Association for Information Systems

AIS Electronic Library (AISeL)

ECIS 2008 Proceedings

European Conference on Information Systems (ECIS)

2008

Specifics of Financial Data Warehousing and Implications for Management of Complex ISD Projects

Wolf Behrmann

Johann Wolfgang Goethe Universitat Frankfurt am Main, wbehrmann@web.de

Marc Rakers

zeb/information.technology, marc@raekers.com

SPECIFICS OF FINANCIAL DATA WAREHOUSING AND IMPLICATIONS FOR MANAGEMENT OF COMPLEX ISD PROJECTS

Behrmann, Wolf, zeb/information.technology, Staufenstr. 34, 60323 Frankfurt, Germany, wbehrmann@web.de

Räkers, Marc, zeb/information.technology, Schlossstr. 22, 48455 Bad Bentheim, Germany, marc@raekers.com

Abstract

Data warehouses play important roles in the IT landscape of the financial industry. In the last years most of the leading banks implemented data warehouse solutions to fulfil regulatory or internal requirements. The institutes have to handle situations, which especially concern data warehouse projects in financial industry. On the basis of three case studies, observing DHW projects for a duration of more than three years, we point out specifics of financial data warehousing and discuss implications to the management of complex information systems development projects. Ensuring a common understanding about the DWH between the involved parties is a major challenge during requirements engineering, data warehouse design, DWH implementation and creation of the ETL processes. In five statements we outline specifics of financial data warehousing and recommend approaches for further research.

Keywords: Financial data warehousing, case study, project management, knowledge transfer, common understanding, communication theory.

1 INTRODUCTION

During the last twenty years research regarding success factors and management of data warehouse (DWH) projects as special information systems development (ISD) projects was conducted to a great extend. These works focus mainly on a technical or a business view of DWHs and not on the integrated process. While the technical papers deal i.e. with physical implementation, physical data models, schema design etc., the business oriented publications focus on a logical design and the specification of information requirements. In addition there are few authors which allow a holistic view of the DWH development process (Kimball & Ross, 2002; Kimball & Caserta, 2004; Inmon, 2005; Jarke et al., 2003; Adelman & Moss, 2000). Although an abundance of modelling languages, process models and success factor analyses were published, DWH projects still cause crucial problems in practice.

„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 p 252). We will follow this approach in our paper and combine rigor and relevance by using approved and broad accepted methods of case study research to identify and to explain real life phenomena. These observations were analysed by using well known theories from other disciplines, e.g. communication and language theory. This enabled the explanation of real life observations. Furthermore areas of action have been identified to increase the performance in DWH projects. Consequently, our findings are based on a strong theoretical basis.

In our case study we describe DWH projects in three different banks. The observations about the DWH projects were acquired over a duration of more than three years. Although data warehousing techniques have been used in the financial industry for several years, there is still need for implementation and extension of integrated informational systems (the term "informational systems" will be used to distinguish them from "operational systems" as purposed by (Gardner, 1998)). The importance of integrated and consolidated risk estimation by using DWH techniques has been recently emphasised by the sub prime crisis. DWH projects in the financial industry or with general financial background are characterized by high complexity and involvement of many different project roles. Based on this case study we will show the specifics of financial data warehousing (FDWH) projects and we will derive implications for FDWH project management.

This paper is structured as follows. Section 2 gives an overview of related work in data warehousing. The section also introduces FDWH as data warehousing with specific requirements. Section 3 summarizes the underlying research methodology of this work. Section 4 presents the case study, while in Section 5 the specific character of FDWH is described. Hence, two major tasks within FDWH projects are focussed in subsections. Section 6 summarizes the findings and introduces five statements about FDWH projects. Finally, section 7 concludes and gives an outlook to future research.

2 RELATED WORK

Since the 1970s comprehensive research took place in the field of data warehousing. The relevant research can be classified into groups with respect to their different focuses. There are fundamental works that explain the overall process of DWH in depth (e.g. Mucksch, 2000; Chamoni & Gluchowski, 2006; Bauer & Günzel, 2004; Kimball & Ross, 2002; Goeken, 2006; Inmon, 2005; Kurz, 1999; Jung, 2001). Most research concerning the overall process of DWH focuses either on conceptual modelling (e.g. Burmester & Goeken, 2006; Böhnlein & Ulbrich-vom Ende, 2000; Goeken, 2004; Jukic, 2006; Hüsemann et al., 2000; Hahne, 2002) or on a rather model-driven and technology-oriented implementation of DWH (e.g. Herden, 2002; Lechtenbörger, 2001). Only few publications integrate both perspectives (e.g. Lehner, 2003).

