

Syndicate Data Incorporation into Data Warehouses: Contrasting Consumer Problems with Supplier Viewpoints

Mattias Strand and Björn Lundell

School of Humanities and Informatics, University of Skövde, Sweden
Box 408, SE-541 28 Skövde,
{mattias.strand,bjorn.lundell}@his.se

Abstract. The research reported in this paper extends current knowledge of problems related to syndicate data incorporation into data warehouses (DWs). It does so by employing an interview study towards syndicate data suppliers (SDSs). The result gives a contextualization of syndicate data incorporation problems, from a supplier perspective, by contrasting the problems experienced by user organizations. The results show that the user organizations and the syndicate data suppliers have unresolved interoperability issues. Specifically with respect to the business and knowledge layers, whereas the interoperation on the ICT-systems layer seems to be prioritized and therefore achieved a higher degree of interoperability.

Keywords. Data Warehousing, Syndicate data, Syndicate data supplier

1 Introduction

Organizations have problems with their initiatives on incorporating syndicate data into data warehouses (DWs) and therefore, they are not able to fully exploit the potential thereof (Strand et al., 2003 and Strand and Wangler, 2004). For clarification, a DW is a: “*subject-oriented, integrated, non-volatile, and time variant collection of data in support of management’s decisions*” (Inmon, 1996, p.33). From a user organization perspective, these problems have been categorized and verified, whereas the supplier side of the problem, i.e. how they interoperate with the user organizations, and tensions between the suppliers and the consumers¹, is unexplored (Strand et al., 2005). The syndicate data supplier (SDS) perspective is important, as the user organizations incorporate most of their external data from these specialized data suppliers (e.g. Strand et al., 2003). Therefore, we present the results of an interview study towards SDSs, partly aimed at contrasting the enhanced list of

¹ For the rest of this paper, *user organization* and *consumer* will be used interchangeably to denote an organization that buys syndicate data from one or several SDSs.

syndicate data incorporation problems experienced by user organizations (Strand et al., 2005), with the viewpoints, experiences, and work routines of the SDSs.

Contrasting the problems experienced by the consumers, from a SDS perspective, is important for at least three reasons. Firstly, it allows for an evaluation of the interoperability between the user organizations and the SDSs. Syndicate data incorporation is a multi-faceted undertaking, covering organizational as well as technological aspects and therefore such an interoperability evaluation may cover all three layers given by Chen and Doumeingts (2003): the business layer, the knowledge layer, and the ICT systems layer. Secondly, it extends the current body of knowledge of syndicate data incorporation into DWs, by describing tensions and collaborations between the consumers and the SDSs. Thirdly, it contributes with important details when developing a hands-on support for organizations incorporating syndicate data into their DWs (as suggested by Strand and Wangler, 2004), by including details on e.g. how they establish business relationships with their customers and to what degree the SDSs allow the user organizations to tailor the syndicate data.

In articulating our findings, we identify a number of important collaborations and tensions between the consumers and the SDSs, indicating that the two parties are rather mature in their interoperability on the ICT-systems layer, whereas more attention should be given the interoperability on both the business- and knowledge layers.

The rest of the paper is organized as follows: Section 2 describes briefly some related work. In Section 3, the research method is described and motivated. Section 4 presents the analysis of the interview study. Section 5 gives the results of the paper. In Section 6, we summarize the paper and outline future work.

2 Related Work

Kimball (1996) gives the first instances of the concepts *syndicate data* and *SDSs*. However, he does not define syndicate data. Therefore, we have adopted the following definition from Strand et al. (2005): “*Business data (and its associated metadata) purchased from an organization specialized in collecting, compiling, and selling data, targeted towards the strategic and/or the tactical decision making processes of the incorporating organization*”.

