

Entwicklung einer Methode zur Stammdatenintegration

DISSERTATION
der Universität St. Gallen,
Hochschule für Wirtschafts-,
Rechts- und Sozialwissenschaften (HSG),
zur Erlangung der Würde eines
Doktors der Wirtschaftswissenschaften

vorgelegt von

Alexander Schmidt

aus

Deutschland / Bulgarien

Genehmigt auf Antrag der Herren

Prof. Dr. Hubert Österle

und

Prof. Dr. Reinhard Jung

Dissertation Nr. 3848

Die Universität St. Gallen, Hochschule für Wirtschafts-, Rechts- und Sozialwissenschaften (HSG), gestattet hiermit die Drucklegung der vorliegenden Dissertation, ohne damit zu den darin ausgesprochenen Anschauungen Stellung zu nehmen.

St. Gallen, den 26. Oktober 2010

Der Rektor:

Prof. Ernst Mohr, PhD

Vorwort

Obwohl die Dissertation nur ein einziger Name ziert, wäre ihre Fertigstellung ohne die menschliche und wissenschaftliche Unterstützung so vieler Personen nicht möglich gewesen. Bei diesen möchte ich mich an dieser Stelle herzlich bedanken.

Der grösste Dank gebührt meiner Familie. Meinen Eltern, Maria und Bernd Schmidt, die für mich stets Vorbild und hilfreiche Unterstützer waren, danke ich für ihre liebevolle Erziehung, die für mich ein unschätzbar wertvolles Gut ist. Ohne sie würde ich nicht an dem Punkt in meinem Leben stehen, an dem ich nun angekommen bin. In gleichem Masse möchte ich meiner lieben Schwester Silvia Weissmann sowie Bruno und Mia Weissmann für ihre Herzlichkeit und ihren moralischen Rückhalt danken. Nicht unerwähnt möchte ich zudem die moralische Unterstützung durch Marianne Rauh, Luise Lorenz, Rudi Schmidt und Alexander Terziev während der gesamten Dissertationszeit lassen und mich dafür herzlich bei ihnen bedanken.

Besonderen Dank schulde ich Murielle Rüdy. Durch ihre Liebe und ihr anhaltendes Verständnis war sie diejenige, der ich während meiner Zeit in St. Gallen am meisten verdanke und mit der ich die schönsten Erinnerungen verbinde. Ihr Wille und ihre Kraft, auch in schwierigen Situationen nicht aufzugeben, sind für mich beispielhaft und haben mir stets dabei geholfen, eigene „Probleme“ während der Promotionszeit zu relativieren. Ihrer Familie danke ich dafür, dass sie wesentlich dazu beitrug, in der Schweiz ein Umfeld vorzufinden, in dem ich mich auch weit weg von zu Hause heimisch und wohl fühlen konnte.

Prof. Dr. Hubert Österle danke ich für die wissenschaftliche Betreuung meiner Dissertation, die hervorragenden Arbeitsbedingungen am Institut und das praxisnahe Forschungsumfeld. Prof. Dr. Reinhard Jung gilt mein Dank für die Übernahme des Korreferats und die Zeit, die er sich für regelmässiges Feedback genommen hat.

Die vorliegende Arbeit entstand im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter des Kompetenzzentrums Corporate Data Quality am Institut für Wirtschaftsinformatik der Universität St. Gallen sowie bei SAP Research CEC St. Gallen. Daher möchte ich mich besonders bei dem Leiter des Kompetenzzentrums Boris Otto nicht nur für die fachliche und persönliche Betreuung bedanken, sondern vor allem für das stets konstruktive Feedback sowie die äusserst angenehme und fruchtbare Arbeitsatmosphäre in einem menschlich und fachlich ausgezeichneten Team. Zu diesem Team zählen unter anderem Alain Benz, Philipp Osl, Andreas Reichert und Ernst Sassen, mit denen ich neben den zahlreichen fachlichen Diskussionen sehr viele menschlich wertvolle Erinnerungen und gemeinsame Abende verbinde, sowie insbesondere Martin Ofner, Frank Höning und Kai Hüner, zu denen ich darüber hinaus in der Dissertationszeit eine sehr enge persönliche Beziehung aufgebaut habe. Anne Cleven, Anke Gericke und Christian Riege danke ich für die schönen Momente jenseits der Lehrstuhl Grenzen am IWI. Besonders hervorheben möchte ich jedoch Tobias Vogel,

der sowohl bei fachlichen als auch bei privaten Problemen immer für mich da war und dessen Unterstützung ich ungemein schätze.

Seitens SAP Research gilt mein besonderer Dank Barbara Flügge. Sie war für mich als Projektleiterin mit ihrer reichhaltigen Berufserfahrung eine grosse Hilfe und insbesondere in der Abstimmung zwischen IWI und SAP Research eine Stütze, auf die ich mich immer verlassen konnte.

Auch die zahlreichen Helfer im Hintergrund möchte ich nicht vergessen. Ich habe bei meiner Arbeit von der hilfreichen Mitarbeit von studentischen Hilfskräften und Diplomanden profitiert. Beispielhaft möchte ich hier Georg Kundert und Andrea Fuchsloch nennen und ihnen für ihr Engagement herzlich danken. Annette Glaus danke ich für das sorgfältige Korrekturlesen, durch das die Dissertation sprachlich abgerundet wurde. Susanne Gmünder und Ernst Ensslin bin ich dafür dankbar, dass sie mir durch ihre ständige Hilfsbereitschaft bei der Lösung organisatorischer und administrativer Fragen den Alltag am IWI sehr erleichtert haben. Ausserdem möchte ich speziellen Dank an Frau Zorka Pajic richten. Ihr täglicher Optimismus und ihr immerwährendes Lächeln tragen viel zu einer angenehmen Atmosphäre am IWI bei und waren für mich stets eine grosse Freude.

Schliesslich möchte ich mich bei meinen langjährigen Freunden aus Dresden bedanken. Mit Simon Theilig, Robin Dittrich, Frieder Jacobi, Roland Möschner, Thomas Hämpel und all den anderen verbinde ich eine einmalige Freundschaft, die trotz der geographischen Distanz während der Dissertationszeit nichts an Herzlichkeit und Intensität verloren hat.

Zusammenfassung

Die Möglichkeit von Unternehmen, flexibel auf veränderte Kunden- und Geschäftsanforderungen reagieren zu können, hängt entscheidend von ihrer Fähigkeit ab, applikations- und geschäftsprozessübergreifend genutzte Stammdaten konsistent und fehlerfrei zur Verfügung zu stellen. Die dezentralen Organisationsstrukturen global agierender Unternehmen führten in der Vergangenheit jedoch dazu, dass Stammdaten in isoliert voneinander entwickelten Applikationen redundant gehalten wurden. Die Beseitigung der Redundanz bedingt die Notwendigkeit, Stammdaten zu integrieren.

Die vorliegende Arbeit entwickelt eine Methode zur Integration von Stammdaten in global agierenden Unternehmen. Durch die Definition eines Vorgehensmodelles mit konkreten Techniken zur Erarbeitung von Ergebnissen, gibt die Methode Datenarchitekten und Stammdatenverantwortlichen eine Handlungsanleitung für die systematische Zusammenführung heterogener Stammdatenbestände. Inhaltliche Schwerpunkte der Methode liegen auf dem Metadatenmanagement, der Integration von Metadaten in Datenmodellen und Architekturmustern zur konsistenten Speicherung und Verteilung von Stammdaten.

Die Arbeit nutzt drei Fallstudien, um Anforderungen an die Methode abzuleiten. Die Anwendung der vorgeschlagenen Methode in zwei Aktionsforschungsprojekten zeigt die Anwendbarkeit in der Unternehmenspraxis und dient der Nutzenbewertung der Methode.

Summary

The possibility of companies to flexibly respond to changing customer and business requirements, essentially depends on their ability to provide consistent and error-free master data across applications and business processes. The decentralized organizational structure of companies acting globally, however, lead to master data being held redundantly in multiple isolated applications. Eliminating this redundancy necessitates the integration of master data.

The thesis therefore develops a method for the company-wide integration of master data. By defining a procedure model with corresponding techniques the method guides data architects and master data managers to systematically integrate heterogeneous master data. Thematically, the main topics of the method comprise metadata management, the integration of metadata in data models and architecture patterns for consistent storage and distribution of master data.

The thesis makes use of three case studies to derive requirements on the method development, while two action research projects show applicability of the proposed method and serve evaluation of its benefit.

Inhaltsübersicht

1 Einführung	1
1.1 Ausgangslage und Handlungsbedarf	1
1.2 Ziele, Adressaten und Nutzen der Arbeit	3
1.3 Forschungsmethodik	4
1.4 Aufbau der Arbeit	7
2 Grundlagen.....	9
2.1 Integration.....	9
2.2 Datenqualitätsmanagement und Stammdatenmanagement	17
2.3 Informationsarchitektur	24
2.4 Zusammenfassung und Beitrag für die Arbeit.....	37
3 Stammdatenintegration in Praxis und Wissenschaft: Bestandsaufnahme und Herausforderungen.....	41
3.1 Integrationsprobleme in der Wirtschaftsinformatik	41
3.2 Standardisierung von Stammdaten	47
3.3 Methodische Ansätze zur Integration von Stammdaten	51
4 Anforderungen aus der Praxis an die Integration von Stammdaten	61
4.1 Auswahl der Fallstudien und Aktionsforschungsprojekte.....	61
4.2 Fallstudien.....	62
4.3 Aktionsforschungsprojekte	91
4.4 Beitrag der Fallstudien und Aktionsforschungsprojekte für die Arbeit	94
5 Entwicklung einer Methode zur Stammdatenintegration.....	97
5.1 Methoden-Engineering	97
5.2 Überblick über Methode und Entwurfsergebnisse	99
5.3 Metamodell	101
5.4 Rollenmodell.....	103
5.5 Methodenfragment 1: Geschäftsobjekttypen in einem BDD beschreiben	107
5.6 Methodenfragment 2: Semantisches Stammdatenmodell ableiten.....	142
5.7 Methodenfragment 3: Stammdatenintegrationsarchitektur gestalten	162
5.8 Dokumentationsmodell.....	187
6 Anwendung und Bewertung der Methode	189
6.1 Anwendung der Methode in Aktionsforschungsprojekten.....	189

6.2	Bewertung der Methode	201
6.3	Aufwand-Nutzen-Betrachtung.....	205
7	Zusammenfassung und Ausblick	213
7.1	Ergebnisse und Erkenntnisse der Arbeit.....	213
7.2	Einschränkungen.....	214
7.3	Weiterer Forschungsbedarf.....	215
Anhang A	Dokumentation zur Forschungsmethodik	221
Anhang A.1	Fallstudieninterviews.....	221
Anhang A.2	Charakterisierung der Aktionsforschungsprojekte	222
Anhang A.3	Workshops und Präsentationen	223
Anhang B	Dokumentation der Methode zur Stammdatenintegration.....	224
Anhang B.1	Definition der Metaentitätstypen der Methode.....	224
Anhang B.2	Funktionendiagramm	229
Anhang B.3	Bezug von Entwurfsergebnissen zu Metaentitätstypen	230
Anhang C	Referenz-Metadatenmodell für ein BDD	231
Anhang C.1	Ableitung der Attribute des Referenz-Metadatenmodells.....	231
Anhang C.2	Klassen und Attribute des Referenz-Metadatenmodells	233
Anhang D	Vorlagen zur Erarbeitung der Entwurfsergebnisse	236
Anhang D.1	Vorlage zur Identifikation von Geschäftsobjekttypen.....	236
Anhang D.2	Vorlage zur Beschreibung von Geschäftsobjekttypen	237
Anhang D.3	Vorlage für einen BDD Change Request	240
Anhang D.4	Vorlage für ein Applikationsverzeichnis.....	241
Anhang D.5	Vorlage für die Erstellung einer CRUD-Matrix	242
Anhang E	Dokumentation zum Aktionsforschungsprojekt Daimler AG	243
Literaturverzeichnis	246

Inhaltsverzeichnis

1 Einführung	1
1.1 Ausgangslage und Handlungsbedarf.....	1
1.2 Ziele, Adressaten und Nutzen der Arbeit.....	3
1.3 Forschungsmethodik.....	4
1.4 Aufbau der Arbeit.....	7
2 Grundlagen	9
2.1 Integration.....	9
2.1.1 Klassifikationsansätze der Integration.....	9
2.1.2 Integrationsebenen in Abhängigkeit vom Integrationsgegenstand.....	11
2.1.3 Daten- und Informationsintegration.....	14
2.2 Datenqualitätsmanagement und Stammdatenmanagement.....	17
2.2.1 Definition und Abgrenzung von Datenarten.....	17
2.2.2 Datenmanagement.....	19
2.2.3 Stammdatenmanagement.....	20
2.2.4 Datenqualitätsmanagement.....	21
2.3 Informationsarchitektur.....	24
2.3.1 Datenarchitektur.....	25
2.3.1.1 Abstraktion der Modellierung.....	25
2.3.1.2 Typen und Darstellungsformen von Datenmodellen.....	28
2.3.1.3 Data Dictionaries.....	31
2.3.1.4 Metadaten.....	32
2.3.2 Applikations- und Integrationsarchitektur.....	34
2.4 Zusammenfassung und Beitrag für die Arbeit.....	37
3 Stammdatenintegration in Praxis und Wissenschaft: Bestandsaufnahme und Herausforderungen	41
3.1 Integrationsprobleme in der Wirtschaftsinformatik.....	41
3.1.1 Herleitung des Integrationsmodells.....	41
3.1.2 Integrationsprobleme in der Praxis.....	44
3.2 Standardisierung von Stammdaten.....	47
3.2.1 Arten von Stammdatenstandards.....	47
3.2.2 Klassifikationsstandards.....	49

3.2.3	Beschreibungsstandards	50
3.3	Methodische Ansätze zur Integration von Stammdaten	51
3.3.1	Auswahl bestehender Ansätze.....	52
3.3.2	Reference and Master Data Management	54
3.3.3	Master Data Management and Customer Data Integration for a Global Enterprise	55
3.3.4	Master Data Management	56
3.3.5	Enterprise Master Data Management.....	57
3.3.6	Unternehmensweite Datenintegration.....	58
3.3.7	Bewertung der Ansätze	59
4	Anforderungen aus der Praxis an die Integration von Stammdaten	61
4.1	Auswahl der Fallstudien und Aktionsforschungsprojekte.....	61
4.2	Fallstudien.....	62
4.2.1	SBB Cargo AG.....	62
4.2.1.1	Unternehmen	62
4.2.1.2	Ausgangslage.....	63
4.2.1.3	Umsetzung.....	65
4.2.1.4	Bisherige Erfahrungen.....	70
4.2.2	Deutsche Telekom AG.....	72
4.2.2.1	Unternehmen	72
4.2.2.2	Ausgangslage.....	73
4.2.2.3	Umsetzung.....	74
4.2.2.4	Bisherige Erfahrungen.....	78
4.2.3	Bosch Rexroth.....	79
4.2.3.1	Unternehmen	79
4.2.3.2	Ausgangslage.....	79
4.2.3.3	Umsetzung.....	81
4.2.3.4	Bisherige Erfahrungen.....	85
4.2.4	Erkenntnisse aus den Fallstudien	87
4.2.4.1	Vorgehensmodell SBB Cargo AG	87
4.2.4.2	Vorgehensmodell Deutsche Telekom AG.....	88
4.2.4.3	Vorgehensmodell Bosch Rexroth.....	90
4.3	Aktionsforschungsprojekte	91
4.3.1	Daimler AG	91
4.3.2	DB Netz AG	92
4.4	Beitrag der Fallstudien und Aktionsforschungsprojekte für die Arbeit	94

5	Entwicklung einer Methode zur Stammdatenintegration	97
5.1	Methoden-Engineering	97
5.2	Überblick über Methode und Entwurfsergebnisse	99
5.3	Metamodell	101
5.4	Rollenmodell	103
5.5	Methodenfragment 1: Geschäftsobjekttypen in einem BDD beschreiben	107
5.5.1	Vorgehensmodell und Aktivitäten	108
5.5.1.1	Gestaltungsbereich definieren	110
5.5.1.2	Stakeholder und Prozesse identifizieren	111
5.5.1.3	Geschäftsprozesse auswerten	112
5.5.1.4	Geschäftsobjekttypen konsolidieren	113
5.5.1.5	Metadatenmodell für BDD spezifizieren	113
5.5.1.6	BDD implementieren	114
5.5.1.7	Geschäftsobjekttypen definieren	115
5.5.1.8	Metadatenmanagementprozesse entwerfen	116
5.5.1.9	Metadatenmanagementprozesse implementieren	117
5.5.2	Referenz-Metadatenmodell für ein BDD	119
5.5.3	Techniken	127
5.5.4	Werkzeugunterstützung	137
5.6	Methodenfragment 2: Semantisches Stammdatenmodell ableiten	142
5.6.1	CCTS als Verfahren zur semantischen Datenmodellierung	143
5.6.2	Vorgehensmodell und Aktivitäten	146
5.6.2.1	Stakeholder und Applikationen identifizieren	148
5.6.2.2	Applikationsanalyse durchführen	149
5.6.2.3	Datenobjektattribute spezifizieren	150
5.6.2.4	Stammdatenlandkarte erstellen	150
5.6.2.5	Daten- und Geschäftsobjekttypen konsolidieren	151
5.6.2.6	BDD aktualisieren	152
5.6.2.7	Semantisches Stammdatenmodell ableiten	152
5.6.3	Techniken	153
5.6.4	Werkzeugunterstützung	159
5.7	Methodenfragment 3: Stammdatenintegrationsarchitektur gestalten	162
5.7.1	Architekturmuster für das Stammdatenmanagement	163
5.7.2	Vorgehensmodell und Aktivitäten	168
5.7.2.1	Stammdatenflüsse und Schnittstellen modellieren	170
5.7.2.2	Priorität der Applikationen festlegen	171
5.7.2.3	Architekturmuster evaluieren	172

5.7.2.4	Stammdatenintegrationsarchitektur definieren.....	173
5.7.2.5	Stammdatenintegrationsarchitektur implementieren.....	174
5.7.2.6	Stammdatenmanagementprozesse spezifizieren	175
5.7.3	Techniken.....	175
5.7.4	Werkzeugunterstützung.....	186
5.8	Dokumentationsmodell.....	187
6	Anwendung und Bewertung der Methode	189
6.1	Anwendung der Methode in Aktionsforschungsprojekten.....	189
6.1.1	Integration von Infrastrukturstammdaten bei der DB Netz AG.....	189
6.1.2	Konsolidierung der Applikationsarchitektur bei der Daimler AG.....	196
6.2	Bewertung der Methode	201
6.2.1	Merkmalsbasierte Evaluation.....	202
6.2.2	Auswertung von Fokusgruppeninterviews.....	203
6.3	Aufwand-Nutzen-Betrachtung.....	205
7	Zusammenfassung und Ausblick	213
7.1	Ergebnisse und Erkenntnisse der Arbeit.....	213
7.2	Einschränkungen.....	214
7.3	Weiterer Forschungsbedarf.....	215
7.3.1	Linked Data	215
7.3.2	Master Data as a Service	217
Anhang A	Dokumentation zur Forschungsmethodik	221
Anhang A.1	Fallstudieninterviews.....	221
Anhang A.2	Charakterisierung der Aktionsforschungsprojekte	222
Anhang A.3	Workshops und Präsentationen	223
Anhang B	Dokumentation der Methode zur Stammdatenintegration.....	224
Anhang B.1	Definition der Metaentitätstypen der Methode.....	224
Anhang B.2	Funktionendiagramm	229
Anhang B.3	Bezug von Entwurfsergebnissen zu Metaentitätstypen.....	230
Anhang C	Referenz-Metadatenmodell für ein BDD	231
Anhang C.1	Ableitung der Attribute des Referenz-Metadatenmodells.....	231
Anhang C.2	Klassen und Attribute des Referenz-Metadatenmodells	233

Anhang D	Vorlagen zur Erarbeitung der Entwurfsergebnisse	236
Anhang D.1	Vorlage zur Identifikation von Geschäftsobjekttypen.....	236
Anhang D.2	Vorlage zur Beschreibung von Geschäftsobjekttypen	237
Anhang D.3	Vorlage für einen BDD Change Request	240
Anhang D.4	Vorlage für ein Applikationsverzeichnis.....	241
Anhang D.5	Vorlage für die Erstellung einer CRUD-Matrix	242
Anhang E	Dokumentation zum Aktionsforschungsprojekt Daimler AG	243
Literaturverzeichnis	246

Abkürzungsverzeichnis

ABIE	Aggregated Business Information Entity
AGLS	Australian Government Locator Service
ALE	Application Link Enabling
ARIS	Architektur integrierter Informationssysteme
ASBIE	Associated Business Information Entity
ASG	Allen Systems Group
B2B	Business-to-Business
BBIE	Basic Business Information Entity
BDT	Business Data Type
BE	Business Engineering
BOM	Business Object Model
CCL	Core Component Library
CCTS	Core Component Technical Specification
CDI	Customer Data Integration
CDQM	Corporate Data Quality Management
CRM	Customer Relationship Management
CRUD	Create, Read, Update, Delete (Datenzugriffsoperationen)
D-U-N-S	Data Universal Numbering System
DAMA	Data Management Association
DB	Deutsche Bahn
DEN	Dictionary Entry Name
DIN	Deutsches Institut für Normung
DQM	Datenqualitätsmanagement
DTAG	Deutsche Telekom AG
DW	Data Warehouse
EAM	Enterprise Architecture Management
EAN	European Article Number

EBIT	earnings before interest and taxes (Gewinn vor Zinsen und Steuern)
ECCMA	Electronic Commerce Code Management Association
eOTD	ECCMA Open Technical Dictionary
EPK	Ereignisgesteuerte Prozesskette
ERD	Entity-Relationship-Diagramm
ERP	Enterprise Resource Planning
ETIM	Elektrotechnisches Informationsmodell
ETL	Extract, Transform, Load
FITS	Finanzielle Transparenz und Steuerung
GDD	Global Data Dictionary
GPC	Global Product Classification
GS1	Global Standards One
GTIN	Global Trade Item Number
IDM	Infrastrukturdatenmanagement
IS	Informationssystem
ISIC	International Standard for Industrial Classification
ISO	Internationale Organisation für Normung
IT	Informationstechnologie
ITAIDE	Information Technology for Administration and Intelligent Design of E-Government
ITIL	IT Infrastructure Library
LL	Living Lab
MBC	Mercedes-Benz Cars
MDM	Master Data Management (Stammdatenmanagement)
ME	Methoden-Engineering
MF	Methodenfragment
MOF	Meta Object Facility
NV	Naamloze Vennootschap (Aktiengesellschaft)
OASIS	Organization for the Advancement of Structured Information Standards
OGC	Office of Government Commerce
OMG	Open Management Group
OWL	Web Ontology Language
PDR	Production Data Replication

PIM	Product Information Management
PLM	Product Lifecycle Management
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
SBB	Schweizerische Bundesbahnen
SID	Shared Information & Data Model
SNIA	Storage Network Industry Association
SOA	Serviceorientierte Architekturen
SQL	Structured Query Language
SRM	Supplier Relationship Management
STEP	Standard for the Exchange of Product Model Data
TDWI	The Data Warehousing Institute
UBL	Universal Business Language
UID	Unique Identifier (Eindeutiger Bezeichner)
UML	Unified Modeling Language
UN	United Nations (Vereinte Nationen)
UN/CEFACT	United Nations Centre for Trade Facilitation and Electronic Business
UNSPSC	Universal Standard Product and Service Specification
UPC	Universal Product Code
W3C	World Wide Web Consortium
XML	Extensible Markup Language

1 Einführung

1.1 Ausgangslage und Handlungsbedarf

In der Vergangenheit entwickelten global agierende Unternehmen aufgrund ihrer dezentralen Aufbauorganisation betriebliche Applikationen häufig isoliert voneinander [Heine 1999, 74]. Die in der Folge historisch gewachsenen Applikationsarchitekturen sind durch eine Vielzahl verteilter Applikationen mit unterschiedlichen Datenbanksystemen und redundanter Datenhaltung charakterisiert. Bedingt durch die Notwendigkeit, flexibel auf veränderte Kunden- und Geschäftsanforderungen reagieren zu können, sehen Unternehmen sich jedoch gezwungen, heterogene Applikationen und Datenbestände zu integrieren [Berson/Dubov 2007, 6]. Die Aufgabe der Integration bestehender Applikationen und ihrer Daten ist indes noch immer sehr aufwendig und verursacht fast die Hälfte der gesamten IT-Kosten [Pezzini 2003, 2; Leser/Naumann 2007, 3].

Voraussetzung für die Möglichkeit, Applikationen bzw. Applikationsfunktionen integrieren zu können, ist die Integration von Daten [Bunjes et al. 2002, 417; Rautenstrauch/Schulze 2003, 221; Jung/Winter 2006, 6]. Vor diesem Hintergrund beschäftigt sich die Forschung bereits seit mehr als 30 Jahren mit dem Problem der Datenintegration. Durch Vereinheitlichung von Datensemantik und durch globale Datenmodelle versuchen vorhandene Forschungsansätze, homogene Sichten über physisch auf heterogene Datenbanken verteilte Daten zu erzeugen und damit strukturelle und semantische Konflikte zu überwinden [Ziegler/Dittrich 2004, 3]. Jedoch bieten Datenbankschemata und Datenmodelle keine ausreichende Möglichkeit, die Semantik von Daten vollständig zu erfassen und sie somit eindeutig zu interpretieren [Sheth/Larson 1990, 187]. Die formalisierte Darstellung von Datensemantik und die semantische Datenintegration bleiben daher bis heute die schwierigsten Herausforderungen im Bereich der Integration [Park/Ram 2004, 596f.; Ziegler/Dittrich 2004, 9].

Stammdaten repräsentieren als spezielle Datenart die Kerngeschäftsobjekte einer Unternehmung [Schemm 2008, 21]. Einheitliche Stammdaten für Aufträge, Artikel, Lieferanten und Kunden bilden die Grundlage für eine leistungsfähige Auftragsabwicklung, weswegen deren Harmonisierung und Integration für den Unternehmenserfolg essenziell ist [Kagermann/Österle 2006, 111; Radcliffe et al. 2006, 2]. Zwar schien die Herausforderung, Stammdaten organisationsweit zu vereinheitlichen, mit dem Aufkommen von unternehmensweiten ERP-Systemen zu Beginn der 90-er Jahre weitgehend überwunden [Pohland 2000, 152]. Jedoch dokumentieren mehrere aktuelle Studien bzw. Veröffentlichungen, dass das Problem noch nicht zufriedenstellend gelöst wurde und Unternehmen die Notwendigkeit erkannt haben, Stammdatenobjekte konsistent zu verwenden und zu speichern (vgl. [Russom 2006, 8f.; Friedman et al.

2008, 9f.; Wise 2008, 10]¹). Befragt nach den genauen Problemen, nennt etwa die Hälfte der Unternehmensvertreter in der Studie von RUSSOM schlechte Entscheidungen aufgrund fehlerhafter Stammdatendefinitionen, fehlende Transparenz über Ursprung und Beziehung der Stammdaten oder die mangelnde Kenntnis über synonyme und homonyme Stammdatenobjekte [Russom 2006, 9].

Nachfolgend werden einige ausgewählte Beispiele für Anforderungen, für deren Erfüllung eine unternehmensweite Stammdatenintegration erfolgskritisch ist, kurz erläutert:

- *Customer Relationship Management (CRM)*. Kundendaten stellen für Unternehmen einen wesentlichen Vermögenswert dar, da sie der Ausgangspunkt für zielgerichtete Kundenansprache, individualisierte Produkt- und Dienstleistungsangebote usw. sind [Leser/Naumann 2007, 9]. Hierfür ist die integrierte und konsistente Datenverfügbarkeit unerlässlich [Nohr et al. 2008, 159]. Aufgrund unterschiedlicher Kundenkontaktkanäle und historisch gewachsener Applikationsarchitekturen (ERP-, CRM-, SRM-Systeme) in einem Unternehmen liegen Kundendaten heutzutage jedoch in der Regel in verschiedenen Applikationen uneinheitlich vor.
- *Fusionen und Akquisitionen*. Ein traditioneller Integrationsbereich ist die Harmonisierung von Prozessen, Applikationen und Daten im Zuge von Unternehmensfusionen [Leser/Naumann 2007, 10] sowie beim Zusammenschluss von Geschäftsbereichen. Die vorher unabhängig voneinander agierenden Unternehmensteile verfügen über abweichende Geschäftsprozesse, eigenentwickelte Applikationen und heterogene Stammdatenbestände, die für einen nahtlosen Übergang zur integrierten Leistungserstellung zu konsolidieren sind (siehe hierzu auch die Fallstudien der Deutschen Telekom und von Bosch Rexroth in Kapitel 4.2).
- *Zwischenbetriebliche Integration*. Ähnliche Anforderungen bestehen im Fall der elektronischen Vernetzung in Ecosystemen und der darauf aufbauenden kooperativen überbetrieblichen Leistungserstellung (vgl. [Stadlbauer 2007; Vogel 2009]), wofür sowohl die Koordination der Geschäftsprozesse als auch die Kommunikation zwischen Informationssystemen notwendig sind. Um „an den System- und Prozessschnittstellen die gleiche Sprache [zu] sprechen“ [Kagermann/Österle 2006, 29], bedarf es einheitlicher Strukturen der ausgetauschten Nachrichten sowie eines gemeinsamen Verständnisses der diesen Nachrichten zugrunde liegenden Geschäftsobjekte.
- *Serviceorientierte Architekturen (SOA)*. Die Standardisierung der Semantik von Datenobjekten als Aufgabe der Datenintegration ist eine Voraussetzung für die Realisierung einer SOA [Herwig 2006, 156; Josuttis 2008, 51ff.]. Sie bildet die

¹ Die dem Artikel von FRIEDMAN ET AL. zugrunde liegende Gartner-Studie befragte 250 IT-Vertreter mit dem Aufgabenbereich Datenintegration, während die Studie von RUSSOM Antworten von 741 Vertretern aus dem Bereich Datenmanagement und Datenqualitätsmanagement umfasst.

Grundlage für ein serviceübergreifend einheitliches Informationsmodell [Papazoglou 2007, 548].

Die vorangegangene Argumentation motiviert den Handlungsbedarf für die Arbeit, die sich ausgehend von allgemeinen Integrationsfragestellungen ausführlich dem Untersuchungsbereich der Stammdaten widmet. Im Verlauf der Arbeit werden folgende Forschungsfragen behandelt:

- Welche allgemeinen Integrationsprobleme gibt es in der Unternehmenspraxis? In welchen Bereichen besteht der grösste Handlungsbedarf bei der Integration?
- Welche Aktivitäten, Techniken und Rollen müssen Bestandteil einer Methode sein, welche die systematische Integration von Stammdaten unterstützt? Mit Hilfe welcher Technologien kann semantische Heterogenität zwischen Stammdaten überwunden werden?

Die bereits erwähnte langjährige Untersuchung des Themas Datenintegration berücksichtigend, baut die Arbeit auf Erkenntnissen der Vergangenheit auf und verknüpft diese mit neueren Ansätzen und Technologien (z. B. neue Modellierungskonzepte, semantische Konzepte) in einem strukturierten, methodischen Ansatz.

1.2 Ziele, Adressaten und Nutzen der Arbeit

Ziel der Arbeit ist es, den Integrationsbedarf in Bezug auf Stammdaten aufzuzeigen und eine Anleitung für deren systematische Integration zu erarbeiten. Zu diesem Zweck entwickelt die Arbeit in einem ersten Schritt ein Modell für Integrationsprobleme, das die noch unzureichend gelösten Aufgaben der Integration identifiziert, kategorisiert und beschreibt und somit der Eingrenzung des Untersuchungsbereiches dient. Basierend auf diesem Integrationsmodell werden offene Fragestellungen für die spezifische Domäne der Stammdatenintegration aufgezeigt. Zur strukturierten Überwindung der identifizierten Stammdatenintegrationsprobleme entwickelt die Arbeit eine Methode gemäss dem Methoden-Engineering. Die Arbeit erfüllt damit die beiden Kernaufgaben der Wirtschaftsinformatik als Wissenschaft – Erklärung und Gestaltung [Heinrich et al. 2007, 18-20]. Im Einzelnen liefert die Arbeit folgende Ergebnisse:

- Einen Überblick über den *Stand der Praxis und Forschung*, der die Grundlagen zur Datenintegration sowie zum Datenmanagement strukturiert zusammenfasst. Der Überblick umfasst zudem ein *Integrationsmodell*, welches aktuelle Integrationsprobleme in der Unternehmenspraxis subsumiert, nach Integrationsobjekten strukturiert und offene Handlungsfelder ableitet.
- Eine *Methode zur Integration von Stammdaten*, welche die systematische Zusammenführung heterogener Stammdatenbestände in einem Unternehmen ermöglicht. Der Schwerpunkt liegt auf der semantischen Integration der Stammdaten.

- Zwei *Aktionsforschungsprojekte* dienen dazu, die Eignung der entwickelten Methode anhand der Anforderungen der Praxis zu bewerten und den geschäftlichen Nutzen integrierter Stammdaten zu belegen.

Gemäss den Prinzipien der Design Science erhebt die Dissertation den Anspruch, einen Erkenntnisgewinn für Vertreter sowohl aus Wissenschaft als auch aus der Praxis zu erzielen, die sich mit dem Themengebiet der Integration und des Stammdatenmanagements beschäftigen. Der Nutzen für die Vertreter der Praxis ergibt sich aus der in dieser Arbeit entwickelten Methode für die vereinfachte Integration von Stammdaten eines Unternehmens. Die Arbeit liefert Stammdaten- und Datenqualitätsverantwortlichen sowie IT-Projektleitern einen systematischen Ansatz, der effektive Techniken und Werkzeuge für die Stammdatenintegration bündelt. Ausgehend von der Identifikation wesentlicher Stammdatenobjekte über deren einheitliche fachliche Definition bis hin zur semantischen Modellierung ermöglicht die Methode eine konsistente Integration von Stammdaten über einzelne Applikationen hinweg. Dies stellt die Grundlage für ein unternehmensweites Datenqualitätsmanagement sowie darin enthaltene Aufgabebereiche (Datenlebenszyklusmanagement, Datenqualitätscontrolling usw.) dar.

Für die wissenschaftliche Diskussion leistet die Arbeit einen Beitrag zur Erfassung und zum Verständnis des Problembereiches Integration im Allgemeinen sowie Stammdatenintegration im Speziellen. Hierzu gab es zwar in der Vergangenheit eine Vielzahl wissenschaftlicher Lösungsvorschläge, diese stifteten in der Praxis jedoch oft nur bedingt den erwarteten Nutzen. Zudem bietet die Dissertation einen strukturierten Überblick über den Stand der Forschung und Praxis zur Stammdatenintegration. Durch die Kombination und Erweiterung bestehender Konzepte sowie deren Übertragung auf den Bereich der Stammdaten wird die wissenschaftliche Wissensbasis um neue Lösungsansätze erweitert. Hierdurch bringt die Arbeit neue Erkenntnisse für das Forschungsgebiet Stammdatenmanagement und Datenqualitätsmanagement.

1.3 Forschungsmethodik

Als Teil der Betriebswirtschaftslehre ist die Wirtschaftsinformatik eine anwendungsorientierte Sozialwissenschaft, die ausgehend von Problemstellungen aus der Praxis das Ziel verfolgt, Modelle und Handlungsempfehlungen zur Gestaltung der betrieblichen Realität zu entwickeln [Ulrich 1984, 178ff.]. Die Qualität der Forschungsergebnisse wird anhand ihrer praktischen Problemlösungskraft sowie der wissenschaftlich fundierten Erarbeitung gemessen.

Im Einklang mit dem obengenannten Ziel der Wirtschaftsinformatik gibt der *Design Science*-Forschungsansatz einen Forschungsrahmen zur Konstruktion anwendbarer und nutzenstiftender Lösungen vor, indem Konstrukte, Modelle und Methoden systematisch entwickelt und evaluiert werden [March/Smith 1995, 256ff.; Wilde/Hess

2007, 281]. Die Lösungen müssen zwei Anforderungen genügen: Sie müssen einerseits einen Beitrag zur Bewältigung der zuvor definierten betriebswirtschaftlichen Problemstellung erbringen, andererseits den wissenschaftlichen Erkenntnisgewinn erhöhen [March/Storey 2008, 726]. Die Arbeit folgt daher den Richtlinien der Design Science als konstruktions- und gestaltungsorientierter Forschungsansatz [Jarvinen 2000, 124f.]. Die konkrete Ausprägung des Design Science-Forschungsansatzes in dieser Arbeit ist in Abbildung 1-1 dargestellt. Die einzelnen Bestandteile werden im Folgenden näher beschrieben.

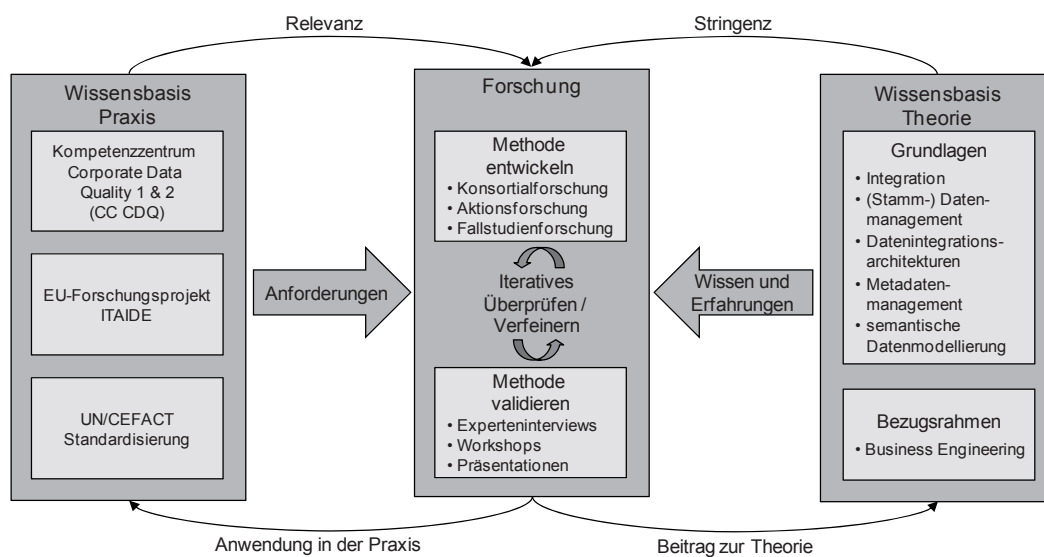


Abbildung 1-1: Ausprägung des Design Science-Forschungsansatzes (in Anlehnung an [Hevner et al. 2004, 80] und [Otto/Österle 2010, 20])

Aus der *Praxis* leitet der Forscher gemäß dem Design Science-Forschungsansatz die Anforderungen für die Forschung ab, wodurch er die praktische Relevanz seiner Ergebnisse gewährleistet. Sie definiert den Untersuchungsbereich der Forschung. Zum Forschungsumfeld der vorliegenden Arbeit zählt als erstes das Forschungsprogramm Business Engineering (BE HSG) des Instituts für Wirtschaftsinformatik der Universität St. Gallen (IWI-HSG). Kernziele des BE HSG sind die Untersuchung von IT-basierten Geschäftslösungen und das Aufzeigen ihrer Potenziale sowie die Entwicklung von Vorgehensweisen zur erfolgreichen Geschäftstransformation [Benz 2001, 6]. Das Dissertationsprojekt baut im Wesentlichen auf Forschungsergebnissen der Kompetenzzentren Corporate Data Quality (2006 bis 2008) sowie Corporate Data Quality 2 (2008 bis heute). In Kompetenzzentren erarbeiten Wissenschaftler gemeinsam mit Fachexperten aus verschiedenen Unternehmen im Rahmen eines Konsortialforschungsprojektes [Back et al. 2007, 94f.; Österle/Otto 2009, 5ff.] Lösungen für Problemstellungen in der Praxis. Dabei wertet der Forscher vorhandenes Praxiswissen der Partnerunternehmen aus und nutzt es für die Konzeption von Lösungsvorschlägen. Die Partnerunternehmen des Kompetenzzentrums definieren im Sinne des Design Science-Forschungsansatzes Anforderungen auf Grundlage ihrer strategischen Ziele und Ge-

schäftstreiber [Österle/Otto 2009, 13f.]. Die Ausrichtung an den Unternehmensanforderungen fördert die Relevanz der Forschungsergebnisse. Darüber hinaus fließen in die Arbeit Erkenntnisse laufender Forschungsaktivitäten aus dem EU-Projekt ITAIDE (Information Technology for Adoption and Intelligent Design for E-Government, 2006-2010, [ITAIDE 2006]) ein, das im 6. Forschungsrahmenprogramm der Europäischen Union verankert ist. Dies betrifft insbesondere die Verwendung der Core Component Technical Specification als Ansatz zur (semantischen) Modellierung und Integration von Daten.

Der Forscher greift zur Erarbeitung seiner Ergebnisse auf etablierte Theorien, Modelle und Methoden (wissenschaftliche Grundlagen) zurück, die gemäss Design Science unter dem Begriff der (*theoretischen*) *Wissensbasis* subsumiert werden [Hevner et al. 2004, 80; Otto/Österle 2010, 20]. Dadurch stellt er die Stringenz des Vorgehens sowie der entwickelten Lösungen sicher und definiert den Bezugsrahmen seiner Arbeit. In der vorliegenden Dissertation unterstützt das *Business Engineering* (BE) als Bezugsrahmen die Entwicklung und Evaluation der Ergebnisse durch einen strukturierten, methoden- und modellbasierten Entwicklungsansatz [Winter 2003, 91 ff.].

Die *Forschung* entwickelt unter Verwendung der zu Verfügung stehenden Wissensbasis Vorschläge zur Gestaltung der betrieblichen Wirklichkeit und zur Lösung der definierten Problemstellungen. Der konstruktionsorientierte Forschungsprozess läuft iterativ mit den beiden Phasen Artefaktentwicklung und Artefaktbewertung (anhand der definierten Anforderungen) ab [Hevner et al. 2004, 79f.], wodurch eine wiederholende Überprüfung und Verbesserung des Artefaktes ermöglicht wird. Zur Erarbeitung der Gestaltungsvorschläge werden in der Wirtschaftsinformatik und speziell im BE qualitativ-empirische Forschungsmethoden eingesetzt (vgl. [Myers 2002]). Für die Entwicklung der Methode zur Stammdatenintegration im Rahmen dieser Arbeit werden drei dieser Methoden genutzt:

- Die *Fallstudienforschung* erlaubt die Untersuchung, Beschreibung und Erklärung komplexer Sachverhalte in ihrem betrieblichen Kontext [Scholz/Tietje 2002, 11f.; Yin 2002, 12ff.]. Zur Erhebung von Fallstudien können verschiedene Techniken genutzt werden; in die Fallstudien der Dissertation flossen Erkenntnisse aus semi-strukturierten, fragebogenbasierten Interviews, Beobachtungen sowie der Inhaltsanalyse von Dokumenten ein. Das Vorgehen beruhte dabei auf der Fallstudienmethodik für Transformationsprojekte des BE, PROMET BECS (vgl. [Senger/Österle 2004]). In dieser Arbeit werden Fallstudien dazu verwendet, um die Anforderungen an die Methode zu identifizieren. Sie sind in Anhang A.1 dokumentiert.
- Die (*kanonische*) *Aktionsforschung* hebt die Grenzen zwischen Forscher und Forschungsgegenstand (der Praxis) auf, indem Wissenschaftler in enger Zusammenarbeit mit Vertretern eines Unternehmens an einer Problemlösung arbeiten [Baskerville/Wood-Harper 1998, 90; Checkland/Holwell 1998, 12]. Dadurch erhält

der Forscher die Möglichkeit, die Auswirkung der Lösung zu beobachten und durch mögliche Anpassungen direkt Einfluss zu nehmen [Susman/Evered 1978, 590; Baskerville/Pries-Heje 1999, 3]. Durch die Anwendung der in dieser Arbeit entwickelten Methode im Rahmen von Praxisprojekten wird diese iterativ bewertet und verbessert [Davison et al. 2004, 72-73].

- Die *Konsortialforschungsmethode* erweitert die Forschungsmethode der Aktionsforschung um ein festgelegtes Konsortium mehrerer Partnerunternehmen, in denen die gemeinsam erarbeiteten Lösungen mehrfach angewendet werden [Otto/Österle 2010, 26f.]. Der Ansatz gliedert sich in Übereinstimmung mit dem Design Science Forschungsprozess [Peppers et al. 2006, 93] in die Phasen Festlegung der Forschungsziele (zu lösende Problemstellungen), Entwicklung neuer Lösungen (Modelle, Methoden, Tools usw.), Anwendung und Bewertung der neuen Lösungen im Unternehmenskontext und Veröffentlichung der Forschungsergebnisse [Österle/Otto 2009, 12ff.]. Sämtliche Schritte werden in enger Zusammenarbeit zwischen Forschern und Vertretern der Partnerunternehmen durchgeführt, wodurch ein intensiver Wissenstransfer zwischen Theorie und Praxis ermöglicht wird [Otto/Österle 2010, 20f.].

Für das in dieser Arbeit definierte Forschungsziel, eine Methode zur Stammdatenintegration zu erarbeiten, wird zusätzlich die Forschungsmethodik des Methoden-Engineering angewendet [Gutzwiller 1994, 11-40], die Kapitel 5.1 genauer beschreibt.

1.4 Aufbau der Arbeit

Die Arbeit gliedert sich in sieben Kapitel. Die abgeschlossene Einleitung in *Kapitel 1* motiviert das Thema und die Zielstellungen, nennt die Adressaten und erläutert den Forschungsansatz der Arbeit.

Kapitel 2 beschreibt die inhaltlichen und methodischen Grundlagen der Arbeit, die anschließend in den Entwurf der eigenen Methode einfließen.

Kapitel 3 legt den derzeitigen Stand zum Thema Stammdatenintegration dar. Das Kapitel diskutiert Treiber und Herausforderungen der Integration im Allgemeinen und fasst die Erkenntnisse in einem Integrationsmodell zusammen. Ausserdem gibt das Kapitel einen Überblick über Standards und methodische Ansätze, die spezifisch die Integration von Stammdaten adressieren.

Kapitel 4 dokumentiert Fallbeispiele, aus denen sich praktische Anforderungen an die zu entwickelnde Methode ableiten lassen. Es beschreibt drei Fallstudien verschiedener Unternehmen, die in der Vergangenheit Projekte zur Integration ihrer Stammdaten durchgeführt haben. Anschliessend führt das Kapitel in zwei Aktionsforschungsprojekte ein, in denen die Methode zur Stammdatenintegration angewendet wurde.

Kapitel 5 ist der Entwicklung der Methode gewidmet. Nach der einführenden Vorstellung des Methoden-Engineering als Bezugsrahmen für die Methodenentwicklung gibt das Kapitel eine Übersicht über Aufbau und Entwurfsergebnisse der Methode. Daraufhin spezifiziert das Kapitel Metamodell und Rollenmodell der Methode sowie die drei Methodenfragmente jeweils mit Vorgehensmodell und Techniken und geht auf Werkzeuge zur Unterstützung der Aktivitäten ein.

Kapitel 6 beschreibt die Anwendung der entwickelten Methoden in den zwei in Kapitel 4 eingeführten Aktionsforschungsprojekten. Das abschliessende *Kapitel 7* fasst die Ergebnisse der Arbeit zusammen, zeigt Grenzen der Untersuchung auf und leitet daraus weitergehenden Forschungsbedarf ab.

Abbildung 1-2 zeigt den Aufbau der Arbeit mit dem Argumentationsfluss zwischen den einzelnen Kapiteln.

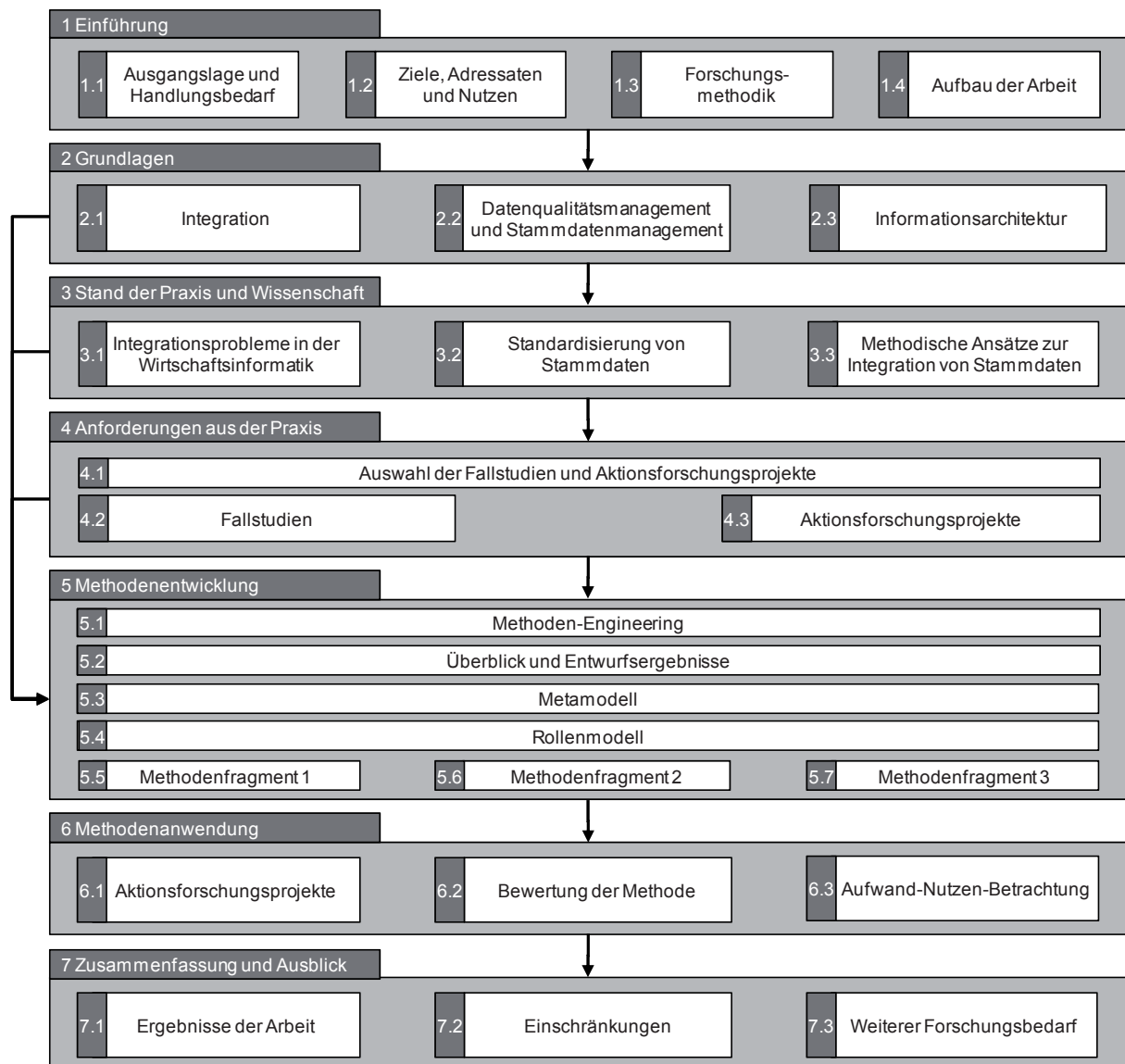


Abbildung 1-2: Aufbau der Arbeit

2 Grundlagen

Ziel dieses Kapitels ist es, die grundlegenden Begriffe und Konzepte dieser Arbeit zu definieren und somit das Thema einzugrenzen. Der Schwerpunkt liegt auf der Einführung der Konzepte; detailliertere Ausführungen werden im Verlauf der Arbeit vorgenommen, wenn die Begriffe im konkreten Zusammenhang verwendet werden. Das erste Teilkapitel behandelt das Thema Integration als allgemeinen Forschungsbereich in der Wirtschaftsinformatik, strukturiert es nach Kategorien und nimmt ausführlicher Bezug auf die Daten- und Informationsintegration. Anschliessend werden im Hinblick auf die Methodenentwicklung in Kapitel 2.2 die Grundlagen des Datenmanagements, des Stammdatenmanagements sowie des Datenqualitätsmanagements eingeführt und in Kapitel 2.3 die Bestandteile einer Informationsarchitektur als wesentliche Gestaltungsbereiche der Stammdatenintegration vorgestellt. Kapitel 2.4 fasst die Bedeutung der vorgestellten Konzepte für die Arbeit abschliessend zusammen.

