

## **Реферат**

магістерської атестаційної роботи на тему:

«Розробка й реалізація метаданих для структуризації сховища даних»

Воробйової Олени Іванівни

### **Актуальність роботи**

Популярність сховищ даних в останні роки суттєво зросла. Конкурентоспроможні організації перебувають на шляху побудови сховищ даних, або розширення, перепроєктування й удосконалення вже існуючих. Метадані вважаються ключовим фактором успіху в проектах по впровадженню сховищ. Вони містять усю інформацію, необхідну для вилучення, перетворення й завантаження даних з вихідних систем, а також для наступного використання й інтерпретації складового сховищ даних.

### **Мета роботи**

Метою роботи було створення програмного продукту для структуризації сховища експериментальних даних і забезпечення застосування стандартних методів сучасної технології Data Mining з використанням середовища розробки Delphi 7. Розглянути основні формати відображення метаданих, такі як HDF, Netcdf, FITS і стандартні формати метаданих. Для структуризації сховища даних використовувати Xml-Документ формату метаданих. Апробувати на прикладі сховища даних, характеристики плазмового джерела ультрафіолету отримані при експериментальних дослідженнях.

### **Завдання, які вирішувалися в роботі**

Для досягнення цілей у роботі вирішувалися наступні завдання:

- огляд існуючих баз даних
- структура сховищ даних
- data mining і його відображення за допомогою XML
- основні формати відображення метаданих
- створення програмного продукту, для структуризації сховища даних

## **Досягнуті результати**

Досягнути результату в розв'язку поставлених завдань автор захищає:

- аналіз розвитку баз даних
- переваги XML для структуризації сховищ даних
- аналіз і перспективи розвитку сховищ даних
- аналіз форматів даних для відображення метаданих
- створення програмного продукту для структуризації сховища даних

## **Наукова новизна роботи**

Науковою новизною роботи є аналіз програмного забезпечення для структуризації сховищ даних і перспективи розвитку технологій для B2B додатків. Це дало можливість зробити висновки про перспективу розвитку метаданих і сховищ даних.

Так само сюди можна віднести розроблений програмний продукт для структуризації сховища даних

## **Практична цінність роботи**

На підставі зробленого аналізу був створений програмний продукт для структуризації Сховища експериментальних даних і забезпечення застосування стандартних методів сучасної технології Data Mining з використанням середовища розробки Delphi 7.

## **Висновки:**

На сьогоднішній день обсяги інформації, які зберігаються на серверах баз даних вимірюються терабайтами. Дуже мало компаній, підприємств цікавляться «сирими даними». Набагато важливіше одержувати знання. Таким чином, необхідно, щоб розроблювачі ПО поставили перед собою завдання вмонтувати в сервера баз даних методи інтелектуального аналізу даних (а краще комбінацію з декількох методів, залежно від завдань), щоб у режимі реального часу можна було проводити аналіз даних, і при оновленні або додаванні початкової інформації користувачам відразу виводилися нові знання отримані з цього. Саме в такий спосіб користувач зможе

перестати думати про фізичну структуру таблиць і колонок. Для роботи зі сховищами даних розроблювачі ПО повинні запропонувати нову технологію, яка на відміну від класичних Ві-Систем, у яких вхідні дані спочатку завантажуються в сховище, очищаються й агрегуються й тільки після цього стають доступні для дослідження за допомогою створеної заздалегідь багатомірної моделі, дозволить почати обробку даних уже в момент їх вступу. Ця технологія повинна буде пропускати вхідні дані через систему фільтрів, виділяючи тільки значиму для конкретного бізнес-процесу інформацію. У результаті загальний обсяг збереженої історичної інформації скоротиться без втрати якості. Було б добре, якщо б технологія могла одночасно аналізувати дані з великої кількості джерел, реалізуючи технологію комплексного аналізу подій (Complex Event Processing, CEP).

Робота містить 110 стор., 22 мал., 1 табл., 20 джерел.

**Ключові слова:** DATA MINING, СХОВИЩА ДАНИХ, МОВА РОЗМІТКИ XML, БАЗИ ДАНИХ