер для забезпечення можливості доступу до них ззовні (через мережу Internet).

На рис. 1.2 показано розташування рисунків на сервері власника онтографу, проте, для розташування рисунків може бути використаний будь який інший сервер (рис. 1.3).

Index of /storage/NASU/tomograf/pics Index of /storage/NASU/tomograf/tabs

Name	Last modified	Size	Description	Name	Last modified	Size	Description
Parent Directory	-	-	-	Parent Directory	-	-	-
1.1a.jpg	17-Feb-2016 12:36	56K		1.1.jpg	17-Feb-2016 12:45	51K	
1.1b.jpg	17-Feb-2016 12:37	38K		1.2.jpg	17-Feb-2016 12:47	85K	
1.2.jpg	17-Feb-2016 12:37	68K		1.jpg	17-Feb-2016 12:44	99K	
1.3a.jpg	17-Feb-2016 12:38	73K		2.1.jpg	17-Feb-2016 12:51	91K	
1.3b.jpg	17-Feb-2016 12:39	73K		2.2.jpg	17-Feb-2016 12:59	108K	
1.4a.jpg	17-Feb-2016 12:40	70K		3.1.jpg	17-Feb-2016 13:15	121K	
1.4b.jpg	17-Feb-2016 12:40	73K		3.2.jpg	17-Feb-2016 13:20	160K	
1.5.jpg	17-Feb-2016 12:41	94K		3.3.jpg	17-Feb-2016 13:27	125K	
1.6a.jpg	17-Feb-2016 12:42	45K		4.1.jpg	17-Feb-2016 13:29	68K	
1.6b.jpg	17-Feb-2016 12:42	61K		4.2.jpg	17-Feb-2016 13:31	143K	
1.6c.jpg	17-Feb-2016 12:43	54K		4.3.jpg	17-Feb-2016 13:34	99K	
1.7.jpg	17-Feb-2016 12:46	134K		4.4.jpg	17-Feb-2016 13:35	44K	
1.8.jpg	17-Feb-2016 12:46	72K		4.5.jpg	17-Feb-2016 13:39	151K	
1.9a.jpg	17-Feb-2016 12:47	88K		4.6.jpg	17-Feb-2016 13:51	81K	
1.9b.jpg	17-Feb-2016 12:47	71K		4.7.jpg	17-Feb-2016 13:52	105K	
2.1.jpg	17-Feb-2016 12:48	52K		5.1.jpg	17-Feb-2016 14:06	112K	

Рис. 1.2. Розміщення файлів з рисунками і таблицями на сервері НАН

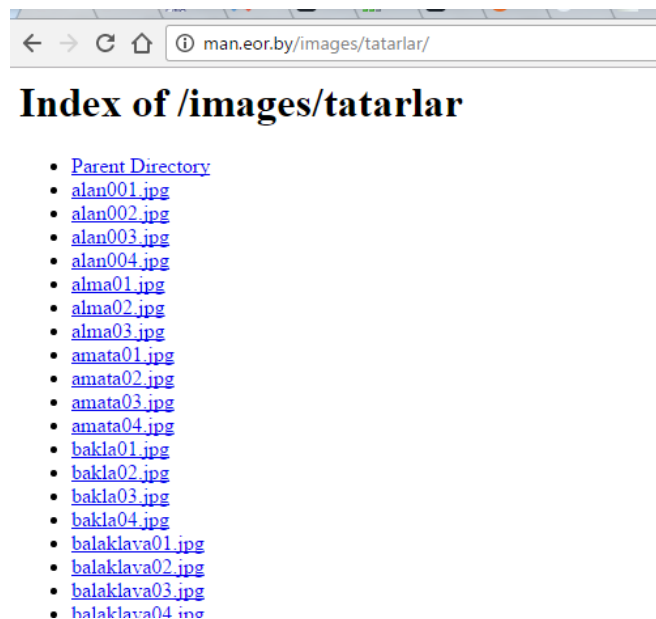


Рис. 1.3. Розміщення файлів з рисунками на сервері бібліотеки eor.by

Етап аналізу тексту

На цьому етапі відбувається виділення термінів і контекстів та групування їх у класи.

Аналіз тексту полягає у створенні таблиці класів термінів (рис. 1.4).

Історія дитячої книги	книга	ілюстрація	література	казка	дитина	мистецтво	худо
книга	книжка	дитяча книга	текст книги	мистецтво кни	майстри мисте	авторська кн	перс
ілюстрація	дитяча ілюстрація	історія дитячої ілюст	історія дитячої к	дитяча книжко	книжкова ілюс	класика кни	мож
література	дитяча література	російська література	історія російськ	література нов	нова літератур	світова літер	худо
казка	літературна казка	сучасна казка	особливості каз	народна казка	казка Пушкіна	казка Ершов	казк
дитина	виховання дітей	світ дитини	знайомство дит	дитяча специф	внутрішній світ	психологія д	особ
мистецтво	вид мистецтва	літографічне мистец	мистецтво книги	майстер мисте	елементи мист	явище мисте	пред
художник	роботи художника	художник Білібін	художники XX ст	художники фор	художники фор	художники ф	поче
століття	третина століття	остання третина сто	століття ілюстра	XVIII століття	XIX століття	XX століття	XXI (
картинка	Азбука в картинка	цікава картинка	гравірована карт	Фізика у карти	Розповіді в кар		
письменник	дитячий письменн	відомий письменник	сучасний письмен	стиль письмен			
ілюстратор	ілюстратори тридц	найпопулярніший ілк	книжковий ілюст	ілюстратор дит			
текст	текст книги	шрифт тексту	автор тексту	стиль тексту			

Рис. 1.4. Таблиця класів термінів

Наповнення таблиці має таку структуру: у верхній лівій комірці знаходиться назва класу всіх термінів документа, що аналізується (це, наприклад, може бути назва самого документу – див.рис4.) – клас термінів A-Z. Далі в комірках йде перелік виділених класів термінів (груп термінів, виокремлених експертами).

Починаючи з другого рядка, перша комірка кожного з рядків теж містить термін, який є ім'ям класу термінів, таким чином, перший рядок і перший стовпчик однакові за змістом. Інші комірки відповідних рядків містять терміни, які є об'єктами даних класів.

(клас термінів A-Z), (термін A), (термін B), (термін C), ..., (термін Z)

(термін A), (термін A₁), (термін A₂), (термін A₃), ..., (термін A_n)

(термін B), (термін B₁), (термін B₂), (термін B₃), ..., (термін B_k)

.....

(термін Z), (термін Z₁), (термін Z₂), (термін Z₃), ..., (термін Z_k).

Для подальшого використання ця таблиця має бути збережена до файлу у форматі *.csv. Для таблиці з рис 4. перші рядки файлу виглядають так (рис. 1.5):

Тобто йде перелік змістів комірок, відокремлених комами. Якщо в комірці таблиці зміст відсутній (наприклад, 13-й рядок) то між комами нічого не ставиться (рис.

1.6):

Історія дитячої книги, книга, ілюстрація, література, казка, дитина, мистецтво, художник, книга, книжка, дитяча книга, текст книги, мистецтво книги, майстри мистецтва книги, автор ілюстрація, дитяча ілюстрація, історія дитячої ілюстрації, історія дитячої книжкової література, дитяча література, російська література, історія російської літератури, літературна казка, сучасна казка, особливості казки, народна казка, казка Пушкіна, дитина, виховання дітей, світ дитини, знайомство дитини з життям, дитяча специфіка, вч мистецтво, вид мистецтва, літографічне мистецтво, мистецтво книги, майстер мистецтва і мистецтва роботи мистецтва мистецтва Білібіна мистецтвах VV століття мистецтвах форм

Рис. 1.5. Файл класів термінів

малюнок, малюнок, тринадцять років, найменування, текст, текст книги, шрифт тексту, автор тексту, стиль тексту, , , , , малюнок, лубочний малюнок, цінність малюнка, малюнок Сутеева, ма

Рис. 1.6. Файл класів термінів, 13-й рядок

Далі файл зберігається в текстовому форматі і розширення змінюється на «csv». Також, такий файл може одразу бути створений за допомогою табличного процесору MS Excel і збережено у форматі «CSV (Разделители - запятыє)».

Засобом автоматизації аналізу тексту тексту, що аналізується, та створення таблиці класів термінів є модуль семантичного аналізу природно-мовних текстів «КОНСПЕКТ» системи ТОДОС (рис. 1.7).

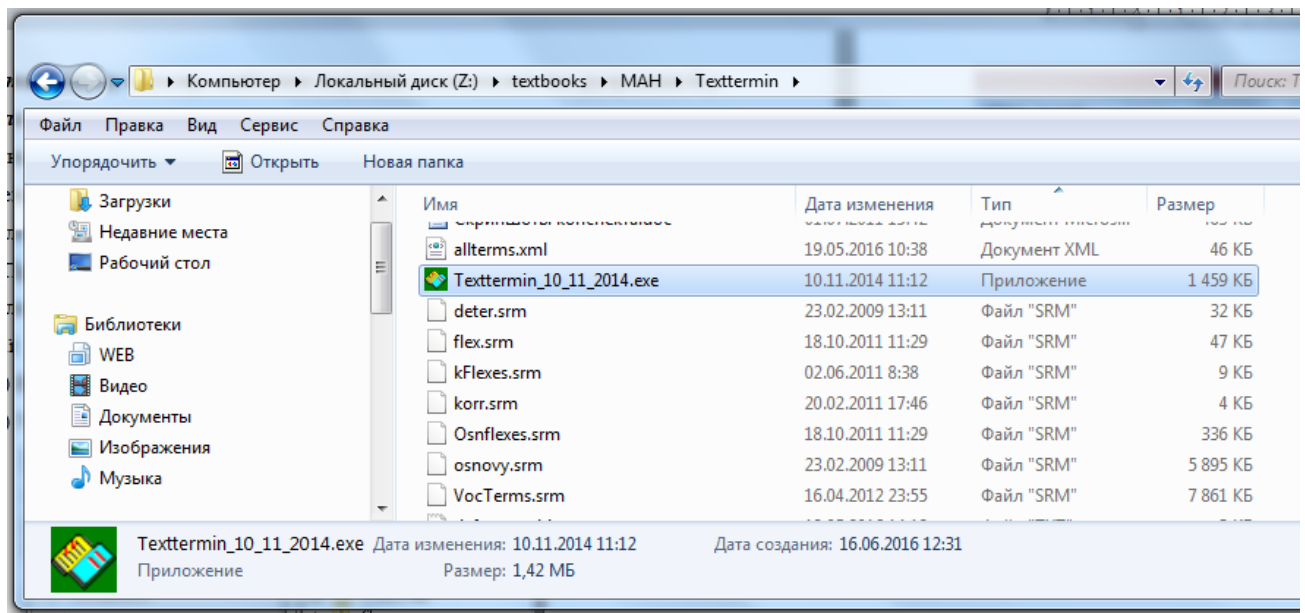


Рис. 1.7. Модуль «КОНСПЕКТ» (TextTermin, версія від 10.11.2014)

За допомогою модулю «КОНСПЕКТ» (TextTermin) будується дерево термінів – фактично відбувається автоматичне виокремлення наявних в тексті термінів та їх ранжування за частотою використання. Визначається контекст використання

термінів та на основі отриманих даних експерт отримує можливість створення таблиці класів термінів.

У разі, розбиття тексту, що аналізується, на окремі файли, після аналізу кожного окремого розділу (пунктів розділів) об'єднуємо отримані таблиці класів термінів в єдину (зведену) таблицю класів.

Побудова таблиці контекстів

	1	2	3	4
1	nodedata			
2	народ	В современном мире насчитывается около 800 наций.	текст	text
3	народ	Курултай крымскотатарского народа	http://man.eor.by/images/tatarlar/narod001.jpg	href
4	народ	День крымскотатарского флага	http://man.eor.by/images/tatarlar/narod002.jpg	href
5	народ	Крымские татары	http://man.eor.by/images/tatarlar/narod004.jpg	href
6	народ	Когда историю неугодного народа умышленно замалч	текст	text
7	народ	Крымские татары	http://man.eor.by/images/tatarlar/narod006.jpg	href
8	народ	Поволжские татары	http://man.eor.by/images/tatarlar/narod007.jpg	href
9	народ	И все же нельзя сказать, что этногенезом не	текст	text
10	народ	Крымские татары: национальная одежда	http://man.eor.by/images/tatarlar/narod005.jpg	href
11	народ	Крымские татары: национальная одежда	http://man.eor.by/images/tatarlar/narod008.jpg	href
12	народ	В современном мире насчитывается около 800 наций.	текст	text
13	народ	Крымские татары 18 век	http://man.eor.by/images/tatarlar/narod003.jpg	href
14	народ	Золотая Орда и Русь	http://man.eor.by/images/tatarlar/narod009.jpg	href
15	тавры	В современном мире насчитывается около 800 наций.	текст	text
16	тавры	Карта территории Тавров	http://man.eor.by/images/tatarlar/tavry01.jpg	href
17	тавры	Таврский менгир	http://man.eor.by/images/tatarlar/tavry02.jpg	href
18	киммерийцы	В современном мире насчитывается около 800 наций.	текст	text
19	киммерийцы	Киммерийские воины	http://man.eor.by/images/tatarlar/kimmerijtsy01.jpg	href
20	киммерийцы	Киммерийцы на греческой вазе	http://man.eor.by/images/tatarlar/kimmerijtsy02.jpg	href

Рис. 1.8. Фрагмент таблиці контекстів

На основі отриманої таблиці класів термінів в середовищі MS Excel проводиться побудова та наповнення таблиці контекстів. До цієї таблиці додаються адреси ілюстрацій термінів – збережених на сервері рисунків та таблиць (рис. 1.8).

За наявності декількох таблиць контекстів вони зводяться до єдиної (зведеної) таблиці контекстів.

Етап генерації таксономії.

Генерація таксономії тексту відбувається автоматично модулем КОНФОР двома способами.

Перший спосіб. Модуль КОНФОР вбудовано в модуль ЕДИТОР, і всі наступні дії виконуються в середовищі останнього (<http://editor3.inhost.com.ua>).

Для генерації таксономії тексту завантажуюмо зведену таблицю класів термінів формату *.xls до модулю ЕДИТОР.

За умови правильного формування зведеної таблиці класів термінів в онтографі не утворюються множинні зв'язки. В протилежному випадку такі зв'язки мають місце і повинні бути усунені.

Наступним кроком є наповнення таксономії контекстами та перетворення її на онтологію.

У випадку правильно сформованої таблиці контекстів термінів система виведе повідомлення про кількість зчитаних об'єктів даних, яка співпадає з кількістю рядків таблиці з файлу *.csv.

За такого способу генерації таксономії модулем КОНФОР можуть бути утворені так звані неозначені вершини. У випадку, коли запропонований варіант представлення результатів НДР призначений не для експертного аналізу (авторів тексту, керівників тощо), а для пересічного користувача (потенційного замовника, молодих науковців (аспірантів), науковців суміжних галузей тощо), неозначені вершини можуть бути видалені. Для цього використовується другий спосіб генерації таксономії тексту.

Другий спосіб. Необхідно перейти за адресою <http://ontology.inhost.com.ua/confor-direct/>, завантажити файл структури і натиснути кнопку «Обробити»..

Етап публікації онтології

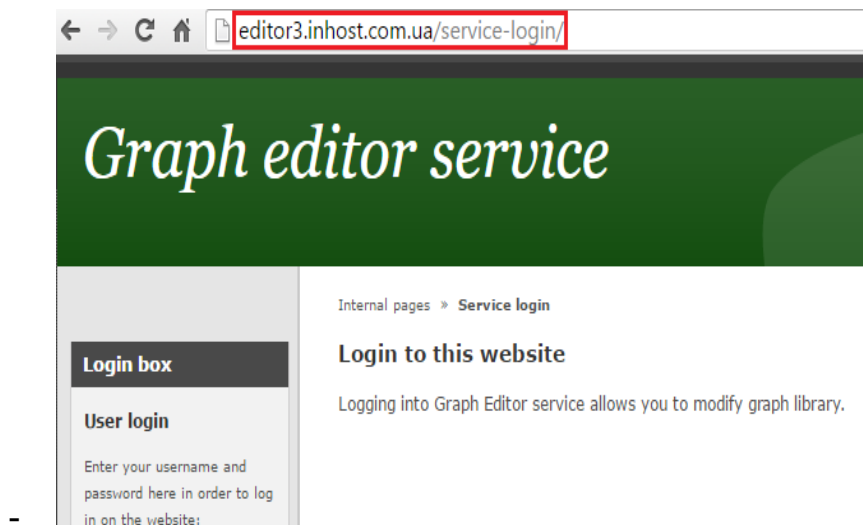


Рис. 1.9. Інтерфейс бібліотеки онтологій на сервері

Для візуалізації онтології у вигляді аплікаційної форми її необхідно опублікувати на сервері в бібліотеці онтологій (<http://editor3.inhost.com.ua/service-login/>) (рис. 1.9).

Для роботи з адміністративною частиною бібліотеки онтографів необхідно авторизуватися за допомогою форми логін-пароль (рис. 1.10).

Login box

User login

Enter your username and password here in order to log in on the website:

Username:

Password:

Stay logged in: ☐

[Login](#)

[Forgot your password?](#)

Рис. 1.10. Форма авторизації

Далі потрібно перейти у вкладку **folder structure**. На сторінці, що після цього відкриється, представлена структура папок бібліотеки (рис. 1.11). Назву кожної з папок можна змінити, натиснувши на кнопку "Edit", зайти всередину (рис. 1.12), або видалити, якщо папка порожня.

