Data Warehousing and Business Intelligence
for E-Commerce
by Alan R. Simon & Steven L. Shaffer

Review by:
Frank G. Goethals, Catholic University Leuven
Faculty of Economics and Applied Economics
Department of Informatics
Frank.Goethals@econ.kuleuven.ac.be

Content of the book

This is no book about classical data warehousing, as the authors point out from the beginning: for information about data modelling, ETL and such, the reader is directed to other sources. This well-structured book is divided into two parts, the first one (120 pages) handling *concepts and business models and* the second one dealing with *building blocks, challenges, and solutions*.

The first chapter is devoted to background and terminology. Terminology in the covered discipline is decidedly non-standard and as a consequence it is well appreciated that the authors give precise definitions of some key concepts.

Chapters two to six cover different business models and discuss how data warehousing enabled Business Intelligence can be applied to these models. First, B2C business models are discussed. The authors differentiate between organisations that offer

- basic products;
- services;
- product and service packaging and brokering;
- portals and communities and
- support for traditional channels.

The second chapter also offers a 'data warehousing content framework for e-commerce', in which the different information-sources for a B2C e-business company are specified. Next the authors discuss *how* data warehousing should be used as part of B2C business operations. Two main categories of data warehouse usage are distinguished, namely customer focused Business Intelligence and operationally focused Business Intelligence. Both are discussed in more detail.

Chapter three looks at C2C and C2B, two derivatives of B2C e-commerce. C2C and C2B are covered in a separate chapter as these models - in contrast to the models covered in the previous chapter - are based on cross-consumer interaction. A closer look at the integration of Business Intelligence into a C2C business model shows the importance of the accordance between the data warehousing strategy and the organization's primary business mission.

B2B is dealt with in chapter four. Business Intelligence Models are introduced for two primary categories of B2B e-commerce:

- Supply-chain-oriented B2B and
- Marketplace-centric B2B

The concepts of the previous chapters can be applied to B2B models, but some significant distinctions should still be considered (e.g. the fact that the customer is not a person).

One might think all business models have been covered by now, but the authors prove the contrary. Chapter five introduces the concept of e-Government. First G2C is discussed. Business Intelligence seems to lag behind in this domain. However, the authors are

convinced that data warehousing will eventually become an integral part of G2C e-commerce and introduce some important topics (e.g. the necessity of a complete customer database). Next B2G data warehousing gets a chance. Many of the functions that comprise B2B also apply to B2G. Knowledge of the processes related to bidding and awarding contracts introduces some specific problems.

In chapter six (B2E models) the authors warn for a repetition of the "ERP syndrome": transactional capabilities might be deployed first, and Business Intelligence will need to be done later in catch-up mode.

The second part of the book (145 pages) is titled *building blocks, challenges, and solutions*. In chapter seven the following building blocks of an e-commerce data-warehousing environment are presented: network-related elements, database technology and vendor web development platforms. For each of the topics (as there are XML, cookies and WAP) a concise description is given, just as the implications for e-commerce data warehousing.

Chapter eight deals with three vendor front-end products for e-commerce intelligence (Vignette, Ithena and Revenio) that are representative of the various types of Business Intelligence tools currently in use.

Data quality and integrity issues come up for discussion in chapter nine. The data quality problem has been around for long time and is now, in the context of e-commerce, getting worse. Two complementary solutions to the problem are offered: a thorough analysis of the source data on the one hand and a close coupling between the data quality processes and the other customer-facing business processes on the other hand.

Chapter ten tackles information privacy and security related questions. First, some laws with respect to privacy are reviewed and the usage of cookies as a means to collect information is discussed. Next, the development of a security strategy is elaborated on, including threats (as there are spoofing, overloading and repudiation) and countermeasures (e.g. encryption). Finally, the data warehousing implications of privacy and security issues are discussed.

The concluding chapter (eleven) brings a fictitious case study in which all above-mentioned topics are considered.

Writing style

The book is well written, covers many real life examples and draws your attention to issues one may not notice in practice. That is why it seems better to read the book right from page 1 up to page 267. Leaving out chapters however would not introduce problems in understanding others. The book includes a handy index and contains many well-explained figures.

Target audience

Given the absence of technical details, this book is in the first place interesting to managers (of diverse companies) who want to know what is happening in the ICT-world. On the other hand, the book should also be read by DB-designers who wish to get a better understanding of business dynamics.

Reviewer's appreciation

Reading this book is a real pleasure and brings you a good view of how ICT, and more specific Data Warehouses and Business Intelligence, can be used to improve your business and your relationship with customers. Many companies have learned through mistake, and their mistakes are offered as lessons to us now. The book gives a good overview of the problem area, but for an in-depth study of a specific topic one might need more specialized literature. Consequently, I am somewhat disappointed the authors do not suggest any further readings. In my opinion there is one topic missing in the book, namely the sharing of information between companies in a collaborative network and the Business Intelligence that can be created through such collaboration.

In conclusion I can say it is an interesting, well-written book, with only some minor imperfections. It is clear the authors are on top of the current evolutions.