2.1 Integration

Wie bereits in der Einleitung angedeutet, werden Integrationsfragestellungen in der Wirtschaftsinformatik und bei der Gestaltung von Informationssystemen (IS) auf zahlreichen Ebenen behandelt. Die Klassifikation von Integrationsfragestellungen ermöglicht es, die Komplexität des Themas zu bewältigen und sich auf einen bestimmten Teilbereich zu konzentrieren. Daher werden im Folgenden bestehende Klassifikationsansätze vorgestellt.

2.1.1 Klassifikationsansätze der Integration

Ein in der deutschsprachigen Wirtschaftsinformatik häufig verwendeter Ansatz zur systematischen Strukturierung von Integrationsaufgaben und -problemen sind die von MERTENS vorgeschlagenen „Ausprägungen der Integrierten Informationsverarbeitung“ [Mertens 2007, 2]. Von den vier vorgeschlagenen Dimensionen sind Integrationsrichtung, Integrationsreichweite und Integrationsgegenstand in der Literatur am weitesten verbreitet [Stadlbauer 2007, 29]. Sie werden im Folgenden kurz charakterisiert. Kapitel 2.4 prägt die einzelnen Dimensionen für die vorliegende Arbeit aus.

Integrationsgegenstand

Die Dimension *Integrationsgegenstand* bezieht sich auf das zu integrierende Objekt². Da bezüglich dieser häufig zur Charakterisierung von Integration verwendeten Dimension in der wissenschaftlichen Literatur zahlreiche Differenzierungsansätze existieren, werden diese in Kapitel 2.1.2 ausführlich beschrieben.

² Daher verwenden einige Autoren auch den Begriff des Integrationsobjektes (vgl. [Vogler 2004, 33]).

Integrationsrichtung

Die Dimension *Integrationsrichtung* unterscheidet abhängig davon, ob Objekte zwischen betriebswirtschaftlichen Funktionsbereichen entlang der Wertschöpfungskette oder zwischen verschiedenen Unternehmensebenen (operativ oder dispositiv) integriert werden. Als Bezugssystem wird dabei häufig die von MERTENS vorgeschlagene Pyramide verwendet [Mertens 2007, 6], wobei auch alternative Vorschläge (z. B. für spezielle Integrationsgegenstände) existieren [Jung 2006, 32]. Ihnen gemein ist die Abgrenzung zwischen:

- *horizontaler Integration*, bei der sich die Integrationsgegenstände (Daten, Funktionen, Applikationen) auf derselben Unternehmensebene befinden (beispielsweise bei der Integration funktional ausgerichteter Applikationen zur Unterstützung operativer Geschäftsprozesse), und
- *vertikaler Integration*, bei der sich die Integrationsgegenstände auf verschiedenen Unternehmensebenen, oft jedoch in demselben Funktionsbereich befinden (beispielsweise bei der Verdichtung und Übertragung von Daten aus operativen Applikationen an Planungs- und Kontrollsysteme der Unternehmensleitung).

Integrationsreichweite

Die Dimension *Integrationsreichweite* definiert den organisatorischen Umfang, innerhalb dessen eine Integration erreicht werden soll. In der Literatur zum Thema Integration besteht weitgehende Einigkeit über die drei Ausprägungen *bereichsbezogen* (innerhalb einer Organisationseinheit eines Unternehmens), *innerbetrieblich* (unternehmensweit und bereichs- bzw. prozessübergreifend) und *zwischenbetrieblich* (über Unternehmensgrenzen hinweg) [Jung 2006, 29; Mertens 2007, 7].

Zusätzlich wird in mehreren Quellen nach der *Art der Integration* unterschieden [Rautenstrauch 1993, 23f.; Rosemann 1999, 6; Winter 2009, 8]. Diesbezüglich gibt es grundsätzlich zwei Möglichkeiten. Bei der Integration durch *Verbinden* werden Objekte eines Systems, welche eine logische Beziehung zueinander aufweisen, bislang jedoch nicht oder nur unzureichend verbunden waren, miteinander verknüpft. Die Integration durch *Vereinigen* führt gleichartige oder inhaltlich zusammengehörende Objekte eines Systems in einem Objekt zusammen. WINTER fügt diesen beiden Varianten, die er als Bindung und Vereinigung bezeichnet, die Integrationsarchetypen *Alignment* und *Ableitung* hinzu [Winter 2009, 25-37]. Während beim Alignment zwei unterschiedliche und sich unabhängig voneinander verändernde Artefakte aneinander ausgerichtet werden (z. B. beim Business/IT-Alignment), beschreibt der Archetyp Ableitung die Ableitung eines Artefaktes des Typs A von einem Artefakt des Typs B mit dem Ziel, die Konsistenz zwischen beiden zu wahren (wenn z. B. die Aufbau- aus der Ablauforganisation abgeleitet wird) [Aier et al. 2009, 101f.].

2.1.2 Integrationsebenen in Abhängigkeit vom Integrationsgegenstand

Das Business Engineering diskutiert Herausforderungen der Integration gemäss dem Drei-Ebenen-Modell in Bezug auf Strategie, Prozesse und Systeme (vgl. [Alt et al. 2001]). Abhängig vom Schwerpunkt ihrer Arbeit differenzieren die Autoren die drei Ebenen unterschiedlich aus. Aufbauend auf dem Core-Business-Metamodell [Österle et al. 2007, 193; Höning 2009, 96-116] ist es möglich, die Prozessebene allgemeiner als Organisationsebene zu bezeichnen und zwischen Aufbauorganisation (organisatorische Strukturen) und Ablauforganisation (Geschäftsprozesse) zu unterscheiden. Dadurch ist eine Abgrenzung ablauforientierter Integrationsfragestellungen von strukturell-organisatorischen möglich. SCHWINN teilt die Systemebene in die Applikationsebene und in die Softwarekomponenten- und Datenstrukturebene auf und ordnet diesen unterschiedliche Herausforderungen der Integration zu [Schwinn 2005, 16]. Verschiedene Autoren, welche sich speziell mit Fragestellungen der Integration von Applikationen sowie Prozessen beschäftigen, identifizieren auf dieser Ebene noch detailliertere Teilebenen der Integration. So führt VOGLER eine zusätzliche Ebene der Desktopintegration ein, die als spezielle Form der Prozessintegration das Ziel verfolgt, dem Benutzer alle zur Erfüllung seiner Aufgabe notwendigen Applikationen an seinem Arbeitsplatz integriert zur Verfügung zu stellen [Vogler 2004, 84]. Im Vordergrund steht hierbei die Integration zwischen System- und Prozessebene. Eine noch detailliertere Unterscheidung trifft HEUTSCHI vor dem Hintergrund serviceorientierter Architekturen, indem er Integration auf Ebene der Applikationen, Workflows und Desktopintegration voneinander abgrenzt [Heutschi 2007, 13].

Auf Grundlage verschiedener Anforderungen bei der Integration von IS unterscheiden BUNJES ET AL. zwischen den drei Integrationsgegenständen Daten, Anwendungen und Prozesse (vgl. [Bunjes et al. 2002]). Während beim Integrationsgegenstand Daten die Vereinigung von Schemata und Datenbeständen verschiedener Anwendungen im Mittelpunkt steht, beschäftigt sich die Anwendungsintegration mit der Verknüpfung der Applikationen selbst. Prozessintegration wird als Realisierung von applikationsübergreifenden Geschäftsprozessen definiert, bei der vor allem die Abstimmung und Steuerung der Ablauflogik (Prozesschoreographie) problematisch ist.

MERTENS differenziert zwischen den fünf Ebenen der Daten-, Programm-, Funktions-, Prozess-/Vorgangs- sowie Methodenintegration [Mertens 2007, 1ff.]. Hierbei entsprechen die ersten beiden den bereits im vorherigen Absatz erläuterten Ebenen der Daten- respektive Anwendungsintegration (entsprechend dem Begriffsverständnis eines Programmes als Softwarebaustein bzw. Applikation). Schliesslich beschreibt SCHEER in seinem Konzept zur Architektur integrierter Informationssysteme (ARIS) fünf Sichten (Daten, Funktionen, Leistungen, Organisation und Steuerung), die sich ebenfalls anhand des Integrationsgegenstandes unterscheiden [Scheer 1998, 33-37]. Während die ersten beiden Sichten mit den gleichnamigen Ebenen der bereits vorgestellten Klassi-

fikationsansätze übereinstimmen, deckt die Steuerungssicht Aspekte der Prozessintegration ab. Leistungen bezeichnen die für die Funktionen notwendigen Inputs bzw. die durch die Funktionen transformierten Outputs. Die Organisationssicht umfasst Elemente der Aufbauorganisation.

Abbildung 2-1 fasst die dargelegten Unterscheidungen nach Integrationsgegenständen zusammen und ordnet sie den Ebenen des Business Engineering als Bezugsrahmen dieser Arbeit zu.

BE-Ebenen	Integrationsgegenstände im BE			[Bunjes et al. 2002]	[Mertens 2007]	[Rosemann 1999]	[Scheer 1998]	
	[Vogler 2004]	[Österle et al. 2007]	[Schwinn 2005]					
Strategie	Strategie	Strategie						
Prozesse	Prozesse	Aufbauorganisation	Ablauforganisation	Prozesse	Prozesse	Prozesse	Prozesse	Organisation Steuerung Leistungen
	Desktop			Applikationen	Anwendungen	Funktionen	Funktionen	
Systeme	IT / IS	Informationssysteme		Softwarekomponenten und Daten	Daten	Programme	Daten	Objekte
					Daten	Daten	Daten	

Abbildung 2-1: Ebenen der Integration in Abhängigkeit vom Integrationsgegenstand

Problematisch, da in der Literatur uneinheitlich, ist das Verständnis des Begriffes der Funktion und dementsprechend der Funktionsintegration [Mertens/Holzner 1992, 6]. Der Interpretation einer Funktion als Applikationsfunktion [Schwinn 2005, 19] folgend adressiert die Funktionsintegration Probleme auf Ebene der Applikationslogik [Alonso et al. 2003, 5; Heutschi 2007, 16]. Dagegen betrachten mehrere Autoren Funktionen aus einer betriebswirtschaftlichen bzw. fachlichen Perspektive, der zufolge eine (Geschäfts-)Funktion einer Aufgabe bzw. einer Verarbeitungseinheit entspricht. Entsprechend stehen bei deren (informationstechnologischer) Integration ähnlich wie bei der Prozessintegration fachliche und ablauforientierte Aspekte im Fokus der Betrachtung [Rosemann 1999, 7f.; Mertens 2007, 2].

Die für diese Arbeit gültigen *Integrationsebenen* leiten sich direkt aus der Konsolidierung der in Abbildung 2-1 dargestellten Klassifizierungen nach Integrationsgegenständen ab. Sie bilden die oberste Gliederungsebene des in Kapitel 3.1 erarbeiteten Integrationsmodells. Relativ detaillierte Unterteilungen einzelner Autoren, welche sich in ihren Arbeiten auf spezifische Integrationsfragestellungen einzelner Ebenen konzentrieren (z. B. Desktopintegration), werden auf dieser Stufe bewusst vernachlässigt.

Im Mittelpunkt der *Applikationsintegration* steht die gemeinsame Nutzung und Wiederverwendung vorhandener Funktionalität durch mehrere Applikationen [Ruh et al. 2001, 27ff.; Kaib 2002, 65ff.]. Integrationsmechanismen zur Applikationsintegration

beruhen primär auf dem Konzept entfernter Funktionsaufrufe und werden unter dem Begriff funktionaler Middleware zusammengefasst. Die Funktionsaufrufe können entweder prozedural (wie bei Remote Procedure Calls und Transaction Processing Monitoren), objektorientiert (wie bei Object Brokern und Object Monitoren) oder nachrichtenorientiert (wie bei Message Oriented Middleware) implementiert werden [Alonso et al. 2003, 33-34]. Mehrere Autoren subsumieren unter dem Synonym der System- bzw. Anwendungssystemintegration entsprechend dem mehrschichtigen Aufbau von Applikationen die drei Teilebenen Daten-, Funktions- und Präsentationsintegration [Vogler 2004, 56; Heutschi 2007, 14-18]. In der vorliegenden Arbeit wird vor dem Hintergrund der Fokussierung auf die Integration von Stammdaten *Datenintegration* bewusst separat betrachtet (siehe folgendes Kapitel). Hierunter fallen Aufgaben wie die Zusammenführung von Datenmodellen bzw. Datenbankschemata, die physische Integration von Datenbeständen oder die ablauforientierte Integration von Applikationen durch Datenweitergabe [Kurbel/Rautenstrauch 1996, 169]. Zudem wird auf die Schaffung eines einheitlichen Datenbestandes durch die Datenintegration als Voraussetzung für die nahtlose Verknüpfung von datennutzenden Applikationsfunktionen sowie Aufgaben verwiesen [Schissler et al. 2004, 16; Mertens 2007, 3].

Auf Organisationsebene ist zwischen statischen (Aufbauorganisation) und dynamischen (Ablauforganisation) Integrationsgegenständen zu unterscheiden [Sperling 2007, 86]. Die Abstimmung von Abläufen mit dem Ziel einer durchgängigen, medienbruchfreien Aufgabenbearbeitung wird unter dem Begriff der *Prozessintegration* zusammengefasst. Dies beinhaltet Massnahmen zur applikations- und organisationsübergreifenden Unterstützung und Steuerung einzelner Prozesse sowie zur Zusammenarbeit zwischen Prozessen [Bunjes et al. 2002, 420; Vogler 2004, 54]. Gemäss der „logische[n] Aneinanderreihung von Funktionen in einer Ablaufsteuerung“ [Winter 2009, 8] umfasst die Prozessintegration auch die Integration von (Geschäfts-)Funktionen. Im Gegensatz zur Integration auf Daten- und Applikationsebene steht bei der Integration auf Prozessebene nicht die Erreichbarkeit von Daten und Applikationslogik einer einzelnen Applikation im Vordergrund, sondern die Automatisierung von Geschäftsprozessen unter Einbeziehung aller in einem Informationssystem vorhandenen Applikationen. Diesem Verständnis folgend können Desktop- und Workflowintegration als spezielle Form der Prozessintegration verstanden und „oberhalb“ der Applikationsebene angeordnet werden [Vogler 2004, 54; Heutschi 2007, 13]. Der Schwerpunkt liegt mit der aufgabenadäquaten Bereitstellung notwendiger betrieblicher Applikationen aus Sicht des Mitarbeiters (Desktopintegration) bzw. der automatisierten Ablaufsteuerung (Workflowintegration) darauf, Applikationen und Informationen entsprechend der Bedürfnisse der Anwender in den Prozessen einzubinden.

Die *Strategieintegration* stellt im Vergleich zu den vorherigen Ebenen einen Sonderfall dar, da sie in erster Linie im überbetrieblichen Kontext sowie bei Unternehmensakquisitionen und -fusionen von Bedeutung ist. Hier gilt es, die strategische

Ausrichtung (Geschäftsfelder, Ressourcen, Zielsysteme) der Unternehmen unter Berücksichtigung der Gegebenheiten auf den darunterliegenden Ebenen in Übereinstimmung zu bringen [Sperling 2007, 83-86].

Wie die bisherigen Ausführungen verdeutlichen, können die Integrationsebenen nicht isoliert voneinander betrachtet werden. Einerseits ist die Abgrenzung nicht immer trennscharf möglich, u. a. zwischen Applikations- und Prozessintegration durch Workflows. Andererseits können die Ebenen als aufeinander aufbauend und folglich einander bedingend verstanden werden [Vogler 2004, 78f.]. Beispielsweise bilden integrierte Datenbestände die Voraussetzung für die Integration von Applikationen, die wiederum Grundlage für harmonisierte Geschäftsprozesse ist [Jung/Winter 2006, 6].

2.1.3 Daten- und Informationsintegration

Der Schwerpunkt der Arbeit liegt auf der Integration des Gegenstandes *Daten*, genau genommen auf der Integration von Stammdaten als spezielle Datenart³. Daher betrachtet dieses Kapitel die Datenintegration näher und grenzt sie zur Informationsintegration ab.

Allgemein wird unter *Datenintegration* die Zusammenführung von Daten aus heterogenen Datenquellen und die damit verbundene vereinheitlichte Sicht auf diese Daten verstanden [Cudré-Mauroux 2006, 17]. JUNG verweist in seiner Habilitationsschrift auf die teilweise stark voneinander abweichenden Definitionen des Begriffes Datenintegration in wissenschaftlichen Publikationen [Jung 2006, 36-39]. Diese resultieren aus den verschiedenen Aspekten, welche die jeweiligen Autoren in ihren Arbeiten adressieren (siehe Tabelle 2-1).

Aspekte der Datenintegration	Beschreibung
1. Einheitlicher Zugriff auf verteilte Datenelemente	Gewährleistung eines einheitlichen Zugriffs für den Nutzer (menschlicher Aufgabenträger) auf in verteilten Datenbanken gespeicherte Datenelemente, ohne zwangsläufig bestehende Datenredundanz zu beseitigen
2. Automatische Datenweitergabe zwischen Applikationen	Vereinheitlichung von Datenelementtypen mit dem Ziel, Daten zwischen verschiedenen Applikationen (maschinelle Aufgabenträger) auszutauschen
3. Konsolidierung von (zuvor verteilten) Datenelementen	Zusammenführung von Datenelementen aus unterschiedlichen Quellen in einem zentralisierten, gemeinsamen Datenbestand, auf den die Aufgabenträger zugreifen
4. Standardisierung von Datenelementtypen	(Ausschliesslich) Logische Integration von Daten durch einheitliche Definition von Datenobjektattributen

Tabelle 2-1: Aspekte der Datenintegration (in Anlehnung an [Jung 2006, 36-39])

Die Aspekte der Datenintegration weisen einen deutlichen Bezug zur Art der Integration (siehe Kapitel 2.1.1) auf. Bei der Datenintegration können die Varianten „Integrieren durch Verbinden“ und „Integrieren durch Vereinigen“ primär dadurch

³ Auf die Unterscheidung von Datenarten wird in Kapitel 2.2.1 genauer Bezug genommen.

unterschieden werden, ob die Integration durch einen konsolidierten Datenbestand erreicht wird, in dem redundante Datenelemente eliminiert werden. Im Gegensatz zu dieser Vereinigung verteilter Datenelemente (wie beim dritten Aspekt in Tabelle 2-1) bleiben bei der Verbindung sämtliche Datenelemente (und ihre Typen) erhalten, ohne die Redundanz zwingend zu beseitigen. Bei einer Integration durch Verbindung werden Daten (lediglich) auf logischer Ebene durch Dokumentation der Beziehungen zwischen Datenelementtypen zusammengeführt. In der Literatur besteht in Bezug auf diese beiden Varianten zur Datenintegration weitgehender Konsens, wobei verschiedene Autoren zur Beschreibung der zwei Möglichkeiten unterschiedliche Terminologien verwenden. Eine umfassende Zusammenstellung der verschiedenen Begrifflichkeiten ist bei JUNG zu finden [Jung 2006, 39].

Uneinheitlich ist in wissenschaftlichen Publikationen die Abgrenzung zwischen den Begriffen der *Datenintegration* und *Informationsintegration*. So unterscheiden einige Autoren anhand des Zweckbezuges: Während Definitionen der Informationsintegration häufig explizit auf den Zweckbezug der Integrationsmassnahmen verweisen (vgl. [Dessloch et al. 2003, 7; Jung 2006, 42]), ist dies bei der Datenintegration in der Regel nicht der Fall. Dies deckt sich mit der Unterscheidung zwischen den Begriffen Daten und Informationen. Daten sind zweckneutrale Fakten, die Attribute von Objekten und Prozessen der Realwelt beschreiben und beispielsweise in einer Datenbank gespeichert werden [Spiegler 2000, 14; Price/Shanks 2005, 89]. Die Verarbeitung (z. B. Strukturierung, Interpretation, Transformation) durch den Nutzer oder eine Applikation weist den Daten eine Struktur und eine Semantik (Kontext der Nutzung) zu, wodurch aus ihnen Informationen werden [Davenport/Prusak 1998, 3f.; Tuomi 1999, 105]. Die Struktur und die Semantik können z. B. in einem Datenmodell definiert sein.

Andere Autoren grenzen die beiden Begriffe abhängig davon ab, ob Daten logisch (Informationsintegration) oder physisch (Datenintegration) integriert werden. LESER und NAUMANN betrachten Informationsintegration auf Typ- bzw. Schemaebene (logische Integration von Datenmodellen und -schemata), während Datenintegration die Integration der physischen Daten (Instanzen) bezeichnet [Leser/Naumann 2007, 317]. In der vorliegenden Arbeit werden hierfür die Begriffe *logische Integration* und *physische Integration* bzw. *Instanzintegration* verwendet. In Bezug auf die logische Integration (auch Schemaintegration) identifizieren die Autoren sechs Heterogenitätsklassen [Leser/Naumann 2007, 58ff.]. Heterogenität besteht demnach zwischen zwei oder mehr IS, wenn diese sich in ihren Methoden, Strukturen und Modellen beim Datenzugriff unterscheiden. Die (logischen) Heterogenitätsklassen decken sich grösstenteils mit den von CONRAD identifizierten Integrationskonflikten der Schemaintegration (vgl. [Conrad 1997; Conrad 2002]). Die Heterogenitätsklassen werden wie folgt definiert und voneinander abgegrenzt [Busse 2002, 16f.; Leser/Naumann 2007, 58ff.]:

- *Technische Heterogenität.* Diese Klasse umfasst rein technische Realisierungsprobleme beim Datenzugriff, insbesondere die Anfragesprache, das Austauschformat und das Kommunikationsprotokoll. Technische Heterogenität kann weiter in Zugriffs- und Schnittstellenheterogenität unterteilt werden.
- *Syntaktische Heterogenität.* Unter dem Begriff der syntaktischen Heterogenität werden Abweichungen bei der (technischen) Darstellung von Daten zusammengefasst. Dazu zählen Unterschiede in den Zahlenformaten und Zeichenkodierungen.
- *Datenmodellheterogenität.* Heterogenität der Datenmodelle bezeichnet Unterschiede bezüglich des verwendeten Modellierungsparadigmas, also beispielsweise objektorientierte, relationale und XML-Modelle, und damit verbundene Abweichungen in Anzahl und Semantik der Modellkonstrukte [Batini et al. 1986, 334; Spaccapietra et al. 1992, 90]. Verschiedene Autoren verwenden hierfür allgemein den Begriff der Heterogenitätskonflikte [Spaccapietra et al. 1992, 90; Conrad 2002, 104].
- *Strukturelle Heterogenität.* Mit strukturellen Heterogenitätsproblemen sind Unterschiede im Aufbau von Datenobjekten gemeint, die aus den Freiheitsgraden des Modellierers resultieren [Spaccapietra et al. 1992, 90-94]. Hierunter fällt die Verwendung verschiedener Modellkonstrukte, wie z.B. die Möglichkeit, Attribute (wie z. B. Adresse) in Unterattribute (Strasse, Nummer, Postleitzahl, Stadt) aufzugliedern.
- *Schematische Heterogenität.* Schematische Heterogenitätsprobleme sind ein Spezialfall der strukturellen Heterogenität. Sie resultieren aus der Tatsache, dass gleiche Sachverhalte mit Modellkonstrukten unterschiedlicher Repräsentationsstärke abgebildet werden (z. B. Modellierung als Objekt oder als Attribut) [Parent/Spaccapietra 1998, 172].
- *Semantische Heterogenität.* Die Klasse der semantischen Heterogenität fasst solche Probleme zusammen, die sich aufgrund der Bedeutung der verwendeten Konzepte ergeben. Ursache der Probleme ist ein Konflikt beim Zusammenspiel von Bezeichner (Name) und intendiertem Konzept, wie dies unter anderem bei Synonymen und Homonymen der Fall ist [Spaccapietra et al. 1992, 89f.; March et al. 2000, 330]. Einige Autoren unterscheiden in Bezug auf semantische Heterogenität zwischen extensionalen Konflikten und Beschreibungskonflikten [Spaccapietra et al. 1992, 89f.; Conrad 1997, 46 und 79f.]. Während extensionale Konflikte auf das Problem verweisen, dass zwei Modellklassen semantisch überlappende (aber nicht exakt dieselben) Weltausschnitte abbilden, umfassen Beschreibungskonflikte u. a. Benennungskonflikte (Synonyme, Homonyme), Datentyp- bzw. Wertbereichs- und Skalierungskonflikte (unterschiedliche Masseinheiten) [Conrad 2002, 104].

Auf Instanzebene werden in der Literatur übereinstimmend Datenwertkonflikte genannt [Kim/Seo 1991, 13; Rahm/Do 2000, 5; Leser/Naumann 2007, 320-321]. Diese resultieren aus falschen oder veralteten Daten (d.h. fehlende, nicht korrekte oder widersprüchliche Werte) sowie aus abweichenden Einheiten und unterschiedlicher Genauigkeit der Datenfelder. Grosse Aufmerksamkeit wird dem Problem der Duplikate gewidmet, d. h. der Repräsentation gleicher Realweltobjekte durch mehrere Datensätze (vgl. [Wang/Madnick 1989; Hernandez/Stolfo 1998]).

2.2 Datenqualitätsmanagement und Stammdatenmanagement

2.2.1 Definition und Abgrenzung von Datenarten

Aus einer betriebswirtschaftlichen Sicht können die in einem Unternehmen gehaltenen Daten anhand der Eigenschaften Änderungshäufigkeit, Volumenkonstanz und Grad der existenziellen Unabhängigkeit in vier Kategorien unterteilt werden [Wedekind 2001, 72; Hansen/Neumann 2005, 8f.; Schemm 2008, 21f.].

Stammdaten sind zustandsorientierte Daten, welche – gerade im Vergleich zu Bewegungsdaten – über eine geringe Änderungshäufigkeit verfügen und über den Zeitverlauf ein relativ konstantes Volumen aufweisen [Mertens 2007, 20; Loshin 2008, 8]. Bestehend aus einem eindeutigen Bezeichner sowie weiteren Attributen beschreiben sie die Kerngeschäftsobjekte eines Unternehmens [White et al. 2006, 2; Smith/McKeen 2008, 63]. Sie bilden die Grundlage sowohl für operative Wertschöpfungsprozesse als auch für analytische Entscheidungsprozesse des Unternehmens. Charakteristisch für Stammdaten ist ihre unternehmensweite Nutzung. In der Master Data Management Survey, einer Umfrage unter 163 Unternehmen verschiedener Branchen und unterschiedlicher Grösse, gaben über 50 % der befragten Unternehmensvertreter an, dass ihre Kundenstammdaten von 6 oder mehr Geschäftsbereichen genutzt werden und diese in 96 % der Fälle auf mindestens zwei Datenquellen verteilt sind [Priglinger/Friedrich 2008, 19]. Die unternehmensweite Nutzung durch verschiedene Geschäftsprozesse, Unternehmensfunktionen und Applikationen stellt erhöhte Anforderungen an die aktuelle Pflege und eindeutige Definition der Stammdaten [Loos 1999, 2]. Diese Problematik wird in der vorliegenden Arbeit adressiert. Eine spezielle Art der Stammdaten sind *Referenzdaten*. Sie dienen ausschliesslich der Klassifikation von Stammdaten [Chisholm 2000, 2f.; Schemm 2008, 21]. Die zugrundeliegenden Codelisten werden teilweise extern (z. B. von Standardisierungsorganisationen) verwaltet und liegen daher nicht immer im Verantwortungsbereich des Unternehmens [Chisholm 2000, 225ff.]. Typische Beispiele für Referenzdaten sind Produktklassifikationen (z. B. eClass) sowie Länder- und Währungscode der Internationalen Organisation für Normung (ISO).

Bewegungsdaten sind abwicklungsorientiert und beschreiben betriebswirtschaftliche Vorgänge. Sie verfügen über einen konkreten Zeitbezug und haben daher eine begrenzte Lebensdauer [Loos 1999, 2]. Typische Beispiele für Bewegungsdaten sind Rechnungen oder Lagerbewegungen [Schemm 2008, 21]. Bewegungsdaten beziehen sich auf Stammdaten und sind daher von diesen existenzabhängig, d. h. ein Bewegungsdatum (z. B. ein Vertriebsauftrag) kann ohne Stammdaten (z. B. einen Kunden) nicht existieren [Dreibelbis et al. 2008, 35]. Da durch den betrieblichen Leistungsprozess eines Unternehmens immer neue Bewegungsdaten entstehen, nimmt ihr Volumen im Laufe der Zeit zu. Die Verarbeitung von Bewegungsdaten bewirkt eine Veränderung von *Bestandsdaten* [Hansen/Neumann 2005, 9]. Bestandsdaten, wie z. B. Lagerbestände oder Kontostände, sind zustandsorientiert und bilden die betriebliche Werte- und Mengenstruktur ab.

Änderungsdaten sind wie Bewegungsdaten abwicklungsorientierte Daten, welche ausschliesslich Stammdaten verändern. Hierzu zählen beispielsweise Änderungen von Kundenadresse oder Kontodaten. Auch ihr Volumen wächst im Zeitverlauf, jedoch wesentlich weniger stark als bei Bewegungsdaten.

In Abbildung 2-2 ist zusammenfassend ein Beispiel für die vier beschriebenen Datenarten sowie ihre Beziehungen zueinander dargestellt.

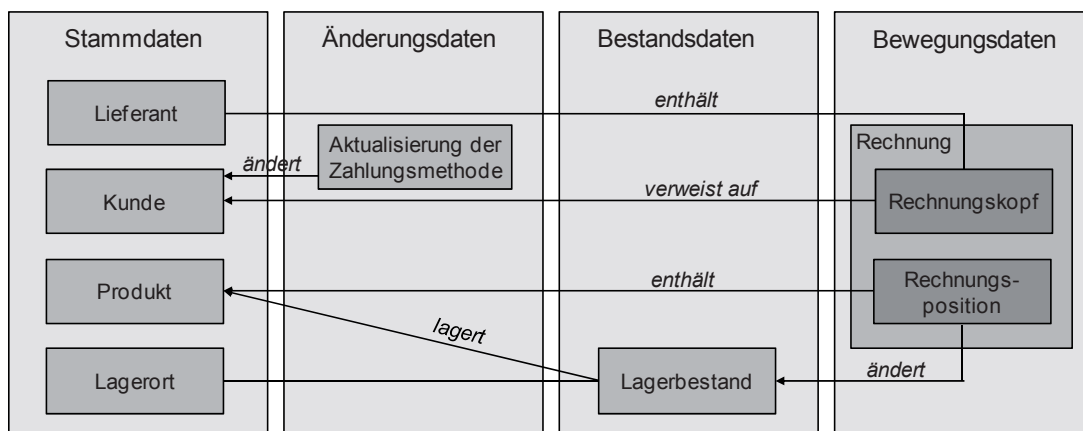


Abbildung 2-2: Beispiel für den Zusammenhang zwischen Datenarten

In der Praxis wird die Klassifikation von Datenobjekten (und ihrer entsprechenden Geschäftsobjekte)⁴ zu den beschriebenen Datenarten uneinheitlich gehandhabt und hängt in erster Linie von der Branche ab. So werden Vertragsdaten im Maschinen- und Anlagenbau als Stammdaten behandelt, während Telekommunikationsunternehmen sie in der Regel als Bewegungsdaten klassifizieren. Die vorliegende Arbeit, in deren Fokus Stammdaten stehen, betrachtet folgende Datenobjekte als Stammdatenobjekte:

- Kunden und Lieferanten sowie zugehörige Verträge und Aufträge,

⁴ Die Unterscheidung zwischen Geschäfts- und Datenobjekten diskutiert Kapitel 2.3.1.1 ausführlich.

- Artikel bzw. Produkte, Produktvarianten, Produkthierarchien, Stücklisten und Materialien,
- Mitarbeiter und Standorte,
- Geschäftseinheiten, Kostenstellen, Konten und Kontenpläne,
- Anlagen.

Durch ihre Eigenschaft, in mehreren Geschäftsprozessen, Organisationseinheiten und Applikationen genutzt zu werden, können die einzelnen Attribute eines Stammdatenobjektes in zwei Gruppen unterteilt werden [Legner/Otto 2007, 562]. Zu den globalen Attributen (oder Grunddaten), die über Abteilungs- und Applikationsgrenzen hinweg Gültigkeit besitzen (sollten), zählen beispielsweise der eindeutige Bezeichner (Unique Identifier, UID) oder der Name des Stammdatenobjektes. Des Weiteren existieren funktions- oder applikationsspezifische Attribute, die nur für bestimmte Teilbereiche wichtig sind. Für das Beispiel eines Materialstammsatzes gehören hierzu unter anderem Preise und Konditionen für den Einkauf oder Losgrößen für die Fertigung (siehe Abbildung 2-3).

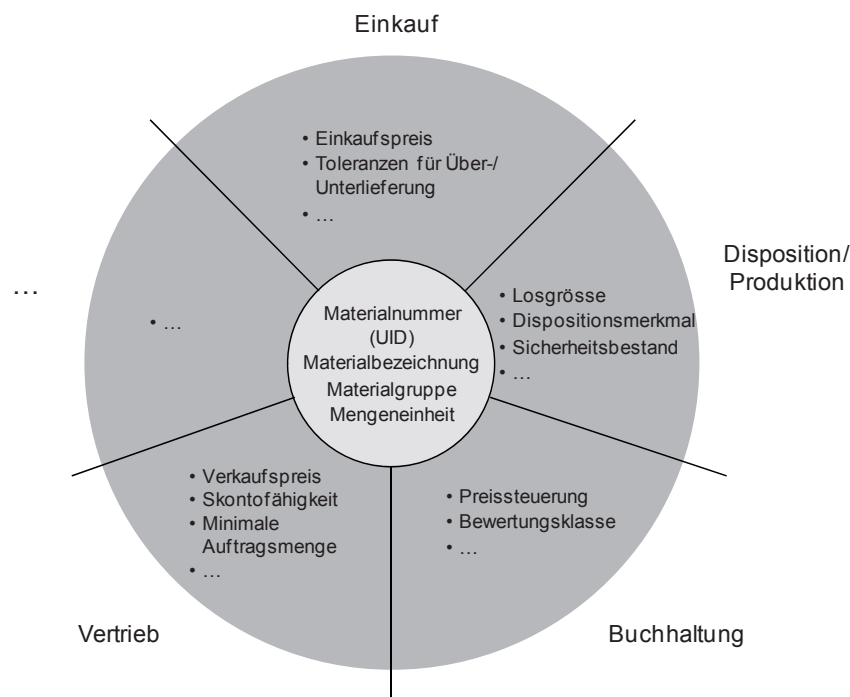


Abbildung 2-3: Beispiel für den Aufbau eines Materialstammsatzes

2.2.2 Datenmanagement

Das *Datenmanagement* hat als Bestandteil des unternehmensweiten Informationsmanagements das Ziel, Daten im Unternehmen optimal zu nutzen [Dippold et al. 2005, 6]. Für das Datenmanagement ist idealerweise eine zentrale Organisationseinheit zu-

ständig. Bei einer dezentralen Datenorganisation, die an betrieblichen Funktionen und deren Fachapplikation ausgerichtet ist, versucht das Datenmanagement schrittweise und systematisch eine unternehmensweite Organisation der Datenressourcen zu etablieren [Ortner 1991, 319]. Dadurch soll das Risiko einer unkontrollierten Redundanz reduziert werden.

Das Datenmanagement umfasst mehrere Aufgabenbereiche zur Erfüllung dieses Ziel: (1) die Definition einer Datenstrategie, (2) die Festlegung der organisatorischen Verantwortung, (3) die Entwicklung eines (unternehmensweiten) Datenmodells bzw. einer Datenarchitektur [Krcmar 2005, 111] sowie (4) die Verwaltung der Daten in Datenbanken [Dippold et al. 2005, 21]. Durch die zunehmende Aufmerksamkeit des Forschungsgebietes in Wissenschaft und Praxis in den vergangenen Jahren wurde der Funktionsumfang des Datenmanagements auf Funktionen wie Data Governance, Metadatenmanagement und Datenqualitätsmanagement (DQM) erweitert [DAMA 2009, 6].

2.2.3 Stammdatenmanagement

Das *Stammdatenmanagement* (engl.: Master Data Management, MDM) prägt die im vorherigen Kapitel genannten Teilaufgaben des Datenmanagements spezifisch für Stammdaten aus: die Beschreibung der Stammdatenstrategie, die Festlegung der organisatorischen Verantwortung, die Entwicklung eines Stammdatenmodells sowie der Entwurf und Pflege der Stammdatenarchitektur [Legner/Otto 2007, 563]. Viele Herausforderungen und Lösungskonzepte des Datenmanagements sind daher speziell auf die Datenart Stammdaten übertragbar, weswegen das Stammdatenmanagement häufig als Teilaufgabe des Datenmanagements gesehen wird [DAMA 2009, 7]. Aus den in Kapitel 2.2.1 beschriebenen spezifischen Eigenschaften von Stammdaten ergeben sich im Gegensatz zu anderen Datenarten jedoch teilweise andere Schwerpunkte.

In der Literatur fällt im Zusammenhang mit dem Stammdatenmanagement häufig der Begriff der „single version of the truth“ [Berson/Dubov 2007, 6]. Der Begriff beschreibt den Bedarf nach einer Erfassung, Integration und Verteilung von vollständigen, konsistenten, fehlerfreien und aktuellen Stammdaten, der in Unternehmen unabhängig von Branchenzugehörigkeit und Grösse besteht [Loshin 2008, 8f.]. Folglich bezeichnet Stammdatenmanagement sämtliche Prozesse, Methoden, Aufgaben und Technologien, welche darauf abzielen, eine verbindliche, nachhaltige und fehlerfreie Stammdatenbasis als „single version of the truth“ zu schaffen und zu pflegen, die applikations- und geschäftsprozessübergreifend genutzt werden kann [Berson/Dubov 2007, 11; Smith/McKeen 2008, 63]. Als Grundlage hierfür sieht das Data Warehousing Institute (TDWI) aus einer fachlichen Sicht das Stammdatenmanagement als die Aufgabe, gemeinsam genutzte Kerngeschäftobjekte eindeutig und einheitlich zu definieren und diese Definitionen über einen längeren Zeitraum über mehrere Applikatio-

nen und Organisationseinheiten hinweg konsistent zu halten [Russom 2006, 4]. Dies erfordert eine gemeinsame Sicht der unterschiedlichen Akteure im Geschäftsprozess auf die Geschäftsobjekte. Aufbauend auf der logischen Integration (auf Typ- bzw. Schemaebene) ist anschliessend die physische Integration der Stammdaten (auf Instanzebene) nötig [Spath et al. 2009, 10f.]. Bei der physischen Integration werden Stammdatensätze aus verschiedenen Applikationen zusammengeführt. Die grösste Schwierigkeit liegt dabei in der Erkennung von Duplikaten, also mehreren Datenobjekten, die jeweils dasselbe Realweltobjekt repräsentieren, und der anschliessenden Datenfusion, um die Duplikate zu beseitigen [Leser/Naumann 2007, 317ff.].⁵ Das Stammdatenmanagement weist somit einen starken Integrationsbezug auf.

Aufgrund der Verwendung von Stammdaten in einer Vielzahl von Applikationen und Funktionen im Unternehmen bestehen die grössten Herausforderungen beim Stammdatenmanagement darin, festzulegen [DAMA 2009, 177f.],

- welche Daten dieselben Entitäten der Realwelt beschreiben,
- welche Applikationen Stammdaten nutzen sowie verändern und welche Applikationen hierfür führend verantwortlich sind bzw. sein sollten,
- welche Personen oder Geschäftseinheiten dafür verantwortlich sind bzw. sein sollten und
- wie eine „single version of the truth“ konsistent an andere Applikationen verteilt werden kann.

Als Teilgebiete des Stammdatenmanagements haben sich in den vergangenen Jahren zudem Ansätze wie Customer Data Integration (CDI) oder Product Information Management (PIM) [Berson/Dubov 2007, 13ff.; Dreibelbis et al. 2008, 13; DAMA 2009, 185] etabliert, die auf spezifische Stammdatenklassen und deren Anforderungen zugeschnittene Lösungskonzepte entwickeln.

2.2.4 Datenqualitätsmanagement

Datenqualitätsmanagement bezeichnet das qualitätsorientierte Management der Unternehmensressource Daten [Weber 2009, 29]. Es umfasst sämtliche Aufgaben und Aktivitäten, die entweder proaktiv oder reaktiv (z. B. Bereinigung von Daten mangelnder Qualität) verfolgt werden, um eine hohe Datenqualität zu gewährleisten [Hüner et al. 2009, 232]. Für das DQM werden in der Literatur von verschiedenen Autoren

⁵ Dies unterscheidet die Stammdatenintegration von der Stammdatenkonsolidierung, bei der verschiedene Datenelemente, die dasselbe Realweltobjekt beschreiben, lediglich konsolidiert werden, ohne die Eindeutigkeit von Stammdatenobjekten zu erzwingen und Redundanz zu beseitigen [Adelman et al. 2005, 42].

zahlreiche Methoden vorgeschlagen⁶. Deren gemeinsamer Kern ist in Analogie zu traditionellen Qualitätsmanagementansätzen ein iteratives Vorgehen, bestehend aus den Phasen Definition, Messung, Analyse und Verbesserung der Datenqualität [Wang 1998, 60ff; English 1999, 70ff.; Eppler 2006, 65ff.]. Dadurch soll eine präventive Verbesserung der Datenqualität als dauerhafte Unterstützungsaufgabe durch Verankerung des Qualitätsprinzips in der Unternehmenskultur erreicht werden [Weber et al. 2009, 590f.]. DQM gibt Richtlinien und Techniken vor, welche die Messung und Verbesserung von Datenqualität eines Unternehmens in definierten Prozessphasen und Entscheidungspunkten ermöglichen [Batini/Scannapieco 2006, 161].

Die Unterscheidung zwischen (Stamm-)Daten- und Datenqualitätsmanagement ist nicht trennscharf möglich: Während DQM in seiner Zielstellung ausdrücklich auf die Qualitätsdimension verweist, ist dies beim Datenmanagement nur implizit der Fall. Letztlich sind Stammdaten jedoch nur verwendbar, wenn sie in einer den Anforderungen der Stammdatennutzer entsprechenden, hohen Qualität verfügbar sind [Legner/Otto 2007, 562]. Daher existiert eine grosse Schnittmenge zwischen den definierten Teilaufgaben beider Forschungsbereiche.

Die Qualität von Daten bzw. ihre Verbesserung wird anhand sogenannter Datenqualitätsdimensionen und entsprechender Kennzahlen operationalisiert, anhand derer die Zielerreichung des DQM gemessen werden kann. Einen umfassenden Überblick über die zahlreichen, in der Literatur vorgeschlagenen Dimensionen geben unter anderem HELFERT [Helfert 2002, 69ff.] sowie WENDE ET AL. [Wende et al. 2009, 54ff.]. Die in Bezug auf Stammdaten am häufigsten genannten Datenqualitätsdimensionen sind in Tabelle 2-2 zusammengefasst [Redman 1996, 254ff.; Scannapieco et al. 2005, 12f.].

Dimension	Definition
Genauigkeit (Accuracy, Correctness)	Grad, in dem Daten korrekt, glaubwürdig und frei von Fehlern sind [Wang/Strong 1996, 31; Olson 2003, 29] bzw. in dem sie den korrekten Zustand eines Realweltobjektes widerspiegeln [Redman 1996, 255; English 1999, 142]
Vollständigkeit (Completeness)	Grad, in dem Daten über die ausreichende Breite, Tiefe und Umfang zur Erfüllung einer Aufgabe verfügen [Wang/Strong 1996, 32] bzw. in dem notwendige Werte oder Attribute eines Datensatzes nicht fehlen [Lee et al. 2006, 56]
Aktualität (Timeliness)	Grad, in welchem das Alter der Daten für deren Nutzung angemessen ist [Redman 1996, 258; Wang/Strong 1996, 32]
Konsistenz (Consistency)	Grad, in dem die Daten mit ihrer Definition übereinstimmen [English 1999, 142]
Eindeutigkeit (Uniqueness)	Grad, in dem es für jedes Realweltobjekt nur eine einzige, eindeutige Repräsentation in den Applikationen gibt [Loshin 2008, 90]
Zugänglichkeit (Accessibility)	Grad, in dem Datennutzer Zugriff auf von ihnen benötigte Daten haben [Wang/Strong 1996, 32; Lee et al. 2006, 59]

Tabelle 2-2: Datenqualitätsdimensionen für Stammdaten

⁶ Einen umfassenden Überblick über Methoden des DQM bieten u. a. WENDE ET AL. [Wende et al. 2009] und BATINI ET AL. [Batini et al. 2009].

Die Herausforderung, eine hohe Datenqualität im Hinblick auf die genannten Dimensionen zu gewährleisten, steigt mit der Anzahl an Unternehmensbereichen, Applikationen und Geschäftsprozessen, die Daten gemeinsam nutzen [Lee et al. 2006, 2f.]. Mit dem Ziel einer höheren Flexibilität gewachsene dezentrale Organisationsstrukturen führen verschärfend dazu, dass Applikationen funktional verteilt sind und über eine eigene, redundante Datenhaltung verfügen. Das Datenqualitätsmanagement muss in solchen Fällen Daten aus heterogenen Datenquellen zusammenführen (z. B. für das konzernweite Berichtswesen) und unterschiedliche Anforderungen an Datenqualität berücksichtigen. OTTO ET AL. adressieren diese Problematik im Rahmen des *Unternehmensweiten Datenqualitätsmanagements* (engl.: Corporate Data Quality Management, CDQM) (vgl. [Otto et al. 2008]) und schlagen hierfür einen Ordnungsrahmen mit sechs Gestaltungsbereichen vor [Otto et al. 2008, 216ff.]: (1) Entwicklung einer CDQM-Strategie (2) Einführung eines Führungssystems (3) Entwurf einer CDQM-Organisation (Data Governance) (4) Definition von CDQM-Prozessen (5) Spezifikation der Informationsarchitektur und (6) Applikationsunterstützung für CDQM. Diese ordnen sie gemäss den Gestaltungsebenen des Business Engineering (Strategie, Organisation, Systeme) zu (siehe Abbildung 2-4).

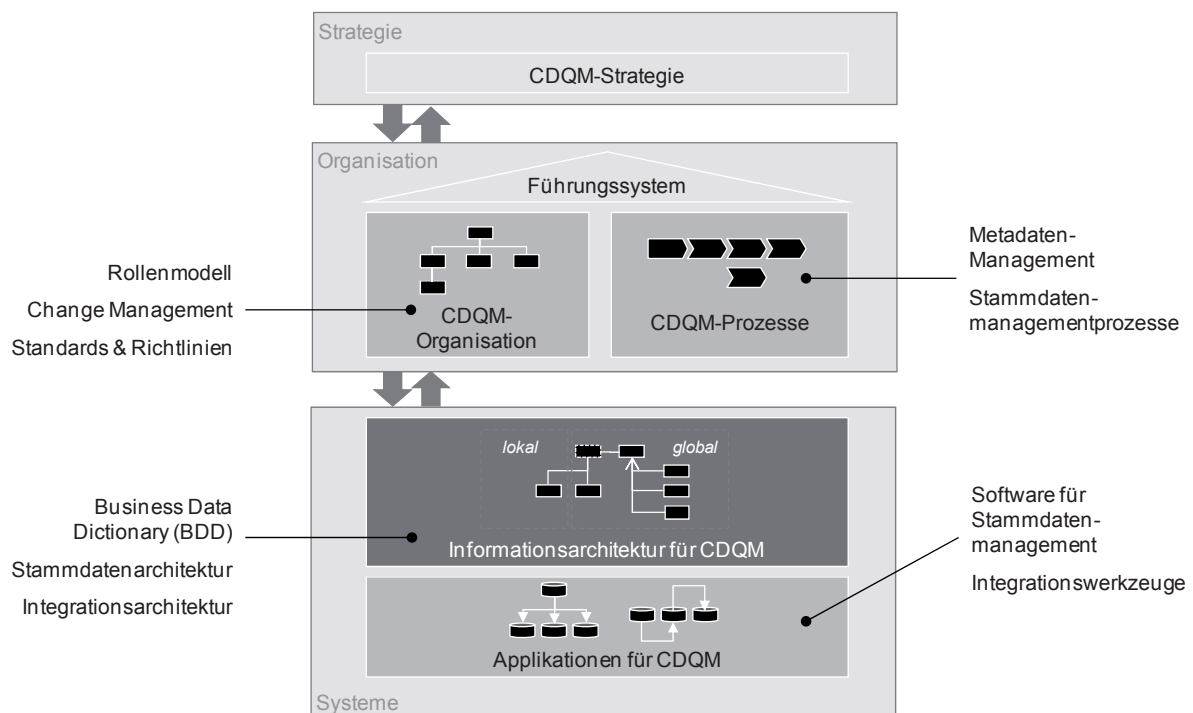


Abbildung 2-4: CDQM-Framework im Kontext des Business Engineering (in Anlehnung an [Otto et al. 2008, 218])

Da der Fokus des CDQM auf unternehmensweit genutzten Daten (sogenannte Corporate Data) liegt und Stammdaten, wie in Kapitel 2.2.1 beschrieben, durch diese Eigenschaft gekennzeichnet sind, dient das Framework als Bezugsrahmen für die vorliegende Arbeit. Die Methode zur Stammdatenintegration kann schwerpunktmässig dem

dunkel eingefärbten Gestaltungsbereich Informationsarchitektur für CDQM zugeordnet werden. Sie muss jedoch auch angrenzende Bereiche wie CDQM-Aufbauorganisation und CDQM-Prozesse sowie die Unterstützung durch Stammdatenmanagementtools berücksichtigen, da das Ziel einer konsistenten Stammdatenbasis nur durch das Zusammenspiel von Architektur, Technologie und Geschäftsprozessen zu erreichen ist [Dreibelbis et al. 2008, 5].