Folder structure »

Internal pages » Service login » Folder structure

Root folder

№	Folder Name	Edit	Delete
2	Misc	Edit	Delete
3	HISTORY	Edit	Delete
4	MINERALS	Edit	Delete
5	NOBEL	Edit	Delete
6	PHYSICS	Edit	Delete
7	RESEARCH	Edit	Delete
8	КЛАССИФИКАТОР	Edit	Delete
9	КЛИМОВА	Edit	Delete
10	Місто майбутнього	Edit	Delete
11	ШЕВЧЕНКО	Edit	Delete
12	МЕДИЦИНА	Edit	Delete
13	SCENARIO	Edit	Delete
14	Ecology	Edit	Delete
15	СППР	Edit	Delete
16	Жемчужинка	Edit	Delete

Login box

Current status

This is your current status:
 Username: feadmin

[Logout](#)

Рис. 1.11. Папка в бібліотеці онтологій, до якої надається доступ за логіном та паролем

Навігація по дереву папок відбувається шляхом натискання на назви папок лівої клавішею мишки. При знаходженні в папці над списком папок є маркер "Up", що переводить режим перегляду в батьківську папку.



Рис. 1.12. Папка «Харчова промисловість» з підпапками та файлами

В папці, доступ до якої надано, можна створити підпапку «New Folder», задавши її ім'я в рядку «name (required)» та натиснувши кнопку «Create new» (рис. 1.13 а), або додати файл з онтографом чи онтологією «New Graph», задавши його ім'я в рядку «Name», символічний код (латинський або кириличний) в рядку «Symbol code» та обравши створений файл на локальному носії при натисненні на кнопку «File» і натиснувши кнопку «Create new» (рис. 1.13 б).

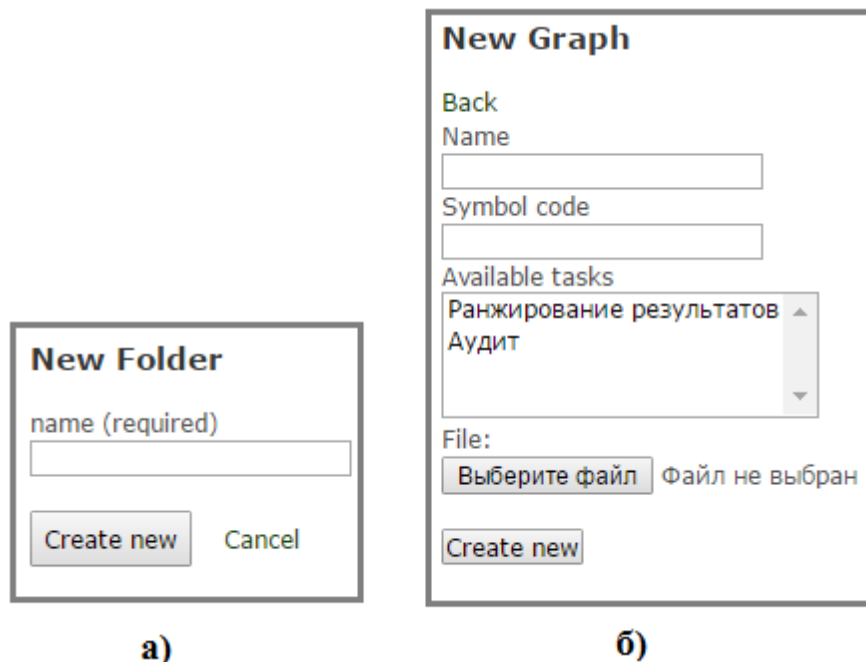


Рис. 1.13. Створення підпапки та нового файлу в бібліотеці онтологій
Завантажений файл відобразиться в папці в бібліотеці онтологій (рис. 1.14).

New Folder New Graph

Рис. 1.14. Завантажений файл в папці в бібліотеці онтологій
Для кожного онтографу в поточній папці існують такі функції:

- 1) Назва онтографа, що відображається у списку Graph Editor. Натискання на це поле відкриває онтограф в новій вкладці. При цьому посилання містить ідентифікатор онтографа;

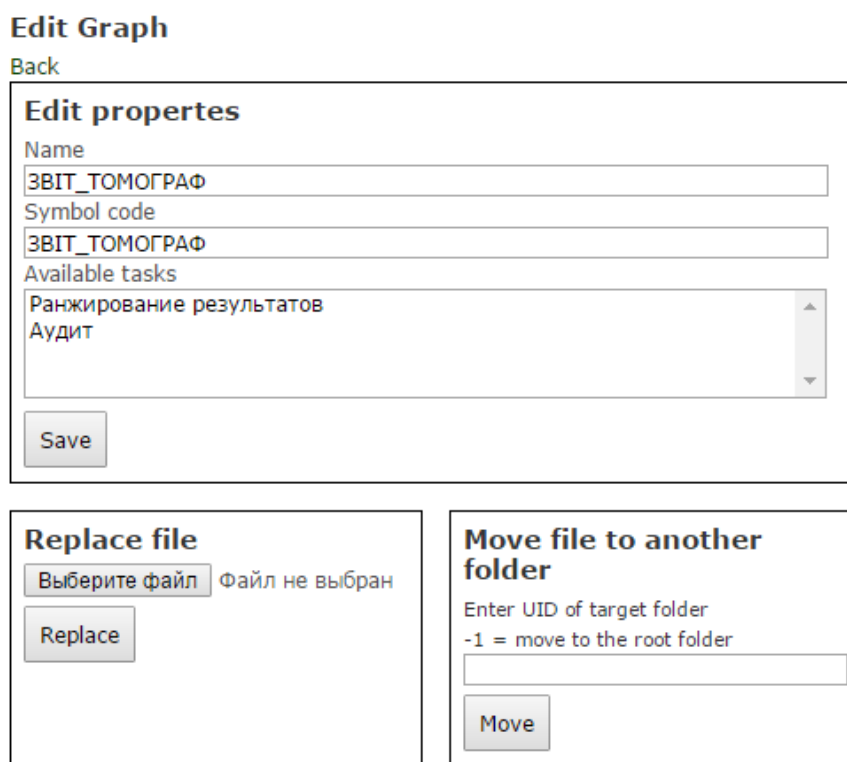


Рис. 1.15. Функція редагування файлу

- 2) Символьний код онтографу,. Це поле унікальне для кожного з онтографів. Натискання на це поле відкриє онтограф в новій вкладці. При цьому посилання міститиме символьний код онтографа;
- 3) HTML5 - гіперпосилання, що дозволяє переглянути онтограф за допомогою HTML5-браузера. (дана функція працює в тестовому режимі);
- 4) Edit – функція редагування онтографу (рис. 1.15):
 - а) Зміна параметрів онтографу - назви, символьного коду і списку задач, які можуть бути до нього застосовані.
 - б) Оновлення XML-файлу.

с) Перенесення файлу в іншу папку шляхом введення її унікального ідентифікатора, що показується в списку папок.

5) Delete – функція видалення онтографу.

Після розміщення файлу з онтологією проаналізованого тексту в папці бібліотеки онтологій (публікації на сервері) онтологія стає доступною для перегляду в режимі апікацій (рис. 1.16).

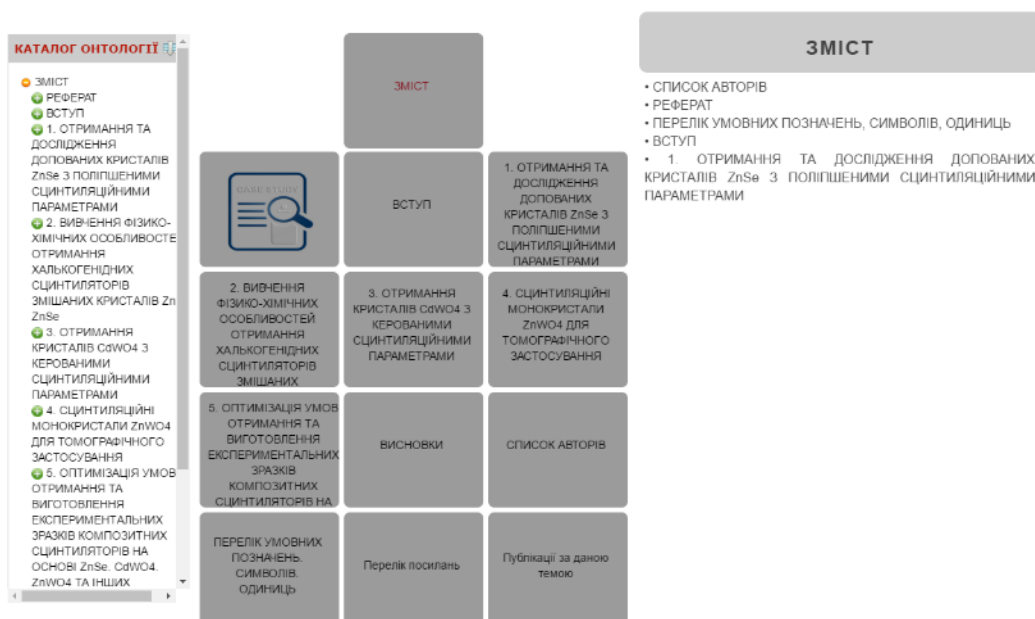


Рис. 1.16. Апікаційна форма представлення онтології (верхній рівень онтології)

2. КОНСПЕКТ (Texttermin) – побудова дерева термінів

2.1. Підготовка ОС Windows до роботи із Texttermin.

Розглянемо підготовку операційної системи до роботи із програмою Texttermin на прикладі ОС Windows-7.

Для роботи з програмою виділення термінів необхідно забезпечити відображення розширень для стандартних типів файлів. Для цього потрібно у сервісному меню папок зняти прапорець «Приховувати розширення для зареєстрованих типів файлів».

Для виклику даного пункту меню необхідно в ОС Windows-7 виконати наступні дії:

- 1) Відкрити будь-яку папку файлової системи (наприклад, «Мой компьютер»).
- 2) У меню «Упорядочить» обрати «Представление» - «Строка меню» (рис. 2.1).

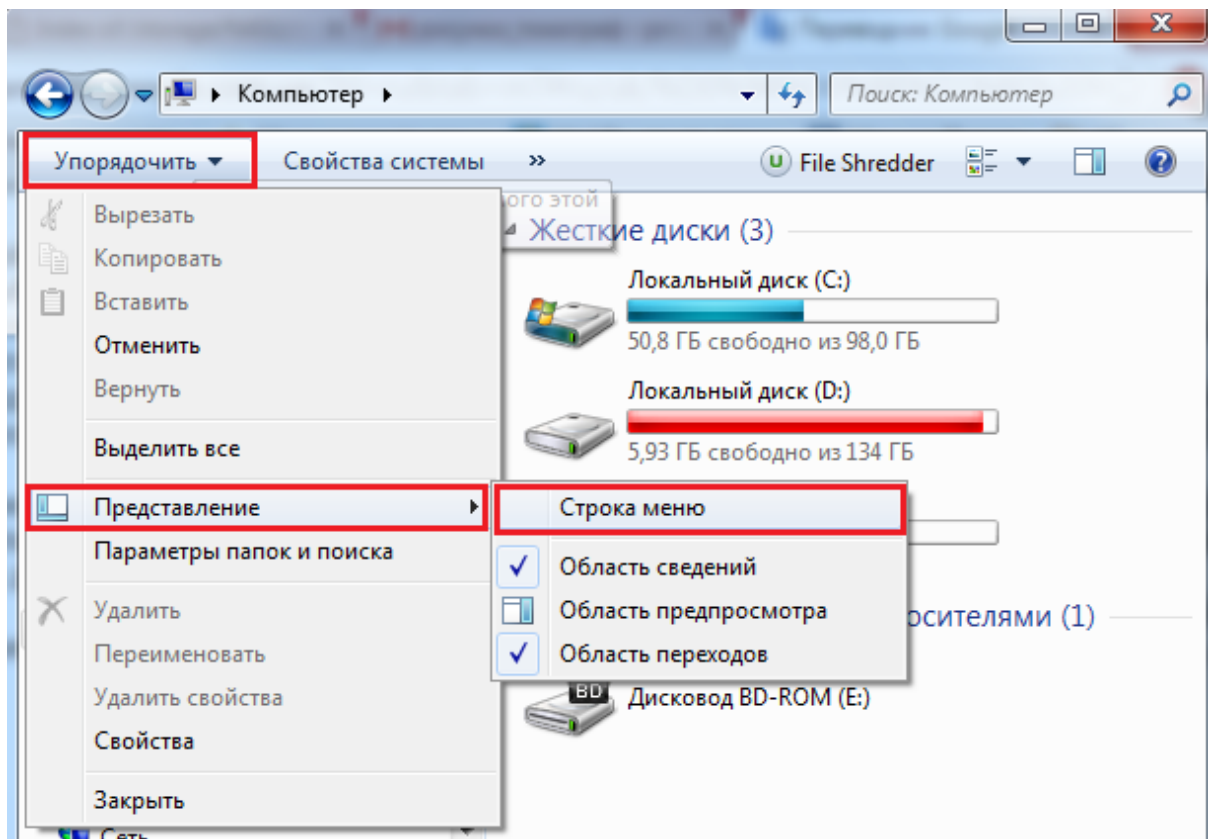


Рис. 2.1. Меню «Упорядочить» - «Представление» - «Строка меню»

- 4) У рядку меню «Сервис» обрати «Параметры папок» (рис. 2.2).

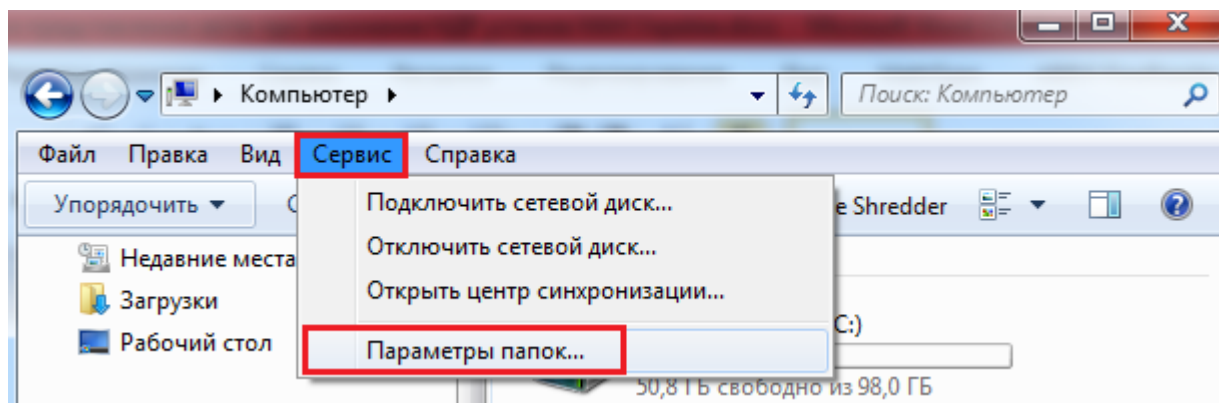


Рис. 2.2. Меню «Сервис» - «Параметры папок»

5) У вікні обрати вкладку «Вид».

6) У пункті «Дополнительные параметры» зняти прапорець «Скрывать расширения для зарегистрированных типов файлов» (рис. 2.3).

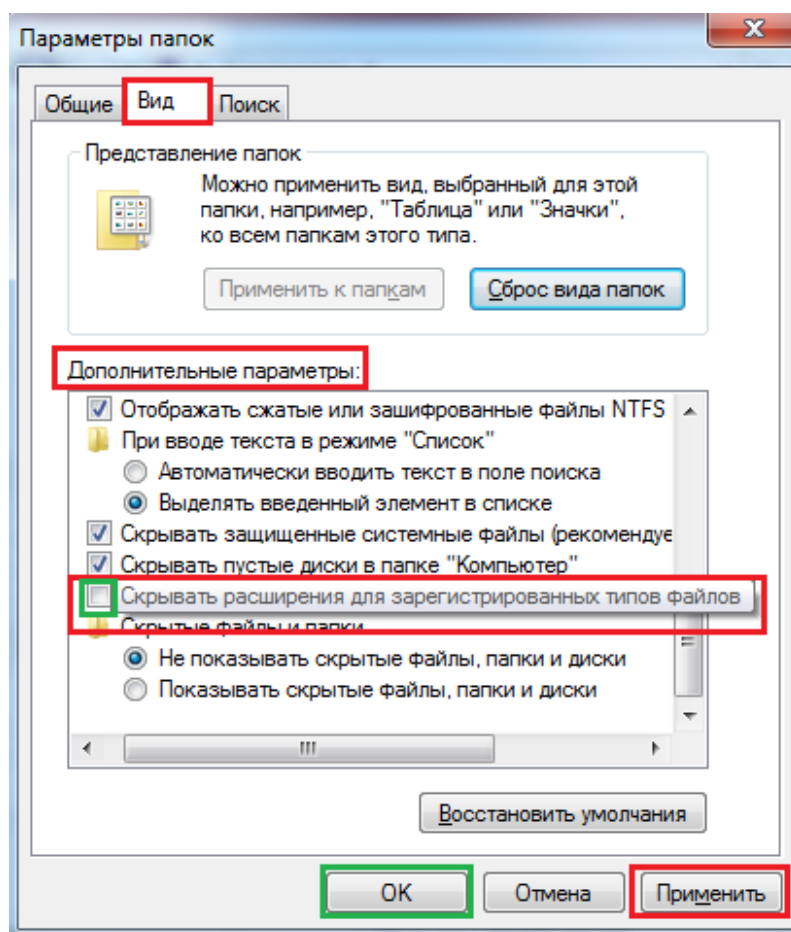
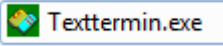


Рис. 2.3. Пункт «Дополнительные параметры» - «Скрывать расширения для зарегистрированных типов файлов» зняти прапорець

7) Натиснути кнопки «Применить» і «Ок».

2.2. Інтерфейс програми Texttermin

Після запуску програми  (Texttermin_10_11_2014.exe) відкривається вікно наступного вигляду (рис. 2.4).

