December 2003

Incorporating External Data into Data Warehouses: Characterizing and Categorizing Suppliers and Types of External Data

Mattias Strand
University of Skövde

Benkt Wangler
University of Skövde

Marcus Olsson
University of Skövde

Incorporating External Data into Data Warehouses: Characterizing and Categorizing Suppliers and Types of External Data[1]

Mattias Strand
University of Skövde
mattias.strand@ida.his.se

Benkt Wangler
University of Skövde
benkt.wangler@ida.his.se

Marcus Olsson
University of Skövde
a98marol@ida.his.se

Abstract

In order for organizations to survive and maintain their competitive edge, they must understand what is going on in their environment. One application, designed to contribute to such an understanding, is the data warehouse (DW). Traditionally, the focus of DW development projects has been on data internal to the organization, but as the competitive forces have become more important to understand and react upon, the importance of external data has been increasingly recognized. Unfortunately, the literature describing DWs provides little detail on what kinds of external data that might be incorporated and how they might support decision-making. Therefore, in this paper we present the results of an interview study conducted among DW consultants, and aimed at outlining the most important external domains about which to incorporate data, what data may be incorporated, from which sources they are acquired, and how they are used as a basis for decision making.

Keywords: External data, data warehouse, syndicate data

Introduction

During the last decade, the business environment of organizations has changed dramatically. Companies are competing in an ever-changing global marketplace. In order for organizations to survive and maintain their competitive edge, they must be flexible and carefully observe what is going on in their environment. If not, they may be overrun and out-competed by companies with a better knowledge of the competitive forces that constantly influence them (Porter, 1985).

To achieve such flexibility, organizations need systems which are able to present a holistic view of the organization and its business in a comprehensive manner (Gray, 1999). For a sales manager to keep track of sales of particular product categories in particular districts, he is usually not interested in keeping track of every single sale by every single salesman. Instead he wants an overall view that, for instance tells him the sales per day and per product category per district. A solution to these needs is to introduce a data warehouse (DW), which may then assist the managers in keeping track of the overall goings on in the business (Inmon, 1996; Kimball, 1996; Connolly and Begg, 2002).

A DW is *a subject oriented, integrated, non-volatile, and time variant collection of data in support of management's decision* (Inmon, 1996, p 33). DWs are designed to support the aggregation and integration of data from various internal, as well as

---

external, data sources. Traditionally, the focus of DW development projects has been on internal data. However, as it has become more and more important to keep track of the competitive forces, it has become apparent that also external information is crucial (Devlin, 1997). For instance, 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"*.

Unfortunately, the literature describing DWs and external data contains little detail on what external data that should be incorporated, what external sources these data may be acquired from, and how the external data may contribute in decision making. The lack of focus on the relationships between external data and DWs is rather paradoxical, since on one hand DW literature often emphasizes the suitability of DW technology for integrating data from various internal, as well as, external sources and, on the other hand, many organizations do use external data outside the data warehouse, as a basis for decision support (Damato, 1999). Therefore, to provide some answers to these questions, we present the results of an interview study conducted among DW consultants, aimed at contributing with an initial understanding of external data incorporation, and to briefly outline the current state of the art within the area. The paper includes descriptions of the most important external domains, a characterization and categorization of types of external data, a categorization of sources from which the data may be acquired, and some general examples on how the data are or may be applied as a basis for decision making.

The outline of the paper is as follows: The next section provides an overview of some related work. The following section defines and discusses briefly external data in general. After this, the interview study is described. The following and largest section presents and analyses the results of the interviews. The last section concludes the paper with a discussion that broadens the perspective of the paper and gives suggestions for future work.

Related Work

As indicated in the previous section, the literature covering external data incorporation into DWs is rather fragmented. Besides the general benefits suggested by e.g. Inmon (1996) and Singh (1998), most literature do not cover many of the most important aspects when incorporating external data into DWs, i.e. how are the data applied, from where are they acquired, which types of external data are there, and which knowledge domains are the data covering. For instance, Damato (1999) states that external data may be applied in the following ways (example queries that may be stated towards the DW are also included):

- *Classification*: Which customer group is generating the highest profit?
- *Explanation*: Why is the return rate exceptionally high for a particular type of product?
- *Meaningful comparison*: How is the company performing if compared to an industry average?
- *Planning*: In order to increase the sales of a particular product, towards which customer group should we target the next marketing?

