

---

## Introduction

*Ce premier numéro de la revue ISI nouvelle formule porte sur la réingénierie des données et des documents. Les applications informatiques constituent un domaine dans lequel les besoins fonctionnels, les environnements techniques et le contexte sont en évolution permanente. Elles sont de complexité sans cesse croissante. Ceci explique que les techniques de réingénierie, à travers toutes leurs facettes, soient plus que jamais d'actualité même si l'urgence de l'an 2000 est passée.*

*Onze articles ont été soumis pour ce numéro, desquels quatre ont été sélectionnés pour leur originalité, leur qualité scientifique et leur portée générale.*

*Le premier article, proposé par I. Comyn-Wattiau et J. Akoka, est un état de l'art complet de la rétroconception des bases de données et des systèmes de fichiers. De tels états de l'art ont été faits dans le passé de façon partielle pour des formats particuliers de données (par exemple, les données relationnelles), mais pas, à notre connaissance, sur l'ensemble des types de données. Les auteurs commencent par définir les divers aspects de la réingénierie des systèmes dans lesquels figure la rétroconception des données. Après un paragraphe dans lequel ils retracent ce que furent les diverses générations de cette rétroconception, ils abordent successivement les principaux travaux concernant les bases de données de la première génération (modèles hiérarchiques et en réseau), les bases relationnelles, et enfin, de façon plus récente, les systèmes multidimensionnels et objets.*

*La seconde contribution, de C. Delcroix, P. Thiran et J.-L. Hainaut, propose un cadre générique pour la réingénierie des données. Celui-ci repose sur une description multiniveau et multiparadigme des données ainsi que sur des opérateurs transformationnels qui permettent d'établir des correspondances entre schémas. La méthodologie décrite est indépendante des technologies. Elle s'appuie sur un modèle générique de données dont les concepts recouvrent ceux de la plupart des modèles existants. Elle a été implantée dans un AGL par le biais de deux types de modules : les convertisseurs qui transforment les données sources en données cibles et les « wrappers » qui fournissent aux convertisseurs une interface transparente d'accès aux données sources. Un exemple de transformation de données de type COBOL en format XML est décrit en dernière partie de l'article.*

*Les deux derniers articles s'intéressent aux données ou applications accessibles sur le web. Dans le premier d'entre eux, S. Castano, V. De Antonellis, S. De Capitani et M. Melchiori se penchent sur le problème de l'intégration sémantique de sources de données web. La problématique de l'intégration de sources de données hétérogènes recouvre partiellement celle de la réingénierie pour ce qui concerne la recherche de la sémantique des données à intégrer et la construction d'ontologies permettant cette intégration. Ce papier met en évidence l'importance de l'analyse sémantique*

*des sources dans le processus d'intégration. A partir des termes décrivant ces sources (par exemple, sous forme de DTD du langage XML), une mesure d'affinité permet de déduire une ontologie et d'effectuer un regroupement des concepts en classes sémantiques, en vue d'une exploitation intelligente des données intégrées. Des techniques d'indexation semi-automatique des descriptions des sources permettent de sélectionner les termes les plus représentatifs de chaque classe.*

*Le dernier papier s'intéresse au problème encore largement ouvert de la maintenance des sites web. La complexité croissante des sites rend la maintenance difficile en l'absence d'outils permettant de localiser les effets d'une modification, en termes de propagation, sur les diverses composantes du site. L'approche choisie par L. Deruelle, N. Melab, M. Bouneffa et H. Basson consiste à modéliser de façon formelle le site sous forme d'un multigraphe à l'aide d'un modèle spécifique traduisant les composantes et leurs interrelations. Une taxonomie des modifications possibles sur le site est définie, qu'il s'agisse de modifications sur un composant logiciel ou sur un schéma de données. Les conséquences d'une modification sont alors propagées sur le graphe représentant le site. Ces propagations se font sous le contrôle d'un moteur qui utilise des connaissances du graphe et qui guide l'utilisateur dans la séquence des modifications en cascade.*

*Je remercie les auteurs de ces articles pour leur contribution à ce numéro spécial, ainsi que les lecteurs qui, par le soin qu'ils ont apporté à leur travail d'analyse et leurs commentaires détaillés, ont permis d'enrichir la qualité de ces publications.*

Jacques KOULOUMDJIAN  
Laboratoire d'Ingénierie des Systèmes d'Information  
INSA de Lyon - Université Lyon 1

**Liste des lecteurs de ce numéro**

Jacky AKOKA – CNAM

Françoise BALMAS – Université Paris VIII

Mohand BOUGHANEM – IRIT

Claude CHRISMENT – IRIT

Jean-Luc HAINAUT – FNDP Namur-Belgique

Reiner KOSCHKE – Université de Stuttgart, Allemagne

Jacques KOULOUMDJIAN – INSA de Lyon

Josiane MOTHE – IRIT

Jean-Marc PETIT – LIMOS

Farouk TOUMANI – LIMOS

Gilles ZURFLUH – IRIT