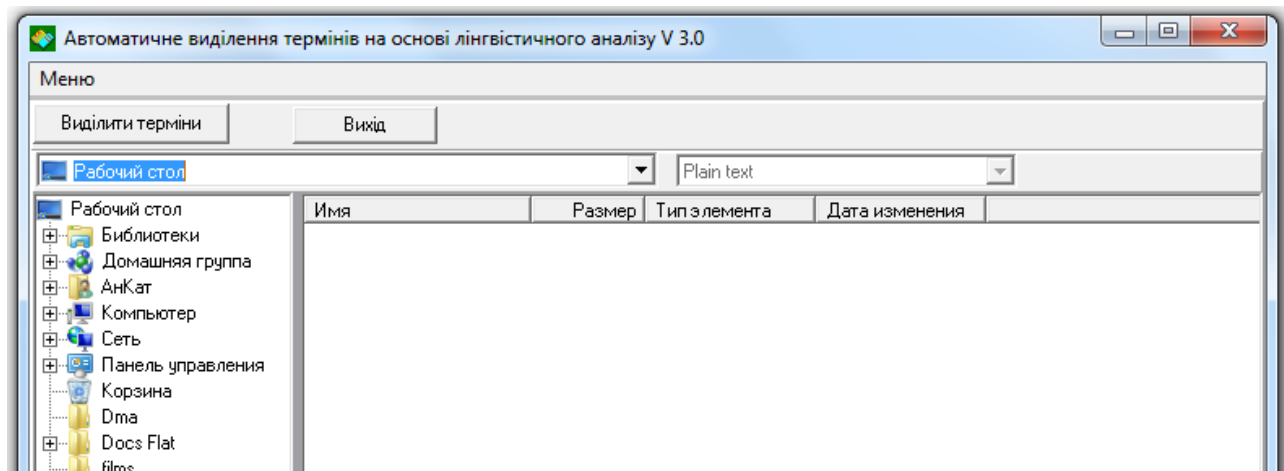


Рис. 2.4. Вікно вибору тексту для аналізу

Після вибору текстового файлу для аналізу натискаємо кнопку

Виділити терміни

. Відкривається вікно, зображене на рис. 2.6.

Інтерфейс користувача умовно розділений на 3 частини:

В першій частині у верхньому вікні «Дерево термінів» (рис. 2.5) виводиться власне термін, що розкривається вниз або вгору, та пов'язані з ним певними зв'язками.

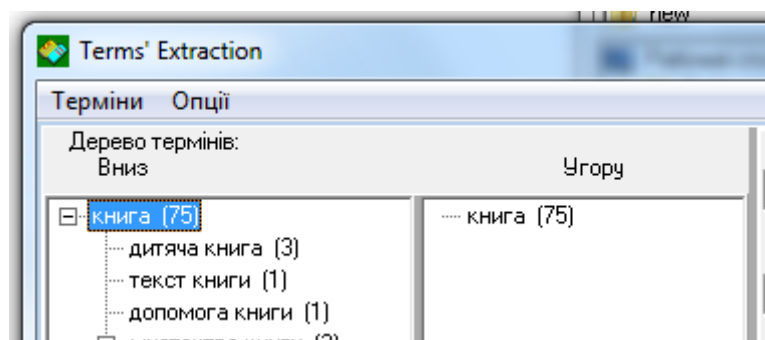


Рис. 2.5. Вікно «Дерево термінів»

В нижньому вікні першої частини інтерфейсу виводяться контексти, в яких зустрічаються обрані терміни, із тексту, що аналізується. При встановленому прапорці «підкреслювати терміни», у вікні контекстів обрані терміни виводяться підкресленими (рис. 2.6).

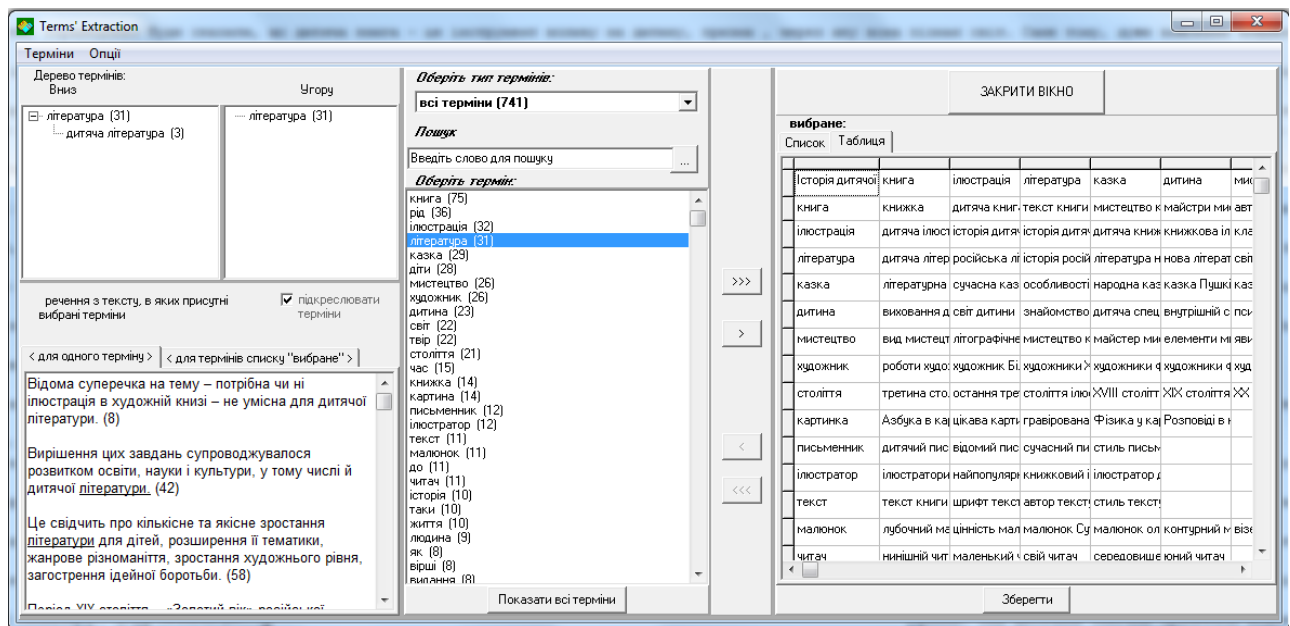


Рис. 2.6. Інтерфейс користувача модулю «КОНСПЕКТ»

В другій частині інтерфейсу з випадаючого меню можна обрати тип терміну (аббревіатура, іменник, іменник+іменник, прикметник+ іменник тощо), що виводитиметься в списку (рис. 2.7) та побачити кількість їх вживання в тексті (цифра в дужках праворуч від типу терміну). За замовчуванням виводяться всі терміни.

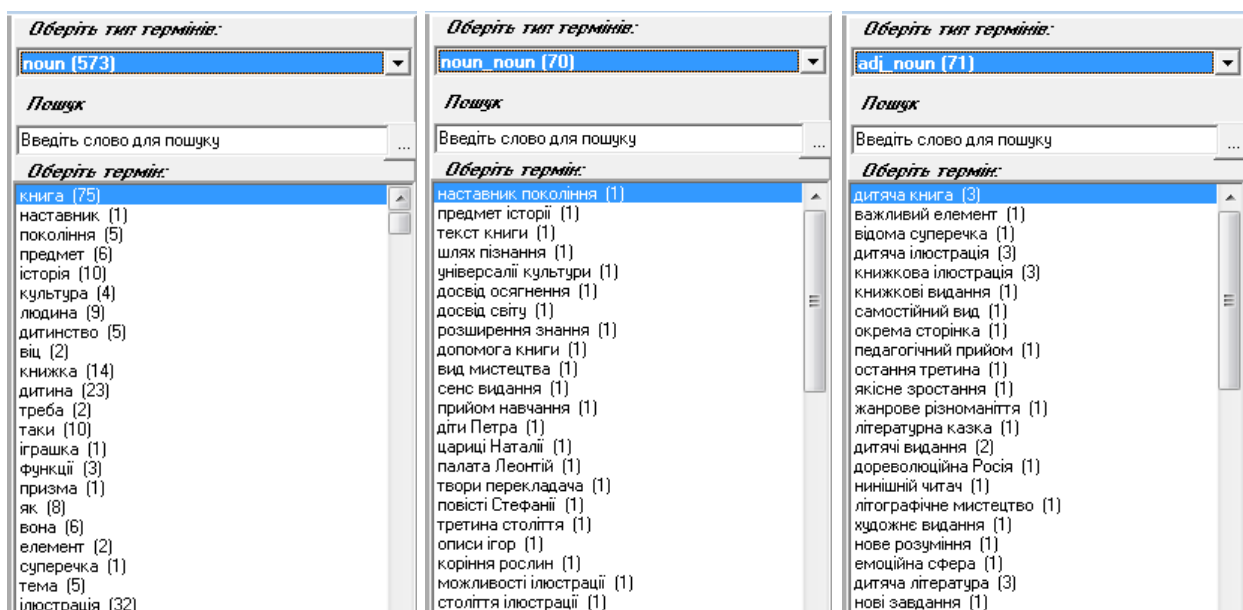


Рис. 2.7. Вибір типу термінів: іменник (573), іменник+іменник (70),
прикметник+іменник (71)

У вікні «Пошук» здійснюється відбір термінів за заданим ім'ям.

В списку термінів останні виводяться в порядку зменшення кількості згадувань в тексті, яка зазначена в дужках праворуч від терміну (рис. 2.6, 3.7).

Третя частина інтерфейсу являє собою таблицю класів термінів, в яку переносяться терміни, виділені з тексту. Схема наповнення таблиці вже наводилась нами раніше:

(клас термінів A-Z), (термін A), (термін B), (термін C), ..., (термін Z)

(термін A), (термін A₁), (термін A₂), (термін A₃), ..., (термін A_n)

(термін B), (термін B₁), (термін B₂), (термін B₃), ..., (термін B_k)

.....

(термін Z), (термін Z₁), (термін Z₂), (термін Z₃), ..., (термін Z_k).

2.3. Наповнення таблиці класів термінів

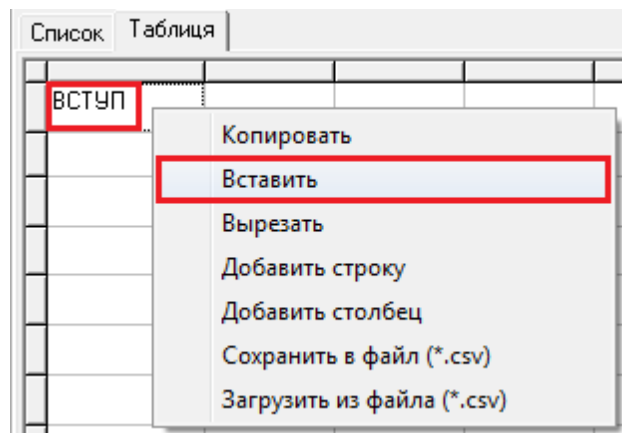


Рис. 2.8. Перший крок наповнення таблиці термінів

Наповнення таблиці класів термінів починається із занесення в першу комірку назви розділу (пункту) тексту, що аналізується, шляхом його копіювання та вставки за допомогою контекстного меню (рис. 2.8).

В наступні комірки першого рядку заносяться терміни, виділені КОНСПЕКТОМ, на розсуд експерта, що проводить аналіз тексту, шляхом перетягування терміну з вікна «Дерево термінів» (рис. 2.9) або його копіювання з вікна контекстів (рис. 2.9).

Контекстне меню, що викликається натисканням правої кнопки миші на таблиці, дозволяє копіювати, вставляти, вирізати терміни в/з таблиці, розширювати таблицю додаванням рядків та стовпчиків, зберігати сформовану таблицю у файл формату *.csv та завантажувати його для редагування таблиці (рис. 2.10).

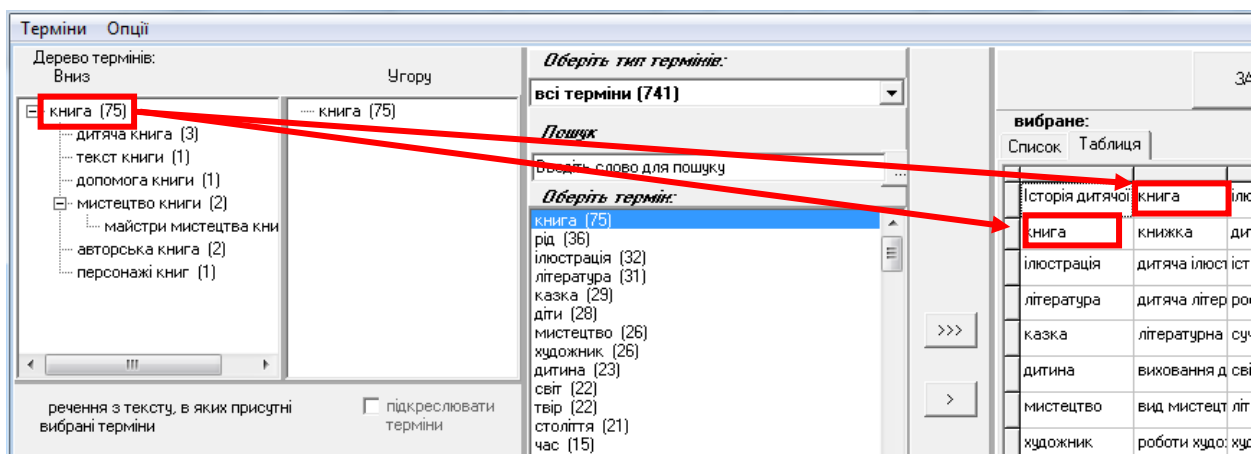


Рис. 2.9. Перетягування терміну з вікна «Дерево термінів» до таблиці класів термінів

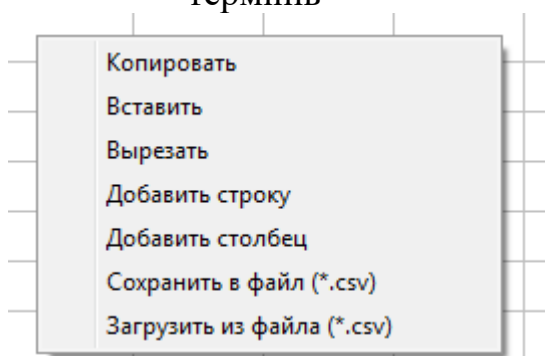


Рис. 2.10. Контекстне меню роботи із таблицею

Копіювання терміну з вікна контекстів до таблиці класів термінів показано на рис. 2.11.

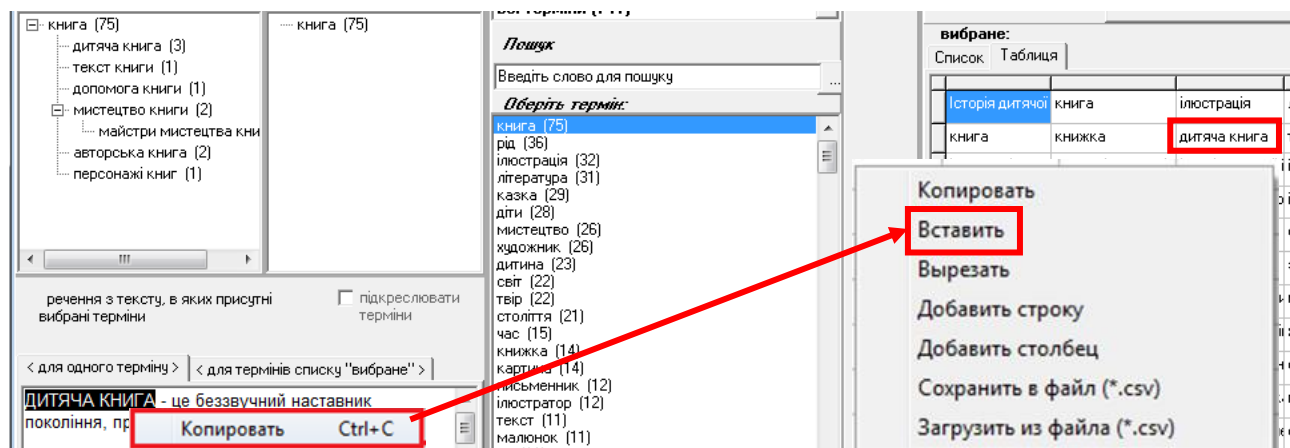


Рис. 2.11. Копіювання терміну з вікна контекстів до таблиці класів термінів

Одночасно із наповненням таблиці класів термінів відбувається наповнення таблиці контекстів, які в подальшому відіграватимуть роль наповнення онтології.

3. Використання MS Excel для побудови таблиці контекстів

Наповнення таблиці контекстів відбувається в середовищі MS Excel за наступною схемою (рис. 3.1).

	A	B	C	D
1	nodedata	Системна команда для модулю Едитор		
2	термін 1	текстовий опис терміну	text	text
3	термін 1	назва рисунку/таблиці, що ілюструє термін	посилання на рисунок/таблицю.jpg (*.JPEG, *.gif, *.png)	href
4	термін 2	текстовий опис терміну	text	text
5	термін 2	назва рисунку/таблиці, що ілюструє термін	посилання на рисунок/таблицю.jpg (*.JPEG, *.gif, *.png)	href
6	термін 2	назва рисунку/таблиці, що ілюструє термін	посилання на рисунок/таблицю.jpg (*.JPEG, *.gif, *.png)	href
7	термін 2	назва рисунку/таблиці, що ілюструє термін	посилання на рисунок/таблицю.jpg (*.JPEG, *.gif, *.png)	href
	Імена термінів з таблиці класів	Текстові описи та назви рисунків і таблиць з вікна контекстів модулю КОНСПЕКТ	Посилання на файли з рисунками і таблицями, збережені на сервері	

Рис. 3.1. Схема наповнення таблиці контекстів

Файл з таблицею контекстів зберігається у форматі «*.CSV (Разделители – запятыє)».

Зверніть увагу, якщо до одного й того самого терміну відноситься наповнення різних типів (текстовий опис та посилання на рисунок/таблицю), воно заноситься у різні рядки, а ім'я терміну в першому стовпчику дублюється (рис. 3.2. Наприклад, рядки 4-8).