The seminal work concerning DWH is accredited to Inmon who followed a data-driven approach (Inmon, 2005). Later publications introduced different approaches, e.g. user-driven (e.g. Bruckner et al., 2001; Holten, 1999; Winter & Strauch, 2003; Holthuis, 1999), goal-driven (e.g. Böhnlein & Ulbrich-vom Ende, 2000; Kimball et al., 1998) or mixed approaches (e.g. Guo et al., 2006). In addition, there are publications which can be seen as surveys each addressing specific aspects of DWH projects. The importance of the requirements engineering phase for DWH success and the specific problems in identifying and modelling of information requirements was identified for both user- and goal-driven approaches (e.g. Holten, 1999; Strauch & Winter, 2002; Winter & Strauch, 2004).

Since several DWH projects failed or significantly exceeded budgets in the past, extensive research concentrates on quantitative or qualitative analyses of success factors as well as contemporary best-practices within the field of DWH. These works focus on organizational as well as methodological research questions (e.g. Herrmann & Melchert, 2004; Watson et al., 2004; Hwang et al., 2004; Hwang & Xu, 2007; Weir et al., 2003; Shin, 2002; Heck-Weinhart et al., 2003; List et al., 2002; Massa & Testa, 2005; Sen & Sinha, 2005; Sanders & Courtney, 1985; Chenoweth et al., 2006; Behme & Mucksch, 2001).

The majority of authors focus on the project management during requirement engineering and conceptual modelling phase. Thereby, implying the assumption that the creation of elaborate data specifications and high formal standards represent the major challenge of DWH projects. It is assumed that if a specification is correct in terms of the applied formal specification language, the implementation will represent the desired result. This functionalist way of thinking has already been criticized by (Iivari et al., 2001) from an IS practice point of view. The main challenge was seen in matching user requirements. Other challenges in DWH projects, e.g. the ETL development process were often mentioned, but only a few researchers work in this subject area (Vassiliadis et al., 2005; Skoutas & Simitsis, 2006; Sellis, 2006), although relevance is assessed to be high (Rizzi et al., 2006). This is remarkable, because the high effort of ETL and interface development in DWH projects – about 50 to 80% of overall budget (Vassiliadis et al., 2002; Kimball & Caserta, 2004) – is without controversy.

None of the aforementioned approaches explicitly address the personal interactions within project teams. Hitherto, only few publications focus on soft factors in DWH or ISD projects, e.g. in the field of knowledge management (e.g. He, 2004), learning processes (e.g. Pirinen & Pekkola, 2006) and communication (e.g. Gallivan & Keil, 2003). All mentioned research concerning DWH process models disregards the specific problem of interpersonal communication within DWH projects beyond the requirements analysis phase.

Recapitulatory, the authors state that no holistic work regarding the specifics of DWH projects in the financial industry exists so far. This work analysis the specifics of projects in FDWH and provides a contribution to the aspect of communication and interaction in complex DWH projects. Furthermore the implications to project management in FDWH projects will be derived.

3 RESEARCH METHODOLOGY: CASE STUDY

The empirical part is the result of an exploratory case study performed according to Yin (Yin, 2003). This methodology was used because we have examined complex organizational phenomena. The purpose of the examination was to understand the interaction of the different project members in large DWH projects with respect to requirements engineering and development issues. Therefore we had to deal with "how and why" questions about a contemporary set of events over which we had little control (Yin, 2003 p 20). Relevance is ensured by focussing on real life phenomena. The application of the broadly accepted method of Yin guarantees the rigor of our results as well.

The described case was documented by the two authors of this document. Both authors are managers at zeb/information.technology, a group member of zeb/ business consulting. At least one of the authors was involved in each of the projects described. The observed facts were documented and reconciled with other project members employed by the banking groups or zeb/. Furthermore the project documentation (minutes of meetings, presentations of lessons learnt sessions, data models, calendars, etc.) was used to confirm the presented facts. Additionally unstructured information and non-formal documentation (i.e. email traffic) was evaluated. To ensure adequate distance and to increase the objectivity of the case study, all findings were discussed in our research team with non-involved people.

4 CASE STUDY: DWH DEVELOPMENT IN FINANCIAL INDUSTRY

The specifics of FDWH projects were derived from experiences acquired by participation in DWH projects in the financial industry supported by zeb/information.technology. During the last decade about 30 FDWH projects were managed or supported by zeb/. To point out FDWH specifics, three characteristic FDWH projects were chosen as examples, to be presented in the following case study, although the same observations were made as well during other FDWH projects.