Furthermore the syndicate data incorporation problems, experienced by the user organizations, are given in Table 1. The problems are categorized according to the activities of the external data incorporation process, i.e. identification, acquisition, integration, and usage. (For a more detailed description of the process activities we refer to Strand and Wangler, 2004). These activities are not unique for syndicate or even external data (external data is a more general concept, including syndicate data as a subtype, due to its acquisition from specialized data suppliers. Other external data may be acquired from e.g. industry organizations or business partners (Strand et al., 2003)). On the contrary, they may be considered as rather generic, since they are also included in most data warehouse development processes, regardless of whether the

data has an internal or external source (e.g. Inmon, 1996; Hammer, 1997; Hessinger, 1997; and Damato, 1999). Still, the process of including syndicate data (or external data) differs from the process of incorporating internal data in two ways: 1) the data is acquired from outside the organizations and thereby crosses organizational boundaries, which may cause other types of problems than those experienced with internal data (Oglesby, 1999, Strand and Wangler, 2004) and 2) syndicate data is bought from special suppliers and therefore have a monetary cost associated (Kelly, 1996, Kimball, 1996).

Table 1. The enhanced list of syndicate data incorporation problems (adopted from Strand et al., 2005).

Identification problems
<i>Id.1 – Identifying new entrants</i>
<i>Id.2 – Overlapping suppliers' capabilities</i>
<i>Id.3 – Overlapping data or products/services</i>
Acquisition problems
<i>Ac.1 – Acquiring incomplete data sets</i>
<i>Ac.2 – Varying data source stability</i>
<i>Ac.3 – The syndicate data is expensive</i>
Integration problems
<i>In.1 – Demanding to design and maintain transformation processes</i>
<i>In.2 – Diverging data representations and structures</i>
<i>In.3 – Assuring data consistency</i>
<i>In.4 – Missing data identifiers</i>
<i>In.5 – Diverging time-stamps</i>
<i>In.6 – Conflicting data from multiple sources</i>
<i>In.7 – Hiding data quality issues in commercial ETL-tools</i>
<i>In.8 – Varying source content</i>
Usage problems
<i>Us.1 – Misunderstanding the meaning of data</i>
<i>Us.2 – Missing metadata</i>
<i>Us.3 – Lacking routines for data quality assurance</i>
<i>Us.4 – Making decisions on outdated data</i>
<i>Us.5 – Trusting the data</i>
<i>Us.6 – Contradicting data from multiple sources</i>
<i>Us.7 – Ignoring syndicate data for DW purposes</i>
<i>Us.8 – Restricting laws and regulations</i>

3 Research Approach

The research approach of this work is to conduct an interview study towards SDSs. The interview study partly focused on contrasting the syndicate data incorporation problems in Table 1, towards the SDSs' perspectives, experiences, and work-routines.

For being able to contrast the answers from the suppliers, with these problems, the interview questions were constructed from an opposite perspective with respect to the problems. For example, one of the problems given in Table 1 is that organisations find it demanding to design and maintain transformation processes Problem – In.1. As a consequence, the question stated towards the SDSs was: *to what degree do you allow your customers to tailor the data so that it fits with their own needs?* (question translated from Swedish). In addition, follow-up questions were also stated, when the respondents answered the questions too generally or in an avoiding manner, or to broaden the material and acquire illustrative examples. Therefore the interviews may be considered as semi-structured (Williamson, 2002).

The interview study included 9 interviews covering 8 different SDSs. Due to the strong competition and strong relation between different companies, all respondents wanted to stay anonymous and they will only be denoted as Respondent 1-9. In addition, as all companies sell data for many different purposes, the respondents have answered the questions with respect to specific data or services/products targeted towards DW incorporation. Furthermore, Company B was covered by two respondents due to their broad coverage of different databases and services. In addition, in one of the interviews, two respondents were participating. Their answers were therefore separated in the transcript and they are denoted as 8A and 8B, respectively. The interviews were conducted via telephone and the interviews were taped and transcribed. The interview lasted in an average of 90 minutes and the transcripts of the interviews ranged from 3560 to 7334 words (5356 words in average). The transcripts were then returned to the respondents for validation, allowing them to, for example, correct misunderstandings or complement their answers. When the transcripts had been validated, they were included in the analysis.

