Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2009 Proceedings

Americas Conference on Information Systems (AMCIS)

2009

Hitting Interface Specifications - The Role of "Common Ground" in

Wolf Behrmann
zeb/information.technology, webhrmann@web.de

HITTING INTERFACE SPECIFICATIONS - THE ROLE OF "COMMON GROUND" IN FINANCIAL DWH PROJECTS

Wolf Behrmann
zeb/information.technology,
Frankfurt am Main, Germany
wbehrmann@web.de

ABSTRACT

Data warehousing (DWH) projects are among the most expensive projects of companies in the financial industry. Problems like budget overruns and low end-user acceptance lead to the assumption that despite intensive research over the last two decades the relevant elements determining the success of such projects are not sufficiently investigated. On the basis of a case study about the implementation of a group-wide DWH the author analysis the existence of those still inadequately considered success factors. One challenging hurdle that project members have to clear is the complete fulfillment of interface specifications elaborated during conceptual modeling phase. The author chooses an interdisciplinary approach that involves established language theories and applies them to real-life observations in the case study to identify the necessary prerequisites. Subsequently rooms for improvement in the project approach are derived to guarantee a successful knowledge transfer among project members. Especially personal communication is essential to ensure the project's success and to increase efficiency in DWH projects.

Keywords

Case Study, Data Warehousing, Interpretivist Research, Knowledge Transfer.

INTRODUCTION

Over the last twenty years many theses dealing with the management and success factors of data warehouse (DWH) projects were published. In practice DWH projects are very challenging due to still existing problems like minor acceptance of the solution by end users, although heavy efforts have spent on the development of modeling languages, process models and on success factor analysis. Today such data consolidation projects are among the most important and expensive projects for companies in the financial industry. One reason for the necessity of such projects is the ongoing expansion of European banks to Central Eastern European countries where banks are confronted with the integration of new subsidiaries into the group. This data consolidation process is essential to enable a group wide controlling. Another reason is the improvement of regulatory and internal reporting methods. Thus, over time a growing quantity of information has to be collected to fulfill these extended requirements. During the last years the *revised international capital framework for Banks* called Basel II (Bank for International Settlement, 2006) was one of the main drivers of large financial data warehousing (FDWH) projects. As a consequence of the current financial crises it is expected that regulatory requirements will be further enhanced.

In this paper the author describes the challenges of FDWH projects on the basis of a group-wide DWH project effected by an international banking institute. These observations summarize a project history of more than five years. „The route to good theory leads not through gaps in the literature but through an engagement with problems in the world that you find personally interesting." (Kilduff, 2006). Besides that approach the paper illustrates how rigor can be brought into accordance with relevance by using approved and broadly accepted methods of case study research combined with well known theories of other disciplines. This interdisciplinary approach allows to explain real-life observations, to identify room for improvement in DWH projects and follows the ideas of Galliers & Land (1987, 1985) and Mingers (2003).

The observations indicated that one of the main influencing factors on the success and performance of IS projects is the efficiency in communication. Therefore the DWH project described in our case will be extended by a language based analysis which importance has been emphasized among others by Lyytinen (1985). The theoretical base for this analysis originates mainly Clark (1996) which has been applied to expert-layperson communication scenarios in the area of knowledge

management by Bromme et al. (2004). They dealt with specific problems occurring in the knowledge transfer in case of high knowledge differences. The same situation can be observed during several steps in the development of information systems (IS). Especially in case of unstructured problems of high semantical complexity like FDWH development many communication defects can be observed (Behrmann and Räkers, 2008).

In this paper the communication defects observed in the case are analyzed. The observations in the case are interpreted by applying well-known concepts of sociology. Further, it is explained which prerequisites are essential for a successful communication in FDWH projects.

RELATED WORK

Since the 1970s comprehensive research has been carried out in the field of data warehousing. While most publications concentrate either on conceptual modeling or on a rather model-driven and technology-oriented implementation of DWHs, only relatively few authors attempt to unify both perspectives. A detailed overview about this kind of DWH literature can be found at Behrmann & Räkers (2008).

