

Evolution de modèle dans les entrepôts de données : existant et perspectives

Cécile Favre, Fadila Bentayeb, Omar Boussaid

Université de Lyon
Laboratoire ERIC - Lyon 2
5 av. Pierre Mendès-France, 69676 Bron Cedex
{cfavre,bentayeb}@eric.univ-lyon2.fr
omar.boussaid@univ-lyon2.fr

Résumé. Un entrepôt de données permet d'intégrer des sources de données hétérogènes à des fins d'analyse. Un des points clés de la réussite du processus d'entreposage de données réside dans la définition du modèle de l'entrepôt en fonction des sources de données et des besoins d'analyse. Une fois l'entrepôt conçu, le contenu et la structure des sources de données, tout comme les besoins d'analyse sont amenés à évoluer et nécessitent ainsi une évolution du modèle de l'entrepôt (schéma et données). Dans cet article, nous présentons un panorama de différents travaux portant sur l'évolution du modèle dans les entrepôts de données. Nous comparons et discutons ces travaux selon les critères qui nous semblent pertinents pour cette problématique. Nous dressons également les perspectives de recherche qui en découlent.

1 Introduction

Les données proviennent aujourd'hui de sources hétérogènes et présentent une volumétrie de plus en plus conséquente. La structuration et le stockage de ces données dans un entrepôt constituent alors un support efficace pour permettre des analyses en vue de prises de décision. Cette structuration est représentée grâce à un schéma caractérisant une organisation dite multidimensionnelle des données. Une organisation multidimensionnelle répond à l'objectif d'analyser des activités au travers de *faits* composés d'indicateurs, appelés *mesures*, ainsi que de *dimensions* qui constituent les différents axes d'observation des mesures. Ces dimensions peuvent présenter des *hiérarchies* qui offrent la possibilité de réaliser des analyses à différents *niveaux de granularité*.

La réussite du processus d'entreposage repose entre autres sur une bonne conception du schéma, puisque c'est ce dernier qui va déterminer les possibilités d'analyse de l'entrepôt. Ainsi, de nombreux travaux de recherche ont été menés sur la conception de schéma des entrepôts de données. Ces travaux témoignent aujourd'hui de la nécessité de prendre en compte à la fois les sources de données et les besoins d'analyse (Nabli et al., 2005), plutôt que d'avoir recours, par exemple, à une approche uniquement guidée par les données telle que celle proposée par Golfarelli et al. (1998).