4 Contrasting Problems Experienced by User Organizations

In this section, we contrast the problems in Table 1 with the viewpoints, experiences, and work-routines of the SDSs. Some of the problems were impossible to respond to by individual suppliers, e.g. *Problem 3.6 – Conflicting data from multiple sources*. (All questions and quotations are translated from Swedish). Instead, these problems have been contrasted with the authors summarized understand of the particular topic, based on the statements of all or some respondents.

4.1 Contrasting the identification problems

Id.1 – Identifying new entrants

All respondents claimed that the industry is under a strong competition, even though a few suppliers are benefiting from a monopoly situation. All respondents also stated that due to the strong competition, they exercise a strong supplier-push towards the consumers. Every supplier had internal resources or bought outsourced capacities for conducting sales initiatives towards prospects or established customers.

Id.2 – Overlapping suppliers' capabilities

The analysis of the suppliers' data, products or services shows that many suppliers are capable of delivering approximately the same data and services. As a consequence, a majority of the suppliers indicated that they naturally compete with data and services offered, but equally much with other, competitive means, like e.g. availability, degree of refinement, and support. Above all, a majority of the respondents claimed that the refinement aspect, i.e. the intelligent combination of the customers' internal data and novel, relevant syndicate data added by the suppliers, was considered as the main competitive edge for the suppliers.

Id.3 – Overlapping data or products/services

The analysis shows that most of the suppliers strived towards offering different standardized sets or packages of data, of which the user organizations must select one or several for procurement. Normally, the data sets or packages are overlapping and the different sets are arranged according to the amount of data contained. As a consequence, the consumers sometimes are forced to procure a more expensive and extensive data set, with respect to content, for being able to acquire a specific parameter or attribute. However, a majority of the suppliers also claimed that they allow a high degree of tailoring of the data, since the customers have such varying demands. To illustrate, Respondent 6 gave the following statement: *“We try to have standardized solutions, but it has been shown that our customers have very varying demands, so for being able to compete, flexibility is a key-word. Therefore, we try to build a generic platform that allows flexibility towards each and every customer”*.

4.2 Contrasting the acquisition problems

Ac. 1 – Acquiring incomplete data sets

The analysis shows that all respondents accounted for a high data quality awareness and claim that it is a prerequisite for being able to survive at the market, because *“without a high quality of the data, you are out of the market”* (Respondent 8b). Most suppliers also conducted tool-based automatic data verifications, e.g. verifying that every record identifier had a corresponding record or that the data sets are complete and a majority also conducted manual data verifications, e.g. contacting organizations if important data are missing.

Ac.2 – Varying data source stability

The analysis shows that the root of this problem is handled very differently by the suppliers. Some of the respondents acquired the internal data from their customer, refined it, and sent it back via FTP to the customers. With this approach, the problems should not appear, as the user organizations are not the active part going in and downloading data from e.g. a web-hotel or a FTP-mailbox. A majority of the suppliers also applied alternative distributed techniques, like CDs/DVDs or e-mail attachments, which also avoids this type of problem. However, a few organizations required the customer to be the active part and access FTP-mail-boxes or Web-hotels.

Ac.3 - The external data is expensive

The analysis shows that most suppliers claimed that the pricing of the external data is under a strong pressure and upon a question related to the competition of the market,

Respondent 5 states: *“it is very hard [...] and the price for the raw data is fast approaching the marginal cost”*. Consequently, and naturally, the suppliers and the user organizations have varying opinions on whether the syndicate data is expensive or not. More interestingly, the suppliers indicated two other reasons for why the consumers may consider syndicate data as expensive: 1) most respondents claimed that the ordering competence is very varying among the user organizations, which may result in data acquisitions that do not meet the actual needs or expectations, and 2) a majority of the suppliers claimed that the consumers do not exploit the full potential of the data they buy.

