Anwendung der Kommunikationstheorie zur Unterstützung der Implementierung von Data Warehouse-Lösungen bei Finanzdienstleistern

Inauguraldissertation

zur Erlangung des Doktorgrades des Fachbereiches
Wirtschaftswissenschaften der Johann Wolfgang Goethe-Universität
Frankfurt am Main

vorgelegt von
Wolf Behrmann
aus Münster (Westf.)

Frankfurt, 2016

Inhaltsverzeichnis

Abbildungsverzeichnis

Abkürzungsverzeichnis

BCBS	Basel Committee on Banking Supervision
BIS	Bank for International Settlements
CSCW	Computer Supported Cooperative Work
DaRT	Data Requirements Tool
DBO	Design Boundary Object
DWH	Data Warehouse
ETL	Extraction, Transformation and Loading
FDL	Finanzdienstleister
FDWH	Financial Data Warehouse
IFRS	International Financial Reporting Standard
HCI	Human Computer Interaction
IS	Informationssystem
VGM	Vorgehensmodell

Inhaltsangabe der Dissertation

Zusammenfassung der Dissertation
Gesamtdarstellung der Dissertation und Einordnung der zugrunde gelegten Veröffentlichungen. Insbesondere wird auf die Motivation, die behandelte Forschungsfrage, das Forschungsvorgehen, den Forschungsbeitrag und die Limitierungen eingegangen. Die Grundlage für die Zusammenfassung bilden vier Veröffentlichungen:

Paper 1
Behrmann, W.; Räkers, M.
Specifics of Financial Data Warehousing and Implications for Management of Complex ISD Projects
In: Proceeding of 16th European Conference on Information Systems (ECIS 2008), Galway, Ireland, Seite 1740-1751

Paper 2
Hoffmann, A.; Behrmann, W.
Classification of Communication Defects in DWH Projects
In: Proceedings of 14th Americas Conference on Information Systems (AMCIS 2008), Toronto, Canada, Seite 296

Paper 3
Behrmann, W.
Hitting Interface Specifications - The Role of 'Common Ground' in Financial DWH Projects
In: Proceedings of 15th Americas Conference on Information Systems (AMCIS 2009), San Francisco, USA, Seite 312

Paper 4
Rosenkranz, C.; Räkers, M.; Behrmann, W.; Holten, R.
Supporting Financial Data Warehouse Development: A Communication Theory-Based Approach
In: Proceedings of 31st International Conference on Information Systems (ICIS 2010), St. Louis, Missouri, USA, Seite 12

1 Einleitung und Motivation

Bei Finanzdienstleitern ist der Einsatz von Informationstechnologie zum wesentlichen Produktionsfaktor geworden: die Erbringung von der Kundendienstleistungen basiert auf der Abwicklung von Prozessen und der Verarbeitung von Informationen (Baldwin et al. 2001; Mulligan and Gordon 2002; Tan and Teo 2000). Dies stellt hohe Ansprüche an die Verfügbarkeit von Daten, die im jeweiligen Prozessschritt in adäquater Form und Qualität zur Verfügung stehen müssen. Zusätzlich unterliegen die Finanzdienstleister einem zunehmenden regulatorischen Druck und müssen z.B. die Richtlinien Sarbanes-Oxley-Act, IFRS oder Basel II/III erfüllen. Die steigenden Berichtsanforderungen resultieren aus einem stetig zunehmenden Informationsbedarf (McLaughlin 2004), der sich auch durch die hohen Anforderungen aus den durch das Basel Committee of Bank Supervision verabschiedeten „Principles for effective risk data aggregation and risk reporting“ (bekannt auch unter BCBS #239) noch einmal bestätigt (BCBS 2013). Die Abdeckung des erweiterten Informationsbedarfs erfordert in der Regel den Aufbau von Data Warehouse (DWH)-Lösungen, die in den letzten zehn Jahren in den Häusern implementiert wurden, um die Bereiche Controlling, Accounting, Risikomanagement und das regulatorische Reporting zu unterstützen (Curko et al. 2007; Gopalakrishnan and Damanpour 2000; Narayanan 2008). Die hierzu eingesetzten analytischen Systeme ermöglichen einen effizienten Zugriff auf die Daten zur direkten Entscheidungsunterstützung. Die hierzu erforderliche benutzerorientierte Aufbereitung und die Erstellung der Reports wird in Praxis häufig mit dem Begriff „Business Intelligence“ bezeichnet (Watson and Wixom 2007).

Trotz der inzwischen langjährigen Erfahrung bei der Entwicklung von DWH-Lösungen, ist die Realisierung immer noch mit erheblichen Projektkosten verbunden und viele Projekte scheitern (Hwang et al. 2004). Die Gründe hierfür liegen dabei in einer nicht mehr handhabbaren Komplexität, die ihre Ursachen u.a. in schlechter Datenqualität in den Quellsystemen, politischen Konflikten hinsichtlich der Verantwortlichkeiten oder der Komplexität der Legacy Systemlandschaft haben kann (Watson et al. 2004; Weir et al. 2003; Wixom and Watson 2001).

Die größte Herausforderung beim Aufbau von DWH-Lösungen stellt die Datenintegration dar; d.h. die technische und vor allem fachliche Vereinheitlichung von Daten aus verschiedenen Quellen (March and Hevner 2007). So wechselt die exakte Bedeutung eines Datums üblicherweise in Abhängigkeit des jeweiligen Kontextes, welcher zwischen Firmen, Fachabteilungen und IT-Systemen variieren kann. Dies kann sich z.B. unterschiedliche Namenskonventionen oder die Verwendung unterschiedlicher Maßeinheiten niederschlagen. Aufgrund der spezifischen Gegebenheiten bei Finanzdienstleistern sind hier die kontextuellen Unterschiede besonders komplex, da Begriffe nicht standardisiert verwendet werden und es so zu unterschiedlichen semantischen Bedeutungen kommt (Behrmann and Räkers 2008). Hieraus ergibt sich bei der Datenintegration im FDL-Umfeld eine besonders hohe semantische Heterogenität (March and Hevner 2007). Für die Begrifflichkeiten und die Datenintegration existiert kein etablierter Standard, die Definition und Nutzung von Begriffen erfolgt unterschiedlich von Bank zu Bank, Fachabteilung zu Fachabteilung oder zwischen den einzelnen Endanwendern. So wird z.B. unter „Limit“, „Produkt“ oder „Buchwert“ von den Fachabteilungen „Vertrieb“, „Risikomanagement“ und „Controlling“ jeweils etwas Unterschiedliches verstanden. Die Entwicklung und Etablierung einer übergreifenden Standardisierung der Begriffe im FDL-Umfeld ist zusätzlich problematisch, weil die Geschäftsmodelle und die angebotenen Produkte einem schnellen Wandel unterliegen, um Wettbewerbsvorteile neu zu generieren oder zu sichern (Corcho et al. 2005). Aufgrund der Immaterialität der Produkte ist für die Einführung neuer Produkte die Abbildung in Prozessen und IT-Lösungen der einzige limitierende Faktor, was zu einer – im Vergleich zur verarbeitenden Industrie – erheblich verkürzten „Time to Market“ führt. Daher ist davon auszugehen, dass die semantische Heterogenität im FDL-Sektor unverändert hoch bleiben wird.

Ein weiterer Faktor, der die Datenintegration bei FDL-Unternehmen erschwert, ist die Verfügbarkeit von Datenquellen. Die aktuellen Systemlandschaften sind häufig historisch gewachsen und die Datenquellen somit heterogen, unterschiedlich organisiert oder schwer bis gar nicht zugreifbar (Zhu et al. 2004). Im Ergebnis sind auch die aktuell vorzufindenden DWH-Lösungen Individualentwicklungen, die mit hohen Entwicklungs- und Betriebskosten verbunden sind (Gray and Israel 1999; Inmon 1997). Hieran haben bislang weder die zunehmende Konzentration durch die Konsolidierung im FDL-Markt noch der erweiterte Einsatz von „Standard-Kernbanksystemen“ etwas geändert. Für eine erfolgreiche Datenintegration ist somit der Einsatz von Experten, die sich im jeweiligen Fach- und IT-Kontext auskennen, essentiell: im Entwicklungsprozess muss anhand der fachlichen Anforderungen über die Wichtigkeit des jeweiligen

Informationsbedarfs entschieden und anschließend die Anforderung mit dem korrespondierenden Informationsangebot auf Systemseite abgeglichen werden (March and Hevner 2007).

Obwohl die Datenintegration allgemein als eine Problemstellung von hoher Relevanz angesehen wird (March and Hevner 2007; Rizzi et al. 2006), haben bisher nur wenige Wissenschaftler dieses Thema eingehend behandelt (Skoutas and Simitsis 2006; Vassiliadis et al. 2005). Dies ist vor allem vor dem Hintergrund bemerkenswert, dass die Datenintegration, also das Laden der Daten in das Data Warehouse schon bei klassischen Anwendungsbereichen einen erheblichen Teil der Umsetzungszeit (ca. 80%) und -kosten (ca. 50%) in Anspruch nimmt (Kimball and Caserta 2004; Vassiliadis et al. 2002; Watson and Wixom 2007). Der genannte Aufwand dürfte im FDL-Umfeld noch höher ausfallen, da sich die überdurchschnittliche semantische Heterogenität insbesondere auf die Komplexität bei der Datenintegration auswirkt.

Auch aktuelle Entwicklungsansätze des Agile Data Warehousing lösen das Problem nicht vollständig, wenngleich die Notwendigkeit einer höheren Interaktion und einer intensiveren Kommunikation in der Praxis grundsätzlich erkannt wurde (Corr and Stagnitto 2011). Die bisher vorliegenden Ansätze bieten für die beschriebenen Umsetzungshürden keinen zufriedenstellenden theoretischen Lösungsansatz auch wurde die Problematik bisher nicht aus der kommunikationstheoretischen Perspektive angegangen.

Genau hier setzt die vorliegende Arbeit an und widmet sich damit einem Thema, das zunehmend an Aufmerksamkeit in der Forschung zu Prozessen und in der Entwicklung von Informationssystemen gewinnt (z.B. Corvera Charaf et al. 2013; Holten and Rosenkranz 2011; Mastrogiacomo et al. 2014; Mingers and Willcocks 2014; Rosenkranz et al. 2013): in einem interdisziplinären Ansatz werden unter Anwendung von sprachtheoretischer Ansätze Artefakte entwickelt, deren Anwendung einen positiven Einfluss auf den Projekterfolg hat. Aus den oben beschriebenen Problemen lassen sich drei wesentliche Ziele für ein verbessertes Vorgehen in DWH-Projekten bei FDL ableiten:

1. Entwicklung eines gemeinsamen Verständnisses von der semantischen Bedeutung der verwendeten Begriffe und der korrespondierenden Datenfelder
2. Sicherstellung einer engen und effizienten Kommunikation zwischen allen am Projekt beteiligten Parteien

3. Ermöglichung einer Früherkennung von Fehlern bei der Konzeption oder der Implementierung der Datenintegration, die sich aus Missverständnissen ergeben (Kommunikationsdefekte)

In der Summe werden durch diese Ziele eine effiziente Datenintegration und die Minimierung des nachträglichen Anpassungsbedarfs zur Fehlerbehebung in der DWH-Lösung erreicht.

Aufgrund der hohen Komplexität und des organisatorischen Fokus des Forschungsgegenstandes sowie einem starken Einfluss der spezifischen Fähigkeiten der beteiligten Personen auf den Projekterfolg, wurde die Forschung nach dem „Design-Science“-Ansatz (Hevner et al. 2004) durchgeführt: Das Vorgehensmodell und das in den Projekten genutzte Tool wurden als wissenschaftliche Artefakte entwickelt. Das theoretische Fundament für diesen Ansatz bildet sprach- und kommunikationstheoretische Ansätze, insbesondere die integrative Theorie der Sprache von Clark (1996), die auf die DWH-Entwicklung übertragen wurde. In DWH-Entwicklungsprojekten als „gemeinsame Aktivität“ stellt die Fähigkeit der Beteiligten zur gemeinsamen Koordination und zur Erreichung eines gemeinsamen Verständnisses ihren „Common Ground“ dar. Der Common Ground bildet dabei die Menge an Wissen, Ausbildung, Erziehung und Annahmen, die zwischen den Beteiligten geteilt wird (Clark 1992; Clark 1996; Clark and Brennan 1991). Der hier vorgestellte Ansatz unterstützt gezielt die Entwicklung eines breiten Common Ground zwischen den Beteiligten. Hierzu wird die gemeinsame Einführung, Diskussion und Verhandlung von fachlichen und technischen Begriffen inkl. ihrer exakten Bedeutung und einer entsprechenden Dokumentation gefördert. Obwohl das Vorgehensmodell im Kontext der Entwicklung von DWH-Lösungen für FDL entwickelt wurde, lassen sich die Erkenntnisse auch auf andere Domänen mit vergleichbarer semantischer Heterogenität übertragen.

2 Überblick und Struktur der Arbeit

Die vorliegende Dissertation besteht aus vier Veröffentlichungen, die aufeinander aufbauen und sukzessive die Entwicklung des Theoriebeitrags darstellen:

Die erste Veröffentlichung gibt eine Einführung in die Problemstellung und die Besonderheiten bei der DWH-Entwicklung im FDL-Umfeld. Unter Nutzung explorativer Fallstudien wird die besondere semantische Heterogenität in FDL-DWH-Projekten aufgezeigt (Behrmann and Räkers 2008). Hierbei werden anhand von drei Fallstudien von DWH-Projekten bei verschiedenen Banken insgesamt fünf Kernthesen aufgestellt, die die Spezifität von DWH-Projekten bei Finanzdienstleistern herausstellen:

- Aufgrund der Silo-Strukturen und der geringen fachliche Integration bei Banken treffen in den Projekten (häufig erstmalig) Personen aus unterschiedlichen fachlichen Domänen und Expertengebieten aufeinander und müssen zu einer gemeinsamen Sicht gelangen.
- Aus den Silo-Strukturen und einer geringen Vereinheitlichung existierender Begrifflichkeiten ergibt sich eine hohe semantische Komplexität. Dies wird bei Finanzdienstleistern zudem noch durch eine immaterielle und dynamische Produktwelt verstärkt.
- Die Probleme in den beobachteten Projekten haben gezeigt, dass primär spezifikations-basierte Vorgehensmodelle nicht erfolgreich sind. Es werden vielmehr Vorgehendmodelle benötigt, die gezielt Wissenstransfer unterstützen und einen hohen Interaktionsgrad aufweisen.
- Diese Besonderheiten von DWH-Projekten bei Finanzdienstleistern sind in den aktuell existierenden Vorgehensmodellen nur unzureichend berücksichtigt.

Aus diesen Spezifika und der identifizierten Lücke bei den bestehenden Vorgehensmodellen und Verfahren resultiert die Motivation für diese Arbeit.

Die Analyse der Beispielprojekte hat eine hohe Anzahl von Missverständnissen in der Kommunikation aufgezeigt. Um für diese Fälle gezielt Abhilfe zu schaffen, wurde in der zweiten Veröffentlichung zum besseren Verständnis dieser immer wieder auftretenden Kommunikationsdefekte unter Anwendung der Sprachtheorie von Kamlah und Lorenzen die „Mikroperspektive" betrachtet und einzelne Kommunikationsdefekte kategorisiert (Hoffmann and Behrmann 2008). Hierbei meint die Mikroperspektive die Betrachtung und Systematisierung einzelner Wissenstransfer-Vorgänge sowie der Frage, mit welchen Strategien eine erfolgreiche Transaktion sichergestellt werden kann. Hieraus wurden insgesamt sechs Voraussetzungen abgeleitet, die für einen erfolgreichen Wissenstransfer gelten müssen. Als Schlussfolgerungen werden eine höhere „Investition" in eine verbesserte Spezifikation und ein höherer Interaktionsgrad empfohlen, statt erheblich höhere Kosten für die Beseitigung der entstandenen Missverständnisse in Kauf zu nehmen.

Aufbauend auf den Erkenntnissen der „Mikroperspektive“ der Kommunikation wurde die Betrachtung unter Anwendung der Sprachtheorie von Clark und unter Nutzung einer erweiterten Fallstudie auf die „Makroperspektive“ erweitert (Behrmann 2009)). Der Kommunikationsprozess wurde anhand der Theorien zum Common Ground ganzheitlich betrachtet und es wurde auf Basis von vier Thesen die entscheidende Rolle direkter Face-to-face-Kommunikation herausgearbeitet:

- Die Antizipation des Adressatenwissens für die Kommunikation ist kritisch. Hierbei kommt es gerade bei einer Experten-Laien-Kommunikation, wie sie in den Financial DWH-Projekten aufgrund der verschiedenen involvierten Parteien üblich ist, schnell zu einer Überschätzung des Empfängerhorizontes. Im Ergebnis versteht der Laie dann die übermittelten Inhalte gar nicht oder ordnet sie in den falschen Kontext ein.
- Face-to-face-Kommunikation ermöglicht aufgrund des direkten verbalen und vor allem auch non-verbalen Feedbacks ein direktes Erkennen und Korrigieren von Kommunikationsdefekten.
- Ohne Face-to-face-Kommunikation können FDWH-Projekte nicht erfolgreich sein – rein spezifikations-orientierte Vorgehensmodelle sind gescheitert.
- Face-to-face-Kommunikation ist proaktiv. Sie kann durch die hohe Effizienz in der Kommunikation ein hohes Maß an gemeinsamem Verständnis sicherstellen und somit zukünftige Probleme im Vorfeld vermeiden.

Ausgehend von diesen Erkenntnissen aus der Übertragung der Kommunikationstheorie auf die DWH-Entwicklung wurden als Ergebnis drei Basisstrategien erarbeitet, die die Grundlage für die Entwicklung einer ersten Version der Artefakte bildeten:

1. Verbesserung der Antizipation des Adressaten-Wissens
2. Verbesserung der Verständlichkeit der Spezifikation
3. Unterstützung bei der Erkennung von Kommunikationsdefekten

Die Ergebnisse und die daraus abgeleiteten Strategien in den ersten drei Veröffentlichungen bildeten die Grundlage für die Entwicklung der Artefakte. Die Erkenntnisse aus der Anwendung der Sprachtheorien und der umfassenden Betrachtung der Projekte in Fallstudien waren Grundlage zur Entwicklung eines Vorgehensmodells in Verbindung mit einem Tool als Artefakte gemäß des Design-Science-Ansatzes. In einer ersten empirischen Prüfung wurde die positive Auswirkung auf DWH-Projekte bei FDL bei Verwendung des Artefaktes belegt (Rosenkranz et al. 2010). In der vierten Veröffentlichung wurden die beiden Artefakte zunächst detailliert beschrieben. Drei weitere Projekte, in denen Teile oder die vollständigen Artefakte zur Anwendung kamen, wurden als Fallstudien aufbereitet und ausgewertet. Durch den Einsatz von strukturierten Interviews mit den involvierten Mitarbeitern und einem Vergleich der insgesamt sechs Fallstudien konnte ein Nutzenbeitrag der Artefakte belegt werden.

Ausgehend von den bereits veröffentlichten Arbeiten wird nun noch einmal ein Gesamtüberblick über die Forschung gegeben: Im Folgenden wird zunächst ein kurzer Überblick über die vorhandene einschlägige Literatur gegeben. Die Forschungsarbeit wurde eng an den Leitlinien des Design Science Research ausgerichtet, was im Kapitel 4 aufgezeigt wird. Anschließend wird der als zentrale Basis genutzte theoretische Rahmen vorgestellt und erläutert, wie die sprachtheoretischen Erkenntnisse auf die Entwicklung von DWH-Lösungen übertragen werden können. In Kapitel 6 wird anschließend dargestellt, wie aus dem theoretischen Fundament die Entwicklung der Artefakte erfolgte und eine kurze Beschreibung des entwickelten Vorgehensmodells und Tools gegeben. Die Validität und der Nutzen der Artefakte wird anhand von verschiedenen Case Studies und einer Umfrage belegt, deren Ergebnisse in Kapitel 7 vorgestellt werden. Am Ende erfolgt die Bewertung der Forschungsergebnisse, der Limitierung und es wird ein Ausblick auf weitergehende Forschungsansätze gegeben.

3 Forschungskontext

Vorgehensmodelle zur Data Warehouse-Entwicklung

In zahlreichen Studien wurden die Wichtigkeit der Ermittlung von Anforderungen sowie einer Informationsbedarfsermittlung in DWH-Projekten aufgezeigt (Watson et al. 2004; Wixom and Watson 2001), die im Ergebnis dokumentiert, welche Daten wem zu welchem Zeitpunkt für die Unterstützung welcher Entscheidung bereitgestellt werden müssen. Dementsprechend wurden in den vergangenen Jahren verschiedene DWH-Vorgehensmodelle entwickelt, die sich explizit mit der Identifikation und der formalisierten Dokumentation des Informationsbedarfs befassen. Hierbei kann zwischen den in den klassischen Standardwerken zu findenden Vorgehensmodellen von Praktikern (Ballard et al. 1998; Inmon 2005; Kimball and Caserta 2004; Kimball and Ross 2002) und wissenschaftlichen Ansätzen (Böhnlein and Ulbrich-vom Ende 1999; Cavero et al. 2001; Chenoweth et al. 2003; Giovinazzo 2000; Golfarelli et al. 1998; Golfarelli and Rizzi 1998; Holten 2003; Mazón and Trujillo 2008; Moody and Kortink 2001; Trujillo et al. 2001; Tryfona et al. 1999) unterschieden werden. Die meisten der wissenschaftlichen Ansätze beschreiben Modellierungsansätze zur logischen Beschreibung des Informationsbedarfs in einer multidimensionalen Umgebung und haben sich als zu komplex für die praktische Anwendung erwiesen (Malinowski and Zimányi 2008). Jeder der aktuell existierenden Ansätze ist auf einen spezifischen Anwendungsbereich fokussiert und weist in anderen Bereichen Schwächen auf. Deshalb hat sich in der Praxis keiner der bisherigen Ansätze breiter etablieren können (Abell et al. 2001; Mazón and Trujillo 2008). Diese stecken somit noch „in den Kinderschuhen“ (March and Hevner 2007). Darüber hinaus folgen alle Vorgehensmodelle einem wasserfallorientierten Ansatz mit einem starken Fokus auf einer formalisierten Dokumentation, wie z.B. Spezifikationen und Datenmodelle (Couture 2014).

Lösungsansätze für die Datenintegration

Die Datenintegration stellt immer noch die Hauptherausforderung bei der Realisierung von DWH-Lösungen dar, was sich insbesondere auf das Design und die Umsetzung der Datenverarbeitungsprozesse Extraktion, Transformation und Laden (ETL) auswirkt (Vassiliadis et al. 2005). Als mögliche Lösung für Datenintegrationsprobleme wird häufig auf die Methode des „Schema-Matching“ verwiesen (Bernstein et al. 2004; Do and Rahm 2007; Ehrig and Staab 2004; Madhavan et al. 2001; Rahm and Bernstein 2001; Shvaiko and Euzenat 2005). Das Schema-Matching basiert auf der Entdeckung von Beziehungen zwischen ähnlichen Elementen in verschiedenen Schemata, z.B. aus unterschiedlichen Quellsystemen. Bei der Anwendung des Schema-Matching in Umgebungen mit hoher semantischer Heterogenität, wie im FDL-Umfeld, erweist sich dieses Vorgehen als zeitintensiv und fehleranfällig: auch beim Einsatz von halbautomatischen Verfahren kann die Korrektheit der Matches erst nach der Bewertung durch Experten bewertet werden (Saleem et al. 2008). Zudem erfordert die Problemstellung beim Aufbau der DWH-Lösung die Integration vieler verschiedener Datenquellen, wozu ein paarweiser Abgleich, wie er normalerweise erfolgt, nicht ausreicht.

Als ein weiteres Konzept zur Lösung des Integrationsproblems und des Schema-Matchings werden Ontologien angesehen (Campbell and Shapiro 1995; Simperl and Tempich 2006; Uschold et al. 1998). Da Ontologien reich an Semantik, computerlesbar und erweiterbar sind, eigenen sie sich auch zur Abbildung komplexer Sachverhalte. Sie können also dazu verwendet werden, die Integration heterogener Datenquellen zu beschreiben – vorausgesetzt, es ist einmal eine robuste Ontologie etabliert. Aktuell existieren diverse (inhaltlich unterschiedliche) Ontologien für den FDL-Sektor (Mäkelä et al. 2007) und es ist nicht ersichtlich, dass sich eine Standard-Ontologie herausbilden wird. Die Problematik der aktuell bereits gegeben heterogenen Struktur des FDL-Sektors wird durch die anhaltende Dynamik durch ständig neue Geschäftsmodelle, Produkte und sich ändernde gesetzliche Rahmenbedingungen vielmehr noch verstärkt (Corcho et al. 2005).

Kommunikationstheorie in IS

Obwohl die große Bedeutung von Wissenstransfer, Abstimmungen und Kommunikation für die Entwicklung von Informationssystemen hinreichend beschrieben wurde (Hansen and Lyytinen 2010; Levina and Vaast 2005; Robillard 1999; Tan 1994) und trotz der Erkenntnis, dass es einen Bedarf für eine Theorie und die Erforschung der Kommunikations-, Abstimmungs- und Lernenprozesse (Kautz et al. 2007) gibt, wurde die Sprach- und der Kommunikationstheorie nur von einzelnen Autoren auf die Domäne der Entwicklung von Informationssystemen angewendet (isnbes. bei: Auramäki et al. 1992; Clarke 2001; Corvera Charaf and Rosenkranz 2010; Flores et al. 1988; Goldkuhl and Lyytinen 1982; Holmqvist 1989; Schoop 2001).

Dagegen fußt ein breiter Forschungsbereich auf dem Gebiet der „human computer interaction“ (HCI) und „computer supported collaboration work“ (CSCW) (Kanda et al. 2004; Maglio et al. 2002; McFarlane and Latorella 2002; Olson and Olson 2000) auf psycholinguistischer Theorie, die zum Teil auf die Sprachtheorie von Clark zurückgeht (Clark 1992; Clark 1996; Clark 2005; Clark and Brennan 1991; Clark and Krych 2004). Obwohl die Theorien von Clark nicht bedingungslos geteilt werden (Koschmann and LeBaron 2003; Nova et al. 2008) findet sein Konzept des Common

Ground z.B. im Bereich der computerbasierten Kommunikation breite Akzeptanz (Carroll et al. 2006; Convertino et al. 2009; McCarthy et al. 1991).

Anders als auf den Themenfeldern der HCI und CSCW wurde die Theorie von Clark in der Forschung über Informationssysteme wenig angewendet. Beispiele hierfür finden sich nur in den Themengebieten des Wissenstransfers in virtuellen Teams (Alavi and Tiwana 2002), in der Verwendung von IT-Systemen in Organisationen (Sjöström and Goldkuhl 2005) oder in sprachbasierter Evaluation von Informationssystemen (Ågerfalk 2004).

Notwendigkeit einer interdisziplinären Sichtweise

Die bisher veröffentlichten Arbeiten liefern somit keinen umfassenden Ansatz zur Beherrschung einer hohen semantischen Komplexität entlang des gesamten Entwicklungsprozesses für Data Warehouses. Zudem werden die Konzepte von Clark nicht auf die Domäne der Informationssystem- oder DWH-Entwicklung angewendet und die DWH-Entwicklung auch nicht als ein Kommunikationsprozess gemäß des in der Arbeit genutzten Verständnisses betrachtet. Dagegen stützt sich der im Rahmen der Arbeit entwickelte Beitrag auf die Kommunikationstheorie. Damit wird eine etablierte Theorie auf das beschriebene Problemfeld einer Unterstützung für die DWH-Entwicklung im FDL-Umfeld übertragen (Vaishnavi and Kuechler 2008).

Da es sich bei der Datenintegration nicht primär um ein IT- oder Formalisierungs-Problem handelt, sondern für den Erfolg von DWH-Projekten eine Einigung auf einen projektindividuellen Standard entscheidend ist, der die im jeweiligen Kontext benötigten fachlichen und technischen Begriffe adäquat beschreibt, wurde ein interdisziplinärer Lösungsansatz gewählt. Für die erfolgreiche Realisierung von DWH-Lösungen im FDL-Umfeld kommt es entscheidend darauf an, eine gemeinsame Sicht über die exakte Bedeutung der Begriffe zu erzielen, nicht nur über die Gleichheit der Bezeichnung (Behrmann and Räkers 2008; Benefelds and Niedrite 2004; Hoffmann and Behrmann 2008).

Zusammenfassend ergibt sich zur Verbesserung der Vorgehensmodelle und Methoden bei der Entwicklung von DWH-Lösungen im FDL-Umfeld ein Bedarf in zwei Punkten:

1. Berücksichtigung des Abstimmprozesses und der damit einhergehenden Diskussionen und Verhandlungen zwischen allen involvierten Parteien im Vorgehensmodell, um ein breites gemeinsames Verständnis über Anforderungen, Spezifikationen und Begriffe zu erlangen.
2. Unterstützung eines durchgängigen und konsistenten Managements der Anforderungen und der gemeinsam definierten Begriffe entlang des Entwicklungsprozesses (Erstellen, Speichern, Aktualisieren).

4 Forschungskonzept: Design Science

Ausgehend von der beschriebenen organisatorischen Problemstellung beim Aufbau von Data Warehouses bei Finanzdienstleistern wurde für das Forschungsdesign ein Vorgehen gemäß des Design-Science-Ansatzes gewählt (Hevner et al. 2004). Ausgangspunkt ist ein beobachtetes, real existierendes Praxisproblem, für das Lösungsansätze identifiziert werden sollten in einem komplexen organisatorischen Umfeld, bei dem der Erfolg stark von menschlichen Faktoren beeinflusst wird. Zudem lassen sich aufgrund der Individualität und Größe der für die Analyse der Fragestellung einzubeziehenden Projekte nur schwer einzelne Faktoren isolieren und eine künstliche Veränderung des Setups kommt nicht in Frage. Somit erfüllt die untersuchte Problemstellung die von Hevner et al. (2004, S. 81) genannten Kriterien für die Anwendung von Design-Science.