As a significant number of DWH projects have failed or have required additional funds, extensive research started to concentrate on quantitative or qualitative analyses of success factors as well as contemporary best-practices within the field of DWHs. In this context, researchers focused mainly on organizational as well as methodological aspects. None of the before mentioned approaches explicitly addresses personal interactions within project teams. Hitherto, only a few publications deal with soft factors in DWH or ISD projects e.g. in the field of knowledge management (e.g. He, 2004), learning processes (e.g. Pirinen and Pekkola, 2006) and communication (e.g. Gallivan and Keil, 2003). In all phases of the DWH development process, except for the requirements engineering phase, existing research concerning DWH process models largely disregards the specific problem of interpersonal communication. The growing importance of personal interaction among project team members is also reflected in the current work of Inmon et al. (2008), although originally Inmon has more engaged in technical oriented approaches (Inmon, 1992).

This paper makes a contribution to close this gap in DWH literature. The basic ideas of language theory are applied to the specific problems described in the case study. The analysis of the case study is mainly based on the publication of Clark (1996). Bromme et al.'s (2005) enhancements of Clark's theory are used to explain the communication defects observed in the case study. While Behrmann & Hoffmann (2008) focus on single communication actions, this paper deals with the basic principles of Clark's language theory.

INTRODUCTION TO LANGUAGE THEROY

In order to facilitate the understanding of the observed problems explained in the case some basics of the used language theory are described in the following. Clark's work focuses on the use of language and communication acts which are considered as joint activities. A communication act consists of two persons, a sender and a receiver, which activities need to be coordinated in order for the communication act to be successful. "In language use, a central problem is coordinating what speakers mean and what their addresses understand them to mean". (Clark, 1996 p 73). The basis for coordination is shared knowledge between the two actors; they share a common ground (CG). "The shared basis (piece of CG) is the key to the coordination problem" (p 99). The kind of an utterance totally depends on the assumptions of the CG made by the sender i.e. he has to assume what already is part of their common knowledge. The sender has the intention to extend the CG step by step by each utterance. Clark defines this process as grounding. If the volume of the new information to the CG gets too big, misunderstandings will occur and CG might be difficult to find. A minor misunderstanding at the beginning might snowball into major ones in the end (p 235). The same problem appears if the assumptions about the CG are wrong which is often caused by an egocentric bias. "Our feeling of other's knowledge does, in fact, have a strong egocentric bias: If I know something, I am more likely to expect others to know it too" (p 111).

Grounding is important whenever people do things together. "To ground a thing, (…) is to establish it as part of CG well enough for current purpose" (p 221). Three principles are essential for the grounding process. The first principle of closure means that "agents performing an action require evidence, sufficient for current purpose, that they have succeeded in performing it" (p 222). In case of joint actions like communication acts the second one, the principle of joint closure, is equally important: "The participants in a joint action try to establish the mutual belief that they have succeeded well enough for current purpose" (p 226). The last one is the principle of least effort: All things being equal, agents try to minimize their

effort in doing what they do to intend (p 224). In other words for the grounding process people should look for the most powerful evidence that is valid, cheap, and timely enough for current purpose (p 225).

The evidence for successful communication can be divided into four layers. The results are action ladders, a hierarchy for communication acts with upward completion and downward evidence (p 147). Upward completion means a next level can only be reached if the former level is completed successfully. Downward evidence implies that the evidence that one level is complete is also evidence that all levels below are complete. The four levels of communications are shown in figure 1.