2.3 Informationsarchitektur

In der Literatur zum Thema *Unternehmensarchitektur* besteht eine grosse Heterogenität bezüglich der Anzahl der betrachteten Architekturebenen, der pro Ebene definierten Architektursichten sowie der diesen zugeordneten Gestaltungsobjekte (vgl. [Aier et al. 2008]). Im Sinne des Business Engineering unterscheidet die Mehrheit der Architektur-Frameworks die drei Ebenen Geschäfts-, Prozess- und IS-Architektur [Krcmar 1990, 399f.; Frank 1994, 167ff.; Winter 2003, 93ff.]. Das Gestaltungsobjekt Daten wird dabei vor allem auf Ebene der IS-Architektur behandelt, die wiederum in mehrere Sichten unterteilt wird.

Das Ziel des (Stamm-)Datenmanagements (siehe Kapitel 2.2.2 und 2.2.3), Daten innerhalb einer Organisation gemäss dem Bedarf bereitzustellen, wird in der Informationsarchitektur adressiert [Wagter et al. 2005, 39]. Die *Informationsarchitektur* setzt sich aus einer Menge von Modellen und Definitionen zusammen, die abbilden, welche Daten ein Unternehmen verwendet, welchen Applikationen diese zugeordnet sind und welche Geschäftsfunktionen sie unterstützen [Brancheau et al. 1989, 9; Periasamy 1993, 255f.].

Eine Informationsarchitektur besteht diesen Definitionen folgend aus zwei Bestandteilen [Periasamy/Feeny 1997, 198; Pienimäki 2005, 39]:

- *Datenarchitektur*. Die Datenarchitektur – auch als (Unternehmens-)Datenmodell bezeichnet – beschreibt die Kernentitäten eines Unternehmens sowie deren Beziehungen zueinander auf konzeptioneller Ebene und bildet damit die Basis für ein semantisch einheitliches Verständnis [Periasamy 1993, 257; Krcmar 2005, 112].
- *Applikationsarchitektur*. Die Applikationsarchitektur bezeichnet die Gesamtheit der Applikationen eines Unternehmens, welche Instanzen der im Datenmodell abgebildeten Entitäten erzeugen, verändern oder löschen. Die Applikationen werden in einem Bebauungsplan als graphisches Modell abgebildet, das sowohl den Istzustand als auch eine Sollarchitektur darstellen kann.

Die Informationsarchitektur ist für das Stammdatenmanagement und das damit verbundene Ziel einer konsistenten Stammdatenbasis essenziell. Zum einen adressiert sie die Aufgabe, eine unternehmensweite Datenarchitektur mit logisch standardisierten

Stammdatenobjekten zu definieren [Ryu et al. 2006, 194f.]. Zum anderen berücksichtigt sie im Rahmen der Applikationsarchitektur die stammdatenhaltenden Applikationen. LOSHIN überträgt das Konzept der Informationsarchitektur auf den Bereich des Stammdatenmanagements und spricht von den beiden Komponenten „Master Data Model“ (im Sinne eines konzeptionellen Stammdatenmodells) und „MDM System Architecture“ als Architektur von stammdatenbezogenen Applikationen [Loshin 2008, 45f.]. Charakteristisch für die Informationsarchitektur sowie deren zwei Teilarchitekturen ist die unternehmensweite, applikationsübergreifende Perspektive, die jedoch auf statische Aspekte beschränkt ist.

Da Unternehmen in der Regel über eine auf mehrere Applikationen verteilte Stammdatenhaltung verfügen, besteht für das Stammdatenmanagement darüber hinaus die Notwendigkeit, Daten zwischen stammdatennutzenden Applikationen in konsistenter Art und Weise zu verteilen bzw. auszutauschen [Hsu et al. 1992, 5]. Diese dynamischen Aspekte werden in der *Integrationsarchitektur* abgebildet [Schwinn 2005, 27] und als wichtiger Bestandteil einer Informationsarchitektur für CDQM betrachtet [EFQM 2009].

Abbildung 2-5 zeigt die Komponenten einer Informationsarchitektur, wie sie auch für die weitere Arbeit massgeblich sind.

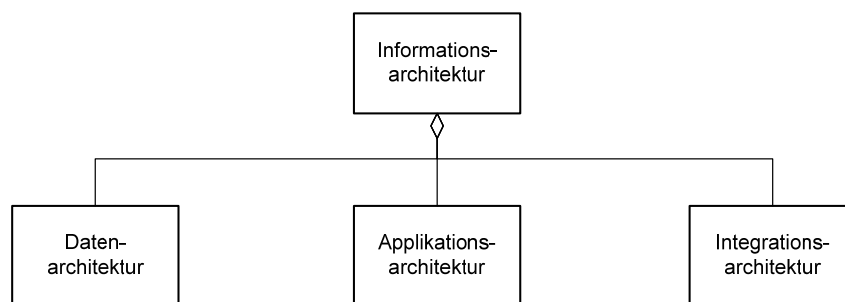


Abbildung 2-5: Bestandteile einer Informationsarchitektur

In den folgenden Teilkapiteln werden die Grundlagen zu den Komponenten Datenarchitektur sowie Applikations- und Integrationsarchitektur detailliert beschrieben.

2.3.1 Datenarchitektur

2.3.1.1 Abstraktion der Modellierung

Das Business Engineering, welches den Bezugsrahmen dieser Arbeit bildet, unterscheidet in seinem Metamodell zwischen Geschäftsobjekten⁷ und Datenelementen

⁷ Aufgrund der synonymen Verwendung der Begriffe Geschäftsobjekt und Informationsobjekt nutzt das Metamodell des Business Engineering den Begriff Geschäfts-(Informations-)objekt [Höning 2009, 47]. Aus Gründen der Vereinfachung sowie der breiteren Akzeptanz im Forschungsumfeld, in dem diese Arbeit entstand, wird im Weiteren der Begriff Geschäftsobjekt verwendet.

[Österle et al. 2007, 193]. *Geschäftsobjekte* (engl.: Business Objects) sind reale oder gedachte Gegenstände der Leistungserstellung, die von Aufgaben in Geschäftsprozessen genutzt und verändert werden [Österle 1995, 87; Höning 2009, 254]. Die Definition legt ein Verständnis als Instanz nahe. Inhalt und Struktur von Geschäftsobjekten werden auf Typebene durch *Geschäftsobjektattribute* und *Geschäftsobjektbeziehungen* eines *Geschäftsobjekttyps* definiert [Vogel 2009, 69]. Geschäftsobjektattribute werden durch fachliche Attributtypen charakterisiert, welche mögliche Wertzuweisungen von Geschäftsobjektattributen einschränken.

Datenelemente repräsentieren Geschäftsobjekte auf Ebene der Informationssysteme [Martin 1977, 50; Brenner 1995, 177]. Sie sind in einer Datenbasis persistent gespeicherte Ausprägungen eines Datenelementtyps [Jung 2006, 17]. Datenelementtypen entsprechen Attributen in relationalen Datenmodellen und werden im Sinne einer einheitlichen Benennung im weiteren Verlauf als *Datenobjektattribute* bezeichnet [Martin 1977, 49f.]. Datenobjektattribute werden durch einen Attributtyp näher spezifiziert. Eine Menge von mehreren Datenelementen bilden zusammen als Ausprägung eines entsprechenden *Datenobjekttyps* ein *Datenobjekt* [Jung 2006, 17]. Der Datenobjekttyp beschreibt ein Datenobjekt z. B. anhand seines Wertebereiches und seiner Klassifizierung [Ferstl/Sinz 2001, 291]. Die Arbeit folgt somit dem Verständnis von Geschäfts- und Datenobjekten als Instanzen. Wird in dieser Arbeit auf die abstrahierten Typen Bezug genommen, so werden explizit die Begriffe Geschäfts- und Datenobjekttyp verwendet.

Die Definitionen zeigen, dass die Begrifflichkeiten anhand zweier Dimensionen unterschieden werden können. Die Dimension der *BE-Ebene* differenziert anhand der beiden Gestaltungsebenen Prozess und System [Österle/Blessing 2003, 81]. Während auf Prozessebene fachlich-konzeptionelle Anforderungen mit dem primären Ziel der Analyse beschrieben werden, bildet die Systemebene die in IT umgesetzten Strukturen und Entitäten ab [Scheer 1998, 38ff.; Sommer 2005, 230]. Geschäftsobjekte sind der ersten Ebene zugeordnet. Sie spielen eine wesentliche Rolle bei der Erfassung der Semantik von Daten und stellen eine Verbindung zwischen Daten und betriebswirtschaftlicher Funktion dar [Scheer 2001, 171f.; Weigand/van den Heuvel 2002, 248 und 250]. Die Differenzierung zwischen Typ und Instanz entspricht dem *Abstraktionsgrad* in der (objektorientierten) Software-Entwicklung [Frank 2003, 16f.]. Die Unterscheidung ist insbesondere in Bezug auf Datenobjekte notwendig, da die Integration von Typen und Instanzen differenziert betrachtet werden muss (vgl. [Batini et al. 1986; Leser/Naumann 2007]). Die definierten Begriffe sind in Tabelle 2-3 anhand der beiden Dimensionen BE-Ebene und Abstraktionsgrad zusammenfassend dargestellt.

		Abstraktionsgrad	
		Typ	Instanz
BE-Ebene	Prozessebene	<ul style="list-style-type: none"> • Geschäftsobjekttyp • Geschäftsobjektattribut 	<ul style="list-style-type: none"> • Geschäftsobjekt • Geschäftsobjektattributwert
	Systemebene	<ul style="list-style-type: none"> • Datenobjekttyp • Datenobjektattribut mit Datenobjektattributtyp 	<ul style="list-style-type: none"> • Datenobjekt • Datenelement

Tabelle 2-3: Unterscheidung zwischen Entitätstypen und Entitätsinstanzen

Ein Beispiel soll den beschriebenen Sachverhalt verdeutlichen:

Bei der DB Netz AG besteht ein unterschiedliches Verständnis über den Geschäftsobjekttyp Gleis abhängig vom Geschäftsprozess, in dem es verwendet wird (z. B. Infrastrukturdatenmanagement, Fahrplan, Anlagenbuchhaltung). Aufgrund der Arbeitsteiligkeit der Organisation haben sich verschiedene fachliche Sichten auf dasselbe Geschäftsobjekt herausgebildet, woraus unterschiedliche, fachspezifische Definitionen resultieren. Das Ergebnis sind Geschäftsobjekttypen, die je nach den in den Prozessen benötigten Geschäftsobjektattributen voneinander abweichen. Beispielsweise benötigt die Abteilung Infrastruktur Geoinformationen, um ein Gleis exakt lokalisieren zu können, während die Anlagenbuchhaltung an der Lebensdauer des Gleises interessiert ist (siehe oberer Teil von Abbildung 2-6).

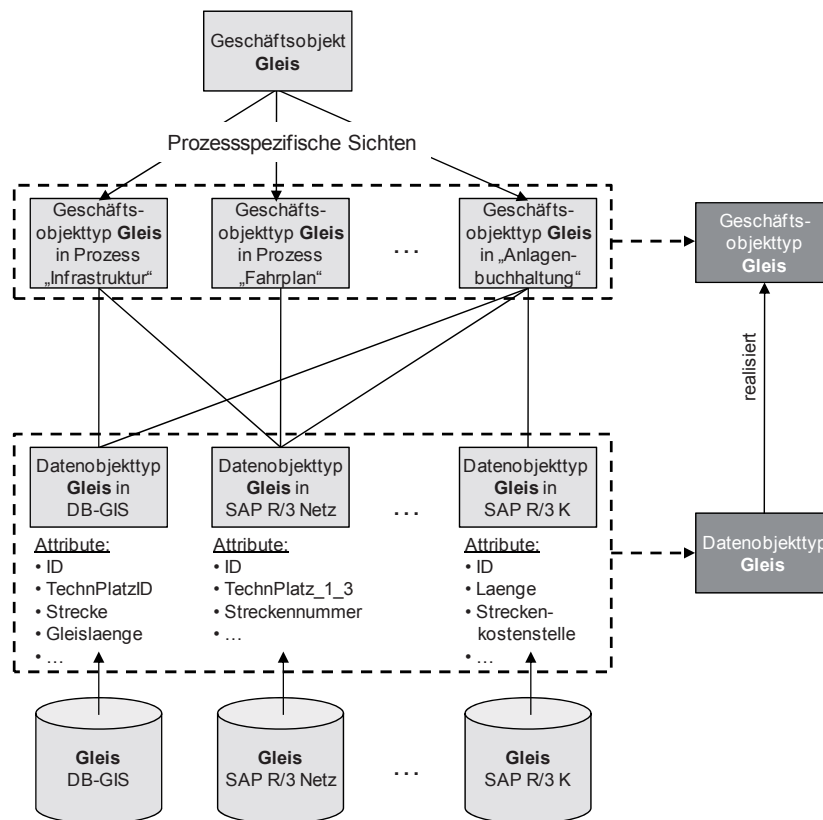


Abbildung 2-6: Beziehung zwischen Geschäftsobjekttypen und Datenobjekttypen

Die Attribute werden in den Datenmodellen der verschiedenen Applikationen (z. B. DB-GIS, SAP R/3 Netz, SAP R/3 K) als Datenobjektattribute mit Typen und Feldlängen spezifiziert, in manchen Fällen werden die Attribute auch durch die Datenmodelle mehrerer Applikationen abgedeckt. Jede Applikation verfügt über ein eigenes Datenmodell, welches applikationsspezifische Datenobjekttypen definiert. In den Datenbanken der Applikationen werden die Datenobjekte als Instanzen der applikationsspezifisch definierten Datenobjekttypen angelegt (siehe unterer Teil von Abbildung 2-6).

2.3.1.2 Typen und Darstellungsformen von Datenmodellen

Datenmodelle dienen im Datenmanagement der ganzheitlichen Darstellung der Datenobjekttypen und ihrer Beziehungen, wie sie sich aus der Datenstrategie und Datenplanung ergeben [Krcmar 2005, 112]. Unter dem Begriff werden häufig implizit *konzeptionelle Datenmodelle* verstanden, die sich von physischen Datenmodellen durch einen höheren Abstraktionsgrad und ihre Implementierungsunabhängigkeit auszeichnen [Simsion 2007, 51].

Ziel eines konzeptionellen Datenmodells (auch semantisches Datenmodell genannt) ist es, durch eine implementierungsunabhängige und semantisch eindeutige Beschreibung von Entitäten aus einer fachlichen Sicht eine gemeinsame Grundlage für die Kommunikation zwischen Vertretern verschiedener Fachbereiche sowie für die systemseitige Implementierung von Datenstrukturen und Datenverteilung zu schaffen [Gemünden/Schmitt 1991, 25; Wand/Weber 2002, 363; Shanks et al. 2003, 85]. Mit der Modellierung geht der Versuch einer Datenstandardisierung einher, die einzelne Datenobjekttypen und Datenobjektattribute über mehrere Applikationen hinweg bereinigt und vereinheitlicht [Ortner 1991, 319]. Zudem bilden Datenmodelle die Basis für eine logische Datenintegration sowie für die Verbesserung von Datenqualität und die Verringerung von Datenredundanz [Schüngel 1995, 1; Shanks/Darke 1999, 20].

In den 80-er und 90-er Jahren wurden mehrmals Versuche unternommen, ein *unternehmensweites Datenmodell* (Enterprise Data Model) zu entwickeln, um funktionale Grenzen in Unternehmen zu überwinden und eine einheitliche Sicht auf die Unternehmensdaten zu erhalten. Diese Versuche scheiterten in der Praxis jedoch aus mehreren Gründen:

- Der Zeitaufwand und die Kosten für die Modellierung sämtlicher Datenobjekte waren zu gross [Schüngel 1995, 161; Moss 2007, 7].
- Die Modelle wurden mit mehreren hundert Datenobjekten und Beziehungen so gross und komplex, dass sie nur schwierig zu pflegen und konsistent zu halten waren [Ortner 2000, 3; Ferstl/Sinz 2001, 144]. Beispielsweise enthält das Datenmodell des SAP R/3-Systems ca. 3.000 Objekte [Scheer/Loos 1996, 8].

- Der Nutzen der Modelle für den Fachbereich blieb marginal, wodurch sie durch diesen kaum Verwendung fanden [Dippold et al. 2005, 80]. Unternehmensweite Datenmodelle wurden primär mit dem Ziel der Anwendungsentwicklung erstellt und blieben in der Regel Aufgabe der IT-Abteilungen, ohne Vertreter des Fachbereiches in die Entwicklung einzubeziehen. Die Verwendung spezieller Datenmodellierungsnotationen führte zudem dazu, dass der Fachbereich die Modelle nicht lesen konnte.

Aufgrund dieser negativen Erfahrungen bei der unternehmensweiten Datenmodellierung fordern neuere Ansätze eine bewusste Beschränkung des Modellierungsumfangs. Hierfür gab es in der Vergangenheit drei Ansätze:

- Der abzubildende Ausschnitt der Realwelt (und somit der Gültigkeitsbereich des Datenmodells) wird eingeschränkt. Dies führt zur Entwicklung sogenannter *Bereichsdatenmodelle*, in dem Datenobjekte eines bestimmten Unternehmensbereiches mit hohem Detaillierungsgrad modelliert sind [Gemünden/Schmitt 1991, 27].
- Anstatt sämtliche Datenanforderungen eines Unternehmens in einem Modell abzubilden, begrenzt man die Anzahl der modellierten Datenobjekte und stellt lediglich die Kernentitäten in einem *Kerndatenmodell* dar [Dippold et al. 2005, 82]. Die damit verbundene Frage ist, welche Entitäten eines Unternehmens als deren „Kern“ definiert werden.
- Der Abstraktionsgrad wird erhöht mit dem Ziel, lediglich eine Begriffsnormierung der wichtigsten Datenklassen in Form von *Unternehmensdatenmodellen* durchzusetzen [Shanks/Darke 1999, 19]. Der Ansatz einer relativ generischen, applikationsübergreifenden Spezifikation von Fachbegriffen (meist in der Sprache des Fachbereiches) wird auch als *kanonisches Datenmodell* bezeichnet [Hoberman 2008, 8]. Da die modellierten Objekte nur relativ grob spezifiziert werden, ist die Möglichkeit begrenzt, diese Modelle zur Applikationsentwicklung zu nutzen [Moss 2007, 7].

Zudem propagieren einzelne Ansätze eine stärker am Fachbereich orientierte Modellierung von Geschäftsobjekten [Fettke 2009, 580], die Eigenschaften von Unternehmensdatenmodellen und Kerndatenmodellen aufweist. Die Ansätze verfolgen primär das Ziel eines einheitlichen Begriffsverständnisses als Grundlage für Datenqualitätsbemühungen. Die resultierenden Modelle werden in der Literatur uneinheitlich als *Fachbegriffsmodell* (vgl. [Kugeler/Rosemann 1998]) oder *High-Level Data Model* (vgl. [Hoberman et al. 2009]) bezeichnet. Sie verwenden häufig vereinfachte Modellierungsnotationen und enthalten Objektdefinitionen bzw. Objektbeschreibungen.

Traditionell werden für die semantische bzw. konzeptionelle Datenmodellierung semi-formale, meist graphische Modellierungssprachen verwendet [Simsion 2007, 12f.], von denen das (erweiterte) *Entity Relationship Diagramm* (ERD) und die Klassendia-

gramme der *Unified Modeling Language* (UML) die am weitesten verbreiteten sind [Fettke 2009, 573]. Diese Modellierungssprachen unterstützen im Gegensatz zu natürlichsprachlichen Beschreibungen (z. B. in Form von Glossaren) durch den geringeren Interpretationsspielraum eine präzisere Definition der Datenobjekte, insbesondere ihres strukturellen Aufbaus (Attribute) sowie ihrer Beziehungen zu anderen Entitäten. Der höhere Grad an Formalisierung in der Darstellung bringt jedoch häufig eine eingeschränkte Verständlichkeit für den Fachbereich mit sich.

Mit dem Aufkommen des Semantic Web wächst in den vergangenen Jahren die Bedeutung von *Ontologien*. Eine Ontologie stellt im allgemeinen Verständnis ein Modell eines Anwendungsbereiches dar, das der Kommunikation zwischen verschiedenen Nutzern dient [Leser/Naumann 2007, 269], indem sämtliche Objekte der Domäne als Konzepte durch Namen und Definitionen spezifiziert werden [Ushold/Gruninger 2004, 59]. Durch eine solche allgemeine Ontologiedefinition können selbst Glossare oder Thesauri als Ontologien bezeichnet werden, weshalb in der Literatur zwischen informellen, primär für den menschlichen Nutzer bestimmten, und formalen Ontologien (für die maschinelle Verarbeitung) unterschieden wird [Gibson et al. 2007, 2]. *Formale Ontologien* beschreiben Objekte eines Anwendungsbereiches unter Nutzung formaler, logikbasierter Modellierungssprachen, die über eine grössere Ausdrucksstärke verfügen. Bedeutende Ontologiesprachen sind das Resource Description Framework Schema (RDFS) [Birkenbihl 2006, 82f.; Hitzler et al. 2008, 66ff.] und die darauf aufbauende Web Ontology Language (OWL) [Birkenbihl 2006, 83ff.; Hitzler et al. 2008, 125ff.], welche beide durch eine W3C-Empfehlung standardisiert sind. Durch die Verwendung logikbasierter Modellierungssprachen mit unterstützenden Inferenzmechanismen wird die Interpretation der Daten aus den datennutzenden Applikationen in die Datenmodelle verlagert, wodurch die Integration der Daten aus heterogenen Quellen vereinfacht werden soll [Mädche et al. 2001, 394; Schemm 2008, 25].

Abbildung 2-7 zeigt verschiedene Darstellungsmöglichkeiten für Datenmodelle in einem Kontinuum unterschiedlicher Grade der Formalisierung und Ausdrucksstärke. Die Auswahl, welche Form zur Darstellung genutzt wird, hängt entscheidend vom Ziel der Modellierung und dem intendierten Nutzer des Datenmodells ab.

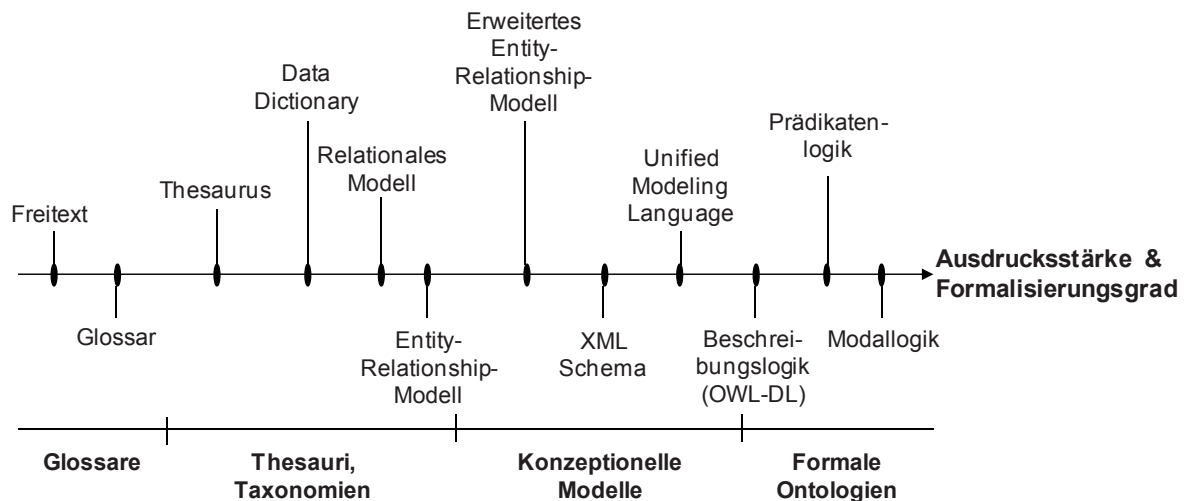


Abbildung 2-7: Möglichkeiten zur Dokumentation von Datenmodellen (in Anlehnung an [Daconta et al. 2003, 157; Uschold/Grüniger 2004, 59])

2.3.1.3 Data Dictionaries

Data Dictionaries (in der Literatur auch als Datenkatalog, Datenlexikon oder Data Repository bezeichnet) haben ihren Ursprung im Bereich der Datenbankmanagementsysteme und Data Warehouses (DWs), wo sie als integriertes Werkzeug Daten über Aufbau und Inhalt von Dateien und Datenbanken – also Metadaten – verwalten und bereitstellen [Ortner 1999, 237]. Durch die Pflege primär technischer Metadaten (siehe folgendes Kapitel) wird das Ziel verfolgt, die Qualität bei der Entwicklung und Pflege von Datenbanksystemen zu erhöhen [Gabriel/Röhrs 1996, 158]. Zu den in einem Data Dictionary gepflegten technischen Metadaten zählen unter anderem Definitionen von Datenbankschema und -tabellen (Speicherstruktur), Konsistenzregeln, Zugriffsrechte, Feldlängen, Nutzungsstatistiken, Fremd- und Schlüsselattribute [Devlin/Murphy 1988, 75; Elmasri/Navathe 2000, 529f.]. Ein Data Dictionary unterstützt als zentrales Verzeichnis aller Inhalte einer Datenbank vor allem Datenbankentwickler und -administratoren (also IT-Anwender) bei ihrer Arbeit [Vossen 1999, 492f.]. DEVLIN und MURPHY erweitern das Konzept des Data Dictionary zu einem sogenannten Business Information Directory, welches technische Informationen aus einem Datenbankkatalog mit fachlichen Beschreibungen der darin enthaltenen Daten und ihrer Anwendungskontexte kombiniert [Devlin/Murphy 1988, 75f.].

Angelehnt an das Konzept der Data Dictionaries haben sich in der Praxis im Rahmen des qualitätsorientierten Stammdatenmanagements in den vergangenen Jahren sogenannte *Business Data Dictionaries* (BDD) als fachliche Metadatenkataloge etabliert [Schemm 2007, 4; Schmidt/Otto 2008, 215ff.]. Diese Repositories fokussieren im Gegensatz zu konventionellen Data Dictionaries auf die Beschreibung von unternehmensweit genutzten Entitäten aus einer fachlichen Perspektive. Entsprechend den BE-Ebenen (siehe Tabelle 2-3) dient ein BDD der Modellierung von Geschäftsobjekten.

Der Schwerpunkt der Modellierung von Geschäftsobjekten liegt auf der unternehmensweiten Vereinheitlichung und einer einvernehmlichen Definition der wichtigsten Entitäten [Ortner 1991, 315]. Dadurch soll eine standardisierte Terminologie im Unternehmen über Unternehmensbereichs- und Applikationsgrenzen hinweg etabliert werden [English 1999, 91f.; Redman 2001, 29]. Die Modellierung dient der Dokumentation der fachlichen Sicht und sollte an der Sprache des Fachbereiches ausgerichtet sein [Fettke 2009, 580]. Hierfür werden fachliche Metadaten (siehe nächstes Kapitel) zu einzelnen Geschäftsobjekttypen im BDD gespeichert, gepflegt und der Nutzergruppe – in erster Linie Datennutzern aus dem Fachbereich – in einer nicht formalisierten Form (Texte, Tabellen) zugänglich gemacht. Auf Basis der im BDD durch Definitionen explizierten Semantik können formalisierte Modelle der Geschäftsobjekte (auf Prozessebene) sowie der Datenobjekte (auf Systemebene) geschäftsgetrieben abgeleitet werden [Loshin 2008, 46; Weber 2009, 157].

Das BDD als Metadaten-Repository erleichtert das Management fachlicher Metadaten, für das bislang nur unzureichende Tool-Unterstützung besteht [Tozer 1999, 123; Shankaranarayanan/Even 2006, 91]. Mit der ISO/IEC 11179 ist ein internationaler Standard für das Management von Metadaten in Metadaten-Repositorys verfügbar, welcher in sechs Teilen Regeln zur Benennung und Definition von Datenelementen sowie ein Metamodell für ein Repository festlegt. Die sechs Teile der ISO/IEC 11179 werden in Tabelle 2-4 kurz charakterisiert.

Teil	Name	Inhalt	Referenz
1	Framework	Überblick über die Standardfamilie mit Kurzbeschreibung der einzelnen Teile sowie Einführung der grundlegenden Konzepte	[ISO/IEC 2004a]
2	Classification	Beschreibung, wie Klassifikationsschemata in einem Metadatenverzeichnis gehandhabt werden	[ISO/IEC 2005a]
3	Registry metamodel and basic attributes	Spezifikation des Metamodells (inklusive Attribute und Beziehungen) für ein Metadatenverzeichnis	[ISO/IEC 2003]
4	Formulation of data definitions	Festlegung von Richtlinien zur Definition von Datenobjekten und Datenelementen	[ISO/IEC 2004b]
5	Naming and identification principles	Definition von Regeln zur standardisierten Benennung von Datenelementen sowie zur eindeutigen Identifikation der Datenelemente im Metadatenverzeichnis	[ISO/IEC 2005b]
6	Registration	Spezifikation des Registrierungsprozesses sowie der notwendigen Rollen	[ISO/IEC 2004c]

Tabelle 2-4: ISO/IEC 11179 und seine Teilstandards

2.3.1.4 Metadaten

Data Dictionarys und konzeptionelle Datenmodelle bilden Eigenschaften von Geschäfts- und Datenobjekten ab: ihre Struktur, ihre Beziehungen zueinander, ihre Attribute und ihre Bedeutung [Maier 1998, 130; Topi/Ramesh 2002, 3]. Sie sind damit eng mit dem Konzept der *Metadaten* verknüpft, mit dem Daten bezeichnet werden, welche die Bedeutung, die Eigenschaften und den Zweck anderer Daten beschreiben [Sie-

gel/Madnick 1991, 134; Burnett et al. 1999, 1212]. MARCO liefert eine umfassende Definition, indem er Metadaten als „all physical data (contained in software and other media) and knowledge (contained in employees and various media) from inside and outside an organization, including information about physical data, technical and business processes, rules and constraints of the data, and structures of the data used by a corporation“ bezeichnet [Marco 2000, 5]. Sie werden genutzt, um Geschäfts- und Datenobjekttypen eines IS standardisiert zu definieren und zu beschreiben. Die konsistente Pflege und Bereitstellung der Metadaten für alle relevanten Nutzer und über den gesamten Lebenszyklus der Entitäten ist Aufgabe des *Metadatenmanagements* [Dippold et al. 2005, 97]. Metadaten finden primär im DW-Umfeld Anwendung, wo sie die Voraussetzung für die korrekte Interpretation und Verarbeitung der Daten und somit für die effiziente Nutzung eines Data Warehouse-Systems (u. a. Suche benötigter Informationen, Unterstützung bei Entscheidungsaufgaben) sind [Auth 2003, 35; Hartmann/Weber 2008, 118].

Das Datenmanagement nutzt Metadaten, um Daten effektiver auffinden, nutzen und pflegen zu können [Burnett et al. 1999, 1212]. Insbesondere verspricht die Verwendung von Metadaten folgende Nutzenpotenziale [Auth 2003, 35ff.; Sen 2004, 151; Berson/Dubov 2007, 128]:

- Unterstützung eines einheitlichen Begriffsverständnisses zwischen mehreren verschiedenen Datennutzergruppen durch konsistente Definitionen der Geschäfts- und Datenobjekttypen und Explikation der Semantik,
- Kontrolle der Datenqualität durch Bereitstellung von Informationen über erlaubte Attributwerte (Datenqualitätsdimensionen Genauigkeit und Vollständigkeit), Datenformate (Datenqualitätsdimension Genauigkeit), die letzte Änderung von Datenobjekten (Datenqualitätsdimension Aktualität) usw. [Marco 2000, 31ff.],
- Entscheidungsunterstützung für den Anwender durch Angabe von Quelle und Qualität der Daten,
- Erleichterung der Datenintegration und -aggregation durch bessere Informationen zu Struktur und Bedeutung der Daten [Thangarathinam et al. 2004, 337].

Metadaten können anhand ihrer Verwendung unterschieden werden [Marco 2000, 49ff.; Lehmann 2001, 87ff.; Inmon et al. 2008, 12ff.]. *Fachliche Metadaten* (auch informative Metadaten [Tozer 1999, 123ff.] genannt) informieren Anwender des Fachbereiches über Bedeutung, Herkunft, Struktur und Qualität der von ihnen verwendeten Daten [Auth 2003, 42; Foshay et al. 2007, 72], wodurch er sie wirksam für seine Arbeit gebrauchen kann. Sie spezifizieren die Semantik, also u. a. den Kontext, in dem Geschäfts- bzw. Datenobjekte verwendet werden [Marco/Michael 2004, 325]. Somit stellen fachliche Metadaten verwendungsorientiertes, betriebswirtschaftliches Wissen über Geschäfts- bzw. Datenobjekte dar. Durch Kontextinformationen kann die Fehler-

anfälligkeit in IS verringert und die Effizienz von IS verbessert werden [Sciore et al. 1994, 255f.]. Zu achten ist hierbei auf eine dem Endnutzer verständliche sprachliche Repräsentation der Metadaten [Lehmann 2001, 88]. Da die vorliegende Dissertation Stammdatenintegration primär aus einer fachlichen Perspektive untersucht, sind für die weitere Arbeit vor allem fachliche bzw. informative Metadaten von Bedeutung. *Technische Metadaten* enthalten Beschreibungen von Architektur- und Applikationskomponenten (z. B. Datenbanken) sowie den technischen Zugriff (z. B. Protokolle), die es Nutzern ermöglichen, Applikationen zu entwickeln und zu betreiben [Busse 2002, 46]. Tabelle 2-5 fasst Beispiele für fachliche und technische Metadaten zusammen.

Fachliche Metadaten	Technische Metadaten
<ul style="list-style-type: none"> • Anwendungsbereich • Dateninhalte, Geschäftsdefinitionen • Beziehungen • Werte und Domänen • Attributnamen • Synonyme / Homonyme • Geschäftsregeln • Datenverantwortung • Dateneigentümerschaft (Data Ownership) • Datenqualitätseigenschaften 	<ul style="list-style-type: none"> • Name der Datenbanktabellen • Technische Datenstruktur • Feldname, Feldeigenschaften • Fremdschlüsselbeziehungen • Zugriffsrechte • Zugriffsprotokolle • Versionsnummer

Tabelle 2-5: Beispiele für fachliche und technische Metadaten

Für die vorliegende Arbeit sind Metadaten vor allem für die einheitliche Identifikation und Definition von Stammdatenobjekten sowie für die Erhöhung der Stammdatenqualität in Bezug auf die in Tabelle 2-2 genannten Dimensionen von Bedeutung (vgl. [Shankaranarayanan et al. 2003], [Ryu et al. 2006]). Durch die Umsetzung eines zentralen Metadatenmanagements erreichen Unternehmen einen höheren Reifegrad des qualitätsorientierten Stammdatenmanagements, da es ihnen mit Hilfe zentral gepflegter Metadaten möglich ist, Stammdatenobjekte und deren Attribute unternehmensweit einheitlich zu spezifizieren [Ryu et al. 2006, 194f.; Loshin 2008, 56ff.].

2.3.2 Applikations- und Integrationsarchitektur

Eine *Applikationsarchitektur* beschreibt die informationstechnischen Komponenten eines Unternehmens zur Unterstützung der Geschäftsprozesse aus einer logischen, funktionalen Sicht [Puschmann 2003, 32; Winter 2004, 317]. Sie bildet Applikationen wie z. B. Datenbanksysteme ab, welche Funktionalität und Geschäftsdaten für betriebswirtschaftliche Aufgaben zur Verfügung stellen [Vogler 2004, 41; Heutschi 2007]. Die Abbildung der Applikationsarchitektur wird im Business Engineering in der Regel ohne Darstellung der Integrationsbeziehungen zwischen den Applikationen als einfacher Bebauungsplan vorgenommen [Pohland 2000, 103f.; Huber et al. 2002, 178-182; Pohland/Fleisch 2002, 384-386]. Der Bebauungsplan ordnet den Unternehmensbereichen oder Geschäftsprozessen Applikationsfunktionalität zu.

Aus Sicht des Stammdatenmanagements können stammdatenführende und stammdatennutzende Applikationen unterschieden werden, die beide bei der Gestaltung der Applikationsarchitektur von Bedeutung sind. Die Klasse der stammdatenführenden Applikationen kann sowohl dedizierte Stammdatenserver als auch betriebswirtschaftliche Applikationen wie beispielsweise ERP-, CRM- oder PLM-Systeme umfassen. Die Applikationsarchitektur muss die konsistente informationstechnische Umsetzung der durch die Datenarchitektur konzeptionell vereinheitlichten Stammdatenobjekte gewährleisten [Loshin 2008, 44]. Dies umfasst die Aufgaben, neue stammdatenführende Applikationen entsprechend den Vorgaben durch die Datenarchitektur zu entwickeln sowie die definierten Standards auf bestehende Stammdatenapplikationen mit minimaler Betriebsstörung zu übertragen.

Uneinigkeit besteht in der Literatur darüber, inwiefern die Applikationsarchitektur auch dynamische Aspekte, wie die Datenverteilung und den Datenaustausch, abbildet. So beziehen mehrere Autoren die Kommunikationsbeziehungen zwischen Applikationen sowie zwischen Applikationen und Datenhaltungssystemen in ihre Definition der Applikationsarchitektur ein [Periasamy/Feeny 1997, 198; Jung 2006, 58]. In dieser Arbeit wird – dem Ansatz neuerer Arbeiten im Umfeld des Business Engineering sowie des Stammdatenmanagements folgend (siehe u. a. [Alt et al. 2004, 46; Heutschi 2007, 12; Winter/Fischer 2007, 2; Schemm 2008, 26]) – eine Unterscheidung zwischen Applikationsarchitektur, welche die statische, funktionale Sicht auf die Applikationen eines Unternehmens abbildet, und Integrationsarchitektur vorgenommen (siehe Abbildung 2-8).

Gegenstand der *Integrationsarchitektur* (auf Systemebene) ist die Kopplung verteilter Funktionen und Daten über Applikationen hinweg mit dem Ziel einer transparenten Kommunikation zwischen Applikationen durch standardisierte Schnittstellen und Protokolle [Schwinn 2005, 27; Heutschi 2007, 12]. Die Kommunikation zwischen Applikationen wird durch *Datenflüsse* dargestellt, die ebenfalls Bestandteil der Integrationsarchitektur sind [Winter/Fischer 2007, 2]⁸. Als Bestandteil der Integrationsarchitektur steuert die *Stammdatenlogistik* die notwendigen Stammdatenflüsse zwischen unternehmensinternen Applikationen [Krcmar 1992, 72-82; Heine 1999, 86f.; Voß/Gutenschwager 2001, 307]. Ziel der Stammdatenlogistik ist es, Stammdaten aus funktionsorientierten Applikationen so zusammenzuführen, dass allen Applikationen die für ihre Aufgabenerfüllung notwendigen Stammdaten in einer konsistenten Form zur Verfügung stehen [Heine 1999, 87]. Sie leistet damit einen Beitrag für die Integration stammdatenführender Applikationen. Primäre Aufgabe der Stammdatenlogistik ist die Strukturierung und Steuerung der Stammdatenverteilung auf Grundlage der

⁸ Im Gegensatz zum Datenfluss, der den Datenaustausch über einen Kommunikationskanal zwischen Anwendungssystemen darstellt, beschreibt der Informationsfluss die Kommunikation zwischen personellen Aufgabenträgern [Heine 1999, 86]. Diese werden in dieser Arbeit nicht näher betrachtet.

Stammdatenhaltung bzw. -speicherung, womit schwerpunktmässig folgende Bereiche tangiert werden [Heine 1999, 88; Voß/Gutenschwager 2001, 309ff.]:

- Stammdatenbereitstellung (Distribution, Allokation, Übertragungskanäle, Standards),
- Stammdatenhaltungskonzepte (anforderungsgerechte Applikationsarchitektur),
- Schnittstellenmanagement,
- Unterstützung mit informationstechnischen Werkzeugen.

Die Aufzählung verdeutlicht, dass die Stammdatenlogistik zur Erfüllung ihrer Aufgabe Artefakte der Applikationsarchitektur (Stammdatenhaltung, Schnittstellen) berücksichtigen muss und daher nicht trennscharf von dieser abgegrenzt werden kann. Beispielhafte Abbildungen der Stammdatenlogistik finden sich unter anderem bei SCHEMM [Schemm 2008, 160, 163, 179].

Ein eng mit der Stammdatenlogistik verknüpftes Konzept ist der *Stammdatenlebenszyklus*, der die prozessorientierte Sicht über sämtliche stammdatennutzenden Applikationen darstellt [Loshin 2008, 46]. Hierfür sind Modelle notwendig, die beschreiben, in welchen Applikationen innerhalb welcher Aufgaben bzw. Prozesse Stammdaten generiert, verarbeitet oder benötigt werden. Diese Modelle schaffen Transparenz über das im Regelfall dezentrale IS mit der Vielzahl verteilter, autonomer Applikationen, welche Stammdaten nutzen, und sind Voraussetzung für die Sicherstellung nachhaltiger Stammdatenqualität [Wang et al. 2000, 10]. Die Modellierung erfolgt unter Zuhilfenahme einer sogenannten CRUD-Matrix⁹, die für jeden Datenobjekttyp zeigt,

- in welcher Applikation ein Stammdatenobjekt ursprünglich entsteht (Quellsystem),
- in welchen Applikation(en) ein Stammdatenobjekt gespeichert und gepflegt wird,
- die Art und Weise, in der Applikationen ein Stammdatenobjekt nutzen.

In der Literatur werden zur Umsetzung des Stammdatenmanagements in einer verteilten Applikationsarchitektur verschiedene Ansätze diskutiert (vgl. [Loser et al. 2004], [Schemm 2008], [Dreibelbis et al. 2008]). Die Ansätze sind abhängig von verschiedenen Eigenschaften der Applikationsarchitektur (Stammdatenhaltung, Applikationstopologie) sowie der Integrationsarchitektur (Stammdatenaustausch bzw. Stammdatenverteilung) ausgeprägt. Ziel der Ansätze ist es, die Probleme der Datenverteilung (Stammdatenbasen sind auf verschiedene technische Plattformen verteilt) sowie der unkontrollierten Datenredundanz (widersprüchliche Datenelemente zu demselben Geschäftsobjekt liegen vor) als Folge der Verteilung zu reduzieren.

⁹ Das Akronym CRUD ergibt sich aus der Kombination der vier grundlegenden Operationen auf Daten: „Create“ (Erzeugen bzw. Anlegen), „Read“ (Lesen), „Update“ (Verändern) und „Delete“ (Löschen).

2.4 Zusammenfassung und Beitrag für die Arbeit

Tabelle 2-6 ordnet die vorliegende Arbeit und die darin zu entwickelnde Methode anhand der in Kapitel 2.1.1 definierten Integrationsdimensionen ein. Die für die Arbeit zutreffenden Ausprägungen sind grau hinterlegt, wobei bei Dimensionen mit mehreren zutreffenden Ausprägungen die primäre fett hervorgehoben ist.

Integrationsdimension	Ausprägungen				
Integrationsgegenstand	Daten	Applikationen	Ablauforganisation	Aufbauorganisation	Strategie
Integrationsrichtung	Vertikal			Horizontal	
Integrationsreichweite	Bereichsbezogen	Innerbetrieblich		Zwischenbetrieblich	
Art der Integration	Verbinden			Vereinigen	

Tabelle 2-6: Einordnung der Arbeit anhand der Integrationsdimensionen

Im Mittelpunkt der Arbeit steht die Integration von Daten bzw. Stammdaten. Die Arbeit betont die Notwendigkeit, Stammdatenintegration ausgehend von den Geschäftsobjekten in den Prozessen und somit aus einer fachlichen Perspektive zu betrachten, weshalb sich die Methode nicht ausschliesslich auf den Gegenstand Daten beschränkt. Die Stammdaten müssen nicht zwangsläufig durch Vereinigung, also durch physische Zusammenführung der Datenelemente in einem redundanzfreien, zentralisierten Datenbestand integriert werden. Vielmehr steht die logische Integration auf Geschäfts- und Datenobjekttypenebene im Vordergrund, die Voraussetzung für eine anschliessende physische Integration auf Instanzebene ist. Bezüglich der Integrationsrichtung liegt der Schwerpunkt der Arbeit auf der Unterstützung operativer Wertschöpfungsprozesse durch die gemeinsame Nutzung konsistenter Stammdaten in den verschiedenen Funktionsbereichen des Unternehmens (horizontal). Ansonsten werden die beiden Anwendungsfälle analytische und operative Stammdatenintegration entsprechend den Integrationsrichtungen vertikal und horizontal in der Arbeit unterschieden. Die vorgeschlagene Methode soll die unternehmensweite Stammdatenintegration über Unternehmensbereichsgrenzen hinweg gewährleisten.

Die vorliegende Arbeit ordnet sich an der Schnittstelle zwischen Datenqualitäts- und Stammdatenmanagement ein. Das auf dem Business Engineering aufbauende Rahmenwerk für CDQM, dessen Ziel die Entwicklung von Methoden, Konzepten und Lösungen für das qualitätsorientierte Management unternehmensweit genutzter Daten ist, bildet den Bezugsrahmen der Arbeit. Die Arbeit nutzt diesen Bezugsrahmen, konzentriert sich aber auf die IS- und teilweise die Prozessebene. Die Methode zur Stammdatenintegration kann schwerpunktmässig dem Gestaltungsbereich „Informationsarchitektur für CDQM“ zugeordnet werden (siehe Abbildung 2-4). Weitere Gestaltungsbereiche eines qualitätsorientierten Stammdatenmanagements, wie die Data Governance, Datenmanagementprozesse oder Führungssysteme, werden in dieser Arbeit nicht detailliert behandelt. Sie finden punktuell dann Berücksichtigung, wenn sie für die Stammdatenintegration einen signifikanten Beitrag leisten (beispielsweise die

Rollen des Data-Governance-Referenzmodells bei der Definition des Rollenmodells der Methode oder das Metadatenmanagement bei der Erstellung eines BDD). Aus dem Datenmanagement bieten bestehende Lösungen in den Bereichen Datenmodellierung und Architekturentwicklung Referenzpunkte für die Übertragung auf den Spezialfall der Stammdatenintegration und die Erarbeitung der Methode.

Die zu entwickelnde Methode fokussiert auf Stammdaten und berücksichtigt die in Kapitel 2.2.1 beschriebenen Eigenschaften von Stammdaten. Vor diesem Hintergrund wird nicht zwischen Daten und Information unterschieden. Das Verständnis von Stammdaten beschränkt sich nicht auf reine Zeichenfolgen im Sinne von Daten, sondern umfasst auch deren Interpretation durch ein bestimmtes Schema oder in einem bestimmten Kontext [Otto et al. 2009, 123]. Einschränkend ist anzumerken, dass die Methode ausschliesslich die Integration strukturierter und semistrukturierter Daten behandelt, während unstrukturierte Daten (z. B. Zeichnungen, Dokumente, Dateien) nicht betrachtet werden. Das primäre Ziel der Integration von Stammdaten ist die Erhöhung ihrer Qualität. Daher muss sich eine harmonisierte und integrierte Stammdatenbasis auch in den in Tabelle 2-2 beschriebenen Datenqualitätsdimensionen niederschlagen, anhand derer Stammdatenqualität operationalisiert und gemessen wird.

Den in diesem Kapitel genannten Definitionen folgend werden innerhalb der Methode zur Stammdatenintegration in dieser Arbeit im Wesentlichen zwei Aspekte behandelt: (1) die Gewährleistung einer konsistenten und eindeutigen Definition der wichtigsten Stammdatenobjekte eines Unternehmens aus fachlicher Sicht als eine Kernaufgabe des Stammdatenmanagements und (2) die Schaffung einer verbindlichen Architektur, die Stammdaten logisch integriert für sämtliche Applikationen bereitstellt. Letzteres umfasst Mechanismen zur konsistenten Verteilung von Stammdatenobjekten. Die Arbeit fokussiert somit auf die Gestaltung der Datenarchitektur sowie der Stammdatenlogistik als Teil der Integrationsarchitektur. Die Einordnung und Abgrenzung der Architekturbegriffe in Abbildung 2-8 verdeutlicht, dass der Schwerpunkt der Betrachtung – und dies entspricht dem Verständnis von Stammdaten – auf unternehmensweit, applikationsübergreifend genutzten Entitäten liegt.

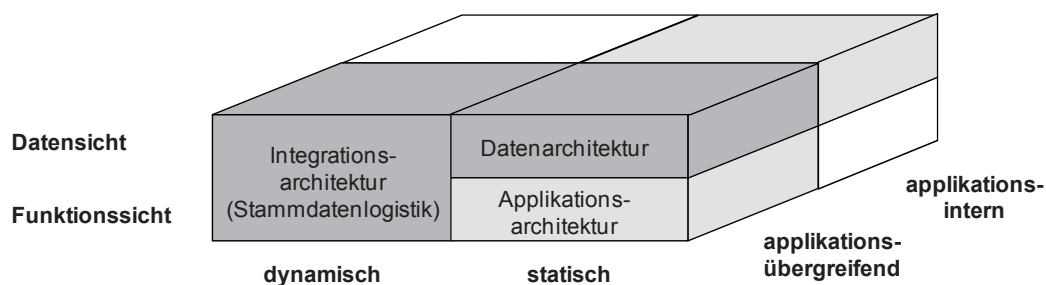


Abbildung 2-8: Betrachtungsumfang der Arbeit

Für das Ziel einer unternehmensweit konsolidierten Datenarchitektur sind jedoch die lokalen, applikationsinternen Datenmodelle zu berücksichtigen (in Abbildung 2-8 auf

Datensicht hellgrau eingefärbte Fläche). Zudem müssen die Modelle der Stammdatenlogistik auf den in der Applikationsarchitektur identifizierten Applikationen aufbauen. Ausgangspunkt der Stammdatenmodellierung ist die unternehmensweite terminologische Vereinheitlichung, d. h. die Stammdatenobjekte müssen aus Sicht der Fachbereiche mit ihrer Semantik definiert und konsolidiert werden. Nur auf dieser Basis ist die anschließende Ableitung von konsistenten Datenmodellen realistisch, die diese Semantik formalisiert abbilden. Die eingesetzten Tools und Notationen zur Dokumentation von Geschäfts- und Datenobjekttypen ist auf die Nutzerzielgruppe zuzuschneiden [Hoberman et al. 2009, 44], wobei die Integration der verschiedenen Modellierungsergebnisse gewährleistet werden muss.