Topic	Bank A	Bank B	Bank C
General Information			
Balance sheet total	>120 billion Euro	>45 billion Euro	>4 billion Euro
Project size (DWH)	>5000 man-days	>3000 man-days	>1500 man-days
Organization of ETL development	Different teams; not involved during conceptual modelling	Different teams; partly involved in conceptual modelling	One team, also doing conceptual modelling
Project goal	Data platform for Basel II	Data platform for the whole banking group	Data platform for Basel II and internal controlling
Number of central core systems	8 large systems	9 large systems	10 systems
Subsidiaries	Detailed non-retail data of 16 subsidiaries; aggregated data of > 150 shareholdings	6 subsidiaries detailed load and aggregated data of approx. 80 shareholdings	-
Load frequency	Monthly	Daily	Daily
Similarities			
DWH approach	Goal-driven approach with end user involvement for specific adaptations		
Standard software	Used for database management system, reporting tools and calculation kernels		
Business focus	Development of a central DWH to fulfil Basel II and regulatory reporting requirements		
Data sources	Data sources are broadly spread over several systems		
Differences			
Existing platform	Existing DWH solution had to be extended		DWH from the scratch
ETL tools	Use of ETL tool	No use of ETL tools	
Project language	Different natives languages – project language: English	Central team: German Subsidiaries team: English	German only
Stability of Source Systems	High - Core system stable	Low - Implementation of new core systems as data sources for the DWH	
Documentation of interface Specification	XSD-based and textual interface specification, PPT presentations for training	Group-wide database containing interface specification	Excel-based and textual interface specification

Table 1: Cases overview

The following observations were made in three DWH projects in the financial industry from now on referred to as Bank A, B and C. All projects were set up to implement a central data supply for at least regulatory reporting. Goal of the projects was to meet the requirements specified by Basel II. One of the main topics of Basel II is the calculation of *Risk Weighted Assets* based on single transaction data for the whole banking groups which requires strong changes in the IT infrastructure. Although the main goal of the three Basel II projects was the same, each project had a specific character. The banks size reached from total assets of four billion Euros to over a hundred billion Euros. The similarities and the differences between the projects are outlined in Table 1. Accordingly the utilised management techniques differ between all projects, but the major goals for the management were equal. They had to deal with deadline achievement, budget compliance and appropriate solution quality, which were endangered by emerging *communication defects* during the projects.

In all three implementation projects problems in communication were observed. These misunderstandings between the project members could not be anticipated ex ante in the projects planning phase. In the following chapter two typical project tasks will be explained. During each task specific communication problems were observed.

5 SPECIFICS OF FINANCIAL DATA WAREHOUSING

During the last 10 years the consolidation of data and the development of central DWHs were major trends in IT projects in the financial industry – especially in the banking sector. One of the main reasons for this was the adoption and extension of regulatory requirements, such as IFRS or Basel II. Calculations that compute the institutions whole portfolio became necessary, requiring group wide data consolidation. The second driver for DWH projects in the banking sector was the enhancement of internal reporting, especially controlling and risk management. For Bank B and C in our case study this was the first reason for starting DWH projects.

Resulting from the legal rules given by regulatory authorities and a common understanding of methods for internal controlling and risk management, the content of the DWH can be derived to a certain degree. While experiences about the business topics were made during the last years, the raw data model designs of DWHs were approved by practice and can be taken as a given starting point. This leads to a goal-driven project approach. The DWH was build to ensure the data supply for a well known specific business application. Although the basic content can be derived from the application-specific requirements, individual adoptions (e.g. for special banking products or organizational structure) became necessary. Therefore an end user involvement is obligatory during the conceptual modelling phase. Thus FDWH projects have elements of goal-driven and user-driven approaches. The usage of common methods allows a reusability of software, which is shown by standard software of different suppliers. The spectrum reaches from solutions specialised on regulatory reporting (such as “SAMBA”, “ABACUS/DaVinci” or “Svenson”) to integrated solutions to cover all informational applications (internal and regulatory reporting) of a banking institute (e.g. “Fermat” or “zeb//control”).

Another reason for the importance of goal- or user-driven approaches is the high complexity of FDWHs. The complexity results from the high amount of data and the company’s organizational structure. Each business area uses its own systems for its specific business. Complexity can be reduced by utilising user- or goal-oriented and evolutionary/incremental DWH development approaches (see also Targett, 1985 p 1001; Benefelds & Niedrite, 2004 p 553; Khirallah, 1999).

The reuse of methods or the implementation of standardised solutions transferred the main challenge in DWH projects from the conceptual data modelling to the ETL and interface design. While content can be easily adapted, the correct data delivery becomes a crucial problem. Moreover the FDWH projects are mostly driven by a specific business topic. During a given project a specific business function (such as Basel II) will be realised. Afterwards the next project will be started and the DWH will be extended or adapted (also identified by Benefelds & Niedrite, 2004 p 553), thereby causing integration scenarios and semantic change. New and changing requirements have to be integrated in an existing DWH solution. Both characteristics have strong impact to DWH project management and will be discussed in detail in the following sections.