	1	2	3	4
1	nodedata			
2	Історія дитячої книги	Історія дитячої книги зі стародавніх часів	text	text
3	Історія дитячої книги	Історія дитячої книги зі стародавніх часів	http://man.eor.by/images/childre	href
4	«Буквар» Каріона Істоміна	Першою ілюстрованою дитячою книгою	text	text
5	«Буквар» Каріона Істоміна	«Буквар» Каріона Істоміна	http://man.eor.by/images/ariclou	href
6	«Буквар» Каріона Істоміна	«Буквар» Каріона Істоміна	http://man.eor.by/images/ariclou	href
7	«Буквар» Каріона Істоміна	«Буквар» Каріона Істоміна	http://man.eor.by/images/ariclou	href
8	«Буквар» Каріона Істоміна	«Буквар» Каріона Істоміна	http://man.eor.by/images/ariclou	href
9	20 століття	Наприкінці 20 століття спостерігається з	text	text

Рис. 3.2. Приклад наповнення таблиці контекстів

По закінченню наповнення таблиці класів термінів і збереження її у форматі *.csv за допомогою контекстного меню модулю КОНСПЕКТ необхідно привести її до формату зчитування модулем КОНФОР для генерації таксономії документу.

Для цього відкриваємо отриманий файл формату *.csv, виділяємо перший стовпчик, переходимо до вкладення «Данные», в пункті «Работа с данными» обираємо «Текст по столбцам» (рис. 3.3).

У вікні, що відкрилося, вказуємо формат даних «с разделителями» та натискаємо кнопку «Далее», в наступному вікні вказуємо, що роздільником є «запятая» та натискаємо кнопку «Далее», в останньому вікні натискаємо кнопку «Готово» (рис. 3.4). Терміни розподілено по комірках.

	A	B	C	D	E	F	G	H	I	J	K	L	
1	Історія ди	книга	ілюстраці	літератур	казка	дитина	мистецтв	художник	століття	картинка	письменн	ілюстратс	тен
2	книга	книжка	дитяча кн	текст кни	мистецтв	майстри	лаворська	персонаж	ілюстров	дитяча іл	художня	і книжкове	пол
3	ілюстрац	дитяча іл	історія ди	історія ди	дитяча кн	книжкова	класика к	можливо	століття	і художня	і виставка	і мистецтв	зн
4	літератур	дитяча лі	російська	історія ро	літератур	нова літе	світова лі	художня	і галузь лі	навчальн	ілюструє	світова ди	сфе
5	казка	літератур	сучасна к	особливо	народна	казка Пуц	казка Ерц	казка Чу	віршован	філософс	особливо	фантастична	
6	дитина	вихованн	світ дитин	знайомст	дитяча сп	внутрішні	психологі	особливо	дитинств	ставленн	естетичне	виховання	д
7	мистецтв	вид мисте	літографі	мистецтв	майстер	л елементи	явище ми	предмет	і образотв	мистецтв	твір мист	Світ мист	рад
8	художник	роботи ху	художник	художник	художник	художник	художник	почерк ху	художник	народний	російськи	знаменит	мо
9	століття	третина с	остання т	століття	і XVIII стол	XIX століт	XX століт	XXI століт	20 століття				
10	картинка	Азбука в	цікава к	гравірова	Фізика у	к Розповіді	в картинках						
11	письменн	дитячий п	відомий г	сучасний	стиль письменника								
12	ілюстрато	ілюстратс	найпопул	книжкови	ілюстратор	дитячих книг							
13	текст	текст кни	шрифт те	автор тек	стиль тексту								
14	малюнок	лубочний	цінність	м малюнок	малюнок	контурни	візерунча	яскравий	малюнок				
15	читач	нинішній	маленьки	свій читач	середови	юний читач							
16	історія	предмет	і шестидес	історія ди	історія іл	біблійна	і історія ро	історія ра	історія ди	історія про	звірів і	птахів	
17	життя	реальне	» сприйнят	т повнота	» характер	подальше	справа жи	активне	» повнота	життя			
18	людина	характер	доля люд	зображен	освічена	і маленька	ставлення	до людей					
19	вірш	улюблені	повчальн	вірш Бло	афористи	тема вірш	вірш Мар	вірш Міл	жартівливі	вірші			
20	видання	книжкове	дитяче ви	оформле	передпла	видання	» ілюстров	функції	дитячого	видання			
21	буквар	«Буквар»	Особовий	російський	буквар								
22	період	період ст	дореволюк	період ст	радянський	період							
23	сторінка	сторінка	сторінка	сторінка	сторінка	сторінка	сторінка	сторінка	сторінка	сторінка	сторінка	сторінка	сторінка

Рис. 3.3. Вкладення «Данные», пункт «Работа с данными» - «Текст по столбцам»

Далі викликаємо контекстне меню натисканням правої кнопки миші на виділеному стовпчику та обираємо пункт «Вставить» (рис. 3.5).

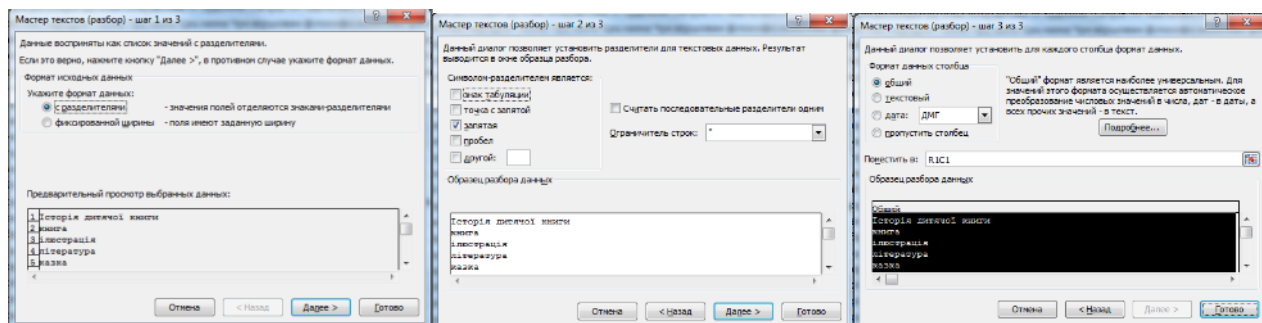


Рис. 3.4. Розподіл даних по комірках

Після цього копіюємо вміст стовпчика В до стовпчика А (рис. 3.6).

	A	B	C	D	E	F	G	H	I	J	K	L
1	Історі	 Вирезать		ратур	казка	дитина	мистецтв	художник	століття	картинка	письменн	ілюстрат
2	книга	 Копировать		т кни	мистецтв	майстри	лаворська	персонаж	ілюстров	дитяча іл	художня	і книжкове
3	ілюст	 Параметры вставки:		рія ди	дитяча кн	книжкова	класика к	можливо	століття	і художня	і виставка	і мистецтв
4	літер			ія ро	літератур	нова літе	світова лі	художня	і галузь лі	навчальн	ілюструє	світова ди
5	казка			бливо	народна	казка Пуц	казка Ерц	казка Чу	віршован	філософс	особливо	фантастичн
6	дитин	Специальная вставка...		іомст	дитяча сп	внутрішні	психологі	особливо	дитинств	ставленн	естетичне	виховання
7	мисте	Вставить		тецтв	майстер	л елементи	явище ми	предмет	і образотв	мистецтв	твір мист	Світ мист
8	худох	Удалить		ожник	художник	художник	художник	почерк ху	художник	народний	російськи	знаменит
9	столі	Очистить содержимое		іття	і XVIII стол	XIX століт	XX століт	XXI століт	20 століття			
10	карти	 Формат ячеек...		ірова	Фізика у	к Розповіді	в картинках					
11	письм			ний	стиль письменника							
12	ілюст	Ширина столбца...		ккови	ілюстратор	дитячих книг						
13	текст	Скрыть		р тек	стиль тексту							
14	малю	Показать		юнок	малюнок	контурни	візерунча	яскравий	малюнок			
15	читач			юний	читач	середови	юний читач					

Рис. 3.5. Додавання першого стовпчика

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	історія	історія	ді	історія	ді	історія	ді	історія	ді	історія	ді	історія	ді	історія	ді
2	книга	книга	книжкова	книжкова	книжкова	книжкова	книжкова	книжкова	книжкова	книжкова	книжкова	книжкова	книжкова	книжкова	книжкова
3	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація	ілюстрація
4	література	література	література	література	література	література	література	література	література	література	література	література	література	література	література
5	казка	казка	казка	казка	казка	казка	казка	казка	казка	казка	казка	казка	казка	казка	казка
6	дитина	дитина	дитина	дитина	дитина	дитина	дитина	дитина	дитина	дитина	дитина	дитина	дитина	дитина	дитина
7	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво	мистецтво
8	художник	художник	художник	художник	художник	художник	художник	художник	художник	художник	художник	художник	художник	художник	художник
9	століття	століття	століття	століття	століття	століття	століття	століття	століття	століття	століття	століття	століття	століття	століття
10	картинка	картинка	картинка	картинка	картинка	картинка	картинка	картинка	картинка	картинка	картинка	картинка	картинка	картинка	картинка
11	письменник	письменник	письменник	письменник	письменник	письменник	письменник	письменник	письменник	письменник	письменник	письменник	письменник	письменник	письменник
12	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор	ілюстратор
13	текст	текст	текст	текст	текст	текст	текст	текст	текст	текст	текст	текст	текст	текст	текст
14	малюнок	малюнок	малюнок	малюнок	малюнок	малюнок	малюнок	малюнок	малюнок	малюнок	малюнок	малюнок	малюнок	малюнок	малюнок
15	читач	читач	читач	читач	читач	читач	читач	читач	читач	читач	читач	читач	читач	читач	читач
16	історія	історія	історія	історія	історія	історія	історія	історія	історія	історія	історія	історія	історія	історія	історія
17	життя	життя	життя	життя	життя	життя	життя	життя	життя	життя	життя	життя	життя	життя	життя
18	людина	людина	людина	людина	людина	людина	людина	людина	людина	людина	людина	людина	людина	людина	людина
19	вірш	вірш	вірш	вірш	вірш	вірш	вірш	вірш	вірш	вірш	вірш	вірш	вірш	вірш	вірш
20	видання	видання	видання	видання	видання	видання	видання	видання	видання	видання	видання	видання	видання	видання	видання

Рис. 3.6. Копіювання вмісту стовпчика В до стовпчика А
Далі зберігаємо файл у форматі *.xls (*.xlsx).

Аналогічним чином аналізуємо інші текстові файли, наповнюючи для кожного розділу (пунктів розділу) таблиці класів термінів та таблиці контекстів. Після аналізу кожного окремого розділу (пунктів розділів) об'єднуємо отримані таблиці класів термінів в єдину (зведену) таблицю класів (рис. 3.10), а таблиці контекстів в єдину (зведену) таблицю контекстів (комірка 1А містить системну інформацію nodedata для всієї таблиці).

44	сцинтилятор	сцинтилятор	композитний сцинтилятор	оксидний сцинтилятор
45	1. ОТРИМАННЯ ТА ДОСЛІДЖЕННЯ ДОПОВАНИХ КРИСТАЛІВ ZnSe 3 ПОЛІПШЕНИМИ СЦИНТИЛЯЦІЙНИМИ ПАРАМЕТРАМИ	1. ОТРИМАННЯ ТА ДОСЛІДЖЕННЯ ДОПОВАНИХ КРИСТАЛІВ ZnSe 3 ПОЛІПШЕНИМИ СЦИНТИЛЯЦІЙНИМИ ПАРАМЕТРАМИ	кристал	параметр
46	кристал	кристал	вироснування кристалів	зразки кристалів
126	2.1. Отримання дослідних зразків сцинтиляторів на основі ZnSe та ZnS	2.1. Отримання дослідних зразків сцинтиляторів на основі ZnSe та ZnS	кристал	зразок
127	кристал	кристал	об'ємний кристал	кристали твердих розчинів ZnSe-ZnS
128				
210	3.5. Катодолумінесценція кристалів вольфрамату кадмію	3.5. Катодолумінесценція кристалів вольфрамату кадмію	кристал	спектр
211	кристал	кристал	імпульсна катодолумінесценція кристалів	спайність кристалів
212				температура кристала

Рис. 3.7. Розкриття класу термінів для терміну «кристал» в різних пунктах тексту

Зверніть увагу, що більшість термінів повторюються в кожному розділі (пунктах розділів) (рис. 3.7), тому немає необхідності в кожній окремій таблиці класів термінів розкривати *зміст класу*, якщо він вже був описаний в попередній таблиці.

Для представлення змісту тексту у зведеній таблиці класів термінів описуються розділи (пункти розділів) за схемою, описаною вище.

(клас термінів A-Z), (термін A), (термін B), (термін C), ..., (термін Z)

(термін A), (термін A₁), (термін A₂), (термін A₃), ..., (термін A_n)

(термін B), (термін B₁), (термін B₂), (термін B₃), ..., (термін B_k)

.....

(термін Z), (термін Z₁), (термін Z₂), (термін Z₃), ..., (термін Z_k).

4. Редактор онтологій Graph Editor

4.1. Загальні відомості про Graph Editor

Graph Editor – це багатофункціональний редактор і переглядач онтологій. Через серверну частину Graph Editor виконується зв'язок з іншими додатками у складі ТОДОС.

4.2. Основні функції

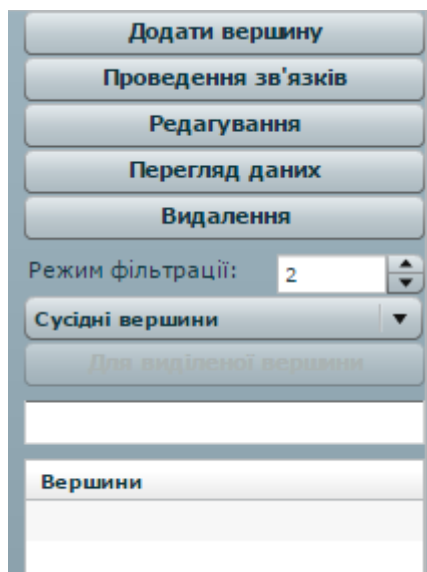


Рис. 4.1. Кнопки виклику основних функцій Graph Editor

Основні функції здійснюються з допомогою елементів керування в правому блоці (рис. 4.1), контекстних меню робочої області або просто з допомогою мишки. Вони надають основний функціонал по редагуванню і перегляду онтографів.

Додати вершину

Вершина додається кліком у відповідному режимі. Потрібно ввімкнути режим створення вершин з допомогою кнопки в правому блоці або через контекстне меню робочої області. Буде відкрита форма редагування вершини, в якій можна встановити значення для нової вершини

Додати зв'язок

Зв'язок додається двома послідовними кліками – на дочірню вершину, а потім на батьківську у відповідному режимі. Потрібно ввімкнути режим створення зв'язків з допомогою кнопки в правому блоці або через контекстне меню робочої області. Буде відкрита форма редагування зв'язка, в якій можна встановити

значення для нового зв'язка (див. Форма редагування зв'язка). Якщо між вибраними вершинами вже існує зв'язок, то він перетвориться в множинний зв'язок (див. Форма множинного зв'язка).

Редагування

Редагування запускається кліком на вершині або зв'язку у відповідному режимі. Потрібно ввімкнути режим редагування з допомогою кнопки в правому блоці або через контекстне меню робочої області. Відкриється відповідна форма (див. Форма редагування вершини, Форма редагування зв'язка). Якщо клік відбувся по множинному зв'язку, відкриється форма множинного зв'язка в варіанті редагування (див. Форма множинного зв'язка). Для редагування конкретного зв'язка необхідно двічі клікнути на відповідну строку в її таблиці

Перегляд даних

Перегляд даних запускається кліком на вершині або зв'язку у відповідному режимі. Потрібно ввімкнути режим перегляду даних з допомогою кнопки в правому блоці або через контекстне меню робочої області. Відкриється відповідна форма (див. Форма перегляду даних). Якщо клік відбувся по множинному зв'язку, відкриється форма множинного зв'язка в варіанті перегляду (див. Форма множинного зв'язка). Для перегляду даних конкретного зв'язка необхідно двічі клікнути на відповідну строку в її таблиці

Видалення

Видалення відбувається кліком на вершині або зв'язку у відповідному режимі. Потрібно ввімкнути режим видалення з допомогою кнопки в правому блоці або через контекстне меню робочої області. Буде відображене повідомлення з запитом підтвердження (рис. 4.2). Якщо видаляється вершина, повідомлення відображатиме факт що разом з вершиною видаляються всі інцидентні зв'язані з нею зв'язки, якщо видаляється множинний зв'язок – що будуть видалені всі дочірні зв'язки. Запит підтвердження можна відключити (див. *Запит на видалення*).

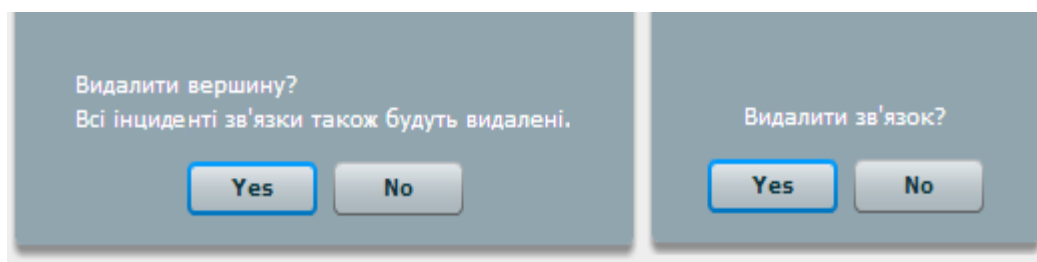


Рис. 4.2. Запит на підтвердження видалення
Топологічна фільтрація вершин

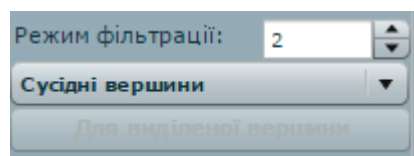


Рис. 4.3. Вибір режиму фільтрації

Топологічна фільтрація відбувається блоком з трьох елементів: селектора глибини, перемикача типу і кнопки «Для виділеної вершини».