Tabelle 2-7 gibt einen Überblick über die Verwendung der Grundlagen in der vorliegenden Arbeit.

Grundlage	Verwendung in der Dissertation
Integration	Die Arbeit behandelt Probleme der logischen und teilweise auch physischen Datenintegration. In die zu entwickelnde Methode fließen bestehende Konzepte zur Datenintegration ein.
Qualitätsorientiertes Stammdatenmanagement	Das auf dem BE aufbauende CDQM-Framework dient als Rahmenwerk für die vorliegende Arbeit. Die Dissertation fokussiert auf den Gestaltungsbereich der Informationsarchitektur.
Informationsarchitektur	Die Methode zur Integration von Stammdaten muss sowohl die Gestaltung der Datenarchitektur als auch der Integrationsarchitektur (Stammdatenlogistik) adressieren. In Bezug auf die Datenarchitektur ist auf die zielgruppenspezifische, semantisch eindeutige Definition und Modellierung von Geschäfts- und Datenobjekten zu achten (BDD, semantisches Datenmodell) und die konsistente Pflege von Metadaten.

Tabelle 2-7: Verwendung der Grundlagen in der Dissertation

3 Stammdatenintegration in Praxis und Wissenschaft: Bestandsaufnahme und Herausforderungen

Dieses Kapitel fasst den derzeitigen Stand in Forschung und Praxis für die Themen Integration im Allgemeinen und Stammdatenintegration im Speziellen zusammen. Nachdem in Kapitel 3.1 bestehende Probleme bei der Integration in Unternehmen näher betrachtet werden, haben die Kapitel 3.2 und 3.3 einen direkten Bezug zur Integration von Stammdaten. Während Kapitel 3.2 das Thema Standardisierung als Voraussetzung für die Stammdatenintegration behandelt, indem es einen Überblick über bestehende Standards gibt, werden in Kapitel 3.3 methodische Ansätze aus Wissenschaft und Praxis beschrieben und einander vergleichend gegenübergestellt.

3.1 Integrationsprobleme in der Wirtschaftsinformatik

Dieses Kapitel fasst, basierend auf den in Kapitel 2.1 dargelegten Grundlagen, bestehende Integrationsprobleme in der Unternehmenspraxis in Form eines *Integrationsmodells* für das Business Engineering zusammen. Das Integrationsmodell soll wesentliche Objekte und Probleme der Integration enthalten und diese strukturiert abbilden. Hierzu werden auf Grundlage der Klassifikation nach dem Integrationsgegenstand Ebenen und Objekte der Integration identifiziert, denen bestehende Integrationsprobleme zugeordnet werden. Das Integrationsmodell wird in einem ersten Schritt auf Basis einer qualitativen Literaturanalyse abgeleitet, anschliessend anhand von Fallstudien an der Praxis gespiegelt, damit die Relevanz der Integrationsprobleme aus Unternehmenssicht beurteilt werden kann. Da das Integrationsmodell als Synthese praktische Integrationsprobleme zusammenfasst, soll es mit dessen Hilfe möglich sein, die dringendsten Herausforderungen der Integration aufzuzeigen und in der weiteren Arbeit – insbesondere der Methodenentwicklung – zu berücksichtigen.

3.1.1 Herleitung des Integrationsmodells

Wie bereits in Kapitel 2.1.2 erwähnt, bilden die *Integrationsebenen* die oberste Gliederungsebene des Modells. Der in mehreren Ansätzen vorgenommenen Unterteilung der Informationssystemebene folgend wird an dieser Stelle zwischen Applikationen und Daten unterschieden. Die *Integrationsobjekte* stellen die nächstfeinere Gliederungsebene dar (zweite Spalte in Tabelle 3-1). Integrationsobjekte sind Gestaltungsobjekte, die Gegenstand der Integration sein können. Mit der Identifikation von Integrationsobjekten beschäftigten sich in der Vergangenheit verschiedene Autoren (vgl. [Vogler 2004, 33f.]), es können aber auch Metamodelle, welche Gestaltungsobjekte abbilden, als Quellen zur Identifikation genutzt werden. Damit eine konsistente Verwendung der Terminologie im Sinne des BE als Bezugsrahmen dieser Arbeit gewährleistet ist, wird

im Folgenden auf das Metamodell des BE-Methodenkerns (vgl. [Österle et al. 2007, 191; Höning 2009, 96ff.]) zurückgegriffen. Zu den einzelnen Integrationsobjekten werden mögliche Integrationsprobleme identifiziert, indem Veröffentlichungen aus Wissenschaft und Praxis ausgewertet werden, die sich mit der Integration mindestens eines der Gestaltungsobjekte beschäftigen.

Tabelle 3-1 fasst mögliche Integrationsprobleme, die im Folgenden erläutert werden, zusammen (siehe dritte Spalte) und ordnet sie den identifizierten Integrationsobjekten zu.

Integrationsebene		Integrationsobjekt	Mögliche Integrationsprobleme
Strategie / Führung		Geschäftsfeld	<ul style="list-style-type: none"> Fehlende Überlappung bzw. Komplementarität
		Zielsystem	<ul style="list-style-type: none"> Unterschiedliche Unternehmensziele / Zielkonflikte Unterschiedliche Kennzahlen(-systeme)
Aufbauorganisation		Organisationsstruktur	<ul style="list-style-type: none"> Unterschiedliche Primärorganisation (Matrix, Funktional, Stablinie) Unterschiedliche Verantwortungsbereiche von Organisationseinheiten
		Rollenmodell	<ul style="list-style-type: none"> Abweichende Anzahl der Rollen Unterschiedlicher Aufgabenumfang von Rollen
Ablauforganisation		Aktivität / Aufgabe	<ul style="list-style-type: none"> Unterschiedliche Granularität der Aktivitäten Abweichende Meilensteine Unterschiedliche Reihenfolge
		Geschäftsobjekt	<ul style="list-style-type: none"> Semantische Heterogenität zwischen Geschäftsobjekten (Existenz von Synonymen und Homonymen)
Informationssysteme	Applikationen	Applikationsfunktion	<ul style="list-style-type: none"> Zugriffe auf externe Daten und Funktionen Unterschiedliche Funktionsgranularität
		Datentransfer / Kommunikation	<ul style="list-style-type: none"> Technische Heterogenität: Unterschiedliche Kommunikationsprotokolle (HTTP, SOAP) und unterschiedliche Nachrichtenformate (Flatfiles, XML, EDI) Abweichende Nachrichtenstruktur (strukturelle Heterogenität)
	Daten	Datenmodell (Datenobjekttyp / Datenobjektattribut)	<ul style="list-style-type: none"> Datenmodellheterogenität Syntaktische Heterogenität Strukturelle Heterogenität Semantische Heterogenität Unterschiedliche Definition der Datenobjektattribute (z. B. Kardinalität)
		Datenobjekt / Datenelement	<ul style="list-style-type: none"> Duplikate / Unkontrollierte Redundanz Datenwertkonflikte bzw. unzureichende Datenqualität

Tabelle 3-1: Identifizierte Integrationsobjekte und mögliche Integrationsprobleme

Die Strategie eines Unternehmens trifft Aussagen über die Tätigkeitsbereiche, allgemeine Entwicklungsrichtungen, zu erreichende Ziele und Kriterien zu deren Erfolgsmessung [Welge/Al-Laham 2008, 19]. Daher resultieren Integrationsprobleme auf dieser Ebene primär bei Unternehmensfusionen und -akquisitionen, bei denen *Geschäftsfelder* (Komplementarität bzw. Überschneidung der Tätigkeitsbereiche) und das *Zielsystem* (konfligierende Ziele und unterschiedliche Kennzahlensysteme) zwischen

vormals unabhängigen Organisationen langfristig aufeinander abgestimmt werden müssen [Sperling 2007, 83-85].

Auf der Ebene der Aufbauorganisation ist die Abstimmung von *Organisationsstrukturen* nötig [Sperling 2007, 86]. Dies betrifft Unterschiede in den Verantwortungsbereichen und Kernkompetenzen einzelner *Organisationseinheiten* genauso wie in den Strukturen der Primärorganisation. Ein weiterer essenzieller Punkt ist der Abgleich voneinander abweichender *Rollenmodelle* [Vogel 2009, 86-87]. Die Definition der Prozessintegration in Kapitel 2.1.2 verweist auf die Notwendigkeit, einzelne *Aufgaben* (respektive Aktivitäten) applikationsübergreifend miteinander zu verknüpfen. Hierzu sind Abweichungen in deren zeitlicher Abfolge sowie in den mit Aufgaben verbundenen Meilensteinen zu überwinden [Weigand/van den Heuvel 2002; McAfee 2005, 79f.]. Zusätzlich sind auf Prozessebene die an den Prozessschnittstellen auszutauschenden Informationen, deren Struktur und Semantik durch *Geschäftsobjekte* repräsentiert werden, aufeinander abzustimmen [Reimers 2001, 235; Vogel 2009, 82]. Das gemeinsame Vokabular, bestehend aus der Gesamtheit der Geschäftsobjekte, soll ein abgestimmtes Begriffsverständnis zwischen den Akteuren gewährleisten [Legner/Vogel 2008, 45]. Da Geschäftsobjekte applikationsunabhängig aus Sicht des Fachbereiches definiert werden (siehe Kapitel 2.3.1.1), besteht das Ziel vornehmlich darin, zu einem einheitlichen Verständnis unter Vermeidung von Synonymen und Homonymen zu gelangen.

Bei der Integration von Applikationen steht die Zusammenführung verteilter *Applikationsfunktionen* im Mittelpunkt (siehe Definition in Kapitel 2.1.2). Analog zu Geschäftsfunktionen bzw. Aufgaben liegt eine Herausforderung im Abgleich der Granularität zwischen Applikationsfunktionen [Papazoglou/van den Heuvel 2006, 415]. Zudem muss der externe Aufruf von bzw. Zugriff auf die Funktionen ermöglicht werden [Linthicum 2000, 38ff.; Boles et al. 2004, 65f.]. Voraussetzung dafür ist die technische Verknüpfung von Applikationen durch einheitliche Kommunikationsprotokolle und Nachrichtenformate [Jain/Zhao 2003, 211-214; Chari/Seshadri 2004, 61f.].

Auf Datenebene können logische und physische Integrationsanforderungen unterschieden werden (siehe Kapitel 2.1.3). Folglich sind die zu integrierenden Objekte *Datenobjekttypen* und *Datenobjektattribute* auf logischer Ebene sowie *Datenobjekte* und *Datenelemente* auf physischer Ebene (siehe Tabelle 2-3). In Bezug auf die logische Ebene, also das Gestaltungsobjekt *Datenmodell*, wurden mögliche Heterogenitätskonflikte bereits in Kapitel 2.1.3 beschrieben. Auf Datenelementebene stehen Datenwertkonflikte im Vordergrund, die u.a. aus Eingabefehlern oder der Verwendung unterschiedlicher Schreibweisen für Datenwerte resultieren [Conrad 2002, 105]. Die grösste Herausforderung besteht auf dieser Ebene jedoch weiterhin in der Bestimmung und Beseitigung von Duplikaten in Datenbanken (vgl. [Wang/Madnick 1989; Hernandez/Stolfo 1998]).

Bis auf *Datentransfer / Kommunikation* und *Datenmodell* sind sämtliche Gestaltungsobjekte dem Metamodell des BE-Methodenkerns entnommen (vgl. [Österle et al. 2007, 193]). Die beiden Gestaltungsobjekte sind nicht explizit im Metamodell abgebildet, sondern implizit über die reflexiven Assoziationen der Applikationen (Datentransfer) sowie der Datenelemente (Datenmodell) zueinander modelliert. Sie sind zudem als Gestaltungsobjekte im Metamodell der Arbeit von VÖGLER sowie im PROMET Methodenhandbuch zur Prozess- und Systemintegration im BE enthalten [IMG 1996, 257; Vogler 2004, 81]. Das Gestaltungsobjekt *Organisationsstruktur* entspricht der Gesamtheit von Organisationseinheiten und Stellen sowie ihrer Beziehungen zueinander.

Gemäss der in Kapitel 2.1.2 beschriebenen Abhängigkeit der Integrationsebenen untereinander ist die Lösung tiefer liegender Integrationsprobleme Voraussetzung für die Lösung darüber liegender Integrationsprobleme. So kann eine Änderung im Prozessablauf dazu führen, dass unterstützende Applikationen modifiziert und neu miteinander interagieren müssen, was die Entwicklung neuer Schnittstellen und die Integration der Datenmodelle zur Folge hat. Ebenso muss beispielsweise bei der Bewertung von Unternehmensfusionen und -akquisitionen der notwendige Aufwand für die Integration der darunter liegenden Ebenen berücksichtigt werden.

3.1.2 Integrationsprobleme in der Praxis

Das deduktiv, aus der Literatur hergeleitete Integrationsmodell soll in einem zweiten Schritt durch Fallstudien an der Praxis erprobt werden. Hierfür wurden einerseits eigene Fallstudien aufgenommen und mit Fokus auf Integrationsaspekte analysiert (wie im Fall von inet-logistics, Bosch Rexroth und PostLogistics), andererseits wurden bestehende Fallstudien aus dem Forschungsumfeld, die ausdrücklich Integrationsprobleme behandeln, ausgewertet (SOA for Automotive, ITAIDE Beer LL). Sie unterlagen keiner Einschränkung bezüglich Industrie oder Integrationsgegenstand bzw. -ebene, sollten jedoch Probleme bei der Integration thematisieren. Die Fallstudien werden an dieser Stelle kurz charakterisiert. Weiterführende Informationen bzw. die vollständigen Fallstudien sind unter den referenzierten Quellen zu finden.

- *SOA for Automotive* [Vogel et al. 2007; Schmidt et al. 2010b]. Das Forschungsprojekt SOA for Automotive untersuchte am Beispiel des Technischen Änderungsmanagements Anforderungen und Lösungen für die Vernetzung von Herstellern und Zulieferern in der Automobilindustrie. Hierfür wurde eine m:n-fähige, servicebasierte Referenzarchitektur für die überbetriebliche Prozess- und Applikationsintegration entwickelt mit dem Ziel, Durchlaufzeiten und Kosten für die Bearbeitung von Änderungsaufträgen zu verringern und den Aufwand für die Einbindung mehrerer Akteure in die Wertschöpfungskette zu reduzieren. Folglich stehen hier Probleme auf Prozess- und Systemebene im Vordergrund der Analyse.

- *inet-logistics [Schmidt 2009a]*. Die Fallstudie untersucht die überbetriebliche Integration von Akteuren in der Transportlogistik. Durch eine Integrationslösung, in deren Zentrum ein koordinierendes Information Hub steht, soll die Einbindung von Flugzeugherstellern, Logistikdienstleistern und Carriern mit geringem zeitlichen und finanziellen Aufwand ermöglicht und die Transparenz über den Fortschritt von Lieferaufträgen erhöht werden. Schwerpunkt der Betrachtung sind Herausforderungen bei der Applikations- und Datenintegration.
- *Bosch Rexroth [Schmidt et al. 2010b]*. Die Fallstudie Bosch Rexroth ist Bestandteil der vorliegenden Arbeit (siehe Kapitel 4.2.3). Im Fokus der Fallstudie steht die Initiative des Unternehmens, seine in einer heterogenen Applikationsarchitektur verteilten Stammdaten über einen zentralen Stammdatenserver zu harmonisieren und zu integrieren. Dadurch soll die Qualität durch weniger fehlerhafte und redundante Stammdaten verbessert sowie die Aggregation für das unternehmensweite Reporting vereinfacht werden. Entsprechend behandelt die Fallstudie in erster Linie Probleme bei der Applikations- und Datenintegration.
- *PostLogistics [Sfintesco 2009]*. Die Fallstudie PostLogistics untersucht die überbetriebliche Integration von Logistikprozessen eines Telekommunikationsunternehmens über den Intermediär Swiss Post E-Business Solutions (Tochterunternehmen der Schweizerischen Post). Sie wurde im Rahmen einer Master-Arbeit aufgenommen und behandelt detailliert Probleme bei der Integration heterogener Applikationen über Unternehmensgrenzen hinweg. Ziele der Integration waren der effiziente, fehler- und medienbruchfreie Informationsaustausch zwischen den Geschäftspartnern sowie die Erhöhung der Prozessflexibilität des Telekommunikationsunternehmens, z. B. bei Nachfrageschwankungen. Im Mittelpunkt der Betrachtung stehen Integrationsprobleme auf Ebene der Informationssysteme.
- *Bank Austria, Kontron, AIG Group [Vogler 2004]*. Die drei Praxisbeispiele entstammen der Habilitationsschrift von VÖGLER zur Prozess- und Systemintegration und fokussieren folglich auf diese Ebenen des BE. Sie sind nicht umfassend als Fallstudien, sondern als Kurzzusammenfassungen von Aktionsforschungsprojekten dokumentiert.
- *ITAIDE Beer LL [Vogel 2009]*. Gegenstand des Aktionsforschungsprojektes ITAIDE Beer LL ist die Kooperation zwischen Unternehmen und Behörden bei der elektronischen Zollabwicklung verbrauchsteuerpflichtigen Waren. Mit den Bemühungen, die unterschiedlichen Zolldeklarationsprozesse europaweit zu harmonisieren und einen medienbruchfreien, elektronischen Austausch von Zolldokumenten zu ermöglichen, geht der Versuch einher, die administrativen Aufwände für Unternehmen zu senken und die Zollabfertigung zu beschleunigen. Der Autor dieser Dissertation war im Aktionsforschungsprojekt aktiv involviert. Herausforde-

rungen und Ergebnisse der Arbeit sind als Bestandteil eines Dissertationsvorhabens am Institut dokumentiert.

Die Auswertung der Fallstudien anhand der aus der Literatur abgeleiteten Klassifikation von Informationsproblemen (siehe Tabelle 3-1) ist in Tabelle 3-2 zusammengefasst.

IntegrationsEbene	Integrationsproblem	SOA for Automotive	inet-logistics	Bosch Rexroth	Post Logistics	Bank Austria, Kontron, AIG	ITAIDE Beer LL
Strategie	Fehlende Überlappung bzw. Ergänzung						
	Zielkonflikte		X		X		
	Unterschiedliche Kennzahlen(-systeme)						
Aufbauorganisation	Unterschiedliche Primärorganisation						
	Unterschiedliche Verantwortungsbereiche (Organisationseinheiten)						
	Abweichende Anzahl der Rollen	X					
	Unterschiedlicher Aufgabenumfang (Rollen)	X					
Ablauforganisation	Unterschiedliche Granularität der Aktivitäten	X					
	Unterschiedliche Reihenfolge	X				X	X
	Semantische Heterogenität zwischen Geschäftsobjekten	X	X				X
Applikationen	Externe Zugriffe auf Daten / Funktionen			X		X	
	Unterschiedliche Funktionsgranularität					X	
	Technische Heterogenität		X		X	X	
	Abweichende Nachrichtenstruktur	X					X
Daten	Datenmodellheterogenität	X				X	
	Syntaktische Heterogenität	X		X	X		
	Strukturelle Heterogenität	X		X	X		
	Semantische Heterogenität	X	X	X	X	X	X
	Unterschiedliche Definition der Attribute		X		X		
	Duplikate / Unkontrollierte Redundanz			X		X	
	Datenwertkonflikte		X				

Tabelle 3-2: Integrationsprobleme der Fallstudien

Die Auswertung zeigt, dass in den Fallstudien zu einem Grossteil der Problemkategorien konkrete Integrationsprobleme identifiziert werden können. Da die Fallstudien vorwiegend Probleme bei der Prozess-, Applikations- und Datenintegration behandeln, sind die unteren Ebenen des Integrationsmodells umfassender abgedeckt. Integrationsprobleme auf strategischer sowie aufbauorganisatorischer Ebene konnten dagegen nur bedingt gefunden werden. Der Grund dafür ist, dass Integrationsprobleme dieser Ebenen in erster Linie bei der vollständigen Vereinigung ganzer Unternehmensteile bzw. Unternehmen von Bedeutung sind (z. B. bei Akquisitionen und Fusionen), diese jedoch nicht Gegenstand der Fallstudien waren.

Besonders auffällig ist die in allen Fallstudien übereinstimmende Nennung und Beschreibung von semantischen Heterogenitätskonflikten bei der Integration von Daten-

objekttypen und Datenobjektattributen (als Bestandteil von Datenmodellen) [Schmidt et al. 2010b, 20]. Das Problem der Semantik ist zudem in ähnlicher Weise auf Prozessebene in Bezug auf Geschäftsobjekte bzw. Geschäftsobjekttypen erkennbar (vgl. [Papazoglou 2007, 549]). Diesbezüglich verweisen die Fallstudien inet-logistics und Bosch Rexroth sowie SOA for Automotive auf das Problem, dass der Verwendungskontext als wesentliches Element für die Integration von Geschäfts- und Datenobjekttypen in den Informationsmodellen zu wenig berücksichtigt wird (vgl. [Madnick 1999, 3]). Die Folge ist, dass bei der Integration von Applikationen und Daten weiterhin aufwendige manuelle Mappings oder Umgehungslösungen erforderlich sind. Die Fallstudien zeigen, dass für die Bewältigung dieses Problems wirkungsvolle Lösungskonzepte trotz intensiver Forschung in den vergangenen Jahren fehlen (vgl. [March et al. 2000, 330; Park/Ram 2004, 596]) bzw. diese in der betrieblichen Praxis aufgrund des hohen Aufwandes (noch) nicht anwendbar sind (wie z. B. das Konzept der Ontologien). Zudem mangelt es an semantischen Standards, welche – in Analogie zu Protokoll- und Formatstandards auf syntaktischer Ebene – zur Überwindung der Heterogenität beitragen. Die Frage der semantischen Integration muss daher auch im Rahmen dieser Arbeit zur Integration von Stammdaten behandelt werden.

Im Hinblick auf die in dieser Arbeit zu entwickelnde Methode und das Integrationsobjekt Stammdaten sind ausserdem die Probleme der syntaktischen und strukturellen Heterogenität (auf Typebene) sowie die Vermeidung von unkontrollierter Redundanz und Datenwertkonflikten (auf Instanzebene) zu beachten.

3.2 Standardisierung von Stammdaten

Voraussetzung für eine effiziente und flexible Integration, unabhängig vom Integrationsgegenstand, ist die Existenz von Standards [Chari/Seshadri 2004, 59]. Standards vereinheitlichen die Objekte eines bestimmten Anwendungsbereiches und vereinfachen damit die Kommunikation zwischen menschlichen oder maschinellen Aufgabenträgern [Buxmann 1996, 10]. Im Stammdatenmanagement sind Standards von grosser Bedeutung, da sie einen Beitrag zur Definition einer gemeinsamen Sprache leisten [Schemm 2008, 37] und Voraussetzung für die perfekte Bedienung des Kundenprozesses sind [Kagermann/Österle 2006, 135].

3.2.1 Arten von Stammdatenstandards

In Bezug auf den Gegenstand der Standardisierung können verschiedene Arten von Standards unterschieden werden (siehe Tabelle 3-3), die – ähnlich wie die Integrationsebenen im vorherigen Kapitel – aufeinander aufbauen. So ist die Standardisierung von Formaten die Voraussetzung für die semantische Standardisierung von Daten und Nachrichten, die wiederum Grundlage für eine Prozessstandardisierung sind.

Art des Standards	Gegenstand der Standardisierung
Formatstandards	Vereinheitlichung der syntaktischen Codierung von Daten durch Spezifikation von Reihenfolge, Länge und Typ der Datenelemente
Datenstandards	Vereinheitlichung der Struktur und Semantik von Daten durch die Spezifikation der Bedeutung von Datenelementen
Nachrichten- / Dokumentenstandards	Vereinheitlichung der Aggregation bzw. Verknüpfung von Datenelementen zu Nachrichten zur Übertragung zwischen Informationssystemen
Prozessstandards	Vereinheitlichung der organisatorischen Abläufe durch Festlegung der Abhängigkeiten zwischen einzelnen Aufgaben
Geschäftsstandards	Einheitliche Regelung rechtlicher Rahmenbedingungen

Tabelle 3-3: Arten von Standards [Heutschi et al. 2004, 134f.; Schemm 2008, 53]

Für die vorliegende Arbeit und für das Ziel der Integration von Stammdaten sind vor allem Datenstandards von Bedeutung. *Katalogaustauschstandards* (wie beispielsweise BMEcat) sowie *Transaktionsstandards* (wie beispielsweise openTRANS) gehören nicht zur Kategorie der Datenstandards. Sie fokussieren auf die Übertragung von Daten zwischen Informationssystemen und werden daher den Nachrichten- bzw. Dokumentenstandards zugeordnet (siehe Tabelle 3-3). Betrachtet man die Datenstandards näher, so können diese gemäss ihrem Zweck weiter in drei Unterkategorien unterteilt werden [Otto et al. 2002, 16ff.; Quantz/Wichmann 2003, 32ff.]:

- *Identifikationsstandards* verfolgen das Ziel, Objekte global mit einer eindeutigen Kennung (z. B. einer Nummer) auszuzeichnen und sie anhand dieser zu erkennen. Die am weitesten verbreiteten Standards zur Identifikation von Produkten sind der Universal Product Code (UPC) in den USA sowie die European Article Number (EAN) und die Global Trade Item Number (GTIN). Für die Identifikation von Unternehmen hat sich das amerikanische Data Universal Numbering System (D-U-N-S) von Dun & Bradstreet durchgesetzt.
- *Klassifikationsstandards* gruppieren Objekte mit gleichen Eigenschaften anhand definierter Kriterien und ordnen die daraus resultierenden Klassen in einer hierarchischen Struktur an, woraus sich eine Taxonomie der Objekte ergibt. Durch die Integration von sogenannten Merkmalsleisten (bestehend aus Merkmalen eines Objektes) unterstützen einige Klassifikationsstandards auch die Beschreibung von Objekten.
- *Beschreibungsstandards* wiederum spezifizieren die Eigenschaften von Objekten eines Anwendungsbereiches, indem sie Datenelemente inhaltlich definieren.

Für die (semantische) Stammdatenintegration und dem damit verbundenen Ziel, ein einheitliches Verständnis von Stammdatenobjekten zu etablieren (siehe Methodenentwicklung in Kapitel 5), sind vor allem Klassifikations- und Beschreibungsstandards von Bedeutung. Sie liefern Informationen über die Semantik von Daten und werden daher im Folgenden näher betrachtet. Die Darstellung beschränkt sich dabei auf die für das Stammdatenmanagement bedeutenden Standards.

3.2.2 Klassifikationsstandards

Sämtliche *Klassifikationsstandards* verfolgen einen ähnlichen Ansatz: Es werden Materialien, Produkte und Dienstleistungen mit gleichen Eigenschaften zu einer Klasse zusammengefasst und diese dann zu übergeordneten Klassen systematisch aggregiert. Der Schwerpunkt dieser Standards liegt auf Artikel- und Materialstammdaten. Die Klassen sind hierarchisch in einer Baumstruktur auf mehreren Ebenen angeordnet, wobei die Produkte auf der untersten Ebene (höchster Detaillierungsgrad) einer Produktklasse zugeordnet werden [Schmidt/Osl 2008, 12]. Jede Hierarchieebene ist mit einem – in der Regel numerischen – Code versehen. So vergeben sowohl der United Nations Standard Products and Services Code (UN/SPSC) und die Global Product Classification (GPC) als auch eCl@ss für jede der vier Hierarchieebenen ihres Klassifikationssystems einen zweistelligen numerischen Schlüssel. Eine Produktklasse auf unterster Ebene ist folglich durch einen achtstelligen Code eindeutig gekennzeichnet. Durch die Hierarchisierung von Produktklassen wird das Einordnen und Auffinden von Produkten vereinfacht.

Ein wichtiges Element vieler Klassifikationsstandards sind sogenannte *Merkmalsleisten* (vgl. [Otto/Beckmann 2001, 352; Hepp 2003, 91]), die den Produktklassen auf unterster Ebene zugeordnet werden können. Sie erlauben die Beschreibung eines Produktes durch Attribute (Merkmale) und zugeordnete Attributwerte [Dorloff et al. 2001, 1529]. Diese beschreibenden Merkmale sind in der Regel auf Basis von ISO- oder DIN-Normen standardisiert. So organisiert und pflegt das Deutsche Institut für Normung (DIN) das deutsche Merkmallexikon. Klassifikationsstandards unterstützen damit die Entwicklung eindeutiger Definitionen von Artikel- und Materialstammdaten, indem sie nicht nur strukturelle Metadaten enthalten (durch die strukturierte Klassifikation), sondern zusätzlich auch semantische Informationen (durch die Merkmalsleisten) umfassen.

Wesentliche Klassifikationsstandards sind in Tabelle 3-4 zusammenfassend charakterisiert (vgl. [Otto et al. 2002, 15-16; Quantz/Wichmann 2003, 188 und 194; ZVEI 2006, 32-47]).

	UN/SPSC ¹⁰	GPC ¹¹	eCl@ss ¹²	ETIM ¹³	proficl@ss ¹⁴
Aktuelle Version	11.1201	08/2009	6.1	4.0	4.0
Organisation	GS1 US (vorher United Nations Development Programme)	GS1 Germany	eCl@ss e. V.	ETIM e. V.	proficl@ss International e. V.
Branchenbezug	Branchenunabhängig	Primär Konsumgüterindustrie	Branchenunabhängig	Elektrotechnik	Handel, Handwerk, Industrie
Geographische Ausrichtung	Global (primär USA)	Global (primär USA)	Global (primär Europa)	Primär Deutschland	Primär Deutschland
Umfang	>21.000 Klassen	Ca. 10.000 Klassen	32.795 Klassen 10.930 Merkmale	2.722 Klassen	4.296 Klassen 3.567 Merkmale
Hierarchieebenen	<ul style="list-style-type: none"> • Segment • Family • Class • Commodity • (Business Function) 	<ul style="list-style-type: none"> • Segment • Family • Class • Brick • 	<ul style="list-style-type: none"> • Sachgebiet • Hauptgruppe • Gruppe • Untergruppe 	<ul style="list-style-type: none"> • Artikelgruppe • Artikelklasse 	Variabel (nicht vorgegeben)
Merkmale/Merkmalsteilen	Nein	Ja	Ja (Basis- und Standardmerkmalsteilen)	Ja	Ja
Verwaltung von Synonymen	Nein	Nein	Ja	Ja (aber nur für Klassen, nicht für Merkmale)	Ja

Tabelle 3-4: Charakterisierung von Klassifikationsstandards

3.2.3 Beschreibungsstandards

Beschreibungsstandards stellen für einen bestimmten Anwendungsbereich einheitliche semantische Objekt- und Attributdefinitionen in Form von Datenkatalogen zur Verfügung. Die eindeutigen Definitionen sind unter anderem Voraussetzung für die richtige Verwendung der Attribute in Nachrichten und Dokumenten. Die Datenkataloge sind dementsprechend in der Regel Bestandteil einer Standardfamilie.

Ein solcher Datenkatalog ist das *Global Data Dictionary (GDD)* aus dem Umfeld der GS1-Standardfamilie. Das GDD enthält Definitionen sämtlicher Stammdatenattribute, die in den XML-Standards von GS1 verwendet werden – in erster Linie Attribute von Artikeln, jedoch auch von Geschäftspartnern und Konditionen (vgl. [GS1 2009]). Die entsprechenden Datenmodelle sind sowohl in Form von UML-Klassendiagrammen als auch tabellarisch dokumentiert [Schemm 2008, 42]. Zur Beschreibung der Stammdatenobjekte nutzt GS1 zwei verschiedene Attributtypen [GS1 2005, 12f.; Schemm 2008, 43]: Während *Kernattribute* (engl.: core attributes) Eigenschaften von Stammdatenobjekten beschreiben, die für sämtliche Branchen und geographische Regionen, al-

¹⁰ Siehe <http://www.unspsc.org/>

¹¹ <http://www.gs1.org/gdsn/gpc>

¹² Siehe <http://www.eclass.de/>

¹³ Siehe <http://www.etim.de/>

¹⁴ Siehe <http://www.proficlass.de/>

so global, gültig sind, ist die Verwendung von *Erweiterungsattributen* (engl.: *extension attributes*) auf spezifische Branchen oder Anwendungsfälle beschränkt.

Die *Core Component Library (CCL)* der UN/CEFACT ist ein Verzeichnis für Datenobjekttypen und Datenobjektattribute, das unter anderem standardisierte Namen und Beschreibungen von Stammdatenobjekten (z. B. Kunde, Lieferant) und von deren Attributen (z. B. Adresse) enthält. Diese sind entsprechend der Core Component Technical Specification (siehe Kapitel 5.6.1) bzw. dem ISO-Standard 11179-5 spezifiziert. Die CCL ist in Form eines Excel-Spreadsheets abrufbar (vgl. [UN/CEFACT 2009a]). In Bezug auf Umfang und Reife der aktuellen Version der CCL ist einschränkend anzumerken, dass sie sich noch in einem relativ frühen Stadium befindet [Vogel 2009, 36]. Zurzeit arbeiten jedoch unterschiedliche Trade and Business Process Groups der UN/CEFACT daran, Objekte und Attribute für branchenunabhängige und -spezifische Anwendungsbereiche zu definieren und in die Core Component Library einzubringen. Für die semantische Konsistenz der CCL-Inhalte ist eine gesonderte Arbeitsgruppe innerhalb der UN/CEFACT zuständig, die durch einen dreistufigen Prozess der Harmonisierung, Bewertung und Bewilligung neuer Definitionen die standardkonforme Benennung und Beschreibung gewährleistet und die Aufnahme redundanter Einträge verhindert [Lemm 2007, 33].

Ein letzter Beschreibungsstandard ist das *ECCMA Open Technical Dictionary (eOTD)*, welches von der Electronic Commerce Code Management Association (ECCMA) herausgegeben wird. Auch eOTD ist ein horizontaler, d. h. branchenübergreifend gültiger Standard, der jedoch relativ wenig verbreitet ist. Das eOTD beschreibt Konzepte (z. B. Produkte, Dienstleistungen, Organisationen) anhand von Definitionen, Eigenschaften und Graphiken und weist ihnen einen eindeutigen „concept identifier“ zu [Arnett 2005, 5-9]. Die einzelnen Konzepte stehen in keiner hierarchischen Struktur, sondern sind nur sehr generischen Produktklassen zugeordnet [Hepp et al. 2006, 92-95].

Zusätzlich zu den vorgestellten, branchenübergreifenden Beschreibungsstandards existieren zahlreiche, branchenspezifische Beschreibungsstandards, wie das *Shared Information & Data Model (SID)* für die Telekommunikationsindustrie oder der *Standard for the Exchange of Product Model Data (STEP)* für die Automatisierungsindustrie.

3.3 Methodische Ansätze zur Integration von Stammdaten

Das Ziel der Arbeit ist die Entwicklung einer Methode zur Integration von Stammdaten. Der Vorschlag eines neuen systematischen Vorgehens ist nur dann sinnvoll, wenn keine Methode vorhanden ist, die diese Problematik bereits behandelt, oder existierende Methoden in Bezug auf die an sie gestellten Anforderungen unvollständig sind.

Kriterien, die eine Methode zur Stammdatenintegration erfüllen muss, ergeben sich zum einen aus den in der Einleitung dargelegten Anforderungen und Gestaltungsberei-

chen der Stammdatenintegration. Anhand dieser Kriterien werden die bestehenden Ansätze bezüglich ihres inhaltlichen Beitrages (*Inhaltliche Charakteristika*) bewertet. Zum anderen fließen Anforderungen an ein methodisches Vorgehen zur Beurteilung der Ansätze ein. Diese lassen sich aus den Elementen des ME (siehe Kapitel 5.1) ableiten, die für den Entwurf von Methoden bzw. für die Gewährleistung eines systematischen Vorgehens massgeblich sind. Die Kategorie *Strukturelle Charakteristika* fasst die methodischen Anforderungen zusammen. Tabelle 3-5 gibt einen Überblick über die verwendeten Bewertungskriterien und deren Bedeutung.

Bewertungskriterium	Beschreibung
Inhaltliche Charakteristika	
Bezug zu Stammdaten	Fokussierung auf Stammdaten und Berücksichtigung ihrer spezifischen Eigenschaften bei der Integration
Eindeutige Definition von Geschäftsobjekttypen	Beschreibung eines systematischen Vorgehens zur semantisch eindeutigen Definition von Geschäftsobjekttypen
Management fachlicher Metadaten	Definition eines Vorgehens für die konsistente Pflege fachlicher Metadaten zu den Geschäftsobjekten (z. B. in einem BDD) sowie zur Integration dieser Metadaten in die Datenpflegeprozesse
Ableitung eines semantischen Stammdatenmodells	Spezifikation eines Ansatzes zur Überführung der in einem BDD gepflegten Definitionen in ein semantisches Stammdatenmodell
Entwicklung der Integrationsarchitektur (Stammdatenlogistik)	Beschreibung von Handlungsempfehlungen zur konsistenten Distribution von Stammdaten in verteilten Applikationsarchitekturen
Werkzeugunterstützung	Aufzeigen möglicher Software-Tools zur Unterstützung der vorherigen Aufgaben
Strukturelle Charakteristika	
Metamodell	Spezifikation eines Metamodells als terminologischen und konzeptionellen Rahmen der Methode
Vorgehensmodell	Definition von Entwurfsaktivitäten und ihrer zeitlich-sachlogischen Abhängigkeit
Techniken	Vorgabe detaillierter Anleitungen zur Erzeugung der Entwurfsergebnisse
Entwurfsergebnisse	Beschreibung der Dokumente und Modelle die pro Entwurfsaktivität erstellt werden
Rollenmodell	Definition der für die Methode notwendigen organisatorischen Rollen und Zuordnung der Rollen zu einzelnen Entwurfsaktivitäten

Tabelle 3-5: Kriterien zur Bewertung bestehender Ansätze

Der folgende Abschnitt beschreibt einleitend die Auswahl bestehender Ansätze anhand definierter Auswahlkriterien, bevor diese Ansätze vorgestellt und auf Erfüllung der in diesem Kapitel spezifizierten Anforderungen überprüft werden.

3.3.1 Auswahl bestehender Ansätze

Für die Bewertung bestehender Methoden bzw. systematischer Ansätze wurden wissenschaftliche und praxisorientierte Veröffentlichungen (Bücher, Dissertationen, Zeitschriftenartikel, Konferenzbeiträge, White Paper, Vortragspräsentationen) berücksichtigt, die sich mit der Integration von Stammdaten oder Teilaspekten, die zur Integration notwendig sind, beschäftigen (siehe Kapitel 2.2 und 2.3). Die Veröffentlichungen wurden mit Hilfe elektronischer Datenbanken durch die Suche im Volltext

nach den Stichworten „Stammdatenintegration“ („master data integration“), „Stammdatenmanagement“ („master data management“) sowie teilweise – bei geringer Anzahl von gefundenen Beiträgen – „Stammdaten“ („master data“) allgemein identifiziert. Anschliessend wurden die Veröffentlichungen inhaltlich anhand der Kriterien geprüft. Hierbei wurde vor allem darauf geachtet, dass die Publikationen ein methodisches Vorgehen beschreiben oder zumindest Methodenbestandteile enthalten. Da die identifizierten Zeitschriftenartikel und Konferenzbeiträge diesem Kriterium in der Regel nur unzureichend genügten, wurden von den jeweiligen Autoren publizierte Bücher in die Analyse einbezogen. Dabei flossen auch Publikationen ein, die nicht explizit die Integration von Stammdaten, sondern allgemeiner von unternehmensweit genutzten Daten behandeln (siehe beispielsweise [Heine 1999]).

Tabelle 3-6 fasst die bewerteten Ansätze zusammen. Eine ausführlichere Beschreibung sowie eine Analyse der Ansätze anhand der in Tabelle 3-5 definierten Kriterien folgt im anschliessenden Kapitel.

Ansatz/Veröffentlichung	Kurzbeschreibung
Reference and Master Data Management [DAMA 2009]	„Reference and Master Data Management“ ist Teil des von der DAMA ¹⁵ veröffentlichtes Handbuches „The DAMA Guide to the Data Management Body of Knowledge“, welches praxisbewährte methodische Ansätze für das (Stamm-)Datenmanagement beschreibt.
Master Data Management and Customer Data Integration for a Global Enterprise [Berson/Dubov 2007]	„Master Data Management and Customer Data Integration for a Global Enterprise“ entwickelt Lösungskonzepte für den Entwurf und die Implementierung von Integrationsarchitekturen für MDM und CDI.
Master Data Management [Loshin 2008]	„Master Data Management“ führt in das Stammdatenmanagement ein und entwirft ein vereinfachtes Vorgehen in MDM-Projekten mit speziellem Fokus auf Integrationsfragestellungen und Anforderungen an die Datenqualität.
Enterprise Master Data Management [Dreibelbis et al. 2008]	„Enterprise Master Data Management“ bietet einen Überblick über grundlegende technische Konzepte für das Stammdatenmanagement und Handlungsempfehlungen für die Entwicklung einer integrierten Stammdatenarchitektur.
Unternehmensweite Datenintegration [Heine 1999]	„Unternehmensweite Datenintegration“ beschreibt das Konzept der Datenlogistik als eine zweckmässige Vorgehensweise zur Integration unternehmensweit genutzter Daten auf Grundlage modularer Daten- und Prozessstrukturen sowie einer Erweiterung des Data Warehouse-Konzeptes.

Tabelle 3-6: Ansätze zur Integration von Stammdaten

Nicht in die Analyse flossen Veröffentlichungen ein, die

- Methoden und Vorgehensmodelle zum Datenqualitätsmanagement allgemein vorschlagen, ohne konkrete Probleme der Stammdatenintegration zu behandeln (z. B. [English 1999], [Wang et al. 2000], [Lee et al. 2006], [Batini/Scannapieco 2006]),
- lediglich auf Probleme beim Management und der Integration von Stammdaten hinweisen, ohne anwendbare Konzepte oder Techniken zu deren Lösung vorzuschlagen (z. B. [Smith/McKeen 2008], [Cleven/Wortmann 2010]),

¹⁵ Die DAMA (www.dama.org) ist eine internationale, anbieterunabhängige Non-Profit-Organisation, die sich aus Fachexperten zusammensetzt und das Ziel verfolgt, Konzepte für das Daten- und Informationsmanagement (weiter) zu entwickeln.

- Stammdatenintegration aus einer rein technischen Perspektive betrachten und sich beispielsweise ausschliesslich auf Probleme bei der Schemaintegration konzentrieren (z. B. [Moraes Batista/Salgado 2007]).

3.3.2 Reference and Master Data Management

Die Data Management Association (DAMA) nennt das Stammdatenmanagement in ihrem Leitfaden *Guide to the Data Management of Knowledge* als eines von insgesamt zehn Gestaltungsfeldern des Datenmanagements [DAMA 2009, 7]. In dem Gestaltungsfeld *Reference and Master Data Management* (vgl. [DAMA 2009, 171-196]) werden zehn Aktivitäten definiert, die im Sinne eines Vorgehensmodells in eine sachlogische Reihenfolge gebracht und vier Kategorien (Planning, Control, Development, Operational) zugeordnet werden. Die Aktivitäten sind jeweils mit der zugehörigen Kategorie in Tabelle 3-7 zusammengefasst.

Kategorie	Aktivität
Planning	Understand Reference and Master Data Integration Needs
	Identify Master and Reference Data Sources and Contributors
	Define and Maintain the Data Integration Architecture
Development	Implement Reference and Master Data Management Solutions
Control	Define and Maintain Match Rules
	Establish "Golden" Records
	Define and Maintain Hierarchies and Affiliations
Development	Plan and Implement Integration of New Data Sources
Operational	Replicate and Distribute Reference and Master Data
	Manage Changes to Reference and Master Data

Tabelle 3-7: Aktivitäten des Stammdatenmanagements nach [DAMA 2009, 172]

Der Leitfaden der DAMA stellt die Herausforderung einer integrierten Stammdatenbasis in den Mittelpunkt der Betrachtung. Der Leitfaden behandelt das Thema der Entwicklung und Pflege von Integrationsarchitekturen für Stammdaten sehr ausführlich, diskutiert verschiedene Architekturmuster und beschreibt die zugrunde liegenden Tools bzw. Technologien. Fragen der Stammdatenverteilung und Replikation (als Teil der Stammdatenlogistik) werden nur sehr oberflächlich behandelt. Die semantisch eindeutige Definition von Geschäftsobjekten als Grundlage für die Implementierung eines „Golden Record“ (als ein integriertes System, welches verbindliche Stammdatenobjekte speichert) wird unter dem Begriff Vocabulary Management behandelt. Hierbei verweist die DAMA vor allem auf die Herausforderung der Erkennung und Dokumentation von Synonymen. Die Anleitungen zum Vocabulary Management sowie zur Pflege von Metadaten beschränken sich zu einem grossen Teil jedoch auf eine Aufzählung wesentlicher Problemstellungen und Aufgaben, ohne konkrete Techniken zu deren Lösung anzubieten. Das Problem der Integration von Semantik aus dem Vo-

cabulary Management in die Datenmodelle bleibt in der Spezifikation gänzlich unberücksichtigt.

Die DAMA bietet zusätzlich ein Glossar, das wesentliche Gestaltungsobjekte kurz definiert (vgl. [DAMA 2008]). Da das Glossar die Relationen zwischen diesen Objekten jedoch nur sehr rudimentär beschreibt und sich zudem auf das gesamte Datenmanagement und nicht speziell auf das Stammdatenmanagement bezieht, kann es nur bedingt als Metamodell zum Vorgehensmodell verstanden werden. Das Kapitel zum Stammdatenmanagement enthält kein eigenes Rollenmodell, sondern greift auf die im Gestaltungsfeld Data Governance definierten Rollen des Datenmanagements zurück (vgl. [DAMA 2009, 171]). Diese werden zusammen mit wichtigen Ergebnisdokumenten (Deliverables) in einer abschliessenden Tabelle den Aktivitäten zugeordnet (vgl. [DAMA 2009, 193-195]). Die Inhalte der Ergebnisdokumente werden nicht detailliert.

3.3.3 Master Data Management and Customer Data Integration for a Global Enterprise

BERSON und DUBOV¹⁶ beziehen sich in ihrer Veröffentlichung dediziert auf Stammdaten mit einem starken Fokus auf Kundenstammdaten. Die vorgeschlagenen Konzepte können aber grösstenteils auf Herausforderungen bei der Integration von Stammdaten allgemein übertragen werden. Sehr ausführlich wird die Entwicklung von Integrationsarchitekturen zur Integration von (Kunden-)Stammdaten thematisiert. Hierfür werden Architekturvarianten vorgestellt, einzelne Komponenten einer Integrationsarchitektur (z. B. Operational Data Stores, Metadaten-Repositorys, Business-Rule-Engine usw.) beschrieben und Technologien zur Erkennung von mehreren Datenobjekten, die dasselbe Geschäftsobjekt abbilden, dargelegt (vgl. [Berson/Dubov 2007, 59-130]). Die Integration auf einer fachkonzeptionellen Ebene (terminologische Vereinheitlichung der Geschäftsobjekte und Ableitung eines semantischen Datenmodells) sowie die konsistente Pflege von Metadaten und deren Integration in Datenpflegeprozesse werden nicht detailliert behandelt. Ein separates Kapitel führt zudem kommerzielle Softwaretools auf, die Funktionen für das Management von (Kunden-)Stammdaten bieten (vgl. [Berson/Dubov 2007, 347-365]).

Die Autoren beschreiben kein systematisches Vorgehen im Sinne des ME, da bereits die beiden wichtigsten Bestandteile einer Methode – das Metamodell und das Vorgehensmodell – nicht spezifiziert werden. Vielmehr werden punktuell Lösungskonzepte für einzelne Probleme (z. B. Architekturvarianten für die Implementierung eines Data Hubs zur Stammdatenintegration) diskutiert, ohne dass diese in eine sachlogische Reihenfolge gebracht werden. Zwar enthält die Veröffentlichung eine Liste von insgesamt 19 Aufgabenbereichen, die in einem Stammdatenmanagementprojekt zu berücksichti-

¹⁶ Die Autoren sind Fachexperten des MDM Institutes (www.tcdii.com), das Lösungskonzepte im Bereich Stammdatenmanagement, Customer Data Integration und Data Governance entwickelt.

gen sind (vgl. [Berson/Dubov 2007, 248]), deren Beziehungen zueinander sowie deren Inhalte sind aber nur unzureichend definiert. Die notwendigen Ergebnisdokumente werden nicht spezifiziert, sondern es wird lediglich auf die Notwendigkeit klar definierter „Deliverables“ verwiesen. Bei den Rollen werden lediglich Data Owner und Data Stewards unterschieden (vgl. [Berson/Dubov 2007, 111f.]), eine Zuordnung zu Aufgaben des MDM bzw. des CDI wird nicht vorgenommen.

3.3.4 Master Data Management

LOSHIN erläutert in seiner Veröffentlichung wesentliche Gestaltungsfelder von MDM-Projekten. Hierbei geht er detailliert auf Aktivitäten des Stammdatenmanagements sowie der Datenintegration ein, die teilweise für die Methodenentwicklung in der vorliegenden Arbeit von Bedeutung sind. Die Aktivitäten sind in Tabelle 3-8 charakterisiert.