5.1 Extension of existing FDWH-Solution

As mentioned before, the data requirements are frequently determined by a standardised data consumer, such as regulatory reporting solutions or calculation engines. Given this case, the goal of a DWH project is to deliver the required data to the final data consumer. If the institutions already deployed a FDWH, the existing solution has to be extended. This applies for Bank A and B. Both had developed an existing FDWH for the provision of controlling and MIS functionalities. For the extension two steps were performed in the projects. First a mapping between the data requirements and the existing FDWH data model was specified and a gap list was derived. In the second step the gap list attributes were added to the FDWH data model. Changes in the data model structure were necessary to close the gaps in addition to merely adding attributes. In both projects different teams managed data requirements specification (DRT) and FDWH development (DEVT). As well during the mapping as the extension phase the project members had to deal with upcoming problems:

Based on the documentation of the FDWH solution the mapping was performed by DRT. The result showed several cases of mapping errors. For example Bank A's FDWH concept contains a table called "counterparty". Due to a translation mistake during concept design, the table was named "counterparty" but contained basic transaction data. However, the table was erroneously expected to contain counterparty (e.g. customer, guarantor) data. Besides this special example usual problems due to homonyms (e.g. "Account") and synonyms (e.g. "Customer" vs. "Counterparty", "Facility" vs. "Counterparty") could be observed as well on entity/table as on an attribute level. The specification on its own proved to be insufficient for quality assurance and correction of the mapping. For a final decision, if a given attribute of the data requirements (DR) equalled one in FDWH, direct communication between DRT and DEVT was necessary in some cases (e.g. exact definition of "book value", "limit amounts", "market values", etc.). This may be surprising, because all these terms are common in the banking domain, but we encountered several similar problems.

After receiving the gap list, DEVT specified the necessary extension of the FDWH solutions and published a new data model. In Bank A's project the extension was complicated, which was caused by different granularity between DR and FDWH. For example, loans and other simple credit transactions were modelled in DR as a single object "Loan". In the physical DWH model a more granular way was used. For different types of loans different tables were modelled ("Account" for current accounts, "Customer Loan", "Limit" for credit lines and "Credit Cards" for credit and debit card transactions). Review and quality assurance (QA) of the extended FDWH model showed some mistakes in the gap lists interpretation by DEVT. For discussion of the findings and correction of the mistakes workshops between DRT and DEVT were performed. Semantics and exact meanings of attributes also were crucial at Bank B. Misunderstandings caused late changes to the securities system interface. The semantics of the specification were interpreted in comprehensible, but different ways. As the specification was the same for every securities source system and some of them connected correctly, it became clear, that specifications are only understandable for a specific group of people that share the same view as the authors of the specification. In most situations we found that direct communication was necessary to ensure correct understanding and an implementation taking the given system complexity into account.

Extensions of FDWH solutions cause lots of coordination efforts between involved experts. Caused by the high number of synonyms and homonyms and the difficulties in giving a conclusive definition of each attribute and its relations lots of misunderstandings can be observed in the mentioned FDWH projects. This fact was reflected in the project management. In both projects the first approach was mainly based on specification documents. After observation of several communication defects the project organization was adapted. Bank A for example had re-organized the collaboration between DRT und DEVT for the next releases: They set up regularly workshops and installed mixed QA-Teams to ensure intensive personal communication.

5.2 ETL and interface design

Looking at the data supply of the DWHs in the three banks, we first have to differentiate between data flows from central systems and interfaces for external data supply. Connecting to central source systems is a standard task for making a DWH run by implementing an ETL process. All banks used the developed data requirement documents to initiate an information exchange between the DWH team and each team of the supplying source systems. Bank A and B additionally had to deal with data sources from outside the central IT environment. All external data suppliers also got the data requirements supplemented by a technical description of the expected data exchange method and process. Bank A used XML to assure the data structure and crude data quality, while Bank B provided a central Lotus Notes database containing structured data requirements all subsidiaries can access. For external data supply the ETL process has an internal and an external part. To keep the interfaces as simple as possible for the external data suppliers, Bank B decided not to demand the same interface data as the parent company needed to deliver from its central systems. Only a subset of data was required for supervisory purposes. So there were extraction and transformation processes running at each subsidiary, transformation and loading processes handling external data interfaces in the central IS infrastructure and ETL processes connecting centrally hosted systems to the DWH. Bank A uses IBM WebSphere DataStage, one of the leading ETL tools available at the market, to support the development of all ETL processes. The other banks opted for individual ETL development. Bank B is connecting the FDWH to it's custom core system and a couple of various other source systems, while Bank C has a SAP core system solution with distributed interfaces (no SAP Financial Database) containing most of the required data.

The requirement documents handover was accompanied by workshops, which were held with all involved stakeholders. Although the data supplier side seemed to be provided with sufficient information for the development process and the involved persons agreed with that verbally and by email, there were great misunderstandings that were uncovered later. The main reason mentioned for misunderstanding was a different understanding of terms used in the requirements document leading to a false interpretation in technical design. Bank B solved this problem by intensifying the contact between the DWH team and the data suppliers. There were only four project members of the bank having insight into the DWH as needed to reconcile the requirement documents with the data suppliers, but these employees also had to run the operational DWH. In Bank C continuous workshops were needed due to team member exchanges on supplier side and complex mapping problems between the source system and the DWH, which both led to misinterpretation by the recent developer.

