

Modeling Data Warehouses and Multidimensional Systems¹

Nicolas Prat
ESSEC Business School, France
prat@essec.fr

Data warehouses are a major component of data-driven decision support systems (DSS). They rely on multidimensional models, thereby easing data navigation and analysis via OLAP (On-Line Analytical Processing) tools. Data warehouse design is a complex task, which requires a systematic method. Few such methods have been proposed to date. In [Prat et al, 2006], we present and illustrate a UML-based data warehouse design method that spans the three design phases (conceptual, logical and physical).

Hierarchies are central to multidimensional modeling and hence their definition is crucial in our data warehouse design method. We have explored this issue in more depth, defining and illustrating rules for deriving multidimensional hierarchies from UML schemas, especially based on aggregation and generalization hierarchies [Akoka et al, 2001].

Since data warehouses are highly strategic in nature, quality issues raised by such systems are crucial. In the past, data warehouse quality has often focused on the issue of data quality. However, since OLAP tool users have direct access to multidimensional schemas, the issue of multidimensional schema quality evaluation is also crucial. We have defined criteria and metrics for assessing the quality of multidimensional schemas [Prat & Si-Said Cherfi, 2003] [Si-Said Cherfi & Prat, 2003]. Multidimensional schemas quality assessment may be used in conjunction with our data warehouse design method, e.g. to guide designers in the choice between alternative multidimensional schemas.

References:

[Akoka et al, 2001] J.Akoka, I.Comyn-Wattiau & N. Prat, "Dimension Hierarchies Design from UML Generalizations and Aggregations", 20th International Conference on Conceptual Modeling (ER 2001), Yokohama, Japan, November 2001.

[Prat & Si-Said Cherfi, 2003] N. Prat, S. Si-Said Cherfi, "Multidimensional schemas quality assessment", Decision Systems Engineering Workshop (DSE'03) in conjunction with the 15th Conference on Advanced Information Systems Engineering (CAiSE'03), Klagenfurt/Velden, Austria, June 2003.

[Prat et al, 2006] N. Prat, J. Akoka & I. Comyn-Wattiau, "A UML-based data warehouse design method", Decision Support Systems, Vol 42/3, December 2006.

[Si-Said Cherfi & Prat, 2003] S. Si-Said Cherfi, N. Prat, "Multidimensional schemas quality: assessing and balancing analyzability and simplicity", International Workshop on Conceptual Modeling Quality (IWCMQ 2003) in Conjunction with the 22nd International Conference on Conceptual Modeling (ER 2003), Chicago, Illinois, October 2003.

¹ This research is pursued in collaboration with Jacky Akoka (CNAM), Isabelle Comyn-Wattiau (CNAM & ESSEC) & Samira Si-Said Cherfi (CNAM).