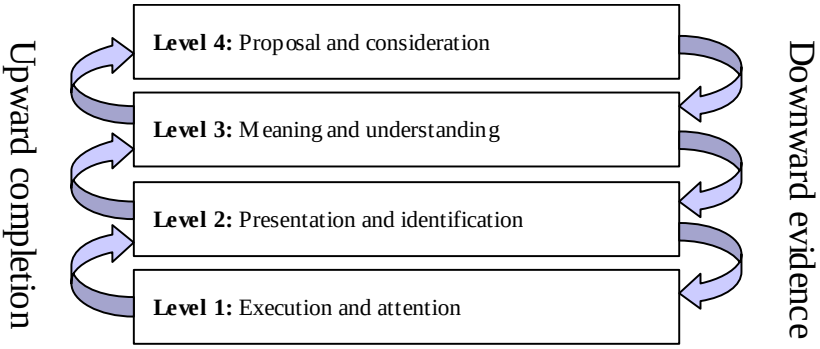

Figure 1. Action ladder in communication acts

Communication at level 1 is successful if the sender executes a behavior and the receiver attends to it. At second level the presentation of a signal has to be identified as such. The correct understanding of a signal on the receiver side as meant by the sender is located at level three. The last level contains the sender's proposal and the receiver's consideration. The receiver can accept the proposal or e. g. ignore it. The grounding process takes place in two phases: presentation and acceptance (p 227). Evidence in face-to-face communication can be achieved by embedded questions and by later corrections after detecting the misconstrual (based on the signals of the receiver) (p 234). During face-to-face communication the receiver expresses his (mis-)understanding by short utterances like "yes" or complete sentences like further questions. These utterances enable a permanent repair of communication defects immediately after detection by sender or receiver. Some communication defects are not detected before having reached level four and the reaction of the receiver does not fit to the sender's expectations.

Clark's theory has been applied by Bromme to two special scenarios: communication in case of high differences in knowledge between sender and receiver, and in case of communication without face-to-face contact (Bromme and Jucks, 2001, Bromme et al., 2004). Expert-layperson communication is characterized by a low CG between the actors at the beginning of the communication process in which the CG will not only be accumulated but also restructured. For a successful knowledge transfer a change in perspective is necessary i.e. the expert must assume the knowledge of the layperson. This is difficult because there is a systematic difference between the perspectives of both. In this context the term systematic means that not only knowledge elements in the layperson's perspective are missing but they are also embedded in a cognitive reference framework (CRF). These CRFs are mainly determined by the participant's disciplines and their specific education. CRFs of laypersons are partly resistant against changes. Utterances of the expert will be embedded in an inaccurate context without stimulating adaptations of the CRF. This may cause an "illusion of evidence". In this situation the expert overestimates the understandability of transferred facts (Bromme and Jucks, 2001). One of the implications on expert-layperson communication is the use of feedback-loops to check if the knowledge transfer has been successful. Additionally the supporting systems and documents must have a structure which is logical and understandable from layperson's and expert's perspective. Both findings implicate the importance of strong interaction within groups and face-to-face communication (Bromme et al., 2004).

The impact of the medium used to communicate was analyzed in another publication of Bromme & Jucks (2001). In case of written communication direct feedback based on gestures and verbal intervention is not applicable. Therefore more effort for the anticipation of the layperson perspective is necessary. Caused by the higher effort for feedback in written form the receiver often gives no response. Due to this fact the probability of misunderstanding and illusion of evidence increases rapidly. So the theoretically existent possibility to signal a missing understanding will often not be used in practice (Bromme and Jucks, 2001).

CASE STUDY

Research Methodology

The empirical part is an exploratory case study which has been conducted according to Yin's (2003) approach. His methodology is suitable to analyze the complex organizational phenomena of the case study. The purpose is to understand the interaction between members of a large DWH project with a special focus on requirements engineering and development issues. This requires posing of "how and why" questions about a contemporary set of events over which the investigator has little control (Yin, 2003). The relevance of the research questions is ensured by focusing on real-life phenomena. The application of the broadly accepted method of Yin guarantees the rigor of the results as well. By using several sources of evidence and fulfilling Yin's data collection rules the evidence of the case study is guaranteed.

The case study contains observations of more than ten sub-projects where the same data mapping problem has been solved. According to Yin the case study can therefore be classified as a single case with embedded multiple units of analysis. The case study contains a high number of different comparable situations to support the transferability of the findings to other FDWH projects. This intended generalizability of theories developed in case studies has been discussed among others by Lee & Baskerville (2003).