Доступні чотири типи фільтрації (рис. 3.3): сусідні вершини, шлях між вершинами, субмножина (всі нащадки вершини, тобто дочірні вершини, дочірні вершини для дочірніх вершин і т.д.) і супермножина (аналогічно для батьківських вершин). Фільтрація запускається вибором вершини в списку вершин (див. Список вершин), або вибором вершини в робочій області і натиском кнопки «Для виділеної вершини».

Список вершин

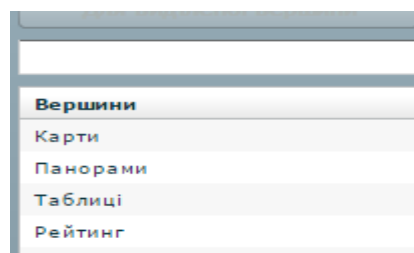


Рис. 4.4. Вибір списку вершин

Список вершин розміщений в нижній частині правого блоку. Він містить перелік всіх імен вершин, присутніх в онтографі, і має поле для пошуку вершин по імені. При введенні послідовності символів в поле фільтрації зі списку приховуються всі вершини крім тих, що містять в назві цю підпослідовність (у випадку якщо ввімкнене налаштування «Точний пошук в списку вершин» – ті, назва яких починається з неї). Клік на вершині в цьому списку виділить цю

вершину певним кольором, центрує робочу область на ній і запустить топологічну фільтрацію, використавши цю вершину в якості стартової точки.

Виділення вершин

Виділення вершин відбувається кліком на вершині у випадку, якщо відключені спеціальні режими (режим створення вершин, режим проведення зв'язків та ін.). Якщо в цей момент кнопка Shift не натиснута, то виділення з інших виділених вершин буде зняте.

Контекстне меню для вершини

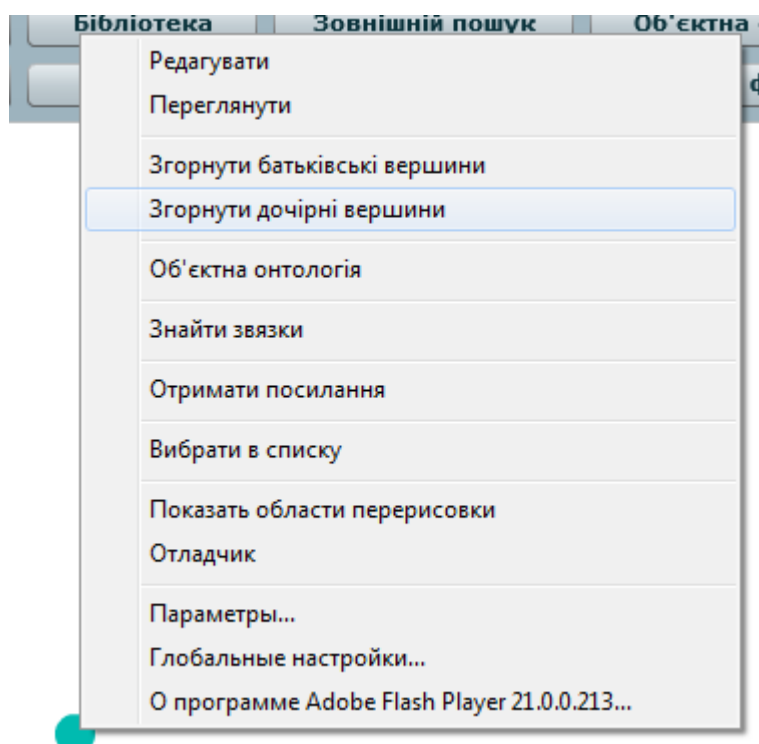


Рис. 4.5. Контекстне меню вершини

Контекстне меню вершини має набір важливих функцій:

Пункти меню «Редагувати» і «Переглянути» викликають форми редагування і перегляду даних, аналогічно до кліку у відповідних режимах.

Пункти «Згорнути батьківські вершини» і «Згорнути дочірні вершини» приховують всі вершини, що є батьківськими або дочірніми для поточної. Всі зв'язані зі згорнутими вершинами зв'язки будуть відображатись як зв'язані з поточною.

Пункт меню «Об'єктна онтологія» відкриє об'єктну онтологію в режимі перегляду даних про поточну вершину (див. Об'єктна онтологія).

Пункт меню «Отримати посилання» видає пряме посилання на виділену вершину, що може в подальшому бути використане для розміщення на сторонніх ресурсах.

Пункт «Вибрати в списку» виділяє вершину в списку вершин в правому блоці, і виконує всі дії, які відбуваються при такому виборі (див. Список вершин).

Зчитування даних

Зчитування даних відбувається з допомогою відповідної кнопки на формі спеціальних функцій (див. Форма спеціальних функцій, Дані – Зчитати дані). Існує три формати зчитування даних (зчитування імен зв'язків, зчитування даних в форматі `nodedata`, Зчитування даних в форматі `nodeproperties`).

Зчитування імен зв'язків

Імена зв'язків можна зчитувати з файлів Confor після побудови онтографу,. Файли Confor являють собою таблиці, кожен рядок яких має наступний формат:

Перша комірка строки містить назву батьківської вершини. Друга комірка містить назву зв'язка (можна залишити пустою). Третя і всі наступні комірки містять назви дочірніх вершин.

Зчитування даних в форматі `nodedata`

В даному форматі в лівій верхній комірці таблиці знаходиться ключове слово «`nodedata`», а в кожній з наступних строк є п'ять комірок, в яких послідовно знаходяться: назва батьківської вершини, текстовий контент елемента даних, посилання елемента даних (тільки для типу даних гіперпосилання), тип даних (`text` для текстових елементів і `href` для гіперпосилань), і клас елемента даних.

Зчитування даних в форматі `nodeproperties`

Даний формат відповідає звичайній електронній таблиці: в лівій верхній комірці знаходиться ключове слово `nodeproperties`, справа від нього – класи елементів даних, знизу – назви вершин, а у всіх інших комірках текстовий контент елементів даних.

4.3. Верхнє функціональне меню

Зчитати файл

Кнопка призначення для зчитування файлів формату *.xml, *.owl, *.viz, *.csv, *.xls, *.xlsx, *.kml. Вона відкриває стандартне для браузера вікно вибору файлу, в якому і потрібно вибрати файл для зчитування.

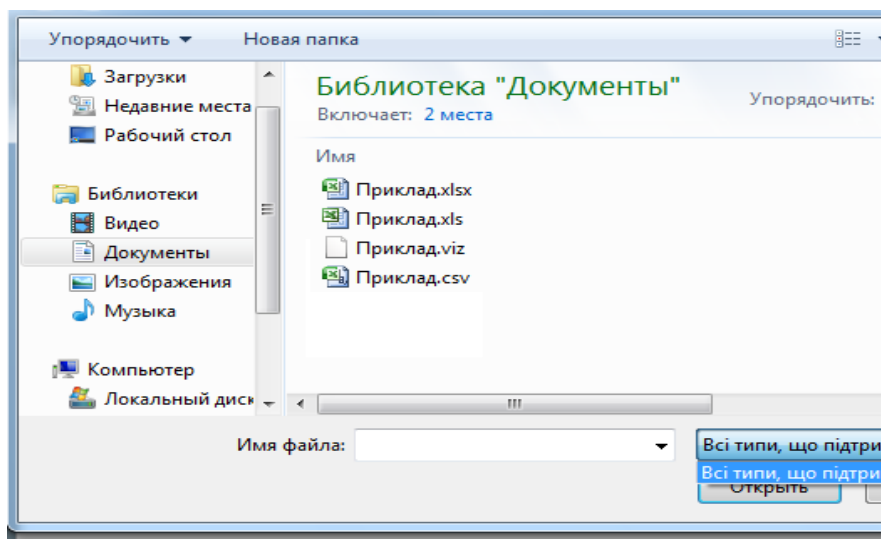


Рис. 4.6. Вікно вибору файлу для зчитування

Файли повинні мати кодування UTF-8. Можливе зчитування файлів з кодуванням windows-1251 (див. Використовувати win1251).

Якщо зчитується файл електронних таблиць (*.csv, *.xls, *.xlsx), то вважається що він представляє онтологію (див. Формат вхідних файлів Confor).

Файли *.owl та *.viz є спеціалізованими і тому використовуються лише в особливих випадках (див. Формат вихідних файлів Confor).

Зберегти онтограф

Кнопка призначена для збереження поточного онтографу у форматі *.xml. Вона також викликає стандартну форму браузера.

Збереження здійснюється в кодуванні UTF-8 або windows-1251 (див. Використовувати win1251).

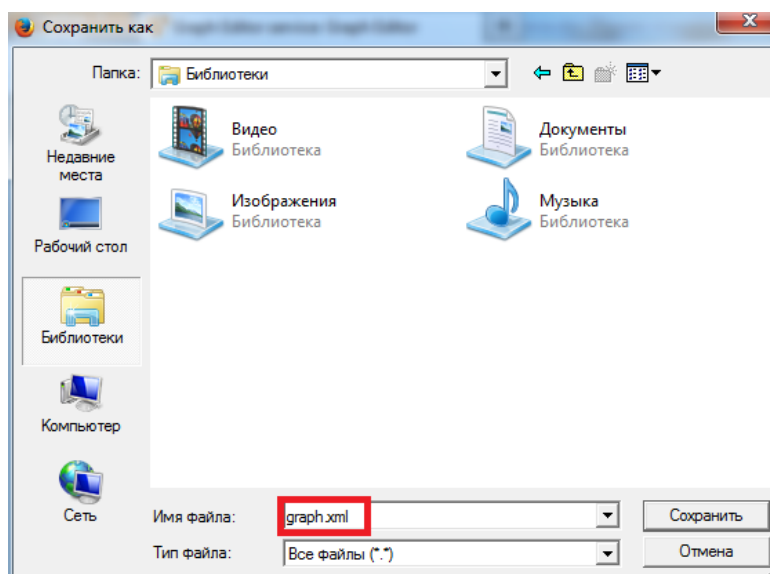


Рис. 4.7. Вікно збереження поточного онтографу,

Розмістити вершини

Перерозміщує вершини у відповідності з налаштуваннями розміщення (див. Налаштування розміщення)

Показати весь онтограф

Знімає всі фільтри, повертає всі вершини на місця, в яких вони знаходились до фільтрації (див. Форма фільтрації зв'язків, Фільтрування вершин, Перерозміщувати вершини після фільтрації).

Фільтр зв'язків

Показує форму фільтрації зв'язків (див. Форма фільтрації зв'язків).

Типи зв'язків

Показує форму, що дозволяє створювати нові типи зв'язків і керувати існуючими (див. Форма керування зв'язками).

Бібліотека

Відкриває бібліотеку онтологій (див. Форма бібліотеки).

Додатки

Відкриває бібліотеку додатків – інформаційних і аналітичних сервісів (див. Форма бібліотеки).

Кнопки «Збільшити» і «Зменшити»

Керують масштабом головного вікна, в якому відображається онтологія.

Кнопка «На весь екран»

Переводить редактор в повноекранний режим (див. Повноекранний режим).

Друк

Виводить поточний онтограф на друк. Поточний онтограф розбивається на прямокутники, кожен з яких друкується на новій сторінці. Для коректної роботи функції потрібно у вікні, що відкриється (його вигляд залежить від браузера) ВРУЧНУ вибрати орієнтацію сторінки «Альбомна».

Налаштування

Відкриває форму налаштувань (див. Форма налаштувань).

4.4. Основні форми

Форма редагування вершини

Дана форма використовується для створення або редагування вершини (рис. 4.8).

Ім'я вершини

Ім'я повинно бути унікальним для даного онтографу,. У випадку, якщо вершина з вибраним іменем вже існує, буде показане повідомлення про помилку. Якщо залишити це поле пустим при створенні, то автоматично згенерується ім'я в форматі «node_X».

Рис. 4.8. Форма редагування вершини

Клас вершини

Метадані вершини. Потрібні для спеціальних випадків, таких, як збереження метаданих. В загальному випадку клас вершини представляється іншою вершиною, вище в ієрархії.

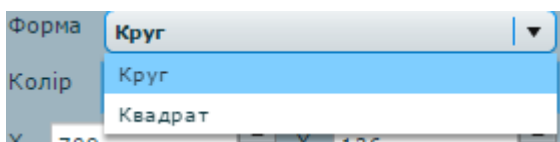
Форма вершини

Рис. 4.9. Вибір вигляду вершини

Круг або квадрат (рис. 4.9). Визначає, як відображатиметься вершина в робочій області.

Колір

Колір, яким відображатиметься вершина за замовчуванням (рис. 4.10). Виділені вершини відображаються іншим кольором, заданим в налаштуваннях.

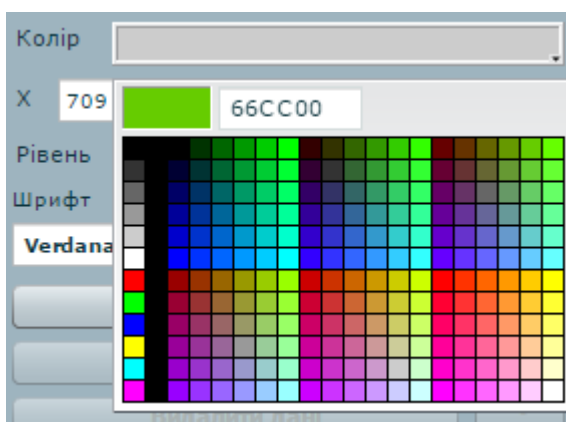


Рис. 4.10. Вибір кольору вершини

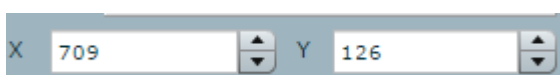
Координати X і Y

Рис. 4.11. Вибір позиції вершини

Позиція вершини в робочій області (рис. 4.11). В загальному випадку дане поле зручніше не редагувати, а замість цього перетягувати вершину за допомогою мишки.

Рівень

Метадані вершини, використовуються в спеціалізованих алгоритмах розміщення (див. Розміщення пірамідальної мережі). В загальному випадку це поле краще залишити без змін.

Шрифт



Рис. 4.12. Вибір шрифту напису біля вершини

Розмір і тип шрифту (рис. 4.12). Визначає, як буде відображатись назва вершини в робочій області.

Блок керування даними вершини

Дозволяє додавати, редагувати і видаляти елементи даних (див. *Форма редагування даних*).

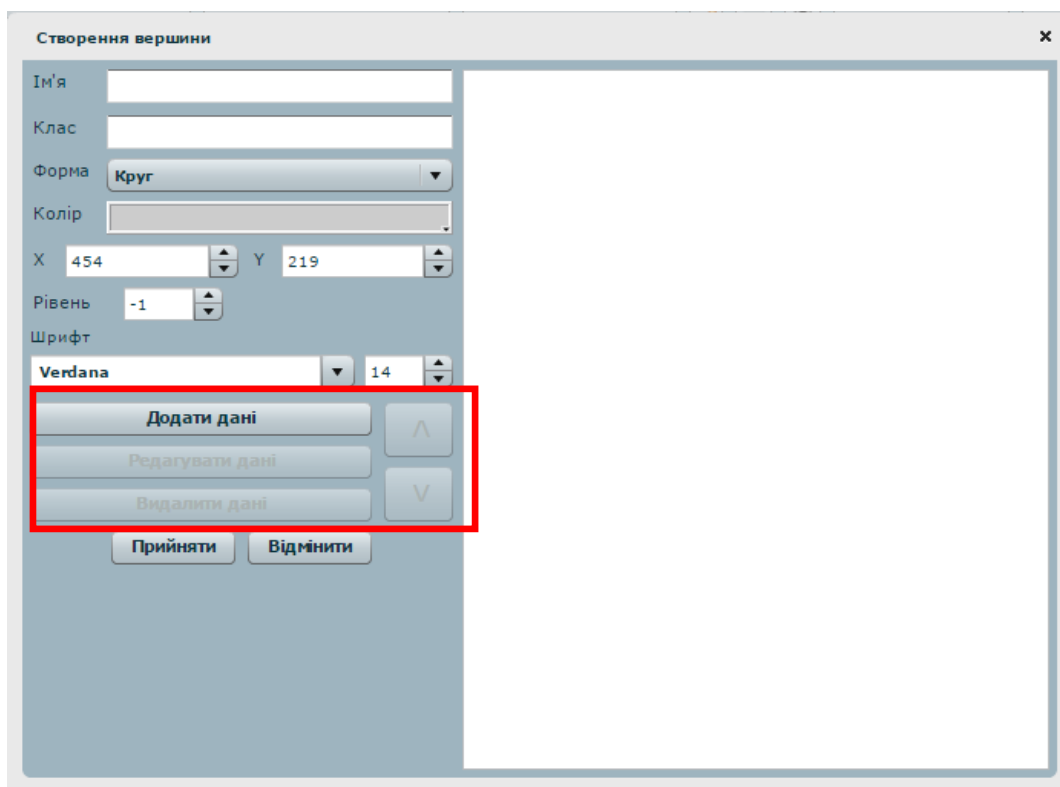


Рис. 4.13. Блок керування даними вершини

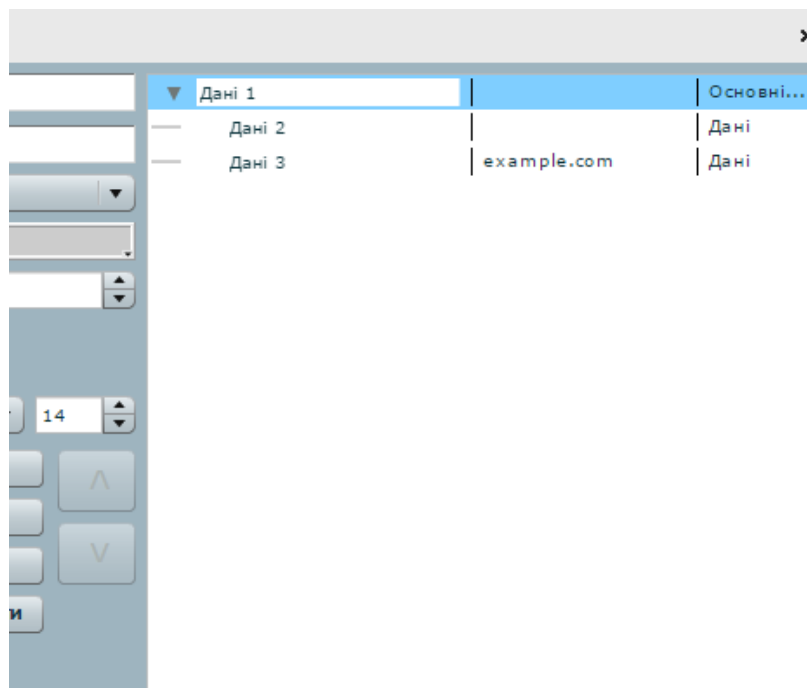


Рис. 4.14. Атрибути вершини

Список даних вершини. Показує атрибути поточної вершини (рис. 4.14). Список поділений на три колонки: колонка з текстом, колонка з посиланням, колонка з класом (див. *Форма редагування даних*). Дані можуть бути вкладеними, тоді зліва будуть відображатись керуючі елементи, з допомогою яких можна розкривати дані вищого рівня для того, щоб показувати дані нижчого.