Unfortunately, Damato (1999) does not give any thorough description of which types of external data that could contribute in such application areas and from where such data may be acquired. Kelly (1996), on the other hand, is very specific on different types of external data, as he gives a categorization of 10 different types of external data, but he does not give any detailed examples on how the external data may contribute and from which sources they may be acquired.

The most complete description that we have found is provided by Oglesby (1999), as he indicates some application areas and relates these areas to relevant external data types and available suppliers. However, Oglesby (1999) has directed his description towards direct marketing and his ideas and examples are therefore illuminated specifically in that light. Unfortunately, this makes his description a bit narrow. Kelly (1996), for instance, claims that incorporating external data has the following three objectives: to recognize opportunities, to detect threats, and to identify synergies. Since, these opportunities, threats, and synergies may be found in many areas, besides direct marketing, Oglesby (1999) is not presenting a complete picture of external data incorporation. The present work, on the other hand, aims to cover all aspects mentioned by Kelly, without any specific focus on particular business areas.

External Data

According to Singh (1998), one of the main reasons for a DW is to set free the information that is locked up in the internal operational databases and to mix it with information from other external sources. Singh (1998) further advocates that organizations should increase their acquisition of data originating from outside their own systems boundaries. Generally, these data include some

level of market information and might, for instance, be information related to economic forecasts, political information, customer demographics, and competitive and purchasing trends. Inmon (1996) is even more outspoken concerning the contribution of external data in a DW. He claims that even though external data do not say anything directly about particular companies, they may still give a lot of valuable information about the environment that the company must work and compete in. When comparing internal data to external data, one of the most useful things is the ability to perform comparisons over a period of time. The comparison allows management to *"see the forest for the trees"* (Inmon, 1996, p. 272).

It is not possible, in the literature, to find one common definition of external data. From a DW perspective, we have found several notions that relate to data that have their origin outside the own organization. In this work, we have chosen to discuss and compare two of them. Firstly, Kimball (1996) refers to *syndicate data*, which are data purchased from data suppliers, such as A.C. Nielsen, IRI, and IMS. However, for this work, we have adopted a definition suggested by Devlin (1997, p. 135). He defines external data as:

> *"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."*

The reason for adopting the definition by Devlin (1997), instead of Kimball's, is that Devlin is more general and his definition may also include data acquired from e.g. business partners and from the Internet. We consider the generality of the definition as important, since the aim of our work was to outline current incorporation of any external data and not specifically external data acquired from syndicate data suppliers. Still, the idea of syndicate data is important, since much external data that are incorporated into DWs are, in fact, acquired from syndicate data suppliers (Kimball, 1996; Strand and Olsson, 2003). The importance of syndicate data has also been shown within other areas. For instance, Salmeron (2002) shows that the incorporation of external data has increased in executive information systems (EIS) and that most of those data are acquired from organizations that Kimball would refer to as syndicate data suppliers.

The Interview Study

Data for the study were primarily collected through interviews. However, although the amount of literature covering this topic is limited, we have also chosen to include some information acquired from literature. This information was mostly used to contrast and discuss the conclusions resulting from the interviews. The following steps were used to guide the preparation and accomplishment of the interviews:

(1) *Selecting the respondents*. Consultants were considered as the most appropriate category of stakeholders for an exploratory study such as this, since consultants have often been working in many different projects and therefore have a broader and more general knowledge of the topic than personnel in specific user organizations. The respondents were selected by searching for consulting companies on the Internet. Different company size, geographical distribution, and a long experience were considered as important parameters for us to be able to get as general knowledge as possible. This resulted in a respondent group with at least 3 years of DW experience, which was geographically distributed and represented both large and internationally well-known consulting firms and national companies with 5-10 employees.