Aktivität	Kurzbeschreibung
1. Establishing a business justification for master data integration and management	Herstellen des Bezuges des MDM-Projektes zu anderen Initiativen, welche von einer integrierten Stammdatenbasis profitieren
2. Developing an MDM road map and rollout plan	Erstellung eines langfristigen Entwicklungsplans für das MDM-Projekt mit wesentlichen Aufgabenbereichen
3. Roles and responsibilities	Definition von Rollen und Verantwortlichkeiten für einzelne Aufgaben
4. Project planning	Festlegen eines detaillierten Projekt- und Meilensteinplanes für das MDM-Projekt auf Grundlage von Roadmap, Rollout-Plan und definierten Rollen
5. Business process models and usage scenarios	Detaillierte Spezifikation der Geschäftsprozesse und Nutzungsszenarien, um relevante Datenobjekte und deren Verwendung zu identifizieren
6. Identifying initial data sets for master integration	Identifikation von stammdatennutzenden Applikationen und der darin enthaltenen Stammdaten auf Grundlage der Geschäftsprozessmodelle
7. Data governance	Erarbeitung von Richtlinien und Standards für das Datenmanagement im Sinne der Unternehmensanforderungen und Überwachung der Einhaltung
8. Metadata	Einführung eines Metadatenmanagements (inklusive Metadatenprozesse und -Repository) zur Speicherung eindeutiger Definitionen und zur kontrollierten Pflege von Stammdaten
9. Master object analysis	Dokumentation der Stammdatenobjekte in verschiedenen Applikationen als Grundlage für die Konsolidierung und Integration über die verschiedenen Applikationen hinweg
10. Master object modeling	Modellierung der (konsolidierten) Stammdatenobjekte
11. Data quality management	Erarbeitung von Techniken zur Gewährleistung von Datenqualität (auf Grundlage der Data Governance-Richtlinien) und Auswahl sowie Implementierung von Datenqualitätstools
12. Data extraction, sharing, consolidation, and population	Spezifikation von Prozessen, Regeln und Tools zur physischen Konsolidierung (Extraktion, Transformation) von Stammdaten
13. MDM architecture	Gestaltung der Integrationsarchitektur durch Auswahl eines präferierten Architekturmusters abhängig von den unternehmensspezifischen Anforderungen und Gegebenheiten (Anzahl der verwalteten Objekte, Kopplung der Applikationen, Integrationsanforderung usw.)
14. Master data services	Implementierung von Stammdatenservices zur Nutzung von Stammdatenobjekten, Zugriffskontrolle usw.
15. Transition plan	Überführung von Stammdatenbeständen aus Legacy-Systemen
16. Ongoing maintenance	Weiterentwicklung der implementierten Prozesse, Richtlinien und Architektur und permanentes Monitoring der Stammdatenqualität

Tabelle 3-8: Aktivitäten eines MDM-Projektes nach [Loshin 2008, 237-256]

In den Aktivitäten 8 bis 10 setzt sich LOSHIN ausführlich mit den in der Einleitung beschriebenen Problemen der Datenarchitektur (unternehmensweit einheitliche Semantik der Geschäftsobjekte, Identifikation und Modellierung von Kerndatenelementen sowie Metadatenmanagement) auseinander und betrachtet die Datenintegration auf fachkonzeptioneller Ebene als wesentlich für den Erfolg des Projektes (vgl. [Loshin 2008, 105-157]). Für die Aktivitäten schlägt er konkrete Arbeitsschritte und Techniken vor. Der Autor verweist auch auf die Notwendigkeit, den Kontext der Stammdatennutzung (Semantik) in die Datenmodelle zu integrieren, wobei er diesbezüglich aber keine Hilfestellung bietet. In Bezug auf Stammdatenintegrationsarchitekturen werden verschiedene Architekturvarianten abhängig von bestimmten Nutzungsszenarien beschrieben und Auswahlkriterien definiert, welche die Entscheidung für oder gegen eine bestimmte Variante beeinflussen (vgl. [Loshin 2008, 159-176]).

Zwar setzt der Autor am Ende der Veröffentlichung die genannten Aktivitäten grob miteinander in Beziehung und nennt auch Ergebnisse bzw. Ergebnisdokumente (Deliverables) für jede der Aktivitäten. Jedoch bleibt der genaue Zusammenhang, wie die Ergebnisse der einzelnen Aktivitäten aufeinander aufbauen, unklar. Zudem werden die genauen Inhalte der Ergebnisse nicht spezifiziert. Ein Metamodell oder zumindest ein Glossar fehlt gänzlich. Für das Rollenmodell identifiziert der Autor acht verschiedene Interessensgruppen (Shareholder) und beschreibt allgemein deren Aufgabe in einem MDM-Projekt (vgl. [Loshin 2008, 27-31]). Eine direkte Zuordnung zu einzelnen Aktivitäten wird nicht vorgenommen.

3.3.5 Enterprise Master Data Management

DREIBELBIS ET AL. konzentrieren sich in ihrer Arbeit hauptsächlich auf die systemtechnischen Aspekte der unternehmensweiten Stammdatenintegration. Entsprechend liegt der Fokus auf der ausführlichen Beschreibung von Lösungskonzepten für die Komponenten einer Stammdatenintegrationsarchitektur, welche auch die Darstellung von Referenzarchitekturen umfasst (vgl. [Dreibelbis et al. 2008, 93-166]). Die Autoren beschreiben detailliert Architekturvarianten für MDM (insbesondere Hub-Architekturen) und setzen sich mit der Konsistenzwahrung bei der Stammdatenverteilung auf mehrere stammdatenführende Applikationen auseinander. In Bezug auf die Architekturalternativen werden unterschiedliche Komplexitätsgrade diskutiert und die Möglichkeit einer schrittweisen Implementierung aufgezeigt. Die Lösung der Integration auf fachlicher Ebene wird kaum behandelt, obwohl die Autoren die Notwendigkeit einer unternehmensweit einheitlichen Semantik der Geschäftsobjekte sowie definierter Metadatenpflegeprozesse durchaus benennen (vgl. [Dreibelbis et al. 2008, 479]).

Die strukturellen Anforderungen werden nur ungenügend erfüllt. Das Vorgehensmodell wird lediglich in Projektphasen untergliedert, einzelne Aktivitäten werden nur teilweise beschrieben und in keine sachlogische Beziehung oder eine Reihenfolge ge-

bracht. In diesem Zusammenhang fehlen auch notwendige Ergebnisse bzw. Ergebnisdokumente vollständig. Die Autoren spezifizieren im Anhang ein Rollenmodell, indem sie einzelne Rollen identifizieren, deren Aufgaben beschreiben und sie den Projektphasen (nicht aber einzelnen Aktivitäten) zuordnen (vgl. [Dreibelbis et al. 2008, 509-532]). Aufgrund dieser Mängel erscheint die Arbeit als Grundlage für ein systematisches Vorgehen ungeeignet, liefert speziell für die Gestaltung und Entwicklung von Integrationsarchitekturen jedoch hilfreiche Handlungsempfehlungen und Techniken.

3.3.6 Unternehmensweite Datenintegration

HEINE entwickelt aufbauend auf dem Konzept der modularen Datenlogistik ein Vorgehen für die unternehmensweite Datenintegration in verteilten IS. Das Konzept bezieht sich nicht ausschliesslich auf Stammdaten, jedoch auf Daten, die über stammdatenähnliche Charakteristika (insbesondere die unternehmensweite Nutzung) verfügen. Wesentlicher Bestandteil ist ein objektorientierter Modellierungsansatz für dezentrale IS mit daten-, funktions- und ablaufbezogenen Modellsichten, die sich an ARIS orientieren. Der Autor erläutert mit dem Ziel, uniforme Datenstrukturen und Datensemantik zu schaffen, die Notwendigkeit einer terminologischen Vereinheitlichung auf fachkonzeptioneller Ebene (vgl. [Heine 1999, 192-216]). Die systematische Speicherung und Pflege von Metadaten der Datenlogistik wird in einer separaten Repository-Modellebene festgelegt (vgl. [Heine 1999, 163-166]). Aus den verallgemeinerten Metadaten von Stamm- und Bewegungsdaten wird ein sogenanntes Unimu-Schema abgeleitet, das die Grundlage für eine integrierte Datenbank ist. Hierfür beschreibt der Autor konkrete Modellierungstechniken und Integrationsansätze. Allerdings fokussiert das geschilderte Vorgehen sehr stark auf analytische Anwendungsszenarien (vertikale Datenintegration zur Unterstützung von Management Support Systemen) und somit auf Data Warehouse-Lösungen als zentrale Datenlogistikkomponente sowie Interaktionsbeziehungen zwischen Data Marts. Auf die Lösung der Heterogenitätskonflikte zwischen Daten geht der Autor kaum ein.

Im Anhang seiner Arbeit zeigt der Autor ein beispielhaftes Vorgehen bei der Konzeption einer integrierten Datenlogistik anhand einer Fallstudie in einem Energieversorgungsunternehmen auf (vgl. [Heine 1999, 312]). Das Vorgehensmodell ist vereinfacht als Ereignisgesteuerte Prozesskette (EPK) mit Funktionen (Aufgaben) und Ereignissen, die wichtigen Teilergebnissen entsprechen, dargestellt. Obwohl er nicht explizit darauf verweist, postuliert HEINE, dass das gewählte Vorgehen auch über die Fallstudie hinaus Gültigkeit besitzt. Die in der EPK modellierten Teilergebnisse können teilweise als Ergebnisdokumente im Sinne des ME verstanden werden, der Autor spezifiziert deren Inhalt und Struktur jedoch nicht näher. Zudem sind die Aktivitäten sehr generisch formuliert und daher wenig aussagekräftig. In Bezug auf die notwendigen

Methodenelemente gemäss ME ist der Ansatz zur unternehmensweiten Datenintegration als unvollständig zu bewerten, da Metamodell und Rollenmodell vollständig fehlen.

3.3.7 Bewertung der Ansätze

Der Vergleich der in den vorherigen Kapiteln untersuchten Ansätze ist in Tabelle 3-9 zusammengefasst. Die Tabelle enthält qualitative Bewertungen anhand kreisförmiger Ideogramme (sogenannte Harvey Balls), bei denen zwischen drei Erfüllungsgraden unterschieden werden kann: von einem vollständig ausgefüllten Kreis, der die umfassende Erfüllung des jeweiligen Kriteriums symbolisiert, bis zu einem nicht ausgefüllten Kreis, der anzeigt, dass das Kriterium nicht erfüllt ist. Die Kriterien entsprechen den in Kapitel 3.3 erläuterten Anforderungen an eine Methode zur Stammdatenintegration.

	Reference and MDM [DAMA 2009]	MDM and CDI for a Global Enterprise [Berson/Dubov 2007]	Master Data Management [Loshin 2008]	Enterprise MDM [Dreibelbis et al. 2008]	Unternehmensweite Datenintegration [Heine 1999]
Inhaltliche Charakteristika					
Stammdaten	◐	●	●	●	◐
Eindeutige Definition	◐	◐	●	○	○
Metadatenmanagement	◐	◐	●	◐	◐
Semantisches Datenmodell	◐	◐	◐	○	◐
Integrationsarchitektur	●	◐	●	●	◐
Werkzeugunterstützung	◐	●	○	◐	○
Strukturelle Charakteristika					
Metamodell	◐	○	○	○	○
Vorgehensmodell	●	○	◐	◐	◐
Techniken	○	◐	◐	◐	◐
Entwurfsergebnisse	◐	○	◐	○	◐
Rollenmodell	●	◐	◐	●	○

Kriterium vollständig erfüllt
 Kriterium teilweise erfüllt
 Kriterium nicht erfüllt

Tabelle 3-9: Vergleich der beschriebenen Ansätze

Aus Tabelle 3-9 lassen sich mehrere, für die weitere Arbeit wichtige Erkenntnisse gewinnen. Als erstes ist festzustellen, dass keiner der untersuchten Ansätze sämtliche Kriterien (sowohl inhaltlich als auch strukturell) erfüllt. Wie zu Beginn des Kapitels 3.3 argumentiert, rechtfertigt dies die Entwicklung einer Methode zur Stammdatenin-

tegration, welche die Kriterien besser bzw. umfassender erfüllt und die bestehende Forschungslücke somit schliesst.

Inhaltlich liegt ein Schwerpunkt der untersuchten Ansätze auf der informationstechnischen Umsetzung der Stammdatenintegration, indem Probleme der Stammdatenlogistik und der Konsistenz von Stammdaten in verteilten Applikationsarchitekturen behandelt werden. Daher bieten die Ansätze am ehesten Konzepte für die Entwicklung von Integrationsarchitekturen. Unzureichend wird bei fast allen Arbeiten das Problem der Stammdatenintegration auf fachkonzeptioneller Ebene behandelt. Es fehlen Vorschläge und Handlungsempfehlungen, wie ein einheitliches Verständnis über wesentliche Geschäftsobjekte durch eindeutige Definitionen erzielt und die dadurch explizierte Semantik für die Ableitung von Datenmodellen genutzt werden kann. Ausserdem wird nicht deutlich, wie eine Organisation die Metadaten zu den Geschäfts- und Datenobjekten auf Dauer konsistent pflegen und sie dem Nutzer in seinen Datenmanagementprozessen einfach zur Verfügung stellen kann. Hierbei sticht die Arbeit von LOSHIN heraus, da sie diese inhaltlichen Anforderungen an eine Methode zur Stammdatenintegration am umfassendsten erfüllt.

Die vergleichende Bewertung zeigt zudem, dass die Ansätze vor allem in Bezug auf die methodischen Anforderungen (strukturelle Charakteristika) Mängel aufweisen. Sämtliche Ansätze spezifizieren ein Metamodell sowie die notwendigen Ergebnisdokumente nur unzureichend oder überhaupt nicht. Ausserdem fällt auf, dass viele der Arbeiten (insbesondere die von DREIBELBIS ET AL. sowie BERSON und DUBOV) lediglich einzelne Problemstellungen bei der Stammdatenintegration und entsprechende Lösungskonzepte beschreiben, diese jedoch nicht in eine sachlogische Abfolge im Sinne eines Vorgehensmodells bringen. Dadurch erfüllen sie nicht die Anforderung an ein systematisches Vorgehen, wie es Ziel der Methodenentwicklung im Business Engineering ist. Besonders auffallend an der Auswahl der Ansätze sowie an der zugrunde liegenden Auswertung der verfügbaren Literatur zum Thema ist die geringe Anzahl an wissenschaftlichen Veröffentlichungen zum Thema Stammdatenmanagement und -integration. Bis auf die Dissertation von HEINE handelt es sich ausschliesslich um Praktikerpublikationen. Dies unterstreicht zusätzlich, dass hier eine Forschungslücke besteht und methodische Ansätze zu diesem in der Praxis bedeutenden Problem fehlen.

Aufgrund der beschriebenen Schwächen wird in Kapitel 5 eine eigene Methode vorgeschlagen, welche die Stärken der zuvor bewerteten Ansätze berücksichtigt, aber die Mängel insbesondere im Hinblick auf die strukturellen Anforderungen (Metamodell, Vorgehensmodell, Techniken) und die Stammdatenintegration auf fachlicher Ebene behebt. Auf diese Punkte wird im Zuge der Methodenentwicklung im Besonderen eingegangen.

4 Anforderungen aus der Praxis an die Integration von Stammdaten

Neben den im vorherigen Kapitel untersuchten theoretischen Beiträgen fließen in die vorliegende Arbeit Erkenntnisse aus der Unternehmenspraxis ein. Diese Erkenntnisse resultieren aus drei Fallstudien und zwei Aktionsforschungsprojekten¹⁷. Aus der Literaturanalyse sowie den Fallstudien werden Anforderungen an die zu entwickelnde Methode abgeleitet. Zudem geben die Fallstudien Hinweise für Handlungsoptionen und für Massnahmen zur Integration von Stammdaten, so dass ein erster Entwurf der Methode respektive der Methodenfragmente erarbeitet werden konnte. Die Aktionsforschungsprojekte dienen der Anwendung der Methodenfragmente sowie ihrer iterativen Verfeinerung bzw. Verbesserung.

Nach einer einführenden Charakterisierung der Fallstudien und Aktionsforschungsprojekte in Kapitel 4.1 stellt Kapitel 4.2 die drei Fallstudien in gekürzter Form mit den für die Methodenentwicklung bedeutsamen Inhalten vor. Das Kapitel schliesst mit einer Synthese der Erkenntnisse aus den Fallstudien insbesondere in Bezug auf das Vorgehensmodell der Methode. Die Aktionsforschungsprojekte stellt Kapitel 4.3 mit Fokus auf die Präsentation des Unternehmens und der Geschäftstreiber für die Stammdateninitiative vor. Die Anwendung der Methode im Rahmen der Aktionsforschungsprojekte und die daraus resultierenden Erkenntnisse behandelt Kapitel 6.1 ausführlicher. In Kapitel 4.4 wird der Beitrag der Fallstudien und Aktionsforschungsprojekte für die anschließende Methodenentwicklung erläutert.

4.1 Auswahl der Fallstudien und Aktionsforschungsprojekte

Die Auswahl von Fallstudien beruht auf vordefinierten Kriterien, die den Untersuchungsgegenstand abbilden und die Definition der Grenzen der Ergebnisse erleichtern sollen [Eisenhardt 1989, 536f.; Perry 1998, 792f.]. Die drei Fallstudien SBB Cargo AG, Deutsche Telekom AG und Bosch Rexroth beschreiben Transformationsprojekte in Unternehmen, deren Schwerpunkt auf der unternehmensweiten Harmonisierung und Integration von Stammdaten lag. Folglich sollten sie zumindest einen Teil der in Tabelle 3-5 definierten inhaltlichen Charakteristika abdecken. Neben dieser inhaltlichen Übereinstimmung waren folgende weitere Kriterien für die Auswahl der Fallstudien von Bedeutung:

- Hinreichende Komplexität des Unternehmens, die sich u. a. in der Unternehmensgrösse, der globalen Tätigkeit, einer Vielzahl von Geschäftsbereichen und Applikationen manifestiert,

¹⁷ Die beiden forschungsmethodischen Ansätze erklärt Kapitel 1.3.

- Langjährige Erfahrung des Unternehmens mit Stammdatenintegrationsprojekten,
- Bereitschaft der Unternehmen zur Aufnahme und Veröffentlichung der Fallstudie.

Die Inhalte der Fallstudien entstammen Interviews mit Verantwortlichen der Transformationsprojekte, Mitgliedern der Projektteams und Vertretern der DQM-Abteilungen in den Unternehmen (siehe Anhang A.1) sowie der Auswertung von Projektunterlagen. Das Vorgehen zur Erstellung der Fallstudien sowie deren Struktur lehnen sich an die Fallstudienmethode für das Business Engineering PROMET BECS an (vgl. [Senger/Österle 2004]). Die Beschreibung ist nach folgenden vier Punkten strukturiert:

- *Unternehmen* stellt die untersuchte Organisation mit ihren Eckdaten vor.
- *Ausgangslage* beschreibt die Situation im Unternehmen zu Beginn des Stammdatenintegrationsprojektes sowie die Treiber zur Initiierung des Projektes.
- *Umsetzung* erläutert die verfolgten Ziele, das gewählte Vorgehen mit getroffenen Massnahmen sowie die neue Lösung im Unternehmen. Die Beschreibung konzentriert sich dabei auf die in Kapitel 3.3 genannten, inhaltlichen Anforderungen.
- *Bisherige Erfahrungen* diskutiert die bisher erreichten Ziele sowie Aufwand und Nutzen und zeigt Überlegungen des Unternehmens zur Weiterentwicklung der Integrationslösung auf.

Die beiden Aktionsforschungsprojekte wurden im Rahmen des CC CDQ von November 2006 bis Juni 2009 durchgeführt. Sowohl bei Daimler als auch bei der DB Netz war die Harmonisierung der Stammdaten als Grundlage für eine konsistente Stammdatenbasis im Unternehmen das wesentliche Handlungsfeld der jeweiligen MDM- bzw. DQM-Initiative. Die Methode unterstützte die Unternehmensvertreter bei der Bearbeitung dieses Handlungsfeldes, indem es ein strukturiertes Vorgehen mit konkreten Techniken zur Verfügung stellte. Aus der Anwendung der Methode konnten zudem Rückschlüsse auf die Anpassung und Verbesserung der Methode gewonnen werden.

4.2 Fallstudien

4.2.1 SBB Cargo AG¹⁸

4.2.1.1 Unternehmen

Die Schweizerische Bundesbahnen Cargo AG (SBB Cargo) mit Sitz in Basel (Schweiz) ist mit über 4.200 Mitarbeitern und einer jährlichen Verkehrsleistung von ca. 12,5 Milliarden Nettotonnenkilometern der marktführende Dienstleister im Schie-

¹⁸ Für die vollständige Fassung der Fallstudie siehe [Schmidt 2009b].

nengüterverkehr der Schweiz. Zudem ist das Unternehmen die Nummer zwei im strategisch wichtigen, alpenquerenden Transport von Gütern von den Nordseehäfen und Deutschland nach Italien, der sogenannten Nord-Süd-Achse. Als vollständiges Tochterunternehmen der Schweizerischen Bundesbahnen (SBB) ist SBB Cargo im Besitz der Schweizerischen Eidgenossenschaft und hat als solches den Auftrag, im Rahmen der Eigenwirtschaftlichkeit ein flächendeckendes Netz für den Wagenladungsverkehr zu betreiben. Die SBB Cargo bietet ganzheitliche Transportlösungen im nationalen und internationalen Wagenladungsverkehr (inklusive Import, Export und Transit), im Ganzzugsverkehr und im kombinierten Verkehr sowie bahnahe Zusatzdienstleistungen (z. B. Rangieren, Flottenmanagement, Beladung) für andere Bahnen an.

	SBB Cargo AG
Gründung	Gründung der SBB am 01. Januar 1902 Gründung der Division Güterverkehr zum 01. Januar 1999; Umwandlung zur SBB Cargo AG am 01. Januar 2001
Firmensitz	Basel, Schweiz
Branche	Logistikdienstleister im Schienengüterverkehr
Firmenstruktur	Aktiengesellschaft mit drei Tochterunternehmen: ChemOil Logistics sowie SBB Cargo Deutschland und SBB Cargo Italien, für die es jeweils eine eigenständige Produktionsgesellschaft und Verkaufagentur gibt
Geschäftsbereiche	Transportgeschäft Schweiz, Transportgeschäft International, Asset Management / Instandhaltung (Service Rollmaterial)
Umsatz (2008)	CHF 1.044,2 Mio.; Verkehrsleistung: 12,53 Mrd. Nettotonnenkilometer
Ergebnis (2008)	CHF -29,9 Mio.
Mitarbeiter (2008)	4.248
Homepage	http://www.sbbcargo.com

Tabelle 4-1: Kurzporträt SBB Cargo AG (vgl. [SBB Cargo 2009])

4.2.1.2 Ausgangslage

Der mit der Reorganisation (aufgrund der Ausrichtung auf das internationale Transitgeschäft im Zuge der Marktöffnung) einhergehende Zeit- und Kostendruck führte zu punktuellen, prozessverbessernden Lösungen, welche jedoch mit zunehmenden Redundanzen und unterschiedlichen Verständnissen von Geschäftsobjekttypen einhergingen. Bei der SBB Cargo gab es bis 2008 keine eindeutig definierte, übergeordnete Verantwortlichkeit im Sinne einer Organisationseinheit innerhalb der bestehenden Organisationsstrukturen, die das Stammdatenmanagement aus Unternehmenssicht steuerte. Die Qualität der Stammdaten wurde entsprechend den jeweiligen Anforderungen projektspezifisch gewährleistet, ohne dass eine koordinierende Instanz projektübergreifend die konsistente Verwendung von Stammdaten sicherstellte bzw. optimierte.

Die fehlende Koordination von Aktivitäten hatte zur Folge, dass in neuen Projekten nicht auf bestehende Funktionalitäten und Daten zurückgegriffen wurde, da dies in der Regel einen zusätzlichen Integrationsaufwand bedeutete. Einerseits wurde die beste-

hende Applikationsarchitektur zunehmend intransparent, da innerhalb einzelner Projekte eigene Applikationen (Individualsoftware) entwickelt wurden, die projektübergreifend verwendete Daten speichern. Diese wurden dann eher ad-hoc in die bestehende Applikationsarchitektur eingebunden, wodurch zahlreiche, grösstenteils unkoordinierte Punkt-zu-Punkt-Verbindungen entstanden. Andererseits wurde bei der Datennutzung und -speicherung nicht auf die konforme Verwendung zum bereits bestehenden Stammdatenbestand geachtet: Stammdatenobjekte wurden mit eigenen Attributen neu definiert, Schlüsselattribute zur Identifikation neu vergeben. Dadurch konnte nicht mehr sichergestellt werden, dass ein Geschäftsobjekt über Applikations- und Projektgrenzen hinweg einheitlich in den Applikationen abgebildet wurde. Dies führte unter anderem dazu, dass für die Planung eine andere Stammdatenbasis (eigene Applikation) verwendet wurde als für die Auswertung der tatsächlich erbrachten Leistung und betriebliche und kommerzielle Daten auch weiterhin in separaten Applikationen gepflegt werden. Zudem war nicht mehr ersichtlich, welche Applikation für ein bestimmtes Stammdatenobjekt führend ist.

Die beschriebenen Probleme schlugen sich vor allem im Reporting und Controlling nieder, für das eine Konsolidierung über den gesamten Wertschöpfungsprozess hinweg notwendig ist. Die redundante Speicherung von Stammdaten sowie das fehlende Wissen über den Ort ihrer Speicherung erschwerte eine Konsolidierung. So konnten beispielsweise Frachtzahler nicht eindeutig den Geschäftsbereichen (Sparten) zugeordnet werden. Die falsche Zuordnung dieser Stammdaten führte zur Berechnung unkorrekter Spartenergebnisse und wirkte sich negativ auf die Qualität der Finanzreports aus. Im Vertriebsbereich schlug sich die unzureichende Datenqualität primär in Kundenreklamationen nieder.

Die wesentlichen Probleme in Bezug auf die Stammdaten lassen sich wie folgt zusammenfassen:

- Fehlendes einheitliches Verständnis, da Begriffe, Geschäftsobjekttypen und deren Strukturen nicht eindeutig definiert waren,
- fehlende Definition von unternehmensweit gültigen Pflege- und Änderungsprozessen,
- fehlende Festlegung der Datenverantwortung (z. B. durch Bestimmung eines Daten-Owners),
- fehlende Transparenz über Entstehung und Verwendung von Stammdaten in den Geschäftsprozessen sowie über deren Herkunft (führendes System).

4.2.1.3 Umsetzung

Projektziele

Aufgrund der Probleme beim Reporting sowie der Planung war der Leidensdruck innerhalb des Finanzbereiches der SBB Cargo am deutlichsten. Als konkreter Auslöser sowie Trägerprojekt für die Stammdateninitiative diente daher das FITS-Projekt (Finanzielle Transparenz und Steuerung), dessen Auftraggeber und Sponsor der Leiter des Finanzbereiches war. Als eines von acht Teilprojekten des FITS startete das Projekt Datenmanagement im Juli 2008 mit einer (vorläufigen) Laufzeit von einem Jahr.

Das Teilprojekt Datenmanagement verfolgte im Wesentlichen vier Ziele:

- Identifikation und eindeutige Definition der Geschäftsobjekttypen des Unternehmens (Geschäftsdatenmodell), Festlegung eines Daten-Owneers und eines Master-Systems für einen Datenobjekttyp sowie die Dokumentation der Verwendung in Geschäftsprozessen und Applikationen,
- Definition eines einheitlichen Vorgehens beim Stammdatenmanagement innerhalb einzelner Projekte, das gewährleistet, dass das einmal definierte Geschäftsdatenmodell in allen Projekten gleichermassen verwendet und konsistent weiterentwickelt wird,
- Etablierung unternehmensweit standardisierter Pflegeprozesse (inklusive Unterstützung durch ein Workflow-Tool),
- Einführung einer Stammdatenschicht, die Daten für Planung und Reporting aus verschiedenen Applikationen und Prozessen konsistent zusammenführt.

Stammdatenarchitektur

Die Stammdatenarchitektur soll die Kerngeschäftsobjekte des Unternehmens aus Sicht der Geschäftsprozesse sowohl als graphisches Modell (Geschäftsdatenmodell) als auch in Form textueller Beschreibungen (Glossare) abbilden und somit eine Grundlage für ein unternehmensweit einheitliches Verständnis bilden.

Zu Beginn der Dokumentation der Stammdatenarchitektur wurde deutlich, dass eine ausschliessliche Beschränkung auf die Stammdaten dem Ziel einer prozessorientierten Dokumentation nur ungenügend gerecht wird, da sie im Geschäftsprozess selten und eher indirekt verändert werden. Der Betrachtungsumfang wurde daher allgemein auf Geschäftsdaten erweitert und eine Dreiteilung der Datenarten vorgenommen, um eine übersichtlichere Darstellung der Objekte und ihrer Beziehungen zu ermöglichen. Folgende Datenarten wurden definiert:

- *Geschäftsvorfalldaten* (Bewegungsdaten gemäss Kapitel 2.2.1) sind Daten, welche direkt im Laufe eines Wertschöpfungsprozesses entstehen, verwendet, verändert

und gelöscht werden. Sie stehen in Beziehung zu anderen Geschäftsvorfalldaten und nutzen Stammdaten sowie Kerndatenelemente.

- *Stammdatenobjekte* werden in den Geschäftsvorfällen verwendet, durch die Geschäftsprozesse aber nicht direkt verändert. Sie bleiben über einen längeren Zeitraum konstant und stehen mit anderen Stammdatenobjekten sowie Kerndatenelementen in Beziehung.
- *Kerndatenelemente* entsprechen am ehesten dem in der wissenschaftlichen Literatur verwendeten Konzept der Referenzdaten (siehe Kapitel 2.2.1), da deren Pflege zum Teil nicht durch das Unternehmen verantwortet wird, sondern im Zuständigkeitsbereich von Organisationen wie dem Internationalen Eisenbahnverband liegt. Sie dienen in der Regel der Klassifikation von Stamm- und Geschäftsvorfalldaten.

Zusätzlich wurde im Laufe der Dokumentation eine vierte Datengruppe gebildet (sogenannte Glossar-begriffe), in der Synonyme zu Objekten dokumentiert werden.

Für die Dokumentation der Geschäftsdaten (bzw. ihrer Geschäftsobjekttypen) wurde ein zweigleisiges Vorgehen gewählt. Als erstes wird jedes Geschäftsvorfalldatum, Stammdatenobjekt, Kerndatenelement sowie jeder Glossar-begriff auf Typebene einheitlich in einem Glossar gepflegt. Zielgruppe des Glossars ist der Fachbereich, dem es über das Intranet zur Verfügung gestellt wird, um damit eine unternehmensweit einheitliche Verwendung der Begrifflichkeiten zu fördern. Zusätzlich wird den Datennutzern der verantwortliche Daten-Owner als Ansprechpartner angezeigt. Innerhalb des Glossars werden folgende Metadaten geführt: Name, unternehmensweit einheitliche Definition, zugehörige Datenklasse, mögliche Werte bzw. Ausprägungen, wesentliche Attribute (keine vollständige Aufzählung), Daten-Owner, Master-Prozess, Nutzer-Prozesse, Master-System und weitere, nutzende Systeme. Ein Beispiel ist in Tabelle 4-2 angeführt.

Name	Beschreibung	Schlüsselbegriff	Business	Master-Prozess (Produziert)	Nutzer-Prozess (Konsumiert)	Master-System	Nutzer-System	Vorschlag Business Owner
Wagentyp	Typisierung / Codierung unterschiedlicher Waren. Die Wagentypen beschreiben die Bauart und die Eigenschaften des Fahrzeugs. Die Codierung ist durch die UIC festgelegt. Beispiele: Eaos, Habbiilnss, Shimms Wagentyp ist eine Teilmenge von Fahrzeug.	KE	Instandhaltung	K4.2 Planung (Produziert)	K1.2 Angebotsplanung (Konsumiert) K2.2 Machbarkeit (Konsumiert) K2.3 Offerte (Konsumiert) K3 Leistung (Konsumiert) K4.1 Life Cycle Management (Konsumiert) K4.3 Bereitstellung (Konsumiert)	SAP/ERP		Mathias Bütikofer

Tabelle 4-2: Ausschnitt aus dem Glossar der SBB Cargo (Geschäftsobjekttyp Wagentyp)

Zusätzlich wird für jedes Geschäftsvorfalldatum ein UML-Klassendiagramm erstellt, welches den Geschäftsobjekttyp mit seinen Schlüsselattributen sowie seinen Beziehungen zu sämtlichen Geschäftsobjekttypen der Stammdaten und Kerndatenelemente zeigt. Das Geschäftsvorfalldatum stellt somit den zentralen Ankerpunkt eines jeden Geschäftsdatenmodells dar. Zum Zeitpunkt der Fallstudienaufnahme konnten 11 Geschäftsdatenmodelle in Abstimmung mit Prozess-/Daten-Ownern und Produktverantwortlichen erstellt werden. Eine Modellierung bzw. vollständige Beschreibung der Geschäftsobjekttypen auf Ebene der Attribute wurde in diesem ersten Schritt bewusst nicht durchgeführt, um die Komplexität (insbesondere für den Fachbereich) möglichst gering zu halten. Derzeit werden pro Geschäftsobjekttyp zum besseren Verständnis wenige, ausgewählte Attribute modelliert; die Ausweitung sowie die Festlegung von globalen und lokalen Attributen werden erst im Rahmen der Architekturgestaltung (Stammdatenverteilung) in Zukunft durchgeführt.

Ausgangspunkt für die Identifikation der Geschäftsobjekttypen war die bestehende Dokumentation der Geschäftsprozesse, die im Management-System Cargo sehr detailliert abgelegt ist. Anhand dieser Geschäftsprozesse konnten erste Geschäftsobjekttypen erkannt und die Prozess-Owner kontaktiert werden. Ausgehend von Gesprächen mit den Prozess-Ownern wurden iterativ Informationen zu einzelnen Geschäftsobjekttypen gesammelt und konsolidiert (wobei wiederum neue Geschäftsobjekttypen identifiziert wurden). Hierfür wurden Interviews und Workshops mit Nutzern der Geschäftsobjekttypen durchgeführt, um konzertierte Definitionen zu finden (im Falle abweichender Verständnisse), Nutzersysteme zu identifizieren, den zugehörigen Prozess zu spezifizieren und den Daten-Owner festzulegen, welcher in Zukunft die Aufgabe hat, über mögliche Veränderungen der Metadaten des Geschäftsobjekttyps zu entscheiden. Eine wesentliche Herausforderung besteht in der Festlegung eines eindeutigen Schlüssels zur Identifikation von Geschäfts- bzw. Datenobjekttypen. Da sich die Schlüssel derzeit applikationsspezifisch unterscheiden (z.B. Kundennummer im CRM- und im SAP-System), ist die Bestimmung des Master-Systems notwendig, dessen Schlüssel dann bei einer zukünftigen Zusammenführung der Datenbestände massgeblich ist.

Das Stammdatenmanagement-Projektteam bereitete die Informationen anschliessend auf, pflegte sie in das Glossar ein und veröffentlichte das aktualisierte bzw. erweiterte Glossar im Intranet. Sukzessiv wurden auch erste UML-Klassendiagramme für einzelne Geschäftsobjekttypen der Geschäftsvorfalldaten entlang der Wertschöpfungskette erstellt. Im Falle uneinheitlicher bzw. inkonsistenter Definitionen sowie zur Verfeinerung der Geschäftsdatenmodelle wurden anschliessend iterativ zusätzliche Abstimmungstreffen durchgeführt.

Für die Metadatenpflege, also die konsistente Fortführung der Informationen in Glossar und Geschäftsdatenmodell, nach fachlicher Freigabe der Änderungen durch den jeweiligen Daten-Owner ist das Stammdatenmanagementteam zuständig. Darunter

fällt die Definition neuer Geschäftsobjekttypen, die Anpassung von Datenformaten oder die Änderung von Hierarchien und Merkmalszuordnungen im Geschäftsobjekttyp. Die zentrale Mutationsstelle pflegt nur Änderungen an den Datenobjektinstanzen.

Die Modelle sowie das Glossar werden derzeit im unternehmensinternen Enterprise Architecture Management-Tool gepflegt und in 14-tägigem Rhythmus aktualisiert. In diesem Tool werden in weiteren Schritten auch sämtliche Prozesse und Applikationen (insbesondere der Bebauungsplan) und der Zusammenhang zwischen Daten, Prozessen und Applikationen abgebildet. Das Tool wird dann nicht nur zur Dokumentation verwendet, sondern erlaubt auch die Ableitung von Reports und Auswertungen über die gepflegten Geschäftsdatenmodelle.

Wie Tabelle 4-3 verdeutlicht, wurden insgesamt 303 Geschäftsdatenobjekte auf Typenebene identifiziert, modelliert und beschrieben. Davon sind 62 Geschäftsvorfalldaten (entspricht 20%) und 47 Stammdatenobjekte (16%). Nach Einschätzung des Projektteams entsprechen diese Zahlen 90-95% der geschätzten Gesamtobjektzahl. In den letzten Wochen ist ein signifikantes Wachstum nur noch bei den Glossarbegriffen festzustellen. Ausserdem ist zu betonen, dass bei mehr als 90% der identifizierten Geschäftsobjekttypen jeweils ein Daten-Owner definiert und der Master-Prozess bestimmt werden konnte.

Beschreibung	Absolute Anzahl	Prozentualer Anteil
Ermittelte Geschäftsdatenobjekte	303	100
davon mit Owner	279	92
davon Owner in Abklärung	23	8
Ermittelte Geschäftsdatenobjekte (pro Ebene)	303	100
Geschäftsvorfalldaten	62	20
Stammdatenobjekte	47	16
Kernelementdaten	114	38
Glossarbegriffe	78	25
noch offen	2	1
Geschäftsdatenobjekte mit Master-Prozess	204	93
Geschäftsdatenobjekte mit bekanntem Master-System	121	54

Tabelle 4-3: Übersicht über definierte Geschäftsdatenobjekte (Stand: 29.05.2009)

Rollen und Verantwortlichkeiten

Im Zuge der Identifikation und Beschreibung der Geschäftsobjekttypen wurde ein standardisiertes Rollenmodell entwickelt, welches für jeden Geschäftsobjekttyp ausgeprägt werden kann. Die darin spezifizierten Rollen sind: a) der Besteller oder Auftraggeber, der eine Veränderung an einem Geschäftsobjekttyp auslöst bzw. beantragt; b) die Prüfstelle, welche die Abhängigkeiten von und zu anderen Geschäftsobjekttypen und damit die Richtigkeit von Veränderungen untersucht; c) der Daten-Owner, der für die Freigabe und konsistente Änderungen verantwortlich ist; d) die Mutationsstelle,

welche die Veränderungen am Geschäftsobjekttyp physisch ausführt, und e) die Kommunikationsstelle, die Veränderungen im Unternehmen publiziert. Mit Veränderungen sind hier sämtliche Modifikationen an einem Geschäftsobjekttypen gemeint, inklusive seiner Entstehung und seiner Löschung.

Projektübergreifende Etablierung von Datenmanagementstandards

Parallel zu der im vorherigen Abschnitt beschriebenen Grundlagenarbeit müssen ständig laufende Projekte, die erste Ergebnisse der Stammdateninitiative nutzen, unterstützt werden, um die konsistente Weiterentwicklung der Stammdatenarchitektur sowie der Stammdatenlandkarte (Zuordnung der Stammdatenobjekte zu Applikationen) zu gewährleisten. Zur Erreichung dieses Ziels wurde schon frühzeitig ein einheitlicher Prozess entwickelt (siehe Abbildung 4-1), der zukünftig in allen Cargo-Projekten Berücksichtigung finden soll. Dieser umfasst datenmanagementspezifische Aufgaben und Meilensteine, die in den allgemeinen Projektplan integriert werden.

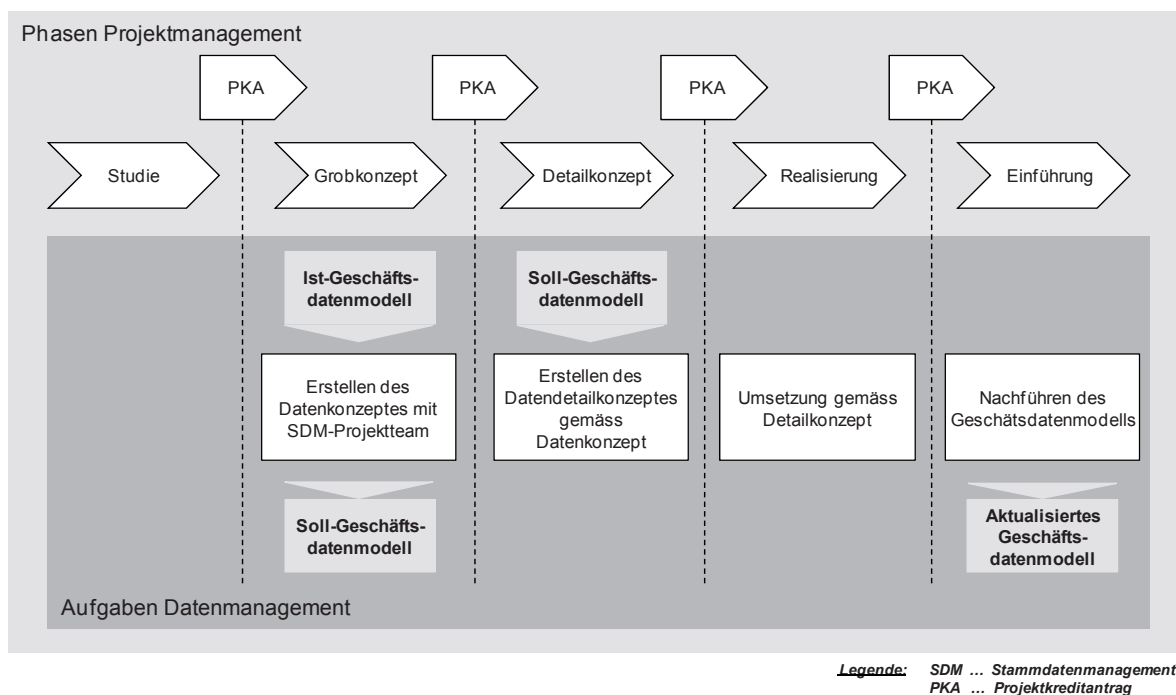


Abbildung 4-1: Einbindung des Stammdatenmanagements in das Projektmanagement

Die Aufgaben des Datenmanagements fokussieren auf die Konsistenzwahrung des Geschäftsdatenmodells. Hierfür muss das Projektmanagement auf Basis des bestehenden Geschäftsdatenmodells frühzeitig ein Datenkonzept erarbeiten und diejenigen Geschäftsobjekttypen identifizieren, welche im Projekt benötigt oder verändert werden. Auf dem Datenkonzept und den darin formulierten Anforderungen aufbauend wird ein konsistentes Soll-Geschäftsdatenmodell in Zusammenarbeit mit dem Stammdaten-Projektteam und den Daten-Ownern erstellt. Zur Etablierung sowie zur Unterstützung der Datenmanagement-Aufgaben in den Projekten wurden Templates erarbeitet, welche die Definition des Datenkonzeptes zu Beginn eines jeden Projektes vereinfachen.

Integrationsarchitektur für Stammdaten

Die Applikationsarchitektur hat sich in Bezug auf das Stammdatenmanagement bis zum derzeitigen Zeitpunkt nicht verändert. Die zielgerichtete (Um-)Gestaltung der Integrationsarchitektur wurde als wichtiges Thema für das Stammdatenmanagement identifiziert, jedoch bewusst nach hinten verschoben. In welcher Form (eigenes Stammdatenprojekt oder im Zuge der konzernweiten Integrationsarchitektur) dieses aufgrund zahlreicher Abhängigkeiten schwierige Handlungsfeld bearbeitet wird, ist derzeit noch nicht geklärt. Voraussetzung hierfür ist die eindeutige Definition und Modellierung der Datenobjekttypen sowie die Identifikation und Dokumentation der entsprechenden Master-Systeme, welche führend für sie verantwortlich sind. Zudem muss festgelegt werden, welche Attribute unternehmensweite Gültigkeit besitzen und welche lediglich über lokale Gültigkeit (z. B. pro Applikation) verfügen sollen. Langfristiges Ziel ist es, die Vielzahl von stammdatenführenden Applikationen auf maximal zwei zentrale Applikationen zu reduzieren, in denen die Stammdaten zentral gepflegt und auf angebundene Applikationen verteilt werden.

4.2.1.4 Bisherige Erfahrungen

Im Vergleich zu früheren Massnahmen zur Erhöhung der Stammdatenqualität stellen die durchgängige Ausrichtung an den Anforderungen und die Einbindung des Fachbereiches entscheidende Erfolgsfaktoren bei der Entwicklung der Stammdatenarchitektur dar. Dementsprechend liegt der Fokus auf der Etablierung eines einheitlichen Begriffsverständnisses als Grundlage für die eindeutige Identifikation und konsistente Verwendung von Stammdatenobjekten. In diesem Sinne wird die Modellierung mit dem Glossar und einer geringen Komplexität der Geschäftsdatenmodelle so gewählt, dass die Nutzung durch den Fachbereich gefördert wird. Erfolgsentscheidend war ausserdem die Orientierung an den Geschäftsprozessen der SBB Cargo, um die funktions- und applikationsspezifischen Sichten auf die Geschäftsobjekttypen zu integrieren und sicherzustellen, dass Mitarbeiter, die Daten verändern, sich der Konsequenzen ihrer Änderungen für Datennutzer zu einem späteren Zeitpunkt im Prozess bewusst sind. Die Prozessorientierung führte zu der Erkenntnis, dass eine ausschliessliche Fokussierung auf Stammdatenobjekte unzureichend ist, um die Datenqualität im Unternehmen zu verbessern. Schlüsselthemen bei der Grundlagenarbeit sind die Festlegung der Daten-Owner (häufig fühlte sich keiner für ein bestimmtes Objekt verantwortlich), der verwendeten Geschäftsprozesse und des Master-Systems.

Aufwand und Nutzen

Vor Beginn des Projektes wurde primär eine qualitative Nutzenbetrachtung durchgeführt, jedoch wurden die Kosten und der erwartete Nutzen in einzelnen Teilbereichen im Sinne einer vereinfachten Business Case-Rechnung grob geschätzt. Der quantifi-

zierte Nutzen wurde auf der Grundlage der reduzierten Anzahl von nicht abrechenbaren Sendungen¹⁹ und des damit einhergehenden, geringeren Zinsverlusts sowie der Reduktion der administrativen Kosten (z. B. durch weniger Nacharbeit) berechnet. Die Verringerung der Administrationskosten wurde auf 7,3 Vollzeitstellen pro Jahr geschätzt. Durch die Verringerung der Anzahl nicht bewertbarer bzw. fakturierbarer Sendungen als Ergebnis der Stammdateninitiative wurde ein Einsparpotenzial für den Zinsverlust von 50% prognostiziert, was ca. 1,8% des monetären Wertes der durchschnittlich nicht abrechenbaren Sendungen entspricht. Zusätzlich werden weitere, qualitative Nutzenpotenziale aufgezählt. Hierzu zählen u. a. die Verminderung manueller Aufwände zur Pflege bzw. Fehlerkorrektur, die Verringerung der benötigten Zeit zur Erstellung finanzieller Reports oder die Reduktion der Anzahl fehlerhafter Reports durch eine höhere Transparenz über die Herkunft der verwendeten Daten.

Obwohl es derzeit noch verfrüht ist, finale quantitative Nutzenbetrachtungen abzuleiten (auch weil Investitionen in das Stammdatenmanagement keine direkten monetären Rückflüsse generieren), lassen mehrere positive Entwicklungen den Schluss zu, dass die erste Projektphase als erfolgreich bewertet werden kann. Die Expertengespräche mit Vertretern der Fachbereiche zeigten, dass insbesondere der Finanzbereich den langfristigen Nutzen positiv bewertet und sich viel von den Ergebnissen verspricht, auch wenn das Stammdatenprojekt für sie im Moment eher mit Aufwand verbunden ist. Auch die Weiterführung der Stammdateninitiative ist ein deutliches Zeichen dafür, dass der Nutzen der Arbeit im Unternehmen erkannt wird.

Weiterentwicklung

Die zukünftigen Herausforderungen liegen im Aufbau einer zentralen Mutationsstelle sowie der Optimierung der Integrationsarchitektur (Zentralisierung der Stammdatenhaltung) zur Erhöhung der Stammdatenqualität. Für die Fertigstellung des Geschäftsdatenmodells wird ein Schwerpunkt (auch im Hinblick auf die Gestaltung der Integrationsarchitektur) auf der Identifikation der Master-Systeme liegen, die für fast die Hälfte der Datenobjekttypen noch nicht durchgeführt wurde (siehe Tabelle 4-3). Ausserdem sollen bis Ende 2010 sämtliche Stammdatenpflegeprozesse definiert werden.

Eine weitere wichtige Aufgabe wird der Aufbau eines kennzahlenbasierten Reportings für die Messung der Stammdatenqualität sein. Diesbezüglich wird es in einem ersten Schritt notwendig sein, aussagefähige Messgrössen und Key Performance Indicators zu identifizieren, um ein Controlling von Datenqualität (Vollständigkeit, Korrektheit, Aktualität, Verfügbarkeit) und Pflegekosten zu etablieren. Dies ist wichtig, um positive Auswirkungen der eigenen Arbeit aufzuzeigen und damit zusätzliche Argumente für eine unternehmensweite Etablierung des Stammdatenmanagements zu haben.

¹⁹ Nicht abrechenbare Sendungen sind Aufträge, welche die SBB Cargo an den Kunden liefert, obwohl sie mangels vollständiger Vertragsdaten oder falscher Angaben der Kunden nicht bewertbar sind und somit nicht fakturiert werden können.

Eine naheliegende Ausweitung des Projektumfangs auf den gesamten SBB-Konzern ist derzeit nicht geplant und wird aufgrund der Heterogenität der einzelnen Konzernbereiche (Personenverkehr, Infrastruktur, Cargo) kritisch bewertet. Allein das Verständnis über einen Kunden unterscheidet sich signifikant zwischen SBB Cargo und SBB Personenverkehr. Sinnvoll erscheint dies lediglich bei gemeinsam genutzten Geschäftsdatenobjekten zwischen den Geschäftsbereichen Cargo und Infrastruktur, wie z. B. Trassen oder Bahnhöfen. Trotz alledem wird das Stammdatenmanagement auch in den anderen Konzernbereichen in naher Zukunft ein Schlüsselthema sein, weshalb der Transfer von Erkenntnissen und Erfahrungen angestrebt wird.

4.2.2 Deutsche Telekom AG²⁰

4.2.2.1 Unternehmen

Die Deutsche Telekom AG (DTAG) mit Konzernsitz in Bonn (Deutschland) beschäftigt an ihren Standorten in rund 50 Ländern weltweit über 260.000 Mitarbeiter. Sie gilt als eines der weltweit führenden Dienstleistungsunternehmen in der Telekommunikations- und Informationstechnologie-Branche. Im Geschäftsjahr 2008 erwirtschaftete sie einen Umsatz von 61,7 Mrd. Euro und ein Betriebsergebnis von 7 Mrd. Euro.