Форма редагування зв'язку

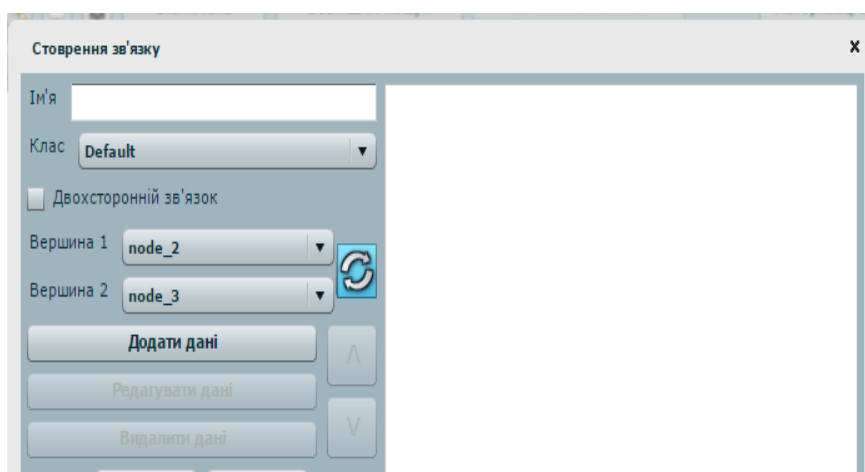


Рис. 4.15. Форма редагування зв'язків

Використовується для створення або редагування зв'язків (рис. 4.15).

Ім'я. Відображається тільки якщо ввімкнена опція «Показувати імена зв'язків». Може бути довільним. Якщо ім'я не вказане, воно згенерується автоматично за шаблоном «link_X».

Клас

Клас зв'язка задає його колір (див. Форма керування зв'язками).

Двосторонній зв'язок



Рис. 4.16. Двосторонній зв'язок між вершинами

Двосторонні зв'язки мають стрілки на обох кінцях. Однак, при розміщенні вони обробляються як односторонні. Якщо множинний зв'язок має дочірній двосторонній зв'язок, він теж вважається двостороннім.

Вершини, які з'єднує зв'язок

Рис. 4.17. Форма задання зв'язку між вершинами

Вершина 1 вважається дочірньою, а Вершина 2 – батьківською. Змінити вершини, між якими проведено зв'язок, не можна – можна тільки поміняти місцями дочірню і батьківську вершину.

Блок керування даними зв'язка

Аналогічний до «Блок керування даними вершини».

Список даних зв'язка

Аналогічний до «Список даних вершини».

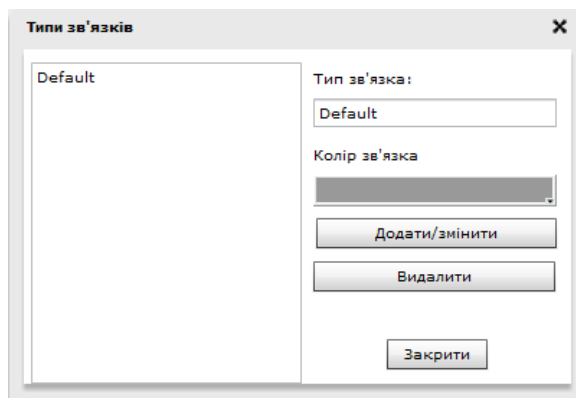


Рис. 4.18. Форма керування зв'язками

Форма керування зв'язками. Дана форма (рис. 4.18) відкривається кнопкою «Типи зв'язків». Форма дозволяє створювати, редагувати і видаляти типи зв'язків (див. Форма редагування зв'язка, Форма фільтрації зв'язків).

Видалення: потрібно вибрати існуючий зв'язок в таблиці і натиснути «Видалити». Для додавання/редагування необхідно ввести ім'я у відповідне поле, і вибрати колір, після чого натиснути відповідну кнопку. Якщо зв'язок з таким іменем вже існує в системі, то замість створення нового зв'язка буде змінено колір існуючого.

За замовчуванням Graph Editor має тип зв'язку з ім'ям Default, який відображається сірим кольором і присвоюється всім новоствореним зв'язкам (див. Проведення зв'язків).

Форма множинного зв'язку

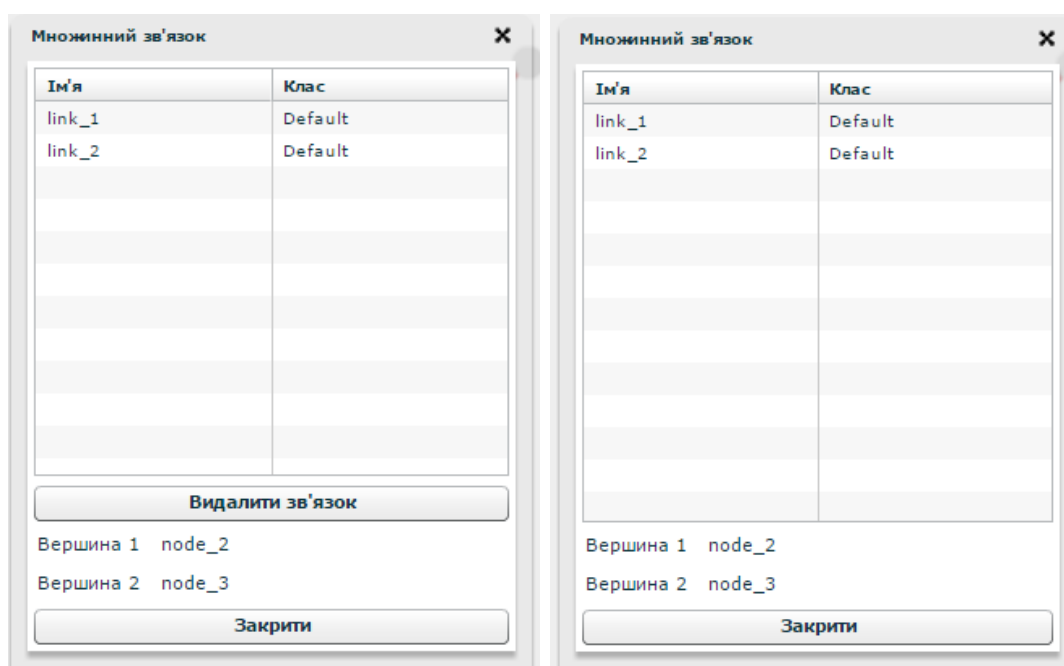


Рис. 4.19. Форма множинного зв'язку

Дана форма доступна в двох варіантах – для перегляду і редагування (рис. 4.19).

Основний елемент форми – таблиця зв'язків, в якій показані всі дочірні зв'язки поточного множинного зв'язку. Подвійний клік на елементі в таблиці в залежності від стану відкриє форму редагування зв'язку або форму перегляду даних.

В режимі редагування додатково доступна функція «Видалити зв'язок».

Форма редагування даних

Рис. 4.20. Форма редагування даних

Дана форма (рис. 4.20) викликається блоками редагування даних зв'язка і вершини, і дозволяє змінювати або створювати нові елементи даних.

Текст

Текстовий контент елемента даних. Відображається при перегляді даних вершини чи зв'язку.

Ієрархія

Рис. 4.21. Форма вибору батьківського елемента

Місце вершини в ієрархії. Даний елемент дозволяє вибрати батьківський елемент даних для поточного (рис. 4.21).

Тип даних

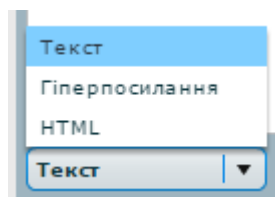


Рис. 4.22. Вибір типу даних

Тип даних, що визначає спосіб відображення елемента даних.

Текст – звичайний текстовий елемент.

Гіперпосилання – текстовий елемент зі зв'язаним гіперпосиланням.

HTML – текст з довільним форматуванням.

Посилання

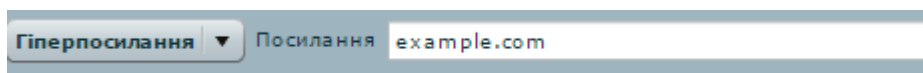


Рис. 4.23. Додавання гіперпосилання

Посилання доступне тільки якщо вибраний тип даних «Гіперпосилання». Дозволяє прикріпити до елемента даних посилання на інший електронний ресурс – картинку, документ, іншу вершину тощо (рис. 4.23).

Клас елемента даних

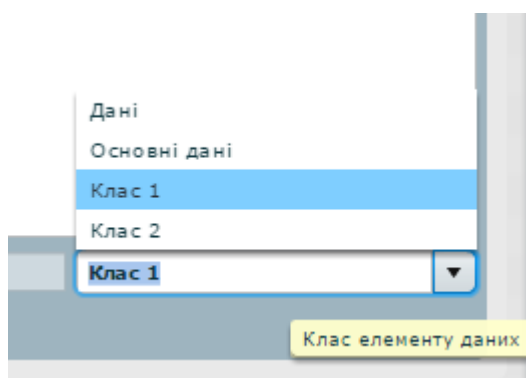


Рис. 4.24. Встановлення класу елемента даних

Клас елемента даних дозволяє фільтрувати елементи даних, а також присвоювати їм спеціалізовані ролі (рис. 4.24). Наприклад, класи формату `arcgis_XXX` визначають посилання на карту, а клас `embedded_graph` визначає

вкладений онтограф. Різні інтерфейси можуть мати різні спеціалізовані класи даних.

Елемент вибору класу являє собою випадаючий список з усіма наявними в онтографі класами даних. З випадаючим списком об'єднано поле для введення нового класу.

Форма перегляду даних

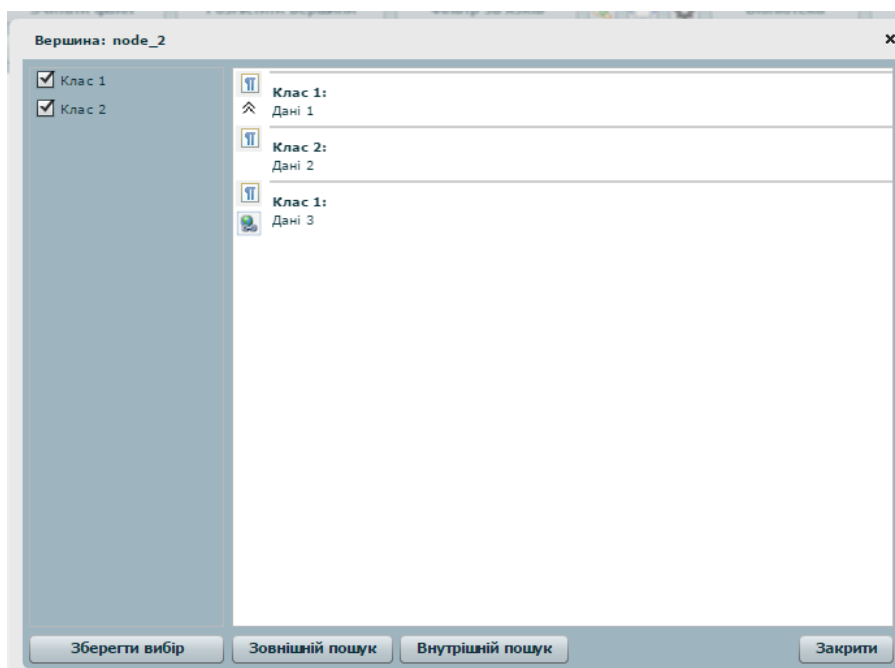


Рис. 4.25. Форма перегляду даних

Форма перегляду даних (рис. 4.25) відображає дані, аналогічно до форм редагування вершин і зв'язків, але в дещо іншому форматі. Кожен елемент даних на цій формі має кнопку для перегляду текстового контенту, яка відкриває вікно з текстовою областю, в якій відображено цей контент.

Також елементи типу «Гіперпосилання» мають кнопку, яка відкриває зв'язане гіперпосилання в новій вкладці браузера.

Якщо елементу даних присвоєно клас, він відображається над текстовим контентом жирним шрифтом.

Форма також містить фільтр даних в правій частині. В ньому можна зняти виділення з деяких класів, і тоді елементи даних з цим класом зникнуть в основній області. Якщо натиснути кнопку «Зберегти вибір» внизу фільтра, то поточний

вибір збережеться для всіх подальших відкриттів даної форми (за замовчуванням вибрані всі доступні класи).

4.5. Допоміжні форми

Форма внутрішнього пошуку

Граф	Вершина	Деталі	Таксономія	Об. онтологія
Спеціалізована школа №304 Свє	Дослідження термоелектричних явищ	Деталі	Таксономія	Об. онтологія
Спеціалізована школа №304 Свє	Дослідження теплової дії електрично	Деталі	Таксономія	Об. онтологія
Спеціалізована школа №304 Свє	Визначення фторидів	Деталі	Таксономія	Об. онтологія
Комунальний заклад «Рішельєвсь	Виготовлення та дослідження фотоел	Деталі	Таксономія	Об. онтологія
Комунальний заклад «Рішельєвсь	Виготовлення та дослідження фотоел	Деталі	Таксономія	Об. онтологія
Комунальний заклад «Рішельєвсь	Дослідження теплової дії електрично	Деталі	Таксономія	Об. онтологія
Комунальний заклад «Рішельєвсь	Вивчення теплових властивостей ме	Деталі	Таксономія	Об. онтологія
Комунальний заклад «Рішельєвсь	Визначення феруму (III)	Деталі	Таксономія	Об. онтологія
Комунальний заклад «Рішельєвсь	Визначення жорсткості води	Деталі	Таксономія	Об. онтологія
Комунальний заклад «Рішельєвсь	Визначення аніаку та іонів амонію	Деталі	Таксономія	Об. онтологія
Комунальний заклад «Рішельєвсь	Визначення сульфатів	Деталі	Таксономія	Об. онтологія
Програма NEED з енергетики	1-2 класи. Крок другий: джерела ен	Деталі	Таксономія	Об. онтологія
Програма стандарту 10 клас хімії	Застосування алюмінію, заліза та їхн	Деталі	Таксономія	Об. онтологія
Програма стандарту 10 клас хімії	Амфотерні властивості алюміній окси	Деталі	Таксономія	Об. онтологія
Програма стандарту 10 клас хімії	Поняття про твердість води (постійн	Деталі	Таксономія	Об. онтологія

Рис. 4.26. Форма внутрішнього пошуку

Дана форма (рис. 4.26) використовується для повнотекстового пошуку по атрибутах всіх вершин всіх онтографів, що знаходяться в системній бібліотеці.

Шукати контексти

Даний перемикач визначає, шукати введену пошукову фразу серед всіх атрибутів вершин, чи тільки серед їх імен.

Навігація форми внутрішнього пошуку

Навігація представлена двома кнопками внизу форми – вперед (>>) і назад (<<).

Пошукові результати внутрішнього пошуку

Граф	Вершина	Деталі	Таксономія	Об. онтологія
Спеціалізована школа №304 Свє	Дослідження термоелектричних явищ	Деталі	Таксономія	Об. онтологія
Спеціалізована школа №304 Свє	Дослідження теплової дії електрично	Деталі	Таксономія	Об. онтологія
Спеціалізована школа №304 Свє	Визначення фторидів	Деталі	Таксономія	Об. онтологія

Рис. 4.27. Результати внутрішнього пошуку

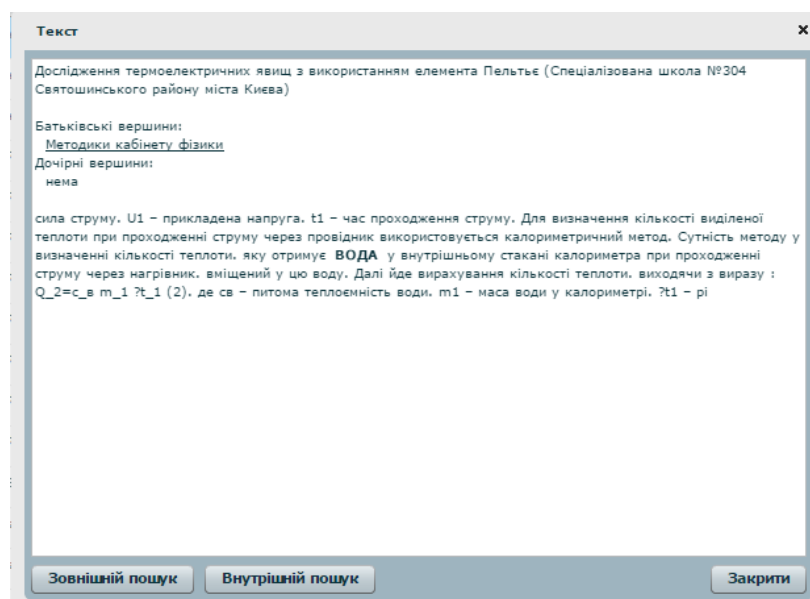


Рис. 4.28. Форма «Деталі»

Пошукові результати відображаються в таблиці з трьома колонками: «Граф», «Вершина» і колонкою з керуючими елементами (рис. 4.27).