4.3 Contrasting the integration problems

Generally, one may claim that the suppliers have a very good understanding of the problems related to integration and some of the suppliers have taken another approach to the data transformation problem, by receiving the data from the user organizations and integrate and refine it on the supplier side, before returning it back to the user organization. A few respondents even claimed that they constantly kept mirrors of the users' data, which they were periodically updating, e.g. customer master dimensions. Still, a majority of the suppliers do apply the more traditional approach of data sales, i.e. the suppliers distribute the data to the consumers, which cater for the integration efforts, either as the only approach or upon the customers' choice. Therefore most issues or examples given further on are related to the traditional approach.

In.1 – Demanding to design and maintain transformation processes

An analysis shows that all suppliers allow for some degree of tailoring of the data, but one should also be aware that the degree of tailoring is vastly shifting. A majority of the suppliers claimed that they allow for a high degree of tailored acquisitions, but it should also be brought forward that most of these tailored acquisitions still are based upon standardized outtakes of data, which do delimit the customers' abilities to totally customize their acquisitions. Respondent 3 was even more outspoken and claimed that: *“The abilities to tailor the data are big, BUT, in practice we rarely do that because it is a question of how we should use our resources. It is always risky to tailor data acquisitions, as they create a maintenance need that becomes very costly for us”*.

In.2 – Diverging data representations and structures

Some of the respondents exemplified on internal codes for e.g. arranging companies according to industries or business, which diverged from the governmentally established code. They proposed structuring and analysis issues for introducing novel codes. A majority of the suppliers also promoted the XML as a key standard, as it allows for more flexible structuring and restructuring of the data. For example, Respondent 5 stated that *“XML in combination with the Internet is, for us that have been writing communication protocols, like a dream come true. It is a complete dream”* Furthermore, A few respondents indicated that they cooperate, or planned to collaborate, with software suppliers on special certificates. The underlying idea was that SDSs and software agree upon different representations of data and thereafter certifies these representations, meaning that a user organization following the

certificate, i.e. procure the software from the particular vendor and the data from the particular SDSs, do not have to transform the syndicate data being incorporated. Thereby, the user organizations drastically reduce the resource they have to spend on data transformations and integration.

In.3 – Assuring data consistency

This issue was very difficult for the suppliers to relate to, as they normally do not know the internal systems structure at the customer side. However, a few respondents briefly indicated that one may question if the consumers really have a full control of their internal systems or if they have designed their systems so that the syndicate data updates are reflected all-over.

In.4 – Missing data identifiers

Besides the general comment given in the introduction to this section, the analysis of the study gives no further details on why this problem occurs. Still, a majority of the suppliers have specifically pinpointed the importance of identifiers for their internal data quality verifications and storage.

In.5 – Diverging time-stamps

Since this is very much related to how the user organizations design their integration processes, the suppliers could not comment upon this. Still, in those cases where the non-traditional integration approach is applied, the suppliers should cater for a correct time-stamping.

In.6 – Conflicting data from multiple sources

Due to the reference to multiple suppliers or sources, the respondents could not contribute with any details to this problem. Still, one may imagine that organizations encounter this problem, if combining demographic data from one supplier and economic data from another, since some of these suppliers have overlapping data/services.

In.7 – Hiding data quality issues in commercial ETL-tools

Since this is very much related to how the consumers design their integration processes, the suppliers could not comment upon this. Still, in those cases where the non-traditional integration approach is applied, the SDSs should cater for a correct time-stamping.

In.8 – Varying source content

All respondents claimed that they are very careful with changing the data and most have contracts with the user organizations, regulating e.g. data content, data format, data structures. Still, based upon the answers of the respondents, one may derive that approximately 50 % of all DW customers acquire the syndicate data on-demand, instead of subscribing to it. Such customer may end up in problems, as they may have ETL-processes that have been out-dated due to the fact that the suppliers have changed the data formatting they allows or the data content they deliver.