	Deutsche Telekom AG
Gründung	1995 mit Privatisierung des Telekommunikationsbereiches der Deutschen Bundespost
Firmensitz	Bonn, Deutschland
Branche	Telekommunikation
Firmenstruktur	Fünf operative Segmente: Mobilfunk Europa, Mobilfunk USA, Breitband/Festnetz, Geschäftskunden, Konzernzentrale & Shared Services
Geschäftsbereiche	Seit 2007 zwei wesentliche Geschäftsbereiche im Privatkundenbereich: T-Home (hervorgegangen aus der vormaligen T-Com und T-Online) Geschäftskundenbereich: T-Systems
Umsatz (2008)	61,7 Mrd. Euro
Ergebnis (2008)	7 Mrd. Euro (EBIT)
Mitarbeiter (Juni 2009)	260.000
Homepage	http://www.telekom.com

Tabelle 4-4: Kurzporträt Deutsche Telekom AG (vgl. [Deutsche Telekom 2009])

Als international ausgerichteter Konzern bietet die DTAG unter der konzernweiten Unternehmensmarke „T“ für ihre zahlreichen Privat- und Geschäftskunden vermehrt ein integriertes Produktportfolio an. Der Geschäftsbereich T-Home bietet klassische Festnetzanschlüsse, Breitband-Internet sowie Kommunikations- und Multimedia-Services für Privat- und Geschäftskunden an. Allein in Deutschland verzeichnet der Konzern ca. 28,5 Mio. Festnetzanschlüsse und rund 13,3 Mio. DSL-Anschlüsse. In der

²⁰ Für die vollständige Fassung der Fallstudie siehe [Schmidt et al. 2010a].

Mobilfunksparte (T-Mobile) telefonieren aktuell rund 128 Mio. Kunden in Europa und den USA. Dies macht T-Mobile zu einem der grössten Mobilfunkanbieter weltweit.

4.2.2.2 Ausgangslage

Ein systematisches, auf Nachhaltigkeit ausgerichtetes Management der Stammdaten war vor 2006 weder unternehmensweit noch in den einzelnen Geschäftsfeldern (T-Com, T-Online) etabliert. Das DQM beschränkte sich in dieser Zeit auf gelegentliche Konsistenz- und Vollständigkeitsmessungen des Datenbestandes für einzelne Applikationen sowie Ad-hoc-Datenbereinigungen bei der Einführung neuer Software-Komponenten oder neuer Produkte im Unternehmen. Dies waren jedoch primär bedarfsgetriebene Massnahmen, bei denen fachliche Anforderungen an die Datenqualität kaum Berücksichtigung fanden. Zudem fehlte es an nachhaltigen Konzepten zur präventiven, unternehmensweit abgestimmten Verbesserung von Datenqualität.

Neben dem allgemeinen Leidensdruck, der beispielsweise aus der Komplexität der zu pflegenden Netzinfrastrukturdaten resultiert, und der Anforderung, diese stets aktuell zu halten, wurde ein akuter Handlungsbedarf durch die Vereinigung zweier Geschäftsbereiche innerhalb der Telekom deutlich. Mit der Zusammenlegung der beiden Bereiche T-Com und T-Online wuchs die Notwendigkeit, DQM unternehmensweit zu etablieren. Um den Kunden kombinierte Produktangebote – also sowohl Festnetzanschlüsse als auch Internetdienstleistungen – bieten und diese zusammen abrechnen zu können, musste z. B. die Konsistenz der Kundenstammdaten mit der Zusammenführung der beiden Datenbestände hergestellt werden. Die in den Bereichen unabhängig voneinander gepflegten Kundenstammdaten waren teilweise redundant, wobei zum selben Kunden gehörende Daten nur mit hohem manuellem Aufwand einander zugeordnet werden konnten. Dies lag zum einen an voneinander abweichenden Kundennummern, zum anderen an der unterschiedlichen Wichtigkeit der Attributdaten für die beiden Bereiche und der damit verbundenen Datenqualität. Während beispielsweise für die T-Com die Adressdaten besonders wichtig waren (auch zur Dokumentation bestehender Festnetzanschlüsse und Infrastruktur) und entsprechend umfassend gepflegt wurden, waren diese für die T-Online zweitrangig.

Durch das fehlende DQM sah sich die DTAG nach der Zusammenlegung von T-Com und T-Mobile mit folgenden wesentlichen Problemen konfrontiert:

- Eine unternehmensweite Transparenz darüber, welche Daten in welcher Qualität unternehmensweit zur Verfügung stehen müssen, war nicht vorhanden.
- Es gab kein einheitliches Verständnis der Stammdatenobjekte (auf Typebene) des Unternehmens und somit keine Kenntnis über die durch das DQM zu verwaltenden Datenobjekttypen.

- Es bestanden keine klar definierten Verantwortlichkeiten für die Pflege und Veränderung von Stammdatenobjekten.
- Unternehmensweit gültige Richtlinien oder Prozesse zur Pflege und Veränderung von Stammdatenobjekten waren nicht definiert bzw. implementiert.
- Herkunft und Verteilung der Stammdatenobjekte in der Applikationsarchitektur waren nicht transparent.

4.2.2.3 Umsetzung

Modellierung von Stammdatenobjekten

Der Ansatz für die Modellierung von Stammdatenobjekten lehnt sich an das Vier-Ebenen-Modell der Unternehmensarchitektur an. Jeder der vier Ebenen sind Modelle der Datenarchitektur zugeordnet, die sich in Detaillierungsgrad und Modellierungsart unterscheiden. Durch eine durchgängige Modellierung auf den vier Ebenen soll die Integration zwischen Fachbereich und IT gewährleistet werden. Tabelle 4-5 zeigt die bei der DTAG verwendeten Modelltypen zur Modellierung von Stammdatenobjekten mit ihrer Zuordnung zu den vier Ebenen des IT-Architekturmodells sowie Beispielmuster der T-Home für jeden Modelltyp.

Architekturebene	Modelltyp	Beispielmodell
Geschäftsprozessarchitektur	Fachbegriffsmodell (Business Object Glossary)	ProuD-Datenmodell
Fachlogische Architektur	Semantisches bzw. Fachlogisches Modell	Business Object Model (BOM)
Applikationen / Applikationsarchitektur	Logische Datenmodelle der einzelnen Applikationen	Datenmodell der Applikation „CRM-Telekom“
Technische Architektur	Physische Datenmodelle (z. B. Datenbankmodelle)	Datenbankmodell der Applikation „CRM-Telekom“

Tabelle 4-5: Modelltypen zur Modellierung von Geschäfts- und Datenobjekttypen

Auf den ersten beiden Ebenen werden die applikationsübergreifend gültigen Geschäftsobjekttypen beschrieben, auf den unteren die Datenobjekttypen der Applikationen. Während es auf den obersten beiden Ebenen für jeden Modelltyp genau eine Ausprägung im Unternehmen gibt, also die Modelle unternehmensweit gültig sind, existieren auf den unteren Ebenen mehrere logische und physische Datenmodelle (jeweils pro Applikation). Die Modellierung folgt einer Kombination von Top-Down- und Bottom-Up-Vorgehen, d. h. es werden einerseits die in den Geschäftsprozessen genutzten Geschäftsobjekttypen sukzessive detailliert. Andererseits werden die Datenmodelle und die in den Applikationen geführten Datenobjekttypen (Ebene 3) bei der Modellierung des semantischen Modells berücksichtigt.

Das Projekt ProuD 2010 (Prozesse und Daten 2010) verfolgt das Ziel einer einheitlichen Darstellung und Gestaltung der Prozessarchitektur der DTAG. Es modelliert die

Unternehmensprozesse auf sechs Detaillierungsebenen. Auf der detailliertesten Modellierungsebene werden die Geschäftsprozesse dem ARIS-Ansatz folgend als ereignisgesteuerte Prozessketten inklusive der Geschäftsobjekttypen modelliert. Die Geschäftsobjekttypen werden im Business Object Glossary in der Sprache des Fachbereiches als Fachbegriffsmodell beschrieben.

Das Business Object Model (BOM) ist das semantische Modell der DTAG. Es definiert und spezifiziert die im Unternehmen relevanten Geschäftsobjekttypen, ihre Eigenschaften und ihre Beziehungen. Als Datenmodell der logischen Architekturebene beschreibt es die von Fachbereich und IT gemeinsam genutzte Begriffswelt, weswegen für das BOM auch der Begriff Integrationsdatenmodell verwendet wird. In die Entwicklung des BOM fließen sowohl Modellierungsergebnisse von ProuD 2010 ein (Fachseite) als auch Datenmodelle der verwendeten Applikationen (IT-Seite). Das BOM regelt die einheitliche, allgemeingültige und redundanzfreie Namensgebung und Verwendung von Begriffen. Es ist ein von Organisationen und konkreten Applikationen unabhängiges Modell. Die Geschäftsobjekttypen des BOM werden den Domänen und Komponenten des Group Domain Models eindeutig zugeordnet. Dem Ziel einer Standardisierung von Terminologie und Modellierung folgend werden sowohl das Fachbegriffsmodell als auch das semantische Modell auf Grundlage des sogenannten SID erstellt. Das SID ist ein vom TeleManagement Forum²¹ entwickeltes generisches Objektmodell, das Entitäten sowie deren Beziehungen zueinander abbildet und definiert, die für ein Telekommunikationsunternehmen von Interesse sind.

Zur Erstellung des BOM wurden durch die Abteilungen MDM und IT2 Modellierungsteams gebildet, die für jeweils einen Geschäftsbereich und für die Modellierung der entsprechenden Entitäten verantwortlich waren. Die Modellierung wurde in Zusammenarbeit mit Vertretern des Fachbereiches und der IT durchgeführt, mit denen regelmäßige Modellierungs-Workshops stattfanden. Dank dieser Workshops konnten sowohl implizites Wissen der Fachbereichsvertreter als auch in den Fachbereichen vorhandene Dokumentationen (u. a. zum ProuD Business Object Glossary sowie Datenmodelle einzelner Applikationen) in die Modellierung einfließen. Anschliessend konsolidierten die Modellierungsteams die Modelle in weitestmöglicher Übereinstimmung mit dem SID.

Das BOM wird als UML-Klassendiagramm im Innovator gepflegt und umfasst in seiner derzeitigen Fassung ca. 600 Klassen. Über eine automatisierte Exportfunktion werden die Informationen des Modells in ein Wiki übertragen, das das Verständnis der BOM-Inhalte für Mitarbeiter ohne Kenntnisse der UML-Modellierungsnotation (z. B. in Fachbereichen) erleichtern soll. Der Grossteil der im Wiki enthaltenen Informatio-

²¹ Das TeleManagement Forum ist eine Arbeitsgemeinschaft von über 700 Unternehmen der IT- und Telekommunikationsindustrie aus mehr als 70 Ländern, deren Ziel die Bereitstellung von Lösungskonzepten für das Management von Informations- und Kommunikationsnetzen ist (siehe <http://www.tmforum.org>).

nen (Attribute, Wertelisten, Assoziationen, Geschäftsregeln) wird direkt aus dem Innovator übernommen. Auch die UML-Klassendiagramme sind für jeden Geschäftsobjekttyp als Graphik hinterlegt. Um Inhalt, Struktur und Verwendungskontext der einzelnen Geschäftsobjekte besser zu dokumentieren, können in dem Wiki zusätzlich Definitionen bzw. Beschreibungen sowie fachliche Metadaten in natürlicher Sprache gepflegt werden.

Die Pflege der Metadaten obliegt zum Zeitpunkt der Fallstudienaufnahme ausschliesslich den Modellierungsverantwortlichen der Abteilungen MDM und IT2. In Zukunft sollen hierfür die Data Architects (siehe nächster Abschnitt) die Verantwortung übernehmen. Eine Öffnung des Wikis mit dem Ziel, Vertreter des Fachbereiches in den Prozess der Beschreibung und Modellierung von Geschäftsobjekttypen zu involvieren, indem diese die Inhalte kommentieren oder verändern können, ist derzeit in Diskussion. Hierdurch könnte die Dokumentation von Inhalt und Verwendungskontext der Geschäftsobjekttypen erleichtert werden. Für die Integration notwendiger Änderungen in das BOM (Hinzufügen von Attributen, Anpassen von Definitionen usw.) wurde ein Change-Request-Prozess definiert, der es Mitarbeitern erlaubt, Änderungsanforderungen zu formulieren. Allerdings bietet die Modellierungsplattform Innovator keine Funktionalitäten zur Versionierung und Änderungshistorie des BOM. Langfristiges Ziel der DTAG ist es, das BOM über seine Rolle als Integrationsdatenmodell hinaus als Richtlinie für die Entwicklung und Einführung neuer Applikation bzw. der zugrunde liegenden logischen Datenmodelle und Schnittstellen (dritte Modellierungsebene) im Unternehmen zu etablieren. Um eine möglichst breite, fachbereichsübergreifende Verwendung zu gewährleisten, ist ein Grossteil der Geschäftsobjekttypen sehr generisch abgebildet. Die Applikationsentwickler prägen diese anschliessend konkret aus.

Auf der dritten Datenmodellierungsebene werden die logischen Datenstrukturen jeder einzelnen Applikation beschrieben. Idealerweise sind diese in Übereinstimmung mit dem BOM definiert. Die Ebene der physischen Datenmodelle, auf der die datenbankspezifische Implementierung modelliert ist, wird im Rahmen der DQM-Massnahmen bei der DTAG nicht betrachtet.

Von besonderer Bedeutung für die Integration im Unternehmen, vor allem zwischen Fachbereich und IT, ist die Ebene der fachlogischen Architektur mit dem Group Domain Model. Im Group Domain Model werden Domänen definiert, die logische Funktionen bündeln. Jeder Domäne werden Geschäftsobjekttypen zugeordnet, die im BOM modelliert sind. Über das Group Domain Model kann zudem abgebildet werden, welche Applikationen die funktionalen Komponenten unterstützen. Dadurch ist das Domänenmodell ein wichtiges Werkzeug für die Integration zwischen Prozess- und Applikationssicht sowie zwischen Prozess- und Datensicht.

Rollen und Verantwortlichkeiten

Derzeit werden bei der DTAG Datenverantwortliche für die Pflege jeweils eines Geschäftsobjekttyps (z. B. Auftrag, Kundensegment) identifiziert. Die Verantwortung liegt dabei nicht nur bei einer Person, sondern verteilt sich auf mehrere Rollen:

- Der Datenverantwortliche ist für den ihm zugewiesenen Geschäftsobjekttyp hauptverantwortlich und trifft finale Entscheidungen in Bezug auf dessen Pflege und notwendige Änderungen. Die Rolle des Datenverantwortlichen wird in der Regel von Bereichs- bzw. Abteilungsleitern aus dem Fachbereich oder der IT besetzt.
- Jedem Geschäftsobjekttyp wird zudem je ein Data Architect aus dem Fachbereich und der IT zugewiesen. Ein Data Architect verantwortet das Management der Metadaten und Datenstrukturen des ihm zugeordneten Geschäftsobjekttyps. Hierzu gehört beispielsweise die aufgrund neuer, fachlicher Anforderungen notwendige Erweiterung von Geschäftsobjekttypen (u. a. im BOM) um benötigte Attribute.
- Sogenannte Data Manager (wiederum jeweils einen seitens des Fachbereiches und der IT) sind für die inhaltliche Datenpflege, also die eigentlichen Datenobjekte (Instanzen) in den Applikationen, im Auftrag der Datenverantwortlichen zuständig.
- Zusätzlich sind fach- und IT-seitige Data Quality Manager für die Steuerung und Kontrolle der Datenqualitätsprozesse verantwortlich.

Integrationsarchitektur für Stammdaten

Eine für das DQM wichtige Aufgabe ist die Gestaltung der Integrationsarchitektur, um redundante Datenhaltung und den damit verbundenen negativen Einfluss auf die Datenqualität zu vermeiden. Die diesbezüglichen Aktivitäten liegen jedoch nicht im Verantwortungsbereich der DQM-Abteilungen, sondern werden durch die Abteilung Enterprise Architecture bzw. die jeweiligen funktionalen Bereiche erfüllt. Grundsätzlich ist derzeit innerhalb der DTAG eine Tendenz zur Konsolidierung und Zentralisierung der Datenhaltung zu erkennen. Diese wird teilweise durch die notwendige Ablösung veralteter Applikationen getrieben.

Voraussetzung für eine systematische Gestaltung der Integrationsarchitektur ist die Kenntnis der bestehenden Applikationsarchitektur und Datenverteilung. Hierzu wird im Rahmen einer konzernweiten Initiative unter Mitarbeit des Bereiches MDM eine sogenannte Datenlandkarte erstellt. Sie umfasst eine Liste der Applikationen der DTAG. Für jede Applikation sind unter anderem die aktuelle Version, der Status, der verantwortliche Bereich, die Wirkbetriebsverantwortung und die Schnittstellen zu anderen Applikationen aufgeführt. Die als Ergebnis der Domänenmodellierung identifizierten Geschäftsobjekttypen, die im BOM abgebildet werden, sind den jeweiligen Applikationen und Domänen mit Hilfe einer CRUD-Matrix zugeordnet. Dadurch wird transparent, welche Applikation wie (Create, Read, Update, Delete) auf ein Objekt zu-

greift bzw. dieses verändert. Die Datenlandkarte erlaubt eine umfassende Analyse der bestehenden Applikationsarchitektur in Bezug auf redundante Datenhaltung (Stammdatenobjekte werden in mehreren Applikationen erzeugt) und (Alt-)Applikationen, die auf keine unternehmensweit verwendeten Stammdatenobjekte zugreifen bzw. diese speichern.

Die Datenlandkarte wird in der Enterprise Architecture Management-Software planningIT gepflegt. Aus dieser besteht – wie beim BOM – ein automatisierter Export in das Wiki, wodurch die Informationen der beiden Modelle direkt miteinander verknüpft sind. Zurzeit dient die Datenlandkarte ausschliesslich der Dokumentation der Ist-Situation. Kurzfristiges Ziel ist es, die Vollständigkeit und somit die Qualität der Datenlandkarte zu verbessern. Eine Verwendung für die Definition einer Soll-Architektur sowie das Ziel einer Konsolidierung der Applikationsarchitektur ist gegenwärtig nicht vorgesehen.

4.2.2.4 Bisherige Erfahrungen

Aufwand und Nutzen

Die MDM-Gruppe Data Governance hat Kosten und Nutzenpotenziale eines durchgängigen Datenmodellierungsansatzes sowie unternehmensweit einheitlicher Datenmodelle bewertet. In die Betrachtung fliessen sowohl qualitative als auch quantitative Annahmen ein. In Bezug auf die qualitativen Nutzenpotenziale werden im Wesentlichen fünf Bereiche beschrieben:

- Ein unternehmensweit einheitliches, logisches Datenmodell definiert eine gemeinsame und konsistente Terminologie, welche die Kommunikation zwischen Fachbereich und IT sowie über Domänengrenzen hinweg erleichtert.
- Die projekt- und domänenübergreifende Definition von Geschäftsobjekttypen als Referenz fördert deren Wiederverwendung, wodurch sich wiederholende Neudefinitionen erübrigen und Kosten und Mehrarbeit reduziert werden.
- Die umfassende Definition der Geschäftsobjekttypen erleichtert deren Verständnis und verringert Fehler durch unsachgemässe Verwendung.
- Einheitliche, logische Datenmodelle schaffen die Voraussetzung für die Konsolidierung verteilter Applikationsarchitekturen und damit verbundene Migrationen.
- Das unternehmensweit einheitliche, logische Datenmodell ist Grundlage für eine effizientere Integration interner und externer Partner, da die Semantik der für angebotene Dienste verwendeten Datenelemente einheitlich ist.

Zusätzlich hat die Abteilung Data Governance einen Business Case auf der Grundlage von Expertenschätzungen berechnet (siehe Kapitel 6.3). Demnach liegt das Gesamt-

einsparpotenzial durch einen über die Architekturebenen durchgängigen Datenmodellierungsansatz bei knapp 0,8 Prozent des gesamten IT-Budgets. Bei einem geschätzten, durchschnittlichen Anteil des IT-Budgets von 4,6 Prozent des Umsatzes bei Unternehmen der Telekommunikationsbranche [Smith/Potter 2009, 31], liesse sich für die DTAG folglich ein Einsparpotenzial in Höhe von mehr als 23 Mio. EUR abschätzen.

4.2.3 Bosch Rexroth

4.2.3.1 Unternehmen

Bosch Rexroth ist einer der weltweit führenden Spezialisten von Antriebs- und Steuerungstechnologien. Das Unternehmen bietet integrierte Lösungen zum Antreiben, Steuern und Bewegen für die Industrieautomatisierung, für mobile Arbeitsmaschinen und Nutzfahrzeuge sowie für erneuerbare Energien. Das angebotene Spektrum ist sehr breit und umfasst pneumatische und hydraulische Produkte, elektrische Antriebe und Steuerungen, Linear- und Montagetechnik, Komponenten für Windkraftanlagen, aber auch Services zur Projektierung. Das Unternehmen entstand im Zuge des Zusammenschlusses der ausgegliederten Automatisierungstechnik aus der Robert Bosch GmbH mit der Mannesmann Rexroth AG im Jahr 2001. Heute ist die Bosch Rexroth hundertprozentige Tochter der Robert Bosch GmbH und erzielte als solche im Jahr 2008 einen Umsatz von mehr als 5,8 Mrd. Euro.

	Bosch Rexroth
Gründung	1. Mai 2001 durch den Zusammenschluss der Mannesmann Rexroth AG mit dem Geschäftsbereich Automatisierungstechnik der Robert Bosch GmbH Gründung des Unternehmens Rexroth: 1795
Firmensitz	Lohr am Main, Deutschland
Branche	Maschinenbau, Automatisierungstechnik
Geschäftsbereiche	Elektrische Antriebe und Steuerungen, Industriehydraulik, Mobilhydraulik, Linear- und Montagetechnik, Pneumatik, Service
Umsatz (2008)	5,88 Mrd. Euro
Mitarbeiter (2008)	35.309
Homepage	http://www.boschrexroth.com

Tabelle 4-6: Kurzporträt Bosch Rexroth

4.2.3.2 Ausgangslage

Auslöser der Initiative war die Entstehung der Bosch Rexroth AG aus dem Zusammenschluss des Bereiches Automatisierungstechnik der Robert Bosch GmbH (ca. 5.000 Mitarbeiter) mit einem Teil der Mannesmann Rexroth AG, welche von Vodafone veräussert wurde (ca. 20.000 Mitarbeiter). Die Problematik des Mergers bestand zu einem grossen Teil in der Integration der weitestgehend unabhängig agierenden Produktbereiche innerhalb der Mannesmann Rexroth AG. Diese Unabhängigkeit führte zu

einer stark heterogenen Prozess- und Applikationsarchitektur, die in den folgenden Jahren standardisiert und harmonisiert werden sollten.

Prozessabläufe sowie Stammdaten wichen zwischen der Mannesmann Rexroth AG und der Bosch Automatisierungstechnik stark voneinander ab. Die alte Applikationsarchitektur der Mannesmann Rexroth AG (zum Zeitpunkt des Unternehmenszusammenschlusses) zeichnete sich daher durch eine hohe Heterogenität aus, bestehend aus SAP-Systemen, AS/400-Systemen sowie eigenprogrammierten, werks- und länderorganisationsspezifischen Applikationen. Abbildung 4-2 zeigt die Applikationsarchitektur von Mannesmann Rexroth zum Zeitpunkt des Mergers.

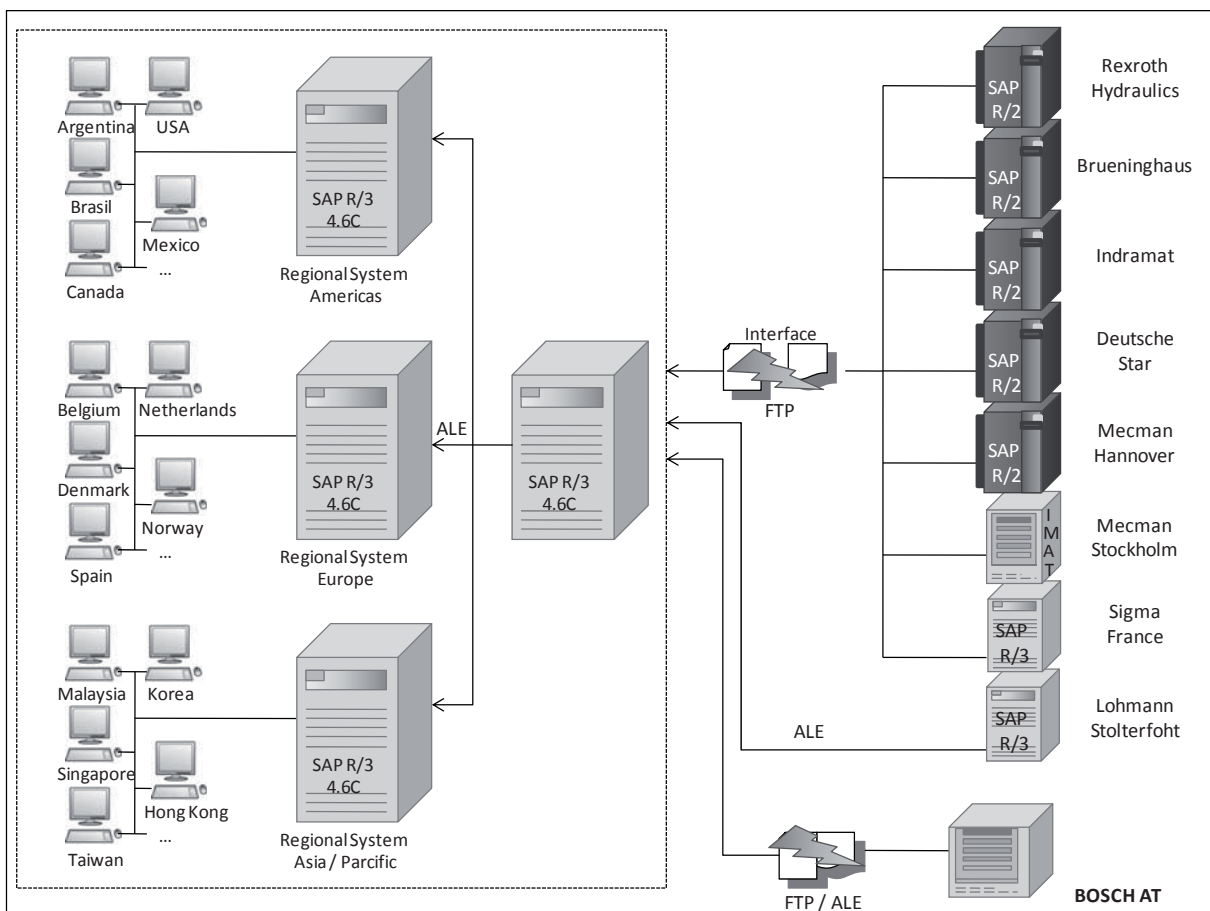


Abbildung 4-2: Applikationsarchitektur der Mannesmann Rexroth AG zum Zeitpunkt des Mergers

Die Datenpflege war hierbei landes- bzw. teilweise sogar werksbezogen organisiert. Das Stammdatenvolumen belief sich zu Beginn auf über 3 Millionen Materialstammdaten und 200.000 Kundenstammdaten, umfasste darüber hinaus aber weitere Stammdatengruppen wie Klassifikationen, Listenpreise, Lieferanten- und Variantenstammdaten.

Die weltweite Geschäftstätigkeit der Bosch Rexroth AG mit globalen Kundenbeziehungen und globalen, unternehmensinternen Geschäftsprozessen führte zu der Not-

wendigkeit einer unternehmensweit harmonisierten Stammdatenbasis. Diesbezüglich bestanden die dringendsten Probleme in:

- der heterogenen Applikationsarchitektur sowohl innerhalb von Mannesmann Rexroth als auch zwischen den beiden fusionierenden Organisationen,
- den uneinheitlich definierten Stammdatenpflegeprozessen, die in den Länderorganisationen bzw. einzelnen Werken voneinander abwichen,
- der unzureichenden Zuordnung von Verantwortlichkeiten für Anlage und Pflege von Stammdatenobjekten,
- dem Fehlen unternehmensweit gültiger Stammdatenstandards für die Beschreibung und Klassifikation von Stammdatenobjekten.

Im Zuge der primären Bemühungen, die Applikationsarchitektur innerhalb der neuen Organisation, Bosch Rexroth, zu harmonisieren, wurde daher speziell die Konsolidierung der Stammdatenbasis angegangen.

4.2.3.3 Umsetzung

Für das Ziel der Stammdatenharmonisierung wurde nicht ein einzelnes Projekt definiert. Vielmehr setzen sich die Harmonisierungsbemühungen aus mehreren Teilprojekten zusammen, die jeweils eine Stammdatenklasse umfassen und grösstenteils nacheinander ablaufen. Aufgrund ihrer Bedeutung für die Geschäftstätigkeit des Unternehmens wurden Kunden- und Materialstammdaten am höchsten priorisiert und waren entsprechend Gegenstand der ersten Harmonisierungsprojekte. Für die schrittweise Harmonisierung der einzelnen Stammdatenklassen wurde ein einheitliches Vorgehen bestehend aus einer Reihe von Standardisierungs-, Bereinigungs- und Migrationsprojekten definiert.

Die ersten Schritte zur Überwindung der beschriebenen Probleme bestanden in der Definition einer Applikationsarchitektur mit einer zentralen Stammdatenverwaltung (separater Stammdatenserver) sowie der Erarbeitung und Einführung von Stammdatenstandards (Standardisierungsprojekt) für jede Stammdatenklasse, die die Grundlage für sämtliche folgende Bereinigungs- und Migrationsprojekte bildete. Mit Stammdatenstandards ist die unternehmensweite Vereinheitlichung von Datenfeldern (Datenobjektattributen) und ihrer Inhalte (Datenelemente) gemeint, durch welche die Qualität der Stammdaten erhöht werden soll. Für die wichtigsten Datenobjektattribute (Nummernkreise zu Materialnummern, Materialarten, Materialartenstatus usw.) wurden auf Grundlage einer ausführlichen Analyse Normen erarbeitet und unternehmensweit veröffentlicht. Für die Normendefinition war hierbei die unternehmensinterne Normenstelle zuständig, in deren Verantwortungsbereich die Vorgabe von Standards für Prozesse und Daten liegt. Diese bildete Arbeitskreise zur Definition der Datenobjekt-

attribute unter Einbezug der verschiedenen Fachbereiche sowie der IT. Durch die Einbindung der Fachbereiche sollen die fachlichen Anforderungen, die sich aus der Verwendung der Stammdatenobjekte in den Geschäftsprozessen ergeben (Verwendungskontext), berücksichtigt werden. Die Arbeitskreise führten zur Festlegung der Definitionen jeweils mehrere Abstimmungsworkshops durch.

Nach der „Urladung“ der Daten (Kunden- und Materialstammdaten, später auch Lieferantenstammdaten) aller Geschäftsbereiche auf den Stammdatenserver wurde der Rollout in mehreren Länderorganisationen durchgeführt. Im Rahmen dieser Rollout-Projekte wurden Altapplikationen sukzessive abgelöst und die restlichen Applikationen schrittweise auf die neue Applikationsarchitektur migriert. Aus der Anbindung weiterer Applikationen, bei denen die Definitionen und Inhalte der Datenobjektattribute vom Stammdatenstandard abweichen (z.B. Länderkennzahlen für den Kundenstamm, Materialarten und -status für den Materialstamm), ergibt sich die fortlaufende Notwendigkeit, diese zu harmonisieren und erforderliche Anpassungen vorzunehmen.

Parallel dazu wurden Teilprojekte zur weltweiten Zentralisierung der bis dahin grösstenteils dezentral in den Werken oder Länderorganisationen verantworteten Stammdatenanlage und -pflege (zentralisierte Datenmanagementprozesse) durchgeführt. Im Zuge dieser Teilprojekte wurden Verantwortlichkeiten in den Länderorganisationen (teilweise bis auf Attributebene) definiert. Bei den Kundenstammdaten, die aufgrund der geringeren Anzahl von Datensätzen im Vergleich zu den Materialstammdaten eine geringere Komplexität aufweisen, wurden zudem bereits Verfahren zur Dublettenerkennung und -bereinigung auf dem Stammdatenserver angewendet. Dank der Implementierung eines in SAP integrierten Tools zur Dublettenprüfung während der Eingabe neuer Kundendaten kann der Stammdatennutzer auf mögliche Dubletten während des Anlageprozesses hingewiesen werden. Die finale Entscheidung über die Anlage liegt dabei weiterhin beim Nutzer. Der letzte Schritt besteht nun in der Spezifikation von Kennzahlen und Prozessen zur Qualitätsmessung und -sicherung.

Die Prozessharmonisierung lief hierbei stets parallel und in enger Verzahnung mit den Stammdaten, wobei immer wieder Anforderungen für die Erweiterung von Datenobjekttypen, für die Definition von Pflegeprozessen oder die Verteilung von Daten entstanden. Insofern waren die harmonisierten Stammdaten stets Voraussetzung und Ergebnis der Prozessintegration.

Rollen und Verantwortlichkeiten

Für die Durchführung der Projekte, insbesondere der Migrationsprojekte, sowie den Betrieb des zentralen Stammdatenservers (siehe folgender Abschnitt) sind die Mitarbeiter der Gruppe Global Data Client and Master Data Management verantwortlich. Die Abteilung ist innerhalb der Unternehmens-IT angesiedelt. Sie begleitet zudem die Arbeitskreise der Standardisierungsprojekte zur Definition von Datenobjektattributen

(fachliche Metadaten) und Pflegeprozessen. Die Arbeitskreise stehen unter der Führung der unternehmensinternen Normenstelle, wobei Vertreter der Geschäftsbereiche (z. B. Prozess-Owner) sowie der IT involviert sind. Die Verantwortung für die Pflege der Stammdatenobjekte (Daten-Owner) bzw. einzelner Attribute liegt in den Fachbereichen.

Stammdatenintegrationsarchitektur

Im Mittelpunkt der neuen Integrationsarchitektur (siehe Abbildung 4-3) steht der Rexroth-Stammdatenserver (Global Data Client) – ein SAP R/3-System, das seit 2001 im produktiven Einsatz ist und auf dem heute jeder Material- und Kundenstamm von Bosch Rexroth angelegt und gepflegt und anschliessend auf die einzelnen Applikationen verteilt wird, wobei die Verteilungskonzepte für die verschiedenen Stammdatenklassen unterschiedlich sind. Die ERP-Prozesse (u. a. Buchhaltung, Fertigung, Vertrieb, Einkauf) laufen jeweils auf den dezentralen SAP-Systemen der einzelnen Regionen (zwei in Deutschland, je eines in Europa ohne Deutschland, in Asien und in Amerika).

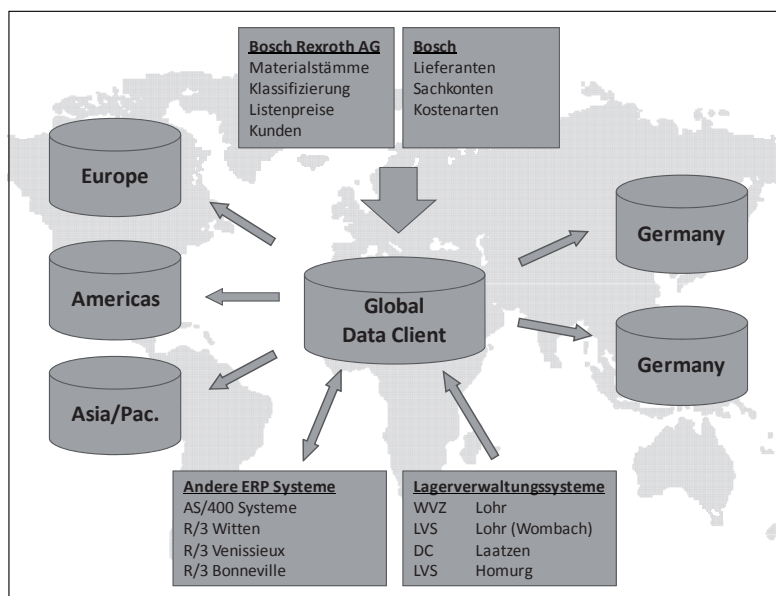


Abbildung 4-3: Neue Applikationsarchitektur mit zentralem Stammdatenserver

Während die Verteilung der Materialstammdaten auf Grundlage eines Abonnement-Ansatzes (basierend auf einem Pull-Prinzip) implementiert ist, werden die Kundenstammdaten über einen Verteilungsmechanismus automatisch verteilt (Push-Prinzip). Grund für diese voneinander abweichenden Distributionsansätze sind die unterschiedlichen Volumina und Änderungsanfragen: Der Materialstamm umfasst heute mehr als 4,2 Millionen Datenobjekte und erfährt täglich ca. 20.000 Änderungen, was bei einer automatisierten Push-Verteilung beträchtliche Datenübertragungen zur Folge hätte. Um die Mengen zu reduzieren, wurde daher zum bereits erwähnten Abonnement-System gewechselt, bei dem jeweils nur die von den angebotenen Applikationen

abonnierten Materialstammdaten automatisch und ohne zeitlichen Verzug (Realtime) übertragen werden. Die zugrunde liegende Technologie für die Verteilung der Daten ist SAP Application Link Enabling (ALE) und SAP Product Data Replication (PDR). Letztere dient primär der Variantenübertragung, d. h. der Auflösung und anschließenden konsistenten Zusammenführung von komplexen Produktstrukturen (mehrstufige Stücklisten, Merkmale, Klassen, Beziehungen usw.), die sich dadurch auszeichnen, dass hier Stammdatenobjekte miteinander in Beziehung stehen. Bei der Verwendung der PDR-Technologie ergaben sich folgende Probleme:

- ein äusserst aufwendiges Schnittstellenmonitoring, welches mittlerweile durch ein eigenentwickeltes Monitoringtool unterstützt wird,
- die Notwendigkeit der Entwicklung zusätzlicher Add-Ons z. B. zur Lastverteilung sowie
- eine geringe Anwenderfreundlichkeit, da die Bedienung der Technologie sehr technisch und bestenfalls für Vertreter der IT geeignet ist (eine Anwenderoberfläche für Fachbereichsanwender musste Rexroth eigenständig entwickeln).

Die beschriebene Ziel-Architektur war zum Zeitpunkt der Fallstudienaufnahme in 16 Länderorganisationen implementiert, wobei weiterhin Rollout-Projekte zur Integration weiterer Länderorganisationen durchgeführt werden. Zudem müssen noch bestehende Altapplikationen (siehe „Andere ERP-Systeme“ in Abbildung 4-3), welche derzeit noch zusätzliche Informationen zu den Stammdaten senden, abgelöst sowie fortlaufend neue Applikationen an den zentralen Stammdatenserver angebunden werden.

Dublettenprüfung und -bereinigung

Wie bereits beschrieben, beinhaltet die neue Lösung für Kundenstammdaten produktive Dublettenerkennungs- und -bereinigungsverfahren und ein in SAP integriertes Tool zur präventiven Dublettenprüfung im Zuge der Anlage neuer Kundenstammdaten. Für die Materialstammdaten ist der Prozess der Dublettenbereinigung um einiges schwieriger. Im Gegensatz zu den Kundenstammdaten sind die Informationen im Materialstamm zu unspezifisch, um einen qualitativ guten Ähnlichkeitsvergleich durchführen zu können. Für einen Dublettenvergleich bleiben eigentlich nur die Materialstammklassifizierung und/oder der Materialstammkurztext. Dies ist jedoch nur möglich, wenn hierzu konzernweit klare Vorgaben und Regeln für den Aufbau und die Pflege dieser Attribute existieren, was zum jetzigen Zeitpunkt bei Bosch Rexroth noch nicht der Fall ist. Dementsprechend sind für Materialstammdaten zurzeit auch noch keine Tools verfügbar. Die Problematik wird dadurch verschärft, dass selbst erkannte Materialdubletten nur sehr schwierig eliminiert werden können, da sie in der Regel mit anderen Objekten verknüpft sind (z. B. in Stücklisten und Zeichnungen) und über eine Historie verfügen. Als erster Schritt zur präventiven Dublettenvermeidung wurden in den Monaten nach Aufnahme der Fallstudie bereits materialverantwortliche Bereiche

für die Stammdatenanlage und -pflege verbindlich festgelegt, um neue Materialien korrekt und einheitlich anzulegen (ca. 900 Materialneuanlagen täglich) und somit neue Materialdubletten zu vermeiden.

Derzeit werden die Kundenstammdaten bei jedem Migrationsprojekt ausführlich und weitestgehend automatisiert auf Dubletten überprüft. Hierfür wurde ein Prüfprozess spezifiziert, dessen Prozessschritte iterativ durchlaufen werden:

- interne Dublettenprüfung in der zu migrierenden Applikation,
- Abgleich mit dem Bestand des Stammdatenservers,
- Festlegung des zu behaltenden Kunden (sogenannter „Mapping-Kunde“) durch den zuständigen Mitarbeiter auf Grundlage einer Vorschlagsliste.

Für den Materialstamm werden die zu migrierenden Datenobjekte manuell gegen den existierenden Datenbestand geprüft. Das Mapping wird dementsprechend in Zusammenarbeit mit dem zuständigen Bereich innerhalb von Bosch Rexroth festgelegt.

4.2.3.4 Bisherige Erfahrungen

Aufwand und Nutzen

Eine ausführliche Kosten-Nutzen-Betrachtung der Harmonisierungsbemühungen bzw. einzelner Teilprojekte für eine Stammdatenklasse steht zum Zeitpunkt der Fallstudienaufnahme bei Bosch Rexroth noch aus.

Die Erfahrungen von Bosch Rexroth zeigen, dass die Harmonisierung von Stammdaten eine sehr aufwendige und langwierige Aufgabe ist, die eine Vielzahl von Ressourcen, Änderungsbereitschaft im Unternehmen und einen hohen Anteil an Eigenentwicklungen erfordert. Hierbei sind Kompromisslösungen aufgrund gegenläufiger Einflüsse und Anforderungen parallel laufender Projekt nicht zu vermeiden. Zudem kann die Harmonisierung von Stammdaten als dauerhafte Herausforderung betrachtet werden, da auch nach einmaliger Bereinigung und Harmonisierung des Stammdatenbestandes ständig neue fachliche Anforderungen (z. B. durch Zukäufe neuer Unternehmen oder Einführung neuer Produkte und Prozesse) entstehen, die eine konsistente Weiterentwicklung bedingen. Als wesentlicher Erfolgsfaktor erwies sich das schrittweise Vorgehen in mehreren überschaubaren Teilprojekten. Dies betraf sowohl die sukzessive Harmonisierung verschiedener Stammdatenklassen entsprechend ihrer Priorität für das Unternehmen als auch die für jede Stammdatenklasse definierten Teilprojekte zur Vereinheitlichung der Datenobjektattribute, zur Migration, zur Zentralisierung der Stammdatenmanagementprozesse usw. Dadurch war es zum einen möglich, die Komplexität der Harmonisierung zu reduzieren, zum anderen aus Erfahrungen bei der Harmonisierung bestimmter Stammdatenklassen für spätere Projekte zu lernen.

Zudem wurde die integrierte Betrachtung von Stammdaten, Applikationen und Geschäftsprozessen als wichtiges Kriterium hervorgehoben. Das Mapping von Stammdaten darf sich nicht auf die Schnittstelle beschränken, sondern muss applikationsübergreifend durchgeführt werden. Dies bedeutet, dass vor allem die Bedeutung der gehaltenen und ausgetauschten Stammdaten, d. h. deren Semantik, applikationsunabhängig definiert werden muss. Dies erfordert die präzise Beschreibung des Geschäftsprozesses und somit des Kontextes, in dem Stammdatenobjekte verwendet werden.

Der hohe Grad an Zentralisierung, der sich nicht nur in der Integrationsarchitektur niederschlägt, die durch den zentralen Stammdatenserver (Global Data Client) gekennzeichnet ist, sondern – darauf aufbauend – auch in den zentral gesteuerten Stammdatenanlage- und -pflegeprozessen, wird durchweg positiv bewertet. Sie erleichtert die konsistente Speicherung der Stammdaten als Grundlage für eine hohe Stammdatenqualität, indem eine Dublette mit geringerem Aufwand erkannt und der Anlageprozess neuer Stammdaten zentral kontrolliert werden kann. Eine derartige Implementierung war jedoch nur möglich, da mit der Zusammenführung von Mannesmann Rexroth und Bosch Automatisierungstechnik ohnehin die Notwendigkeit bestand, die heterogene Prozess- und Applikationsarchitektur zu harmonisieren. Insofern hat sich die Möglichkeit, Massnahmen zur Stammdatenharmonisierung in einen grösseren Projektkontext einzubetten, als nicht unwesentlicher Erfolgsfaktor erwiesen.

Weiterentwicklung

Zum Zeitpunkt der Fallstudienaufnahme waren die Aktivitäten zur Harmonisierung der Material- und Kundenstammdaten weitestgehend abgeschlossen. Aufgrund der Geschäftstätigkeit (z. B. Einführung neuer Produkte und Geschäftsprozesse) entstehen fortlaufend neue fachliche Anforderungen an eine Erweiterung bestimmter Stammdatenobjekte (z. B. im Falle fehlender Datenobjektattribute). Beispiele für derartige Anforderungen sind u. a. die Definition eines Feldes für einen globalen Exportstatus für ein Material oder die Integration von Qualitätsdaten in den Materialstamm infolge neuer Qualitätsprozesse. Die Erweiterung der Stammdatenobjekte um notwendige Attribute (bzw. Felder in den Applikationen) erfordert in der Regel jedoch nur einen geringfügigen Anpassungsaufwand, wenn die Anpassungsfähigkeit der Felder durch die Applikation nicht eingeschränkt ist. Zur Gewährleistung einer konsistenten unternehmensweiten Erweiterung werden jeweils Struktur und Inhalt eines neuen Feldes durch einen speziell dafür zuständigen Arbeitskreis standardisiert vorgegeben. Gleichzeitig wird für das neu definierte Feld der Pflegeprozess inklusive Verantwortlichkeit festgelegt, so dass Stammdaten- und Prozessharmonisierung stets Hand in Hand gehen.

Die harmonisierten Materialstämme bilden die Grundlage für die zukünftige Harmonisierung ganzer Produktstrukturen (Stücklisten und Änderungsdienste). Hierzu wurde ein Projekt initiiert, in welchem die ca. 1,5 Millionen Stücklisten per PDR auf den zentralen Stammdatenserver übertragen und anschliessend wieder verteilt werden. Pa-

rallel dazu wird das bewährte Vorgehen auf weitere Stammdatenklassen übertragen, wodurch sich der Umfang der durch den zentralen Stammdatenserver verwalteten Datenobjekte schrittweise erweitert. Mit der Harmonisierung der Lieferantenstammdaten wurde bereits begonnen, anschliessend sollen Transferpreise, Kostenarten, Sachkonten, Variantenstammdaten und Stücklisten folgen.

4.2.4 Erkenntnisse aus den Fallstudien

Die Fallstudien geben Hinweise für Anforderungen an ein strukturiertes Vorgehen, indem sie Aktivitäten aufzeigen, die im Zuge eines Projektes zur Stammdatenintegration durchzuführen sind. Daher werden die Fallstudien im Folgenden mit Fokus auf das verfolgte Vorgehen zusammengefasst. Zu diesem Zweck werden die einzelnen Aktivitäten identifiziert, zur Bestimmung des jeweiligen Vorgehensmodells in eine zeitliche Ablauffolge gebracht und anhand der inhaltlichen Bewertungskriterien aus Kapitel 3.3 analysiert. Die fallstudien-spezifischen Vorgehensmodelle dienen als Ausgangspunkt für die Ableitung des Vorgehensmodells der eigenen Methode in Kapitel 5.

4.2.4.1 Vorgehensmodell SBB Cargo AG

Abbildung 4-4 zeigt das bei der SBB Cargo verfolgte Vorgehen wie in der Fallstudie beschrieben. Der Schwerpunkt des Vorgehensmodells liegt auf der Identifikation und der semantisch eindeutigen Beschreibung der Geschäftsobjekttypen des Unternehmens, d. h. der Stammdatenobjekte und davon abhängiger Bewegungsdaten auf Typebene aus einer fachlichen Sicht. Die zu den Geschäftsobjekttypen gepflegten, fachlichen Metadaten, die vor Beginn der Definition in einem Metadatenmodell festgelegt wurden, umfassen dabei auch Kontextinformationen zur Verwendung der Geschäftsobjekte in Geschäftsprozessen und Applikationen. Die Geschäftsobjekttypen werden zum einen in einem Glossar beschrieben, zum anderen in einem UML-Klassendiagramm modelliert, wobei in letzterem Metadaten zum Verwendungskontext nicht abgebildet werden. Mit der Festlegung von Stammdatenobjekt-Ownern und Aktivitäten zur Pflege der in Glossar und Datenmodellen verwalteten Metadaten wird das Metadatenmanagement explizit als Tätigkeit des Stammdatenmanagements betont. Die Pflegeprozesse von Daten und Metadaten wurden für jeden Geschäftsobjekttyp gemeinsam mit den Daten-Ownern genau spezifiziert und beinhalteten u. a. die Festlegung von Kriterien für eine erlaubte Veränderung des Stammdatenobjektes, der Geschäftsregeln zur Überprüfung der Abhängigkeiten zu anderen Datenobjekten sowie der Mutationsregeln für jede betroffene Applikation. Diese Regeln werden schrittweise in ein Workflow-Management-System integriert, welches die Änderungs- und Pflegeprozesse unterstützt [Schmidt 2009b, 30-31].

