

# СТВОРЕННЯ АНАЛІТИЧНОЇ ВЕБ-СИСТЕМИ НАКОПИЧЕННЯ ТА АНАЛІЗУ ВІДКРИТИХ ДАНИХ

ВИКОНАВ: КОХАНСЬКИЙ МАКСИМ

НАУКОВИЙ КЕРІВНИК: БОЦУЛА МИРОСЛАВ  
ПАВЛОВИЧ

# ОБ'ЄКТ – ПРЕДМЕТ – МЕТА ДОСЛІДЖЕННЯ

- **Об'єктом дослідження** є процес розробки аналітичної веб-системи накопичення та аналізу відкритих даних.
- **Предметом дослідження** є аналітична веб-система накопичення та аналізу відкритих даних.
- **Метою дослідження** є побудова та опис реалізації веб-системи.

# ІДЕЯ РОЗРОБКИ

- Для уникнення проблем з авторськими правами зберігати тільки гіперпосилання на веб-ресурс, який безпосередньо містить певний набір даних;
- Для забезпечення зручності пошуку, усі дані формалізуються за трьома категоріями:
  - Час;
  - Простір;
  - Ключові слова.
- Кожна з категорій формалізується окремо;

# ІДЕЯ РОЗРОБКИ

- Формалізацію просторових об'єктів та ключових слів пропонується здійснювати у вигляді окремих семантичних мереж;
- Семантичні мережі пропонується будувати за принципом mesh-мереж, коли кожен вузол з'єднується з багатьма сусідніми і вже, через них, — з іншими вузлами.
- Інші характеристики: назва, власник даних тощо.

# КОМПЛЕКСНА МОДЕЛЬ ІНФОРМАЦІЙНО-ПОШУКОВОЇ АНАЛІТИЧНОЇ ВЕБ-СИСТЕМИ

«Час»:

$$T_i = [T_{i11}, T_{i12}] \cup [T_{i21}, T_{i22}] \cup [T_{i31}, T_{i32}] \cup \dots \cup [T_{iw1}, T_{iw2}]$$

«Простір»:

$$O_p = \langle X_p, R, F_p \rangle$$

«Ключові слова»:

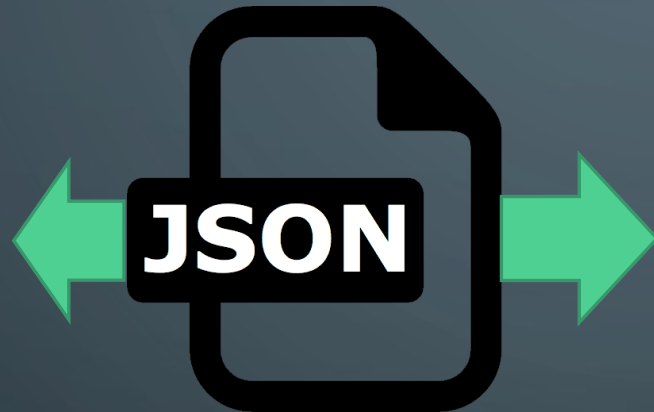
$$O_k = \langle X_k, R, F_k \rangle.$$

# АЛГОРИТМ ПОБУДОВИ ОНТОЛОГІЧНОЇ БАЗИ ЗНАНЬ ДЛЯ КОЖНОГО $i$ -ГО НАБОРУ

1. Визначається прив'язка в часі за моделлю і зберігається у базі даних, наприклад  $T_i = [15.04.2016, 15.10.2016]$ .
2. Визначається прив'язка у просторі у вигляді назв просторових об'єктів у кількості  $N_{p_i}$ , наприклад ( $N_{p_i} = 3$ ):  $x_{p_i1} = \text{«Вінниця»}$ ,  $x_{p_i2} = \text{«Транспортна мережа»}$ ,  $x_{p_i3} = \text{«басейн р. Південний Буг»}$ .
3. Для кожного концепту прив'язки у просторі визначаються концепти із загальної множини  $X_p$  та визначаються відношення між ними з множини  $R$ .
4. Задається інтерпретація із множини  $F_p$ .
5. Визначаються ключові слова для опису предметної області даного набору у кількості  $N_{k_i}$ , наприклад ( $N_{k_i} = 3$ ):  $x_{k_i1} = \text{«Парникові гази»}$ ,  $x_{k_i2} = \text{«Атмосферне повітря»}$ ,  $x_{k_i3} = \text{«Довкілля»}$ .
6. Для кожного концепту прив'язки у просторі визначаються концепти із загальної множини  $X_k$  та визначаються відношення між ними із множини  $R$ .
7. Задається інтерпретація із множини  $F_k$ .