4.4 Contrasting the usage problems

Us.1 – Misunderstanding the meaning of data

In order to prevent that the consumers misunderstand the meaning of the data, a majority of the suppliers worked together with the user organizations in projects, when starting up a business relationship. Since the incorporation of syndicate data into DWs is a rather extensive undertaking, the project form was considered as a necessity for being able to establish all the needs of the user organizations. To exemplify, Respondent 6 stated that *“most important in such a project is to establish a specification of the customers’ needs, since we may then use the specification as a validation instrument and show the customer how the data contributes and which needs it fulfil”*. However, some of the respondents gave the impression of working in more loosely coupled supplier-consumer constellations, where no detailed contacts are established between the two parties and in such cases one may assume that misunderstandings arise, especially since most respondents claimed the order competence of the consumers as strongly varying.

Us.2 – Missing metadata

All respondents claimed that metadata was included with the data that was distributed to the customers. In addition a few respondents claimed that they have a metadata service, which the users could use for verifying the meaning of the data incorporated.

Us.3 – Lacking routines for data quality assurance

As indicated previously, all respondents were shown to be very data quality aware and a majority of the suppliers spent a lot of resources on conducting automatic and manual data quality verifications. A few respondents even claimed that they hire external organizations, labelled data quality verifiers, to manually verify that the data is correct. For example, by phoning private persons and ask if their names are spelled correctly or by contacting companies and verify that a certain person still is the manager of a certain department or division.

Us.4 – Making decisions on outdated data

This problem transcends from the data sources of the SDSs and some of the respondents claimed that their sources are sometimes rather slow on generating procurable data. For example, it may take one and a half year before a SDS gets the annual accounts of an organizations or it may take more than a month before an established organization is registered in the SDSs’ databases and possible for user organizations to get informed of. Respondent 8 further exemplified that: *“it happens that our customers inform us that they have found an organization or a part of an organization that we do not have in other systems”*. Still, a majority of the suppliers have not acknowledged this as a problem and that it mostly comes down to the routines of the user organization on whether they make decisions on out-dated data or not.

Us.5 – Trusting the data

The suppliers had problems to give any particular details to this, but they indicated on an average that 50% of their customers were subscribing the data and that the customers acquiring on-demand tended to be rather faithful, which could be seen as an indication that the consumers trust their suppliers. Naturally, the selection of the respondents effects the response regarding this problem, since the suppliers participating were established and with a good reputation. Still, most of the respondents claimed that they were actively working for having a constructive

relationship with their customers and a majority of the respondents exemplified on novel data or products/services that were a result of long business relations.

Us.6 – Contradicting data from multiple sources

Due to the reference to multiple suppliers or sources, the respondents could not contribute with any details to this problem. Still, one may imagine that organizations encounter this problem, if combining demographic data from one supplier and economic data from another, since some of these suppliers have overlapping data/services.

Us.7 – Ignoring external data for DW purposes

A few respondents indicated that many organizations have problems bridging different parts of the organizations as the systems are stow-piped, making the syndicate data updates isolated or fragmented. Respondent 1 gave the following illustrative example: “An organization may have 7-8 registries that do not interact, so there are many integration benefits that may be achieved. For example, an organization may have a customer registry, a supplier registry, a prospect registry, and a competitor registry, which are not integrated in any way. If merging these together, into e.g. a DW, one may find that a supplier is also a customer and a supplier may also be a prospect in other occasions. There I think it is a lot that needs to be done”. A majority of the respondents also claimed that incorporating syndicate data into DWs is a rather novel concept, compared to acquiring syndicate data for more operative purposes, e.g. verifying the customer address of a single customer or acquiring the solidity of a particular organization. Thereby, many user organizations are rather immature when it comes to incorporating syndicate data in DWs and therefore do not have the experience for being able to fully exploit the potential thereof. However, most of the respondents claimed that the syndicate data incorporation will inevitably steer into a DW direction, as the combination of DW technology and syndicated data caters for novel and powerful abilities to perform various types of business analysis. Furthermore, a majority of the respondents also claimed that they try to influence organizations to make better use of the syndicate data, by indicating the possibilities of incorporating into DWs, if the user organizations were not already exploiting that possibility.