All banks made the ETL processes run after a phase of more intense communication between the involved teams. To increase the intensity of communication the project management decided to send high-skilled employees or consultants to the central and local source system teams. By using face-to-face communication the teams could cope with all remaining problems in a short period of time compared to earlier attempts using enhanced requirement documents.

6 DISCUSSION

The observations made during the analyses of the FDWH implementation projects are the base for five statements about specifics of FDWH projects.

Statement 1: *"FDWH projects have to deal with several business domains and fields of knowledge"*

In financial institutes, different business models in different departments and subsidiaries are used. For example the treasury is concerned with inter-bank business while the retail department has to deal primarily with private customers and loans. In Bank A and Bank B special business activities were separated in own subsidiaries. This was done as well to service different markets (e.g. CEE activities of Banks A and B) as to provide special business (e.g. leasing or factoring). Therefore different departments and subsidiaries have different thematic emphases in IT support on operational and informational level. This theme-oriented grouping is also reflected in the IT Architecture of the financial institutes. Often business domain oriented building blocks were used as well for operational as for informational systems which lead to several source systems for the FDWH. In the IT Architecture of all three banks separate systems for loans and current accounts, security transactions, treasury transaction, collaterals, securities lending, etc. were in place. This continuous theme-orientation leads to different domain knowledge in different departments and subsidiaries and specific fields of knowledge arise. To build up a centralized and integrated system (like a FDWH) a common understanding between all involved department and subsidiaries has to be reached.

Statement 2: *"FDWH Projects are characterized by a high semantic complexity"*

As mentioned in statement 1 several business domains can occur within a single financial institution. Due to the different fields of knowledge a high number of context-dependent homonyms (e.g. limit, facility, book value, market price) and synonyms (e.g. debt security, bond, obligation) can be observed. Due to this variety, the definition of FDWH terms (e.g. dimensions and measures) is challenging. For providing the required measures in the FDWH, often complex calculations are necessary (e.g. cash-flow generation, calculation of present values, value at risk, risk weighted assets). These calculations have to be implemented during the ETL process. During the "T" (transformation) all calculations are computed, transforming source data to the final form for the data consumer. The spectrum reaches, depending on the source systems, from simple mapping rules up to usage of calculation kernels. In all analysed FDWH projects several source systems had to be connected. The number reaches from 10 central systems in Bank C up to more than 8 central source systems, 16 subsidiaries with heterogeneous source systems and more than 150 small institutes with aggregated data in Bank A. Besides this variety, another driver of complexity is the stability of the IT environment. In Bank B the core system as the main data source was replaced during the time period involved in the case study. Therefore no experts for the new systems were available and the project team had to learn the semantic of the new source system. The raised amount of changes within the source systems demanded flexibility and adaptability of the ETL process. On the other hand the software that had to deliver regulatory measures calculated from the DWH to the supervisory authorities was in its implementation phase as well. Especially at Bank C changes of requirements occurred several times. So a reduced stability in specifications has to be considered. The high complexity of FDWH solutions was also emphasized by (Benefelds & Niedrite, 2004), although we do not agree with their valuation of source system similarity. As shown in our cases, financial source systems are varied (i.e. systems for derivatives, loans, equity, securities) and additionally are subjected to short release cycles due to ongoing product innovation in the financial services sector. This semantic complexity increases the difficulties in reaching an overview and a common understanding about FDWH solutions.

Statement 3: *"Specification-based approaches are not sufficient in FDWH projects"*

In all presented FDWH projects sufficient accuracy of requirements was an issue. During the QA and test phases lots of interpretation mistakes were observed. In the project of Bank A lots of defects could be detected during the analysis of the data delivered by the subsidiaries. Surprisingly the same parts of a specification were interpreted correctly by some subsidiaries but misunderstood by another. That leads to the conclusion that understanding of specification is dependent on the personal knowledge of the addressee. This fact, the importance of a "common ground" in communication acts, was broadly discussed in the language theory of Clark (Clark, 1996). For a complete and correct understanding of a specification a perfect anticipation of the addressee's knowledge is necessary, but this prerequisite is not realistic. Another issue was the timely detection of interpretation mistakes. Usually it is expected that the addressee will ask in case of problems in understanding specifications, but during the projects in Bank A and B only few questions were raised by the subsidiaries. Due to homonyms and specification based collaboration situations of "illusion of evidence" occur which are described in expert-layperson communication scenarios (Bromme & Jucks, 2001; Bromme et al., 2004; Bromme et al., 2005). In these situations sender and addressee assume a correct understanding. To avoid "illusion of evidence" situations and to ensure correct understanding by the addressee an exclusive usage of specification is not sufficient. Beside the specification additional methods are required. To overcome the specification-based problems face-to-face-communication was increased in all observed projects.