Die mit der Art der Problemstellung einhergehenden Problemen bei der objektiven Messung von Einflüssen wurden gezielt berücksichtigt und die bei Gregor and Hevner (2013) für die Entwicklung und Validierung der Artefakte angeführten Verfahren wurden angewendet. Unter Nutzung von explorativen Fallstudien wurden teilweise neue Problembereiche identifiziert und mit der Übertragung der Kommunikationstheorie auf die Domäne der IS-Forschung neue Lösungen (Artefakte) entwickelt. Somit stellt das Ergebnis gemäß der Klassifikation von Goes (2014) mit einer Mischung aus „Improvement“ und „Exaptation“ einen relevanten Beitrag zur Forschung dar. Das Vorgehen wurde konsequent an den sieben Leitsätzen für Design-Science ausgerichtet und die hier vorgestellten Forschungsergebnisse decken diese vollständig ab:

1. Design as an artifact

Zur Lösung der identifizierten Problemstellungen wurden im Rahmen der verschiedenen Studien drei Artefakte entwickelt (Rosenkranz et al. 2010). Erstens ein formalisiertes Template zur Informationsbedarfsermittlung, zweitens ein Vorgehensmodell, in dem Abläufe, Organisation und Zusammenarbeitsweisen definiert werden und drittens ein Software-Tool, in dem das Template umgesetzt ist, das die im Vorgehensmodell beschriebenen Prozesse und Methoden unterstützt und für die verschiedenen Stakeholder-Gruppen unterschiedliche Sichten zur Darstellung der Inhalte anbietet. Zudem wird durch das Tool die frühzeitige Analyse von Echtdaten unterstützt und der jeweilige Datenqualitätsstatus nachvollziehbar. Das Vorgehen zur Ableitung sowie die Artefakte selbst werden in Kapitel 6 beschrieben.

2. Problem relevance

Die Motivation für die Entwicklung der Artefakte entstand aus real beobachteten Problemstellungen im Rahmen der Entwicklung von Data Warehouses für Finanzdienstleister. Diese weisen eine hohe semantische Komplexität auf, der gezielt durch Vorgehensmodelle und Methoden begegnet werden muss (Behrmann and Räkers 2008). Die Artefakte wurden in verschiedenen Praxisprojekten entwickelt, angewendet und der positive Effekt auf den Erfolg der Projekte grundsätzlich bestätigt. Somit weisen die untersuchte Problemstellung und die Lösungsansätze eine hohe Relevanz auf. Die grundsätzliche Motivation und der Bedarf zur Lösung des Problems wurde in Kapitel 1 bereits erläutert. Die Einordnung in verwandte Forschungsarbeiten und Literatur sowie die Darlegung des immer noch bestehenden Forschungsbedarfs wurde im vorherigen Kapitel bereits dargestellt.

3. Research rigor

Bei der Entwicklung des Ansatzes wurden ausgewählte Konzepte der Software- und DWH-Entwicklung in Verbindung mit der Kommunikations- und Sprachtheorie interdisziplinär verknüpft. Hierzu wurden auf Basis der Beobachtungen und der theoretischen Grundlagen Strategien abgeleitet (Behrmann 2009; Hoffmann and Behrmann 2008). Die Entwicklung des Lösungsansatzes geht also von einem soliden theoretischen Fundament aus. Für die Identifikation der Problemstellung und für die Evaluation der Artefakte kamen Case Study Research und strukturierte Interviews zum Einsatz (Rosenkranz et al. 2010). Das verwendete theoretische Fundament wird in Kapitel 5 und das Vorgehen zur Ableitung der Artefakte in Kapitel 6.1 genauer erläutert.

4. Design as a search process

Ausgehend von den Erkenntnissen der explorativen Fallstudien (Behrmann and Räkers 2008) sowie der theoretischen Grundlagen (Behrmann and Räkers 2008; Hoffmann and Behrmann 2008) wurden in einem iterativen Prozess die Artefakte entwickelt. In mehreren Stufen wurden diese erweitert, ergänzt, angepasst und wieder in Projektkonstellationen erprobt (Rosenkranz et al. 2010). Das Vorgehen zur Ableitung der Artefakte auf Basis der aus den Fallstudien und der Theorieanalyse entwickelten Strategien ist in Kapitel 6.1 beschrieben. Die iterative Vorgehensweise und das stetige Weiterentwickeln wird aus der Beschreibung der Fallstudien in Kapitel 7.2 deutlich.

5. Design evaluation

Für die Bewertung des Designs und der Konfiguration und zur Identifikation von Verbesserungen während des iterativen Designprozesses wurden multiple Case Studies eingesetzt. Es wurden dabei jeweils die Validität und der Nutzen der drei Kern-Artefakte als Hauptkriterien geprüft. Es wurde in multiplen Case Studies beschrieben, beobachtet und evaluiert, ob der Ansatz beim Umgang mit der hohen semantischen Komplexität hilfreich war. Darüber hinaus wurde die finale Konfiguration durch die Anwendung von strukturierten Interviews ebenfalls im Hinblick auf Validität und Nutzen geprüft (Rosenkranz et al. 2010). Eine zusätzliche Evaluierung unter Anwendung einer Fragebogen-basierten Umfrage wurde bereits durchgeführt und bestätigt der Ergebnisse aus den strukturierten Interviews (Rosenkranz et al. (under review)). Zur Verbreiterung der Evaluierungs-Basis sind weitere Untersuchungen in besser kontrollierbaren Umgebungen, was eine bessere Vergleichbarkeit gewährleitstet, in Planung. Die ausführliche Beschreibung der angewendeten Evaluationsverfahren ist in Kapitel 7 zu finden.

6. Research Contribution

Der Hauptbeitrag zur Forschung liegt in der Kombination der beiden bisher getrennt betrachteten Bereiche Softwareentwicklung und Kommunikationstheorie. Die sprachliche Koordination und Abstimmung zwischen Menschen spielt die entscheidende Rolle für eine erfolgreiche Kommunikation, die für den Erfolg der analysierten DWH-

Entwicklungsprojekte essenziell ist. Anhand der entwickelten Artefakte wird deutlich, dass eine gezielte Unterstützung der kritischen Kommunikationssituationen in den Projekten eine positive Wirkung auf den Gesamterfolg hat. Der Forschungsbeitrag und die Limitierungen werden im Detail in Kapitel 8 erläutert.

7. Research Communication

Während der Erarbeitung der Forschungsergebnisse wurden die Zwischenschritte im Rahmen eigener Veröffentlichungen auf Konferenzen präsentiert und diskutiert – diese Paper bilden nun auch die Basis für die kumulative Dissertation. Darüber hinaus fand permanent ein intensiver Austausch innerhalb des Forscherteams statt und eine weitere Veröffentlichung, die die Gesamtergebnisse der Gruppe darstellt, befindet sich aktuell in Vorbereitung. Durch das gemischte Team aus Universitäts-Mitarbeitern und in den Projekten tätigen Beratern war stets ein enger Austausch zwischen der Praxis- und der Theorieperspektive gewährleistet. Die gemischten Projektteams, die teilweise aus IT-, Fach- und Projektmanagementexperten bestanden, ermöglichten eine kontinuierliche Kommunikation und Rückmeldung aus IT-technischer- und Management-Sicht. Details zur Zusammensetzung der Teams in den jeweiligen Projekten und zu den im Rahmen des Evaluationsprozesses interviewten Personen sind in Kapitel 7.1 zu finden.

5 Theoretisches Fundament

5.1 DWH-Entwicklung als Kommunikationsprozess

Aufgrund des breiten Scopes, der umfassenden Größe und der heterogenen IT-Architektur sind in der DWH-Entwicklung im FDL-Umfeld zahlreiche verschiedene Stakeholder und Experten involviert. Verstärkt wird dieser Effekt durch eine häufig etablierte interne Organisation entlang fachlicher Silos. Zum Aufbau einer integrierten DWH-Lösung müssen dann zahlreiche Fachexperten aus den einzelnen Silos, IT-Experten für die Liefersysteme, alle relevanten Entscheidungsträger und aufgrund der Größe der Projekte üblicherweise auch noch externe Ressourcen in das Projekt involviert werden. Jede dieser Gruppen verfügt dabei über spezifisches Wissen, das bei der DWH-Entwicklung berücksichtigt werden muss (March and Hevner 2007).

Den Ausgangspunkt für die Entwicklung von DWH-Lösungen bilden die fachlichen Anforderungen, also der fachliche Informationsbedarf. Nach der Festlegung der Grobinhalte (Scope) erfolgt eine sukzessive Detaillierung der fachlichen Anforderungen und zunehmend der Einbezug der IT-Perspektive zur Ableitung des logischen Designs. Dazu müssen die Fach- und IT-Experten zu einem gemeinsamen Verständnis über die umzusetzenden Anforderungen gelangen, es handelt sich also um eine Problemstellung der Erlangung gemeinsamen Wissens („mutual kowledge") (Cramton 2001). Die DWH-Entwicklung kann daher im Wesentlichen als Kommunikationsprozess zwischen den verschiedenen involvierten Stakeholdern verstanden werden (Behrmann 2009; Behrmann and Räkers 2008).

5.2 Die Kommunikationstheorie von Clark

Eine wesentliche Grundlage für die Forschungsarbeit bildet die Kommunikationstheorie von Clark. Aus den Arbeiten von Clark und seinen Fachkollegen ergeben sich drei grundsätzliche Prinzipien für den Kommunikationsprozess (Clark 1992; Clark 1996; Clark and Brennan 1991; Clark and Krych 2004):

Prinzip 1: Kommunikation ist eine gemeinsame Aktivität (Joint Acitivity)

Kommunikation ist schlicht eine gemeinsame Aktivität, in der die konventionelle Sprache eine prominente Rolle spielt (Clark 1996, S. 58). Um gemeinsam Dinge zu erledigen, ist eine Koordination erforderlich und es bedarf der Kommunikation, die notwendige Koordination zu erreichen (Clark 2005). Eine Aktivität, wie die DWH-Entwicklung, wird zu einer Joint Activity wenn, eine Koordination in zwei Ebenen erfolgt: Erstens die des Inhaltes, also dessen was die Beteiligten intendieren zu tun und zweitens die des Prozesses, also des physischen und mentalen Systems, dass zur Erreichung der Intention verwendet wird (Clark 1996, S. 59). Joint Activities erfordern ein aktives Einbringen und eine permanente Verifikation bei allen Beteiligten. Sie lassen sich nicht ohne das Verständnis der allgegenwertigen Interaktion zwischen Inhalten und Prozess in Joint Activities erklären: Konversationen können ohne die Koordination von beidem, Inhalt und Prozess, nicht funktionieren (Clark 1996, S. 59, 319). Dabei sind Inhalt und Prozess voneinander abhängig: je komplizierter oder bedeutsamer der Inhalt, je länger der Prozess (Clark 1996, S. 60). Dies wird im Alltag z.B. deutlich, wenn eine Telefonnummer weitergegeben wird: da sich jeder kleine Fehler direkt auswirkt, wird der Prozess durch das Wiederholen der Ziffern verlängert, aber die Validität erhöht. Es kann dabei bei umfassenden Inhalten zusätzlich problematisch sein, dass ein kleines Missverständnis am Anfang im Laufe des Prozesses zu einem großen Missverständnis am Ende führen kann (Clark 1996, S. 235).

Prinzip 2: Kommunikation hängt vom „Grounding“ Prozess ab

Die Basis für die Kommunikation und Koordination ist geteiltes Wissen zwischen den handelnden Personen, das Clark als „Common Ground“ bezeichnet (Clark 1996, S. 120). Der Common Ground wird sukzessive auf Basis der Historie von Joint Actions zwischen den Kommunizierenden aufgebaut. Für den Erfolg des Ausbaus des Common Ground ist es entscheidend, das richtige Maß bei der Erweiterung zu wählen. Hierbei ist die Annahme über den mit dem Empfänger existierenden Common Ground entscheidend, der als Koordinationsbasis für die Erweiterung gilt (Clark 1996, S. 99). Der Sender muss bei der Mitteilung den Common Ground antizipieren und entsprechende Nachrichten formulieren. Für den Erfolg der Kommunikation – und somit des Groundings – ist die richtige Abschätzung entscheidend. Dabei kann der Common Ground in zwei Kategorien unterschieden werden: in den „prozessualen Common Ground“ und in den „inhaltlichen Common Ground“ (Clark and Brennan 1991). Der inhaltliche Common Ground enthält „Ich weiß, dass Du weißt, dass ich weiß was (…gemeint ist)“ und der prozessuale Common Ground „Ich weiß, dass Du weißt, das ich weiß wie (…etwas zu tun ist)“ (Convertino et al. 2009, S. 2340). Der inhaltliche Common Ground ist das gemeinsame Verständnis über das Thema und den Fokus der Arbeit als Ergebnis des inhaltlichen Austauschs und des wechselseitigen Signalisierens des Verstehens. Der prozessuale Common Ground besteh aus dem gemeinsamen Verständnis über die Regeln, das Vorgehen, das Timing und die Form, in der die Interaktion durchgeführt wird.

Prinzip 3: Kommunikation ist ein Prozess auf mehreren Ebenen

Die Kommunikation umfasst mehr als nur Wörter und verbale Kommunikation (schriftlich und mündlich). Schon der Begriff „Common *Ground*“ deutet an, wie wichtig die physische Co-Präsenz und ein geteiltes physisches Setting für die Erreichung eines gemeinsamen Verständnisses sind (Cramton 2001, S. 346). Auch die weniger im Fokus stehenden Formen der Kommunikationen, wie Gesten, Gesichtsausdrücke, Blicke und Posen helfen den Akteuren beim Austausch von Nachrichten (Clark 1996, S.

180; Clark and Brennan 1991; Clark and Krych 2004). Gerade die non-verbale Kommunikation gibt direkte Rückmeldung hinsichtlich des Verständnisses auf Seiten des Empfängers. Einschränkungen in den Kommunikationsebenen führen direkt zu einem spezifischen Kommunikationskontext, der auch als spezifische Beschränkung des Grounding („Grounding Contraints“) bezeichnet wird (Clark and Brennan 1991; Clark and Krych 2004). Hierbei ist die Wahl des Mediums entscheidend: je besser das Medium auf die Grounding Contraints ausgerichtet ist, umso besser ist es geeignet, eine Erweiterung des Common Ground zu unterstützen. Dies spiegelt sich auch in der HCI und CSCW Forschung wider, wo ebenfalls verschiedene Medien für unterschiedliche Kommunikationskontexte Anwendung finden.

5.3 Common Ground und Experten-Laien-Kommunikation

Die Theorien von Clark wurden von Bromme und seinen Mitarbeitern in zwei Szenarien angewendet und validiert (Bromme 2000; Bromme and Jucks 2001; Bromme et al. 2005a; Bromme et al. 2005b; Bromme et al. 1999; Jucks et al. 2008), die sich auch in der DWH-Entwicklung wiederfinden:

Experten-Laien-Kommunikation

Kommunikation mit einem starken Wissensunterschied zwischen Sender und Empfänger (Experten-Laien-Kommunikation) ist bei der DWH-Entwicklung im FDL-Umfeld permanent gegeben. Nicht nur in offensichtlichen Situationen, in denen sich Fach- und IT-Experten austauschen müssen, die jeweils im anderen Fach Laien sind, ist das Szenario beobachtbar. Dies gilt häufig auch für den Austausch zwischen den Fachabteilungen, weil sich aufgrund der starken fachlichen Spezialisierung Wissens-Silos herausgebildet haben. Experten-Laien-Kommunikation zeichnet sich durch einen geringen Common Ground zwischen den Akteuren aus, welcher in der Regel auf einen abweichenden Bildungs- und Erfahrungshintergrund zurückgeht. In diesen Fällen muss der Common Ground häufig nicht nur – wie bereits bei Clark beschrieben – sukzessive erweitert, sondern in Teilen auch restrukturiert werden. Die Empfänger ordnen sonst die Informationen in einen falschen Kontext ein und es entstehen somit Missverständnisse. Zusätzlich verstärkt wird dieser Effekt durch eine bei den Akteuren üblicherweise zu findende egozentrierte Verzerrung: wenn ich etwas weiß, erwarte ich mit einer höheren Wahrscheinlichkeit, dass andere dies auch wissen (Clark 1996, S. 111).

Die erschwert die korrekte Antizipation des Common Ground, da ein bewusster Perspektivwechsel erforderlich ist, um den Wissenstransfer erfolgreich zu gestalten: der Experte muss sich in die Lage des Laien versetzen und eine Annahme über sein vorhandenes Wissen treffen.

Im Experten-Laien Szenario ist dieser Perspektivwechsel aufgrund systematischer Unterschiede zwischen beiden Perspektiven problematisch. Systematisch heißt in diesem Zusammenhang, dass nicht nur Wissenselemente in der Laien-Perspektive fehlen, sondern dass das vorhandene Wissen auch noch in einen kognitiven Bezugsrahmen eingebettet ist, der wesentlich durch den Hintergrund und die spezifische Bildung geprägt ist (Bromme et al. 2004). Der kognitive Bezugsrahmen ist, insbesondere bei Laien, in gewisser Weise resistent gegenüber Veränderungen, da die Äußerungen des Experten in den unpassenden Kontext eingeordnet werden, ohne eine eigentlich notwendige strukturelle Veränderung des Bezugsrahmens auszulösen. Dies führt im Ergebnis häufig zu einer „Illusion of Evidence“ (Bromme et al. 2005a), in der der Experte die Verständlichkeit seiner Kommunikation überschätzt und der Laie die Information scheinbar verstanden hat (Bromme and Jucks 2001). Diese Kommunikationsdefekte sind aufgrund des vermeintlichen Verstehens schwer zu identifizieren und zu heilen.

Kommunikation ohne direkte physische Co-Präsenz

Gerade in großen DWH-Projekten mit zahlreichen involvierten Parteien und einer Verteilung auf unterschiedliche regionale Standorte, wie in Bankkonzernen üblich, wird die Kommunikation aus Kostengründen hauptsächlich über schriftliche Medien abgewickelt. Bei dieser Art der Kommunikation ist eine direkte Rückmeldung der Empfänger z.B. durch Gesten oder verbale Intervention nicht verfügbar. Bedingt durch höhere Hürden (z.B. eine schriftliche Antwort zu verfassen oder extra einen Anruf zu tätigen) wird ein Großteil der spontanen Rückmeldungen gefiltert. Durch diesen Umstand steigt das Risiko von Missverständnissen und Illusion-of-Evidence-Fällen stark an, sodass Im Ergebnis ein fehlendes gemeinsame Verständnis nicht bewusst wird (Bromme and Jucks 2001).

5.4 Implikationen für die DWH-Entwicklung

Überträgt man zunächst die von Clark formulierten Prinzipien auf die DWH-Entwicklung, folgt aus dem ersten Prinzip (Kommunikation ist eine Joint Activity), dass das gesamte Projektteam hinsichtlich der Bedeutung von fachlichen Begriffen und der Datenfelder ein gemeinsames Verständnis erlangen muss, um die Datenintegration erfolgreich umzusetzen (Inhalt). Hierzu müssen die fachlichen und technischen Analysen gemeinsam unter Einbeziehung fachlicher und technischer Experten durchgeführt und am Ende in einer integrierten Spezifikation über die Datenversorgung dokumentiert werden (Prozess). Das zweite Prinzip (Kommunikation hängt vom Grounding-Prozess ab) und Clarks zentrales Konzept des Common Ground haben die größten Implikationen für die DWH- und IS-Entwicklung. Das Schaffen eines gemeinsamen Verständnisses und eines breiten Common Ground zwischen den Mitgliedern des Projektteams ist entscheiden, um ein korrektes Datenmapping und somit ein fehlerfreies ETL-Design umzusetzen. Um dieses Ziel zu erreichen, ist ein tiefes Verständnis der fachlichen Anforderungen auf Seiten der IT-Mitarbeiter entscheidend. Ebenso müssen die Teammitglieder aus den Fachabteilungen die Korrektheit des Designs bewerten können und auch dieses nachvollziehen. Um diesen breiten inhaltlichen Common Ground zu schaffen, muss sich das Team auf einen Prozess verständigen, der die gemeinsame Diskussion fördert und die Ergebnisse dokumentiert (prozessualer Common Ground). Hierbei stellt sich ein Lerneffekt ein: wenn die Gruppe weiß, wie sie funktioniert, hilft dies, spezifische Handlungen auszulösen, die die Erreichung des gemeinsamen Zieles fördern (Convertino et al. 2009). Aus dem dritten Prinzip (Kommunikation ist ein Prozess auf verschiedenen Ebenen) wird deutlich, dass die alleinige Verwendung von schriftlicher Spezifikation nicht ausreichend ist, um ein Problem wie die DWH-Entwicklung mit der ihr eigenen erheblichen semantischen Komplexität, zu lösen. Eine Reduktion auf die Schriftform führt dabei unweigerlich zu Missverständnissen in der Kommunikation, die zudem nur schwer identifiziert werden können. Hierbei schaffen auch formalisierte Formen, wie Datenmodelle oder ETL-Spezifikationen keine Abhilfe, da ein Großteil der Semantik (z.B. die genaue Bedeutung eines Datenfeldes) weiterhin in herkömmlicher Schriftform dokumentiert wird.

Die ausgeführten Erkenntnisse aus der Kommunikationstheorie haben wichtige Implikationen für das Vorgehen in DWH-Entwicklungsprojekten. Wenn man die Interaktion zwischen den involvierten Parteien im Projektteam als eine Form der Experten-Laien-Kommunikation ansieht, muss ein Vorgehensmodell gezielt die zu erwartenden Kommunikationsprobleme aufgreifen. Eine effiziente Erarbeitung und ein breiter Common Ground zwischen allen Beteiligten ist das Ziel, das mit dem entwickelten Ansatz verfolgt wird. Aufgrund der Rahmenbedingung ist im Projektverlauf mit einer

Vielzahl von Illusion-of-Evidence-Fällen zu rechnen, die durch das Einziehen expliziter Feedback-Loops zur Prüfung des richtigen gemeinsamen Verständnisses reduziert werden können (Bromme et al. 2004; Te'eni 2001). Zusätzlich kann das gemeinsame Verständnis durch eine gut strukturierte und möglichst für alle involvierten Parteien nachvollziehbare Dokumentation gefördert werden. Bei der Verwendung von schriftlicher Dokumentation z.B. (Spezifikationen, logische Datenmodelle, Tools, Datenbanken) muss eine erhöhte Aufmerksamkeit auf die Antizipation des Common Ground des Empfängers (Laien) gelegt werden. Zusätzlich müssen Rahmenbedingungen geschaffen werden, die die Abgabe von Feedbacks von den Beteiligten fördern, was im Ergebnis die Wichtigkeit für einen hohen Interaktionsgrad in den Projektteams und einen hohen Anteil an direkter Face-to-Face-Kommunikation verdeutlicht (Bromme et al. 2004).

Auf Basis der Analyseergebnisse lassen sich drei Strategien ableiten, um ein effizientes Entstehen des prozessualen und des inhaltlichen Common Ground zu ermöglichen und somit eine erfolgreiche DWH-Entwicklung sicherzustellen (Behrmann and Räkers 2008; Rosenkranz et al. 2010).

Strategie	Beschreibung
1. Verbesserung der Anforderungsspezifikation	Erweiterung der Spezifikation unter Nutzung von formalen Beschreibungen (z.B. Formularen) zur Strukturierung und Reduktion der Interpretationsmöglichkeiten, wo möglich. Kontinuierliche Verbesserung der Inhalte in den nicht formalisierbaren Bereichen entlang der Diskussion im Projekt. Dies gilt insbesondere für die Beschreibung der exakten Semantik von Begriffen.
2. Reduktion der Illusion of Evidence Szenarien	Hoher Grad an direkter Interaktion und ausreichende Berücksichtigung von Co-Präsenz Szenarien. Bilden von gemischten Teams und Definition von direkten Ansprechpartnern zum Abbau möglicher Feedback-Hürden.
3. Frühzeitiges Erkennen verbleibender Kommunikationsdefekte	Durch die ersten beiden Strategien lässt sich das Auftreten von Kommunikationsdefekten verringern, aber nicht verhindern. Zur Vermeidung hoher Korrekturkosten in einer späten Implementierungsphase müssen diese Defekte frühzeitig erkannt werden. Bei DWH-Entwicklung kann dies durch ein möglichst frühes Arbeiten mit Echtdaten in Verbindung mit verschiedenen Datenanalyse-Zyklen erfolgen.

Abbildung 1: Strategien zur Unterstützung der Common-Ground-Entwicklung

6 Verbesserter Ansatz zur DWH-Entwicklung

6.1 Herleitung und Zusammenspiel der Artefakte

Die Artefakte, die im Folgenden kurz vorgestellt werden, wurden iterativ entwickelt, d.h. sukzessive weiterentwickelt bis eine stabile Konfiguration entstand (Baskerville et al. 2009). Das Vorgehensmodell für die DWH-Entwicklung wurde kontinuierlich verbessert, indem die jeweiligen Versionen in Projekten validiert und begleitende Fallstudien analysiert wurden. Die direkte und begleitende Anwendung der Methoden führte zu Erweiterungen, Anpassungen und teilweise Re-Designs des Vorgehensmodells, was dem Leitsatz „design as a search process“ Rechnung trug (Hevner et al. 2004).

Zentraler Ausgangspunkt für das initiale Design der Artefakte waren die drei aus der Analyse der Kommunikationstheorie abgeleiteten Strategien (vgl. Abbildung 1):

1. Ein formalisiertes **Template zur Informationsbedarfsermittlung**
2. Ein **Vorgehensmodell**, in dem Abläufe, Organisation und Zusammenarbeitsweisen definiert werden.
3. Ein **Software-Tool**, in dem das Template umgesetzt ist, das die im Vorgehensmodell beschriebenen Prozesse und Methoden unterstützt und für die verschiedenen Stakeholder-Gruppen unterschiedliche Sichten zur Darstellung der Inhalte anbietet. Zudem wird durch das Tool die frühzeitige Analyse von Echtdaten unterstützt und der jeweilige Datenqualitätsstatus nachvollziehbar.

Eine ausführliche Beschreibung der Artefakte ist im Paper 4 der Dissertation zu finden (Rosenkranz et al. 2010). Eine Zusammenfassung des iterativen Zusammenspiels der einzelnen Bausteine des Ansatzes ist in Abbildung 2 dargestellt.

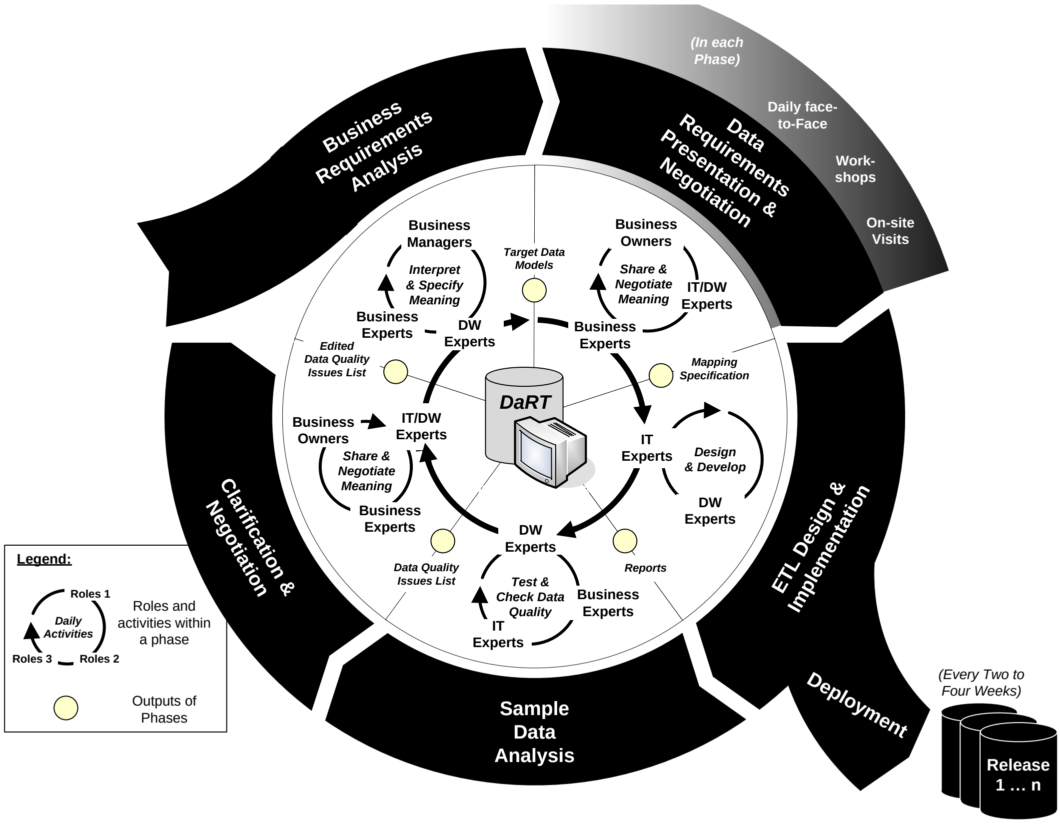

Abbildung 2: Überblick der durch Artefakte unterstützten FDWH-Entwicklung

Am Anfang eines DWH-Entwicklungsprojekts werden der Informationsbedarf und die daraus resultierenden Datenanforderungen von den Endbenutzern, also den fachlichen Entscheidungsträgern (Business Managers) und den Fachexperten (Business Experts) erhoben. Hierzu führen die DWH-Experten (DW Experts) Interviews und Workshops durch und dokumentieren die Anforderungen in strukturierter Form (Phase fachlicher Informationsbedarfsermittlung). Die Ergebnisse aus der Phase werden wie im domänen-spezifischen Template vorgegeben im DaRT erfasst und gespeichert. Auf Basis des ersten Entwurfes der Datenanforderungen erfolgt ein Konsistenzabgleich und die einzelnen Datenanforderungen werden in ein integriertes Zieldatenmodell überführt, das die zur Abdeckung des Informationsbedarfs benötigten Daten vollständig beschreibt.

Ausgehend vom Zieldatenmodell werden in der nächsten Phase die einzelnen Datenanforderungen präsentiert und mit den verfügbaren Informationen in den Quellsystemen abgeglichen. In gemeinsamen Workshops mit den fachlichen Systemverantwortlichen (Business Owners), IT- und DWH-Experten (face-to-face) werden so die Datenquellen und Mapping-Regeln abgestimmt. Hierbei wird das Zieldatenmodell schrittweise durch die „Business Experts“ vertreten und auf Basis der Erkenntnisse sukzessive angepasst und erweitert. Während der Workshops werden im Projektteam die Datenanforderungen diskutiert, die Beschreibungen in der Dokumentation begleitend angereichert und der jeweilige Stand der Abstimmung festgehalten. Auf diese Weise werden die Teammitglieder dazu gezwungen, im Fall von bestehenden Kommunikationsdefekten (und nur dann) deutlich mehr Hintergrundinformationen zu dokumentieren, um am Ende ein gemeinsames Verständnis zu erreichen und Transparenz über die kritischen Abstimmungsbereiche zu erzielen. Auswertungen über den Status der einzelnen Felder geben dem Projektleiter einen Überblick über den Gesamtstatus und den verbleibenden Restaufwand.