# ВИБІР ТЕХНОЛОГІЙ

- HTML
- CSS
- AngularJS
- jQuery
- Bootstrap



- PHP
- SLIM
- RESTfull API
- MySQL



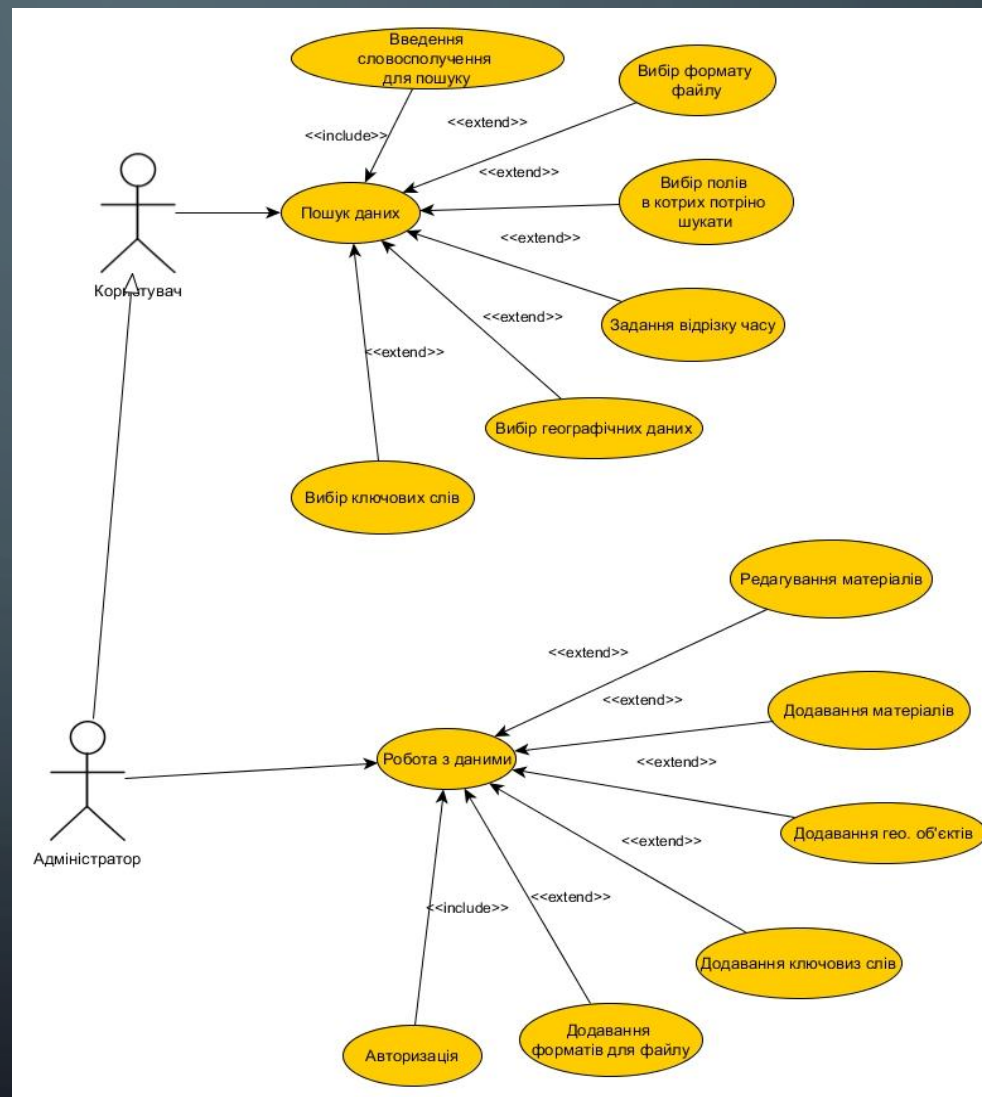
# Діаграма використання

## Можливості користувача:

- Введення запиту для пошуку;
- Вибір потрібного формату;
- Вибір відрізка часу;
- Вибір географічного розміщення;
- Вибір ключових слів.

## Можливості адміністратора:

- Авторизація;
- Робота з форматами файлів;
- Робота з ключовими словами;
- Робота з географічними об'єктами;
- Робота з наборами даних.