Statement 4: *"FDWH projects require knowledge transfer methods and strong interaction"*

For a successful FDWH implementation a common understanding between all involved project members has to be ensured. This cannot be ensured by merely interchanging a written specification. In all presented FDWH implementation projects misunderstandings about the specification documents occurred. These misunderstandings could be detected and eliminated easily by using personal communication because non-verbal correction-mechanisms can be used and direct feedback-loops are possible (Bromme & Jucks, 2001). In the projects, the level of personal interaction was raised during project progress. At the beginning most communication was done by written documents. After detection of several misunderstandings during test phases workshops were set up to discuss specifications together with the involved departments. Additionally in projects A and B regular visits in local units were performed. During these discussions many additional "illusion of evidence" situations could be detected and solved easily which allowed a proactive avoidance of interpretation mistakes and therefore implementation errors, i.e. the language theory of Clark and Bromme was supported by our observations. Another problem the projects had to cope with was the fluctuation in the project teams. Within the duration of multiple years project members having important knowledge left the company. Although successors had time to familiarize with the environment, there are significant consequences to the individual understanding and occurring new problems. In Bank C responsible core system developers changed a couple of times, which lead to additional coordination overhead and knowledge transfer. Therefore, changes in the project team had to be considered in the projects as well. To ensure the success of FDWH projects personal interaction and knowledge transfer have to be reflected in process models and project management. To assure effective knowledge transfer, all three banks teamed up the experts for specific topics.

Statement 5: *"FDWH specifics are non-satisfactory reflected in DWH process models"*

We already mentioned in chapter 2 that the specifics of FDWH are not satisfactory reflected in DWH process models and project management approaches. The main problem discussed in literature is elaboration of the specification for a DWH. The complexity of this task was often discussed, but the later tasks, like correct implementation of specifications are not covered. In other words: the importance of understanding a specification is not emphasized. DWH research identified high skills as mandatory, but specific knowledge regarding the individual systems is needed. Our observations show that this is one of the main issues for success in FDWH implementation projects, but knowledge transfer was not explicitly advanced by actual published DWH process models and project management approaches.

7 CONCLUSION & OUTLOOK

Our case studies show that a common understanding between all involved parties of FDWH projects is the prerequisite for a successful implementation of FDWHs. This applies as well for the different involved departments and subsidiaries as for project members with different skills (e.g. business expert, database designer, ETL programmer, source system administrator). The important role of personal knowledge can be observed in all phases of the project from requirements analysis up to implementation and test.

Our main finding is that personal problem-relevant knowledge exists, which cannot be expatiated in specifications. Specifications will be misinterpreted by people who are not part of the "community" and have a different knowledge context, which can be observed in complex FDWH implementation projects. Complex ETL processes and interface implementations need personal communication as long as no common ground for inter-subjective understanding is reached. But how can the threshold between conceptual modelling and ETL be crossed? A common solution is the recruiting of specialists to assure knowledge transfer. In every solution common understanding needs communication. These findings will be stronger funded in future work by applying communication and language theory e.g. of Clark , Kalmlah/Lorenzen and Bromme (Bromme & Jucks, 2001; Bromme et al., 2004; Bromme et al., 2005; Clark, 1996; Kamlah & Lorenzen, 1984). Therefore it is planned to cover as well the basics of successful communication and knowledge transfer processes (macro perspective) as the prerequisites for a common understanding of single terms (micro perspective).

Another idea to avoid misunderstanding is the contextualization and extension of the specification documents which causes documentation effort. Even if the requirements are not formalizeable as in FDWH scenarios it is difficult to find an appropriate level of detail. Because degree of misunderstandings of a specification is low (about 10%) and misunderstandings occur depending on the addressee's personal context, personal communication is expected be a more efficient way to clarify the remaining problems. Looking at these points, there seems to be a point in each project, where formal specification is less efficient than personal knowledge transfer. We assume the existence of such a point and want to find evidence for this idea.

Besides building a theoretical base for the observed factors further research has to cope with methods for a successful project-wide knowledge transfer in FDWH projects. These findings have to be picked up in the management of the development process. We expect implications for project management of FDWH during different phases for all of these investigations. Existing process models and project management approaches have to be extended to fulfil the requirements in FDWH projects.

Our findings were derived from specific observations in FDWH projects. These projects are suitable since because of their complexity lots of communication problems can be observed. We assume that our conclusions are valid for other IS domains, where complex and low formalized problems in heterogeneous business domains need to be solved.

References

Adelman, S. and Moss, L. T. (2000) *Data warehouse project management.* Addison-Wesley Boston, MA.

Bauer, A. and Günzel, H. (Eds.) (2004) *Data-warehouse-systeme: Architektur, entwicklung, anwendung.* dpunkt-Verl., Heidelberg.

