

Modélisation Intégrée de la Dynamique des Systèmes d'Information Décisionnels

Estella Annoni, Franck Ravat
Olivier Teste, Gilles Zurfluh

IRIT Equipe SIG-ED,
F-31062 Toulouse Cedex 9, France
{annoni, ravat, teste, zurfluh}@irit.fr,
<http://www.irit.fr>

Résumé. Les systèmes d'information décisionnels (SID) sont des systèmes d'information (SI) qui ont pour objectif de faciliter la prise de décision à partir d'information résultant de processus complexes de dérivation et de préparation des données de SI sources. Ces processus sont généralement peu modélisés et sont directement implantés avec des logiciels spécifiques au cours des projets décisionnels bien que trois modèles particuliers ont été proposés pour représenter ces processus. En effet, ces modèles utilisent de nouvelles notations distinctes de celles de la modélisation des données qu'ils proposent. Ils requièrent deux schémas distincts pour les données et les processus alors que les schémas conceptuels des SID sont déjà nombreux et énormes en raison de la taille des projets et des spécificités des domaines.

Ainsi, nous proposons des outils pour la prise en compte de la dynamique des SID et une modélisation intégrée des processus dans le schéma des SID. Nous définissons les spécificités de la dynamique des SID. L'objectif est de prendre en compte dans un unique schéma aussi bien les processus liés à la dérivation des données étudiés par les travaux précédents que ceux liés à la préparation des données de l'environnement de prise de décision.

1 Introduction

Les systèmes d'information décisionnels (SID) résultent du besoin de réaliser des analyses complexes de données pour la prise de décision. Ces analyses sont appelées des analyses OLAP (On-Line Analytical Processing) et elles requièrent une modélisation multidimensionnelle (Kimball (1996)). La modélisation multidimensionnelle consiste à représenter les données comme un point dans un espace à plusieurs dimensions. Cette modélisation repose sur :

- le concept de **fait** qui représente le sujet de l'analyse. C'est l'information qui va retenir toute l'attention des décideurs. Il est caractérisé par des données appelées **mesures**,
- le concept de **dimension** qui représente un axe de l'analyse. Il est composé de **paramètres** en fonction desquels le fait sera exprimé. Ces paramètres sont organisés en hiérarchies et dans ce cas, ils sont dénommés **niveau** d'une **hiérarchie**.

Les données des SID dérivent de sources de données existantes. La fiabilité et la valorisation de ces données sont indispensables pour assurer une bonne prise de décision car elles permettent une analyse à partir d'informations justes et pertinentes.

D'une part, pour la fiabilité des données, il convient de définir les processus à l'origine de la dérivation des SID à partir des systèmes sources. Les travaux sur la dynamique des SID définissent ces processus, appelés ETL (Extraction, Transformation et Loading), après