Im nächsten Schritt erstellen DWH-Experten auf Basis der im DaRT erarbeiteten Mapping-Spezifikation ein erstes ETL-Design (Phase ETL Design, Implementierung und Deployment). Auf Basis des Erst-Designs erfolgen eine sukzessive Detaillierung, die Klärung offener Punkte und die Verbesserung der Spezifikation. Dieses dynamische Vorgehen bei der ETL-Entwicklung erfordert eine zeitnahe Dokumentation zur Vermeidung von Inkonsistenzen und Missverständnissen zwischen den verschiedenen beteiligten Projektmitgliedern. Hierzu werden die Ergebnisse aus den Abstimmungen zwischen DWH- und Quellsystemexperten direkt im DaRT dokumentiert. Im Anschluss an das Deployment werden so schnell wie möglich Beispieldaten in das DWH geladen und im Rahmen der Phase „Analyse Beispieldaten“ geprüft. Die durch die Analyse identifizierten Probleme werden mit allen beteiligten Parteien besprochen und gehen nach „Klärung und Abstimmung“ in den nächsten Iterationszyklus ein.

6.2 Artefakt 1: Template zur Informationsbedarfsermittlung

Die Definition eines formalisierten Template für die Informationsbedarfsermittlung hilft den verschiedenen Stakeholdern, im hier vorliegenden Experten-Laien-Szenario alle relevanten Informationen zu explizieren. Im Unterschied zur Nutzung freier natürlicher Sprache ist so sichergestellt, dass alle relevanten Informationen abgefragt und notwendige Festlegungen gleich getroffen werden. Hierdurch wird der Interpretationsspielraum sowohl bei der Erstellung als auch beim späteren Lesen durch Dritte reduziert und somit die Entstehung eines Common Ground gefördert. Während der Diskussion wird durch die Nutzung des Template die notwendige Explizierung des Wissens der einzelnen Stakeholder durch das formalisierte Niederschreiben gefördert (Boisot and Canals 2004). Erst wenn versucht wird die Bedeutung von Dingen oder Konzepten zu formalisieren oder zu codieren, werden unterschiedliche Bedeutungen transparent und es kann eine Diskussion der für den Kontext korrekten Bedeutung erfolgen. Im Kontext der DWH-Entwicklung muss das Template alle notwendigen Informationen abdecken, die eine spätere Implementierung der ETL-Prozesse ermöglichen (Rizzi et al. 2006; Vassiliadis et al. 2005) und die zugrunde liegenden fachlichen Anforderungen beschreiben (z.B. Felddefinitionen für die Datenfelder), sowie alle weiteren Metainformationen zur Prozessteuerung berücksichtigen.

6.3 Artefakt 2: Vorgehensmodell

Zur Erreichung des prozessualen Common Ground ist das Wissen der Stakeholder über den beabsichtigen Kommunikationsprozess entscheidend. Dies war die grundlegende Motivation für die Ausgestaltung und Festlegung eines Vorgehensmodells, das insbesondere auf die Interaktion zwischen den Beteiligten fokussiert. Jüngere Forschung zum Themenfeld „Entscheidungen in Teams“ und „Zusammenarbeitstechnik“ hat den Schwerpunkt von „team mental models“ auf transaktive Modelle des Wissenstransfers verlagert: das Wissen über „wer weiß was im Team“. Diese Modelle scheinen passender für Gruppenaufgaben mit einem hohen Grad an Interaktion und Spezialisierung der Gruppenmitglieder (Convertino et al. 2009). Die zugrundeliegende Annahme für das Vorgehensmodell ist, dass das Wissen über die intendierten Kommunikationsverfahren, also die Expertise und die genaue Rolle der jeweiligen anderen Teammitglieder den Bedarf für unintendierte Kommunikation innerhalb des Teams deutlich reduziert. Hierfür regelt das entwickelte Vorgehensmodell die nach vorn gerichteten und die feedback-orintierten Kommunikationskanäle zwischen den einzelnen Rollen entlang der verschiedenen Entwicklungsphasen.

Dieser Ansatz folgt der Idee der transaktiven Modelle für den Wissenstransfer (Cooke et al. 2000): die Entstehung des inhaltlichen Common Ground in Entwicklungsprojekten für Financial DWHs kann durch Befähigung der Teammitglieder zur spezifischen Antizipation, welche Informationen zu welchem Zeitpunkt für eine erfolgreiche Zusammenarbeit geteilt werden sollten, gefördert werden (Convertino et al. 2009, S. 2341). Um dies zu unterstützen, sind für jede Phase des Vorgehensmodells konkrete

Ergebnisse formuliert. Diese können gezielt zur Entwicklung des inhaltlichen Common Ground beitragen, wenn sie als „Boundary Objects“ verwendet werden. Boundary Objects sind in diesem Zusammenhang Verbindungen über Wissensgrenzen hinweg und können dazu dienen, die Grenzen zwischen verschiedenen Expertengruppen zu überbrücken (Bergman et al. 2007, S. 662; Brown and Duguid 2001, S. 105; Pawlowski and Robey 2004, S. 209). Um aktiv ein Produktdesign unter Einbeziehung verschiedener Stakeholder, wie im Falle eines DWH-Projektes, z.B. durch das Hindeuten auf „erfolgreichen Fortschritt“ oder „drohende Probleme“ während des Designs, zu unterstützen, müssen Boundary Objects nach Bergman et al. (2007) vier Eigenschaften erfüllen: (1) Förderung einer geteilten Repräsentation, (2) Transformation des Design-Wissens, (3) Mobilisierung von Design-Aktionen und (4) Legitimierung des Design-Wissens (Bergman et al. 2007, S. 551). Die Autoren definieren Objekte, die diese vier Eigenschaften erfüllen als „Design Boundary Objects“ (DBO). Die im Rahmen des Vorgehensmodells definierten Ergebnistypen entsprechen diesen vier Voraussetzungen und stellen somit DBO dar (Rosenkranz et al. 2010). Das Modell folgt einem iterativen und agilen Entwicklungsansatz (Cao et al. 2009) und weicht somit von den traditionellen Wasserfall-orientierten Modellen ab, die aktuell häufig in DWH-Projekten Anwendung finden (Giorgini et al. 2008; Inmon 2005; Kimball and Caserta 2004; Mazón and Trujillo 2008; Moody and Kortink 2001).

6.4 Artefakt 3: Software Tool

Die Funktionen des Data Requirement Tool (DaRT) sind auf eine unmittelbare Unterstützung des Vorgehensmodells und der in Abbildung 1 formulierten Strategien ausgelegt. Durch die formalisierte Eingabemaske, die aus dieser resultierenden Festlegung konkreter Ergebnistypen und die permanente Dokumentation der Ergebnisse während der Abstimm-Diskussionen, unterstützt das DaRT direkt Strategie 1: Es reduziert Missverständnisse, fördert die Bildung des inhaltlichen Common Ground und verbessert kontinuierlich die Spezifikation. Während der engen Zusammenarbeit mit dem DaRT in den gemischten Teams ist eine Verbindlichkeit von allen Beteiligten und durch die direkte Dokumentation auch ihre Vollständigkeit und Korrektheit sichergestellt, was das Vorkommen von „Illusion of Evidence“-Fällen reduziert (Strategie 2). Durch die Anwendung des Vorgehensmodells in Verbindung mit dem DaRT werden die Projektbeteiligten praktisch zum direkten und persönlichen Austausch über mögliche Probleme und zu einer umfassenden Interaktion gezwungen. Das daraus resultierende direkte Feedback aus der Kommunikation hilft bei der Erkennung verbleibender Kommunikationsdefekte und der Reduktion von „Illusion of Evidence“-Fällen (Strategien 2 und 3). Durch das iterative Vorgehen und das frühe Arbeiten mit Echtdaten zur Kontrolle werden Kommunikationsdefekte, die trotz des deutlich erweiterten inhaltlichen Common Ground weiterhin bestehen, direkt sichtbar und können in der nächsten Iteration behoben werden (Strategie 3). Hierbei hilft die Versionierung und die Dokumentation der Abstimmhistorie, um schnell wieder in die weitere Abstimmung einsteigen zu können. Der Fortschritt der Abstimmung kann ebenfalls jederzeit aus dem DaRT reportet werden, sodass trotz des iterativen Ansatzes stets Transparenz über die noch ausstehenden Spezifikationsarbeiten besteht.

Mit dem DaRT als zentrales Repository für den Status und die Inhalte des ETL-Designs kann die Dokumentation allen Stakeholdern zugänglich gemacht werden. In Verbindung mit den Trackinginformationen können so frühzeitig bestehende Problemfelder identifiziert und den Fachexperten zur Klärung zugewiesen werden (Strategie 3). Über die bereits beschriebenen Inhalte können im DaRT auch Datenqualitätsregeln hinterlegt werden. Diese Regeln werden im Rahmen der ETL-Prozess-Implementierung direkt als technische Prüfregeln in die Datenbank übernommen. Im Rahmen verschiedener Projekte wurden bereits zahlreiche Regeln aus den unterschiedlichen Business Areas gesammelt, die in Folgeprojekten wiederverwendet werden können und eine frühzeitige Erkennung von Mappingfehlern fördern (Strategien 1 und 3).

7 Evaluierung und Weiterentwicklung der Artefakte

Nach den Grundsätzen des Design Science muss die gewollte (oder auch ungewollte) Wirkung von Artefakten geprüft und ihr Beitrag bei der Lösung des definierten Problems belegt werden. Dies kann den Vergleich zwischen ursprünglichen Zielen und beobachteten Ergebnissen bei Praxisfällen beinhalten (Hevner et al. 2004; Winter 2008).

Die entwickelten Artefakte wurden im Rahmen der Arbeit nach Kriterien wie Validität, Nutzen, Qualität oder Wirksamkeit evaluiert (Gregor and Hevner 2013, S. 351). In der Evaluierung der vorgestellten Artefakte wird hier auf die Validität und den Nutzen fokussiert und geprüft, ob die Artefakte funktionieren und die gesetzten Ziele erreichen (Validität). Erhoben wird auch, ob das Erreichen dieser Ziele einen positiven Einfluss außerhalb der Designumgebung bei der Anwendungen in der Praxis haben (Nutzen). Folgende These übersetzt die aus der Theorie abgeleiteten Design-Strategien in eine Erwartung in Bezug auf beobachtbare Auswirkungen in DWH-Entwicklungsprojekten und eignet sich somit für die Evaluation von Validität und Nutzen für die entwickelten Artefakte:

Die Anwendung der Artefakte unterstützt die Entwicklung des prozessualen und inhaltlichen Common Ground in FDWH-Entwicklungsprojekten. Somit, fördert sie die frühe Erkennung von Kommunikationsdefekten und Fehlern bei der Datenintegration (T1), verbessert deutlich die Projekt-Performance (den Entwicklungsprozess) (T2) und führt zu einer klaren Verbesserung der Qualität des Endproduktes (also des DWH) (T3).

Im Folgenden wird der Ansatz für die Case Study-basierte Analyse während der iterativen Weiterentwicklung und für die erste Evaluierung der Artefakte vorgestellt.

7.1 Evaluations- und Interviewdesign

Zur iterativen Überprüfung des Nutzens der Artefakte während der Entwicklungsphase kam ein „embedded multiple case study design“ (Yin 2003, S. 49) zur Anwendung. Hierzu wurden sechs Entwicklungsprojekte für Financial-DWHs in Banken ausgewählt, die typisch für die vergangenen Projekte im europäischen Bankensektor sind. Der Auslöser für sämtliche untersuchten Projekte waren die regulatorischen Anforderungen, die unter dem Namen „Basel II“ bekannt sind. Diese Anforderungen wurden in den meisten europäischen Ländern im Jahr 2007 in nationales Recht überführt und sind somit seitdem seitens der Banken zu erfüllen. Dies erfordert die Berechnung spezifischer Risikokennzahlen, sowie die Einführung von Risikomanagement- und -controllingprozessen. Eine Übersicht über die analysierten Cases und der zentralen Eckdaten ist in Abbildung 3 zu finden.

	Bank A	*Bank B*	*Bank C*	*Bank D*	*Bank E*	*Bank F*
Bilanzsumme (in Mrd. EUR)	100	70	7	4	2	230
Anzahl Länder in der Gruppe	> 13	5	1	1	2	>5
Anzahl Finanztöchter	18	6	-	-	-	15
Anzahl geleistete Projekttage	~13.500	~5.000	~5.000	~4.000	~7.000	~3.000
Regionaler Projektscope	EU	EU	DE	DE	DE	EU
Projektdauer (Jahre)	> 5	> 2,5	> 2	> 1,5	> 2	> 1,5
Anzahl Teammitglieder	> 100	> 150	> 20	> 25	> 15	> 25
Wechsel des Kernbankensystems	nein	ja	ja	nein	nein	nein
Komplexitäts-einschätzung[1]	sehr hoch	hoch	mittel	sehr hoch	mittel	hoch
Verfahren (eingesetzte Artefakte)	-	-	-	Vorgehen & Template	Vorgehen & DaRT Prototyp	Vorgehen & DaRT

Abbildung 3: Überblick und Kennzahlen der betrachteten Fallstudien

Sämtliche Projekte wurden von zeb begleitet, einer auf Finanzdienstleister spezialisierten Unternehmensberatung. Beim ersten Block (Banken A-C) handelt es sich um DWH-Projekte nach klassischen Vorgehensmodellen. Im zweiten Block (bei Banken D-F) kamen einzelne oder sämtliche Artefakte in unterschiedlichen Designstadien zum Einsatz. Die Analyse der gesammelten, unstrukturierten Basisdaten fand durch den Einsatz von Coding-Techniken und Checklisten statt.

Bereits etablierte „natürliche Kontrollen“ können für das Forschungsdesign genutzt werden (Lee 1989). Beispielsweise können durch das Einbeziehen von Personen, die an mehr als einem der betrachteten Fälle beteiligt waren, persönliche Faktoren für einen Projektvergleich konstant gehalten werden. Im vierköpfigen Wissenschaftlerteam, das die Fallstudienanalyse durchgeführt hat, befinden sich zwei Praktiker, die zum Zeitpunkt der Analyse über sieben Jahre als Berater für zeb tätig waren. Sie waren an allen drei Projekten der Banken A, B und C beteiligt. Im Anschluss waren sie auch am Design der Artefakte, jedoch nicht direkt an den Projekten bei Banken D, E und F beteiligt, in denen die Artefakte zur Anwendung kamen. Die anderen zwei Mitglieder des Wissenschaftlerteams waren an keinem der Projekte beteiligt. Während der Eva-

[1] Gemeinsam klassifiziert auf Basis einer Skala von 1 (sehr niedrig) bis 5 (sehr hoch) im Rahmen eines Kernteam-Treffens mit allen beteiligten Projektleitern.

luation der Artefakte agierten sie als Beobachter und haben die Interviews durchgeführt. Die Analyseergebnisse aus den Fällen D bis F wurden für die Weiterentwicklung des Designs verwendet.

Direkte Beobachtungen der zwei Praktiker des Teams, unstrukturierte und semi-strukturierte Interviews, strukturierte Selbsteinschätzungs-Umfragen unter den Projektmitgliedern, die Projektdokumentationen sowie ausgewählter Email-Verkehr wurden in einer Fallstudien Sammlung zusammengetragen. Es wurden Interviews sowohl mit Projektleitern als auch mit Projektmitarbeitern durchgeführt und zusätzlich wurde die Projektdokumentation zur Überprüfung und für das bessere Nachvollziehen der Interviewaussagen verwendet. So konnten z.B. Begriffe wie komplexe oder einfache DWH-Architektur anhand von konkreten Datenflussdiagrammen verdeutlicht werden. Der Zeitraum der Datenhebung erstreckte sich von März 2007 bis Dezember 2008. Am Ende wurden unter Verwendung Coding-Techniken und Checklisten die erhobenen Daten den einzelnen Thesen zugeordnet (Miles and Huberman 1994; Yin 2003).

Aufgrund der Tatsache, dass zum Zeitpunkt der Interviews die Projekte selbst oder direkte Nachfolgeprojekte noch liefen, wurde das Risiko für eine retrospektive Verfälschung oder das Vergessen von Teilaspekten in der Interaktion mit anderen minimiert. Für eine Steuerung des Interviewverlaufs durch den Interviewer trotz einer von der interviewten Person ausgehenden Gesprächsführung kam ein Interviewleitfaden zum Einsatz, der gezielt auf Beispiele und konkrete Situationen abstellt. Der Interviewleitfaden war den Kandidaten unbekannt und diente als Rahmen mit hohen Freiheitsgraden aber dennoch ausreichend Raum für Kontrollfragen. So wurde zum Beispiel nach Fällen von unzureichender Datenqualität, den Ursachen die dazu führten und den durchgeführten Gegenmaßnahmen gefragt. Die Interviews wurden teilweise telefonisch und teilweise in direkten face-to-face Interviews in den Büros von zeb von zwei Mitgliedern des Wissenschaftlerteams durchgeführt. Die Interviews dauerten zwischen 45 und 120 Minuten, wurden aufgezeichnet und anschließend transkribiert. Im Ergebnis entstand ein Interview-Manuskript mit ca. 130 Seiten. Im Anschluss an die Interviews wurden zu offenen Fragen und Unklarheiten Emails an die Kandidaten versandt, was die Möglichkeit für Feedback und Klarstellungen gab. Nach der Durchführung und Auswertung von 16 Interviews mit 11 Projektmitgliedern wurde gemeinsam festgestellt, dass eine ausreichende Datenverfügbarkeit für die meisten Bereiche erreicht war (Guest et al. 2006): die meisten Interviews waren sehr ergiebig, neue Aspekte traten nicht mehr auf und die Zielsetzung war die Prüfung der Validität und des

Nutzens der Artefakte. In ähnlichen Fällen, z.B. bei Anwendertests können bis zu fünf Benutzer bereits ausreichend sein (Nielsen and Landauer 1993). Am Ende wurde ein Treffen mit der Fokusgruppe durchgeführt, in dem über Gleichheiten und Unterschiede in den einzelnen Projekten diskutiert wurde. Die Diskussion wurde von zwei Mitgliedern des Wissenschaftlerteams moderiert.

In den Interviews wurde eine Variante der „Critical Incident“-Technik verwendet (Flanagan 1954). Die Interviews waren strukturiert, um die kritischen Fälle in DWH-Entwicklungsprojekten in Bezug auf Datenqualität, den Common Ground und Kommunikationsdefekte zu identifizieren. Das Konzept der Interviews und der Leitfaden wurden hinsichtlich der Reihenfolge der Fragen geprüft (Bouchard 1976; Mishler 1986). Obwohl die Interviews offen geführt wurden und es somit auch zu Abweichungen vom Leitfaden kam, wurden so wenig direkte Fragen wie möglich formuliert und lenkende Fragen vermieden (Loftus 1975).

In Abbildung 4 ist eine Übersicht über die geführten Interviews zu finden.

Fall	Name*	Details	Rolle im Projekt
Bank A	Christian	Männlich, 27 Jahre, 4 Jahre Erfahrung als Berater	Projektmitarbeiter
	Thomas	Männlich, 31 Jahre, 5 Jahre als Berater und Projektleiter	Projektmitarbeiter
	Michael	Männlich, 32 Jahre, 4 Jahre Erfahrung in der Linie, vorher als IT-Berater tätig	Projektleiter
	Nicole[a]	Weiblich, 46 Jahre, 17 Jahre Erfahrung als Berater und in der Linie	Projektmitarbeiter
Bank B	Jan	Männlich, 32 Jahre, 7 Jahre Erfahrung in der Linie	Projektmitarbeiter
	Hans[b]	Männlich, 40 Jahre, 8 Jahre Erfahrung als Berater und Projektleiter	Projektleiter
	Nicole[a]	s.o.	Projektmitarbeiter
Bank C	Lucas	Männlich, 29 Jahre, 3 Jahre Erfahrung als Berater	Projektleiter
	Daniel[c]	Männlich, 29 Jahre, 3 Jahre Erfahrung als Berater	Projektmitarbeiter
Bank D	Hans[b]	s.o.	Projektleiter
	Peter[d]	Männlich, 32 Jahre, 5 Jahre Erfahrung als Berater und Projektleiter	Projektleiter
	Stefan[e]	Männlich, 32 Jahre 4 Jahre Erfahrung und Berater und Projektleiter, 6 Jahre Erfahrung in der Linie	Projektmitarbeiter
Bank E	Peter[d]	s.o.	Projektleiter
	Wolfgang	Männlich, 28 Jahre, 2 Jahre Erfahrung als Berater	Projektmitarbeiter
Bank F	Peter[d]	s.o.	Projektleiter
	Daniel[c]	s.o.	Projektmitarbeiter
	Stefan[e]	s.o.	Projektmitarbeiter

*Alle verwendeten Namen sind Pseudonyme. [a] Mitarbeiter in Projekten bei Bank A und Bank B. [b] Mitarbeiter in Projekten bei Bank B und Bank D. [c] Mitarbeiter in Projekten bei Bank C und Bank F. [d] Mitarbeiter in Projekten bei Bank D, Bank E und Bank F. [e] Mitarbeiter in Projekten bei Bank D und Bank F.

Abbildung 4: Übersicht über die durchgeführten Interviews

7.2 Fallstudienanalyse der FDWH Projekte

Sämtliche Banken haben ihren Hauptsitz und die Tochtergesellschaften in Deutschland oder dem europäischen Raum. Die zugrundeliegenden Projekte haben alle das Ziel verfolgt, die regulatorischen Anforderungen gemäß Basel II umzusetzen und im

Ergebnis die geforderten Berichte an die Aufsicht zu übermitteln. Hierzu ist die Kalkulation neuer regulatorischer Kennzahlen erforderlich (insbesondere die risikogewichteten Aktiva (RWA)), die sowohl auf Einzelinstituts- als auch auf Gruppenebene ausgewiesen werden müssen. Für die Durchführung dieser Kalkulationen ist die Sammlung umfassender, konzernweiter Daten erforderlich, was den Aufbau zentraler DWHs in den Banken erforderlich machte und die Hauptaufgabe in den betrachteten Projekten darstellte. Der Fokus der zu integrierenden Daten lag dabei auf Geschäften der einzelnen Geschäftsarten, Sicherheiten, Kreditnehmer und Ratinginformationen. Um die Vollständigkeit der Daten zu gewährleisten, mussten auch Datenanlieferungen aus den Töchtern realisiert werden, was die Implementierung lokaler Extraktionsroutinen und die Transformation in das geforderte Anlieferungsformat erforderlich machte. Eine schematische Darstellung der Gesamtarchitektur, die in dieser Form in allen Projekten zur Anwendung kam, ist in Abbildung 5 dargestellt. Die in den Projekten realisierten Architekturen sind jedoch deutlich komplexer.

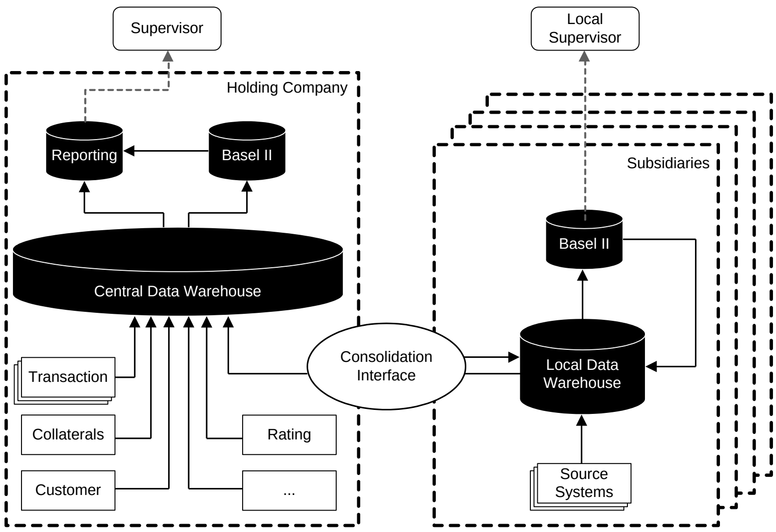

Abbildung 5: Schematische Architektur für Basel II-Datenkonsolidierung

Erwartungsgemäß wurde die semantische Komplexität innerhalb aller sechs Projekte als hoch bewertet. Unterschiedliche Bedeutungsinterpretationen zentraler IT- und fachbezogenen Konzepte waren in allen sechs Banken anzutreffen, es war also in allen Projekten zum Start ein geringer inhaltlicher Common Ground vorhanden.

Detailanalyse zum Rollout-Prozess bei Bank A

Das geringe Vorhandensein eines inhaltlichen Common Ground und die Wichtigkeit eines hohen Interaktionsgrads und einer face-to-face-Kommunikation wurde bei der

Detailanalyse des Rollout-Prozesses bei Bank A deutlich (Behrmann 2009). Erst nachdem nach mehreren Eskalationsschritten eine Workshop-orientierte Diskussion der Datenanforderungen in den jeweiligen Töchtern durchgeführt wurde, hat sich die Qualität der Datenanlieferung signifikant verbessert. Auf Basis der Beobachtungen wurden die grundsätzlichen Ansätze (entspricht den Strategien in Abbildung 1) für die Verbesserungen abgeleitet (Behrmann 2009):

- Verbesserung der Antizipation des Common Ground des Adressaten: Voraussetzung für die Erweiterung des Common Ground ist die korrekte Abschätzung des Wissens des Adressaten, sonst kommt es zu Missverständnissen oder „Illusion of Evidence“-Fällen (entspricht Strategie 2).
- Verbesserung der Verständlichkeit der Dokumentation: die Verständlichkeit der Spezifikation kann durch Weiterentwicklung und Verbesserung bis zu einem gewissen Grad erhöht werden, aber für sich alleine nicht die vollständige Semantik beschreiben (entspricht Strategie 1).
- Unterstützung bei der Erkennung von Kommunikationsdefekten: Eine späte Erkennung der noch bestehenden Kommunikationsdefekte führt zu einer späten Korrektur von Fehlern und zu überproportionalen Kosten für die Korrektur. Die einfachste Art, diese Kommunikationsdefekte zu erkennen, ist, eine Umgebung zu schaffen, die direkte Feedbackschleifen ermöglicht (Strategie 3).

Auf Basis derselben Fallstudie wurden die Voraussetzungen für eine erfolgreiche Kommunikation (also eine erfolgreiche Weiterentwicklung eines Common Ground) untersucht und mögliche Fälle für Kommunikationsdefekte systematisiert (Hoffmann and Behrmann 2008). Ausgehend von dieser Klassifikation wurden mögliche Gegenmaßnahmen formuliert, die gezielt in die Entwicklung der Artefakte eingeflossen sind. Haupterkenntnis war hierbei, dass neben einer besseren Formalisierung der Spezifikation eine (wenn auch teurere) direkte Interaktion zwischen den Projektteilnehmen sichergestellt werden muss.

Auf Basis dieser Ersthesen wurde der Untersuchungsbereich auf weitere Projekte ausgedehnt und die Fälle B und C in die Analyse mit einbezogen.

Vergleich aller Fälle

Aufgrund der unterschiedlichen Projektkonstellationen und den komplexen organisatorischen Rahmenbedingungen lässt sich die Performance der Projekte nicht durch Daten, wie beispielsweise Durchlaufzeiten, beobachtete Verzögerungen oder Budgets messen. Daher wurde für den Vergleich des Erfolges und der Effizienz der Projekte, also der Projektperformance, aus den sieben Key-Playern der Projekte eine Fokusgruppe gebildet. Sämtliche Teilnehmer dieser Gruppe waren in alle der sechs Projekte involviert und wurden zu einer zusammenfassenden Diskussion eingeladen, die zwei Ziele verfolgte: Erstens sollten noch einmal die einzelnen Projekte betrachtet und zweitens sollte auf Basis der Expertenmeinungen eine Rangfolge über die Performance der Projekte entwickelt werden. Aufgrund einer im Schnitt fünfjährigen Erfahrung und der direkten Beteiligung an zwei bis fünf Implementierungen von FDWHs, verfügten alle Teilnehmer der Fokusgruppe über ein umfangreiches Maß an Expertise.

Während des mit der Fokusgruppe durchgeführten Workshops wurde zunächst jedes der sechs Projekte durch das Forscherteam erläutert. Im nächsten Schritt nahmen alle Teilnehmer eine Bewertung der Projekte unter Berücksichtigung der Qualität des Prozesses, des Endproduktes und der Projekteffizienz auf einer Skala von 1-6 vor (Nidumolu 1995). Dabei wurde das Projekt, das aus Sicht des Experten im Vergleich die höchste Performance hatte mit 1 bewertet, das Projekt mit der schlechtesten mit 6. Sämtliche sechs Projekte mussten dabei in eine Reihenfolge gebracht werden – eine einheitliche Bewertung mehrerer Projekte war nicht zulässig, um eine vollständige Rangfolge zu erreichen. Die einzelnen Ergebnisse wurden im Anschluss im Rahmen des Workshops noch einmal reflektiert und in der Gesamtrunde besprochen. Das Ergebnis der Diskussion wurde in einer gemeinsam getragenen Gesamtreihenfolge zusammengefasst, die in Abbildung 6 zu finden ist.

Im Ergebnis ist festzustellen, dass alle sechs Projekte in Bezug auf den Scope und die Komplexität vergleichbar waren und ähnliche Fragestellungen behandelten. Die später durchgeführten Projekte, die einige oder alle Artefakte verwendeten, schneiden bei der Bewertung besser ab, als die ersten drei, bei denen zu Beginn weder das Vorgehensmodell noch das Tool zum Einsatz kamen. Der beobachtete Unterschied in der Performance wurde auch seitens der Experten zumindest anteilig auf den Einsatz der Artefakte zurückgeführt. Es kann nicht ausgeschlossen werden, dass hierbei auch Effekte aus der natürlichen Lernkurve eingingen, da sämtliche Projekte durch dasselbe Beratungsunternehmen begleitet wurden und sich im Zeitverlauf auch die Fähigkeiten der handelnden Personen aufgrund der Erfahrungen weiterentwickelten. Durch das Design der Interviews wurden diese Effekte abgemildert: Die Interviewpartner wurden sowohl unter nur an einem Projekt Beteiligten als auch unter Mitarbeitern, die in mehrere Projekte involviert waren, ausgewählt. Zudem wurden die Interviews teilweise in noch laufenden Projekten geführt. Doch auch diese „natürlichen Kontrollen" schließen eine Verzerrung der Ergebnisse nicht vollständig aus.

