

Data Management in eCommerce: The Good, the Bad, and the Ugly

Avigdor Gal
Department of Management Science and Information Systems
Rutgers University
avigal@rci.rutgers.edu
<http://www.rci.rutgers.edu/~avigal>

1 Description

Electronic Commerce (eCommerce) is conceived as one of the major channels for performing commerce on a global scale, and the area is rapidly evolving. Electronic commerce is based on electronic transfer of data as a main vehicle of “doing business,” while the usage of conventional methods (paper, telephone, etc.) is dramatically reduced in such an environment. In particular, electronic exchange of data becomes the most cost-effective channel of performing low-cost transactions.

Enabling the transfer of data, in and by itself, is hardly sufficient to support efficient and correct processing of transaction in an eCommerce environment. Especially in B2B, enterprises have evolved independently, creating different data ontologies, and thus giving rise to semantic heterogeneity. This problem, which has been investigated for a few decades now, have received a new vigour in a computerized environment that is constantly changing.

This tutorial is aimed at introducing the data integration problem in eCommerce and to discuss three types of solutions to it, namely the good, the bad, and the ugly. We start by reviewing the problem of data integration for distributed heterogeneous information sources. Next, we identify the bad and the ugly that is currently used in eCommerce. Finally, we identify the good, or rather, discuss the research role in this field.

Proceeding in a reverse order, we classify the use of glue code to connect ECommerce tools with enterprise databases (e.g., EDI-aware COTS) to be the ugly. The bad category includes half-baked, communication-based solutions (e.g., EDI and XML), deployed indiscriminately across the enterprise. Finally, we define the good to be well-integrated data management sys-

tem to support translation, integration, and coordination of data transfer, combining a global standard with local management systems. While such systems do not exist in any generic form (and glue coding them is considered to be ugly), a good basis for such systems exist in enterprises in the form of database and workflow management systems. Therefore, the last part of the tutorial provides some ideas as to the establishment of good data integration in eCommerce. These ideas come from the research areas of federated databases and inter-enterprise workflows.

The tutorial is based on a half-day seminar, given at CAISE*99, and a full-day seminar, which was presented by the presenter at Twente University, the Netherlands on October 1999.

2 Instructor

Avigdor Gal is a faculty member at the MSIS Department at Rutgers University. He received his D.Sc. degree from the Technion-Israel Institute of Technology in 1995 in the area of temporal active databases. He has published more than 30 papers in journals (e.g. IEEE Transactions on Knowledge and Data Engineering), books (Temporal Databases: Research and Practice) and conferences on the topics of information systems architectures, active databases and temporal databases. Together with Dr. John Mylopoulos, Avigdor has chaired the “Distributed Heterogeneous Information Services” workshop at HICSS’98 and he was the guest editor of a special issue by the same name in the International Journal of Cooperative Information Systems. Avigdor is a member of the ACM and the IEEE computer society.

Avigdor has a practical, as well as academic, experience with Ecommerce. He worked as a consultant with MyZebra.com, a virtual enterprise specializing in customer service over the Internet. Also, he has conducted research in the area of data integration of federation of cooperative and non-cooperative databases, with applications to Web-based information services.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
ACM SIGMOD 2000 5/00 Dallas, TX, USA
© 2000 ACM 1-58113-218-2/00/0005...\$5.00