(2) *Constructing the interview questions*. The set of interview questions was split into three groups. The first group was introductory and aiming to collecting background material and to verifying that the respondent accepted the definitions of external data and DWs chosen for this work. The second group was the main part of the interview and aimed at collecting material for the analysis. Since we had found no previous study on external data incorporation that we could use as a basis for formulating the interview questions, we decided to build the interview around the aspects mentioned before, i.e. important external domains, sources of external data, applications of external data, and suppliers of external data. Questions that were included concerned e.g. what external data have been incorporated in the DW development projects that you have been involved in; what are the most common external data sources to acquire data from; and how is the selection of external data done? The last questions aimed at letting the respondents introduce additional ideas and aspects not covered by the main questions. The analysis accounted for in this paper is based on material from this group.

(3) *Initiating the interviews*. The interview questions were sent to the respondent in advance so as to let them read through the questions and reflect upon them before the actual interview and to consider whether they accepted our definitions. If the respondents were having different definitions, the interview results might not have been comparable and not possible to use in the analysis. A personalized cover letter explaining the purpose of the study accompanied the interview questions.

(4) *Conducting the interviews*. Due to the timeframe of the study (February - May, 2002) and due to limitations in the financial resources available for the study, the interviews were done by telephone. Every interview lasted for approximately 45 minutes. After the interviews, the answers were written down. The compiled material was sent back to the respondent for reviewing and authorization. In this way, we avoided errors in the material and misinterpretations were corrected.

The result of the interviews gave us answers from 12 respondents. None of these was excluded due to differing views on definitions. Every respondent returned their reviewed notes, along with a clearance for including them in the study. The final response rate from the scheduled interviews was 86% (12 out of 14). The missing two respondents were scheduled, but at the time for interviewing them, they were unavailable. After repeated, but unsuccessful, attempts to contact them, they were excluded. The number of respondents may be on the low side for being able to generalize upon the findings and that some important knowledge is, hence, missing. However, the study was not aimed at given a complete picture of external data incorporation. Instead, the intention was to acquire an initial understanding of the main types of sources, suppliers, and their relationships, for being able to briefly outline the current state of the art and to propose future work.

Analysis and Results

In the following sections, the analysis and the results of the interview study will be presented. The chapter is divided into the following sections: important external knowledge domains, external data classes and sample applications, and external data sources.

Important External Knowledge Domains

As a starting point for the discussion of which external data to incorporate into DWs, we began with identifying which are the most important knowledge domains to be informed about. The analysis of the collected material showed that the external domains identified as most desirable and interesting for organizations were: *Customers, competitors, business partners, and business market*. These domains were deemed directly related to explicit and important knowledge needs concerning the organizations' performance regarding e.g. customer satisfaction, competitive positioning, etc. In other words, these are knowledge domains where external data is a prerequisite for being able to stay informed.

To be able to generalize upon our findings, we intended to compare our results with previously conducted studies, but due to the small amount of literature related to the incorporation of external data into DWs, we have found no similar study related to DWs. However, in other research areas, related to data warehousing, such studies have been conducted. One such is a questionnaire study, conducted by Alavi and Leidner (1999). Their study included more than 500 top-level managers from 5 continents and was partly aimed at identifying which domains that were considered most important for organizations which had or were building knowledge management systems (KMS). Since DWs are often considered as KMSs or as an important component in KMSs (e.g. Gray and Watson, 1998; Alavi and Haley, 1999; Gray, 1999), we regard it as relevant to compare the outcome of the Alavi and Leidner study, with the results from our study.

The managers in the Alavi and Leidner study were asked to assess the importance of business domains, by giving the domains an integer from 7 down to 1 in falling importance (most important = 7). The most important domains were (resulting average in parentheses):

- Customer service (5.14)
- Business partners (4.83)
- Internal operations (4.62)
- Competition (4.57), Marketing/Sales (4.57)
- Suppliers (4.56)

Comparing the results of our study, with the results from Alavi and Leidner (1999), it is clear that they are well aligned. Even though other terms are used, the domains of our study are in alignment with the results of the Alavi and Leidner study. The only domain from our study, which is not explicitly stated, is the market domain. However, since most of the domains given by Alavi and Leidner (1999) are external to organizations and constitute important components of an organizations' market environment, we still would like to claim that also that aspect is covered.