# РОЗРОБКА БАЗИ ДАНИХ

5 – головних таблиць

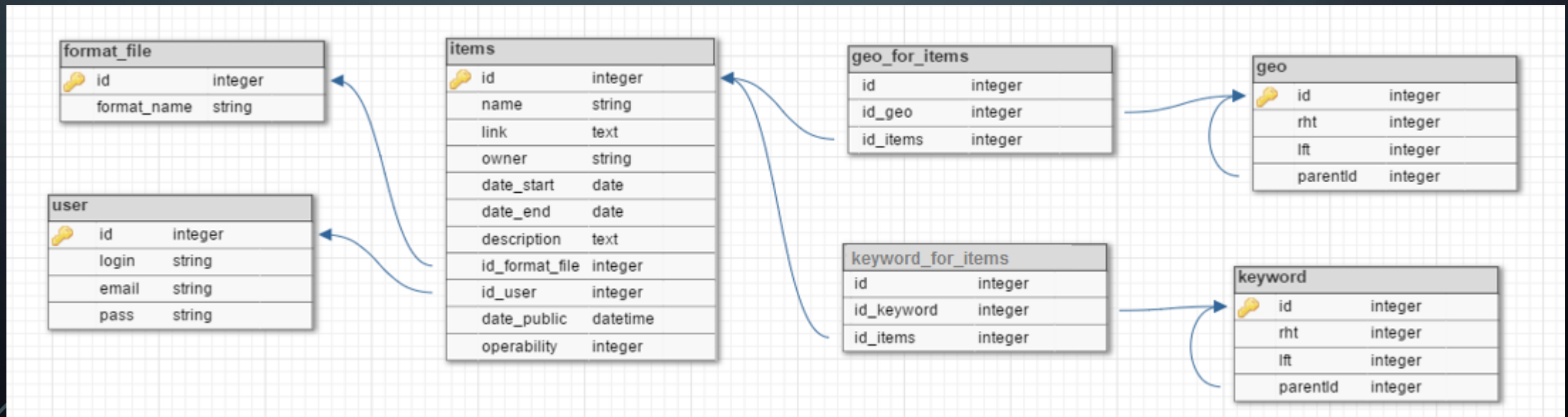
2 – додаткові таблиці

Основні таблиці:

- Items – набори даних;
- Geo – географічні об'єкти (NestedSet);
- Keyword – ключові слова (NestedSet);
- Format\_file – формати файлів;
- User – користувачі.

Додаткові таблиці

- Geo\_for\_items
- Keyword\_for\_items



# КОРИСТУВАЦЬКИЙ ІНТЕРФЕЙС

Пошук по обраному формату

csv, JSON


Відсутній

csv ✓

doc

xls

JSON ✓



Наприклад: Викиди забруднюючих речовин в атмосферне повітря

Пошук по обраному формату: csv, doc, JSON


Параметри дати: 09.12.2017 - 09.12.2017

Виберіть потрібний регіон

- Європа
- Україна
- Житомирська область
- Вінницька область
- Київ
- Черкаська область
- Дніпропетровська область

Виберіть ключові слова

- забруднюючі речовини
- діоксид сірки
- моніторинг довкілля
- викиди
- викиди забруднюючих речовин
- забруднювачі довкілля
- природно-заповідний фонд



Наприклад: Викиди забруднюючих речовин

Пошук по обраному формату: csv, doc, JSON

Параметри дати: 09.12.2017 - 09.12.2017

Виберіть потрібний регіон

- Європа
- Північна Америка
- Південна Америка
- Антарктида

Виберіть ключові слова

- забруднюючі речовини
- моніторинг довкілля
- викиди
- викиди забруднюючих речовин
- забруднювачі довкілля
- природно-заповідний фонд
- ...

Будь-коли

16.01.2018 21.02.2018 Обрати Очистити

Січень 2018      Лютий 2018

Пн	Вт	Ср	Чт	Пт	Сб	Нд	Пн	Вт	Ср	Чт	Пт	Сб	Нд
31	1	2	3	4	5	6	28	29	30	31	1	2	3
7	8	9	10	11	12	13	4	5	6	7	8	9	10
14	15	16	17	18	19	20	11	12	13	14	15	16	17
21	22	23	24	25	26	27	18	19	20	21	22	23	24
28	29	30	31	1	2	3	25	26	27	28	1	2	3
4	5	6	7	8	9	10	4	5	6	7	8	9	10

# АДМІНІСТРАТИВНА ЧАСТИНА САЙТУ

Додавання матеріалу

Редагування матеріалу

Додавання географічного об'єкта

Переміщення географічних об'єктів

Редагування географічних об'єктів

Додавання ключового слова

Переміщення ключового слова

Редагування ключового слова

Додавання формату файлу

### Редагування матеріалу

Назва даних\*:

Веб-посилання\*:

Власник даних:

Формат файлу:

Період часу, за який ці дані\*:

Просторові об'єкти\*

- Європа
- Україна
- Північна Америка
- Південна Америка
- Антарктида

Ключові слова\*

- забруднюючі речовини
- моніторинг довкілля
- викиди
- викиди забруднюючих речовин
- забруднювачі довкілля
- природно-заповідний фонд

Опис:

Список матеріалів, для редагування

Редагування та додавання матеріалу

Додавання матеріалу

Редагування матеріалу

Додавання географічного об'єкта

Переміщення географічних об'єктів

Редагування географічних об'єктів

Додавання ключового слова

Переміщення ключового слова

Редагування ключового слова

Додавання формату файлу

### Редагування матеріалу

Пошук

id	Назва	Статус посилання	Дата створення	Редагування
127	Обсяги заготівель ліквідної деревини, тис. куб м	1	27.06.2017	<a href="#">Edit</a> <a href="#">Delete</a>
126	Основні показники використання й охорони лісу	1	27.06.2017	<a href="#">Edit</a> <a href="#">Delete</a>
125	Обсяги вилову риби та добування інших водних біоресурсів станом на 01.04.16	1	27.06.2017	<a href="#">Edit</a> <a href="#">Delete</a>
124	Обсяги вилову риби та добування інших водних біоресурсів станом 01.07.2016	1	27.06.2017	<a href="#">Edit</a> <a href="#">Delete</a>
123	Обсяги вилову риби та добування інших водних біоресурсів станом на 01.10.16	1	27.06.2017	<a href="#">Edit</a> <a href="#">Delete</a>
122	Обсяги вилову риби та добування інших водних біоресурсів за січень -березень 2017 року	1	27.06.2017	<a href="#">Edit</a> <a href="#">Delete</a>
121	Новозбудовані об'єкти, що введені в експлуатацію ФО та ФОП в Одеській області за вересень 2016 р.	1	26.06.2017	<a href="#">Edit</a> <a href="#">Delete</a>
120	Новозбудовані об'єкти, що введені в експлуатацію ФО та ФОП в Одеській області за серпень 2016 р.	1	26.06.2017	<a href="#">Edit</a> <a href="#">Delete</a>
119	Зведена інформація про об'єкти, що будуються фізичними особами в Одеській області	1	26.06.2017	<a href="#">Edit</a> <a href="#">Delete</a>
118	Об'єкти, що здані в експлуатацію юридичними особами в Одеській області	1	26.06.2017	<a href="#">Edit</a> <a href="#">Delete</a>

Admin Вийти

# ЗАПУСК ВЕБ-СИСТЕМИ

Аналітичну веб-систему накопичення та аналізу відкритих даних було розгорнуто на сервері і запуснено на сайті кафедри Системного аналізу, комп'ютерного моніторингу та інженерної графіки під доменом [isodata.mmss.vntu.edu.ua](http://isodata.mmss.vntu.edu.ua)

# ЕКОНОМІЧНА ЧАСТИНА

**На основі зроблених підрахунків в економічній частині магістерської кваліфікаційної роботи досягнуті наступні результати:**

- витрати на розробку та її впровадження складають 42 000 грн.;
- абсолютний ефект розробки складає 492,75 тис. грн. за три роки або 1 478 000 грн. щорічно;
- за здійсненими підрахунками було виявлено, що внутрішня норма дохідності була досягнута і складає 88%;
- термін окупності системи, що розробляється складає 1,13 років.

# ВИСНОВКИ

- Дана розробка запущена за адресом [isodata.mmss.vntu.edu.ua](http://isodata.mmss.vntu.edu.ua);
- Розглянуто проблему формалізації та ідентифікації інформаційних моделей відкритих числових даних будь-якого призначення для забезпечення їх швидкого комплексного пошуку одночасно за прив'язкою у часі, у просторі та по ключових словах;
- Запропоновано нову комплексну модель інформаційно-пошукової аналітичної веб-системи накопичення та аналізу відкритих даних з прив'язкою одночасно у часі, у просторі та по ключових словах на основі онтологічних моделей і семантичних мереж;

# ВИСНОВКИ

- Продемонстровано працездатність розробленої моделі та підходів до формалізації відкритих даних;
- Охарактеризовано її інтерфейс та особливості функціонування;
- Доведено, що продукт є економічно доцільним;
- Опубліковано 2 тези доповідей на наукових конференціях і одна науково-технічна стаття.

ДЯКУЮ ЗА УВАГУ