Behme, W. and Mucksch, H. (2001) *Data warehouse-gestützte anwendungen: Theorie und praxiserfahrungen in verschiedenen branchen.* Gabler, Wiesbaden.

Benefelds, J. and Niedrite, L. (2004) Comparison of approaches in data warehouse development in financial services and higher education. In *Proceedings of the 6th International Conference on Enterprise Information Systems*, pp 552-557, Porto.

Böhnlein, M. and Ulbrich-Vom Ende, A. (2000) Business process oriented development of data warehouse structures. *Data Warehousing*, 3-21.

Bromme, R. and Jucks, R. (2001) Wissensdivergenz und kommunikation: Lernen zwischen experten und laien im netz. *Partizipation und Interaktion im virtuellen Seminar*, 81-103.

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.

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.

Bruckner, R., List, B. and Schiefer, J. (2001) Developing requirements for data warehouse systems with use cases. *Proc. 7th Americas Conf. on Information Systems*, 329–335.

Burmester, L. and Goeken, M. (2006) Method for user oriented modelling of data warehouse systems. In *ICEIS 2006 - Proceedings of the Eighth International Conference on Enterprise Information Systems: Databases and Information Systems Integration* (Manolopoulos, Y. and Filipe, J. and Constantopoulo, P. and Cordeiro, J., Eds), pp 366-374, Paphos.

Chamoni, P. and Gluchowski, P. (Eds.) (2006) *Analytische informationssysteme: Data warehouse, online analytical processing, data mining.* Springer Bln, Berlin.
Chenoweth, T., Corral, K. and Demirkan, H. (2006) Seven key interventions for data warehouse success. *Communications of the ACM* 49 (1), 114-119.
Clark, H. H. (1996) *Using language.* Cambridge University Press, Cambridge.
Gallivan, M. J. and Keil, M. (2003) The user-developer communication process: A critical case study. *Information Systems Journal* 13 (1), 37-68.
Gardner, S. R. (1998) Building the data warehouse. *Communications of the ACM* 41 (9), 52-60.
Goeken, M. (2004) Referenzmodellbasierte einführung von führungsinformationssystemen grundlagen, anforderungen, methode. *WIRTSCHAFTSINFORMATIK* 46 (5 S 353), 365.
Goeken, M. (2006) *Entwicklung von data-warehouse-systemen : Anforderungsmanagement, modellierung, implementierung.* Deutscher Universitäts-Verlag, Wiesbaden.
Guo, Y., Tang, S., Tong, Y. and Yang, D. (2006) Triple-driven data modeling methodology in data warehousing: A case study. *Proceedings of the 9th ACM international workshop on Data warehousing and OLAP*, 59-66.
Hahne, M. (2002) *Logische modellierung mehrdimensionaler datenbanksysteme.* Dt. Univ.-Verl., Wiesbaden.
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.
Heck-Weinhart, G., Mutterer, G., Herrmann, C. and Rupprecht, J. (2003) Entwicklung eines angepassten vorgehensmodells für data-warehouse-projekte bei der w&w ag. In *Data warehouse management: Das st. Galler konzept zur ganzheitlichen gestaltung der informationslogistik* (Von Maur, E. and Winter, R., Eds), Springer, Heidelberg.
Herden, O. (2002) *Eine entwurfsmethodik für data warehouses.* Oldenburg, Univ., Diss.,.
Herrmann, C. and Melchert, F. (2004) Sponsorship models for data warehousing: Two case studies. *Proc. AMCIS*,
Holten, R. (1999) *Entwicklung von führungsinformationssystemen. Ein methodenorientierter ansatz.* Dt. Univ.-Verl., Wiesbaden.
Holthuis, J. (1999) *Der aufbau von data warehouse-systemen: Konzeption-datenmodellierung-vorgehen.* Duv.
Hüsemann, B., Lechtenbörger, J. and Vossen, G. (2000) *Conceptual data warehouse design.* Univ., Münster.
Hwang, H.-G., Ku, C.-Y., Yen, D. C. and Cheng, C.-C. (2004) Critical factors influencing the adoption of data warehouse technology: A study of the banking industry in taiwan. *Decision Support Systems* 37 (1), 1.
Hwang, M. I. and Xu, H. (2007) The effect of implementation factors on data warehousing success: An exploratory study. *Journal of Information, Information Technology, and Organizations* 2,
Iivari, J., Hirschheim, R. and Klein, H. K. (2001) A dynamic framework for classifying information systems development methodologies and approaches. *Journal of Management Information Systems* 17 (3), 179-218.
Inmon, W. H. (2005) *Building the data warehouse.* Wiley, Indianapolis.
Jarke, M., Vassiliou, Y., Lenzerini, M. and Vassiliadis, P. (2003) *Fundamentals of data warehouses.* Springer, Berlin, New York.
Jukic, N. (2006) Modeling strategies and alternatives for data warehousing projects. *Communications of the ACM* 49 (4), 83-88.
Jung, R. (Ed.) (2001) *Data warehousing 2000 : Methoden, anwendungen, strategien.* Physica-Verl., Heidelberg.
Kamlah, W. and Lorenzen, P. (1984) *Logical propaedeutic: Pre-school of reasonable discourse.* University Press of America.
Khirallah, K. (1999) Building a data warehouse at chase manhattan bank. *Bank Accounting & Finance* 12 (3), 40.