The following case was developed by a mixed research team. Two members are management consultants at zeb/information.technology of whom one was directly involved in the project described. The observed facts were documented and reconciled among other project members of the banking group and zeb/. Furthermore the project documentation (minutes of meetings, presentations of lessons learned sessions, data models, calendars, etc.) was used to confirm the presented facts. Additionally unstructured information and non-formalized documentation like email traffic was evaluated. To reduce personal bias and to increase the objectivity of the case study, all findings were discussed among the extended research team including non-involved persons.

Roll out of interface specification

The following observations have been made during a Basel II project at an international banking group (BG) with total assets of more than € 100 billion. The project goal is fulfillment of the requirements specified by Basel II. One of the main topics of Basel II is the calculation of *Risk Weighted Assets* (RWA) based on single transaction data for the consolidated BG which requires strong changes in the IT infrastructure. The BG consists of a head office (HO) and 15 major subsidiaries. The Basel II project started in 2001 with a small number of project members. After the initiation of the main activities in 2004 more than 100 persons were temporarily involved.

To fulfill the requirements a central DWH has been implemented to serve as the single source of data for the Basel II calculations. Thus, the DWH has to contain all relevant raw data required for the calculation of the BG's RWAs. Therefore, data of the HO and all subsidiaries has to be loaded into the central DWH. Each subsidiary has to develop procedures (so-called "extraction jobs") to extract data from their local systems and to transform it to flat files that meet the interface specification for transfer to the central DWH. The data supply chain has been implemented in several releases.

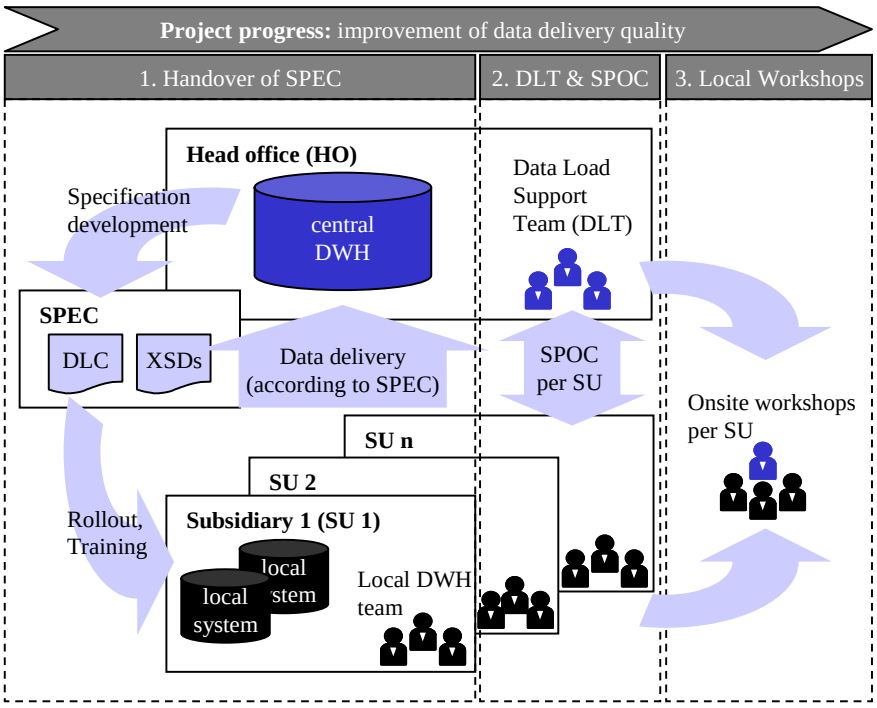

Figure 2. Case study – illustration of architecture and organisation.

Following technical implementation of the central DWH an interface specification (SPEC) based on the physical DWH structure was developed in HO. To document all relevant interface information and to enable machine readability i.e. for data quality the SPEC was described as an XML schema file (XSD). To check XSD conformity of interface flat files a toolset was developed by HO. By using the described XML technique an attribute-oriented description was covered. Further definitions for a sufficient SPEC i.e. links between tables were needed to enable a correct data load into the DWH. To close this gap a data load concept (DLC) was written and became part of SPEC. The SPEC was sent to the subsidiaries to create extract jobs for their local databases and to deliver flat files meeting the XSD and the textual specification. The subsidiaries' feedback brought problems in understanding the SPEC to the surface. Therefore a workshop about the principles of the SPEC was arranged for all subsidiaries. Furthermore some examples for data records were discussed. After the workshop the subsidiaries began to implement their interfaces. HO stayed in permanent contact with the subsidiaries to answer their questions (see figure 2, step 1).

