Syndicate Date Suppliers: Their Business Environment, the Industry, and Their Core Business Process

Mattias Strand
University of Skövde, Sweden

Sven A. Carlsson
Lund University, Sweden

INTRODUCTION

It has become increasingly important for firms to monitor the competitive forces affecting their business and competitiveness (Arnott & Pervan, forthcoming; Davenport & Harris, 2007). As a consequence, more and more attention has been directed towards data originating external to the own organizations, that is, *external data*. The increased interest is also shown in current literature and many authors have expressed the benefits thereof. The following quotations illustrate the perceived benefits of incorporating external data: (1) Oglesby (1999, p. 3) claims that "Companies who use external data systems have a strategic advantage over those who don't, and the scope of that advantage is growing as we move deeper into the information age," (2) Stedman (1998, p. 2) states that "external data helps us understand our business in the context of the greater world," and (3) Inmon (1996, p. 272) argues that "the comparison of internal and external data allows management to see the forest for the trees."

In alignment, a majority of companies incorporate their external data from organizations specialized in collecting, compiling, refining, and selling data (Strand, Wangler, & Lauren, 2004; Strand, Wangler, & Olsson, 2003). Kimball (1996) refers to these specialized and commercial data suppliers as *syndicate data suppliers* (SDSs).

The research area of external data incorporation is currently expanding and light is being spread on different aspects of external data incorporation. Still, the supplier side of the supplier-consumer constellation of external data is only fragmentarily covered. As noted above, most organizations incorporate their external data from SDSs. Therefore, this article intends to give a comprehensive description of SDSs' business environment, the industry they are competing in, and their core business process. The motive for describing the supplier side is two-fold. First, it fills a gap in current literature and contributes in making current research regarding external data incorporation more complete. Second, in describing the SDSs, organizations may increase their ordering-competency by creating an increased understanding of the SDSs work process, their industry, and other actors within the industry.

The material creating a foundation for this work originates from five interview studies, as well as two extensive literature reviews. The interview studies covered include: data warehouse (DW) consultants (two studies), consumer organizations (two studies), and one study towards SDSs. The studies were originally conducted within the scope of establishing a state of practice description regarding external data incorporation into data warehouses. The total number of interviews comprised 34 different respondents, all representing a unique company. The respondents were distributed as 12 DW consultants, 13 consumer organizations (banking, automotive, media, groceries, petroleum, and medical), and 9 SDSs. The interviews lasted on an average for 75 minutes and the transcripts ranged from 1370 to 7334 words (4214 words on average). Here, it is important to state that although the SDSs were, naturally, able to give the most detailed information regarding their industry, the two other groups of respondents contributed with details and aspects not mentioned by the SDSs. Finally, although current literature only contains fragmentary details regarding external data suppliers, a thorough review according to the instructions of Webster and Watson (2002), gave further details and contributed in making this work more comprehensive.

BACKGROUND

Literature accounts for two main directions related to the concept external data. First, external data may concern data crossing organizational boundaries, that is, the data is acquired from outside the organization's boundary (e.g., Kimball, 1996). Second, external data may also refer to any data stored or maintained outside a particular database of interest, that is, the data is external to the database but internal from an organizational point of view (e.g., Morzy & Wrembel, 2003). Since this work focuses on data which is exchanged between organizations, the direction accounted for by, for example, Kimball (1996), was adopted. In defining such data, the definition suggested by Devlin (1997, p. 135) was adopted, stating that external data is:

Business data (and its associated metadata) originating from one business that may be used as part of either the operational or the informational processes of another business.

External data may be acquired from different types of suppliers (or sources). Strand et al. (2003, p. 2466) account for the most comprehensive categorization of different suppliers. According to them, external data may be acquired from the following suppliers (or sources):

- Syndicate data supplier
- Statistical institute
- Industry organization
- County councils and municipality
- The Internet
- Business partner
- Biproduct data supplier

The different types of suppliers are described as follows. *Syndicate data suppliers* are organizations with the very core business idea of collecting, compiling, refining, and selling data to other organizations. Since they are the main focus of this article and will be extensively describe further on, the description thereof in this section is deliberately kept short. Different types of governmental *statistical institutes* deliver statistics concerning, for example, the labor market, trade, population, and welfare. Some of the data delivered from statistical institutes may be acquired for free based on legislative rights, but occasionally these institutes takes a commission for processing the data and for consulting. *Industry organizations* are also delivering data. Naturally, this data is specific and therefore often only interesting for a particular industry or even a subsection of an industry. Often, these industry organizations deliver industry averages concerning, for example, performance and sales, for comparisons with internal measures. *County councils and municipalities* may occasionally also deliver data. The data they deliver is similar to what governmental statistical institutes deliver, but narrower in its scope due to their geographic size. *The Internet* is considered as an unexplored source of data. Cao, Lim, and Ng (2003), Nørvåg (2002), Stolba and List (2004), Walter and Bjorking (2004), Zhu (1999), Zhu, Bornhövd, Sautner, and Buchmann, (2000), Zhu, Bornhövd, and Buchmann, (2001), and Zhu and Buchmann (2002) describe different applications for acquiring and sharing external data from Web pages. For example, the following applications are introduced: product pricing via competitors Web pages, preparation of a marketing campaign based on weather forecasts, and personnel planning based on promoted events advertised on the Internet. Still, the quality of the data acquired from the Internet is questionable and therefore many organizations hesitate in applying Internet data as a base-line for decision making (Hackathorn, 1999). *Business partners* are also possible external data sources. Normally when data is exchanged, the organizations are cooperating and the data may therefore be very specific. Therefore, this specific type of data supplier should not be considered as an "open" supplier for everyone else to buy from. Instead, a business partner is a very specific type of external data supplier. In addition, although the data is external according to the definition introduced above, it may be value-chain internal, making it even more difficult to, from a business perspective, consider it as an external data supplier. Finally, Strand et al. (2003) accounts for biproduct data suppliers. These organizations are generating large amounts of data as a result of their core business. This data may then be interesting for other organizations to procure. Strand et al. (2003) introduces an example adopted from Asbrand (1998), describing how the National Data Corporation/Health Information Services (NDC/HIS) in Phoenix, Arizona, sells its medical data to, for example, advertising agencies and stock analysts.