Fall	*Rang*	*Einsatz Template*	*Einsatz VGM*	*Einsatz DaRT*	*Beobachtete Effekte*
A	6	-	-	-	Illusion of Evidence und späte Fehlererkennung (geringer inhaltlicher Common Ground). Nach Umstellung des Vorgehensmodells auf interaktions- und kommunikationszentrierte Methoden ist der inhaltliche Common Ground entstanden.
B	5	-	-	-	Kommunikationsprobleme und späte Fehlererkennung (geringer inhaltlicher Common Ground). Wechsel auf kommunikationsintensiveres Vorgehen. Probleme wurden eher erkannt, als bei Bank A.
C	4	-	-	-	Illusion of Evidence und später Fehlererkennung (geringer inhaltlicher Common Ground). Das Vorgehen wurde schneller als bei Bank A und Bank B auf interaktions- und kommunikationszentriert umgestellt.
D	3	+ (in Excel)	+ (ohne Tool)	-	Frühe Bildung von prozessualem und inhaltlichem Common Ground. Frühe Erkennung von Fehlern, bessere Kommunikation und Projekt-Performance im Vergleich zu Banken A-C.
E	2	++ (im Prototyp)	++	+ (Prototyp)	Template und VGM förderten prozessualen und inhaltlichen Common Ground, der DaRT-Prototyp unterstützte den prozessualen Common Ground. Bessere Kommunikation und Projekt-Performance im Vergleich zu Banken A-D.
F	1	++ (im DaRT)	++	++	Template und VGM förderten prozessualen und inhaltlichen Common Ground, der DaRT Prototyp den prozessualen Common Ground. Fehler wurden früh erkannt. Bessere Kommunikation und Projekt-Performance im Vergleich zu Banken A-E.

Abbildung 6: Zusammenfassung der Fallstudien

7.3 Zusammenfassung der Validierung

Die Überprüfung von Validität und Nutzen stützte die Eingangsthese, dass die Verwendung der Artefakte bei der Bildung eines prozessualen und inhaltlichen Common Ground hilft, indem diese zu einer früheren Erkennung von Kommunikationsdefekten und Datenintegrationsfehlern führen, signifikant positiven Einfluss auf eine bessere Projektperformance und die Endproduktqualität haben. Die Erwartungen in Bezug auf das beobachtbare Verhalten in realen DWH Entwicklungsprojekten aus der Praxis wurden durch folgendes Vorgehen bestätigt: die betrachteten Fallstudien zeigten, dass die entwickelten Artefakte (Template, Vorgehensmodell und Software-Tool) zu einer früheren Erkennung von Kommunikationsdefekten und Datenintegrationsfehlern führen (unterstützt These T1). Darüber hinaus wurde aus den Fallstudien deutlich, dass durch die Artefakte die Projekt-Performance verbessert wurde, indem die notwendige Gesamtzeit für die Kommunikation der Anforderungen in den Entwicklungsprojekten reduziert werden konnte (unterstützt These T2). Schlussendlich wurde in den Fallstudien und in der Umfrage eine steigende Qualität des Endproduktes durch den Einsatz der drei Artefakte während des Entwicklungsprozesses bestätigt (unterstützt These T3). Obwohl die Validierungsergebnisse den positiven Einfluss der im Rahmen des Design-Science Ansatzes entwickelten Artefakte untermauern, hat die in der Praxis durchgeführte Validierung limitierende Faktoren. Bei den durchgeführten Fallstudien handelt sich nicht um einen kontrollierte und vollständige Überprüfung im Rahmen eines kontrollierten Setups, wie es in Projekten des Behavioral Research üblich ist, bei denen die Artefakte durch Dritte entwickelt wurden (Gregor and Hevner 2013, S. 351). Dieser Aspekt muss bei der Einschätzung der Forschungsergebnisse berücksichtigt werden.

8 Forschungsbeitrag, Limitierungen und weitere Forschungsansätze

Theorien zur Entwicklung von FDWH-Projekten unbefriedigend

Die vier übergreifenden Ziele für die Unterstützung des Entscheidungsprozesses mit Hilfe von Data Warehouses, die March und Hevner (2007) identifiziert haben – Integration, Implementierung, Intelligenz und Innovation – , erfordern eine enge Vernetzung und Interaktion zwischen den Mitgliedern in einem DWH Entwicklungsprojekt. Bis heute fehlen allerdings effektive Ansätze, um erstens den Informationsbedarf zu erheben und zweitens die erhobenen Anforderungen in ein konzeptuelles Modell mit einer gemeinsamen Sprache zwischen Fach- und IT-Experten zu überführen (Jarke et al. 2009; Rizzi et al. 2006).

Anwendung von Kommunikationstheorien in neuem Forschungsfeld

Der Hauptbeitrag der Dissertation zur Forschung liegt in der Kombination der beiden, bisher getrennt betrachteten Bereiche Softwareentwicklung und Kommunikationstheorie. Die sprachliche Koordination und Abstimmung zwischen Menschen spielt die entscheidende Rolle für eine erfolgreiche Kommunikation. Es gibt nun klare Anzeichen, dass diese Erkenntnisse auch für die Interaktion zwischen Computern und Menschen gelten und das diese natürlich auch zunehmend in größerem Umfang auftritt (Branigan et al. 2010). Trotzdem wurden die Kommunikations- und Sprachtheorie bisher nur selten in dieses Feld übertragen. Die Anwendung dieser Theorien in der Softwareentwicklung schafft neue Forschungsmöglichkeiten, zum Beispiel für die Empfängergerechte Kommunikation für alle Beteiligten in einem Softwareentwicklungsprozess (Experten-Laien-Kommunikation). Das Explizieren des Expertenwissens und die gezielte Unterstützung von Experten-Laien-Kommunikation, die ohne die Bildung eines Common Ground nicht funktionieren kann, sind vielversprechende Ansätze zur Anwendung in der IS-Forschung. Im Endeffekt öffnet die Übertragung der Kommunikationstheorie auf die Forschung zu Informationssystemen ein neues Feld für die Anwendung und Validierung dieser Konzepte. Das Beobachten und Nachverfolgen der Entstehung eines Common Ground kann sowohl Wissenschaftlern als auch Praktikern dabei helfen, die Diversität in Softwareentwicklungsprojekten zu verstehen, insbesondere die Abhängigkeit des Projekterfolges vom Wissen des Entwicklers und dem Zusammenhang zwischen Struktur und Softwareentwicklungs-Praxis (Kautz et al. 2007). Die Anwendung der Konzepte aus der Kommunikationstheorie kann eine große Hilfe dabei sein, die Entstehung von formalisierten, hilfreichen Anforderungen und die Mechanismen des Wissenstransfers während des Softwareentwicklungsprozesses besser zu durchdringen.

Gleichzeitig liefert die Arbeit aus der Analyse real durchgeführter DWH Entwicklungsprojekte in der Finanzdienstleistungsindustrie erste umfassende Erkenntnisse für die praktische Anwendung. Es wurde aufgezeigt, dass eine funktionierende Kommunikation keinen statischen Prozess darstellt, sondern permanente Aufmerksamkeit und Anpassung erfordert. Dies wiederum zeigt die Notwendigkeit für ein aktives Management der Kommunikationsprozesse auf, was sowohl die Auswahl und Anwendung eines passenden Vorgehensmodells als auch die richtige Zusammensetzung von Projektteams umfasst. Darüber hinaus zeigt die Evaluation deutlich, dass die entwickelten Konzepte eine direkte praktische Anwendbarkeit haben und einen Nutzenbeitrag bei der Entwicklung der DWHs haben. Sowohl das Vorgehensmodell als auch das DaRT werden weiterhin in DWH-Entwicklungsprojekten eingesetzt und permanent weiterentwickelt.

Erste Evaluierung untermauert positiven Beitrag der Forschungsergebnisse

Wie die Evaluierung bestätigt, werden durch den entwickelten Ansatz diese Probleme in Entwicklungsprojekten für Financial-DWHs mit einer hohen semantischen Komplexität erfolgreich aufgegriffen. Die Berücksichtigung und das Einbeziehen von Erkenntnissen, die aus der Sprachtheorie stammen, sind vielversprechend, um die Performance in FDWH Entwicklungsprojekten zu steigern und Missverständnisse bei der Explizierung und dem Transfer des Wissens zu vermeiden. Zwei Punkte sind dabei vor allem entscheidend:

- Die Etablierung von Verfahren zur Unterstützung der Bildung eines inhaltlichen Common Ground direkt zu Beginn der FDWH-Entwicklung ist essenziell für die Vermeidung und frühe Erkennung von Kommunikationsdefekten. Dies wiederum ermöglicht eine frühzeitige Erkennung und Vermeidung von Datenintegrationsfehlern, anstelle diese erst im Nachhinein zu lösen, wodurch der Anteil von aufwandsintensiven Tests im Nachgang und insbesondere auch das teure Fixing von Fehlern deutlich reduziert wird.
- Eine enge Führung der Projektteams während des Entwicklungsprozesses und somit die Bildung eines prozessualen Common Ground ist von zentraler Bedeutung. Dadurch lässt sich die konsequente Anwendung des Vorgehensmodells sicherstellen und somit die schnelle Bildung eines inhaltlichen Common Ground.

Der Ansatz hat demnach eine positive Wirkung auf die kritischen Erfolgsfaktoren Zeit und Kosten. Eine Gesamtbetrachtung der Kost-Nutzen-Effekte erfordert die Einbeziehung weiterer Faktoren, die auf den ersten Blick die Vorteile des entwickelten Vorgehens schmälern. Den Kern des Vorgehensmodells bilden eine hohe Interaktion und ein intensiver persönlicher Austausch als Ergänzung zu schriftlichen Spezifikationen. Die Umsetzung erfordert zum einen mehr Zeit, als der bloße Versand der Dokumentation und ist vor allem nur schwer skalierbar. Zudem erfordert die persönliche Interaktion, dass sich die Beteiligten an einem Ort befinden, was in internationalen Projekten zu steigenden Reisekosten führt. Der erhöhte Zeit- und Kostenbedarf durch das Präsentieren, Diskutieren und Abstimmen der Datenanforderungen hat sich auch in den vorgestellten Fallstudien gezeigt. Für eine Gesamtbetrachtung müssen dabei jedoch drei weitere Faktoren berücksichtigt werden. Ersten fällt in jedem DWH-Projekt ebenfalls ein Spezifikationsaufwand an, auch wenn die Anforderungen auf klassische Art und Weise erstellt werden – dieser Aufwand ist dem der Abstimmungsprozesse gegenüber zu stellen. Zweitens werden durch ein höheres Maß an inhaltlichem Common Ground Kommunikationsdefekte und somit Datenintegrationsfehler proaktiv vermieden. Es müssen also während der Testphase weniger Fehlerfälle behandelt und gefixt werden. Drittens werden durch die gezielte Erkennung von noch verbleibenden Kommunikationsdefekten auch die verbleibenden Fehler frühzeitig erkannt, was zu einer erheblichen Reduktion der Fixing-Kosten führt. Alle diese gegenläufigen Kosteneffekte kompensieren die zunächst hoch erscheinende Investition in die höhere Interaktion. Neben diesen Kosteneffekten innerhalb des Projekts ermöglicht die durch das Vorgehensmodell explizit berücksichtigte kontinuierliche Verbesserung der Spezifikation zum einen die Wartbarkeit der implementierten Lösungen und zum anderen wird ein hohes Maß an möglicher Wiederverwendung erreicht.

Bei den im Rahmen der Fallstudien analysierten Entwicklungsprojekten konnte eine durch die Anwendung des Vorgehensmodells evtl. vermutete Verzögerung nicht beobachtet werden. Die Durchführung der einzelnen Phasen inklusive Definition, Diskussion und Abstimmung der Anforderungen wurde in den Projekten so schnell durchlaufen, dass dies seitens der Projektmitglieder nicht als Verzögerung wahrgenommen wurde. Auf der anderen Seite wurden aber die „Illusion of Evidence“-Fälle und die Datenintegrationsprobleme nachhaltig reduziert. Insgesamt fühlten sich die Projektmitglieder durch den Ansatz nicht gebremst, sondern nahmen umgekehrt den Diskussionsfortschritt sogar beschleunigt wahr, da man sich stärker auf die Datenanforderungsdiskussion und das Erreichen eines gemeinsamen Verständnisses konzentrieren konnte, ohne zu stark über mögliche spätere Datenintegrationsfehler und Datenqualitätsprobleme nachdenken zu müssen.

Der vielversprechendste Aspekt des Ansatzes ist jedoch, dass die in allen Projekten anzutreffende semantische Heterogenität durch die intensive Kommunikation und die Entwicklung eines Common Ground zwischen den Beteiligten überwunden werden kann. In Bezugnahme auf die Anwendung von DBOs wie bei Bergman et al. (2007) dargestellt können wir bestätigen, dass der Einsatz von DBOs dabei hilft, Missverständnisse und Unsicherheiten, die aus der semantischen Heterogenität und komplexen Anforderungen resultieren, deutlich zu reduzieren. Bergman et al. beschäftigen sich intensiv mit der Frage, über welche Eigenschaften DBOs verfügen müssen, um als DBO zu wirken, geben aber keine praktischen Hinweise darauf, wie diese Eigenschaften in konkreten DBOs umgesetzt werden können. Im Rahmen der Forschungsarbeit wird hierzu ein konkretes Beispiel gegeben, wie DBOs entwickelt und innerhalb von Softwareentwicklungsprojekten eingesetzt werden können.

Limitierungen und weitere Forschungsansätze

Aus den Fallstudien konnte ein erstes Verständnis über die Vorteile des entwickelten Ansatzes abgeleitet werden, welches perspektivisch unter Verwendung einer größeren Stichprobe weiter erhärtet werden sollte. Aufgrund der Tatsache, dass die Fälle in der jeweiligen „natürlichen“ Praxisumgebung betrachtet wurden, kann nicht vollständig ausgeschlossen werden, das es andere – im Rahmen der Analyse nicht erfassten – Faktoren gibt, die nicht mit dem angewendeten Ansatz ursächlich im Zusammenhang stehen, aber die beobachteten Ergebnisse beeinflusst haben. Dieses mögliche Problem sollte in einer weiteren, vertiefenden Analyse verfolgt bzw. ausgeschlossen werden. Zusätzlich ist zu berücksichtigen, dass sich die hier vorgestellte Forschungsarbeit ausschließlich auf die Kommunikationsprozesse fokussiert. Andere Faktoren, wie Fähigkeiten der Anwender, Macht- und Kulturfragen, institutioneller Kontext oder Zielkonflikte können für den Projekterfolgt ebenfalls entscheidend sein und sollten demnach auch durch die Anwendung von DBOs oder anderen Konzepten explizit berücksichtigt werden.

Eine weitere offensichtliche Limitierung ergibt sich aus dem verwendeten Schema zur Festlegung der Reihenfolge der Projekt-Performance im Rahmen der Diskussion mit der Fokusgruppe. Natürlich basiert das Ranking auf einer dünnen empirischen Datenmenge, da es auf den Angaben von lediglich sieben Projektmanagern beruht. Dies zeigt die Notwendigkeit für vertiefende Studien, um ein höheres Maß an Sicherheit bei der Evaluation, z.B. durch die Durchführung von Experimenten zu erreichen. Bei der Betrachtung derart großer und komplexer Praxisprobleme ergibt sich jedoch immer das Auseinanderklaffen des aufwändigen Analyseverfahrens mit einer einer geringen Zahl von Fällen. Um diese Limitierung gezielt aufzugreifen wurde eine erste Umfrage zur Bewertung der stabilen Konfiguration durchgeführt (Rosenkranz et al. (under review)). Da beide Analysewege zu demselben Ergebnis führten, nämlich dass die Verwendung der Artefakte einen positiven Einfluss auf die Projekt-Performance besitzt, ist die Wahrscheinlichkeit der Richtigkeit der formulierten Thesen trotzdem sehr hoch.

Literaturverzeichnis

Abell, A., Samos, J., and Saltor, F. 2001. "A Framework for the Classification and Description of Multidimensional Data Models", *12th International Conference on Database and Expert Systems Applications (DEXA'01)*, Munich, Germany, S. 668-677.

Ågerfalk, P. J. 2004. "Investigating actability dimensions: a language/action perspective on criteria for information systems evaluation", in: *Interacting with Computers* (16:5), S. 957-988.

Alavi, M., and Tiwana, A. 2002. "Knowledge integration in virtual teams: The potential role of KMS", in: *Journal of the American Society for Information Science and Technology* (53:12), S. 1029-1037.

Auramäki, E., Hirschheim, R., and Lyytinen, K. J. 1992. "Modelling Offices Through Discourse Analysis: The SAMPO Approach", in: *Computer Journal* (35:4), S. 342-352.

Baldwin, L. P., Irani, Z., and Love, P. E. D. 2001. "Outsourcing Information Systems: Drawing Lessons from a Banking Case Study", in: *European Journal of Information Systems* (10:1), S. 15-24.

Ballard, C., Herreman, D., Schau, D., Bell, R., Kim, E., and Valencic, A. 1998. *Data Modeling Techniques for Data Warehousing*. San Jose, CA, USA: IBM Corporation.

Baskerville, R., Pries-Heje, J., and Venable, J. 2009. "Soft design science methodology", *4th International Conference on Design Science Research in Information Systems and Technology*, Philadelphia, PA, USA: ACM, S. 9.

BCBS. 2013. *Principles for effective risk data aggregation and risk reporting*. Basel, Switzerland: Bank for International Settlements, Basel Committee on Banking Supervision (BCBS).

Behrmann, W. 2009. "Hitting Interface Specifications -The Role of 'Common Ground' in Financial DWH Projects", *15th Americas Conference on Information Systems (AMCIS 2009)*, San Francisco, USA, S. 312.

Behrmann, W., and Räkers, M. 2008. "Specifics of Financial Data Warehousing and Implications for Management of Complex ISD Projects", *16th European Conference on Information Systems (ECIS 2008)*, Galway, Ireland, S. 1740-1751.

Benefelds, J., and Niedrite, L. 2004. "Comparison of approaches in data warehouse development in financial services and higher education", *6th International Conference on Enterprise Information Systems*, Porto, Portugal, S. 552-557.

Bergman, M., Lyytinen, K., and Mark, G. 2007. "Boundary Objects in Design: An Ecological View of Design Artifacts", in: *Journal of the Association for Information Systems* (8:11), S. 546-568.

Bernstein, P. A., Melnik, S., Michalis, P., and Quix, C. 2004. "Industrial-strength schema matching", in: *ACM SIGMOD Record* (33:4), S. 38-43.

Böhnlein, M., and Ulbrich-vom Ende, A. 1999. "Deriving Initial Data Warehouse Structures from the Conceptual Data Models of the Underlying Operational Information Systems", *Second ACM International Workshop on Data Warehousing and OLAP*, Kansas City, MO, USA, S. 15-21.

Boisot, M., and Canals, A. 2004. "Data, information and knowledge: have we got it right?", in: *Journal of Evolutionary Economics* (14:1), S. 43-67.

Bouchard, T. 1976. "Field Research Methods: Interviewing, Questionnaires, Participant Observation, Systematic Observation, Unobtrusive Measures", in:

Handbook of lndustrial and Organizational Psychology, M.D. Dunnette (ed.). Chicago, IL, USA: Rand McNally.

Branigan, H. P., Pickering, M. J., Pearson, J., and McLean, J. F. 2010. "Linguistic alignment between people and computers", in: *Journal of Pragmatics* (42:9), S. 2355-2368.

Bromme, R. 2000. "Beyond one's own perspective: The psychology of cognitive interdisciplinarity", in: *Practising interdisciplinarity*, P. Weingart and N. Stehr (eds.). Toronto, Canada: Toronto University Press, S. 115-133.

Bromme, R., and Jucks, R. 2001. "Wissensdivergenz und Kommunikation: Lernen zwischen Experten und Laien im Netz", in: *Partizipation und Interaktion im virtuellen Seminar*, F.W. Hesse and H.F. Friedrich (eds.). Münster, Germany: Waxmann, S. 81-103.

Bromme, R., Jucks, R., and Rambow, R. 2004. "Experten-Laien-Kommunikation im Wissensmanagement", in: *Der Mensch im Wissensmanagement: Psychologische Konzepte zum besseren Verständnis und Umgang mit Wissen*, G. Reinmann and H. Mandl (eds.). Göttingen, Germany: Hogrefe, S. 176-188.

Bromme, R., Jucks, R., and Runde, A. 2005a. "Barriers and Biases in Computer-Mediated Expert-Layperson-Communication", in: *Barriers and Biases in Computer-Mediated Knowledge Communication,* R. Bromme, F.W. Hesse and H. Spada (eds.). New York, NY, USA: Springer, S. 89-118.

Bromme, R., Jucks, R., and Wagner, T. 2005b. "How to refer to lsquodiabetesrsquo? Language in online health advice", in: *Applied Cognitive Psychology* (19:5), S. 569-586.

Bromme, R., Nückles, M., and Rambow, R. 1999. "Adaptivity and Anticipation in Expert-Laypeople Communication", *AAAI Fall Syposium on Psychological Models of Communication in Collaborative Systems*, North Falmouth, MA, USA.

Brown, J. S., and Duguid, P. 2001. "Knowledge and Organization: A Social-Practice Perspective", in: *Organization Science* (12:2), S. 198-213.

Campbell, A. E., and Shapiro, S. C. 1995. "Ontologic Mediation: An Overview", *IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing*, Montreal, Canada, S. 16-25.

Cao, L., Mohan, K., Peng, X., and Ramesh, B. 2009. "A framework for adapting agile development methodologies", in: *European Journal of Information Systems* (18:4), S. 332-343.

Carroll, J. M., Rosson, M. B., Convertino, G., and Ganoe, C. H. 2006. "Awareness and teamwork in computer-supported collaborations", in: *Interacting with Computers* (18:1), S. 21-46.

Cavero, J., Piattini, M., and Marcos, E. 2001. "MIDEA: A Multidimensional Data Warehouse Methodology", *3rd International Conference on Enterprise Information Systems (ICEIS'01)*, Setubal, Portugal, S. 138-144.

Chenoweth, T., Schuff, D., and Louis, R. 2003. "A method for developing dimensional data marts", in: *Communications of the ACM* (46:12), S. 93-98.

Clark, H. H. 1992. *Arenas of Language Use*. Chicago, IL, USA: Chicago University Press.

Clark, H. H. 1996. *Using Language*. New York, NY, USA: Cambridge University Press.

Clark, H. H. 2005. "Coordinating with each other in a material world", in: *Discourse Studies* (7:4-5), S. 507-525.

Clark, H. H., and Brennan, S. E. 1991. "Grounding in Communication", in: *Perspectives on Socially Shared Cognition*, L.B. Resnick, J.M. Levine and S.D. Teasley (eds.). Washington, D.C., USA: APA, S. 127-149.

Clark, H. H., and Krych, M. A. 2004. "Speaking while monitoring addressees for understanding", in: *Journal of Memory and Language* (50:1), S. 62-81.

Clarke, R. J. 2001. "Studies in organisational semiotics: an introduction", in: *Australasian Journal of Information Systems* (8:2), S. 75-82.

Convertino, G., Mentis, H. M., Rosson, M. B., Slavkovic, A., and Carroll, J. M. 2009. "Supporting content and process common ground in computer-supported teamwork", *27th International Conference on Human Factors in Computing Systems*, Boston, MA, USA: ACM, S. 2339-2348.

Cooke, N. J., Salas, E., Cannon-Bowers, J. A., and Stout, R. e. J. 2000. "Measuring Team Knowledge", in: *Human Factors* (42:1), Spring 2000, S. 151-173.

Corcho, O., Losada, S., Martínez Montes, M., Bas, J. L., and Bellido, S. 2005. "DIP: Data, Information and Process Integration with Semantic Web Services, WP10: Case study eBanking. Financial Ontology. DIP deliverable D10.3." Retrieved 2009-06-16, from http://dip.semanticweb.org/documents/D10.3.pdf.

Corr, L., and Stagnitto, J. 2011. *Agile data warehouse design: Collaborative dimensional modeling, from whiteboard to star schema*. Leeds, UK: DecisionOne Press.

Corvera Charaf, M., and Rosenkranz, C. 2010. "Natural Language Alignment As a Process – Applying Functional Pragmatics in Information Systems Development", *18th European Conference on Information Systems (ECIS 2010)*, Pretoria, South Africa.

Corvera Charaf, M., Rosenkranz, C., and Holten, R. 2013. "The emergence of shared understanding: applying functional pragmatics to study the requirements development process", in: *Information Systems Journal* (23:2), S. 115-135.

Couture, N. 2014. "Best Practices for Adopting Agile Methodologies for Data Warehouse Development", in: *Business Intelligence Journal* (18:2), S. 8-17.

Cramton, C. D. 2001. "The Mutual Knowledge Problem and Its Consequences for Dispersed Collaboration", in: *Organization Science* (12:3), S. 346-371.

Curko, K., Bach, M. P., and Radonic, G. 2007. "Business Intelligence and Business Process Management in Banking Operations", *29th International Conference on Information Technology Interfaces (ITI 2007)*, Cavtat, Croatia, S. 57 - 62

Do, H.-H., and Rahm, E. 2007. "Matching large schemas: Approaches and evaluation", in: *Information Systems* (32:6), S. 857-885.

Ehrig, M., and Staab, S. 2004. "QOM – Quick Ontology Mapping", in: *The Semantic Web – ISWC 2004. Proceedings of the Third International Semantic Web Conference, Hiroshima, Japan, November 7-11, 2004*, S.A. McIlraith, D. Plexousakis and F.v. Harmelen (eds.). Berlin, Germany et al.: Springer, S. 683-697.

Flanagan, J. C. 1954. "The critical incident technique", in: *Psychological Bulletin* (51:4), S. 327-358.

Flores, F., Graves, M., Hartfield, B., and Winograd, T. 1988. "Computer systems and the design of organizational interaction", in: *ACM Transactions on Information Systems* (6:2), S. 153-172.

Giorgini, P., Rizzi, S., and Garzetti, M. 2008. "GRAnD: A goal-oriented approach to requirement analysis in data warehouses", in: *Decision Support Systems* (45:1), S. 4-21.

Giovinazzo, W. 2000. *Object-Oriented Data Warehouse Design. Building a Star Schema*. Upper Saddle River, NJ, USA: Prentice Hall.

Goes, P. B. 2014. "Editor's comments: design science research in top information systems journals", in: *MIS Quarterly* (38:1), S. iii-viii.

Goldkuhl, G., and Lyytinen, K. 1982. "A language action view of information systems", *3rd International Conference on Information Systems (ICIS 1982)*, Houston, TX, USA.

Golfarelli, M., Maio, D., and Rizzi, S. 1998. "Conceptual Design of Data Warehouses from E/R Schemes", *31st Hawaii International Conference On Systems Science (HICSS 1998)*, Kohala Coast, HI, USA, S. 334-343.

Golfarelli, M., and Rizzi, S. 1998. "A Methodological Framework for Data Warehouse Design", *ACM First International Workshop on Data Warehousing and OLAP (DOLAP 1998)*, Washington, D. C., USA, S. 3-9.

Gopalakrishnan, S., and Damanpour, F. 2000. "The impact of organizational context on innovation adoption in commercial banks", in: *IEEE Transactions on Engineering Management* (47:1), S. 14-25.

Gray, P., and Israel, C. 1999. "The Data Warehouse Industry. Center for Research on Information Technology and Organizations, UC Irvine." *I.T. in Business* Retrieved 2010-10-15, from http://escholarship.org/uc/item/1hp1k5m7

Gregor, S., and Hevner, A. R. 2013. "Positioning and Presenting Design Science Research for maximum Impact", in: *MIS Quarterly* (37:2), S. 337-356.

Guest, G., Bunce, A., and Johnson, L. 2006. "How Many Interviews Are Enough?", in: *Field Methods* (18:1), February 1, 2006, S. 59-82.

Hansen, S., and Lyytinen, K. 2010. "Challenges in Contemporary Requirements Practice", *43th Hawaii International Conference on System Sciences (HICSS 2010)*, Koloa, HI, USA, S. 1-11.

Hevner, A., March, S., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research", in: *MIS Quarterly* (28:1), S. 75-105.

Hoffmann, A., and Behrmann, W. 2008. "Classification of Communication Defects in DWH Projects", in: *14th Americas Conference on Information Systems (AMCIS 2008)*. Toronto, Canada: S. 296.

Holmqvist, B. 1989. "Work, Language and Perspective: An Empirical Investigation of the Interpretation of Computer-Based Information Systems", in: *Scandinavian Journal of Information Systems* (1:1), S. 72-96.

Holten, R. 2003. "Specification of Management Views in Information Warehouse Projects", in: *Information Systems* (28:7), S. 709-751.

Holten, R., and Rosenkranz, C. 2011. "Designing Viable Social Systems: The Role of Linguistic Communication for Self-Organization", in: *Kybernetes* (40:3/4), S. 559-580.

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", in: *Decision Support Systems* (37:1), S. 1-21.

Inmon, B. 1997. "The Data Warehouse Budget. DM Review Magazine, January 1997." Retrieved 2010-10-15, from www.dmreview.com/master.cfm?NavID=55&EdID=1315

Inmon, W. H. 2005. *Building the data warehouse*, (4 ed.). Indianapolis, IN, USA: Wiley.

Jarke, M., Loucopoulos, P., Lyytinen, K., Mylopoulos, J., and Robinson, W. 2009. "High-Impact Requirements for Software-Intensive Systems: A Manifesto", *Perspectives Workshop: Science of Design: High-Impact Requirements for Software-Intensive Systems*, M. Jarke, K. Lyytinen and J. Mylopoulos (eds.), Dagstuhl, Germany: Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany.

Jucks, R., Becker, B.-M., and Bromme, R. 2008. "Lexical Entrainment in Written Discourse: Is Experts' Word Use Adapted to the Addressee?", in: *Discourse Processes* (45:6), S. 497-518.

Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H. 2004. "Interactive robots as social partners and peer tutors for children: a field trial", in: *Human-Computer Interaction* (19:1), S. 61-84.

Kautz, K., Madsen, S., and Nørbjerg, J. 2007. "Persistent problems and practices in information systems development", in: *Information Systems Journal* (17), S. 217–239.

Kimball, R., and Caserta, J. 2004. *The Data Warehouse ETL Toolkit, Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. Indianapolis, IN, USA: Wiley.

Kimball, R., and Ross, M. 2002. *The data warehouse toolkit: the complete guide to dimensional modeling*, (2. ed.). New York [u.a.]: Wiley.