Analysis of the data initially delivered showed many violations against the XSD and the textual part of the SPEC. The late detection of those violations during the test phase led to critical time pressure and to an escalation of the project. HO team assumed three main reasons for this deficient implementation. The first one is a low interaction between the local subsidiary project teams and the HO DWH team. The other reasons are a minor understanding of the DWH solution in subsidiaries and knowledge differences among the HO DWH team itself. Therefore a dedicated data load support team (DLT) was nominated by HO. The DLT's task was to support the subsidiaries intensively and to streamline the communication. For each subsidiary a contact person and a deputy were nominated as single point of contact (SPOC) to get deep insight into local specifics. Since then communication was streamlined and based on email and telephone contact. Regular visits of the subsidiaries by the HO team were not planned due to high traveling costs (figure 2, step 2).

Unfortunately some subsidiaries still made insufficient progress and were not able to deliver the data according to SPEC. That counts for attribute based violations as well as for structural problems, i.e. incorrect references across data records etc. After another escalation an onsite support was installed (figure 2, step 3) and the contact persons of the DLT visited the subsidiaries regularly. They discussed open problems in workshops and tracked progress. To guarantee the availability of all required skills DLT was extended with business and IT experts familiar with the central DWH. The implementation of close collaboration between the DLT and the subsidiaries lead to a rapid increase in the quality of data delivered.

A discussion within the HO team regarding the advantages and disadvantages of the chosen approach lead to the following results:

- The usage of email as the main way of communication caused many misunderstandings and mail "ping-pong".
- The clarification of a specific topic was easier by phone especially when it was supported by visualization such as workshop presentations. The communication on the phone was very straight forward and focused on a specific problem.
- Before entering into the tests phase only little feedback was given by the subsidiaries. The late detection of misunderstandings on the subsidiaries' side was the reason for high fixing efforts and a critical delay in the implementation.
- During on-site visits several experts of the subsidiaries joint the meetings. A clarification of the problems was mostly possible with low effort. In addition the meetings enabled the HO team to get local insights and to learn the specifics of the subsidiary. The discussions were broader and more intensive than via email or phone. Many new issues were raised but could be clarified in the workshops directly.
- The areas of the SPEC that caused misunderstandings were not uniform between subsidiaries. A systematic, structural or common problem in understanding the SPEC could not be identified.

DISCUSSION

The case starts after the SPEC was fixed by HO i.e. the conceptual modeling phase was finished. Subsidiaries were not involved in the modeling process and just had to fulfill the SPEC. In order to succeed two facts are essential: first, the engineer must understand the requirements of the customer completely and correctly, so the customer and the engineer have a common understanding of how to interpret the SPEC. Second, there must be sufficient information in the SPEC for a complete implementation. That means: All relevant information for the problem solution is expatiated in the conceptual model.

Hence, the first challenge during the implementation process is a successful communication about the meaning of the SPEC, i.e. to transfer the SPEC's inherent knowledge from HO team to the subsidiaries. This problem is an expert-layperson communication scenario. The first approach, transferring knowledge by exclusive usage of written documents, was not successful because the understandability of the SPEC was over-estimated by HO. According to Clark and Bromme a common reason for this observation is a wrong assumption about the CG and the skills of the addressees. The high degree of observed knowledge differences had been surprising because the project members of subsidiaries had experiences in banking business as well as in DWH technology. Due to the late detection the effect of the misunderstandings on the project timetable was heavy. Before the test phase began the HO team assumed that the SPEC was correctly understood by the subsidiaries. This is an example for an illusion of evidence and shows that feedback about the understandability is poor in case of written communication. By using written communication it can be assumed that levels one and two of the communication action ladder are successful completed. On the third level of the action ladder – the correct understanding of the meaning – the observed problems were located.

