Chapter I
Development of Data Warehouse Conceptual Models: Method Engineering Approach

Laila Niedrite
University of Latvia, Latvia

Maris Treimanis
University of Latvia, Latvia

Darja Solodovnikova
University of Latvia, Latvia

Liga Grundmane
University of Latvia, Latvia

ABSTRACT

There are many methods in the area of data warehousing to define requirements for the development of the most appropriate conceptual model of a data warehouse. There is no universal consensus about the best method, nor are there accepted standards for the conceptual modeling of data warehouses. Only few conceptual models have formally described methods how to get these models. Therefore, problems arise when in a particular data warehousing project, an appropriate development approach, and a corresponding method for the requirements elicitation, should be chosen and applied. Sometimes it is also necessary not only to use the existing methods, but also to provide new methods that are usable in particular development situations. It is necessary to represent these new methods formally, to ensure the appropriate usage of these methods in similar situations in the future. It is also necessary to define the
Development of Data Warehouse Conceptual Models

contingency factors, which describe the situation where the method is usable. This chapter represents the usage of method engineering approach for the development of conceptual models of data warehouses. A set of contingency factors that determine the choice between the usage of an existing method and the necessity to develop a new one is defined. Three case studies are presented. Three new methods: user-driven, data-driven, and goal-driven are developed according to the situation in the particular projects and using the method engineering approach.

INTRODUCTION

Data warehouses are based on multidimensional models which contain the following elements: facts (the goal of the analysis), measures (quantitative data), dimensions (qualifying data), dimension attributes, classification hierarchies, levels of hierarchies (dimension attributes which form hierarchies), and attributes which describe levels of hierarchies of dimensions.

When it comes to the conceptual models of data warehouses, it is argued by many authors that the existing methods for conceptual modelling used for relational or object-oriented systems do not ensure sufficient support for the representation of multidimensional models in an intuitive way. Use of the aforementioned methods also ensures a waste of some of the semantics of multidimensional models. The necessary semantics must be added to the model informally, but that makes the model unsuitable for automatic transformation purposes. The conceptual models proposed by authors such as Sapia et al. (1998), Tryfona et al. (1999) and Lujan-Mora et al. (2002) are with various opportunities for expression, as can be seen in a comparison of the models in works such as (Blaschka et al., 1998), (Pedersen, 2000) and (Abello et al, 2001). This means that when a particular conceptual model is used for the modeling of data warehouses, some essential features may be missing. Lujan-Mora et al. (2002) argue that problems also occur because of the inaccurate interpretation of elements and features in the multidimensional model. They say that this applies to nearly all conceptual models that have been developed for data warehousing. The variety of elements and features in the conceptual models reflect differences in opinion about the best model for data warehouses, and that means that there is no universal agreement about the relevant standard (Rizzi et al., 2006).

There are two possible approaches towards the development of a conceptual model. One can be developed from scratch, which means additional work in terms of the formal description of the model’s elements. A model can also be developed by modifying an existing model so as to express the concepts of the multidimensional paradigm.

The conceptual models of data warehouses can be classified into several groups in accordance with how they are developed (Rizzi et al., 2006):

- Models based on the E/R model, e.g., ME/R (Sapia et al., 1998) or StarE/R (Tryfona et al., 1999);
- Models based on the UML., e.g., those using UML stereotypes (Lujan-Mora et al., 2002);
- Independent conceptual models proposed by different authors, e.g., Dimensional Fact Model (Golfarelli et al., 1998).
Related Content

Data Warehouse Architecture: Practices and Trends
[www.igi-global.com/chapter/data-warehouse-architecture/38216?camid=4v1a](www.igi-global.com/chapter/data-warehouse-architecture/38216?camid=4v1a)

Data Warehouse Refreshment
[www.igi-global.com/chapter/data-warehouse-refreshment/7821?camid=4v1a](www.igi-global.com/chapter/data-warehouse-refreshment/7821?camid=4v1a)

On Association Rule Mining for the QSAR Problem
[www.igi-global.com/chapter/association-rule-mining-qsar-problem/10802?camid=4v1a](www.igi-global.com/chapter/association-rule-mining-qsar-problem/10802?camid=4v1a)

Specification of Components Based on the WebComposition Component Model
[www.igi-global.com/chapter/specification-components-based-webcomposition-component/7875?camid=4v1a](www.igi-global.com/chapter/specification-components-based-webcomposition-component/7875?camid=4v1a)