Koschmann, T., and LeBaron, C. D. 2003. "Reconsidering common ground: examining Clark's contribution theory in the OR", *8th European Conference on Computer Supported Cooperative Work*, Helsinki, Finland.

Lee, A. S. 1989. "A Scientific Methodology for MIS Case Studies", in: *MIS Quarterly* (13:1), March 1989, S. 32-50.

Levina, N., and Vaast, E. 2005. "The emergence of boundary spanning competence in practice: implications for implementation and use of information systems", in: *MIS Quarterly* (29:2), S. 335-363.

Loftus, E. 1975. "Leading Questions and the Eye Witness Report", in: *Cognitive Psychology* (7:4), S. 560-572.

Madhavan, J., Bernstein, P. A., and Rahm, E. 2001. "Generic Schema Matching with Cupid", *27th International Conference on Very Large Data Bases (VLDB 2001)*, Rome, Italy.

Maglio, P. P., Matlock, T., Gould, S. J., Koons, D., and Campbell, C. S. 2002. "On understanding discourse in human-computer interaction", *24th Annual Conference of the Cognitive Science Society*, W.D. Gray and C. Schunn (eds.), Fairfax, VA, USA, S. 602-607.

Mäkelä, T., Rommel, K., Uskonen, J., and Wan, T. 2007. "Towards a Financial Ontology – A Comparison of e-Business Process Standards." Retrieved 2009-06-16, from http:////www.soberit.hut.fi/T-86/T-86.5161/2007/FinancialOntology_final.pdf.

Malinowski, E., and Zimányi, E. 2008. *Advanced Data Warehouse Design: From Conventional to Spatial and Temporal Applications*. Berlin, Germany: Springer.

March, S. T., and Hevner, A. R. 2007. "Integrated decision support systems: A data warehousing perspective", in: *Decision Support Systems* (43:3), S. 1031-1043.

Mastrogiacomo, S., Missonier, S., and Bonazzi, R. 2014. "Talk Before It's Too Late: Reconsidering the Role of Conversation in Information Systems Project Management", in: *Journal of Management Information Systems* (31:1), Summer2014, S. 44-78.

Mazón, J.-N., and Trujillo, J. 2008. "An MDA approach for the development of data warehouses", in: *Decision Support Systems* (45:1), S. 41-58.

McCarthy, J. C., Miles, V. C., and Monk, A. F. 1991. "An experimental study of common ground in text-based communication", *SIGCHI Conference on*

Human Factors in Computing Systems (CHI'91), New Orleans, LA, USA: ACM.

McFarlane, D. C., and Latorella, K. A. 2002. "The Scope and Importance of Human Interruption in Human-Computer Interaction Design", in: *Human-Computer Interaction* (17:1), S. 1-61.

McLaughlin, R. 2004. "Data in financial institutions", in: *Journal of Financial Transformation* (11), S. 75-80.

Miles, M. B., and Huberman, A. M. 1994. *Qualitative Data Analysis: A Sourcebook of New Methods*. Beverly Hills, CA, USA: Sage.

Mingers, J., and Willcocks, L. 2014. "An integrative semiotic framework for information systems: The social, personal and material worlds", in: *Information and Organization* (24:1), S. 48-70.

Mishler, E. G. 1986. *Research interviewing: context and narrative*. Cambridge, MA, USA: Harvard University Press.

Moody, D., and Kortink, M. 2001. "From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design", *3rd International Workshop on Design and Management of Data Warehouses (DMDW'01)*, Interlaken, Switzerland, S. 1-10.

Mulligan, P., and Gordon, S. R. 2002. "The impact of information technology on customer and supplier relationships in the financial services", in: *International Journal of Service Industry Management* (13:1), S. 29-46.

Narayanan, L. V. 2008. "Data Warehousing and Analytics in Banking: Concepts", in: *Data Warehousing and Mining: Concepts, Methodologies, Tools, and Applications*, J. Wang (ed.). Hershey, PA, USA: IGI Global, S. 1825-1839.

Nidumolu, S. 1995. "The Effect of Coordination and Uncertainty on Software Project Performance: Residual Performance Risk as an Intervening Variable", in: *Information Systems Research* (6:3), S. 191-219.

Nielsen, J., and Landauer, T. K. 1993. "A mathematical model of the finding of usability problems", *INTERACT '93 and Conference on Human Factors in Computing Systems (INTERCHI '93)*, Amsterdam, The Netherlands, S. 206-213.

Nova, N., Sangin, M., and Dillenbourg, P. 2008. "Reconsidering Clark's Theory in CSCW", *8th International Conference on the Design of Cooperative Systems (COOP '08)*, Carry-le-Rouet, France.

Olson, G. M., and Olson, J. S. 2000. "Distance Matters", in: *Human-Computer Interaction* (15:2), S. 139-178.

Pawlowski, S., and Robey, D. 2004. "Bridging User Organizations: Knowledge Brokering and the Work of Information Technology Professionals", in: *MIS Quarterly* (28:4), S. 645-672.

Rahm, E., and Bernstein, P. A. 2001. "A survey of approaches to automatic schema matching", in: *The VLDB Journal* (10:4), S. 334-350.

Rizzi, S., Abell, A., Lechtenbörger, J., and Trujillo, J. 2006. "Research in data warehouse modeling and design: dead or alive?", *9th ACM International Workshop on Data Warehousing and OLAP (DOLAP)*, Arlington, VA, USA, S. 3-10.

Robillard, P. N. 1999. "The role of knowledge in software development", in: *Communications of the ACM* (42:1), S. 87-92.

Rosenkranz, C., Corvera Charaf, M., and Holten, R. 2013. "Language Quality in Requirements Development: Tracing Communication in the Process of Information Systems Development", in: *Journal of Information Technology* (28:3), S. 198-223.

Rosenkranz, C., Räkers, M., Behrmann, W., and Holten, R. 2010. "Supporting Financial Data Warehouse Development: A Communication Theory-Based Approach", *31st International Conference on Information Systems (ICIS 2010)*, St. Louis, Missouri, USA, S. 12.

Rosenkranz, C., Räkers, M., Behrmann, W., and Holten, R. (under review). "Supporting Financial Data Warehouse Development: A Communication Theory-Based Approach", in: *European Journal of Information Systems* (under third round of review (major revisions)).

Saleem, K., Bellahsene, Z., and Hunt, E. 2008. "PORSCHE: Performance ORiented SCHEma mediation", in: *Information Systems* (33:7-8), S. 637-657.

Schoop, M. 2001. "An Introduction to the Language-Action Perspective ", in: *CM SIGGROUP Bulletin* (22:2).

Shvaiko, P., and Euzenat, J. 2005. "A Survey of Schema-Based Matching Approaches", in: *Journal on Data Semantics IV*, S. Spaccapietra (ed.). S. 146-171.

Simperl, E. P. B., and Tempich, C. 2006. "Ontology Engineering: A Reality Check", in: *On the Move to Meaningful Internet Systems 2006: CoopIS, DOA, GADA, and ODBASE* R. Meersman and Z. Tari (eds.). Berlin, Germany: Springer, S. 836-854.

Sjöström, J., and Goldkuhl, G. 2005. "The Semiotics of User Interfaces", in: *Virtual, Distributed and Flexible Organisations: Studies in Organisational Semiotics*, K. Liu (ed.). New York, USA et al.: Kluwer Academic, S. 217-236.

Skoutas, D., and Simitsis, A. 2006. "Designing ETL processes using semantic web technologies", *9th ACM International Workshop on Data Warehousing and OLAP (DOLAP)*, Arlington, VA, USA, S. 67-74.

Tan, M. 1994. "Establishing Mutual Understanding in Systems Design: An Empirical Study", in: *Journal of Management Information Systems* (10:4), S. 159-182.

Tan, M., and Teo, T. S. H. 2000. "Factors Influencing the Adoption of Internet Banking", in: *Journal of the Association for Information Systems* (1:1), S. 1-42.

Te'eni, D. 2001. "Review: A Cognitive-Affective Model of Organizational Communication for Designing IT", in: *MIS Quarterly* (25:2), S. 251-312.

Trujillo, J., Palomar, M., Gómez, J., and Song, I. 2001. "Designing Data Warehouses with OO Conceptual Models", in: *IEEE Computer* (34), S. 66-75.

Tryfona, N., Busborg, F., and Borch Christiansen, J. G. 1999. "starER. A Conceptual Model for Data Warehouse Design", *ACM Second International Workshop on Data Warehousing and OLAP (DOLAP'99)*, Kansas City, MO, USA: ACM, S. 3-8.

Uschold, M., Healy, M., Williamson, K., Clark, P., and Woods, S. 1998. "Ontology reuse and application", S. 179–192.

Vaishnavi, V. K., and Kuechler, W. 2008. *Design Science Research Methods and Patterns. Innovating Information and Communication Technology*. Boca Raton, FL, USA: Auerbach Publications, Taylor & Francis Group.

Vassiliadis, P., Simitsis, A., Georgantas, P., Terrovitis, M., and Skiadopoulos, S. 2005. "A generic and customizable framework for the design of ETL scenarios", in: *Information Systems* (30:7), S. 492-525.

Vassiliadis, P., Simitsis, A., and Skiadopoulos, S. 2002. "Conceptual modeling for ETL processes", *5th ACM International Workshop on Data Warehousing and OLAP (DOLAP)*, McLean, VA, USA: ACM, S. 14-21.

Watson, H. J., Fuller, C., and Ariyachandra, T. 2004. "Data warehouse governance: best practices at Blue Cross and Blue Shield of North Carolina", in: *Decision Support Systems* (38:3), Dec2004, S. 435-450.

Watson, H. J., and Wixom, B. H. 2007. "The Current State of Business Intelligence", in: *Computer* (40:9), S. 96-99.

Weir, R., Peng, T., and Kerridge, J. 2003. "Best practice for implementing a data warehouse: A review for strategic alignment", *5th International Workshop on Design and Management of Data Warehouses (DMDW'2003)*, H.-J. Lenz, P. Vassiliadis, M. Jeusfeld and M. Staudt (eds.), Berlin, Germany.

Winter, R. 2008. "Design science research in Europe", in: *European Journal of Information Systems* (17:5), S. 470-475.

Wixom, B. H., and Watson, H. J. 2001. "An Empirical Investigation of the Factors Affecting Data Warehousing Success", in: *MIS Quarterly* (25:1), S. 17-41.

Yin, R. K. 2003. *Case Study Research: Design and Methods*, (3 ed.). Thousand Oaks, CA, USA et al.: SAGE Publications.

Zhu, K., Kraemer, K. L., Xu, S., and Dedrick, J. 2004. "Information Technology Payoff in E-Business Environments: An International Perspective on Value Creation of E-Business in the Financial Services Industry", in: *Journal of Management Information Systems* (21:1), Summer2004, S. 17-54.

Paper 1: Specifics of Financial Data Warehousing and Implications for Management of Complex ISD Projects

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. AA

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 de-A ployed 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.

Paper 2: Classification of Communication Defects in DWH Projects

Classification of Communication Defects in DWH Projects

Alexander Hoffmann
Johann Wolfgang Goethe-University,
Frankfurt am Main, Germany
alexhfm@wiwi.uni-frankfurt.

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

ABSTRACT

Having completed the requirements engineering phase in a software engineering project, we usually refer to requirements as unambiguous, comprehensible and objective. In a software development case study conducted over a period of more than three years we observed several problems within and between the project teams resulting from different understandings of the software specification. Aim of the project was the implementation of a DWH that collects and evaluates significant data of an international banking institute in order to fulfill the requirements specified by Basel II. In order to classify and explain the observed communication defects that occurred during the implementation phase, we consider findings from communication theory and knowledge transfer (on the interaction setting). As a result, we propose a general schema for the classification of communication defects that may make a significant contribution to the handling of problems in communication issues.

Keywords

Software Engineering, Knowledge Transfer, Communication Theory, Communication Defects, Common Ground, Language Community

INTRODUCTION

During the last twenty years many theses regarding success factors and management of data warehouse (DWH) projects were published. They mainly focus either on a technical or a business view of DWHs and not on the development process as a whole. While the technical papers deal 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 give a holistic view of the DWH development process. Although lots of modeling languages, process models and success factor analyses were published, in practice DWH projects remain critical and many well-known problems still occur.

In a case study we accompanied a DWH project of an international banking group over a period of more than three years. During this time we recognized several problems concerning communication between the participating parties and within teams. Since the observed problems lead to higher project costs and communication related issues are not being considered by DWH literature, we set up a research project in order to analyze communication problems within the project. Our aim was to structure and classify the observed communication defects and to find reasons for their occurrence.

As we will show in this paper, it is possible to combine rigor and relevance by using approved and broad accepted methods of case study research joined with well known theories from other domains like knowledge transfer and communication theory. By doing so it is possible to better understand our real life observations and to identify areas of action to increase performance in DWH projects and software engineering projects in general. Or as (Kilduff, 2006) puts it: „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." By using this approach our conclusion base on a strong theoretical fundament.

This paper is structured as follows: The next section gives an overview of related work of the area of DWH projects. Afterwards, we present our case study and describe our research methodology. The fourth section gives a short introduction into communication theory and presents findings from the areas of knowledge transfer and language theory, which we use in the next section to analyze and to classify communication defects. Finally, we summarize our results and show the need for further research.

LITERATURE REVIEW

In 1968 the term “software engineering” initially appeared (Naur & Randell, 1968), but in spite of all subsequently invented methods to improve software engineering (Berry, 2004; Sommerville, 2001), software development still suffers from efficiency problems. 23% of software development projects are cancelled due to failure and 49% exceed project resources (Standish Group International, 2001), e.g. projects do not meet deadlines or simply get too expensive. Similarly, (Keil, 1995; Keil et al., 2000) state that between 30% and 40% of all software projects exhibit some degree of escalation.

It seems obvious, that a precise inquiry and specification of requirements is necessary for any software development project to be successful. Incomplete, wrong, or changing requirements lead to longer project runtime and higher costs. Thus, a lot of research has been done in the field of requirements engineering (RE). For example, (Kavakli & Loucopoulos, 2005) give a broad overview of current goal modeling approaches in RE. These approaches propose and offer procedures, methods and tools to cope with the process of RE. However, these approaches do not explicitly address the specific problems that occur after the specification is created. In contrast to existing approaches in RE, this work emphasizes a communication-based perspective that focuses on the time after RE, where the development team has to put the requirements into practice.

Since the 1970s a lot of publications have been made in the field of DWH. Most research addresses modeling languages and different process models to create the specification of a DWH project. Furthermore, these approaches refer to the organization of such projects. This implies the assumption that the major challenges of DWH projects lie in the elaboration and formal specification of the data. It is assumed that if a specification is correct in terms of the used formal specification language, the implementation will represent the desired result.

None of the aforementioned approaches explicitly address the personal interaction within project teams. Hitherto, only few publications focus on soft factors in DWH or on software development projects, e.g. in the field of knowledge management (He, 2004), learning processes (Pirinen & Pekkola, 2006) and communication (Gallivan & Keil, 2003). All mentioned research concerning DWH process models does not cope with the specific problems of interpersonal communication within DWH projects after RE phase is finished. This work gives a contribution to the aspect of communication within the field of DWH and software development.

CASE STUDY

Data Warehouse Development in Financial Industry

The following observations have been made during a Basel II project at an international banking group with a balance sheet total of more than € 100 billion. The banking group consists of a head office (HO) and 15 major subsidiaries. Goal of the project is the fulfillment of requirements specified by Basel II, which is a list of recommendations on banking laws and regulations issued by the Basel Committee on Banking Supervision. The project started in 2001 and is planned to end in December 2007, so that the project has a running time of 6 years. Currently, the IS solution is being approved by the regular authority.

To fulfill the requirements of Basel II and to enable a group-wide calculation of the required values one main task of the project was to implement a central DWH. It is designed to serve as the data source for the Basel II calculation engine and has to store all relevant raw data required for the calculation of so called Risk Weighted Assets of the banking group. Primarily, data of transactions, collaterals, customers and rating information of both HO and subsidiaries have to be delivered into the central DWH. To achieve this, each subsidiary had to develop extraction jobs for their local databases which create specification conform flat files that are then sent to HO and imported into the DWH.

The specification of the flat files was developed by a core team in HO consisting of five persons, including business and DWH experts to ensure a broad support of the local units. The team also supported and controlled the implementation process of the database extraction jobs. The specification was created using XML Schema Definition (XSD) in order to define the required data in a formal and unambiguous way. At the beginning of the project, an initial training workshop was conducted where the XSD format was explained to representatives of the subsidiaries. Since the specification was assumed to be unambiguous due to the use of XSD, it was sent to each subsidiary via e-mail without further descriptions. Furthermore, HO assumed that the persons in the subsidiaries share the same knowledge about the banking domain, that each of the used terms and descriptions would be understood without any problems.

During the implementation phase, a contact person in HO was nominated for each subsidiary to clarify misunderstandings and technical problems. Support through HO was realized by using several communication channels including e-mail and telephone. Since after the distribution of the specification all support requests of the subsidiaries could be satisfied, HO team assumed that each subsidiary had understood the specification in the intended way. But completeness checks of the flat files that were delivered by the subsidiaries showed several problems resulting from a different understanding of the specification. These misunderstandings led to delays in the project timetable and unforeseen, additional costs, since on-site-visits had to be arranged about once per month in order to discuss revealed problems, the actual implementation status, and acute problems in face-to-face situations. The different understandings could neither be predicted nor managed via telephone and e-mail. It seemed as if HO and subsidiaries each spoke different languages, since the interpretations of the published specification varied to a considerable degree, although the required data and data structures were formalized unambiguously due to the use of XSD and additionally extensive textual descriptions were used.

Aim of our research was to structure the observed problems and to find reasons for their occurrence. During our research we did not use findings from the area of requirements engineering, since a specification was already at hand and problems appeared not until the publication of the specification. Instead, we classified and explained the observed misunderstandings using findings from communication theory and language theory, since the observed problems occurred due to misunderstandings of the already finalized specification.

Research Methodology

Our research is performed according to the case study method described by (Yin, 2003). This methodology was used because we have examined complex organizational phenomena in their natural setting. The goal of our examination is to understand communication related problems between different project members in a large DWH project. Therefore we have to deal with how and why questions about a contemporary set of events over which we have only little control (Yin, 2003 p 20). By analyzing real life phenomena it is ensured that we deal with relevant questions and by using the broad accepted method of Yin the rigor of our results is guaranteed as well. These criteria led use to choose case study as our preferred research method.

One of the authors is still involved as a consultant in the project described in chapter 3.1. The observations concerning communication defects were documented by him and reconciled with other involved employees of the banking groups. The observations contain descriptions of the occurred problems, the identified causes for the problems (normally communication defects) and details about how and when the problems were discovered as well as involved persons or units. The facts were collected on-site and through interrogations of subsidiary contact persons and HO team members by one of the authors. Furthermore the project documentation (meeting protocols, presentations of lessons learnt sessions, data models, calendars, etc.) was used to confirm the documentations. Additionally unstructured information and non-formal documentation (e.g. e-mail correspondence) was evaluated. The observed facts were then discussed and analyzed by both authors. To ensure an adequate distance and to increase the objectivity of the case, all findings were also discussed in our research team with non-involved people.

THE “COMMON GROUND” IN COMMUNICATION

In our case study we observed communication between HO team and the subsidiaries necessary to fulfill the specified requirements. During this communication, knowledge was transferred from HO to the subsidiaries. Most problems we observed in our case study are based on misunderstandings concerning the concepts that are represented by data structures in the specification. That is, defects in knowledge transfer from HO to subsidiaries occurred. In order to explain the observed communication defects and to be able to put them into a structure that allows us to find strategies for more successful communication in software engineering projects, we will explain the idea of common ground, which is a prerequisite for successful knowledge transfer, and give a short insight into the basics of human communication insomuch as they are necessary to understand the evolution of common ground..

(Clark, 1996) emphasizes that communication acts can be seen as joint activities with two actors, a sender and a receiver, being involved. For successful communication they need to coordinate their activities. Basis for this coordination is shared knowledge between the two actors: They must share a common ground. Any utterances will be formed in relation to the assumptions made by the sender about the common ground. By each of his or her utterances the sender wants to extend the common ground step by step. This extension of the common ground is called grounding. If the contribution to the common ground gets too big or is based on wrong assumptions, misunderstandings occur that might be difficult to detect. Incorrect assumptions about the common ground are often caused by an egocentric bias: “If I know something, I am more likely to expect others to know it too”.

There are three principles that are related to the grounding process. The principle of closure states that agents performing an action require evidence that they have succeeded in performing that action in a way that is sufficient for current purposes. In case of joint actions (e.g., communication acts) the principle of joint closure becomes important: The participants in a joint action try to establish the mutual belief that they have succeeded well enough for current purpose. At last, the principle of least effort states that all things being equal, agents try to minimize their effort in doing what they intend.

Clark’s theory about common ground in communication was applied by (Bromme et al., 2004) to communication scenarios with both sender and receiver having major knowledge differences. This so-called expert-layperson communication is characterized by a low common ground between the actors at the beginning of the communication process. Parts of knowledge in the layperson perspective might be missing and knowledge elements might be embedded in the layperson’s “amateurish” cognitive reference framework that holds the risk of resistance against changes. For example, the expert might receive positive feedback from the layperson, while in fact the layperson did not understand the transferred knowledge in the correct way (Bromme & Jucks, 2001). This may cause an illusion of evidence to the expert. In our case study we observed expert-layperson communication between HO team and the subsidiaries. Both parties share a common ground concerning knowledge related to the banking domain, but only HO team consists of experts regarding the specific requirements of Basel II. Since after the distribution of the software specification the subsidiaries, which are laypersons regarding Basel II requirements, did not make extensive support requests besides minor issues, HO team interpreted subsidiary’s “silence” as prove of a correct understanding of the specification, i.e. HO team suffered from an illusion of evidence.

We now see that a common understanding or common ground respectively is an important success factor of knowledge transfer. But in order to fully understand the communication defects which we have observed in our case study, more in-detail explanations concerning the process of knowledge transfer are necessary.

For knowledge to be communicated to other persons, the knowledge holder has to represent its mental concepts in a way that they are perceivable for other persons. The result of this representation process is data that can be stored by and shared between persons and machines. Since knowledge is what people have in their minds and data is external perceivable structures, transfer of data is not identical to knowledge transfer. Or, as (Garavelli et al., 2002) put it: “Even when knowledge can be materialized in an object […], the transfer of that object does not necessarily fulfill the knowledge transfer process.” That is because data has to be interpreted by people. Data is inherently meaningless, „It simply exists […] – all waiting to be interpreted, all waiting to have meaning attached – by people.” (Miller, 2002). It is in the moment of interpretation when new knowledge is created in a person’s mind. But one can never be sure that two persons associate the same representation with the same object. „Identical [data] almost invariably provokes (or evokes) different meanings in each of us. […] Although the words we use stand for meaning, we should not assume that those words necessarily provoke the same meaning in others” (Miller, 2002). Therefore, knowledge transfer is not a well defined process.

Problems concerning knowledge transfer occur, if persons speak different languages, whereas language not only refers to natural languages like English, German or Spanish, but also jargons like "mechanical engineer" or "field sales" (Davenport & Prusak, 1998). These persons may have different ways of codifying and communicating their knowledge. A group of people sharing the same representations associated with the same objects is called language community (Kamlah & Lorenzen, 1984). Obviously, for persons to be members of a language community, they must share the same knowledge, i.e., they must have knowledge about the same objects independently of their representations. Therefore, the existence of a common ground is a prerequisite for the existence of a language community.

CLASSIFICATION OF COMMUNICATION DEFECTS

Based on the theory of knowledge transfer, we analyzed and classified communication defects that we observed in our case study. In this section we consider a sender-receiver-situation with the sender ("A") formulating a software specification (HO team in our case study) and the receiver ("B") having to understand the specification and transforming the requirements into concrete action or software respectively (the subsidiaries in our case study). The sender codifies a specific requirement, which we call X, using a representation R.

A proposition of the considered situation is that the sender's aim is an efficient and reasonable communication of X, i.e., the sender pursues a communication strategy that maximizes the probability that the receiver's understanding of X is correct and minimizes the costs that arise from the codification and transfer of X (see principle of least effort). The communication process is successful, if after the transmission of R both sender and receiver have the same idea of X in mind. We identified six basic conditions that must be fulfilled any communication process to be successful and applied these conditions to the communication defects that occurred in our case study.

	General description	Case study
Condition 1	The sender must know at least one representation R of X in order to communicate X.	Due to agreed specification language a fulfillment of condition 1 was always observed.
Condition 2	Both sender and receiver must have knowledge of X independently of any representation.	Some communication defects based on non-fulfillment of condition 2 have been observed.
Condition 3	R must not be manipulated on its way from the sender to the receiver.	By using textual specification and electronic data transfer a fulfillment of condition 3 was always observed.
Condition 4	The receiver must see R as a part of a communication process and must feel addressed by R.	Due to a clearly defined communication process a fulfillment of condition 4 was always observed.
Condition 5	The receiver must know at least one interpretation of R and one of these interpretations must correspond to X.	Some communication defects based on non-fulfillment of condition 5 have been observed.
Condition 6	If the receiver knows more than one interpretations of R, he or she must choose the one that corresponds to X.	Some communication defects based on non-fulfillment of condition 6 have been observed.

Table 1: Conditions for successful communication

Condition 1

At the beginning of a communication process the sender must know a set of representations with each element representing X in a way that is appropriate and communicable to the receiver (condition 1). This set is called $P_A\left(X\right)$, since it contains predications (P) that a sender (A) knows for a certain object (X). $P_A\left(X\right)$ must contain at least one element or representation respectively. What elements are in $P_A\left(X\right)$ depends on the sender's capability of expression and on external circumstances, e.g., language guidelines in an organization. In the case of $\left|P_A\left(X\right)\right|>1$ the sender knows several representations of X of which he or she must choose one. According to the principle of least effort, the sender will choose the most efficient representation that the receiver is assumed to understand. The representation that the sender chooses from $P_A(X)$ and communicates to the receiver is called R. In our case study we did not find situations, where the sender was not able to choose an appropriate representation for X. Therefore, the observed communication defects did not result from non-fulfillment of condition 1.

Condition 2

The second condition demands, that the receiver must know X independently of its representations, i.e., the receiver must know the object that is represented by R. If the sender knows for sure or assumes, that the receiver has knowledge about X, R can be used without further definitions of X. Additionally, conditions 3 to 6 have to be fulfilled as well. Alternatively, if the sender is not sure about the receiver's knowledge about X, it must be explained to ensure successful communication. Explaining or defining an object X means to combine several other representations that the receiver is assumed to understand and to equate these combined representations with a shorter representation R that will be used further on to achieve a more efficient communication (Seiffert, 1996). In our case study, definitions were only conducted if new and project-specific ideas and concepts were introduced, that the receiver was unlikely to know. For example, the specification contained an attribute group named "multi source, multi entity fields" that was represented with the term "MSME fields". Since the sender, i.e., HO team, did assume that the receiver neither knows the underlying concept nor the used representation, a detailed definition of "MSME fields" was given. Due to the elaborate definition, no communication defects occurred regarding "MSME fields" during the project duration. In another situation, the word "limit", which was included in the specification, was misunderstood by the subsidiaries. The reason for this misunderstanding was that the subsidiaries did not know the concept that was represented by the word "limit", which is in a Basel II context an external committed credit line. "Limit", however, was known by the subsidiaries only for internal risk limitation. Due to an egocentric bias HO team did not define "limit" with explicit focus on the Basel II context. As a consequence subsidiaries used their own interpretation of "limit" without knowledge of the object that the sender actually wanted to represent and developed program code that was not compatible with the specification. Therefore, some communication defects in our case study are based on non-fulfillment of condition 2.

Condition 3

According to the classical sender-receiver-model published by Shannon and Weaver, communication comprises the transmission of a signal from sender to receiver over a communication channel (Shannon and Weaver, 1963). During transmission a signal can be manipulated by an external noise source, e.g. background noise during a face-to-face conversation or technical problems in case of e-mail communication. In our case study we did not observe any communication defects resulting from noise that changed the transmitted representations (condition 3).

Condition 4

For the communication process to be successful, the receiver must realize the intention of communication and feel addressed if receiving R (condition 4). We assume that this is always the case if commonly known symbols like letters, numbers, or elements of a well known modeling language are used, since these symbols do not exist naturally but are created by men with intend to communicate. But even so, the intended receiver might, for example, see R as „internal notes" of the sender that are not addressed to him or her. In our case study we did not find any situations where the receiver did not feel addressed or did not see R as part of a communication process. Therefore, the observed communication defects did not result from a non-fulfillment of condition 4.

The receiver must be able to interpret R so that X is among the possible interpretations (condition 5). Furthermore, if more than one interpretation is available, the receiver must choose the correct interpretation (condition 6). Given a fulfillment of condition 2, the sender must make an assumption about the objects that the receiver associates with R. The interpretations of R that the receiver is assumed to know are called $I_{A \to B}(R)$ (Interpretations).

Condition 5

Regarding $I_{A \to B}(R)$, we must at first consider if the sender assumes X to be in $I_{A \to B}(R)$ or not (condition 5). In the latter case, i.e., $X \notin I_{A \to B}(R)$, the sender must choose a more appropriate representation, given that $|P_A(X)| > 1$. In our case study, the specification team tried to avoid such terms or phrases that the receiver was likely to misinterpret by choosing more appropriate representations if $|P_A(X)| > 1$. Given that $|P_A(X)| = 1$ and $X \notin I_{A \to B}(R)$, the sender is not able to choose a more appropriate representation of X that the receiver is less likely to misinterpret. For example, in our case study the specification contained a data field named "*original*_principal_amount" that had to store the *current* principal amount, which is calculated in a different way than the original principal amount. That is, the name of the data field suggested another purpose of use as the specification actually required. Due to technical reasons, HO team (sender) was not able to choose another representation (i.e., another name) of the data field. Unfortunately, the specification did not contain adequate explanations, so that some subsidiaries (receiver) used this data field in a wrong context, resulting in faulty implementations. As opposed to the example given for condition 2, here the receiver had knowledge about both *original* and *current* principle amount so that the misunderstanding was only based on the usage of an inappropriate representation. Therefore, communication defects occurred due to a non-fulfillment of condition 5. In order to increase the chance for successful communication, the sender could additionally have added examples or analogies, since the receiver was likely to interpret R in a not intended way.

Since the sender supposes that the receiver knows X independently of any representation (see condition 2), support for the correct interpretation of R should be given, if $|I_{A \to B}(R)| \geq 1$ is assumed. For example, the sender could give a synonym or a list of synonyms of R. The sender could also list and exclude other possible interpretations in order to draw the receiver's attention to a possible misunderstanding. Neither of these strategies was observed in our case study.