The second challenge is the development of a "complete" SPEC that contains all implementation-relevant information. The correct anticipation of the CG is crucial to be able to decide which elements have to be added and what can be assumed as known. This requires that the expert puts himself into the layperson's perspective. The high number of upcoming questions on the subsidiary's-side showed that the anticipation of the CG was insufficient and the local knowledge had been overestimated for business as well as for DWH topics. This problem was also observed by Bromme. Especially the degree of familiarity with terms on the threshold between common and domain language was overestimated by experts (Bromme and Jucks, 2001). In the case this effect was enforced by the fact that basic business and DWH terms were assumed as common knowledge. The extent of the two observed problems varied from subsidiary to subsidiary. This can be explained by different skills and knowledge of the subsidiary's project teams. The heterogeneous CG of the subsidiaries complicated the anticipation of addressees' knowledge. The detection such lacks was followed by an active improvement of the SPEC.

After having extended the SPEC a mixture of several documentation types like formal XSDs, training presentation, mapping examples, etc. was available. All documents were presented during the SPEC training. The level of interaction between DLT and subsidiaries was increased by establishing direct telephone and email communication. The installation of a SPOC allowed a better anticipation of local specifics. The approach allowed the correction of concrete misunderstandings, i.e. a CG could be extended by few elements which were topic of the communication. The communication was characterized by a problem-orientated and straightforward style. In email as well as in telephone communication only one local expert was involved at the same time. The clarification of upcoming issues was not sufficient to ensure the project's success because the quality of the implementation remains low in some areas. A possible reason for this observation is the existence of CRFs. A

common understanding about the topics seems to be achieved by discussing specific issues and questions but they may not be embedded in the holistic context correctly.

The remaining implementation problems in the subsidiaries brought the project management to arrange regular onsite visits despite additional travelling and staff costs. This strong personal interaction in local workshops led to a significant and immediate reduction of open issues because several interdependent topics were tackled in contrast to limited telephone and email communication. The participation of local experts with different skills led to a common understanding quickly as many communication defects could be identified and corrected simultaneously and the "teachers" of the DLT could easily anticipate the knowledge of local teams. To sum it the face-to-face discussion enabled to detect and to solve illusion of evidence situations. In this context the face-to-face communication can be considered as a proactive action because it prevents future implementation problems.

CONCLUSION & OUTLOOK

The explanations of the author underline that effective communication and knowledge transfer is a key success factor of FDWH projects. Based on language theory possible reasons for the problems observed in the case study were identified. The following conclusions show that face-to-face communication and strong interaction play a decisive role in all phases of complex FDWH projects and cannot be fully replaced by a conceptual model:

- *Anticipation of addressee's knowledge is crucial:* The anticipation is essential for the design of understandable specifications. Therefore the sender has to put himself in the addressee's position which is difficult in case of high knowledge differences. This problem increases in case of heterogeneous groups with varying knowledge: The risk to overestimate the addressee's knowledge in FDWH projects is high, especially if project members have apparently the same educational background.
- *Face-to-face communication allows early corrections:* In case of written communication little feedback is given by addressees that easily leads to an illusion of evidence. This late detection and correction of such communication defects cause higher efforts and longer projects. In contrast direct feedback loops in face-to-face communication allow early detection of communication defects.
- *Without face-to-face communication an FDWH project cannot succeed:* A pure specification based knowledge transfer without face-to-face communication is not sufficient in case of FDWH-projects. Both premises, completeness and understandability, could not be met although the specification were continuously improved.
- *Face-to-face communication is proactive:* Face-to-face communication allows a proactive strategy meaning that broad discussions and considerations of CRFs allow effective knowledge transfer and the avoidance of future problems.

These findings are the basis for optimization strategies that will be part of further research. In the following three possible strategies will be described:

- *Improve anticipation of addressee's CG:* A better anticipation of the addressee's CG allows a receiver-oriented documentation and avoids misunderstandings. This can be achieved by establishing personal discussions at an early stage of the project.
- *Improve understandability of documents:* It can be assumed that the observed problems occur independently of the specification design. Understandability can be improved to a certain degree by enhancing the methodology whereas a complete coverage of the semantic cannot be achieved.
- *Support in the detection of communication defects:* A late detection of communication defects and the resulting late correction of errors in implementations lead to disproportionate high fixing costs (Boehm, 1981). The easiest way to detect communication defects is to create an environment which allows direct feedback loops.