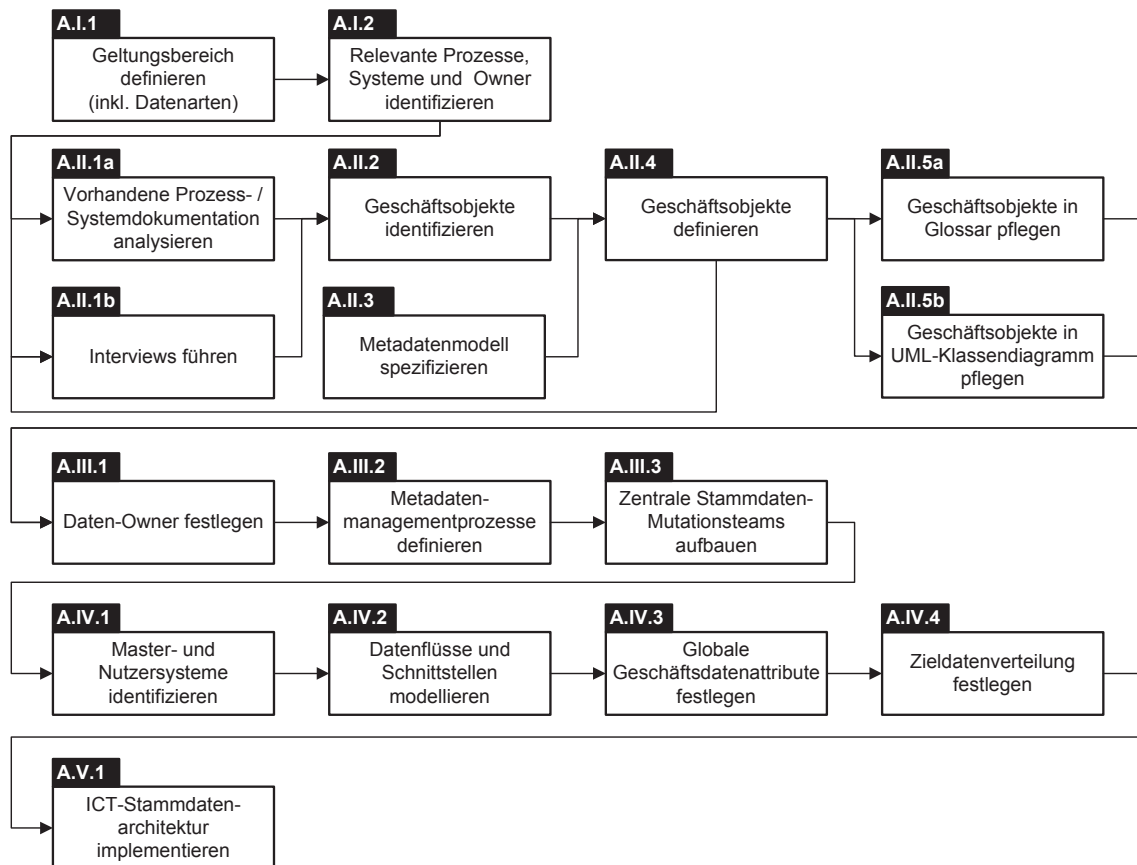


Abbildung 4-4: Vorgehensmodell aus der Fallstudie SBB Cargo AG

Durch die Definition von Nutzer- und Master-Systemen für die Stammdatenobjekte wurde ein erster Schritt zur Gestaltung der Integrationsarchitektur²² unternommen. Die weiteren Aktivitäten zur konzeptionellen Gestaltung der Integrationsarchitektur (Aktivitäten A.IV in Abbildung 4-4) sind im Projektplan als eigene Handlungsfelder definiert [Schmidt 2009b, 17], jedoch wurde mit der Umsetzung zum Zeitpunkt der Fallstudienaufnahme noch nicht begonnen. Hierzu zählen insbesondere die Erstellung der Stammdatenlandkarte, welche sowohl die Zuordnung der Stammdatenobjekte zu den nutzenden Applikationen als auch die Modellierung der Datenflüsse und Schnittstellen zwischen den Applikationen umfasst, sowie die Festlegung globaler Attribute und der Zieldatenverteilung. Langfristiges Ziel ist die Implementierung der konzeptionell erarbeiteten ICT-Stammdatenarchitektur bis Ende 2011. Die dazu notwendigen Aktivitäten sind derzeit noch nicht detailliert beschrieben.

4.2.4.2 Vorgehensmodell Deutsche Telekom AG

Die zeitliche Ablauffolge der für die Stammdatenintegration bedeutsamen Aktivitäten bei der DTAG ist in Abbildung 4-5 dargestellt.

²² Innerhalb der SBB Cargo wird für die Integrationsarchitektur der Begriff ICT-Stammdatenarchitektur verwendet (Information and Communication Technology). Die ICT-Stammdatenarchitektur definiert notwendige Anwendungen für das Stammdatenmanagement sowie Schnittstellen und Datenflüsse zwischen diesen.

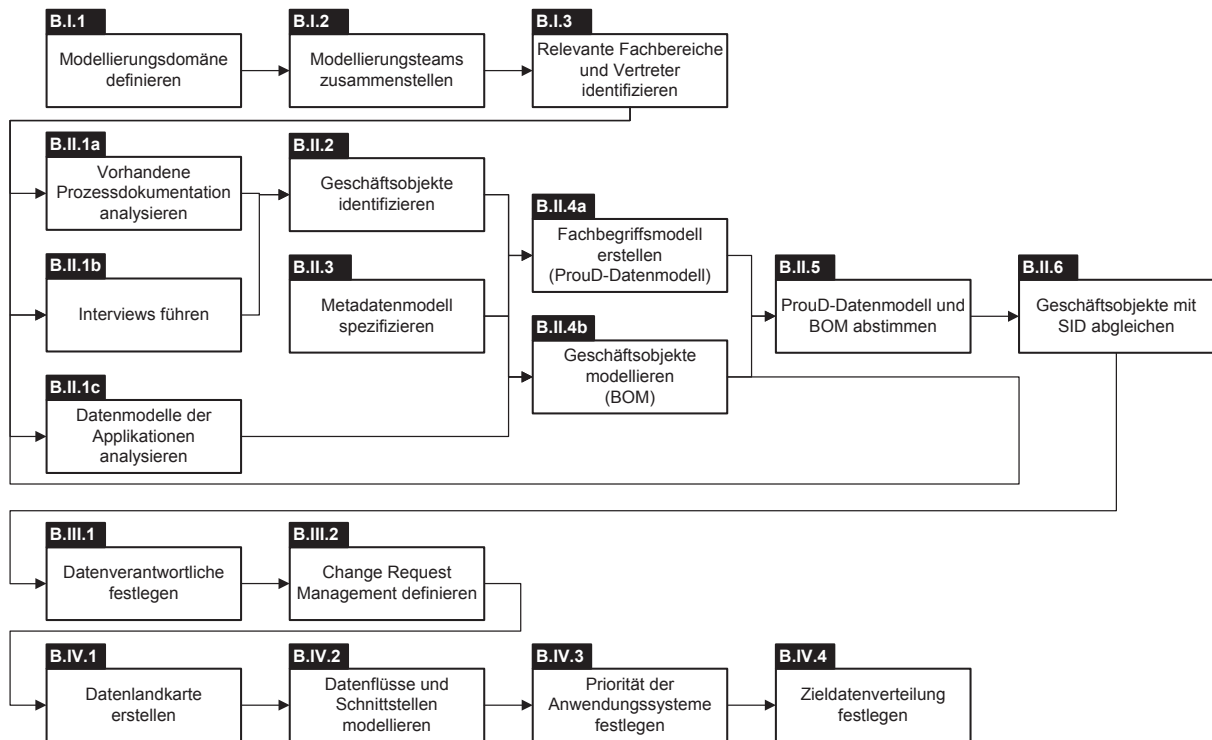


Abbildung 4-5: Vorgehensmodell aus der Fallstudie Deutsche Telekom AG

Analog zum Vorgehen bei der SBB Cargo verfolgt die DTAG eine vom Fachbereich ausgehende Modellierung der Geschäftsobjekttypen (Top-Down), indem basierend auf den Prozessen ein Fachbegriffsmodell (ProuD-Datenmodell) erstellt wird. Dem Fachbegriffsmodell werden die in den Applikationsdatenmodellen enthaltenen Datenobjekttypen (Bottom-Up) gegenübergestellt. Das BOM integriert beide Sichten auf Ebene der fachlogischen Architektur. Dieses kombinierte Vorgehen unterstützt ein einheitliches Verständnis zwischen Fachbereich und IT. Die im BOM gepflegten Metadaten sind in einem Metadatenmodell festgelegt, in dem keine Metadaten zum Verwendungskontext enthalten sind. Das BOM nutzt als Referenz ein spezifisch für die Telekommunikationsbranche definiertes Informationsmodell (das SID) und stützt sich damit auf branchenweit gültige Definitionen für Geschäftsobjekttypen.

In Bezug auf das Metadatenmanagement, also die Pflege der im ProuD-Datenmodell und im BOM gespeicherten Metadaten, sind die Verantwortlichkeiten und Prozesse bisher noch unzureichend festgelegt. Die DTAG unternimmt derzeit Anstrengungen die notwendigen Rollen (Datenverantwortliche, Data Architects, Data Manager) zu definieren und zu besetzen. Die Erstellung unternehmensweit gültiger Metadatenpflegeprozesse ist vorgesehen. Die zuständigen Modellierer (Data Architects) pflegen sämtliche Änderungen in den Modellen, wofür Change-Request-Prozesse, die es dem Metadatenutzer erlauben, der zuständigen Person notwendige Änderungen der Metadaten direkt zu melden, entwickelt werden.

Die Gestaltung der Integrationsarchitektur liegt nur bedingt im Aufgabenbereich der für das Stammdatenmanagement verantwortlichen Abteilung. Die erstellte Datenlandkarte bildet den Ist-Zustand ab und ist ein Ausgangspunkt für den Entwurf der zukünftigen Integrationsarchitektur. Sie enthält im Unterschied zur Stammdatenlandkarte der SBB Cargo jedoch nur die statische Zuordnung von Datenobjekttypen zu Applikationen, ohne Datenflüsse abzubilden.

4.2.4.3 Vorgehensmodell Bosch Rexroth

Anhand des in Kapitel 4.2.3.3 beschriebenen Vorgehens bei Bosch Rexroth kann das in Abbildung 4-6 dargestellte Vorgehensmodell abgeleitet werden.

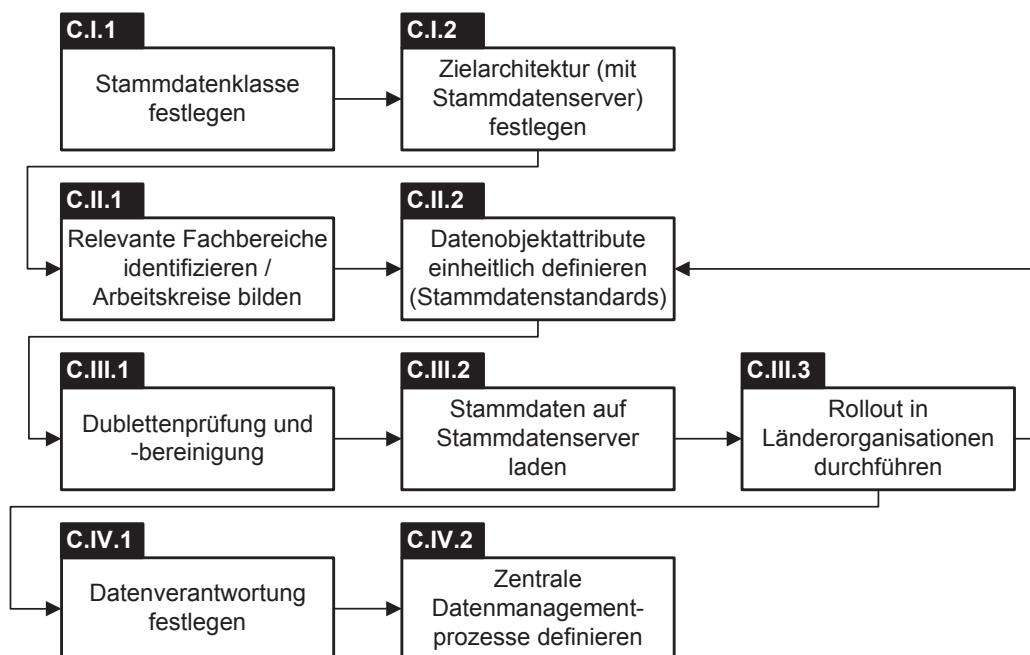


Abbildung 4-6: Vorgehensmodell aus der Fallstudie Bosch Rexroth

Im Gegensatz zum Vorgehen bei der SBB Cargo folgt Bosch Rexroth mit dem Ziel der Zentralisierung der Stammdatenhaltung eher einem Bottom-Up-Ansatz. Ausgangspunkt ist die Harmonisierung der Definitionen der Datenobjekttypen und Datenobjektattribute der verschiedenen Applikationen, die an den zentralen Stammdatenserver (Global Data Client) angebunden werden. Hierfür ist die unternehmensweite Normenstelle verantwortlich. Auf Grundlage der harmonisierten Definitionen erstellt die für das MDM verantwortliche Abteilung ein einheitliches Stammdatenmodell für den Global Data Client sowie für die Schnittstellen zu den lokalen ERP-Systemen der Länderorganisationen. Durch das Entfernen von Duplikaten in der anschließenden Phase werden die Stammdatenbestände der lokalen ERP-Systeme schrittweise bereinigt und auf den Stammdatenserver migriert. Im Zuge der einzelnen Migrationsprojekte können dabei neue Anforderungen an die Definition der Datenobjektattribute entstehen, die eine Anpassung des Stammdatenmodells zur Folge haben. Der letzte

Schritt besteht darin, die Verantwortlichkeiten für einzelne Datenobjekttypen sowie neuer Datenanlage- und -pflegeprozesse auf Basis der zentralen Applikationsarchitektur festzulegen. Das Vorgehen ist insofern iterativ, als einzelne Harmonisierungsprojekte pro Stammdatenklasse (erst Kunden- und Materialstammdaten, später Produktstrukturen, Lieferantenstammdaten usw.) definiert werden. Innerhalb der Projekte werden dann jeweils die gleichen Aktivitäten durchgeführt.

4.3 Aktionsforschungsprojekte

4.3.1 Daimler AG

Die Daimler AG ist der dreizehntgrösste Automobilhersteller weltweit sowie der grösste Produzent von Lastfahrzeugen. Mit seinen mehr als 270.000 Mitarbeitern erwirtschaftet der Konzern einen jährlichen Umsatz von fast 100 Milliarden Euro. Wichtigster Geschäftsbereich der Daimler AG ist Mercedes-Benz Cars (MBC), die PKW-Sparte, welche unter anderem die Kernmarke Mercedes-Benz umfasst und für die Hälfte des Konzernumsatzes verantwortlich ist.

	Daimler AG
Gründung	17. November 1998 (als DaimlerChrysler AG), Umbenennung 2007
Firmensitz	Stuttgart, Deutschland
Branche	Automobilhersteller
Geschäftsfelder	Mercedes-Benz Cars; Daimler Trucks; Daimler Financial Services; Vans, Buses, Other
Unternehmensstruktur	5 Funktionalressorts (Vorstand, Personal, Finanzen & Controlling, Konzernentwicklung & IT, Konzernforschung & Entwicklung) und 41 Produktionsstätten
Umsatz (2008)	95,873 Mrd. Euro
Ergebnis (2008)	2,73 Mrd. Euro (EBIT)
Mitarbeiter (2008)	273.216
Homepage	http://www.daimler.com

Tabelle 4-7: Kurzporträt Daimler AG (vgl. [Daimler 2009])

Mit mehr als 2.000 Applikationen, die ein Ergebnis der vormals werksspezifischen Applikationsplanung sind, verfügt die bestehende Applikationsarchitektur bei MBC über eine beträchtliche Komplexität. Diese führt zu zahlreichen Punkt-zu-Punkt-Verbindungen zwischen Applikationen, erheblichen IT-Kosten, fehlender Transparenz über Quell- bzw. Master-Systeme für Stammdaten und zu einer erhöhten Heterogenität zwischen den Datenmodellen der verschiedenen Applikationen. Vor dem Hintergrund des Kostendrucks in der Automobilindustrie sowie des Bedarfes an einheitlichen Reporting-Strukturen, um die Ergebnisse der einzelnen Standorte vergleichen zu können, bestand zunehmend die Notwendigkeit, die Applikationsarchitektur zu harmonisieren und die Anzahl der Applikationen zu reduzieren. Allerdings ist dieses Ziel ohne das Wissen darüber, welche Applikationen Stammdaten führen, unmöglich, da unklar ist, wie sich die Beseitigung von Applikationen auf die Informationsversorgung des Un-

ternehmens auswirkt. Daher muss das Datenmanagement vorbereitend Transparenz darüber schaffen,

- welche Daten für das Unternehmen erfolgskritisch sind,
- welche Datenobjektattribute welche Geschäftsobjekttypen repräsentieren,
- welche Prozesse die Geschäftsobjekte nutzen und welche Applikationen die zugehörigen Stammdatenobjekte speichern und verändern,
- welche Applikationen für welche Datenobjektattribute führend sind.

Aufgabe des Datenmanagements ist damit die Erarbeitung von Grundlagen, mittels derer einerseits die Datenqualität verbessert, andererseits die IT-Bebauungsplanung im Sinne der beschriebenen Ziele modifiziert werden kann.

4.3.2 DB Netz AG

Die DB Netz AG betreibt das Eisenbahnnetz in Deutschland. Als eines von neun Geschäftsfeldern der Deutsche Bahn (DB) AG gehört das Unternehmen zu 100% der Bundesrepublik Deutschland. Zu den wesentlichen Aufgaben der DB Netz AG gehören Betrieb, Instandhaltung und Erneuerung der Eisenbahninfrastruktur (z. B. Tunnel, Brücken, Gleise, Weichen, Oberleitungen). Mit etwas mehr als 40.000 Angestellten erwirtschaftet das Unternehmen einen Umsatz von ca. 5,5 Milliarden Euro, wobei der Verkauf von Trassen an die 320 deutschen und internationalen Eisenbahnverkehrsunternehmen einen Grossteil des Umsatzes generiert.

	DB Netz AG
Gründung	1998 (Gründung der DB AG im Jahr 1994)
Firmensitz	Frankfurt am Main, Deutschland
Branche	Schienennetzbetreiber
Geschäftsfelder	Betrieb der Eisenbahninfrastruktur (Fern-/Ballungsnetz, Regionalnetz, Zugbildungs- und -behandlungsanlagen), Trassenvermarktung, Fahrplanerstellung, Instandhaltung und Instandsetzung, Weiterentwicklung der Infrastruktur
Unternehmensstruktur	Tochterunternehmen der DB AG (Managementholding) des Ressorts Infrastruktur und Dienstleistungen („DB Netze“) 4 Funktionen (Produktion, Finanzen/Controlling, Personal, Vertrieb), 7 Regionalbereiche und 5 Stabsabteilungen (u.a. Technologie)
Umsatz (2008)	5.551 Mio. Euro, davon 4.826 Mio. Euro Innenumsatz
Ergebnis (2008)	670 Mio. Euro (EBIT)
Mitarbeiter (2008)	40.974
Homepage	http://www.db.de/netz

Tabelle 4-8: Kurzportrait DB Netz AG (vgl. [Deutsche Bahn 2009])

Die wichtigsten Daten der DB Netz sind die Infrastrukturdaten, also diejenigen Stammdaten, welche die Eisenbahninfrastruktur beschreiben. Typische Stammdatenobjekte sind beispielsweise Gleise (ca. 65.000 Kilometer), Weichen und Kopplungen

(ca. 86.000 Objekte), Eisenbahnbrücken (ca. 31.000 Objekte), Bahnübergänge (ca. 24.000 Objekte) oder Tunnel (ca. 700 Objekte). Zwei Aspekte prägten das Management der Infrastrukturdaten in den vergangenen Jahren. Zum einen zeichnen sich die Geschäftsprozesse sehr stark durch die Notwendigkeit aus, dass mehrerer Unternehmensfunktionen zusammenarbeiten müssen. So verlangt ein Grossteil der Wartungs- und Bauarbeiten die Kooperation zwischen Bauplanung, Fahrplan, Instandhaltung und Anlagenmanagement bzw. -buchhaltung. Die funktionale Organisationsstruktur erschwert jedoch die dafür erforderliche Integration der Unternehmensfunktionen sowie der Applikationen. Die Konsequenzen sind längere Durchlaufzeiten, mehr manuelle (Nach-)Arbeit, höhere Prozesskosten durch Doppelarbeiten sowie fehlende Transparenz über den aktuellen Betriebszustand der Infrastruktur.

Zum anderen hat sich die Vergabe der Finanzmittel durch den Bund für die Instandhaltung der Infrastruktur verändert. Während früher die Mittel fallweise vergeben wurden, wird der DB Netz nun jährlich ein pauschaler Betrag zur Verfügung gestellt. Die Höhe des Betrages ist in einem Rahmenvertrag zwischen Bund und DB AG abhängig von der Güte der Infrastruktur festgelegt. Grundlage für die Beurteilung der Güte ist ein Kataster, in dem für jedes Jahr die Anzahl, das Alter und der Zustand der Infrastrukturanlagen zu dokumentieren ist. Da die Informationen über das Inventar zum Stichtag aktuell, konsistent und vollständig vorhanden sein müssen, um daraus u. a. den Infrastrukturzustandsbericht zu erstellen, ergeben sich für das Unternehmen erhöhte Anforderungen an die Qualität der Infrastrukturdaten. Mangelnde Datenqualität kann im schlimmsten Fall die Minderausstattung des Unternehmens mit finanziellen Mitteln zur Folge haben. Als wesentliche Voraussetzung für eine hohe Qualität des Infrastrukturkatasters hat die DB Netz die Verwendung gleicher Definitionen für die Infrastrukturanlagen in allen Regionen und Unternehmensfunktionen erkannt, da nur so Informationen zu ein und demselben Geschäftsobjekt zusammengeführt werden können. Das zur eindeutigen Definition der Geschäftsobjekttypen benötigte Wissen ist jedoch entsprechend der bereits beschriebenen funktionalen Trennung im Unternehmen verteilt und liegt nicht in einer einzelnen Fachabteilung vor. Für das Management der Infrastrukturstammdaten ergeben sich demnach folgende Fragen:

- Was ist eine Definition der Geschäftsobjekttypen Gleis, Weiche, Tunnel usw., die in sämtlichen Geschäftsbereichen gemeinsam verwendet werden kann?
- Welche Attribute eines Geschäftsobjekttyps müssen global für alle Geschäftsprozesse definiert und prozessübergreifend einheitlich verwendet werden?
- Wie können die Informationen über die Geschäftsobjekttypen (Metadaten) dauerhaft und in konsistenter Form im Unternehmen zur Verfügung gestellt werden?
- Wer ist der Verantwortliche für jedes dieser Geschäftsobjekttypen und in welchen Applikationen werden die zugehörigen Datenobjekte gespeichert und verändert?

4.4 Beitrag der Fallstudien und Aktionsforschungsprojekte für die Arbeit

Die Fallstudien und Aktionsforschungsprojekte dieser Arbeit dienen der Ableitung von Anforderungen an die zu entwickelnde Methode. Neben den Folgerungen bezüglich des Vorgehensmodells der zu entwickelnden Methode (siehe Kapitel 4.2.4) zeigen die Fallstudien und Aktionsforschungsprojekte, dass das Ziel der Integrationsprojekte den Umfang der Stammdatenintegration beeinflusst. Grundsätzlich können abhängig von Geschäftstreibern und Zielen der Stammdatenintegration drei Anwendungsszenarien unterschieden werden: die operative Stammdatenintegration, die analytische Stammdatenintegration sowie Stammdatenintegration für die Konsolidierung der Applikationsarchitektur [Dayton 2007, 15-16; Longman 2008, 18; Loshin 2008, 160-165]. Sie unterscheiden sich wesentlich in der Art der Nutzung der Stammdaten und im Aufwand für die Stammdatenintegration.

Ziel der *operativen Stammdatenintegration* ist es, dass Stammdaten in unterschiedlichen operativen Applikationen übereinstimmen. Folglich müssen Stammdaten der operativen Applikationen zum Zweck der Konsistenz der Stammdatenbasis synchronisiert werden [Longman 2008, 18]. Die operative Stammdatenintegration dient in der Regel übergeordneten Zielen wie beispielsweise der Prozessharmonisierung oder der Steigerung der Prozesseffizienz, indem Akteure innerhalb eines Geschäftsprozesse auf konsistente Stammdaten zugreifen. Auch bei der Zusammenlegung von Geschäftsbereichen mit vormals eigener Stammdatenhaltung, wie im Fall der DTAG, müssen Stammdaten zwischen verschiedenen operativen Applikationen (z. B. mehreren MDM-Systemen) synchronisiert werden. Der operative Ansatz ist häufig auf eine bestimmte Stammdatenklasse, z. B. Kundenstammdaten im Rahmen von CDI, beschränkt.

Die *analytische Stammdatenintegration* fokussiert auf die Verbesserung der Qualität des Reporting und somit der Entscheidungsfindung im Unternehmen. Im Gegensatz zur operativen Stammdatenintegration wird hier keine Synchronisation zwischen operativen Applikationen angestrebt, sondern lediglich das Erzeugen eines harmonisierten Referenzdatensatzes in einem Repository [Longman 2008, 18]. Hierfür sind lediglich ein Mapping der Datenelemente sowie die Bereinigung von Redundanzen im Repository, jedoch nicht in den operativen Applikationen notwendig (so wie im Fall des Infrastrukturkatasters bei der DB Netz). Unabhängig davon, ob die Analyse und somit die Aggregation der Stammdaten in einem separaten Business-Intelligence-System oder als Funktion des MDM-Systems durchgeführt wird, ist die Eindeutigkeit von Stammdatenobjekten entscheidend [Dreibelbis et al. 2008, 21]. Da der Datenfluss in diesem Anwendungsfall unidirektional ist und keine Veränderungen in den operativen Applikationen notwendig sind, wird der Aufwand im Vergleich zur operativen Stammdatenintegration geringer eingeschätzt.

Ein in der Literatur selten beachtetes Szenario lässt sich anhand der beschriebenen Praxisfälle identifizieren. Konzerne mit globaler Ausrichtung, in denen Applikationsarchitekturen in der Vergangenheit aufgrund ihrer tendenziell dezentralen Organisation heterogen gewachsen sind, stehen vor der Herausforderung, ihre Applikationsarchitektur zu konsolidieren, um IT-Kosten einzusparen. Damit ist – wie im Fall der Daimler AG – die Aufgabe verbunden, die Anzahl der Applikationen zu reduzieren. Um eine gezielte *Konsolidierung* der Applikationsarchitektur zu ermöglichen, muss transparent sein, welche Applikationen Stammdaten erzeugen und speichern [Russom 2006, 11]. Dieses Szenario stellt insbesondere Anforderungen an die unternehmensweit eindeutige Identifikation von Stammdatenobjekten, auf deren Grundlage führende Systeme bestimmt werden können. Die Integration auf Instanzebene beschränkt sich auf die Übernahme derjenigen Datenelemente, die in abzulösenden Applikationen gehalten werden.

Tabelle 4-9 ordnet die Fallstudien sowie die Aktionsforschungsprojekte den identifizierten Anwendungsszenarien der Stammdatenintegration zu.

Fallstudie / Aktionsforschungsprojekt	Operativ	Analytisch	Konsolidierung
SBB Cargo AG	---	Finanzreporting	---
Deutsche Telekom AG	Zusammenschluss von zwei Geschäftsbereichen 360°-Sicht auf Kunden	---	Ablösung von Legacy-Systemen
Bosch Rexroth	Prozessharmonisierung	---	Harmonisierung der Applikationsarchitektur (Stammdaten)
Daimler AG	Applikationsübergreifende Stammdatenharmonisierung	---	Reduktion der IT-Kosten
DB Netz AG	Prozessharmonisierung	Compliance-Reporting	---

Tabelle 4-9: Kategorisierung der Fallstudien und Aktionsforschungsprojekte anhand der Anwendungsszenarien

Je nach Anwendungsszenario unterscheiden sich die für die Stammdatenintegration notwendigen Aktivitäten, der Umfang der zu harmonisierenden Stammdatenattribute sowie die Entwurfsmuster für die Integrationsarchitektur (siehe Methodenfragment 3), was bei der Entwicklung der Methode zu berücksichtigen ist. Im Sinne des konfigurativen Methoden-Engineering (vgl. [Becker et al. 2007b, 7f.]) dienen die Anwendungsszenarien als Konfigurationsparameter der Methode. Dem Konfigurationsmechanismus der Elementselektion folgend (vgl. [Becker et al. 2002, 73f.]) werden abhängig vom Anwendungsszenario bestimmte Aktivitäten der Vorgehensmodelle sowie die für jede Aktivität definierten Techniken als optional definiert [Schelp/Winter 2006, 568f.].

5 Entwicklung einer Methode zur Stammdatenintegration

In den vorangegangenen Kapiteln wurde die mangelnde methodische Unterstützung zur Lösung von Problemen der Stammdatenintegration identifiziert und die Anforderungen an eine entsprechende Methode erarbeitet. Die Fallstudien untermauern die Anforderungen aus einer praktischen Perspektive. Die folgenden Kapitel entwickeln nun eine Methode, die diesen Anforderungen gerecht wird. Dazu beschreibt Kapitel 5.1 das Methoden-Engineering (ME) als konzeptionellen Rahmen zur Entwicklung von Methoden im Business Engineering. Kapitel 5.2 gibt einen Überblick über die gesamte Methode mit ihren Methodenfragmenten und den dazugehörigen Ergebnissen. Anschliessend werden die Bestandteile der Methode gemäss ME spezifiziert. Das Metamodell (siehe Kapitel 5.3) und Rollenmodell (siehe Kapitel 5.4) werden übergreifend für sämtliche Methodenfragmente definiert, während Vorgehensmodell, Ergebnisdokumente und Techniken anschliessend pro Methodenfragment beschrieben werden. Beispiele und Vorlagen für einzelne Ergebnisdokumente sind den Anhängen D und E beigelegt und werden jeweils in den Aktivitätsbeschreibungen in Kapitel 5 referenziert. Bei der Beschreibung des Vorgehensmodells wird auf die Notwendigkeit und den Umfang der durchzuführenden Aktivitäten und Techniken je nach Anwendungsszenario (siehe Kapitel 4.4) eingegangen. Zusätzlich stellen die Kapitel zu den Methodenfragmenten mögliche Software-Tools vor, die Aktivitäten zur Zielerreichung unterstützen. Die Methodenbeschreibung schliesst mit dem Dokumentationsmodell in Kapitel 5.8, das eine Übersicht über sämtliche Entwurfsergebnisse der Methode sowie über ihre Beziehungen zueinander gibt. Konkrete Beispiele aus der Anwendung der Methode sind teilweise bei der Beschreibung der Techniken enthalten, ausführlicher geht darauf Kapitel 6 ein.

5.1 Methoden-Engineering

Das *Methoden-Engineering* (ME) hat seinen Ursprung im Bereich des Software-Engineerings, in dem es der Entwicklung und Anpassung von Methoden durch eine systematische und strukturierte Beschreibung der Methodenelemente und ihrer Beziehungen dient [Heym 1993, 5; Nuseibeh et al. 1996, 267]. Methoden selbst werden als ein integriertes, systematisches Vorgehen zur Entwicklung von IS verstanden, indem sie Handlungsempfehlungen liefern, durch welche Aktivitäten und Techniken ein bestimmtes Ergebnis zu erreichen ist [Heym/Österle 1993, 345; Brinkkemper 1996, 275f.]. In der Vergangenheit wurden basierend auf dem ME mehrfach erfolgreich Vorgehensmodelle und Methoden für das Business Engineering entwickelt (vgl. beispielsweise [Kaiser 2000], [Pohland 2000], [Schwinn 2005]). Das ME wendet Prinzipien von Methoden auf den Entwurf von Methoden selbst an [Heym/Österle 1993,

347; Gutzwiller 1994, 11] und wird in dieser Arbeit zur Entwicklung der Methode zur Stammdatenintegration angewendet.

Der Aufbau einer Methode gemäss dem ME lehnt sich an die von GUTZWILLER anhand einer Analyse verschiedener Ansätze zum ME identifizierte Struktur an [Gutzwiller 1994, 13]. Die Elemente sind im Metamodell des ME (siehe Abbildung 5-1) mit ihren Beziehungen zueinander abgebildet. Auf die einzelnen Elemente wird an dieser Stelle nicht im Detail eingegangen, da sie bereits in zahlreichen Publikationen definiert sind (vgl. u. a. [Gutzwiller 1994, 12-14], [Österle/Blessing 2003, 80f.], [Höning 2009, 13f.]).

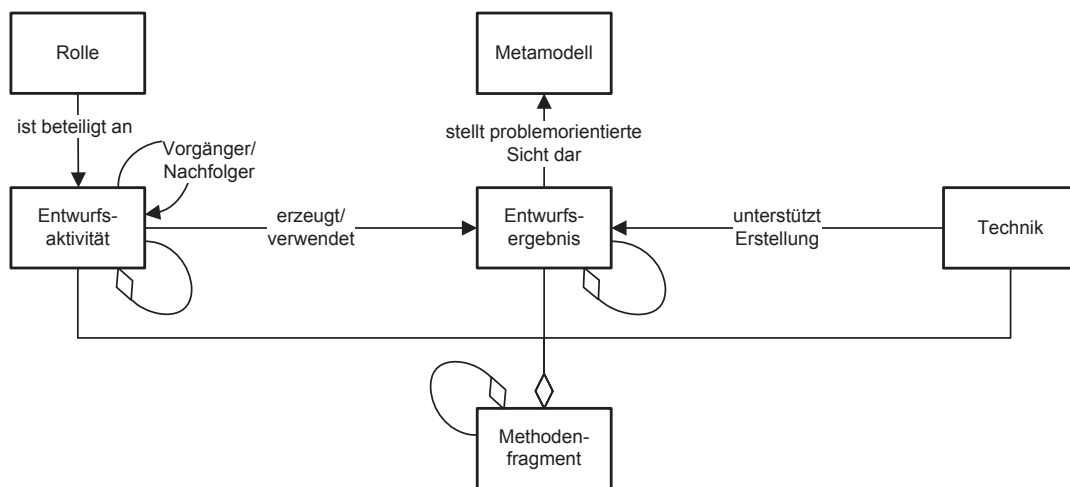


Abbildung 5-1: Elemente der Methodenbeschreibung (in Anlehnung an [Gutzwiller 1994, 13] und [Bucher 2009, 17])

Aufbauend auf dem Metamodell definiert BUCHER den Begriff der *Methodenfragmente*, der sich stark am Konzept der Methodenkomponente anlehnt (vgl. [Karls-son/Wistrand 2006, 85ff.]). Ein Methodenfragment wird durch genau ein Entwurfsergebnis einer Methode charakterisiert und kann durch dieses eindeutig identifiziert werden. Zusätzlich besteht es aus jeweils einer oder mehreren Techniken, Entwurfsaktivitäten und Rollen [Bucher 2009, 48]. Das Konzept der Methodenfragmente wird in dieser Arbeit verwendet, um die Methodenbeschreibung logisch sinnvoll anhand wesentlicher Entwurfsergebnisse aufzuspalten. Die Dissertation verfolgt nicht das mit Methodenfragmenten oft anvisierte Ziel, eine Methode situativ, d. h. abhängig von bestimmten Kontingenzfaktoren (z. B. Projekttyp, Umweltparameter), zu konstruieren und anzuwenden [Kumar/Welke 1992, 258; ter Hofstede/Verhoef 1997, 402; Bucher et al. 2007, 45-54]. Vielmehr soll die zu entwickelnde Methode derart spezifiziert werden, dass der Nutzer sie ohne Formulierung von Adaptionsmechanismen kontext- und projektabhängig anpassen und anwenden kann.

5.2 Überblick über Methode und Entwurfsergebnisse

Ziel einer Methode zur Stammdatenintegration muss die Schaffung einer integrierten, verbindlichen Stammdatenbasis sein (siehe Kapitel 2.2.3), welche die konsistente Verteilung qualitativ hochwertiger Stammdaten auf die stammdatennutzenden Applikationen im Unternehmen gewährleistet. Der technologiegetriebene Ansatz zur Etablierung einer integrierten Stammdatenbasis, z. B. über die Implementierung neuer MDM-Lösungen, ist in der Vergangenheit vielfach gescheitert [Martin/Leben 1989, 8; White et al. 2006, 2; Loshin 2008, 14f.]. Dies ist u. a. damit zu erklären, dass es für die Integration einer sorgfältigen Analyse der Stammdaten, ihres Inhaltes und ihres Verwendungskontextes bedarf, die ein Softwaretool allein nicht bewältigen kann [Adelman et al. 2005, 35]. Ausgangspunkt muss die Integration der Entitäten auf einer fachlichen Ebene sein. Diese Anforderung lässt sich nicht nur aus der Literatur ableiten (vgl. [Russum 2006, 5; Kentouris 2009, 14]), sondern wurde auch anhand der im vorherigen Kapitel beschriebenen Fallstudien bestätigt. Die fachliche Integration bezeichnet die Etablierung und Pflege eindeutiger und konsistenter Definitionen der Stammdaten aus Sicht der Fachbereiche bzw. der Prozesse, in denen diese als Geschäftsobjekte repräsentiert werden. Hierfür ist die Harmonisierung der Struktur und der Semantik der prozessspezifisch verwendeten Geschäftsobjekttypen erforderlich (siehe Punkt 1 in Abbildung 5-2). Für die Harmonisierung müssen die fachlichen Metadaten der Geschäftsobjekttypen unternehmensweit standardisiert verwaltet werden [Ryu et al. 2006, 191; Baumöl/Meschke 2009, 65].

Erst auf Grundlage dieser Definitionen der Geschäftsobjekttypen als gemeinsame Sprache ist eine Überführung in konsistente, implementierungsnähere Datenmodelle möglich. Dazu bedarf es einerseits der Zuordnung der identifizierten Geschäftsobjekttypen zu den entsprechenden Datenobjekttypen der Systemebene, andererseits eines Ansatzes zur Abbildung der in den Geschäftsobjekttypdefinitionen enthaltenen Semantik in den Datenmodellen (siehe Punkt 2 in Abbildung 5-2). Auf diese Weise werden die fachlichen Anforderungen an Stammdaten für die informationstechnische Umsetzung vorgegeben. Das semantische Stammdatenmodell bildet die Basis für die Verteilung auf verschiedene Applikationen und somit für die Gestaltung der Integrationsarchitektur, in deren Rahmen Fragen der Stammdatenverteilung und des Stammdatenaustausches beantwortet werden (siehe Punkt 3 in Abbildung 5-2).

Das Vorgehen mit den beschriebenen Gestaltungsbereichen der Stammdatenintegration ist in Abbildung 5-2 zusammenfassend dargestellt.

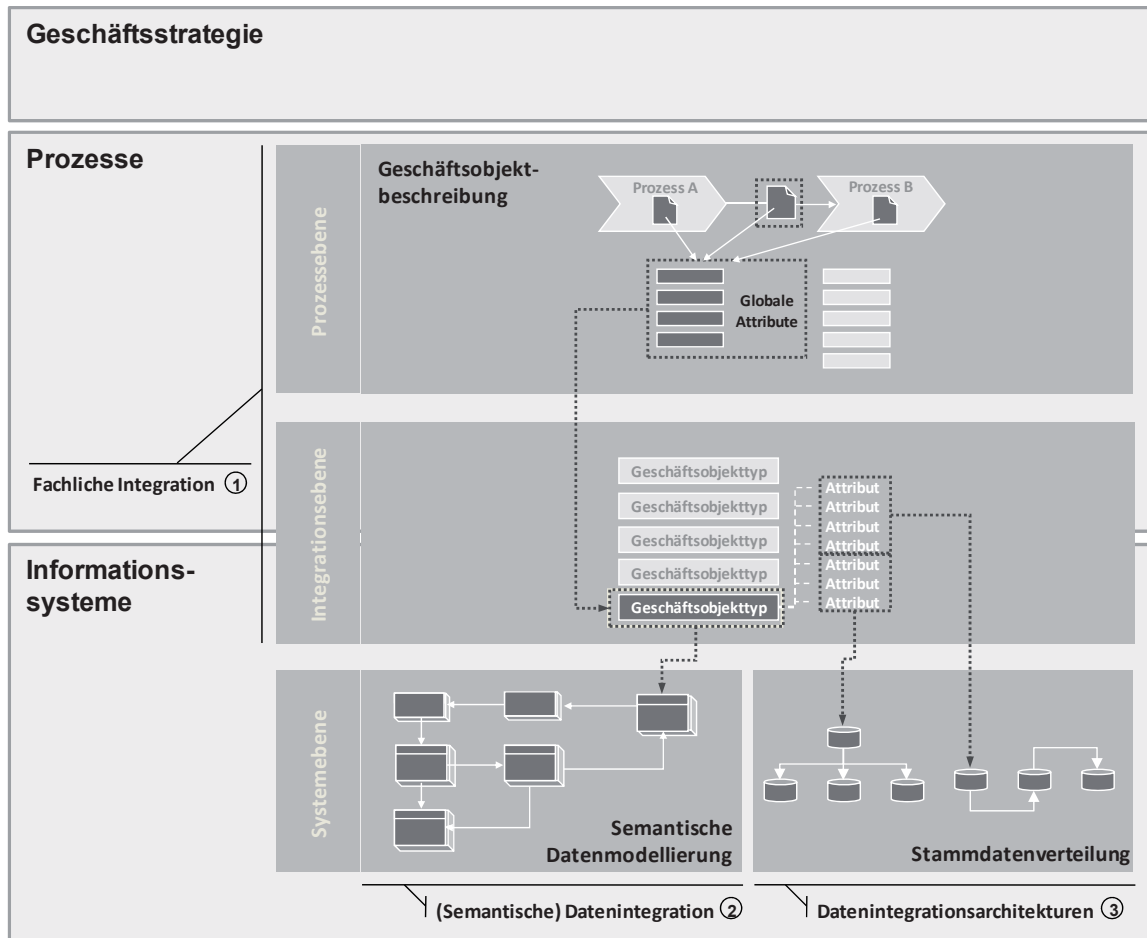


Abbildung 5-2: Bestandteile und Vorgehen bei der Integration von Stammdaten

Die Methode zur Stammdatenintegration folgt diesem Vorgehen und ist entsprechend den wesentlichen Entwurfsergebnissen in drei Methodenfragmente (MF) unterteilt. Methodenfragment 1 beinhaltet die Integration von Geschäftsobjekttypen, Methodenfragment 2 behandelt die Integration von Datenobjekttypen und Methodenfragment 3 umfasst die Integration der Datenobjekte. Die Methodenfragmente verwirklichen in ihrer vollständigen Umsetzung das Ziel einer integrierten Stammdatenbasis auf fachlicher und systemtechnischer Ebene. Tabelle 5-1 fasst die Methodenfragmente mit ihren Entwurfsergebnissen zusammen.

Nummer	Name	Entwurfsergebnis
Methodenfragment 1	Geschäftsobjekttypen in einem BDD beschreiben	BDD als Repository für fachliche Metadaten mit definierten Geschäftsobjekttypen
Methodenfragment 2	Semantisches Stammdatenmodell ableiten	Applikationsübergreifendes, semantisches Stammdatenmodell
Methodenfragment 3	Stammdatenintegrationsarchitektur gestalten	Stammdatenintegrationsarchitektur

Tabelle 5-1: Methodenfragmente der Methode zur Stammdatenintegration und ihre Entwurfsergebnisse

Für die Unterteilung in Methodenfragmente sprechen zwei Gründe. Zum einen reduziert sie die einem Stammdatenintegrationsprojekt inhärente Komplexität, indem sie

die Gesamtaufgabe in überschaubare Teilaufgaben mit definierten Meilensteinen (Entwurfsergebnisse) gliedert. Dies vereinfacht nicht nur deren Beschreibung, sondern auch die Anwendbarkeit durch potenzielle Nutzer. Zum anderen werden Teile der Methode flexibler anwendbar, da sie abhängig von den bereits in einem Unternehmen bestehenden Entwurfsergebnissen (z. B. bestehende Geschäftsobjekttypdefinitionen) nicht zwingend vollständig umgesetzt werden muss [Bucher 2009, 49].

Die Entwurfsergebnisse der Methodenfragmente sind als aufeinander aufbauend zu betrachten, d. h. sie stehen in einer zeitlichen Reihenfolge, wobei Rücksprünge bzw. Iterationen möglich sind. Abbildung 5-3 zeigt die zeitliche Reihenfolge der Methodenfragmente sowie die Fallstudien, die für die Ableitung des jeweiligen Methodenfragmentes massgeblich waren (schwarz hinterlegte Buchstaben).

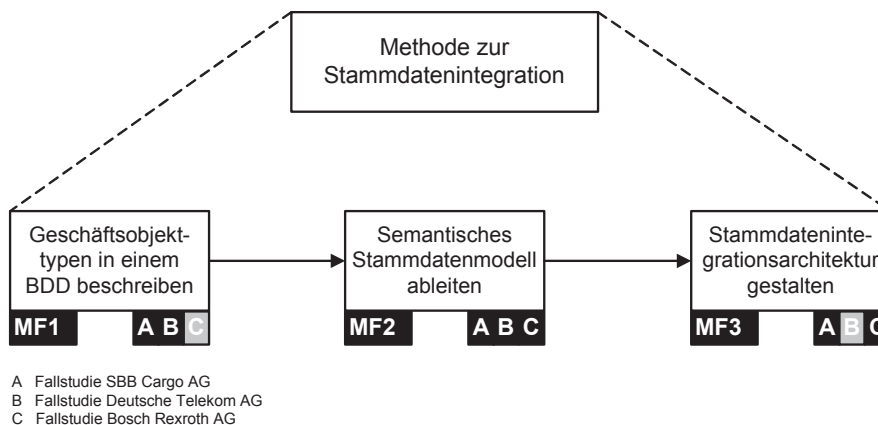


Abbildung 5-3: Zeitliche Abfolge der Methodenfragmente

5.3 Metamodell

Das Metamodell legt den terminologischen und konzeptionellen Rahmen der Methode eindeutig fest. Daher steht die Beschreibung des Metamodells zu Beginn der Ausführungen, auch wenn dessen Entwicklung auf den Ergebnissen der übrigen Entwicklungsschritte beruht. Als konzeptionelles Datenmodell beschreibt das Metamodell die Gestaltungsobjekte bzw. Entwurfsergebnisse der Methode in Form von Metaentitätstypen, definiert deren Bedeutung (Semantik) und deren Beziehungen zueinander sowie Konsistenzbedingungen für die Verwendung der Gestaltungsobjekte [Sinz 1999, 1035]. Das Metamodell der Methode legt den Schwerpunkt folglich auf die Spezifikationen der verwendeten Begriffe und somit den konzeptionellen Aspekt [Strahinger 1996, 44f.], wodurch das Verständnis der folgenden Methodenbestandteile vereinfacht wird. Die Auswahl der Metaentitätstypen und ihrer Beziehungen basiert auf den Erfahrungen der Fallstudien und Aktionsforschungsprojekte (siehe Kapitel 4).

Als Bezugsrahmen für die Metamodellierung dient in dieser Arbeit die von der Object Management Group (OMG) definierte Meta Object Facility (MOF). Sie bietet ein

mehrstufiges Metamodellierungsrahmenwerk, um mit einer minimalen Anzahl an Konzepten andere Metamodelle zu beschreiben [Schnabel 2007, 49]. Zu diesen Konzepten gehören Klassen, Assoziationen und Multiplizitäten. Wegen der Ähnlichkeit der Meta Object Facility mit den Strukturmodellen der UML eignen sich UML-Klassendiagramme zur Modellierung von Metamodellen und haben als Notationsstandard in Wissenschaft und Praxis weite Verbreitung gefunden. Sie wird daher für die Darstellung des Metamodells verwendet. Während Klassen die Gestaltungsobjekte des Metamodells repräsentieren, stellen Beziehungen in Form von Assoziationen, Aggregationen oder Generalisierungen Abhängigkeiten zwischen Gestaltungsobjekten dar [Braun 2007, 21]. Um die Eigenschaft einer Beziehung zwischen Gestaltungselementen zu verdeutlichen, kann die Beziehung zusätzlich mit einem Stereotyp und mit der Angabe von Multiplizitäten qualifiziert werden [Oestereich 2004, 57].

Zusätzlich zur graphischen Darstellung des Metamodells (siehe Abbildung 5-4) werden in Anhang B.1 alle Gestaltungsobjekte (Metaentitätstypen) alphabetisch aufgeführt und textuell beschrieben. Zur Beschreibung der Metaentitätstypen wird folgende Struktur verwendet:

- Bezeichnung: Name des Metaentitätstyps,
- Definition: präzise Beschreibung des Metaentitätstyps,
- Beziehungen: Beschreibung der Beziehungstypen zu anderen Metaentitätstypen.

Die Definitionen der Metaentitätstypen stützen sich auf wissenschaftliche Literatur aus dem Bereich der Wirtschaftsinformatik und des Business Engineering. Die in Abbildung 5-4 mit dem Präfix „BE::“ versehenen Metaentitätstypen sind in ihrer Bedeutung direkt dem Metamodell des Methodenkerns des BE entnommen [Österle et al. 2007, 192f.; Höning 2009, 96-116].

CDQM auf (siehe Abbildung 5-5). Dies reduziert zum einen den Aufwand für die Konstruktion des Rollenmodells der Methode und erlaubt es zum anderen, einen konsistenten Begriffsapparat, welcher die heterogenen Stellen- und Rollenbezeichnungen in den verschiedenen Unternehmen konsolidiert²³, zu verwenden. Die Rollen werden im Folgenden mit ihren Aufgaben und Verantwortlichkeiten bei der Stammdatenintegration erläutert. Sie werden im Zuge der Vorgehensmodellbeschreibung der Methodenfragmente den einzelnen Aktivitäten zugeordnet. Ein zusammenfassendes Funktionsdiagramm, welches für sämtliche Aktivitäten der Methode die Verantwortlichkeiten dokumentiert ist Anhang B.2 beigefügt.

