

SISRO^{M2C} : Un Outil de Modélisation Conceptuelle à base Ontologique d'un entrepôt de données

Selma Khouri*, Ladjel Bellatreche**, Chimène Fankam**

*Ecole Nationale Supérieure en Informatique, Alger – Algérie
s_khouri@esi.dz

*LISI/ENSMA – Université de Poitiers, Futuroscope 86960 France
(fankamc, bellatreche)@ensma.fr

Résumé. Concevoir un entrepôt de données passe souvent par trois niveaux de modélisation: conceptuel, logique et physique. Contrairement aux deux derniers niveaux, le niveau conceptuel n'a pas eu le même intérêt de la part de la communauté des entrepôts de données, et peu de méthodes de conception d'un entrepôt ont été proposées. Délaisser cette modélisation a fortement contribué à l'échec de plusieurs projets d'entrepôt. Dans cet article, nous proposons une méthode, appelée SISROM2C et un outil de conception multidimensionnelle d'un entrepôt de données qui suppose l'existence d'une ontologie de domaine décrite en OWL (Web Ontology Language) couvrant toutes les sources participant dans le processus de construction de l'entrepôt de données. Elle confronte les besoins et les sources à priori, au niveau ontologique, et offre plus de présence des concepteurs dans la génération du schéma conceptuel.

1 Introduction

Un entrepôt intègre des données hétérogènes provenant de sources réparties. Avant de les charger, les données doivent être extraites, nettoyées, transformées selon le format cible du schéma de l'entrepôt. La plupart des travaux sur les entrepôts de données se concentrent principalement sur les niveaux logique (dé-normalisation) et physique (la sélection des techniques d'optimisation). Générer directement à partir d'un cahier des charges des schémas logique et physique sans passer par une modélisation conceptuelle pose plusieurs problèmes : (i) les schémas logique et physique n'offrent pas un niveau d'abstraction indépendant de toute implémentation, (ii) ils ne facilitent pas une communication entre les concepteurs et les utilisateurs qui est primordiale dans la conduite des projets, et est reconnue comme un risque important du rejet du projet final (Gam et Salinesi, 2006) et (iii) ils omettent la détection des éventuelles erreurs de conception qui doivent être décelées et corrigées dès le commencement du projet.

Pour donner plus d'importance à la modélisation conceptuelle dans la construction des projets d'entrepôt de données, trois approches de conception existent : *ascendantes*, *descendantes* et *mixtes*. Les approches *ascendantes* (Golfarelli et al., 1998), (Moody et Kortnik, 2000) ont été les premières à être proposées. Elles découlent de la définition d'Inmon (Inmon, 1992), qui indique que le développement des entrepôts repose sur les données des sources par opposition au développement des bases de données traditionnelles