After having understood the opportunities and constraints of each strategy the economic aspects have to be taken into account. The early detection of defects and the understanding of the addressee's CG require personal communication. Due to high staff and travel costs this kind of communication is the most expensive one. In contrast a documentation based interaction is cheaper especially in multi receiver scenarios but carries the risk of misunderstandings and high fixing costs. Finally a mixture of personal communication and document based interaction has to be found considering cost-benefit aspects to ensure the success of FDWH projects.

REFERENCES

1. Bank for International Settlement (2006) "Basel II: Revised international capital framework," Bank for Interantional Settlement, http://www.bis.org/publ/bcbsca.htm (13.04.2009).
2. Behrmann, W. and Hoffmann, A. (2008) Classification of Communication Defects in DWH Projects, in *Proceedings of the 14th Americas Conference on Information Systems*, August 14-17, Toronto, Canada.
3. Behrmann, W. and Räkers, M. (2008) Specifics of Financial Data Warehousing and Implications for Management of Complex ISD Projects, in *Proceedings of the 16th European Conference on Information Systems*, June 9-11, Galway, Ireland.
4. Boehm, B. W. (1981) Software Engineering Economics, Englewood Cliffs, NJ: Prentice-Hall.
5. Bromme, R. and Jucks, R. (2001) Wissensdivergenz und Kommunikation: Lernen zwischen Experten und Laien im Netz, *Partizipation und Interaktion im virtuellen Seminar*, 81-103.
6. Bromme, R., Jucks, R., and Rambow, R. (2004) Experten-Laien-Kommunikation im Wissensmanagement, *Der Mensch im Wissensmanagement: Psychologische Konzepte zum besseren Verständnis und Umgang mit Wissen*, 176-188.
7. Bromme, R., Jucks, R., and Runde, A. (2005) Barriers and biases in computer-mediated expert-layperson-communication, *Barriers, biases and opportunities of communication and cooperation with computers-and how they may be overcome*, 89-118.
8. Clark, H. H. (1996) Using Language, Cambridge University Press, Cambridge.
9. Galliers, R. D. and Land, F. F. (1987) Viewpoint: choosing appropriate information systems research methodologies, *Communications of the ACM*, 30, 11, 901-902.
10. Gallivan, M. J. and Keil, M. (2003) The user-developer communication process: a critical case study, *Information Systems Journal*, 13, 1, 37-68.
11. He, J. (2004) Knowledge impacts of user participation: a cognitive perspective, *Proceedings of the 2004 SIGMIS conference on Computer personnel research: Careers, culture, and ethics in a networked environment*, 1-7.
12. Inmon, W. H. (1992) Building the data warehouse, 1. edition, QED Technical Pub. Group, Boston.
13. Inmon, W. H., Strauss, D., and Neushloss, G. (2008) DW 2.0 - The Architecture for the Next Generation of Data Warehouse, Burlington.
14. Kilduff, M. (2006) EDITOR'S COMMENTS: PUBLISHING THEORY, *Academy of Management Review*, 31, 2, 252-255.
15. Land, F. (1985) Is an Information Theory Enough?, *The Computer Journal*, 28, 3, 211-215.
16. Lee, A. S. and Baskerville, R. L. (2003) Generalizing generalizability in information systems research, *Information Systems Research*, 14, 3, 221-243.
17. Lyytinen, K. (1985) Implications of theories of language for information systems, *MIS Quarterly*, 9, 1, 61-74.
18. Mingers, J. (2003) The paucity of multimethod research: a review of the information systems literature, *Information Systems Journal*, 13, 3, 233-249.
19. Pirinen, A. and Pekkola, S. (2006) Different Perspectives on Learning in Information System Development, University of Jyväskylä.
20. Yin, R. K. (2003) Case Study Research: design and methods, Sage Publications Inc, Beverly Hills.