External Data Classes and Sample Applications

The respondents provided several examples of what external data that are incorporated into DWs. We will start by describing the two types of external data and applications that were given most attention. Thereafter, other types of data and applications will be exemplified and described.

Firstly, the interviews indicate that industry codes are frequently incorporated in DWs. Most respondents have mentioned that they were using SNI[2] codes. SNI is an index that describes which industry an organization belongs to. This index further divides the industry into several levels of industrial sectors, and may be used to sort out companies in a step by step search. The SNI codes are also used when designing the aggregation levels of star-schema dimensions related to companies/customers. Secondly, address information was frequently mentioned and several applications were exemplified. In general, externally acquired address information was considered as important, since it allows the organization to keep track of the whereabouts of its customers, suppliers, and other actors in the marketplace. Staying updated on addresses was considered as almost impossible without the incorporation of external data, since customers and other actors tend to move around without much discipline on notifying address changes. Such applications are also given attention to in literature. For example, Oglesby (1999) indicates that address/phone verification is one out of four application areas that you may support by integrating external data. On a more detailed level, a number of different applications were given:

(1) External address data is used when analyzing market shares in specific regions. Address data allow organizations to derive demographic data which may be used for e.g. identifying regional coverage. For example, a newspaper may want to identify the proportion of households covered in a particular city or a particular neighborhood. External address data may be used to derive the number of households in the particular area of interest. That number may then be compared to the number of customers derived from internal data. The quotient gives the newspaper's coverage of households in the area.

(2) In order to avoid "bad will" external address data are used during the preparation of targeted marketing campaigns and in customer relationship management (CRM) efforts. Since the aim of both these is to create an intimate relationship to the customer, it is important that the address is correct and that names and addresses are correctly spelled. Otherwise, the customer may get a sloppy impression of the advertising organization and thereby develop a non-buying attitude.

(3) Updating and refreshing the customer dimension with external data may also save money for the organization, since correct and updated address data may reduce the number of costly returns of e.g. payment notifications, member information, or indirect advertising that were sent to an outdated address.

Besides industry codes and address data, a few other examples of external data and applications were suggested by the respondents. In the following paragraphs, these will be exemplified and described.

First of all, various kinds of statistical data and data from marketing investigations were suggested during the interviews. These data are private customer oriented and relate strongly to organizations selling business-to-consumer (B2C) products or services. Examples are e.g. customer target groups, population statistics, age groups, and citizen educational levels. These data are incorporated to fully understand the customers with respect to e.g. movement, behaviors, or purchasing-patterns. Such insights are important when planning and designing one-to-one marketing campaigns, since they assist organizations in identifying potential customers and in marketing products aligned to the needs of these potential customers.

The interviews also gave examples of external data that were important for organizations acting in business-to-business (B2B) constellations. Here, annual accounts data, risk class data, credit ratings and risk forecasts have been stated as particularly important to acquire. Risk forecast, for instance, includes information about the risk of future insolvency of a customer or business partner. The liability of a company is important as a support in making strategic decisions about e.g. what companies are suitable to initiate partnerships or to do other business with. The risk class data covers both credit ratings and risk forecasts, and is rated 1 to 5, where 1 indicates the highest risk and 5 indicates the lowest risk of becoming insolvent within a two-year period. General information on other organizations and their products was also considered as interesting, in order to fully understand the conditions and needs of customers, competitors and business partners.

---

[2]Svensk Näringsgrenindelning (in Swedish).

Moreover, data related to frequently changing rates, such as stock prices, currencies, and oil prices, were also mentioned. However, the applications of such data have been rather vaguely expressed and we have not been able to determine their contribution. Still, one might imagine that such data are collected and incorporated into DWs for convenience reasons, i.e. to have all data at one (logically) access point, instead of manually acquiring such data at several, distributed websites.