Kilduff, M. (2006) Editor's comments: Publishing theory. *Academy of Management Review* 31 (2), 252-255.
Kimball, R. and Caserta, J. (2004) *The data warehouse etl toolkit: Practical techniques for extracting, cleaning, conforming, and delivering data.* John Wiley & Sons.
Kimball, R., Reeves, L., Thornthwaite, W., Ross, M. and Thornwaite, W. (1998) *The data warehouse lifecycle toolkit: Expert methods for designing, developing and deploying data warehouses with cd rom.* John Wiley & Sons, Inc. New York, NY, USA.
Kimball, R. and Ross, M. (2002) *The data warehouse toolkit: The complete guide to dimensional modeling.* Wiley, New York [u.a.].
Kurz, A. (1999) *Data warehousing: Enabling technology.* mitp, Bonn.
Lechtenbörger, J. (2001) *Data warehouse schema design.*
Lehner, W. (2003) *Datenbanktechnologie für data-warehouse-systeme: Konzepte und methoden.* dpunkt-Verl., Heidelberg.
List, B., Bruckner, R. M., Machaczek, K. and Schiefer, J. (2002) A comparison of data warehouse development methodologies case study of the process warehouse. *Proc. DEXA*, 203–215.
Massa, S. and Testa, S. (2005) Data warehouse-in-practice: Exploring the function of expectations in organizational outcomes. *Information and Management* 42 (5), 709-718.
Mucksch, H. (2000) *Das data-warehouse-konzept: Architektur - datenmodelle - anwendungen; mit erfahrungsberichten.* Gabler, Wiesbaden.
Pirinen, A. and Pekkola, S. (2006) Different perspectives on learning in information system development. In *IRIS*, LO-skolen.
Rizzi, S., Abelló, A., Lechtenbörger, J. and Trujillo, J. (2006) Research in data warehouse modeling and design: Dead or alive? *Proceedings of the 9th ACM international workshop on Data warehousing and OLAP*, 3-10.
Sanders, G. L. and Courtney, J. F. (1985) A field study of organizational factors influencing dss success. *MIS Quarterly* 9 (1), 77-93.
Sellis, T. (2006) Formal specification and optimization of etl scenarios. *Proceedings of the 9th ACM international workshop on Data warehousing and OLAP*, 1-2.
Sen, A. and Sinha, A. P. (2005) A comparison of data warehousing methodologies. *Communications of the ACM* 48 (3), 79-84.
Shin, B. (2002) A case of data warehousing project management. *Information & Management* 39 (7), 581-592.
Skoutas, D. and Simitsis, A. (2006) Designing etl processes using semantic web technologies. *Proceedings of the 9th ACM international workshop on Data warehousing and OLAP*, 67-74.
Strauch, B. and Winter, R. (2002) Vorgehensmodell für die informationsbedarfsanalyse im data warehousing. *Vom Data Warehouse zum Corporate Knowledge Center; Heidelberg*, 359-378.
Targett, D. (1985) An experience in the design of a management information system for a bank. *Journal of the Operational Research Society* 36 (11), 999-1007.
Vassiliadis, P., Simitsis, A., Georgantas, P., Terrovitis, M. and Skiadopoulos, S. (2005) A generic and customizable framework for the design of etl scenarios. *Information Systems* 30 (7), 492-525.
Vassiliadis, P., Simitsis, A. and Skiadopoulos, S. (2002) Conceptual modeling for etl processes. *Proceedings of the 5th ACM international workshop on Data Warehousing and OLAP*, ACM Press, McLean, Virginia, USA.
Watson, H. J., Fuller, C. and Ariyachandra, T. (2004) Data warehouse governance: Best practices at blue cross and blue shield of north carolina. *Decision Support Systems* 38 (3), 435-450.
Weir, R., Peng, T. and Kerridge, J. (2003) Best practice for implementing a data warehouse: A review for strategic alignment. *Proc. DMDW*,
Winter, R. and Strauch, B. (2003) A method for demand-driven information requirements analysis in data warehousing projects. *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on*, 9.
Winter, R. and Strauch, B. (2004) Information requirements engineering for data warehouse systems. *Proceedings of the 2004 ACM symposium on Applied computing*, 1359-1365.
Yin, R. K. (2003) *Case study research: Design and methods.* Sage Publications Inc.