Us.8 – Restricting laws and regulations

A majority of the respondents informed the consumers on what they may legally do with the data, whereas some of the suppliers clearly stated that it is up to the user organizations to stay updated on law and regulations. In addition, all respondents established contracts with the consumers, in which they, amongst other things, regulated for what purposes the syndicate data may be applied. For example, all suppliers stated in their contracts that the data they sell may not be applied for establishing competing business, i.e. the consumers are not allowed to sell the data in their turn. The laws and regulations, combined with the contracts established by the suppliers, may also be hinderers for the user organizations to fully exploit the potential of the data acquired (see problem Ac.3), as the consumers may not dare to apply the data for other purposes than those regulated in the contracts, since the regulations and laws are very complex and requires a lot of resources to stay updated with.

5 Results

The analysis of the problems gives a lot of details related to the suppliers' viewpoints, experiences, and work-routines related to their part of the problem and thereby contextualize the problems from a SDS perspective. In addition, the analysis also reveals some collaborations and tensions on all three interoperability layers of the organizations, which facilitate or hinder interoperability between the user organizations and the SDSs. These collaborations and tensions are given below:

Collaborations:

- The user organizations and SDSs collaborate in projects, when incorporating syndicate data into DWs, as these initiatives are rather demanding undertakings, stating technological challenges, as well as organizational challenges.
- The consumers and data suppliers are collaborating on developing novel services, based upon the demands of the user organizations or based upon the competencies of the suppliers, knowing what types of data or services the organizations may benefit from.
- The suppliers are very data quality aware and collaborate with the consumers on identifying missing data.
- The suppliers know that the user organizations have problems with their data quality verifications and therefore have a strong focus on assisting the consumers with these issues.
- The user organizations seem to be rather pleased with their suppliers and satisfied with their services, as approximately half of the consumers are subscribing on syndicate data for longer time-periods and the suppliers consider their customers as rather faithful.
- The data suppliers collaborate with the user organizations in order to assist them in exploiting the potential of the syndicate data as efficiently as possible.

Tensions:

- The SDSs strive towards standardized data sets or packages, since it facilitates the maintenance thereof and is cost efficient, as tailored solutions become expensive for the SDSs to maintain, especially if every single customer would like to tailor their acquisitions of data. Unfortunately, the standardization sometimes forces the user organizations to procure more expensive and extensive standardized data sets, than they actually need, for being able to acquire a specific parameter or attribute.
- The suppliers are conducting a strong marketing and sales push towards the consumers for buying novel data, despite the fact that they know that the user organizations do not exploit the potential of syndicate data already being incorporated.
- The suppliers regulates the consumers with contracts, which may hinder the user organizations to fully exploit the potential of the data acquired, as they may not dare to apply the data for other purposes than those exactly stated in the contracts, since the contracts are very complex and requires a lot of resources to stay updated with.

- The consumers and the suppliers have very diverging opinions on whether the syndicate data is expensive or not and pricing issues may naturally cause tensions between the user organizations and the supplier.
- The suppliers state that the ordering competency is very varying among the user organizations, but still some of them outsource their marketing and sales initiatives. Thereby, it becomes even harder for the consumers to actually acquire the most appropriate data, since this intermediate actors distances the user organizations from suppliers.
- User organizations spend a lot of internal resources on transforming syndicate data and pays for non-relevant data, as the suppliers strive towards standardized data sets.

