

# Ontology Evolution and Source Autonomy in Ontology-based Data Warehouses

Dung Nguyen Xuan\* Ladjel Bellatreche\* Guy Pierra \*

\* LISI/ENSMA - Poitiers University  
1 Avenue Clément Ader 86960 Futuroscope - France  
(nguyenx, bellatreche, pierra)@ensma.fr

**Abstract.** Ontology-based integration systems (OBIS) use ontologies in order to describe the semantic of sources and to make the content explicit. Two major architectures of OBISs are available: (i) those using an unique ontology, and (ii) those using multiple ontologies. In the first architecture, all sources are related to one shared ontology. This architecture suffers from changes of ontology and sources which can affect the conceptualization of the domain represented in the ontology. Any change in the ontology may affect the sources. Therefore, sources are not really autonomous. In hybrid ontologies, each source is described by its own ontology, called local ontology. Each one references/maps a shared ontology in order to guarantee that each source shares the same vocabulary. The articulation between local ontologies and the shared ontology can be done either a posteriori or a priori. Two major issues are raised in this architecture: (i) evolution of the shared ontology and its consequence on the integrated system, and (ii) autonomy of the ontology and the local schema of each source. In this paper, we propose an approach and a model to manage asynchronous evolution of warehouse integrated systems where the articulation is done in an a priori manner. The fundamental hypothesis of our work, called *principle of ontological continuity*, supposes that an evolution of an ontology does not make false an axiom that was previously true. This principle allows to manage each old instance using the actual ontology. Therefore, it simplifies significantly the management of the evolution process and allows a complete automation of the whole integration process. Our work is motivated by the automatic integration of catalogs of industrial components in engineering databases. It has been validated by a prototype using ECCO environment and EXPRESS language.

## 1 Introduction

It is widely recognized that an automatic integration of heterogeneous data sources is one of the keys to improve management and productivity of several application domains like data warehouse, bio-informatic, e-commerce, etc. Generally, a data integration system combines the data residing at different heterogeneous and autonomous sources, and provides an unified, reconciled view of these data, called global schema, which can be queried by the users. A