In conclusion, the external data identified during the interviews, may be classified into the following categories:

- *B2C customer data*
- *Business-to-business data*
- *Industry/Industry Sector data*
- *Governmental/State data*
- *Products/Services data*

As indicated above, it is deemed very important to incorporate information about customers into the DW. Hence, many examples of customer oriented data were given. However, since the data related to customers are covering a myriad of different aspects, we have chosen to split them into two separate categories. We have made a distinction between private persons and organizations. This was regarded as important, since the data covering the two classes are very different and the diversity in the data coverage also affects the applicability of the data. For example, *B2C* data cover names, addresses, as well as purchasing behavior and population statistics, and is mostly aimed at supporting the design and follow-up of marketing campaigns. In contrast, *B2B* data relate to branch index, credit ratings, number of employees, and business partners, etc. Such data are also used in marketing initiatives, but they are also widely used when assessing current or possible business partners. That is, B2B data do not only concern customers, but may equally well be related to suppliers or other business partners, e.g. in a network or a value chain of organizations.

In addition, the next two classes: *Industry/Industry sector* data and *Governmental/State* data cover more general data, aimed at broadening the perspective of the organization and its relationship to the marketplace. These data are rarely used in isolation. Instead, they are used to contrast the internal numbers. For example, consider a sales manager: If she derives from the DW that the sales of all products for the last years are constantly decreasing, it may be interesting to compare these figures with averages for the industry sector. Such averages are often possible to acquire from industry organizations. In addition, governmental/state data may also be included for comparison. Governmental/state data may be e.g. demographic data related to the whole country or for particular regions, or data on expected and actual economical growth for different regions. In our example, such data might indicate that the figures follow the decline of economical growth for the whole region.

Finally, we suggest *products/services* as a particular class. Products and services are the results of the production that generates a value for the organization. Therefore, it is important to stay informed on products/services and their constituent components. Such data may relate to weights, quality certificates, testability/reliability rates, etc.

External Data Sources

In this section, sources of external data will be categorized, characterized, and exemplified. First of all, one of the most common categories of external data sources is the syndicate data suppliers. The interviewees gave some examples including both national and international suppliers. The national sources that were mentioned were PAR, SPAR and UC (these are all marketed and sold by Swedish syndicate data suppliers). PAR contains information about addresses, and other data about companies such as which industry sector they are active in. This could be used to assure that all addresses are coupled to a correct postal number, and that all changes in localities are updated. SPAR is a state public names and addresses register. SPAR addresses organizations that want to keep their customer or member register updated. It may be concluded that these kinds of sources are used mainly in DWs that are meant to be used for campaigns or as support in customer relations management.

UC AB is a syndicate data supplier owned by Swedish banks. It is Sweden's largest business and credit information agency, and offers reports to support credit and commercial decisions, credit monitoring, and qualified financial analysis. The reports are produced from a database, which contains information on all enterprises registered in Sweden and all Swedish citizens of an age above 16. UC is accessible online and the customers to UC are subscribing to an on-line membership, which means that the information can be acquired when needed.

Among international suppliers of syndicate data, organizations such as Dun & Bradstreet have been mentioned. Dun & Bradstreet deliver economical information about companies and they store data related to more than 70 million companies worldwide. Their data are often used to assist organizations in reducing credit risks, finding profitable customers and managing vendors efficiently.

Several respondents mention SCB (Statistics Sweden) as a common external source. SCB offers official statistics concerning various fields, such as e.g. the labor market, the economy, trade and industry, prices, population and welfare, housing and construction. SCB also takes on commissions for specially processed data, data collection and consulting. SCB supplies a great deal of different data that may be incorporated in a DW. Often, changes in a turbulent market have various external causes, and the type of data that are possible to acquire from SCB may provide basic information, which supports decision-making in that context. The information could, for instance, be used in customer segmentation and target marketing.

Two of the respondents also remark that there are marketing investigators, which supply statistical information concerning consumers and products etc. Also industry sector organizations may be appropriate to mention in this context. They offer statistics specific for a certain industry sector. These data may be valuable when comparing internal numbers with e.g. performance or sales averages for the whole industry sector. In addition, the industry organizations often have data on the organizations related to the industry, as well as legal and governmental laws and rules regulating the industry. Unfortunately, some respondents had experienced that it is often difficult to acquire such data, in particular if you are interested in industry data that originate from an industry outside your own. The impression is that there seems to exist an agreement in some industries that the industry internal data are not to be spread externally.