It is interesting to reflect upon the tensions and collaborations as some of them contradict each other, e.g. on one hand the suppliers conduct a strong push for selling novel data and at the same time they assist the consumers to fully exploit the potential in already acquired data. In addition, five out of six tensions are relatable to the business and knowledge layers. Together these indicate that the two parties are rather immature in their interoperability on these layers, whereas on the ICT-systems layer, covered by the acquisition and identification problems, only one tension were identified, and it was a consequence of incorporating data according to the traditional approach, instead of letting the suppliers handle the integration and refinement. As another example for a higher maturity on ICT-systems layer, the mirroring of the customers data could be given, showing that the suppliers and the user organizations interoperate to produce the best possible result.

6 Summary and Future Work

As indicated in the introduction, a SDS contextualization of the syndicate data incorporation problems experienced by the consumers served several purposes. Firstly, the contextualization allowed for evaluating the maturity of the interoperability between the SDSs and the user organizations. As shown in the results, the more organizational oriented aspects need more attention, as most of the tensions were of such a kind. Therefore, it is reasonable to conduct further research aimed at improving the interoperability between them. Consequently, one should focus on the tensions and try to solve the causes generating these tensions, by e.g. trying to develop models for calculating and making the pricing of the syndicate data for DW incorporation more obvious. Secondly, the contextualization gave general tensions and collaborations between the two parties. Future work could be devoted to deepened and detail the descriptions of and causes for these tensions. Finally, the contextualization contributed with important details for a hands-on support for organizations incorporating syndicate data into their DWs (as suggested by Strand et al., 2005), by including details on e.g. how they establish business relationships with customers, to what degree they allow the organizations to tailor the data, and how they respond to suggestions for novel services. The development of such a hands-on support is now being undertaken and will be presented in the shape of detailed

guidelines or heuristics, accompanied with descriptions of the problem context and possible solutions.

References

- Chen, D. and Doumeingts (2003) European initiatives to develop interoperability of enterprise applications – basic concepts, frameworks and roadmap. *Annual Reviews of Control* (27) pp. 153-162.
- Damato, G. M. (1999) *Strategic information from external sources: a broader picture of business reality for the data warehouse*. Available at Internet: http://www.dwway.com/file/20020726170552_get_ext_data.pdp [Accessed 03.02.20]
- Hammer, K. (1997) "Migrating data from legacy systems", in *Building, using, and managing the data warehouse*, in Ramon Barquin & Herb Edelstein (Eds), New Jersey: Prentice Hall PTR, pp. 27-40.
- Hessinger, P. (1997) "A renaissance for information technology" in *Data warehouse practical advice from the experts*, Joyce Bischoff and Ted Alexander (Eds), New Jersey: Prentice Hall PTR, pp. 16-29.
- Inmon, W. H. (1996) *Building the data warehouse, 2nd edition*. New York: John Wiley & Sons.
- Kelly, S. (1996) *Data warehousing: the route to mass customization*. New York: John Wiley & Sons.
- Kimball, R. (1996) *The Data Warehouse Toolkit*. New York: John Wiley & Sons.
- Oglesby, W. E. (1999) *Using external data sources and warehouses to enhance your direct marketing effort*. DM Review. Available at Internet: http://www.dmreview.com/editorial/dmreview/print_action.cfm?EdID=1743 [Accessed 03.02.21].
- Strand, M., Wangler, B. & Olsson, M. (2003) Incorporating external data into data warehouses: characterizing and categorizing suppliers and types of external data. In *Proceedings of the Americas Conference on Information Systems (AMCIS'03)*, 4-6 August, 2003, Tampa, Florida, USA, pp-2460-2468.
- Strand, M. & Wangler, B. (2004) Incorporating external data into data warehouses - problems identified and contextualized. Presented at the *7th International conference on information fusion (Fusion'04)*, June 28- July 1, Stockholm, Sweden.
- Strand (2005) Syndicate data incorporation in data warehouses: a categorization and verification of problems. Submitted to 13th European Conference on Information Systems (ECIS2005), 26-28 May, Regensburg, Germany.
- Williamson, K. (2002) *Research methods for students, academics and professionals*. 2nd Edition, Thousand Oaks: Sage Publications, Inc.