Condition 6

If the sender assumes $X \in I_{A \to B}(R)$, i.e., if the sender assumes that both sender and receiver are members of the same language community, depending on $|I_{A \to B}(R)|$ different communication strategies can be pursued: Given that $|I_{A \to B}(R)| = 1$, the sender assumes that the receiver interprets R exactly as intended and is therefore likely to use R without any further explanation. If $|I_{A \to B}(R)| > 1$ is assumed, R can be used without further explanation, if the sender trusts in the receiver's ability to choose the interpretation that corresponds to X among several others (condition 6). This trust might lead to communication defects, if the receiver is uncertain about the correct interpretation and chooses the wrong interpretation of R. In our case study, for example, the specification postulates that each subsidiary has to deliver book values according to a specific data format. But in the financial world different definitions of the term "book value" exist, each having different calculation instructions, e.g. according to the German Handelsgesetzbuch (HGB) or the International Financial Reporting Standards (IFRS). Although HO team (sender) knew about the different definitions in principle, it unconsciously assumed that the subsidiaries (receiver) know all of these definitions and would interpret "book value" in the intended way. In fact, the subsidiaries were uncertain about what interpretation was intended by HO. As a consequence, HO had to answer further inquiries from the subsidiaries, leading to higher communication costs. Therefore, some communication defects in our case study occurred due to a non-fulfillment of condition 6. If the receiver is not assumed to choose the correct interpretation, the sender can list and explicitly exclude other possible interpretations in $I_{A \to B}(R)$. This communication strategy is, of course, only possible for interpretations of R that are known to the sender, i.e. that are elements of $I_A(R)$. For example, in our case study a recapitulatory entity for all banking transactions was introduced and called "facility". But the term "facility" was usually also used in other contexts, e.g. credit lines. In order to avoid wrong interpretations of the term, other possible interpretations were listed and excluded. Because $|I_{A \to B}(R)|$ can be greater than $|I_A(R)|$, the sender can not be sure to have excluded all other possible interpretations. Therefore, additional interpretation support through the use of examples or analogies should be given.

CONCLUSION

Aim of our research was to classify and give reasons for communication defects that we observed in our case study. We therefore identified six conditions that must be fulfilled for communication processes to be successful. These conditions are based upon findings from the areas of knowledge transfer and communication theory. When applying these conditions to our case study, we found that fulfillment of three out of six conditions for successful communications can easily be ensured by basic communication setup and that most misunderstandings occurred due to inappropriate phrasing in the specification and the sender's unconscious trust in the receiver's ability to choose the intended interpretation among several possible. As a consequence, we recommend accepting higher costs at setting up the specification either by stronger interaction or a more detailed specification in order to reduce follow-up costs resulting from misunderstandings. The proposed way of structuring communication defects allowed us to precisely classify the problems we observed in our case study and therefore enables us to develop more efficient communication strategies in software development projects. Further research will include appliance of the classification schema to other domains and the extension of the schema to communication transactions in total.

REFERENCES

1. Berry, D. M. (2004) The Inevitable Pain of Software Development: Why There Is No Silver Bullet. In Radical Innovations of Software and Systems Engineering in the Future (Wirsing, M. and Knapp, A. and Balsamo, S., Eds), pp 50-74, Venice, Italy.
2. Bromme, R. and Jucks, R. (2001) Wissensdivergenz und Kommunikation: Lernen zwischen Experten und Laien im Netz. Partizipation und Interaktion im virtuellen Seminar, 81-103.
3. 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.
4. Clark, H. H. (1996) Using Language. Cambridge University Press, Cambridge.
5. Davenport, T. H. and Prusak, L. (1998) Working Knowledge - How Organizations Manage What They Know. Boston, Massachusetts.
6. Gallivan, M. J. and Keil, M. (2003) The user-developer communication process: a critical case study. Information Systems Journal 13 (1), 37-68.
7. Garavelli, A. C., Gorgoglione, M. and Scozzi, B. (2002) Managing knowledge transfer by knowledge technologies. Technovation 22, 269-279.
8. 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.
9. Kamlah, W. and Lorenzen, P. (1984) Logical Propaedeutic. Pre-School of Reasonable Discourse. University Press of America, Lanham, MD.
10. Kavakli, E. and Loucopoulos, P. (2005) Goal Modelling in Requirements Engineering: Analysis and Critique of Current Methods. Information Modeling Methods and Methodologies, 102–124.
11. Keil, M. (1995) Pulling the Plug: Software Project Management and the problem of Project Escalation. MIS Quarterly 19 (4), 421-447.
12. Keil, M., Mann, J. and Rai, A. (2000) Why Software Projects Escalate: An Empirical Analysis and Test of Four Theoretical Models. MIS Quarterly 24 (4), 631-664.
13. Kilduff, M. (2006) EDITOR'S COMMENTS: PUBLISHING THEORY. Academy of Management Review 31 (2), 252-255.
14. Miller, F. J. (2002) I=0 (Information has no intrinsic meaning). Information Research 8 (1),
15. Naur, P. and Randell, B. (1968) Software Engineering: Report an a Conference Sponsored by the NATO Science Commission. Scientific Affairs Division, NATO, Garmisch, Germany.
16. Pirinen, A. and Pekkola, S. (2006) Different Perspectives on Learning in Information System Development. In IRIS, LO-skolen.
17. Seiffert, H. (1996) Einführung in die Wissenschaftstheorie 1. Verlag C. H. Beck, München.
18. Shannon, C. E. and Weaver, W. (1963) The Mathematical Theory of Communication. University of Illinois Press, Urbana.

19. Sommerville, I. (2001) Software Engineering. Pearson Education Limited, Essex.
20. Standish Group International, I. (2001) Extreme CHAOS. Research report, ordering information available at www.standishgroup.com,
21. Yin, R. K. (2003) Case Study Research: design and methods. Sage Publications Inc.

Paper 3: Hitting Interface Specifications -The Role of 'Common Ground' in Financial DWH Projects

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.

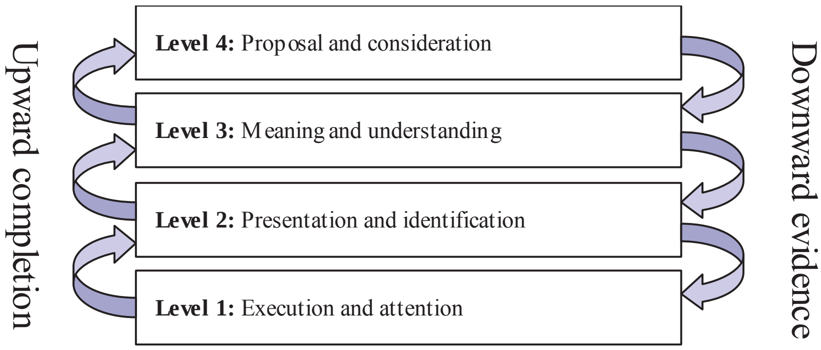

Figure 1. Action ladder in communication acts

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.

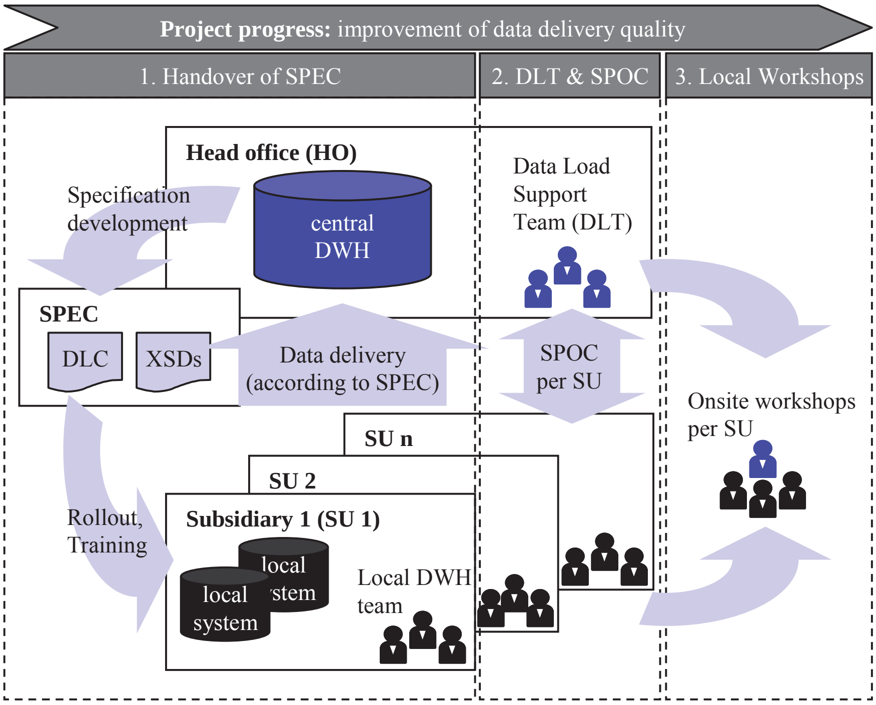

Figure 2. Case study – illustration of architecture and organisation.

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.

Paper 4: Supporting Financial Data Warehouse Development: A Communication Theory-Based Approach

SUPPORTING FINANCIAL DATA WAREHOUSE DEVELOPMENT: A COMMUNICATION THEORY-BASED APPROACH

Completed Research Paper

Christoph Rosenkranz
Goethe University
Frankfurt am Main, Germany
rosenkranz@wiwi.uni-frankfurt.de

Marc Räkers
zeb/information.technology
Bad Bentheim, Germany,
mraekers@zeb.de

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

Roland Holten
Goethe University
Frankfurt am Main, Germany
holten@wiwi.uni-frankfurt.de

Abstract

Data warehouses increasingly play important roles in the information technology landscape of the financial industry. However, semantic heterogeneity is high in banking – data is defined differently by different banks, business units, and users. Therefore data integration in financial data warehouse development projects relies on the knowledge, know-how, and judgment of human experts. Up to now, methodical support is missing for the communication process among experts that determine and negotiate a shared understanding of requirements. In contrast to ontology-driven or schema-matching approaches proposing the automatic resolution of differences ex-post, we introduce an approach that addresses data integration already in early project phases. Our approach supports developing shared understanding of domain concepts and data fields in financial data warehouse projects, good communication of all participants while the project progresses, and early detection of errors within projects. This way, we prevent problems that result from the ex-post resolution of semantic heterogeneity.

Keywords: Communication theory, Business intelligence (BI), Data warehouse development, Knowledge sharing, Common ground, Language

Introduction

Data warehouses (DWH) increasingly play important roles in the information technology (IT) landscape of the financial industry. IT has become one of the main factors of production in banking because the production process of a bank mainly consists of information processing (Baldwin et al. 2001; Tan and Teo 2000). That is why it is critical for financial institutions to be able to access relevant data when it is needed most. Analytical applications, in practice often described as “business intelligence”, allow users to access data and use it to make decisions in order to realize the full value from their DWH costs (Watson and Wixom 2007). Moreover, financial institutions increasingly have to fulfill extensive regulatory requirements such as Sarbanes-Oxley Act, IFRS, or Basel II (McLaughlin 2004). These regulations demand of banks to conduct calculations that include the institutions’ whole portfolio and detailed subsidiary data. The importance of such integrated and consolidated risk estimation has been emphasized by the recent financial crisis which has been triggered by the rise in delinquencies in the United States. Therefore many banks have implemented DWH solutions during the last decade to enhance decision-making, controlling, and risk management.

However, the adoption of DWH technology, which requires huge capital spending and also consumes a good deal of development time, has a very high possibility of failure (Hwang et al. 2004). The challenge stems from multiple causes such as poor data quality in the operational source systems, politics around data ownership, and legacy technology (Hwang et al. 2004; Hwang and Xu 2007; Watson et al. 2004; Weir et al. 2003; Wixom and Watson 2001). One of the biggest challenges for DWH designers is *data integration* – the consolidation of data from disparate sources into one consistent body of data (March and Hevner 2007, p. 1036). In general, the meaning of a specific data field in its context varies from company to company, from business unit to business unit, and from department to department. Such differences can be as simple as naming conventions or units of measures (March and Hevner 2007, p. 1038). In the financial industry, the differences frequently are much more complex and involve different meanings ascribed to captured data (Behrmann and Räkers 2008). For example, what is the meaning of terms such as “limit”, “bond”, “interest payment” and so forth, and which data fields of what source system map to these meanings? Therefore *semantic heterogeneity* (March and Hevner 2007, p. 1038) is high in banking – data is defined differently by different banks, business units, or users. No common standards for data integration exist until today in the financial industry. Such an universal, consistent, and unitary description of all types of financial businesses is not a realistic option because financial institutions constantly update their business models by rapidly developing new products and services in order to gain a competitive advantage (Corcho et al. 2005). Consequently, semantic heterogeneity is likely to stay high in the financial industry. Moreover, banks struggle to integrate all relevant data because data sources are often heterogeneous, disorganized, or even inaccessible. For these reasons financial DWH systems still are tailor-made, cost-intensive applications, even if standardized core banking systems do exist. Therefore effective data integration can only ever be achieved by human experts who have access to knowledge about internal and external business contexts – “What questions are momentarily important?” – and access to knowledge pertaining operational systems that have to be integrated – “Which operational systems provide data for answering those questions?” (March and Hevner 2007, p. 1041).

Although the relevance of data integration in general is assessed to be high (March and Hevner 2007; Rizzi et al. 2006), only few researchers work in this subject area (Skoutas and Simitsis 2006; Vassiliadis et al. 2005). This is remarkable because the high effort of data integration seems to be without controversy – getting data into a DWH is the most difficult aspect of business intelligence, requiring about 80 % of the time and effort and generating more than 50 % of the project costs (Kimball and Caserta 2004; Vassiliadis et al. 2002; Watson and Wixom 2007). The goal of this paper is to introduce an approach that primarily consists of a tool-supported procedure for designing, managing, and tracking requirements for data integration in financial DWH development projects. Our approach aims to support (1) developing shared understanding of domain concepts and data fields, (2) good communication of all participants while the project progresses, and (3) early detection of errors within projects. This way, we prevent data integration problems that result from the ex-post resolution of semantic heterogeneity.

We follow a Design Science approach (Hevner et al. 2004; Hevner and March 2003) that deals with the construction of scientific artifacts for solving a problem. As a theoretical basis for our approach, we make use of linguistic communication theory and Clark (1996)’s central thesis that “language use is really a form of joint action” (p. 3). In projects as a joint activity, the source of project participants’ ability to coordinate and create shared understanding is their *“common ground”*, the set of knowledge, beliefs and suppositions that they believe they share (Clark 1992; Clark 1996; Clark and Brennan 1991). Our approach supports development of common ground by jointly introducing, discussing, and negotiating concepts and terms, including descriptions of their meanings. Although the approach has been developed for financial DWH development projects, which we also use for illustrational purposes, our approach is in principle generic and can be applied to any domain being equally prone to semantic heterogeneity.

The remainder of the paper is structured as follows. First, we analyze related work on DWH development and the research gap that led to the development of our approach. Next, we introduce a conceptual framework for supporting financial DWH development that builds on communication theory. Then we present our approach and the design components. The feasibility of the approach is demonstrated with a detailed application scenario and evaluated using multiple case studies. We finish the paper in a "Conclusions and Outlook" section and motivate further research.

Related Work and Problem Awareness

Several studies have revealed the importance of determining information requirements in DWH development (Watson et al. 2004; Wixom and Watson 2001). To develop DWHs, it is necessary to identify what kind of data has to be provided to whom for what kind of management decision. Consequently, numerous researchers have presented different *approaches for DWH design, DWH development, and DWH engineering* over the recent years. Some of these have been suggested by practitioners (Ballard et al. 1998; Inmon 2005; Kimball and Caserta 2004; Kimball and Ross 1996). In addition, researchers and academics have proposed a variety of approaches (Böhnlein and Ulbrich-vom Ende 1999; Cavero et al. 2001; Chenoweth et al. 2003; Giovinazzo 2000; Golfarelli et al. 1998; Golfarelli and Rizzi 1998; Holten 2003; Mazón and Trujillo 2008; Moody and Kortink 2001; Trujillo et al. 2001; Tryfona et al. 1999). Many of these target a specific conceptual modeling approach (for example, star schemas with ER models) and are often too complex to be used in real-world environments (Malinowski and Zimányi 2008). None of these approaches has been widely accepted and all feature some deficits (Abell et al. 2001; Mazón and Trujillo 2008). The development of a DWH has been traditionally guided by an in-depth analysis of the underlying operational data sources, thus overlooking an explicit development phase in which information requirements of decision-makers are addressed (Mazón et al. 2007). March and Hevner (2007) conclude that current methodologies for DWH engineering are in their infancy.

Even with a fitting DWH engineering approach, semantic heterogeneities and data integration continue to pose enormous challenges for DWH developers, especially in the context of data extraction, transformation, and loading (ETL) (Vassiliadis et al. 2005). *Schema matching* is often proposed as a solution in data integration contexts and relies on discovering correspondences between similar elements in a number of schemas, for example, from different operational source systems. Researchers have proposed several different approaches for schema matching (Bernstein et al. 2004; Do and Rahm 2007; Doan et al. 2003; Ehrig and Staab 2004; Madhavan et al. 2001; Rahm and Bernstein 2001; Shvaiko and Euzenat 2005). However, schema matching in environments with heterogeneous data is time-consuming and error-prone, as existing mapping tools employ semi-automatic techniques for mapping two schemas at a time. Saleem et al. (2008) conclude that even sophisticated state-of-the art semantic matching approaches cannot guarantee that the mappings are 100 % correct, and ultimately still have to rely on human judgment to select the best candidate as the mapping for the source schema elements to the target schema elements.

Ontologies (Campbell and Shapiro 1995; Simperl and Tempich 2006; Uschold et al. 1998) seem like a possible solution for the schema-matching problem. They offer the potential for interoperability because they are semantically rich, computer interpretable, and inherently extensible. Ontologies can be applied for integration of heterogeneous data sources once a robust domain ontology is established. However, several ontologies exist in the financial industry (Mäkelä et al. 2007), and it is unlikely that a common, parsimonious ontology will ever be developed. The financial domain is constantly changing and developing new products, and innovative financial institutions are not willing to share their knowledge, being afraid of losing their competitive advantage (Corcho et al. 2005).

Fundamentally, the problem of data integration in financial DWH development projects is not a pure technological problem. What is really needed in financial DWH development projects is a standardization of the concepts and terms themselves. This results in involved stakeholders from business and IT that communicate and discuss with questions along the lines of "What does this mean?", "Is this the same as one of those?", or "What does this model actually show?" For example, does the term "redemption" for a bond mean the same thing as for a fund? Are the meanings of the term "redemption" in both contexts related or are they completely different? And which data fields of operational source systems capture those meanings? Therefore data integration relies on the knowledge, know-how, and judgment of human experts (Behrmann and Räkers 2008). What financial DWH development projects really need to capture is the semantics of the terms themselves, not only the words or data field names. To sum up, we identify the following need for the development of DWH in the financial industry: up to now, methodical support is missing for (1) the communication process among experts that determine and negotiate a shared understanding of requirements, concepts, and terms relevant for the financial DWH, and (2) for managing (creating, storing, updating) a consistent terminology based on a shared understanding of the requirements.

Design Rationale and Conceptual Framework

Caused by the broad scope, the large size, and the heterogeneous IT infrastructure and environment, a multitude of different stakeholders is involved in financial DWH development projects, for example, DWH experts, operational source system specialists, business (subject matter) experts, or business managers and decision-makers. Each group of stakeholders owns specific knowledge which has to be reflected in DWH specifications (March and Hevner 2007, p. 1035). During a financial DWH development project, the data integration requirements usually are analyzed starting from the business perspective (business information requirements). After defining the basic scope of the financial DWH, the business content is detailed, specified and merged with the IT perspective (conceptual and logical design). Therefore both business experts and IT experts have to reach a shared understanding and develop an integrated specification of the business information requirements. Financial DWH development can therefore fundamentally be understood as a *communication process* between stakeholders with expertise in different fields of knowledge.

Although the importance of knowledge transfer, negotiation, and communication for IS development (ISD) are recognized (Hansen and Lyytinen 2010; Levina and Vaast 2005; Robillard 1999), linguistics and communication theory have only seldom been applied for studying or supporting ISD (Auramäki et al. 1992; Clarke 2001; Corvera Charaf and Rosenkranz 2010; Flores et al. 1988; Goldkuhl and Lyytinen 1982; Holmqvist 1989; Schoop 2001).

In psycholinguistics, Clark's and collaborators' research on human language (Clark 1992; Clark 1996; Clark and Brennan 1991; Clark and Krych 2004) has provided a communication theory that has been useful in research on human-computer interaction (HCI) and computer-supported collaboration work (CSCW) (Kanda et al. 2004; Maglio et al. 2002; McFarlane and Latorella 2002; Olson and Olson 2000). Although the use of Clark's work as a theoretical framework is not without controversy (Koschmann and LeBaron 2003; Nova et al. 2008), the concept of common ground, for example, has been extensively applied for designing computer-mediated communication (Carroll et al. 2006; Convertino et al. 2009; McCarthy et al. 1991).

In contrast, Clark's theory has sparsely been used in the IS discipline, for example, for studying knowledge integration in virtual teamwork (Alavi and Tiwana 2002), for examining the use of IT systems in organizations (Sjöström and Goldkuhl 2005), or for language-based IS evaluation (Ågerfalk 2004). To the best of our knowledge, however, these concepts have never been applied in the context of examining or supporting ISD. In particular, we are not aware of any study that applies these concepts to deal with DWH development as a communication process.

Some well-established theoretical principles about the communication process follow from research by Clark and collaborators (Clark 1992; Clark 1996; Clark and Brennan 1991; Clark and Krych 2004).

Principle 1: Communication is a joint activity

What makes an action such as DWH development a joint one is the coordination of both *content*, what the participants intend to do, and *processes*, the physical and mental systems they recruit in carrying out those intentions (Clark 1996, p. 59). Joint actions require active involvement and constant verifications by all participants. Joint actions such as DWH development cannot be accounted for without understanding the interplay between content and process, and their place in overall joint activities; conversations cannot work without coordination of both content and process (Clark 1996, pp. 59, 319). Content and process are interdependent: the more complicated the content, the longer the process (Clark 1996, p. 90). Moreover: a minor misunderstanding at the beginning might snowball into major ones in the end (Clark 1996, p. 235). In the context of DWH development, this principle suggest that the whole project team has to establish early transparency regarding the meaning of data fields in order to achieve data integration (content); the available operational source systems have to be jointly analyzed by business experts and IT experts, and subsequently have to be consolidated into an integrated specification of the data supply (process).

Principle 2: Communication depends on a grounding process

The basis for coordination is shared knowledge (common ground) between actors. The grounding process, the process of establishing mutual understanding, is always adaptive to the current context of communication (Clark 1996, p. 99) and totally depends on the assumptions of the common ground made by the sender. In general, two types of common ground do exist – process and content (Clark and Brennan 1991). *Content common ground* includes "I know that you know that I know *what*", *process common ground* encompasses "I know that you know that I know *how*" (Convertino et al. 2009). Content common ground is the shared understanding on the subject and focus of work, resulting from exchanging content and mutually checking and signaling understanding. Process common ground is the shared understanding of the rules, procedures, timing, and manner in which the interaction will be conducted. The principle of grounding and the concept of common ground are perhaps the most central concepts from Clark's theory for DWH development. Common ground is incrementally built on the history of joint actions between communicators. The creation of a shared understanding and common ground is the prerequisite for a correct data mapping and ETL design. To reach this goal, the business information requirements must be understood by all involved team members. In financial DWH projects, content common ground is related to the need to selectively share information about meanings: for example, team members must know what businesses, operational source systems, and data fields the group is discussing, why they are discussing it, what the meaning of the concepts and terms is, and what information is known by themselves and others. At the same time, process common ground allows team members to be more effective in their information sharing: once they know how the group is working, this process knowledge helps them to know when they should perform specific actions that can help to make progress toward the shared goal (Convertino et al. 2009).

Principle 3: Communication is a multi-modal process

Communication involves more than just words or verbal, written or spoken, conversations. Less conventional forms of language such as vocal gestures, facial expressions, eye gaze, and postures also help people to make signals and exchange messages (Clark 1996, p. 180). It has been proposed that specific communication contexts can be described in terms of specific sets of grounding *constraints* (Clark and Krych 2004): the more constraints a media can provide the better the media is for facilitating common ground. This is mirrored in research regarding the use of media for specific situations (Daft et al. 1987; Kock 2005). In the context of complex DWH development, this principle also implies that mere written specifications of the business information requirements, the conceptual and logical data models, or the ETL concept may not be enough.

Clark's theory and principles have been tested in psycholinguistics by Bromme and collaborators (Bromme 2000; Bromme and Jucks 2001; Bromme et al. 2005a; Bromme et al. 2005b; Bromme et al. 1999; Jucks et al. 2008) in two special situations that are very similar to scenarios of ISD and DWH development projects:

- *Communication in case of high differences in knowledge between sender and receiver* (expert-layperson communication as can also be observed in financial DWH development projects between business users and IT experts). Expert-layperson communication is characterized by low common ground between the actors at the beginning of the communication process in which the common ground will not only be accumulated but also restructured. People usually have an "egocentric bias": if I know something, I am more likely to expect others to know it too (Clark 1996, p. 111). For a successful knowledge transfer a change in perspective is necessary, that is, 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). In particular Bromme and Jucks (2001) found that these CRFs are mainly determined by the participants' background and their specific education. Moreover 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, p. 93).
- *Communication in case of communication without face-to-face contact* (distributed projects as can also be observed in large multi-national financial DWH development projects). In case of written communication without face-to-face contact, direct feedback based on gestures and verbal intervention is not applicable. 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 missing understanding will often not be signaled in practice (Bromme and Jucks 2001, p. 93; see also Garrod and Anderson 1987).

We argue that these insights from communication theory can be transferred to financial DWH development projects. The financial industry shows a complex problem situation, many involved stakeholders, and lots of misunderstandings, especially regarding the exact meaning of terms and data fields (semantic heterogeneity). If we regard the interaction between the involved stakeholders as an instance of expert-layperson communication then our approach has to address corresponding communication issues. Achieving common ground therefore is the main goal of our approach. However, there is a high risk of illusion of evidence scenarios. These situations are difficult to detect because none of the involved persons are aware of the lack in common ground. One implication of the findings on expert-layperson communication is the use of feedback-loops to check if the knowledge transfer has been successful. Additionally, supporting application systems, tools, and documents must have a structure which is logical and understandable from both layperson's and expert's perspective. In case of written communication (requirements documents, logical and conceptual models, tools and databases) more effort for the anticipation of the layperson perspective is necessary. We possibly even have to encourage and motivate people to signal a missing understanding. This implicates the importance of strong interaction within groups and face-to-face communication (Bromme et al. 2004, p. 186).

Based on this we propose to follow *three strategies* that support and facilitate the development of both process and content common ground in financial DWH development projects. The three strategies are described in Table 1.

Table 1. Strategies for Supporting Common Ground in Financial DWH Development Projects

Strategy	Description	Principle addressed
(1) Improve requirements specification with regard to development of common ground	It may be useful to extend the specification and use formal, standardized descriptions whenever possible as to provide constraints, but to retain and encourage lots of not formalized areas as well as mechanisms and space for discussions and negotiations regarding meaning of requirements, concepts, and terms.	Principle 2 (selectively share information), principle 3 (provide constraints)
(2) Detect and signal remaining communication defects early	Early detection and signaling of remaining communication defects is necessary to avoid high fixing cost. For example, operational systems data analysis at early stages might allow early detection of errors.	Principle 1 (joint activity), principle 2 (share know process)
(3) Reduce scenarios with potential for illusion of evidence	Workshops, personal interaction, and on-site visits are expensive but may be needed, possibly coupled with mechanisms for signaling missing understandings.	Principle 1 (joint activity), principle 3 (multimodality)

Specification of the Approach

We iteratively developed the final configuration of the artifacts that are presented here (Baskerville et al. 2009). Our approach for financial DWH development was continuously refined as the method was put into use and exposed to environmental constraints in field studies (cf. section "Application and Evaluation"). The continuous application of the method led to adaptations and enhancements and resulted in redesigns of the approach, which is a common occurrence in Design Science research (Vaishnavi and Kuechler 2008, p. 25).

To establish common ground in financial DWH development, the strategies (cf. Table 1) are implemented as three artifacts:

1. A *template* for specifying information requirements for the domain of financial DWH completely and understandable wherever possible.
2. A *procedure model* comprising process definitions (i. e., what to do in a project phase) and organizational aspects (i. e., roles and responsibilities) for financial DWH development.
3. A *software tool* implementing template and procedure model. The tool provides views to support different stakeholder groups and communities of practice in financial DWH development projects.

Template for Information Requirements (Specification Format)

Project stakeholders from several (sub-) domains are either experts or laypersons on aspects of the domain in question (here: financial DWH domain). We propose that in such scenarios, a *domain-specific template* helps to define content common ground between stakeholders (design rationale: improve requirements specification, strategy 1, cf. Table 1). In order to discuss the meaning of things or concepts, we need to explicate our knowledge first, for example, by writing it down (Boisot and Canals 2004). Only if we try to formalize and codify the meaning of a thing or concept, we begin to notice that different meanings exist for this concept, and we can engage into discussions of those meanings. In general a template has to instantiate a meta-structure for content common ground in the domain specified by superior categories needed for requirements specification. For financial DWH, the template has to specify the relevant categories for ETL design (Rizzi et al. 2006; Vassiliadis et al. 2005), including basic business information requirements (for example, attribute definitions for data fields) and several additional facets (for example, meta-data for ETL processing). Appendix A presents the data model of the template and Table 2 gives an overview of the superior categories. A different template is needed for scenarios in other domains.

Table 2. Overview of Template Specification Format

Field Categories	Description
Supplementing structures and content of attributes	Some information about the content of the attributes cannot be covered easily by a simple description, for example, product catalogues, currency conversion, limit structures, customer structures, and so forth. Requirements pertaining these categories need to be specified and discussed.
Data quality needs	To fulfill both the expectation of the end user and legal requirements, data quality is of uttermost importance in financial DWH. For example, this category includes the specification of mandatory attributes, enhanced data quality rules, and concepts for checking the completeness of DWH data.
Data-processing meta data	The data load strategy and the technical design of issues such as load frequency play important roles for the development of the ETL job framework. Different parts of the information demand are processed on different levels, for example, daily (profit and loss (PNL) data), weekly (internal risk report), monthly (regulatory Basel II reporting), or yearly (balance sheet information). This data-processing meta-data has to be discussed.
Historical data requirements	For a historization concept, the definition of time periods for the storage of the data is needed, for example, for time series analysis or for the fulfillment of archiving requirements.
Granularity of data	Depending on the business information requirements, the granularity for different data may differ. For some analyses, transactional levels must be available (PNL calculations); others can be done on aggregated data.