Керуючі елементи представлені трьома кнопками: «Таксономія» і «Об’єктна онтологія» відкривають в новому вікні вибраний онтограф, тоді як «Деталі» показує додаткову форму (рис. 4.28). Ця форма містить наступні дані: назву вершини і онтографа, в якому вона знаходиться, її батьківські і дочірні вершини а також контексти, в яких було знайдено пошукову фразу або її фрагменти.

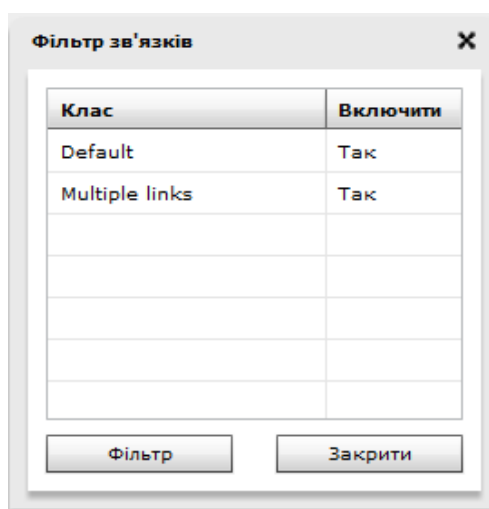


Рис. 4.29. Форма фільтрації зв'язків

Форма фільтрації зв'язків. Відкривається кнопкою «Фільтр зв'язків». Дозволяє обрати типи зв'язків, що будуть відображатися під час перегляду онтографу,. Також впливають на розміщення вершин (див. Розміщення вершин).

Подвійне натиснення лівої кнопки миші на назві типу зв'язку в стовпчику «Клас», на слові «Так» в стовпчику «Включити» або власне на кнопці «Включити» перемикає індикатор праворуч від назви типу зв'язку на положення «Ні», який означає, що ребро онтографу із обраним зв'язком не відображатиметься. Після вибору типу зв'язку, який відображатиметься під час перегляду онтографу,, необхідно натиснути кнопку «Фільтр». На екрані відобразяться лише ті ребра онтографу,, чий тип зв'язків в таблиці фільтрації відмічений словом Так» (рис. 4.30).

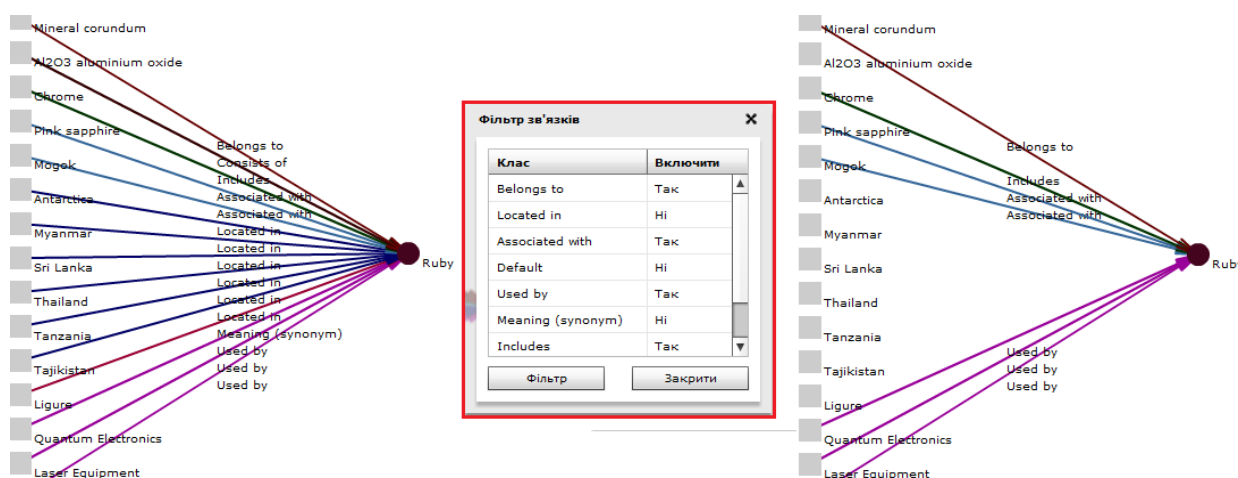


Рис. 4.30. Відображення фільтрації ребер



Рис. 4.31. Форма бібліотеки

Форма бібліотеки. Форма бібліотеки (рис. 4.31) відкривається кнопками «Бібліотека» і «Додатки». При цьому відкриваються бібліотека онтологій і додатків відповідно. Бібліотека онтологій – це сховище вже існуючих онтологій, впорядкованих по папках за тематикою, галуззю застосування, іменем автора тощо,

призначених для перегляду. Бібліотека додатків містить різноманітні аналітичні та інформаційні сервіси, які, як правило, розташовуються на сторонніх серверах.

В бібліотеці передбачений пошук (рис. 4.32), що здійснюється шляхом введення ключового слова в пошукову строку в верхній частині вікна бібліотеки.

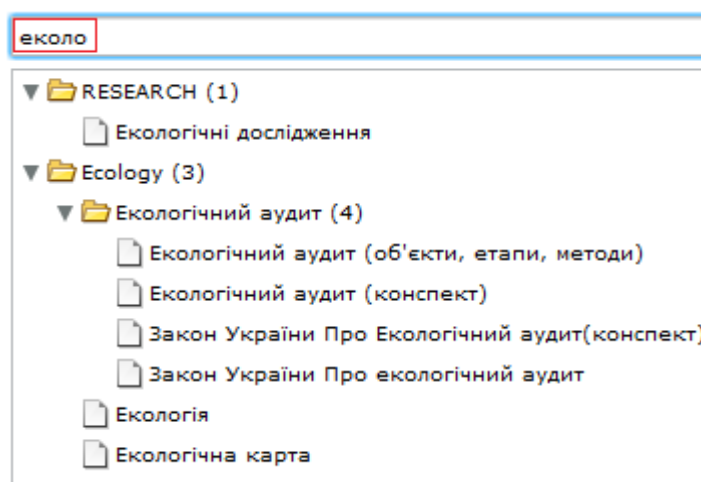


Рис. 4.32. Пошук у бібліотеці

Спеціальні функції:

Форма спеціальних функцій дозволяє виконувати допоміжні, вузькоспеціалізовані або просто небезпечні функції, що не призначені для використання пересічними користувачами.

Експорт – Protégé (файл owl). Дозволяє експортувати поточний онтограф у формат Ontology Web Language.

Експорт – Confor (файл viz). Дозволяє експортувати поточний онтограф у внутрішній формат системи Confor.

Експорт – Виділені вершини (файл xml). Дозволяє зберегти в файл не всі вершини, а тільки вибрану підмножину. Зв'язки між даними вершинами, їх дані і налаштування також зберігаються.

Дані – Зчитати дані. Виконує Зчитування даних.

Дані – Очистити дані. Видаляє всі дані з усіх вершин. Перед видаленням видається запит підтвердження.

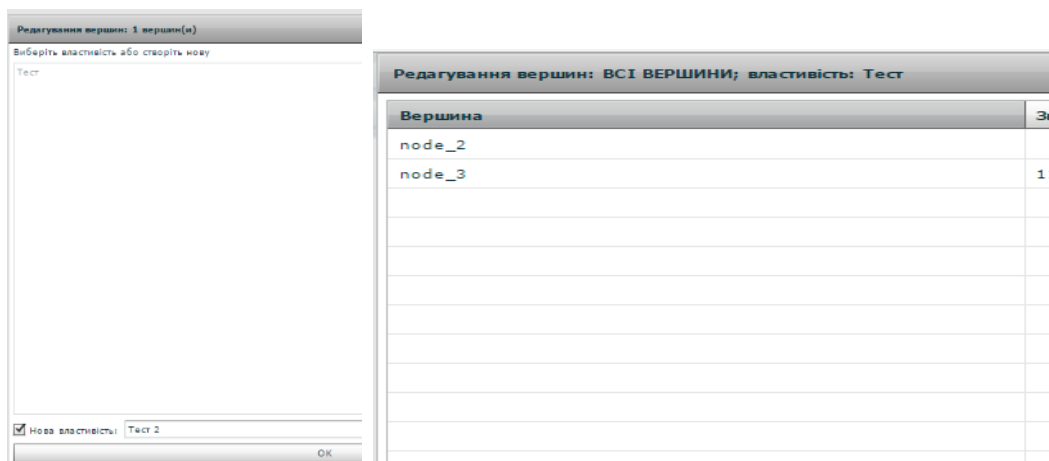


Рис. 4.33. Функція редагування властивостей даних визначеного типу

Спеціальні дії – Редагування властивостей. Дана функція (рис. 4.33) дозволяє редагувати дані вибраного типу у всіх вершинах.

Перед початком редагування можна вибрати одну або кілька вершин, тоді редагуватимуть тільки дані з цих вершин. Інакше редагуватимуться дані з усіх вершин.

Необхідно вибрати один з існуючих або створити новий клас даних. Буде показано список всіх елементів даних, що належать даному класу. Подвійним кліком на правій половині списку можна починати редагування даних.

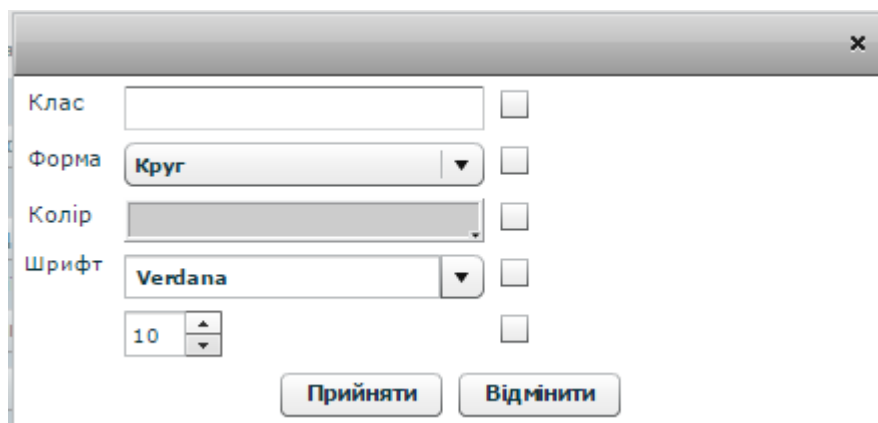


Рис. 4.34. Функція редагування відображення вершин

Спеціальні дії – Редагування відображення. Перед відкриттям форми необхідно вибрати одну або кілька вершин в онтографі. На формі є набір властивостей, що керують відображенням вершини (рис. 4.34). Після вибору властивостей (з допомогою чек боксів в центрі) і самих значень цих властивостей і

натискання «Прийняти» вибрані значення будуть присвоєні всім вибраним вершинам.

4.6. Форма налаштувань

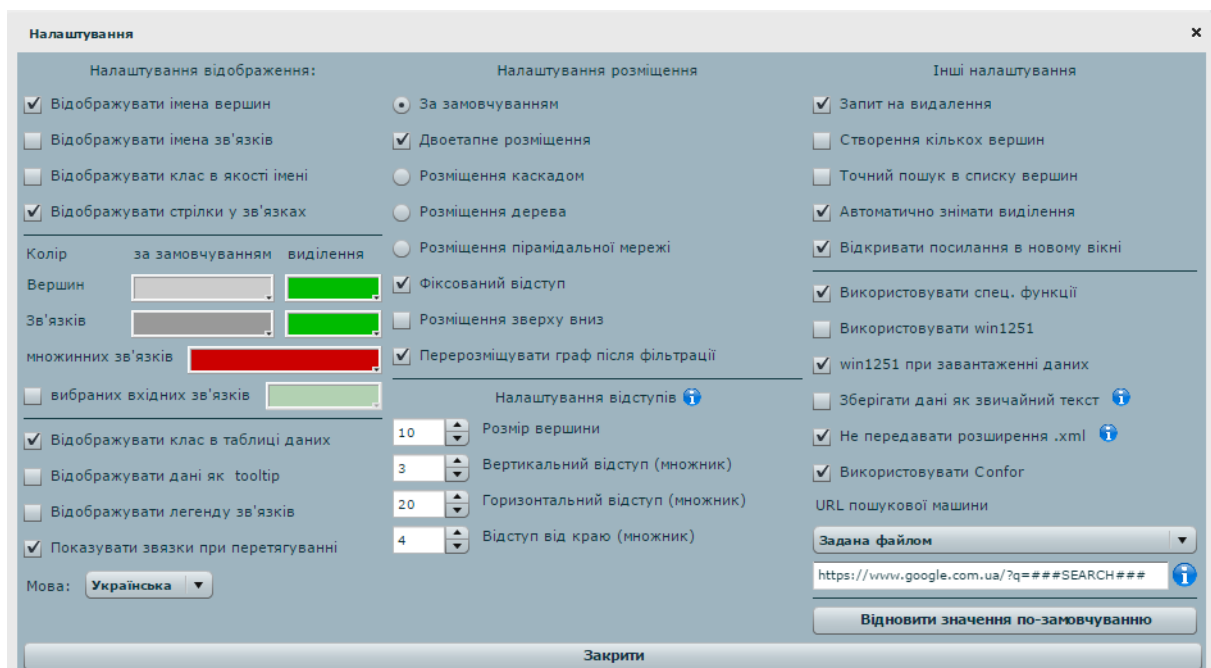


Рис. 4.35. Форма налаштувань

Дана форма (рис. 4.34) складається з чотирьох блоків: налаштування відображення, налаштування розміщення, налаштування відступів та інші налаштування.

Налаштування відображення:

Даний блок призначений для керування тим, як виглядають вершини і зв'язки, а також керує роботою деяких функцій.

Показувати імена вершин. Режим (рис. 4.36) відображає або відключає відображення імені вершини онтографу (див. Форма редагування онтографу).

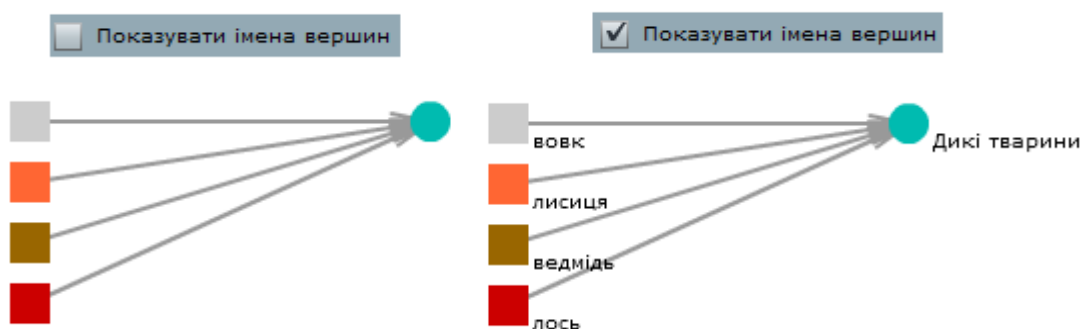


Рис. 4.36. Відображення імені вершини онтографу,

Показувати імена зв'язків. Режим (Рис. 4.37) вмикає або вимикає відображення імені зв'язку між вершинами онтографу.

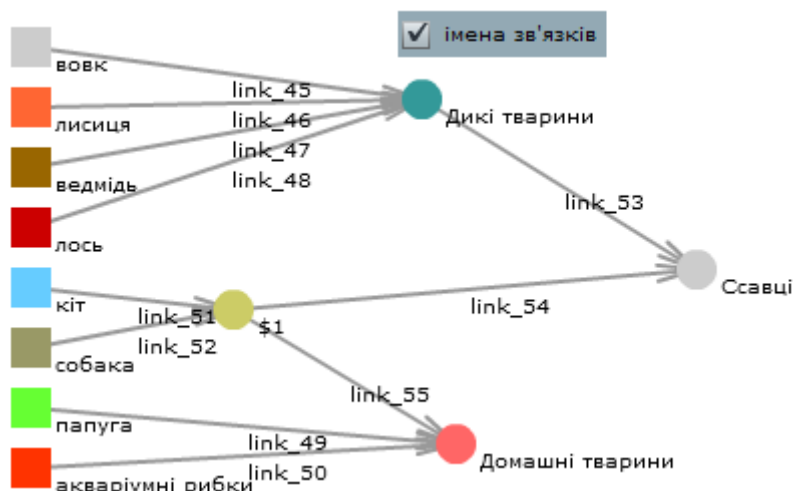


Рис. 4.37. Відображення імені зв'язку між вершинами онтографу,
Відображувати клас в якості імені. При ввімкненні даного режиму (рис. 4.38) разом з режимом «Показувати імена зв'язків», замість імен зв'язків будуть показані їх класи (див. Форма редагування зв'язка).

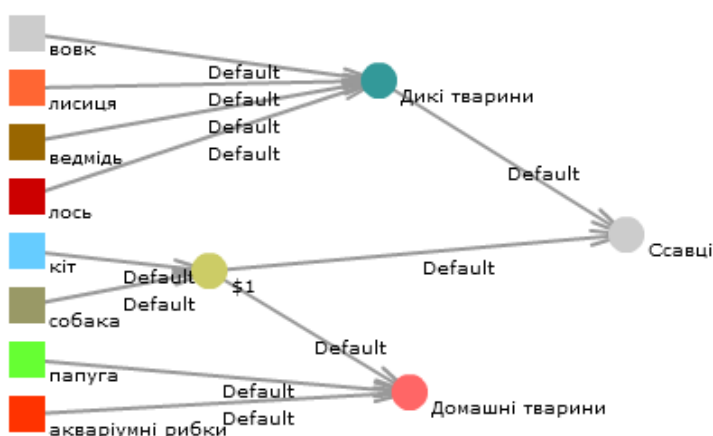


Рис. 4.38. Відображення назви класів замість імен зв'язків
Показувати стрілки у зв'язках. Режим (рис. 4.39) показує або приховує стрілки на кінцях зв'язків (див. Зв'язок).