Information from other enterprise systems such as those of your business partners are also possible external sources. Often, these organizations constitute links in value chains or cooperate in networks. Moreover, public organizations may also supply information. A benefit, in that context, is that the information is often free to collect. The information may, for instance, concern statistics from county councils and municipalities. Furthermore, some responses have pointed to the Internet as an enormous source of external data. From a DW perspective, web technology is extensively used as a distribution channel for transferring data between organizations, but there are also efforts aiming to automating the acquisition of web content into DWs such as e.g. *Web Farming* (Hackathorn, 1999).

In the following, we summarize and briefly describe the external data sources that were identified during the interviews.

• *Statistics Institutes*: Such as SCB (Statistics Sweden) that offer different kinds of statistics, related to industries, persons, and organizations.
• *Syndicate data suppliers*: Provide economical company data e.g. credit refinements on companies. Dun & Bradstreet offer data on more than 70 million companies worldwide – for credit, marketing and purchasing decisions. The aim is to help businesses to reduce credit risk, find profitable customers and manage vendors efficiently.
• *Industry organizations:* These organizations offer information about particular industry sectors. The data are often rather detailed, but may be difficult to acquire if you are not in the particular industry yourself.
• *County councils and municipalities:* External data related to population statistics and demographics. These data are often free, but may need major transformations before being possible to incorporate in a DW.
• *The Internet*: The Internet may be used to look for competitors, by watching what they offer and to what prices. The Internet was considered as an enormous resource, if it was possible to sort out the useful information among the myriads of data that it contains.
• *Business Partners:* Data from other companies, or data acquired from suppliers' systems were also considered as contributory, since they may help in getting an overview of the performance of the environment.
• *Bi-product data suppliers*: There are companies, which create huge amounts of data as a bi-product of their core business, and that are selling these data to other organizations, as a way of letting the large amounts of data cover their own costs.

We have included *Bi-product data suppliers* as a new class of external data suppliers, although it was not mentioned in the interviews. We include this class for completeness, since there are in fact many organizations that are selling data that are a bi-product of their core business. We have chosen not to consider these organizations as pure syndicate data suppliers, since they have a core business which is not targeted towards selling data. Asbrand (1998) exemplifies with medical organizations which are selling pharmaceutical data to e.g. advertising agencies and stock analysts. In particular, she mentions the National Data Corporation/Health Information Services (NDC/HIS) in Phoenix, U.S., which has opened its 3.5-terabyte DW to any company who wants to get access to medical data. From a Swedish perspective, examples are Swedish Meteorological and Hydrological Institute (SMHI) and Svensk Bilprovning (Swedish vehicle inspection institute).

Table 1. The Data Suppliers' Coverage of Different Classes of External Data

Data class Data supplier type	B2C	B2B	Industry	Gov./State	Products/ Services
Statistics Institutes	X	X	X	X	
Syndicate data suppliers	X	X	X		
Industry organizations		X	X	X	X
County councils and municipalities	X			X	
The Internet	(X)	(X)	(X)	(X)	(X)
Business Partners		X	X		X
Bi-product data suppliers		X	X		X

In Table 1 we summarize the data suppliers and relate them, with respect to coverage, to the classes of external data suggested in the previous section. The Internet coverage is considered as total. However, identifying relevant data on the Internet is difficult and the quality of the data is also difficult to assess (Hackathorn, 1999). Still, we include the Internet as a possible source of all classes of data, but we indicate the identification and quality issues with brackets.

Concluding Remarks

When comparing the external data identified during the interviews, with the competitive forces suggested by Porter (1985), it is clear that organizations are currently incorporating external data related to customers, suppliers, and competitors. The data are, however, not easily covering new entrants or substitute products. This is not surprising, since the knowledge needed for being able to identify or predict new entrants or substitute products is simply not available in a direct and evident form. Instead it has to be discovered by combining information from various sources such as newspapers, magazines, WWW, e-mails, memos etc. Such data formats are not suitable for DW incorporation, without extensive manual efforts (Kimball, 1996). In other forms they may, on the other hand lend themselves to techniques such as data mining.

