Abstract

As an effective method to build Decision Support System, data warehouse has been one of the major topics in IT fields in recent years, but its applications in China are not as good as expected. Therefore, we always pay great attention to the combination of data warehousing methodology and practical projects. In this thesis, we discuss various theoretical and engineering problems in the design and implementation of data warehouse systems based on our prototype system development.

The dissertation starts with the architecture of data warehouse system and is focused on the data warehouse design, the data population and the metadata management. The advantages and disadvantages of each design method are explained and a more feasible strategy is presented. Meanwhile the new idea of using the data warehouse methodology in the CIMS data environment design is put forward and its necessity and feasibility is discussed. As to the data population, two difficulties, namely updates population and slowly changing dimensions, are first addressed. Then we present a novel population method combining the DTS tool and the DBMS programming. Its flexibility and efficiency is shown by examples. In the discussion of metadata management, we focus on the self-developed metadata management system based on the OIM and Microsoft Repository. For the completeness of our system, some important topics, such as data analysis, data quality and data security, are also investigated.

The design and implementation of Human Resource Data Mart in HY-CIMS is discussed as an example throughout the dissertation and clusters all the chapters.

In the end we discussed the marriage of data warehouse and the Web technologies. The architecture of Web-based data warehouse using XML is presented and its powerful functions are shown.

Key words: Data Warehouse, CIMS, Metadata, XML