Рис. 4.39. Відображення стрілок на кінцях зв'язків

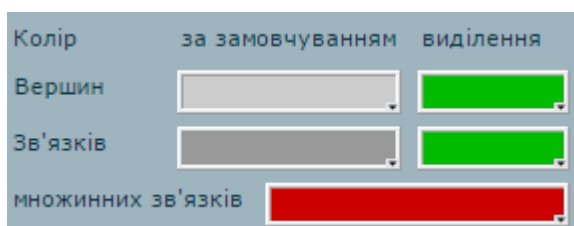


Рис. 4.40. Блок керуючих елементів «Колір»

Колір. Блок керуючих елементів «Колір» (рис. 4.40) дозволяє керувати кольорами вершин і зв'язків в різних ситуаціях. Доступні наступні елементи:

Колір за замовчуванням вершин – призначається вершині автоматично всім новоствореним вершинам.

Колір за замовчуванням зв'язків використовується, якщо зв'язку призначена некоректна група (наприклад, якщо відповідну йому групу видалили).

Колір виділення вершин/зв'язків використовується для виділених кліком миші зв'язків (див. Виділення).

Колір множинних зв'язків задає колір відображення зв'язка, якщо він складається з кількох підзв'язків (див. Зв'язок).

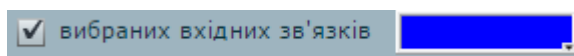


Рис. 4.41. Задання кольору вхідних зв'язків

Колір вибраних вхідних зв'язків. Даний блок (рис. 4.41) з перемикача і селектора кольорів дозволяє відображати зв'язки, що входять у вибрану вершину, іншим кольором, відмінним від кольору вихідних зв'язків (рис. 4.42). Це може бути корисним для складних онтографів, де зв'язків занадто багато. Якщо вибрано кілька вершин, то зв'язки між виділеними вершинами відображаються звичайним кольором.

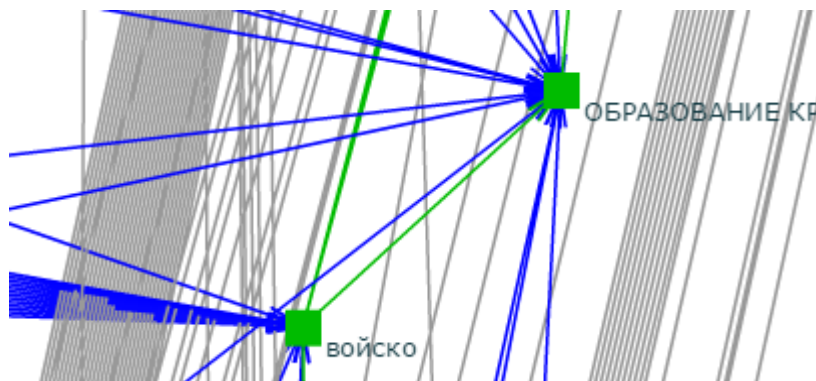


Рис. 4.42. Відображення вибору декількох вершин

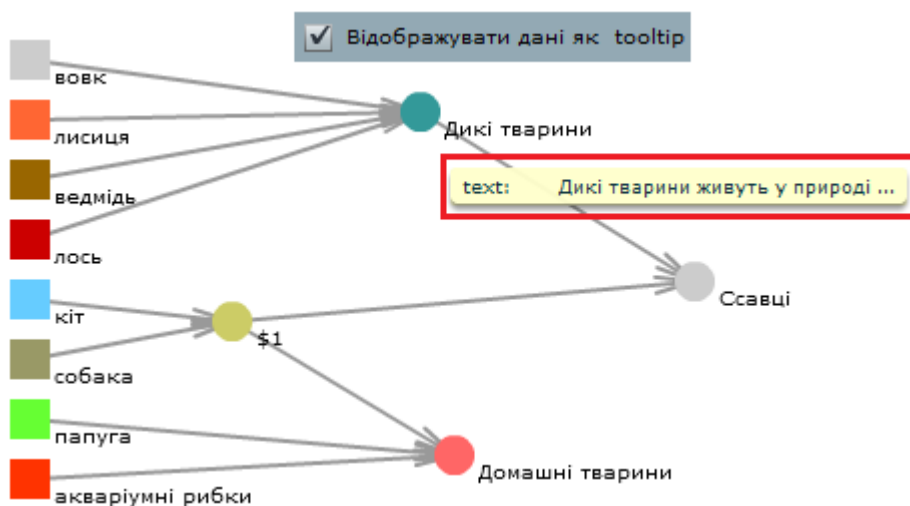


Рис. 4.43. Відображення підказки зі списком атрибутів вершини

Відображувати дані як tooltip. Якщо дана опція ввімкнена, то при наведенні вказівника миші на вершину буде показана підказка зі списком атрибутів вершини (рис. 4.43).

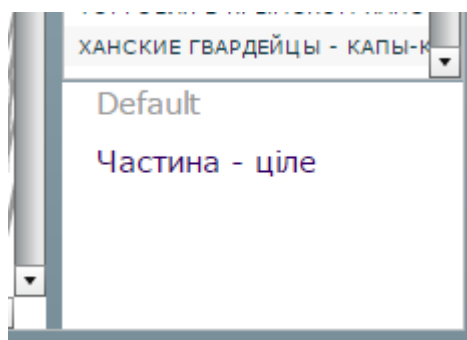


Рис. 4.44. Відображення списку всіх типів зв'язків з їх кольорами

Показувати легенду зв'язків. Якщо ввімкнути дану опцію, то справа, під списком вершин буде показано список всіх типів зв'язків з їх кольорами (рис. 4.44).

Показувати зв'язки при перетягуванні. Якщо вимкнути дану опцію, то при перетягуванні вершини інцидентні їй зв'язки не будуть автоматично перемальовуватись, і будуть приховані. Корисно для великих онтографів, якщо є відчутні затримки оновлення робочої області при перетягуванні.

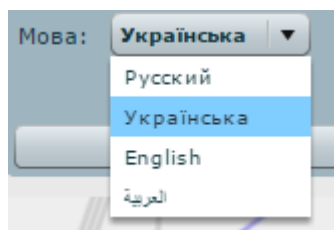


Рис. 4.45. Вибір мови інтерфейсу

Мова. Даний перемикач (рис. 4.45) дозволяє вибрати мову інтерфейсу модуля. Доступні українська, англійська і арабська.

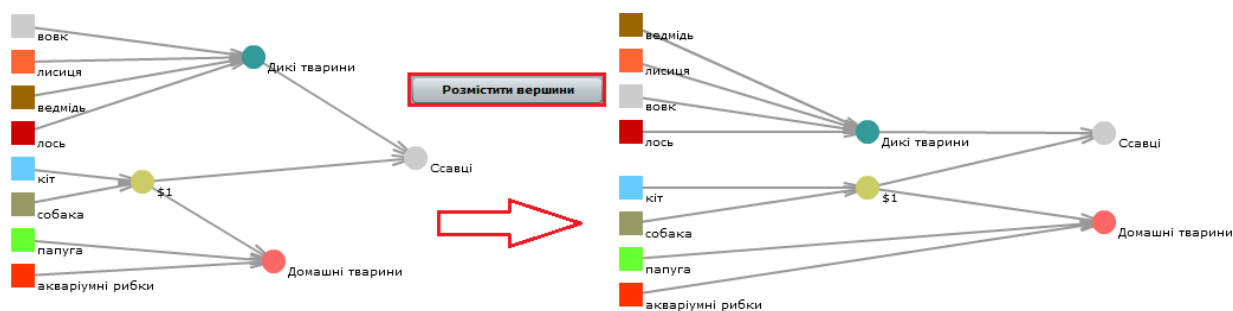


Рис. 4.46. Розташування онтографу за замовчуванням
Налаштування розміщення:

Розташування за замовчуванням. Даний режим (рис. 4.46) розміщення користується простим критерієм – сумарна довжина зв'язків між вершинами має бути мінімальною.

Двоетапне розміщення. При ввімкненні даної опції в режимі розміщення за замовчуванням розміщення вершин буде відбуватись двічі. В загальному випадку це дає кращі результати, але існують конфігурації, для яких в результаті двох етапного розміщення будуть гірші, ніж одноетапного.

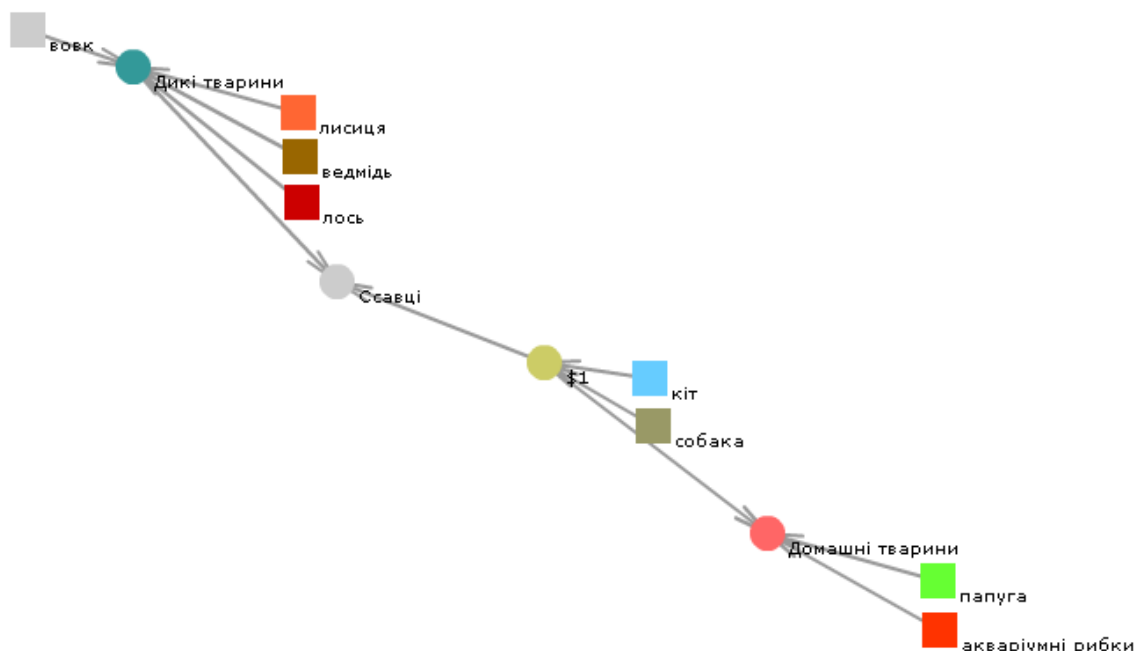


Рис. 4.47. Розташування онтографу каскадом

Розміщення каскадом. Даний режим (рис. 4.47) розміщення призначений для циклічних онтографів. Він розміщує вершини довільному порядку, з фіксованими

відступами по вертикалі і горизонталі. Як правило, результати роботи даного алгоритму достатньо погані.

Розміщення дерева. Даний алгоритм розміщує деревовидні структури. Він виділяє піддерева і розміщує їх елементи таким чином щоб вони формували дерево візуально. Застосування до не-деревовидних структур може мати непередбачуваний результат.

Розміщення пірамідальної мережі. Алгоритм розміщення розміщує пірамідальну мережу за наступним правилом: відстань від лівого краю у вершини залежить від кількості листових вершин, що є нащадками даної вершини.

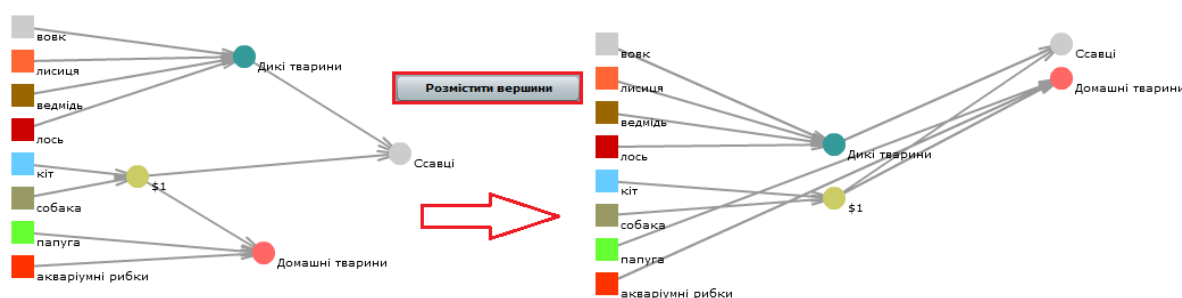


Рис. 4.48. Розташування батьківські вершини зверху, з фіксованим відступом *Фіксований відступ*. При ввімкненні даної опції всі батьківські вершини будуть розміщені зверху, з фіксованим відступом (рис. 4.48). Якщо опція вимкнена, кореневі вершини будуть розміщені за тим же алгоритмом, що і всі інші.



Рис. 4.49. Зміна напрямку розміщення онтографу, *Розміщення зверху вниз*. Міняє напрямок розміщення онтографу,. При ввімкненій опції кореневі вершини розміщуються справа, а листові – зліва (рис. 4.49). При вимкненій опції кореневі вершини розміщуються зверху, листові – знизу.

Налаштування відступів. В даному блоці є наступні селектори:

Розмір вершини – розмір вершини у пікселях. На її основі (в результаті множення на цю величину) визначаються всі відступи.

Вертикальний відступ (множник) – величина, з допомогою якої вираховується мінімальна вертикальна відстань між вершинами.

Горизонтальний відступ (множник) – величина, з допомогою якої вираховується відступ між рівнями (див. Розміщення вершин).

Відступ від краю (множник) – величина, з допомогою якої вираховується відступ від верхнього і лівого краю.

Інші налаштування:

Запит на видалення. Перемикач дозволяє вимкнути показ діалогового вікна «Видалити вершину/зв'язок?», що відкривається при спробі видалити зв'язок або вершину. Дане налаштування може бути корисне при видаленні великої кількості вершин (див. Видалення).

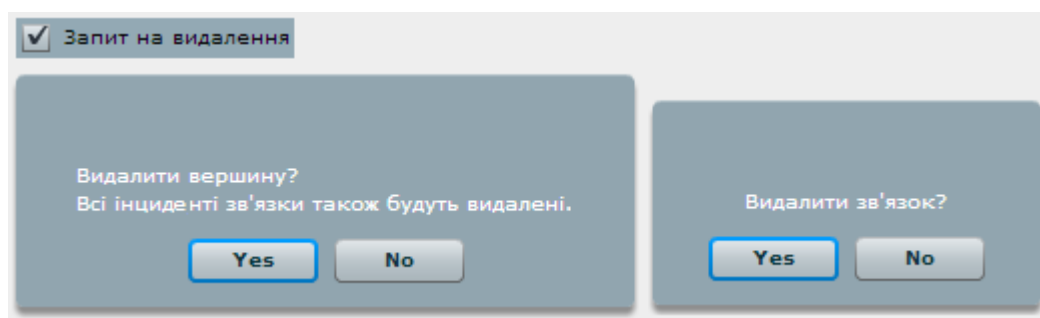


Рис. 4.50. Перемикач відображення запиту на видалення

Створення кількох вершин. Якщо ввімкнути дану опцію, то режим «Додати вершину» залишиться ввімкненим після створення вершини. Корисно час при створенні великої кількості вершин (див. Додати вершину).

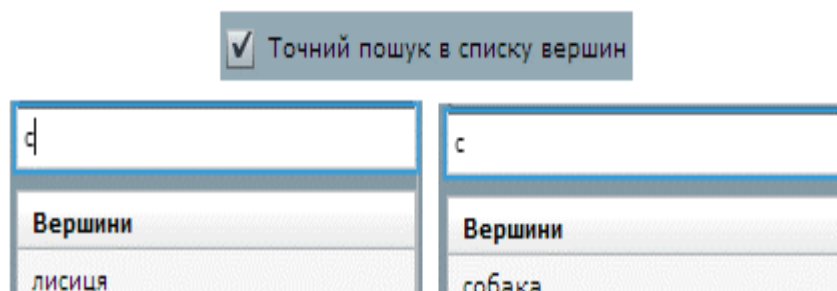


Рис. 4.51. Опція точного пошуку вершин

Точний пошук в списку вершин. При ввімкненні даної опції (рис. 4.51) пошук в списку вершин вибиратиме тільки ті вершини, що починаються з введеної послідовності символів. Інакше пошук шукатиме входження в будь-якому місці назви (див. Список вершин).

Автоматично знімати виділення. Якщо це налаштування ввімкнене, то виділення вершини автоматично знімається при кліку на робочу область чи іншу вершину. Інакше – знімати виділення потрібно вручну, затиснувши клавішу Shift і клікнувши.

Відкривати посилання в новому вікні. Якщо ця опція ввімкнена, то всі посилання (переходи з бібліотеки, на зовнішній пошук та ін.) Направлятимуть користувача в нову вкладку браузера. Інакше використовується поточна вкладка.

Використовувати спец. Функції. Показує на верхній панелі кнопку «Спец. функції» (див. *Спеціальні функції*).

Використовувати win1251. За замовчуванням онтографи зчитуються і записуються в кодуванні UTF-8. Даний режим дозволяє відкривати і зберігати онтографи в кодуванні windows-1251.

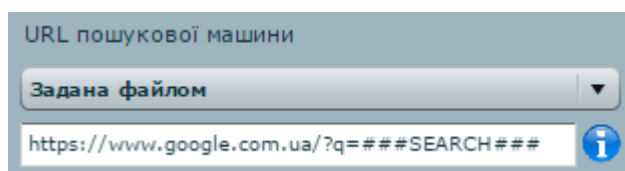


Рис. 4.52. Задання адреси пошукової машини

URL пошукової машини. Даний блок складається з двох полів (рис. 4.52):

Вибір пошукової машини з кількох наперед визначених варіантів, таких, як Google або Yandex.

У випадку, якщо в попередньому селекторі вибрано опцію «Задана файлом», то активується поле вводу специфічної пошукової машини. Значення даного поля зберігається в XML-файлі.

Посилання, що водиться в друге поле, мусить мати маркер `###SEARCH###`, який і буде замінений пошуковим запитом.