Finally, the initial findings show that there are some issues that may be considered for future work. A similar study could be directed towards the investigation of external data incorporation among the users of external data. From the users' point of view, there are many issues to be resolved, such as technical and conceptual integration of the external data with internal figures and numbers, for the incorporation to be truly successful and beneficial. In conjunction, the experiences of the suppliers of external data, such as syndicate data suppliers and semi-syndicate data suppliers, would also be interesting to investigate, since their view on external data incorporation may suggest new perspectives on important issues and challenges. In other words, contrasting the experiences of the consumers with the experiences of the data suppliers could be synergetic, since it is not unrealistic to believe that challenges experienced by the consumers may be opportunities for the suppliers, and vice versa. For example, if the user organizations find it difficult to technically integrate the external data with their own internal data, the suppliers may see this as an opportunity and increase the degree to which they would offer to tailor the data and align them to the particular customer's needs.

References

Alavi, M., Haley, B. "Knowledge Management Systems: Implications and Opportunities for Data Warehousing", Journal of data warehousing (4:1), 1999, pp. 2-7.

Alavi, M., Leidner, D. "Knowledge Management Systems: Emerging Views and Practices from the Field", In Proceedings of the 32nd Hawaii International Conference on System Sciences (HICSS), 5-8 January, Maui, Hawaii, 1999.

Asbrand, D. "Making Money from Data", Datamation article, acquired from: **wysiwyg://56/http://datamation.earthweb.com/ dataw/11tran.html**, printed 000621, 1998.

Connolly, T., Begg, C. "Database Systems: a Practical Approach to Design, Implementation and Management", 3rd edn., Addison-Wesley Longman, Harlow, 2002.

Devlin, B. "Data Warehouse: from Architecture to Implementation", Addison Wesley Longman, Harlow, 1997.

Damato, G. M. "Strategic information from external sources – a broader picture of business reality for the data warehouse", acquired from **http://www.dwway.com/file/20020726170552_get_ext_data_pdp**, printed 030320, 1999.

Gray, P. "Data Warehousing – Three Major Applications and Their Significance", In Kendall, K. E. (ed.), Emerging Information Technologies: Improving Decisions, Cooperation, and Infrastructure. Thousand Oaks, Calif, 1999.

Gray, P., Watson, H. J. "Present and future directions in data warehousing", The DATA BASE for Advances in Information Systems (29:3), 1998, pp. 83-90.

Hackathorn, R. D. "Web Farming for the Data Warehouse", Morgan Kaufmann Publishers, Boston, 1999.

Inmon, W. H. "Building the Data Warehouse", 2nd edn., John Wiley & Sons, New York, 1996.

Kelly, S. "Data Warehousing the route to customization – updated and expanded, John Wiley & Sons, New York, 1996.

Kimball, R. "The Data Warehouse Toolkit", John Wiley & Sons, New York, 1996.

Oglesby, W. E. "Using external data sources and warehouses to enhance your direct marketing effort", acquired from **http://www.dmreview.com/editorial/dmreview/print_action.cfm?EdID=1743**, printed 030221, 1999.

Porter, M., E.: Competitive Advantage: Creating and Sustaining Superior Performance. Free Press, New York, 1985.

Salmeron, J. L. "EIS data: findings from an evolutionary study", The Journal of Systems and Software (64), 2002, pp. 111-114. Elsevier.

Strand, M., Olsson, M. "The Hamlet dilemma on external data in data warehouses", In Proceedings of the 5th International Conference on Enterprise Information Systems (ICEIS) - Part 1, Camp, O., Filipe, J., Hammoudi, S. and Piattini, M. (eds.), 23-26 April, Angers, France, pp. 570-573, 2003.

Singh, H. "Data Warehousing: Concepts, Technologies, Implementations, and Management", Prentice Hall, New Jersey, 1998.