

Abstract

Master's atestation work on:

"Development and implementation of metadata for structuring the data warehouse"

Vorobiova Olena

Relevance

The popularity of data warehousing in recent years has increased substantially. Competitive organizations are on the way of building data warehouses, or expansion, redesign and improve existing ones. Metadata are the key success factor in projects to implement repositories. They contain all the information needed to extract, transform and load data from source systems, as well as for future use and interpretation of the contents of data warehouses.

Objective

The aim of the work was the creation of software for structured storage of experimental data and ensure the application of standard methods of modern technologies of Data Mining using the IDE Delphi 7. Consider the basic displaying formats of metadata, such as HDF, NetCDF, FITS standard metadata formats. For the structuring of the data warehouse using XML-document format metadata. Approved by the example repository of data from experimental studies of the characteristics of the plasma source of ultraviolet radiation.

The problems that have been solved in the work

In order to achieve the objectives in the following tasks were:

- Review of existing databases
- The structure of Data Warehouse
- data mining and its representation using the XML
- major display formats of metadata
- creation of a software product for structuring data warehouse

Achieved results

Upon reaching the result in the task author is protecting:

- analysis of the development of databases
- benefits of XML for structuring Data Warehouse
- analysis and development prospects of the Data Warehouse
- analysis of data formats for mapping metadata
- creation of software for structured data storage

Scientific novelty

Scientific novelty of the work is analysis software for structuring data warehousing and prospects of technology development for B2B applications. This made it possible to draw conclusions about the future development of metadata and data warehousing. Also here can be attributed developed software for structuring data warehouse

The practical value

Based on the analysis made by the software was created to structuring the data warehouse experimental data and ensure the application of standard methods of modern technology Data Mining using the IDE Delphi 7.

Conclusions:

To date, the volume of information that are stored on the database servers Terabytes. Very few companies, businesses interested in "raw data". Much more important to gain knowledge. Thus, it is necessary that developers set themselves the task to embed in your database server data mining techniques (preferably a combination of several methods, depending on the task), so in real time, it was possible to analyze data, and modifying or add basic information to users at once deduced new knowledge gained from this. In this way the user can not stop thinking about the physical structure of tables and columns. To work with data warehousing software developers have to offer a new technology, which unlike classic BI-systems in which incoming data is first loaded into the store, cleaned and aggregated, and only then become available for research using pre-established multidimensional model, will start processing at the moment they arrive. This technology will have to pass the incoming data through a filtering system, allocating only meaningful for a specific usiness process information. As a result, the total volume of stored

historical data all be reduced without loss of quality. It would be nice if technology could simultaneously analyze data from multiple sources, implementing a comprehensive.

The work contains 110 p., 22 fig., 1 table, 20 Sources.

Keywords: DATA MINING, REPOSITORY, XML, DATABASE