Procedure Model

Our basic design rationale for the procedure model is that establishing *process* common ground requires stakeholders' knowledge of the intended (formal and informal) communication processes. Recent research on team decision-making and collaborative technology has shifted attention from team mental models to transactive models of sharing – knowledge of who knows what within a team; such models appear more appropriate for group tasks involving interdependency and role specialization (Convertino et al. 2009). We claim that knowing intended communication channels and all other team members' expertise, roles, and responsibilities in the project reduces the need for unintended communication between team members. Therefore we specify the sequence of a set of selected development phases, with explicit forward and feedback communication channels between roles in these phases.

Following transactive models of knowledge sharing, the development of *content* common ground in financial DWH projects can be enhanced by enabling team members to selectively determine what knowledge should be shared in the process. Therefore we specify output objects (for example, files and documents) for every single phase of our procedure model. We propose that the development of content common ground can be enhanced by using these outputs as *boundary objects*. Boundary objects are a form of boundary connection to bridge knowledge boundaries separating stakeholder groups and communities of practice (Bergman et al. 2007, p. 662; Brown and Duguid 2001, p. 105; Pawlowski and Robey 2004, p. 209). In order to actively participate in multi-stakeholder product design, such as the design of a DWH, and to actively support the design, for example, by indicating "successful progress" or "impending issues" in the design, Bergman et al. (2007) argue that boundary objects need to possess four features: (1) promote shared representation, (2) transform the design knowledge, (3) mobilize for design action, and (4) legitimize the design knowledge (Bergman et al. 2007, p. 551). Bergman et al. (2007) define objects embodying these four features as *design boundary objects* (DBO). Therefore our output objects are designed to conform to these four features.

Our procedure model (cf. Figure 2) follows an iterative, agile development approach (Cao et al. 2009) and differs from traditional "waterfall-based" methodologies often used in DWH development (Giorgini et al. 2008; Inmon 2005; Kimball and Caserta 2004; Mazón and Trujillo 2008; Moody and Kortink 2001). Sprints-like stages are used for procedure phases and are coupled with constant forward and feedback loops, data quality checks, and a focus on communication for detecting errors near the beginning through early deployment and quality testing of data integration and ETL models. Therefore *explicit* and extensive communication phases for presenting and negotiating the meaning of data requirements by all stakeholders are designed for the procedure model:

- *Business Information Requirements Analysis*. Design rationale: improve requirements specification (strategy 1, cf. Table 1). By interviewing business managers, other decision-makers, and business experts, DWH experts determine the project goal from an end user perspective and iteratively collect needed information. Output: end users' information requirements are documented as formally as possible as *target data models* using the domain-specific template.

- *Data Requirements Presentation and Negotiation*. Design rationale: detect communication defects early, reduce illusion of evidence, and improve requirements specification (strategies 1 to 3, cf. Table 1). DWH experts discuss all target data models with business experts, business owners of operational source systems, and corresponding IT experts during Joint Application Design (JAD)-like workshops (Dennis et al. 1999). Joint discussions with these "data suppliers" and operational experts are crucial because they allow early detection of misunderstandings, different meanings, and other communication problems such as illusion of evidence. If necessary, team members have to step back to the "business information requirements analysis" phase to clarify end user needs. Output: *mapping specification* between target data models and data sources.
- *ETL Design and Implementation*. Design rationale: detect communication defects early; reduce illusion of evidence (strategies 2 and 3, cf. Table 1). Technical as well as business problems concerning data sources and transformation are clarified during implementation between DWH experts and operational IT experts on a very detailed level (for example, issues such as load frequency for a specific data field). Open questions concerning these technical and operational details force team members to step back to the "data requirements presentation and negotiation" phase and improve and correct the specifications. The specification is gradually improved until all open issues are deemed to be addressed. Finally, an interim version of the DWH is deployed. Output: *reports* based on the data that is loaded into the DWH.
- *Sample Data Analysis*. Design rationale: detect communication defects early; reduce illusion of evidence (strategies 2 and 3, cf. Table 1). The reports as results of the previous phase are jointly analyzed in detail by DWH experts, operational IT experts, and business experts as regards completeness, correctness, and accuracy. Sample data analyses are conducted automatically and manually (for example, it is checked if reported measures account for the same numbers as a manual calculation of business experts). This is done to prevent illusion of evidence and to detect communication defects that still remain after the previous sprint cycle early on. This check of data quality is especially important if regulatory and legal requirements need to be fulfilled or if the data are used for risk management. Output: *data quality issues list* that compiles remaining data integration conflicts.
- *Clarification and Negotiation*. Design rationale: detect communication defects early; reduce illusion of evidence (strategies 2 and 3, cf. Table 1). The remaining data quality issues have to be resolved with all stakeholders. The resulting discussions feed directly back into the "business information requirements analysis" phase of the next cycle. Output: *edited data quality issue list* that is annotated with comments from all stakeholders.

The phases are repeated in cycles until all data quality issues are sufficiently resolved and the final DWH is constructed.

Software Tool

Bergman (2009, p. 405) suggests that specifications in a development process need to conform to the four features of a DBO. This can be supported by a software tool that manages the output DBOs and that affords selective sharing, specifying, and negotiating of information meanings. We argue that such a software tool (IT artifact) supports faster discovery of misunderstandings between all team members and thereby accelerates the creation of content common ground. This enables the transfer of knowledge across boundaries between stakeholder groups. Moreover, a software tool can act as a bracket connecting the different steps of the procedure model, thereby effectively supporting process common ground through guiding team members along the stages of the procedure model (design rationale: implement template and support procedure model, thus addressing all three strategies). In this, it is a DBO in its own right. The so-called *Data Requirements Tool (DaRT)* was iteratively developed to support the procedure model (cf. Figure 1Figure 1Figure 1 for a screen shot of the user interface, showing an exemplary form). DaRT implements the template (specification format) described above (cf. Appendix A for the underlying data model). Its navigation tree supports navigation of data requirements (left side of Figure 1); completeness of the requirements specification is enforced by the structured user interface for discussing and editing content (right side of Figure 1).

Database tables are defined in the technical configuration overview. The tool encourages the user to create a business view which contains a selection of fields coming from data tables. A combination of those views is grouped to business areas. Figure 1 shows client table 'Deals - External rating deal' and business area 'Standardized Approach'. The focused attribute 'RATING_AGENCY_ID' is shown in detail on the right side. Views can be set as mandatory or not in every business area on its own. General comments on table attributes can be enhanced by view-related comments. This supports discussions and negotiations of meanings directly in DaRT. A field status indicates the reached status of negotiation. Multiple status values can be used and defined individually to support each user organization best. Further information is provided to document the identified data source for the data requirements in this specific business area (source system, table, and field name). Additionally, a free text field allows further documentation (for example, meeting minutes, involved persons, time references, and so forth). These features make the tool a reliable information pool containing details of the financial DWH development process.

Every suggestion in DaRT reduces misunderstandings, enhances content common ground, and improves specification quality. When working together with the tool and documenting every single step, all involved people commit to the results of their work immediately and also control whether other involved team members document their work correctly. This helps detecting illusion of evidence scenarios. By applying the procedure model and using DaRT, people are literally "forced to talk to each other" about the problems and get into deep interaction. DaRT also allows versioning of these negotiation steps. The formalized structure of the tool also leads to a better documentation and eases tracking and control for project managers. The current status and status changes can promptly be reported. DaRT thus provides a central repository for the DWH design status. As a financial DWH is a central system with lots of stakeholders, lots of people can access this documentation. Project managers have a good overview regarding deficient data requirement negotiations by tracking the field status over time. Therefore problems can be detected early and directed to specialists. In addition DaRT is used to configure data quality rules. Those rules are directly derived into the database management system to provide technical checks in the ETL process. During the application in real-life field situations (cf. section "Application and Evaluation") lots of standard rules were implemented in DaRT. This knowledge about data sources can now be reused throughout future financial DWH development projects dealing with similar business areas. Furthermore, existing data requirement descriptions can be reused as well.

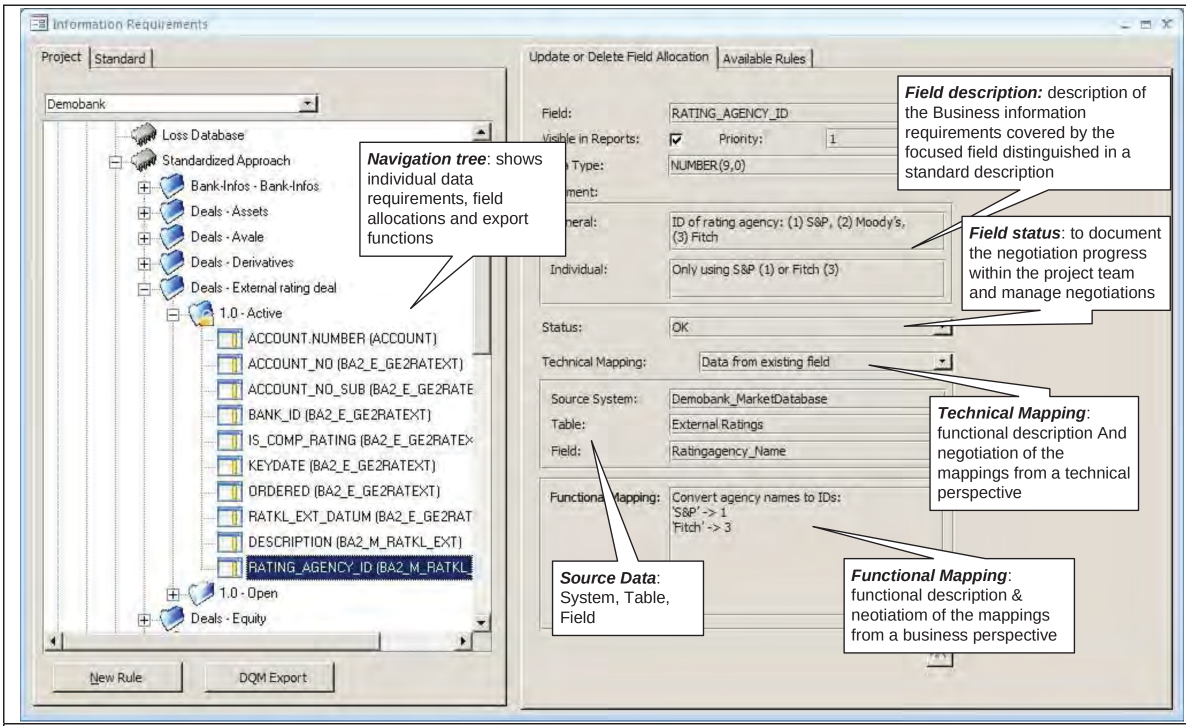

Figure 1. Exemplary Screenshot of DaRT Interface

Recap of the Approach

Figure 2 summarizes our approach and illustrates the iterative interplay of artifacts and roles in each cycle. At the beginning of a financial DWH development project, the business information requirements and the resulting data requirements are collected and codified as best as possible in a structured way, using mostly interview-based techniques (phase "business information requirements analysis").

These requirements are described in the specification format according to the domain-specific template and stored in a structured database within DaRT. An initial draft version of the data requirements specification is used to align all data requirements in the bigger picture and to derive the draft target data model. For each requirement the relevant data sources need to be identified in JAD-like workshops and face-to-face discussions (phase “data requirements presentation and negotiation”). The target data model is amended stepwise by detailing the data requirements and designing data flows from different source systems. During the workshops the project team discusses the data requirements, adds details to the descriptions to document their understanding of the requirements, and defines a status for each data requirement. People are forced to write-up more documentation for data requirements that lead to defects in communication, thereby allowing others to gain a better understanding of the defects. The data requirement status is also used by project managers to validate effort estimations and forecast the progress of the project.

Next, DWH experts and source system specialists use DaRT to document first rough ideas of how the source system data needs to be transformed to fit in the DWH (phase “ETL design and implementation”). The team uses these initial designs to add details, negotiate open issues, and improve the specification. The highly dynamic ETL implementation process requires a timely documentation to avoid misunderstandings. Therefore DaRT is used to document decisions taken by DWH experts and source system specialists.

As soon as possible, the team loads sample data in the DWH using the implemented data flows from one source system each (phase “sample data analysis”). This allows finding problems regarding design, implementation, or data quality as early as the first cycle.

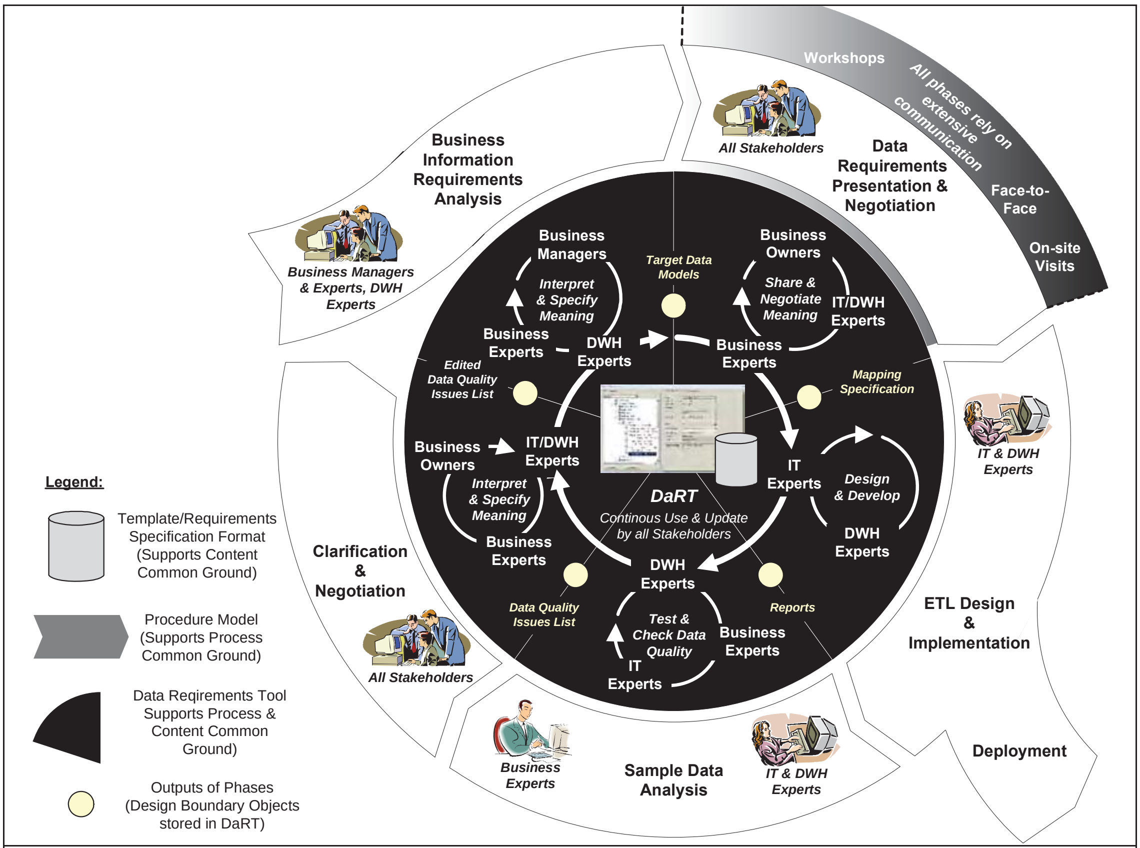

Figure 2. Summary of Approach for Financial DWH Development Projects and IT Artifact Support

Application and Evaluation

During design science research the intended (and unintended) impact of the artifacts needs to be scientifically evaluated to show their usefulness in solving the intended problems (Hevner et al. 2004; Winter 2008). This can involve comparing objectives and observed results in real-world settings. We therefore translate our theoretically motivated design into expectations about observable behavior in financial DWH development projects to enable its empirical evaluation. This results in the following propositions for evaluating effectiveness and utility of the artifacts:

The use of the artifacts supports the development of process and content common ground in financial DWH development projects. Therefore (1) it leads to earlier detection of errors, (2) it significantly enhances project performance (the development process), and (3) it significantly enhances the quality of the final product (the DWH).

Research Strategy and Design

For the evaluation of our artifacts in organizational settings, adequate research methods are required (Gregg et al. 2001). As a first step, we employed an *embedded multiple case study design* (Yin 2003, p. 49) in order to examine the usefulness of the artifacts during their iterative development. The six financial DWH projects of various banks in Europe that were selected as cases are typical for recent projects in the European financial services sector. The trigger for all selected projects was the supervisory requirements known as Basel II (BCBS 2006). These regulations demand specified risk calculation and risk treatment processes that have an impact on the whole structure of financial institutions and became law in most member states of the European Union in 2007. Table 3 gives an overview of the cases. All projects were conducted with the help of zeb/information.technology, a German consultancy which focuses on IT in the financial services industry.

The first set of cases (projects at Bank A-C) did not employ any of the developed artifacts whereas the second set of cases (projects at Bank D-F) did use some or all of the artifacts in different states of their development (cf. row "Treatment (artifact)" in Table 3). Following Lee (1989), natural controls that were already in place were also utilized. For example, focusing on persons who were participants in more than one of the cases held constant the people factors while comparing projects. Therefore control in the research design was already in place. Other examples of natural controls are the scope of the project and the parallel implementation of a new core banking system. Two of the authors are practitioners that have worked for more than seven years for zeb/information.technology. They participated in all three of the selected projects that were conducted without the approach (Bank A to C). Afterwards both participated in the design of the artifacts but not in the subsequent projects (Bank D to F). The other two authors were not involved in any project. They acted as unbiased observers and neutral interviewers in the case studies and during the evaluation of the artifacts. The analyses of the case studies in Bank D to F evaluated the artifacts and led to refinements of the design (for example, DaRT was a response to a need for managing the template).

Table 3. Cases Overview and Comparison

	Bank A	*Bank B*	*Bank C*	*Bank D*	*Bank E*	*Bank F*
Balance value	100 bn. EUR	70 bn. EUR	7 bn. EUR	4 bn. EUR	2 bn. EUR	230 bn. EUR
No. of countries	> 13	> 5	1	2	2	> 5
No. of subsidiaries	18	6	-	-	-	15
Person days (CONS)	~13,500	~5,000	~5,000	~4,000	~7,000	~3,000
Project scope	Europe	Europe	Germany	Germany	Germany	Europe
Project duration	> 5 years	> 2.5 years	> 2 years	> 1.5 years	> 2 years	> 1.5 years
No. of team members	> 100	> 150	> 20	> 25	> 15	> 25
Core system change	no	yes	yes	no	no	no
Perceived complexity[1]	very high	high	medium	very high	medium	high
Treatment (artifact)	-	-	-	procedure & template	procedure & DaRT prototype	procedure & DaRT

[1] collectively determined on a scale from 1 (very low) to 5 (very high) during a focus group meeting of all involved project managers

Direct observations by the first two authors, unstructured and semi-structured interviews, structured self-estimation surveys of project members, project documentation, and e-mail exchanges were used to generate data and collected in a case study database. We interviewed both project managers and project workers. Since at the time of the interviews most of the selected projects or related successor projects were still ongoing, this minimized the risk that participants displayed retrospective bias or that they had already forgotten something in their interactions with other project members. The collection of the data was started in March 2007 and ended in December 2008. During data collection, we conducted 17 interviews with 11 project members, spread over all six projects. The interviews were controlled by a semi-structured interview guideline and were either conducted by phone if the informants worked at the time in geographically remote locations or in face-to-face sessions at the organizations' site. The interview guideline was not shared with interviewees and was only used as a general outline. Coding techniques and checklists were used to connect data with the propositions (Miles and Huberman 1994, pp. 170-244; Yin 2003, pp. 109-138).[1]

Case Studies' Narratives and Analyses

All projects deal with the following typical setup. All banks consist of a head office and subsidiaries in Germany or different European countries. The projects aim to develop an application system which meets the regulatory requirements of Basel II and delivers reports to the financial supervisory authorities. One main task of the projects is to implement a central financial DWH. The focus of the projects lies in determining the relevant amount of capital required for credit risk for the total international bank portfolio of the banks. Primarily, data of transactions, collaterals, customers, and rating information of both head office and subsidiaries have to be delivered and consolidated into the central financial DWH. To achieve this, each subsidiary has to develop extraction jobs for their local databases and operational source systems (for example, for managing collaterals); the extracted data are then sent to head office and imported into the financial DWH.

As expected, semantic heterogeneity was high in all six cases. Different meanings of important IT-related and business-related concepts and terms existed in all banks. Several concepts were not immediately interpreted in the same way by all involved stakeholders (low initial content common ground in all six cases).

> *"[...] we had problems concerning demarcations of conceptions, what really is meant by turnovers and settlements of accounts in the context of ... in connection with credit cards. There, they always understood it differently in parts because you can definitely argue 'Is the disposal on the credit card a turnover or is it not a turnover until I have the credit card debit booked to the account?'"* (Interview with project member DF, translated by the authors)

> "*A typical example is the 'Buchwert'* [German: book value]*, which is really interpreted differently by each department. Whether these are people from risk management or the controlling people, or also the different units, everybody understands something a little bit different, so whether some interests are included or not or whether these are outlined separately ....*" (Interview with project member BK, translated by the authors)

Cross-Case Analysis for Bank A, Bank B, and Bank C (Cases with no Treatments)

Several problems resulting from a different understanding of the meaning of concepts surfaced in the projects that did not employ the approach but instead followed a more traditional, specification-driven methodology for DWH development. For most of the projects' runtime, it seemed as if project members from head office, different departments, and subsidiaries each "spoke different languages" because the interpretations of the published specification varied to a considerable degree (slow creation of content common ground in the cases of Bank A, B, and C).

> "*In fact, there are really thirteen data warehouses that have to be loaded identically, but which weren't loaded identically because how should one load them identically if nobody takes care that this happens? This was the case in the first implementation phase, that one just gave the documents to everybody and everybody did interpret them. And the results were thirteen interpretations. [...] This only surfaced in the discussions later [...]. There it made kind of 'Click! Ah, that is why so very much is not working here!' "* (Interview with project member JS, translated by the authors)

Moreover, throughout Bank A to C, the waterfall-based, specification-driven approach caused similar problems and made communication and knowledge exchange processes very difficult by being too constrained and too formalized (low process common ground in the cases of Bank A, B, and C).

> *"... one simply used examples to reach an agreement, if possible, and then one came out of the workshop. And then, naturally, every time there were callbacks and further inquiries, one met on the floor. Thus it was not concluded with the workshop; but then, it always was this special business case, and then they called us by phone or we met somewhere and they say 'By the way, I have another question' and 'How do we do this exactly?', and we discussed this together.*" (Interview with project manager WB, translated by the authors)

However, the mode-of-procedure within all three projects was changed from being specification-driven to being more communication-driven as the projects progressed. Self-estimated efforts for on-site meetings were up to three times higher after this than before. For example, so-called task forces were established for subsidiaries that had not been able to deliver data on time because their interfaces did not work properly. The project managers at head office

---

1 Due to space constraints, details on the interviewees and interview guideline are available from the authors on request. MaxQDA (http://www.maxqda.com/) was used for coding interview transcripts and for linking other data from the database.

voiced concern that just sending out the interface specification had not been enough to clarify the meaning of all important concepts. Personal face-to-face contact and even on-site visits were necessary for solving these misunderstandings. The creation of content common ground by intensive face-to-face communication and physical inspection was one of the central factors for project success according to the key informants within the projects (low initial process common ground, slow creation of process common ground by change to communication-driven procedure that helped to establish better content common ground in the cases of Bank A, B, and C).

> "*This then changed from the initial procedure of 'We send you a concept; look what you got and send us back the results' to a much more workshop-oriented procedure. So one did say 'We have to support you much more, we come over to you and look everything through in workshops together with you' and in a second step one conscripted the colleagues respectively and said 'We let our project member sit with you and he will support you for the next days and weeks'.*" (Interview with project manager TA, translated by the authors)

> "*... but the real turning point, where this issue had full management attention, came actually in April 2007, where one said the data delivery does not work at all ... and only then one saw what a shit this was ... and then did decide that we need so-called task forces, and only then I came back to the project, very different persons now became involved, who really were on-site [at the subsidiaries] and went through the data with them weekly, looked at the data fields and discussed every single thing.*" (Interview with project member BK, translated by the authors)

Consequently, the specification and the few initial workshops were not enough to generate a shared understanding. The technical face-to-face coaching and on-site visits were very important so that the stakeholders discussed the requirements and could analyze and negotiate if the specified requirements really were reasonable and meaningful for everybody (creation of content common ground). To recap, we observed slow creation of content and process common ground at Bank A, B, and C. This resulted in errors, quality issues with delivered data, and project performances were ranked lower by project managers compared to the other three projects (cf. row "Ranking" in Table 4).

Cross-Case Analysis for Bank D, Bank E, and Bank F (Cases with Treatments)

As shown above, high semantic heterogeneity and low initial content common ground also characterized the projects in Banks D, E, and F. The major difference is that these projects were using some or all of the approach's artifacts in some version (early design, prototype, or final configuration; cf. row "Treatment (artifact)" in Table 3). During the course of the projects, shared understanding emerged due to the daily interactions of all project members and stakeholders (early development of content common ground). The employed procedure model relied on the initial workshops of all involved stakeholders, and a lot of direct face-to-face communication ensured that ambiguity gradually was reduced. Most of the time was spent for specifying the data requests in face-to-face communication.

> *"[The procedure model] makes it easier in the implementation phase, but is more complicated during the coordination of data requests. [...] Because you have to discuss this longer so that they understand it correctly, that you make the point. [...] And that's why the coordination just took months, especially those events concerning loss control, this is so complicated, till you understand that yourself and because they had to implement that in the source system themselves, this took ages."* (Interview with project manager SK, translated by the authors)

The specification format was first used in the project at Bank D. At Bank E, the complete course of the project was accompanied by a prototype of DaRT for the first time, implemented by zeb/information.technology's internal software department.[2]

> *"At Bank D the tool was not existing [...] we had Excel sheets containing the data requests, and there not two persons could work simultaneously, you have version conflicts. Then in customizing a lot of people work with views, and data quality management is in the views and nobody knows why, if an error occurs [...] And that's why I figured it could be useful to have such a tool [...] where you have standard data requests and can describe the mappings."* (Interview with project manager SK, translated by the authors)

Due to the employed procedure supported by DaRT and the communication-intensive "data presentation and negotiation" phase, content common ground quickly emerged. After the first cycle, additional meetings and discussions of DWH experts, IT experts and business experts became necessary, for example, if data field descriptions had to be supplemented in order to make the meaning of terms clear and to reduce ambiguity. The prototype was used intensively for specifying data requests in these phases and for quality checks in the "sample data analysis" phase.

---

[2] DaRT is commercially available as a module of "zeb/data.quality-manager", see http://www.zebcontrol.com/.

As described above (cf. "Software Tool" section) DaRT allows for detailed consistency checks in both testing phases and operational loading phases, and generates reports showing errors and open issues. All project members including employees of Bank E and Bank F used DaRT for specifying the business view mappings according to source systems. Moreover, all project members used the tool for checking the consistency of the data fields and to mark open issues and problems. The use of the tool, coupled with the intensively communication-oriented procedure model, provided process common ground and facilitated the creation of content common ground within the project.

> *"So, DaRT, we basically enter all data requests that we have into it. [...]And in the next step, for each source system, a specialist [...] looked where he could get those data fields in his source system. And then you naturally always had inquiries [...] by setting a status marker accordingly and by sending the questions regarding those fields back to the business department, who could really use DaRT for directly giving the answers. And the next step was that those who should get the data out of the source systems, into the data warehouse of the bank, again looked into DaRT and there they found their ... yes, they quasi found the mappings already described there and could implement them relatively quickly."* (Interview with project member MK, translated by the authors)

During the coordination phase and the accompanying discussions DaRT proved to be extremely useful in order to really understand the situation and relationships at Bank E and Bank F. The tool documented the history of common ground creation and all project team members could refer back to it and use it as a guideline for open issues, problems and misunderstandings. Basically, the tool allowed for semantically mapping the meaning of concepts at Bank E and Bank F to data fields in the DWH and the operational source systems, and to engage into clarifying discussions about understanding issues. Together with the procedure model this provided process common ground and simplified the creation of content common ground, speeding up the process of sense-making. Open issues could be tracked and questions regarding meanings of terms could be sent to the business department. Employees in the business department tried to answer those questions directly in the tool and afterwards, the IT developers could continue directly with those fields. This was perceived as an enormous support for the procedure model.

Comparing the cases of Bank D, Bank E, and Bank F, which all used the new procedure model, DaRT – which was only used as a first prototype at Bank D – especially helped to follow through with the distinctive phases:

> *"Eventually, if you make such alignments, then this leads to the uncovering of any errors or that we ask any questions, and then they [the employees of the bank] give thought to those issues and then new solutions appear, so 'you don't really need this, this possibly is still wrong in our system', so you get more clarity in this, clearly. [At Bank E] we put a lot of emphasis on the standardization of the data requests, that wasn't that standardized before [...] and at Bank E the customer worked with the tool, continuously, and then a process came into life. [...] it isn't an agony to consolidate the data requests but a part of the job. [At Bank F it took] two and a half months, then we had the first connection, the complete system. At the others this took seven, eight months in parts [...]."* (Interview with project manager SK, translated by the authors)

Cross-Case Summary

For ranking the projects' performance we have taken the following approach. After conducting the case studies we contacted seven project managers involved in all six projects to invite them to participate in a focus group discussion with us. The purpose was to (a) discuss the projects and (b) to establish a ranking scheme on the basis of their expert opinions. These professionals had an average experience of five years with financial DWH development and created two to four financial DWH each during this period, indicating a considerable level of expertise.

During the workshop, a presentation was given by the researchers, explaining each project. Next, all participants were asked to rank the six projects with respect to the project's performance, understood as the quality of the process, the quality of the final product, and the project's efficiency using a scale of 1 to 6. For this scale, a rank of 1 indicates that a project is perceived as having the highest relative performance; a rank of 6 indicates the other end of the scale, representing the lowest relative performance. The participants were asked to give a full ordering, that is, assigning an equal rank to two or more project was not allowed. Afterwards the single rankings were presented, discussed among all workshop participants, and consolidated into one final ranking (presented in Table 4).

Summing up, all six projects are comparable in scope and complexity. The latter projects that used some or all of the components of the presented approach are ranked higher in performance than the first three projects that did not; the quality of product and process of the latter three projects were stated to be superior. The projects managers also ascribed this difference in performance, at least partially, to the artifacts. Based on our findings we cannot completely rule out that, as all projects have been conducted with the help of zeb/information.technology, the DWH consultancy's many aspects of abilities on DWH application will be increased with the experiences of the previous projects. Although we tried to employ natural controls such as interviewing project team members for both single and across projects, we cannot completely eliminate factors like the improvement on project abilities. These abilities, however, are also clearly supported by the artifacts.

> *"[…] and only at this point in time [the start of the project at Bank F] we finally had finished all the tools. So, DQM [data quality management], DaRT, and we had people and so on, which extremely look to it that if something doesn't work you still continue […] instead of botching. […] DaRT assists us, personally, very strongly because we do the DQM process with it, that is, [we] maintain all data quality management rules with it, keep hold of all data requirements with it. Because we thus have had four or five controller loops so far, we always could extract the data requirements quite fast. Everybody knows exactly where they have to check and have the same approach."* (Project manager SK during workshop, translated by the authors)

The experience and knowledge about data sources, for example, is at least partially stored within DaRT, existing data requirement descriptions are reused in lather phases of the financial DWH development project, and DaRT and the procedure let all team members know about their tasks, roles, and responsibilities. The cumulated expert knowledge about and experience with the procedure is taken to other financial DWH development projects by the involved team members of the consultancy. This important issue has to be addressed in further performance and utility evaluations of the artifacts.

Table 4. Summary of Cases Evaluation and Design Cycles

	Bank A	Bank B	Bank C	Bank D	Bank E	Bank F
Ranking	6th	5th	4th	3rd	2nd	1st
Treatment A (Template)	-	-	-	+ (in spreadsheets)	++ (in prototype)	++ (in DaRT)
Treatment B (Procedure)	-	-	-	+ (without tool)	++	++
Treatment C (DaRT)	-	-	-	-	+ (prototype)	++ DaRT
Observed Effects	Illusions of evidence & late detection of errors were observed (low content common ground). Once the procedure shifted from a specification-driven approach to a more communication-driven approach common ground was created.	Communication problems & late detection of errors were observed (low content common ground). The procedure shifted to a more communication-driven approach. Problems were realized earlier than in Bank A.	Communication problems, illusions of evidence & late detection of errors were observed. The procedure was shifted to a more communication-driven approach at an earlier stage than in Bank A or Bank B.	Early existence of content & process common ground. Early detection of errors, better communication and better project performance compared to Bank A-C.	Template & procedure helped create content and process common ground DaRT prototype helped create process common ground. Better communication and better project performance compared to Bank A-D.	Template & procedure helped create content and process common ground. DaRT helped create process common ground. Early detection of errors, better communication and better project performance compared to Bank A-F.

Conclusion and Outlook

The four overarching objectives for DWH support of management decision-making processes identified by March and Hevner (2007) – integration, implementation, intelligence, and innovation – assume a high affiliation between the stakeholders participating in a DWH development project. Until today effective techniques (a) for collecting information needs and requirements and (b) for translating those requirements into conceptual models based on a common vocabulary between IT experts and decision-makers are missing (Jarke et al. 2009; Rizzi et al. 2006). As demonstrated through our evaluation, our approach successfully addresses these problems for the case of financial DWH development projects. Integrating principles grounded in communication theory is promising for increasing the performance of financial DWH development projects and avoiding ambiguities in knowledge representation due to semantic heterogeneity. Two characteristics are significant to avoid common problems:

- Defining and providing mechanisms for supporting content common ground prior to financial DWH development project start is the basis for avoiding errors early on rather than resolving them later. Therefore time-consuming ex-post data quality checks and correction become dispensable.

- Guiding project team members during financial DWH development and creating process common ground is of substantial importance, since we can assure the compliance with the procedure model and thus the creation of content common ground only in this way.

Thus, the main contribution to practitioners concern projects' time and costs. However, analyzing costs-benefits, several issues have to be taken into account, which seem to decrease the benefit of the approach at a first glance. The procedure model relies on intensive communication of all stakeholders, not just on specifying information requirements in written form, and this can be very time-consuming. We found that defining, presenting, discussing and negotiating the requirements is indeed quite time-consuming. Taking into account that in most financial DWH development projects requirements are defined anyway (although in absence of methodical support), and that early detection of errors due to early creation of content common ground trade-offs early communication costs against late coordination and fixing costs of errors, this cost issue is moderated. Moreover, once the requirements specification format is defined, it is reusable for further projects in the organization or domain.

Looking at the financial DWH development process itself, the suspected slow-down turned out not to be significant in the case studies. The execution of the procedure model including defining, discussing, and negotiating is fast enough not to be recognized by project team members as a burden and leads to less overall communication because of significantly reduced errors and illusions of evidence. After our first field studies, project team members expressed that they were not slowed down by the approach overall, but even sped up, because they did not need to think about later errors and issues of data quality any more. However, the most promising aspect of the approach is that semantic heterogeneity is overcome by intensive communication and development of common ground between all stakeholders. With regard to the use of DBOs as presented by Bergman et al. (2007), we can confirm that DBOs help to resolve the ambiguity and uncertainty associated with semantic heterogeneity and complex requirements. However, Bergman et al. (2007) mainly concentrate on the features that boundary objects need to reflect in order to become DBOs; this gives no concrete guidance for practitioners on how to implement these features into DBOs. Our research shows an example of how DBOs can be designed and used within ISD.

The case studies we have conducted provide a first understanding of the benefit of the approach, but they have to be extended to a significant population of projects to be able to score it precisely. Since the cases took place in real-life environments, we cannot rule out that better performance was achieved because of other causes than due to our approach. Likewise, we focused on communication processes. Other factors (users' capabilities, characteristics and goals, institutional contexts, power, or culture) may be important and might need to be supported by other means as well. Another obvious limitation of the evaluation relates to the ranking scheme used for performance measurement. Clearly, the derived ranking has a small empirical basis as it relies on the involvement of seven project managers only. This raises the importance of further studies to arrive at a more rigorous evaluation, for example, through experimentation and surveys. However, as the qualitative content analysis reveals, project members attribute the better performance to the artifacts and are quite positive about it. Further evaluations concerning applicability and acceptance as well as efficiency will be subject of empirical studies to be performed in the short term.

Our main contribution to research is the combination of the two so far separated research areas of ISD and communication theory. Linguistic alignment in interactions between humans plays a critical role in achieving successful communication; there is now growing evidence that the same processes are operative in interactions between humans and computers, and indeed that they occur to a greater extent (Branigan et al. 2010). Linguistics and communication theory, however, have until now only seldom been applied for studying or supporting ISD. Applications of communication theory in ISD open new research opportunities, for example, in the area of ISD stakeholder communication (expert-layperson). The explication of expert knowledge and support of expert-layperson communication in ISD, which are nearly impossible without mechanisms and constraints for creating common ground, are further very promising areas of research. Finally, applying concepts from communication theory and linguistics in IS research permits a new field of application and the evaluation of these concepts.

Our future research will focus on further evaluating the proposed approach. In the short-term, the approach will be instantiated for different DWH development projects and for other application scenarios that are characterized by high semantic heterogeneity as well. In particular, the capability of our approach to increase the efficiency of distributed ISD and its acceptance will be evaluated. Long-term research will evaluate on a larger scale to what extent the development of common ground can improve the knowledge representation and communication processes in ISD projects.

Acknowledgements

We are grateful to the editors and the anonymous reviewers for very helpful feedback and advice. In addition we would like to thank Sven Krämer and Thomas Abel (zeb/information.technology) for their support and constructive contribution to this work.

Appendix A: Data Model

Figure 3 shows a simplified Entity Relationship Diagram (ERD) for an excerpt of the DaRT relational database that implements the specification format (domain-specific template). Martin's notation (crow's foot notation) for cardinality and optionality is used to show the relationships between the tables (Finkelstein 1989). Open circles indicate an optional relationship, while vertical slashes indicate a required relationship. The 'crow's feet' indicate that many tuples (i. e., many rows) can take part in the relationship, while a single slash indicates that only one tuple (i.e., one row) may take part. Dotted lines are used to indicate relationships that are not completed.

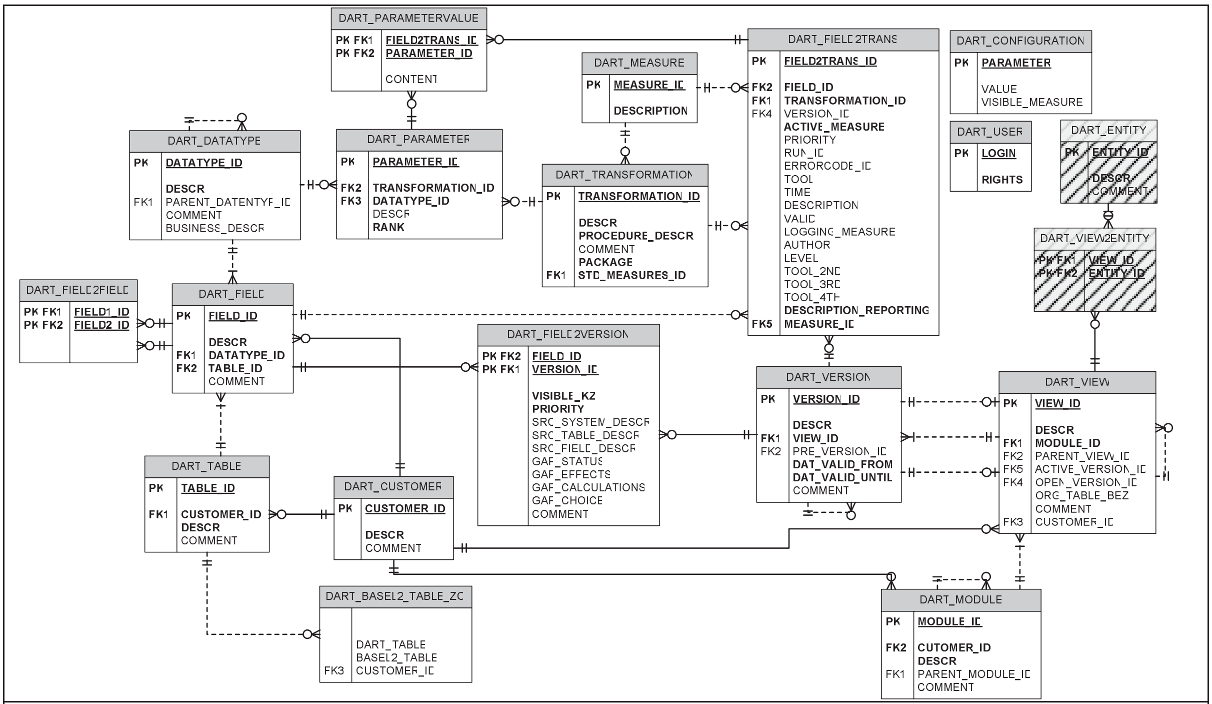

Figure 3. Data Model of Template in DaRT (Excerpt)

References

Abell, A., Samos, J., and Saltor, F. 2001. "A Framework for the Classification and Description of Multidimensional Data Models," *12th International Conference on Database and Expert Systems Applications (DEXA'01)*, Munich, Germany, pp. 668-677.

Ågerfalk, P.J. 2004. "Investigating actability dimensions: a language/action perspective on criteria for information systems evaluation," *Interacting with Computers* (16:5), pp. 957-988.

Alavi, M., and Tiwana, A. 2002. "Knowledge integration in virtual teams: The potential role of KMS," *Journal of the American Society for Information Science and Technology* (53:12), pp. 1029-1037.

Auramäki, E., Hirschheim, R., and Lyytinen, K.J. 1992. "Modelling Offices Through Discourse Analysis: The SAMPO Approach," *Computer Journal* (35:4), pp. 342-352.

Baldwin, L.P., Irani, Z., and Love, P.E.D. 2001. "Outsourcing Information Systems: Drawing Lessons from a Banking Case Study," *European Journal of Information Systems* (10:1), pp. 15-24.

Ballard, C., Herreman, D., Schau, D., Bell, R., Kim, E., and Valencic, A. 1998. *Data Modeling Techniques for Data Warehousing. IBM Redbooks SG24-2238-00*.

Baskerville, R., Pries-Heje, J., and Venable, J. 2009. "Soft design science methodology," in: *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*. Philadelphia, PA, USA: ACM.

BCBS. 2006. *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version*. Basel, Switzerland: Bank for International Settlements, Basel Committee on Banking Supervision (BCBS).

Behrmann, W., and Räkers, M. 2008. "Specifics of Financial Data Warehousing and Implications for Management of Complex ISD Projects," *16th European Conference on Information Systems (ECIS 2008)*, Galway, Ireland.

Bergman, M. 2009. "Requirements' Role in Mobilizing and Enabling Design Conversation," in: *Design Requirements Engineering: A Ten-Year Perspective,* K. Lyytinen, P. Loucopoulos, J. Mylopoulos and B. Robinson (eds.). Springer, pp. 394-407.

Bergman, M., Lyytinen, K., and Mark, G. 2007. "Boundary Objects in Design: An Ecological View of Design Artifacts," *Journal of the Association for Information Systems* (8:11), pp. 546-568.

Bernstein, P.A., Melnik, S., Michalis, P., and Quix, C. 2004. "Industrial-strength schema matching," *ACM SIGMOD Record* (33:4), pp. 38-43.

Böhnlein, M., and Ulbrich-vom Ende, A. 1999. "Deriving Initial Data Warehouse Structures from the Conceptual Data Models of the Underlying Operational Information Systems," *Second ACM International Workshop on Data warehousing and OLAP*, Kansas City, pp. 15-21.

Boisot, M., and Canals, A. 2004. "Data, information and knowledge: have we got it right?," *Journal of Evolutionary Economics* (14:1), pp. 43-67.

Branigan, H.P., Pickering, M.J., Pearson, J., and McLean, J.F. 2010. "Linguistic alignment between people and computers," *Journal of Pragmatics* (42:9), pp. 2355-2368.

Bromme, R. 2000. "Beyond one's own perspective: The psychology of cognitive interdisciplinarity," in: *Practising interdisciplinarity,* P. Weingart and N. Stehr (eds.). Toronto, Canada: Toronto University Press, pp. 115-133.

Bromme, R., and Jucks, R. 2001. "Wissensdivergenz und Kommunikation: Lernen zwischen Experten und Laien im Netz," in: *Partizipation und Interaktion im virtuellen Seminar,* F.W. Hesse and H.F. Friedrich (eds.). Münster, Germany: Waxmann, pp. 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*), pp. 176-188.

Bromme, R., Jucks, R., and Runde, A. 2005a. "Barriers and Biases in Computer-Mediated Expert-Layperson-Communication," in: *Barriers and Biases in Computer-Mediated Knowledge Communication,* R. Bromme, F.W. Hesse and H. Spada (eds.). New York, NY, USA: Springer, pp. 89-118.

Bromme, R., Jucks, R., and Wagner, T. 2005b. "How to refer to lsquodiabetesrsquo? Language in online health advice," *Applied Cognitive Psychology* (19:5), pp. 569-586.

Bromme, R., Nückles, M., and Rambow, R. 1999. "Adaptivity and Anticipation in Expert-Laypeople Communication," *AAAI Fall Syposium on 'Psychological Models of Communication in Collaborative Systems*, North Falmouth, MA, USA.

Brown, J.S., and Duguid, P. 2001. "Knowledge and Organization: A Social-Practice Perspective," *Organization Science* (12:2), pp. 198-213.

Campbell, A.E., and Shapiro, S.C. 1995. "Ontologic Mediation: An Overview," *IJCAI Workshop on Basic Ontological Issues in Knowledge Sharing*, pp. 16-25.

Cao, L., Mohan, K., Peng, X., and Ramesh, B. 2009. "A framework for adapting agile development methodologies," *European Journal of Information Systems* (18:4), pp. 332-343.

Carroll, J.M., Rosson, M.B., Convertino, G., and Ganoe, C.H. 2006. "Awareness and teamwork in computer-supported collaborations," *Interacting with Computers* (18:1), pp. 21-46.

Cavero, J., Piattini, M., and Marcos, E. 2001. "MIDEA: A Multidimensional Data Warehouse Methodology," *3rd International Conference on Enterprise Information Systems (ICEIS'01)*, Setubal, Portugal, pp. 138-144.

Chenoweth, T., Schuff, D., and Louis, R. 2003. "A method for developing dimensional data marts," *Communications of the ACM* (46:12), pp. 93-98.

Clark, H.H. 1992. *Arenas of Language Use*. Chicago, IL, USA: Chicago University Press.

Clark, H.H. 1996. *Using Language*. New York, NY, USA: Cambridge University Press.

Clark, H.H., and Brennan, S.E. 1991. "Grounding in Communication," in: *Perspectives on Socially Shared Cognition,* L.B. Resnick, J.M. Levine and S.D. Teasley (eds.). Washington D.C.: American Psychological Association, pp. 127-149.

Clark, H.H., and Krych, M.A. 2004. "Speaking while monitoring addressees for understanding," *Journal of Memory and Language* (50:1), pp. 62-81.

Clarke, R.J. 2001. "Studies in organisational semiotics: an introduction," *Australasian Journal of Information Systems* (8:2), pp. 75-82.

Convertino, G., Mentis, H.M., Rosson, M.B., Slavkovic, A., and Carroll, J.M. 2009. "Supporting content and process common ground in computer-supported teamwork," *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, Boston, MA, USA: ACM, pp. 2339-2348.

Corcho, O., Losada, S., Martínez Montes, M., Bas, J.L., and Bellido, S. 2005. "DIP: Data, Information and Process Integration with Semantic Web Services, WP10: Case study eBanking. Financial Ontology. DIP deliverable D10.3." Retrieved 2009-06-16, from http://dip.semanticweb.org/documents/D10.3.pdf.

Corvera Charaf, M., and Rosenkranz, C. 2010. "Natural Language Alignment As a Process – Applying Functional Pragmatics in Information Systems Development," *18th European Conference on Information Systems (ECIS 2010)*, Pretoria, South Africa.

Daft, R.L., Lengel, R.H., and Trevino, L.K. 1987. "Message Equivocality, Media Selection, and Manager Performance: Implications for Information Systems," *MIS Quarterly* (11:3), pp. 355-366.

Dennis, A.R., Hayes, G.S., and Daniels Jr., R.M. 1999. "Business process modeling with group support systems," *Journal of Management Information Systems* (15:4), pp. 115-142.

Do, H.-H., and Rahm, E. 2007. "Matching large schemas: Approaches and evaluation," *Information Systems* (32:6), pp. 857-885.

Doan, A., Madhavan, J., Dhamankar, R., Domingos, P., and Halevy, A. 2003. "Learning to match ontologies on the Semantic Web," *The VLDB Journal* (12:4), pp. 303-319.

Ehrig, M., and Staab, S. 2004. "QOM – Quick Ontology Mapping," in: *The Semantic Web – ISWC 2004. Proceedings of the Third International Semantic Web Conference, Hiroshima, Japan, November 7-11, 2004,* S.A. McIlraith, D. Plexousakis and F.v. Harmelen (eds.). Berlin, Germany et al.: Springer, pp. 683-697.

Finkelstein, C. 1989. *An Introduction to Information Engineering. From Strategic Planning to Information Systems*. Reading, MA, USA: Addison-Wesley Publishing.

Flores, F., Graves, M., Hartfield, B., and Winograd, T. 1988. "Computer systems and the design of organizational interaction," in: *ACM Transactions on Information Systems*. pp. 153-172.

Garrod, S., and Anderson, A. 1987. "Saying what you mean in dialogue: A study in conceptual and semantic coordination," *Cognition* (27:2), pp. 181-218.

Giorgini, P., Rizzi, S., and Garzetti, M. 2008. "GRAnD: A goal-oriented approach to requirement analysis in data warehouses," *Decision Support Systems* (45:1), pp. 4-21.

Giovinazzo, W. 2000. *Object-Oriented Data Warehouse Design. Building a Star Schema*. Upper Saddle River, NJ, USA: Prentice Hall.

Goldkuhl, G., and Lyytinen, K. 1982. "A language action view of information systems," *3rd International Conference on Informations Systems (ICIS 1982)*, Houston, TX, USA.

Golfarelli, M., Maio, D., and Rizzi, S. 1998. "Conceptual Design of Data Warehouses from E/R Schemes," *31st Hawaii International Conference On Systems Science (HICSS1998)*, Kohala Coast, HI: IEEE, pp. 334-343.

Golfarelli, M., and Rizzi, S. 1998. "A Methodological Framework for Data Warehouse Design," *ACM First International Workshop on Data Warehousing and OLAP (DOLAP)*, Washington, D. C., USA: ACM, pp. 3-9.

Gregg, D.G., Kulkarni, U.R., and Vinzé, A.S. 2001. "Understanding the Philosophical Underpinnings of Software Engineering Research in Information Systems," *Information Systems Frontiers* (3:2), pp. 169-183.

Hansen, S., and Lyytinen, K. 2010. "Challenges in Contemporary Requirements Practice," *Proceedings of the [43th] Annual Hawaii International Conference on System Sciences*, pp. 1-11.

Hevner, A., March, S., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp. 75-105.

Hevner, A.R., and March, S.T. 2003. "The Information Systems Research Cycle," *Computer* (36:11), pp. 111-113.

Holmqvist, B. 1989. "Work, Language and Perspective: An Empirical Investigation of the Interpretation of Computer-Based Information Systems," *Scandinavian Journal of Information Systems* (1:1), pp. 72-96.

Holten, R. 2003. "Specification of Management Views in Information Warehouse Projects," *Information Systems* (28:7), pp. 709-751.

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), pp. 1-21.

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).

Inmon, W.H. 2005. *Building the data warehouse*, (4 ed.). Indianapolis: Wiley.

Jarke, M., Lyytinen, K., Mylopoulos, J., Kappel, G., Leite, J., Mark, G., Ramesh, B., Schmitz, D., and Sutcliffe, A.G. 2009. "High-Impact Requirements for Software-Intensive Systems:Seminar Outlines and Working Group Summaries," *Perspectives Workshop: Science of Design: High-Impact Requirements for Software-Intensive Systems,* M. Jarke, K. Lyytinen and J. Mylopoulos (eds.), Dagstuhl, Germany: Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany.

Jucks, R., Becker, B.-M., and Bromme, R. 2008. "Lexical Entrainment in Written Discourse: Is Experts' Word Use Adapted to the Addressee?," *Discourse Processes* (45:6), pp. 497-518.

Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H. 2004. "Interactive robots as social partners and peer tutors for children: a field trial," *Human-Computer Interaction* (19:1), pp. 61-84.

Kimball, R., and Caserta, J. 2004. *The Data Warehouse ETL Toolkit, Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. Wiley & Sons.

Kimball, R., and Ross, M. 1996. *The Data Warehouse Toolkit. Practical Techniques for Building Dimensional Data Warehouses*. New York, NY, USA et al.: John Wiley & Sons.

Kock, N. 2005. "Media Richness or Media Naturalness? The Evolution of Our Biological Communication Apparatus and its Influence on Our Behavior Toward e-Communication Tools," *IEEE Transactions on Professional Communication* (48:2), pp. 117-130.

Koschmann, T., and LeBaron, C.D. 2003. "Reconsidering common ground: examining Clark's contribution theory in the OR," *8th European Conference on Computer Supported Cooperative Work*, Helsinki, Finland: Kluwer Academic Publishers.

Lee, A.S. 1989. "A Scientific Methodology for MIS Case Studies," *MIS Quarterly* (13:1), pp. 32-50.

Levina, N., and Vaast, E. 2005. "The Emergence of Boundary Spanning Competence in Practice: Implications for Implementation and Use of Information Systems," *MIS Quarterly* (29:2), pp. 335-363.

Madhavan, J., Bernstein, P.A., and Rahm, E. 2001. "Generic Schema Matching with Cupid," in: *Proceedings of the 27th International Conference on Very Large Data Bases*. Morgan Kaufmann Publishers Inc.

Maglio, P.P., Matlock, T., Gould, S.J., Koons, D., and Campbell, C.S. 2002. "On understanding discourse in human-computer interaction," *24th Annual Conference of the Cognitive Science Society,* W.D. Gray and C. Schunn (eds.), Fairfax, VA, USA, pp. 602-607.

Mäkelä, T., Rommel, K., Uskonen, J., and Wan, T. 2007. "Towards a Financial Ontology – A Comparison of e-Business Process Standards." Retrieved 2009-06-16, from http:////www.soberit.hut.fi/T-86/T-86.5161/2007/FinancialOntology_final.pdf.

Malinowski, E., and Zimányi, E. 2008. "Designing Conventional Data Warehouses," in: *Advanced Data Warehouse Design*. pp. 251-313.

March, S.T., and Hevner, A.R. 2007. "Integrated decision support systems: A data warehousing perspective," *Decision Support Systems* (43:3), pp. 1031-1043.

Mazón, J.-N., Pardillo, J., and Trujillo, J. 2007. "A Model-Driven Goal-Oriented Requirement Engineering Approach for Data Warehouses," in: *Advances in Conceptual Modeling – Foundations and Applications*. pp. 255-264.

Mazón, J.-N., and Trujillo, J. 2008. "An MDA approach for the development of data warehouses," *Decision Support Systems* (45:1), pp. 41-58.

McCarthy, J.C., Miles, V.C., and Monk, A.F. 1991. "An experimental study of common ground in text-based communication," in: *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*. New Orleans, LA, USA: ACM.

McFarlane, D.C., and Latorella, K.A. 2002. "The Scope and Importance of Human Interruption in Human-Computer Interaction Design," *Human-Computer Interaction* (17:1), pp. 1-61.

McLaughlin, R. 2004. "Data in financial institutions," *Journal of Financial Transformation* (11), pp. 75-80.

Miles, M.B., and Huberman, A.M. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*, (2 ed.). Thousand Oaks, CA, USA: Sage.

Moody, D., and Kortink, M. 2001. "From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design," *3rd International Workshop on Design and Management of Data Warehouses (DMDW'01)*, Interlaken, Switzerland, pp. 1-10.

Nova, N., Sangin, M., and Dillenbourg, P. 2008. "Reconsidering Clark's Theory in CSCW," *8th International Conference on the Design of Cooperative Systems (COOP'08)*, Carry-le-Rouet, France.

Olson, G.M., and Olson, J.S. 2000. "Distance Matters," *Human-Computer Interaction* (15:2), pp. 139-178.

Pawlowski, S., and Robey, D. 2004. "Bridging User Organizations: Knowledge Brokering and the Work of Information Technology Professionals," *MIS Quarterly* (28:4), pp. 645-672.

Rahm, E., and Bernstein, P.A. 2001. "A survey of approaches to automatic schema matching," *The VLDB Journal* (10:4), pp. 334-350.

Rizzi, S., Abell, A., Lechtenbörger, J., and Trujillo, J. 2006. "Research in data warehouse modeling and design: dead or alive?," in: *Proceedings of the 9th ACM international workshop on Data warehousing and OLAP*. Arlington, Virginia, USA: ACM.

Robillard, P.N. 1999. "The role of knowledge in software development," *Communications of the ACM* (42:1), pp. 87-92.

Saleem, K., Bellahsene, Z., and Hunt, E. 2008. "PORSCHE: Performance ORiented SCHEma mediation," *Information Systems* (33:7-8), pp. 637-657.

Schoop, M. 2001. "An Introduction to the Language-Action Perspective " *CM SIGGROUP Bulletin* (22:2).

Shvaiko, P., and Euzenat, J. 2005. "A Survey of Schema-Based Matching Approaches," in: *Journal on Data Semantics IV,* S. Spaccapietra (ed.). pp. 146-171.

Simperl, E.P.B., and Tempich, C. 2006. "Ontology Engineering: A Reality Check," (4275), p. 836.

Sjöström, J., and Goldkuhl, G. 2005. "The Semiotics of User Interfaces," in: *Virtual, Distributed and Flexible Organisations: Studies in Organisational Semiotics,* K. Liu (ed.). New York, USA et al.: Kluwer Academic, pp. 217-236.

Skoutas, D., and Simitsis, A. 2006. "Designing ETL processes using semantic web technologies," in: *Proceedings of the 9th ACM International Workshop on Data Warehousing and OLAP*. Arlington, Virginia, USA: ACM.

Tan, M., and Teo, T.S.H. 2000. "Factors Influencing the Adoption of Internet Banking," *Journal of the Association for Information Systems* (1:1), pp. 1-42.

Trujillo, J., Palomar, M., Gómez, J., and Song, I. 2001. "Designing Data Warehouses with OO Conceptual Models," *IEEE Computer* (34), pp. 66-75.

Tryfona, N., Busborg, F., and Borch Christiansen, J.G. 1999. "starER. A Conceptual Model for Data Warehouse Design," *ACM Second International Workshop on Data Warehousing and OLAP (DOLAP'99)*, Kansas City, MO, USA: ACM, pp. 3-8.

Uschold, M., Healy, M., Williamson, K., Clark, P., and Woods, S. 1998. "Ontology reuse and application," pp. 179–192.

Vaishnavi, V.K., and Kuechler, W. 2008. *Design Science Research Methods and Patterns. Innovating Information and Communication Technology*. Boca Raton, FL, USA: Auerbach Publications, Taylor & Francis Group.

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), pp. 492-525.

Vassiliadis, P., Simitsis, A., and Skiadopoulos, S. 2002. "Conceptual modeling for ETL processes," in: *Proceedings of the 5th ACM International Workshop on Data Warehousing and OLAP*. McLean, Virginia, USA: ACM.

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), Dec2004, pp. 435-450.

Watson, H.J., and Wixom, B.H. 2007. "The Current State of Business Intelligence," *Computer* (40:9), pp. 96-99.

Weir, R., Peng, T., and Kerridge, J. 2003. "Best practice for implementing a data warehouse: A review for strategic alignment," *Proceedings of the 5th Intl. Workshop (DMDW'2003),* H.-J. Lenz, P. Vassiliadis, M. Jeusfeld and M. Staudt (eds.), Berlin, Germany.

Winter, R. 2008. "Design science research in Europe," *European Journal of Information Systems* (17:5), pp. 470-475.

Wixom, B.H., and Watson, H.J. 2001. "An Empirical Investigation of the Factors Affecting Data Warehousing Success," *MIS Quarterly* (25:1), pp. 17-41.

Yin, R.K. 2003. *Case Study Research: Design and Methods*, (3 ed.). Thousand Oaks, CA, USA et al.: SAGE Publications.