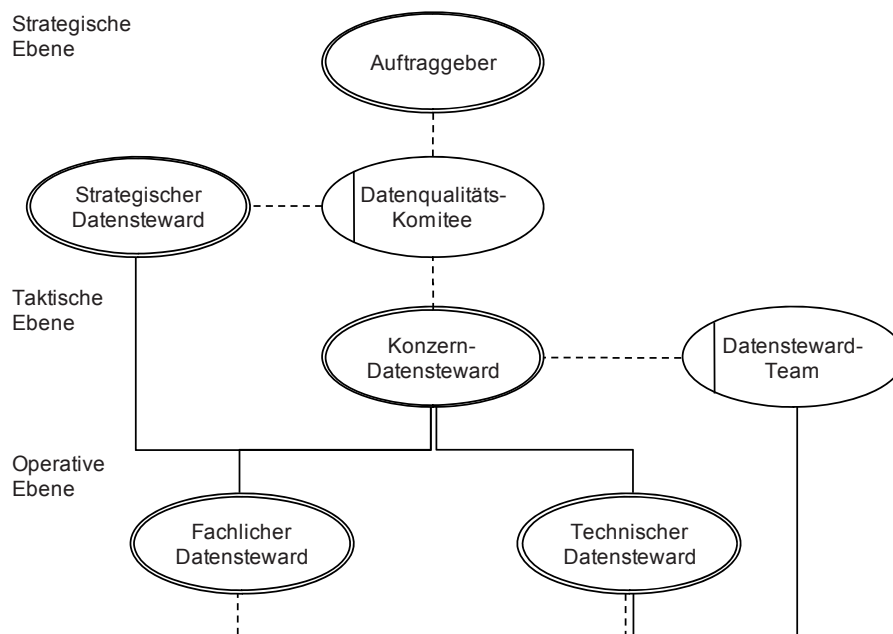


Abbildung 5-5: Referenz-Rollenmodell des CDQM [Weber 2009, 112]

Die Rollen (doppelt umrandete Ellipsen) und Gremien sind für langfristige, unternehmensweite DQM-Entscheidungen, die Durchführung und Steuerung von Datenqualitätsinitiativen im Unternehmen sowie die operative Umsetzung von DQM-Massnahmen zuständig [Weber 2009, 111]. Die Integration von Stammdaten (und die damit verbundene Gestaltung der Informationsarchitektur) stellt eine solche Datenqualitätsinitiative dar. Daher werden für die Methode ausschliesslich diejenigen Rollen übernommen, die einen direkten Bezug zu dieser Aufgabe haben. Dies sind im Einzelnen:

- Der *strategische Datensteward* (engl.: Strategic Data Steward) verfügt über höchste Entscheidungsbefugnis für einen Teil der Kerngeschäftsobjekte [English 2006, 9f.]. Er verantwortet insbesondere die Metadaten der Geschäfts- und Datenobjekttypen (Definition, Struktur, Verwendung), legt Datenpflegeprozesse fest, entschei-

²³ Eine Gegenüberstellung der verschiedenen Bezeichnungen mit dazugehörigen Quellenverweisen gibt WEBER [Weber 2009, 112-113].

det über Standards und Geschäftsregeln und vermittelt im Fall von Unstimmigkeiten zwischen den operativen Stewards.

- Der *fachliche Datensteward* (engl.: Business Data Steward) ist auf operativer Ebene für die Identifikation und Definition der Geschäftsobjekttypen (Geschäftsterminologie), der mit ihnen verbundenen Geschäftsregeln und Standards in seinem Verantwortungsbereich aus fachlicher Sicht zuständig [Dyché/Levy 2006, 167f.; Marco/Smith 2006, 17]. Als Schnittstelle zwischen Fachbereich und Datenqualitätsmanagement verfügt er unter anderem über Wissen darüber, welche fachlichen Anforderungen bestehen und welche Geschäftsobjekte in welchen Geschäftsprozessen verwendet werden.
- Der *technische Datensteward* (engl.: Technical Data Steward) übersetzt als Vertreter der IT fachliche Terminologien und Geschäftsregeln (Geschäftsobjekttypen) in informationstechnische Repräsentationen (Datenobjekttypen). Insbesondere entwickelt und pflegt er Datenmodelle und spezifiziert Datenformate und Datentypen [Weber 2009, 119]. Zudem besitzt er umfassendes Wissen über stammdatenführende Applikationen sowie die Datenflüsse zwischen den Applikationen [Dyché/Levy 2006, 169].
- Der *Konzern-Datensteward* (engl.: Corporate Data Steward) ist als Pendant des strategischen Datenstewards auf operativer Ebene massgeblich für die Leitung und Durchführung von DQ-Initiativen verantwortlich und koordiniert die Arbeit der einzelnen fachlichen und technischen Datenstewards.

Das *Datenqualitäts-Komitee* führt zwar keine operativen Tätigkeiten aus, ist als strategisches Entscheidungsgremium des DQM jedoch für wichtige Beschlüsse bei der Umsetzung der Methode (z. B. Freigabe von Definitionen, Umsetzung von Metadatenpflegeprozessen) zuständig. Das *Datensteward-Team* ist ein Gremium, das der Koordination der Arbeit sowie dem Austausch zwischen den einzelnen Datenstewards dient, für die Methode zur Stammdatenintegration jedoch von untergeordneter Bedeutung ist.

Da Stammdatenintegration nicht ausschliesslich Aufgabe des Datenqualitätswesens ist, sondern auch Fragestellungen der Prozess- und Applikationsarchitektur sowie des Metadatenmanagements behandelt, müssen weitere Unternehmensrollen in den Prozess integriert werden [Loshin 2008, 27].

- Sämtliche stammdatenführenden Applikationen sind potenziell von einer Stammdatenharmonisierung bzw. -integration betroffen. Daher gilt es, *System-Owner* (engl.: Application Owner) sowie *Datenbankadministratoren* in die Initiative zu involvieren. Sie bringen das Fachwissen über die applikationsspezifischen Datenobjekte in ihren Applikationen bzw. Datenbanken und folglich Anforderungen für die Erstellung eines konsolidierten Datenmodells ein [Loshin 2008, 28f.].

- Der *Datenarchitekt* (engl.: Enterprise Data Architect, Data Engineer oder Data Modeler) gewährleistet die Versorgung des Unternehmens mit benötigten Daten. Er ist für die Erstellung und Pflege von Informations- und Datenmodellen im Unternehmen sowie die Implementierung und Veränderung von Datenbanken verantwortlich. Dadurch spielt er bei der logischen Datenintegration eine wichtige Rolle [Adelman et al. 2005, 142].
- Wie bereits mehrmals erwähnt, haben (fachliche) Metadaten bei der Integration von Stammdaten eine grosse Bedeutung. Der *Metadatenadministrator* (engl.: Metadata Analyst oder Metadata Administrator) ist für die Pflege und Aktualisierung der Metadaten zu Geschäfts- und Datenobjekttypen in einem Metadaten-Repository zuständig [Dyché/Levy 2006, 171; English 2006, 10]. Diese Rolle kann ein fachlicher Datensteward oder ein Datenarchitekt übernehmen.

Zudem können weitere Vertreter des Fachbereiches, z. B. der Prozess-Owner, als Informationsquellen zur Identifikation und Definition von Geschäfts- und Datenobjekttypen in einzelne Aktivitäten des Vorgehensmodells involviert werden.

In Bezug auf ihren Aufgabenbereich ähnliche Rollen, die jedoch in ihrer Benennung vom vorgestellten Rollenmodell abweichen, lassen sich in den einzelnen Fallstudien und Aktionsforschungsprojekten identifizieren. Eine Gegenüberstellung der Rollenmodelle der Methode sowie der Fallstudien und Aktionsforschungsprojekte, in denen Rollen und Gremien für das DQM bereits spezifiziert wurden, zeigt Tabelle 5-2.

Methode zur Stammdatenintegration	SBB Cargo	DTAG	DB Netz
Strategischer Datensteward	Daten-Owner	Datenverantwortlicher	---
Konzern- Datensteward	---	Datenverantwortlicher	DB Netz Datenmanager
Fachlicher Datensteward	Mutationsstelle	Data Manager (Fachbereich)	Fachlicher Datenmanager
Technischer Datensteward	Mutationsstelle	Data Manager (IT)	Technischer & Operativer Datenmanager
Datenqualitäts-Komitee	Prüfstelle	---	IDM Board
System-Owner	System-Owner	Systemverantwortlicher	Systemverantwortlicher
Datenarchitekt	Stammdatenmanagement-Team	Data Architect	---
Metadatenadministrator	Stammdatenmanagement-Team	Data Architect	---

Tabelle 5-2: Vergleich der Rollenmodelle von Methode und Fallstudien bzw. Aktionsforschungsprojekten

5.5 Methodenfragment 1: Geschäftsobjekttypen in einem BDD beschreiben

Wie bereits zu Beginn des Kapitels erläutert, muss der Fachbereich der Ausgangspunkt für eine Integration der Stammdaten sein. Zum einen erlaubt dies die Fokussierung auf die tatsächlich genutzten Kerngeschäftobjekte anstelle einer nur schwer zu realisierenden Bestandsaufnahme sämtlicher Datenelemente eines Unternehmens [Loshin 2008, 79]. Zum anderen können die Fachbereiche den Verwendungskontext und die Semantik der Geschäftsobjekte am besten beschreiben [Russom 2006, 3; Kentouris 2009, 14].

Die zentrale, werkzeuggestützte Steuerung fachlicher Metadaten ist für das Stammdatenmanagement und das damit verbundene Ziel einer integrierten Stammdatenbasis erfolgskritisch (vgl. [Blechar/White 2009], [Loshin 2008, 48]). Bereits in Kapitel 2.3.1.4 wird darauf verwiesen, dass durch die Umsetzung eines zentralen Metadatenmanagements Unternehmen einen höheren Reifegrad des DQM erreichen, indem sie Geschäftsobjekttypen und deren Attribute unternehmensweit einheitlich spezifizieren [Ryu et al. 2006, 194f.]. Voraussetzung dafür ist die Zusammenführung der im Unternehmen verteilt vorliegenden Metadaten.

Obwohl gemäss aktueller Marktstudien lediglich ein Prozent der befragten Unternehmen ihre fachlichen Metadaten im Rahmen des Stammdatenmanagements pflegen, wird der potenzielle Nutzen für die Zukunft als hoch bewertet [White et al. 2008, 10]. Für Grossunternehmen, welche aufgrund ihrer Grösse und ihres historischen Wachstums keine Transparenz über die Kerngeschäftobjekte besitzen, ergeben sich durch einen zentral gesteuerten fachlichen Metadatenkatalog mehrere Nutzenpotenziale [Informatica 2006, 9f.; Blechar/White 2009, 4-6]:

- Der fachliche Metadatenkatalog, insbesondere seine Erstellung, fördert die organisationseinheitsübergreifende Zusammenarbeit und Harmonisierung der verschiedenen Unternehmensfunktionen.
- Der fachliche Metadatenkatalog leistet einen Beitrag für das Business/IT-Alignment. Fachliche Anforderungen an die Stammdaten werden durch Explikation der Semantik (also der Bedeutung und des Kontextes, in dem der Fachbereich Stammdaten nutzt) nachvollziehbar formuliert und können entsprechend in der IT-technischen Umsetzung berücksichtigt werden. Damit unterstützt ein fachlicher Metadatenkatalog die Integration der verschiedenen Sichten und der abweichend verwendeten Terminologien zwischen Fachbereich und IT [White 2008, 6].
- Der fachliche Metadatenkatalog erhöht das Vertrauen des Fachbereiches in die von ihm genutzten Stammdaten, da Mitarbeitern eine verbindliche Quelle zur Verfügung gestellt wird, in der sie zu einem Geschäftsobjekt benötigte Informationen (Metadaten) vorfinden.

- Der fachliche Metadatenkatalog unterstützt die Etablierung von Verantwortlichkeiten für Stamm- und Metadaten, da die für die Pflege zuständigen Rollen für jeden Geschäftsobjekttyp im fachlichen Metadatenkatalog hinterlegt werden.
- Der fachliche Metadatenkatalog dient als wichtige Referenz für Entwurf und Implementierung der Stammdatenintegrationsarchitektur.

Das erste Methodenfragment beabsichtigt daher, Geschäftsobjekttypen mittels fachlicher Metadaten eindeutig zu spezifizieren und die fachlichen Metadaten konsistent in einem Metadatenkatalog, dem Business Data Dictionary (siehe Kapitel 2.3.1.3), zu pflegen. Diese Integration auf fachlicher Ebene bildet die Voraussetzung für eine systemseitige Integration der Stammdaten.

5.5.1 Vorgehensmodell und Aktivitäten

Um das Vorgehensmodell des Methodenfragmentes *Geschäftsobjekttypen in einem BDD beschreiben* systematisch abzuleiten, werden die Aktivitäten der Fallstudien mit Bezug zum Methodenfragment vergleichend analysiert und konsolidiert, indem gleichartige zusammengefasst werden. Massgeblich für die Zusammenfassung ist das Ergebnis der Aktivitäten. Tabelle 5-3 fasst die Gegenüberstellung sowie die daraus abgeleiteten Aktivitäten des ersten Methodenfragmentes zusammen. Die Aktivität MF1: III.2 beschreibt die Implementierung des BDD als Werkzeug für das Management der fachlichen Metadaten. Sie wird in den Fallstudien nicht explizit genannt, da in diesen Fällen auf bestehende IT-Lösungen innerhalb der Unternehmen (z. B. Wikis, Datenmodellierungstool Innovator) zurückgegriffen wurde.

Nr.	Aktivität	Ergebnisse	Aktivität aus Fallstudie		
			SBB Cargo	DTAG	Bosch Rexroth
I Vorbereitung					
I.1	Gestaltungsbereich definieren	Abgegrenzter Gestaltungsbereich	A.I.1	B.I.1	C.I.1
I.2	Stakeholder und Geschäftsprozesse identifizieren	Liste Ansprechpartner	A.I.2	B.I.2, B.I.3	C.II.1
II Bestandsaufnahme					
II.1	Geschäftsprozesse auswerten	Geschäftsprozessarchitektur	A.II.1a, A.II.1b	B.II.1a, B.II.1b	–
II.2	Geschäftsobjekttypen konsolidieren	Liste der zu definierenden Geschäftsobjekttypen	A.II.2	B.II.2	–
III Definition					
III.1	Metadatenmodell für BDD spezifizieren	Metadatenmodell des BDD, Anforderungsspezifikation BDD	A.II.3	B.II.3	–
III.2	BDD implementieren	Implementierte BDD-Applikation	–	–	–
III.3	Geschäftsobjekttypen definieren	Geschäftsobjekttypdefinitionen	A.II.4, A.II.5a, A.II.5b	B.II.4a	(C.II.2)
IV Metadatenpflege					
IV.1	Metadatenmanagementprozesse entwerfen	Definierte Metadatenmanagementprozesse	A.III.1, A.III.2	B.III.1	–
IV.2	Metadatenmanagementprozesse implementieren	Implementierte Metadatenmanagementprozesse	A.III.3	B.III.2	–

Tabelle 5-3: Aktivitäten des induzierten Vorgehensmodells von Methodenfragment 1

Die zeitlich logische Folge der einzelnen Aktivitäten ist in Abbildung 5-6 dargestellt. Die gestrichelten Pfeile verdeutlichen inhaltliche Beziehungen zwischen Aktivitäten, ohne dass zwischen diesen eine zeitliche Abhängigkeit besteht. Die mit römischen Ziffern nummerierten Phasen gruppieren inhaltlich zusammenhängende Aktivitäten.

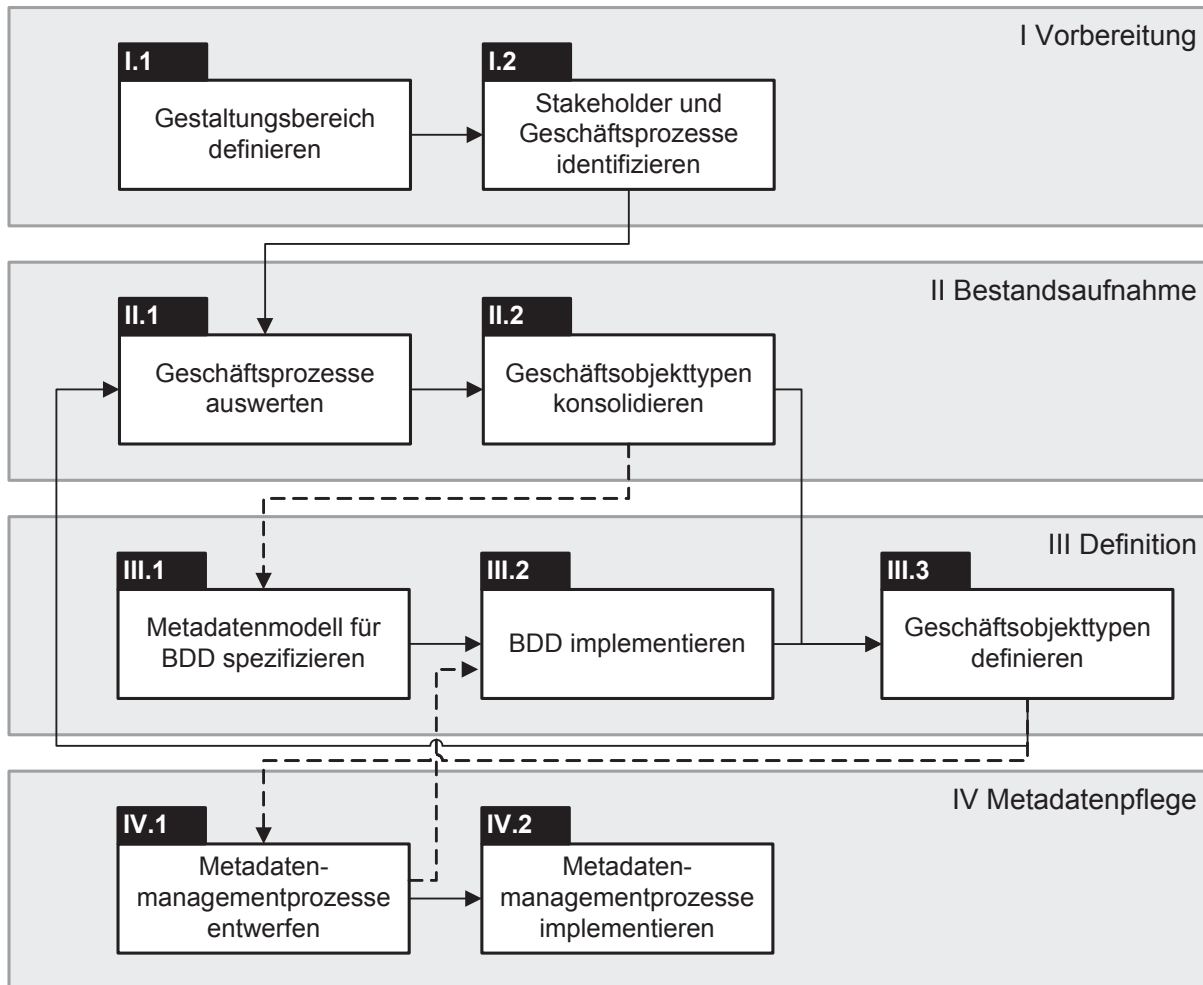


Abbildung 5-6: Vorgehensmodell des Methodenfragmentes 1

Die einzelnen Aktivitäten werden im Folgenden kurz textuell beschrieben und jeweils durch eine Tabelle, die verantwortliche Rollen, Ergebnisse, notwendige Vorbedingungen und Techniken nennt, charakterisiert. Die Rollen sind dem in Kapitel 5.4 definierten Rollenmodell entnommen. Die Techniken erläutert Kapitel 5.5.3.

5.5.1.1 Gestaltungsbereich definieren

Die vorherrschende Komplexität global tätiger Unternehmen in Bezug auf Grösse, Organisationsstruktur, Anzahl und Verteilung der stammdatennutzenden Applikationen bedingt eine bewusste Beschränkung des initialen Anwendungsbereiches der Methode [Adelman et al. 2005, 109-112; Longman 2008, 18]. Dies ermöglicht es, den Aufwand für die Identifikation und die einheitliche Definition von Geschäftsobjekttypen im Sinne einer Standardisierung in kleinen Schritten²⁴ zu minimieren. Die iterative Erweiterung des Gestaltungsbereiches kann sich noch innerhalb des ersten Methoden-

²⁴ ROSENTHAL ET AL. sprechen diesbezüglich von „Standardize in small granules“ [Rosenthal et al. 2004, 48].

fragmentes oder mit abgeschlossener Aktualisierung des BDD in Methodenfragment 2 (siehe Abbildung 5-14) vollziehen.

Der Gestaltungsbereich kann anhand zweier Kriterien eingegrenzt werden:

- Fokussierung auf bestimmte Stammdatenklassen (z. B. ausschliesslich Material- oder Kundenstammdaten) oder
- Auswahl eines funktions- und applikationsübergreifenden Geschäftsprozesses und der darin genutzten bzw. veränderten Stammdaten.

Um einen bestimmten Anwendungsbereich festzulegen, sollten die Stammdatenklassen oder die Geschäftsprozesse anhand definierter Relevanzkriterien (z. B. Bedeutung für das Unternehmen, erwarteter Einfluss auf Datenqualitätsverbesserung) bewertet und priorisiert werden (siehe Techniken T.I.1 und T.I.2). Der initiale Gestaltungsbereich der Methode sollte vor Projektbeginn durch den strategischen Datensteward in Abstimmung mit dem Datenqualitäts-Komitee definiert werden. Auch die vorgesehenen schrittweisen Erweiterungen sollten hierbei im langfristigen Projektplan festgelegt werden.

MF1: I.1 Gestaltungsbereich definieren	
Ergebnis	Abgegrenzter Gestaltungsbereich
Rollen	Strategische Datenstewards [ausführend, verantwortlich] Datenqualitäts-Komitee [genehmigend/freigebend]
Techniken	T.I.1 Bewertung der Gestaltungsbereiche
Input/Vorbedingungen	Projekt (mit Budget) zur Stammdatenintegration ist freigegeben

Tabelle 5-4: Aktivität MF1: I.1 Gestaltungsbereich definieren

5.5.1.2 Stakeholder und Prozesse identifizieren

Nach Definition des Gestaltungsbereiches der Methode gilt es, einerseits die betroffenen Prozesse zu bestimmen (falls im vorherigen Schritt ein bestimmter Geschäftsprozess ausgewählt wurde, entfällt diese Aufgabe), andererseits die Experten aus den Fachbereichen zu ermitteln, die über Wissen bezüglich der in den Prozessen verwendeten Geschäftsobjekte verfügen. Erste Kontaktperson ist in der Regel der Prozess-Owner, da diese Rolle in vielen Unternehmen heute schon etabliert ist und sie am ehesten einen Überblick über im Prozess verwendeten Geschäftsobjekte und deren Verantwortliche besitzt. Sind aufgrund vorheriger Massnahmen zur Etablierung einer Data Governance im Unternehmen bereits Rollen wie beispielsweise fachliche Datenstewards oder Daten-Owner spezifiziert und besetzt, sind diese als Ansprechpartner für die Durchführung der folgenden Aktivitäten mit Namen und Rollenbezeichnung zu dokumentieren.

MF1: I.2 Stakeholder und Geschäftsprozesse identifizieren	
Ergebnis	Liste Geschäftsprozesse und Ansprechpartner
Rollen	Konzern-Datensteward [ausführend, verantwortlich] Fachliche Datenstewards bzw. Prozess-Owner [unterstützend]
Techniken	T.I.2 Identifikation der Daten- bzw. Prozess-Owner
Input/Vorbedingungen	Gestaltungsbereich der Methodenanwendung ist definiert

Tabelle 5-5: Aktivität MF1: I.2 Stakeholder und Prozesse identifizieren

5.5.1.3 Geschäftsprozesse auswerten

Mit Abschluss der Identifikation von Gestaltungsbereich und relevanten Prozessen werden die fachlichen Datenstewards mit der Analyse der Geschäftsprozesse mit Fokus auf die Ermittlung von Geschäftsobjekttypen beauftragt. Ist die Rolle des fachlichen Datenstewards im Unternehmen nicht etabliert, erfüllt eine Rolle mit entsprechendem Tätigkeitsprofil (z. B. der Daten-Owner) oder der Prozess-Owner diese Aufgabe. Der Konzern-Datensteward koordiniert die Analyse der Geschäftsprozesse.

MF1: II.1 Geschäftsprozesse auswerten	
Ergebnis	Geschäftsprozessarchitektur (Fokus: Geschäftsobjekte)
Rollen	Konzern-Datensteward [verantwortlich] Fachliche Datenstewards [ausführend] Prozess-Owner, Fachbereichsvertreter [unterstützend]
Techniken	T.II.1 Inhaltsanalyse Geschäftsprozessdokumentation T.II.2 Einzelinterviews
Input/Vorbedingungen	Identifizierte Geschäftsprozesse und Ansprechpartner

Tabelle 5-6: Aktivität MF1: II.1 Geschäftsprozesse auswerten

Ziel der Analyse ist es, den Informationsbedarf aus fachlicher Sicht, also aus Geschäftsprozessperspektive, zu ermitteln [Loshin 2008, 134]. Für die Geschäftsprozessanalyse stehen im Wesentlichen zwei Techniken abhängig von den vorhandenen Informationsquellen zur Verfügung [Biazzo 2000, 106; McComb 2003, 158]. Zum einen können die in Technik T.I.2 identifizierten Ansprechpartner im Rahmen von Interviews direkt zu ihren Tätigkeiten befragt werden (siehe Technik T.II.2), zum anderen besteht die Möglichkeit, im Unternehmen verfügbare Dokumentationen zu Geschäftsprozessen auszuwerten (siehe Technik T.II.1). In beiden Fällen ist das Hauptaugenmerk der Analyse auf die Geschäftsobjekte zu legen (Auswertung beispielsweise der Informationsflüsse zwischen einzelnen Prozessschritten oder der Bearbeitung von Geschäftsdokumenten). Die Techniken T.II.1 und T.II.2 sind sowohl für die Aktivitäten der Identifikation (Aktivität MF1: II.1) als auch der Beschreibung von Geschäftsobjekttypen (Aktivität MF1: III.3) anzuwenden.

5.5.1.4 Geschäftsobjekttypen konsolidieren

Aus der Auswertung der Geschäftsprozesse können potenzielle Geschäftsobjekttypen für die Definition in einer Liste zusammengetragen werden. In der Liste sollten unterschiedliche Bezeichnungen für denselben Geschäftsobjekttyp (Synonyme) bereits vermerkt werden. Zusätzlich ist für jeden Geschäftsobjekttyp ein Verantwortlicher (z. B. ein fachlicher Datensteward oder Daten-Owner) aus den einzelnen Fachbereichen, die entsprechende Geschäftsobjekte nutzen, zu dokumentieren (siehe Anhang D.1). Diese sind in den anschließenden Definitionsprozess (siehe Aktivität MF1: III.3) zu involvieren.

MF1: II.2 Geschäftsobjekttypen konsolidieren	
Ergebnis	Liste der zu definierenden Geschäftsobjekttypen (mit jeweiligen fachlichen Datenstewards)
Rollen	Strategische Datenstewards [verantwortlich] Konzern-Datensteward [ausführend] Datenqualitäts-Komitee [genehmigend/freigebend]
Techniken	T.II.3 Identifikation von Synonymen und Homonymen / Homogenisierung T.II.4 Bewertung der Geschäftsobjekttypen
Input/Vorbedingungen	Vollständige Geschäftsprozessanalyse

Tabelle 5-7: Aktivität MF1: II.2 Geschäftsobjekttypen konsolidieren

Die Aufgabe der Konsolidierung umfasst die Bewertung und Priorisierung der zu definierenden Geschäftsobjekttypen [Loshin 2008, 79]. Hierzu sollten analog zu Aktivität MF1: I.1 Relevanzkriterien als Grundlage der Bewertung definiert werden. Zusätzlich zu eigendefinierten Kriterien ist es möglich, unternehmensübergreifend verwendete Klassifikationsansätze zu nutzen. Ein Beispiel hierfür ist der Klassifikationsstandard der Storage Network Industry Association (SNIA), der Geschäftsdaten abhängig von ihrer Kritikalität für den Geschäftserfolg in fünf Kategorien (von unwichtig bis betriebsnotwendig) unterteilt [Thome/Sollbach 2007, 181; Schmidt/Osl 2008, 29-30]. Die Priorisierung der Geschäftsobjekttypen ist massgeblich für deren anschließende Definition.

5.5.1.5 Metadatenmodell für BDD spezifizieren

Die konsistente und gleichzeitig effiziente Pflege der fachlichen Metadaten der Geschäftsobjekttypen ist nur durch eine zentrale, werkzeuggestützte Steuerung möglich [Loshin 2008, 48]. Daher muss das Methodenfragment die Bereitstellung eines BDD gewährleisten, welches die Metadaten speichert, aktualisiert und unternehmensweit zugänglich macht. Diese Aufgabe wird mit den Aktivitäten MF1: III.1 und MF1: III.2 in zwei Schritten erfüllt.

Im ersten Schritt wird das Metadatenmodell erstellt, das den Kern des BDD bildet. Es spezifiziert sämtliche fachlichen Metadaten, die zur eindeutigen Definition eines Ge-

schäftsobjekttyps spezifiziert werden müssen. Das Metadatenmodell kann entweder unternehmensspezifisch modelliert oder durch Anpassung eines Referenz-Metadatenmodells an die eigenen Anforderungen erstellt werden. Dies hängt entscheidend von den Anforderungen an die Metadaten im Unternehmen ab. Kapitel 5.5.2 zeigt die Ableitung eines solchen Referenz-Metadatenmodells für ein BDD. Es sollte die fachlichen Anforderungen an das BDD berücksichtigen und daher federführend von den fachlichen Datenstewards erstellt und – durch den Datenarchitekt unterstützt – modelliert werden.

Abhängig vom Anwendungsszenario der Stammdatenintegration (siehe Kapitel 4.4) bestehen Unterschiede im Umfang des Metadatenmodells. Während für die analytische Stammdatenintegration die Definition weniger Attribute, anhand derer ein Geschäftsobjekttyp eindeutig identifiziert werden kann, ausreichend ist, müssen im operativen Anwendungsszenario sämtliche globalen Attribute gepflegt werden. Dies wird auch anhand des Umfangs der gepflegten Metadaten in den Fallstudien ersichtlich. Während bei der SBB Cargo (analytisches Anwendungsszenario) eine relativ geringe Anzahl an Metadatenattributen mit Fokus auf die stammdatennutzenden und stammdatenführenden Applikationen gepflegt wird, ist das Metadatenmodell der DTAG (operatives Anwendungsszenario) sehr viel umfangreicher (siehe Kapitel 5.5.2). Dadurch unterscheidet sich auch der zeitliche Aufwand für die Sammlung fachlicher Metadaten in den Folgeaktivitäten (beispielsweise bei Aktivität MF1: III.3).

MF1: III.1 Metadatenmodell des BDD spezifizieren	
Ergebnis	Metadatenmodell des BDD (mit sämtlichen Attributen, die zur eindeutigen Spezifikation eines Geschäftsobjekttyps im Unternehmen definiert werden müssen)
Rollen	Strategische Datenstewards [verantwortlich] Konzern-Datensteward, Fachliche Datenstewards [ausführend] Datenarchitekt [unterstützend] Datenqualitäts-Komitee [genehmigend/freigebend]
Techniken	T.III.1 Datenmodellierung bzw. Auswahl relevanter fachlicher Metadatenklassen
Input/Vorbedingungen	Anforderungen des Fachbereiches (bedingt auch die Liste der zu definierenden Geschäftsobjekttypen)

Tabelle 5-8: Aktivität MF1: III.1 Metadatenmodell des BDD spezifizieren

Die Aktivität MF1: III.1 sowie die darauf aufbauende Folgeaktivität der BDD-Implementierung können grundsätzlich unabhängig von anderen Aktivitäten durchgeführt werden. Wichtig ist die rechtzeitige Bereitstellung des BDD-Tools mit Beginn von Aktivität MF1: III.3, in der die Geschäftsobjekttypen definiert und die fachlichen Metadaten angelegt werden.

5.5.1.6 BDD implementieren

Aufbauend auf dem spezifizierten Metadatenmodell kann das BDD implementiert werden. Voraussetzung für die Implementierung ist die Kenntnis der Rahmenbedin-

gungen für die BDD-Nutzung. Diese umfassen u. a. die Anzahl der Nutzer, das Rollenkonzept und die Anzahl der zu pflegenden Geschäftsobjekttypen. Die Anforderungen sind zu Beginn der Aktivität aufzunehmen und in Form von Kriterien für die Evaluation möglicher BDD-Lösungen zu berücksichtigen. Von besonderer Bedeutung ist die Entwicklung und Implementierung eines langfristigen Rollenkonzeptes für das BDD bzw. für das Metadatenmanagement, das den verschiedenen Nutzergruppen unterschiedliche Rechte für den Zugriff auf fachliche Metadaten gewährt. An dieser Stelle sind die Rollen, die beim Entwurf von Metadatenmanagementprozessen definiert werden (siehe Aktivität MF1: IV.1), zu berücksichtigen.

Bei der Evaluation von Software-Lösungen ist insbesondere die Möglichkeit der Verwendung bereits implementierter Software im Unternehmen (Wiki-Lösungen, MDM- oder EAM-Werkzeuge) zu prüfen. Diese bieten häufig einfache Anpassungsmöglichkeiten, um darin fachliche Metadaten zu Geschäftsobjekttypen mit geringer Komplexität (einfaches Metadatenmodell) zu pflegen.

MF1:III.2 BDD implementieren	
Ergebnis	Anforderungsspezifikation/Lastenheft Implementiertes BDD
Rollen	Strategische Datenstewards [verantwortlich] Software-Entwickler [ausführend] Konzern-Datensteward, Fachliche Datenstewards [unterstützend]
Techniken	T.III.2 Bewertung von Software-Lösungen für ein BDD
Input/Vorbedingungen	Metadatenmodell des BDD Rollen des Metadatenmanagements

Tabelle 5-9: Aktivität MF1: III.2 BDD implementieren

Abhängig von der verwendeten Software zur Umsetzung des fachlichen Metadatenkataloges ergeben sich unterschiedliche Anforderungen an die Implementierung (Neueinführung einer Software oder Anpassung bestehender Lösungen) sowie an die Notwendigkeit, potenzielle Nutzer zu schulen. Konkrete Werkzeuge für die Realisierung eines BDD werden im Kapitel zur Werkzeugunterstützung für dieses Methodenfragment (siehe Kapitel 5.5.4) behandelt.

5.5.1.7 Geschäftsobjekttypen definieren

Aktivität A.III.3 ist die Kernaktivität des ersten Methodenfragmentes und gleichzeitig die zeitaufwändigste. Ihr Ziel ist es, die identifizierten Geschäftsobjekttypen und deren Attribute entsprechend der vorgegebenen Klassen des Metadatenmodells (siehe Kapitel 5.5.1.5) zu beschreiben. Hierbei gilt es, die voneinander abweichenden Verständnisse für einen Geschäftsobjekttyp, die aus fachbereichs- oder prozessspezifischen Sichten resultieren, abzustimmen und in Einklang zu bringen. Für die abgestimmte Definition eines Geschäftsobjekttyps werden Workshops bzw. Gruppeninterviews mit den fachlichen Datenstewards der verschiedenen Fachbereiche durchgeführt (siehe Er-

gebnis von Aktivität MF1: II.2). Anhang D.2 enthält einen Interviewleitfaden, der in den Aktionsforschungsprojekten als Vorlage genutzt wurde, um die notwendigen Informationen zur Beschreibung von Geschäftsobjekttypen zu dokumentieren²⁵. Bei der Beschreibung wesentlicher Attribute des Geschäftsobjekttyps ist besonderes Augenmerk auf die Festlegung bzw. Kennzeichnung global gültiger Attribute, die in mehreren Applikationen bzw. Organisationseinheiten verwendet werden (siehe Kapitel 2.2.1), zu legen [Pohland 2000, 153f.]. Die Harmonisierung ist jeweils nur für die global verwendeten Geschäftsobjektattribute durchzuführen.

MF1: III.3 Geschäftsobjekttypen definieren	
Ergebnis	Eindeutig definierte Geschäftsobjekttypen (für jeden Geschäftsobjekttyp ausgefüllter Interviewleitfaden)
Rollen	Strategische Datenstewards [ausführend, verantwortlich] Konzern-Datensteward, Fachliche Datenstewards, Metadatenadministrator [ausführend] (falls nötig auch Prozess-Owner) Datenqualitäts-Komitee [genehmigend/freigebend]
Techniken	T.III.3 Workshops / Gruppendiskussion T.II.2 Einzelinterviews
Input/Vorbedingungen	Implementiertes BDD zur Pflege der Metadaten Metadatenmodell des BDD (zu pflegende fachliche Metadaten)

Tabelle 5-10: Aktivität MF1: III.3 Geschäftsobjekttypen definieren

Als Referenz für die eigenen Definitionen können bereits existierende, standardisierte Glossare, die beispielsweise Fachbegriffe einer bestimmten Industrie beschreiben, dienlich sein und den Abstimmungsaufwand zur Definitionsfindung begrenzen. Bietet die implementierte BDD-Lösung Mechanismen zur toolgestützten Moderation kollaborativer Definitionsprozesse, wie dies bei wiki-basierten BDDs der Fall ist (siehe Kapitel 5.5.4), kann die Anzahl der Workshops bzw. Interviews zur Abstimmung der Definitionen reduziert werden.

5.5.1.8 Metadatenmanagementprozesse entwerfen

Inhalt und Struktur des fachlichen Metadatenkataloges unterliegen im Geschäftsverlauf regelmässigen Änderungen. Unternehmensfusionen, die Anpassung der Geschäftsstrategie oder die Einführung neuer Applikationen und Produkte beeinflussen die Verantwortlichkeiten für Geschäftsobjekttypen sowie die stammdatenführenden Applikationen (inhaltliche Veränderung auf Ebene der Metadateninstanz) und führen zur Notwendigkeit, für Geschäftsobjekttypen zusätzliche Metadaten zu pflegen (strukturelle Veränderung auf Ebene des Metadatenmodells) [Blechar/White 2009, 6; Khathri/Brown 2010, 151]. Diese Beispiele verdeutlichen, dass das Management der

²⁵ Die in der Vorlage enthaltenen Daten entstammen einem Projekt bei der ETA S.A. (siehe Kapitel 5.5.2) und beschreiben den Geschäftsobjekttyp „Uhrwerk“.

Metadaten²⁶ eine kontinuierliche und dauerhafte Aufgabe ist, die einen definierten Metadatenmanagement-Prozess und flexible Strukturen zu dessen Umsetzung bedingen [Tannenbaum 2001, 429ff.; Auth 2003, 3].

Die Spezifikation der Metadatenmanagementprozesse umfasst die Definition einzelner Aufgaben (u. a. Anlage bzw. Erzeugung, Änderung, Löschung bzw. Bereinigung fachlicher Metadaten) und ihrer Reihenfolge sowie die langfristige Festlegung der Verantwortlichkeiten für diese Aufgaben. Da ein Metadatenmanagementprozess pro Geschäftsobjekttyp spezifiziert wird, müssen die Geschäftsobjekttypen vorher zumindest identifiziert sein. Die Aufgaben des Prozessentwurfs können sich an etablierten Techniken des Business Process (Re-)Engineering orientieren. So lässt sich beispielsweise die Entwurfstechnik der Ablaufplanung der Methode PROMET-BPR für den Entwurf betrieblicher Prozesse (vgl. [Hess 1996, 202-210; IMG 1997, 230-244]) in vereinfachter Form auf den Entwurf von Metadatenmanagementprozessen übertragen.

MF1: IV.1 Metadatenmanagementprozesse entwerfen	
Ergebnis	Definierte Metadatenmanagementprozesse (Aktivitätsdiagramm, Aufgabenverzeichnis, Rollenmodell)
Rollen	Strategische Datenstewards [verantwortlich] Konzern-Datensteward [ausführend] Datenqualitäts-Komitee [genehmigend/freigebend]
Techniken	T.IV.1 Ablaufplanung
Input/Vorbedingungen	Identifizierte Geschäftsobjekttypen

Tabelle 5-11: Aktivität MF1: IV.1 Metadatenmanagementprozesse entwerfen

Die Aktivitäten zum Entwurf (MF1: IV.1) sowie zur Implementierung (MF1: IV.2) der Metadatenmanagementprozesse müssen nicht zwangsläufig, sollten aber spätestens nach der erstmaligen Definition von Geschäftsobjekttypen im BDD durchgeführt werden. Eine vorherige, parallel zur BDD-Implementierung (MF1: III.2) laufende Ausführung ist zu empfehlen, um die sich aus den Metadatenmanagementprozessen ergebenden Anforderungen an eine Werkzeugunterstützung (z. B. Zugriffs- und Änderungsrechte verschiedener Rollen) möglichst frühzeitig berücksichtigen zu können (siehe gestrichelter Pfeil in Abbildung 5-6).

5.5.1.9 Metadatenmanagementprozesse implementieren

Wesentliches Instrument für die Implementierung der Metadatenmanagementprozesse ist die Etablierung eines formalisierten und kontrollierten Prozesses für die Bearbeitung von Änderungsanforderungen (sogenannte Change Requests). Diese sichern Aktualität und Richtigkeit der fachlichen Metadaten im BDD und sind ein wichtiges

²⁶ Das Metadatenmanagement „bezeichnet die Gesamtheit der Aktivitäten, die den koordinierten, nutzenorientierten Umgang mit den Metadaten einer Anwendungsdomäne gewährleisten und dauerhaft in der Unternehmensorganisation verankern“ [Melchert 2006, 40].

Werkzeug für die Integration von Stammdaten- und Metadatenmanagement (siehe Abbildung 5-7).

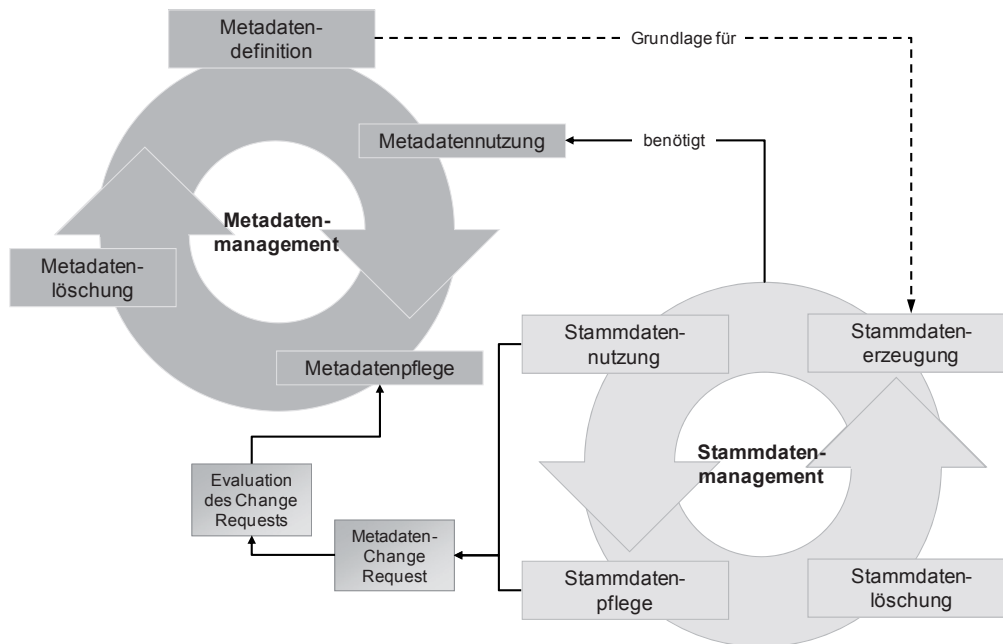


Abbildung 5-7: Integration von Stamm- und Metadatenmanagement

Die Änderungsanforderungen sollten von jedem Stammdatennutzer mittels Ausfüllen eines Change-Request-Templates (z. B. über das Intranet) angestossen werden können, in dem der Antragsteller Geschäftsobjekttyp und -attribut sowie Änderungswunsch formuliert und an den zuständigen strategischen Datensteward weiterleitet (siehe Anhang D.3). Der Change-Request-Prozess hängt dabei wesentlich von der implementierten BDD-Lösung ab (siehe Kapitel 5.5.4).

	MF1: IV.2 Metadatenmanagementprozesse implementieren
Ergebnis	Implementierte Metadatenmanagementprozesse (inklusive Change Requests)
Rollen	Strategische Datenstewards [verantwortlich] Konzern-Datensteward, Metadatenadministrator [ausführend] Datenqualitäts-Komitee [genehmigend/freigebend]
Techniken	T.IV.2 Definition von Change-Request-Prozessen
Input/Vorbedingungen	Definierte Metadatenmanagementprozesse (Aktivitätsdiagramm, Aufgabenverzeichnis, Rollenmodell) Implementiertes BDD

Tabelle 5-12: Aktivität MF1:IV.2 Metadatenmanagementprozesse implementieren

Zusätzlich zur Implementierung der Change-Request-Prozesse muss gewährleistet werden, dass fachliche Datenstewards die Metadaten der ihnen zugewiesenen Geschäftsobjekttypen in regelmässigen Abständen (z. B. halbjährlich) auf ihre Aktualität überprüfen. Dies ist notwendig, um veraltete, nicht mehr gültige Metadaten aus dem BDD zu entfernen. Sowohl für die periodische Überprüfung als auch für die Änderung

infolge von Change Requests ist eine automatisierte Benachrichtigung der Metadatenpfleger (z. B. per E-Mail) zu garantieren.

5.5.2 Referenz-Metadatenmodell für ein BDD

Wie das Metamodell der Methode (siehe Abbildung 5-4) verdeutlicht, speichert und verwaltet ein BDD (fachliche) Metadaten. Dies führt zu der Frage, welche Metadaten in einem Unternehmen für Geschäftsobjekttypen gepflegt werden müssen, um diese semantisch eindeutig zu beschreiben und gleichzeitig die korrekte Verwendung der Geschäftsobjekte in den Geschäftsprozessen zu gewährleisten [Adelman et al. 2005, 8]. Diese sollten in einem globalen *Metadatenmodell* spezifiziert werden, welches den Kern des fachlichen Metadatenkatalogs darstellt. Als Hilfsmittel für diese Aufgabe leitet das Kapitel ein Referenz-Metadatenmodell für ein BDD ab, das als Grundlage für unternehmensspezifische Metadatenmodelle genutzt werden kann, indem relevante Metadatenklassen selektiert werden.

Für die Ableitung des Referenz-Metadatenmodells wurden vier verschiedene Quellen verwendet:

- Internationale Standards, welche die Verwendung von Metadaten zur eindeutigen Definition von Objekten allgemein bzw. speziell von Geschäftsobjekten und Datenelementen regeln;
- Wissenschaftliche und praxisorientierte Publikationen;
- Erkenntnisse aus den Fallstudien der Dissertation;
- Ergebnisse von Fokusgruppeninterviews mit Fachexperten des Kompetenzzentrums CDQ.

Internationale Metadatenstandards

Die Kategorie der internationalen Metadatenstandards beschränkt sich nicht ausschließlich auf die Beschreibung von Geschäftsobjekten, sondern allgemein von Informationsressourcen (z. B. Geschäftsdokumente bzw. geschäftsdokumentartige Informationsobjekte). PÄIVÄRINTA ET AL. liefern in ihrer Arbeit einen umfassenden Vergleich von insgesamt 19 Metadatenstandards zur Beschreibung von Geschäftsdokumenten (vgl. [Päivärinta et al. 2002, 1155-1158]), von denen der Dublin Metadata Core Element Set (vgl. [DCMI 2008]) sowie der Australian Government Locator Service (vgl. [AGLS 2002]) die weiteste Verbreitung haben. Ausserdem spezifiziert der ISO/IEC-Standard 11179 als internationaler Standard zur Abbildung von Metadaten eines Unternehmens in einem Metadatenkatalog in Teil 3, welche Attribute und Beziehungen das Metamodell eines Metadatenkatalogs enthalten sollte, um Datenelemente umfassend zu definieren (siehe Kapitel 2.3.1.4). Die in den Standards festgelegten Me-

tadatenattribute können – abgesehen von den dokumentenspezifischen – als Referenz für den vorliegenden Anwendungsfall verwendet werden.

Der Vergleich der drei genannten Standards führt zu einer Liste von 16 Attributen zur Beschreibung von Metadaten (siehe Tabelle 5-13), die potenzielle Kandidaten für das Referenz-Metadatenmodell eines BDD sind.

Metadatenattribut	Beschreibung	Quelle
Coverage / Context	Eingrenzung des vom Objekt/Element inhaltlich abgedeckten Bereichs bzw. des Bereiches, in dem es verwendet wird	AGLS, Dublin Core, ISO/IEC 11179-3
Data type	Datentyp des Objektes/Elementes	Dublin Core, ISO/IEC 11179-3
Date	Zeitpunkt der Erstellung bzw. letzten Änderung des Objektes/Elementes	AGLS, Dublin Core
Definition / Description	Kurze Beschreibung des Objektes/Elementes bzw. seines Inhaltes	AGLS, Dublin Core, ISO/IEC 11179-3
Examples	Beispiele für Ausprägungen des Objektes/Elementes	AGLS
Function	Geschäftsfunktion, zu der das Objekt/Element gehört	AGLS
Identifier	Eindeutiger Bezeichner für ein Objekt/Element	AGLS, Dublin Core, ISO/IEC 11179-3
Max. Occurrence	Anzahl der Werte, die Objekt/Element annehmen kann	Dublin Core, ISO/IEC 11179-3
Name / Title	Name des Objektes/Elementes, unter dem es bekannt ist	AGLS, Dublin Core, ISO/IEC 11179-3
Obligation	Verbindlichkeit des Objektes/Elementes: zwingend notwendig (obligatorisch), bedingt notwendig oder optional	Dublin Core, ISO/IEC 11179-3
Producer(s) / Creator(s)	Organisationseinheit, Rolle oder Person, die Inhalt des Objekt/Element erstellt und somit dafür verantwortlich ist	AGLS, Dublin Core, ISO/IEC 11179-3
Purpose and comments	Begründung, warum Objekt/Element notwendig ist, sowie weitere Anmerkungen zu seiner Verwendung	Dublin Core, ISO/IEC 11179-3
Relation / Sub-elements	Andere Objekte/Elemente, mit denen das zu beschreibende Objekte/Elemente in Beziehung steht	Dublin Core, AGLS
Source	Verweis auf Ursprung des Objektes/Elementes	Dublin Core
Subject / Keyword	Thema bzw. Schlagwort, dem Objekt/Element zugeordnet ist	AGLS, Dublin Core
Value domain	Name einer Werteliste bzw. einer Menge möglicher Werte, die dem Objekt/Element zugeordnet werden kann	AGLS, ISO/IEC 11179-3

Tabelle 5-13: Metadatenattribute der drei Metadatenstandards

Wissenschaftliche und praxisorientierte Publikationen

In wissenschaftlichen und praxisorientierten Publikationen werden (fachliche) Metadaten vor allem im Umfeld von Data Warehouses sowie bei der Erstellung einheitlicher Fachbegriffsmodelle diskutiert. Entsprechend dienen sie der Beschreibung von Fachbegriffen (vgl. [Auth 2003; O'Neil 2005]) und Datenelementen (vgl. [Adelman et al. 2005; Inmon et al. 2008]). Tabelle 5-14 zeigt eine gegenüberstellende Analyse der in Veröffentlichungen genannten fachlichen Metadatenattribute.

Metadatenattribut	[Auth 2003]	[DAMA 2009]	[Inmon et al. 2008], [O'Neil 2005]	[Marco 2000], [Marco/Michael 2004]	[Ortner et al. 1990]	[Adelman et al. 2005]
Eindeutiger Name, Fachbegriff	X	X	X	X	X	X
Gängige Abkürzung		X			X	
Definition	X	X	X	X	X	X
Alternative / ältere Definitionen		X				
Attribute (inkl. Definition, Datentyp)	X	X	X	X		X
Identifizierende Attribute (Distinguishing Characteristics)			X	X	X	
Gültigkeit (global vs. lokal)			X			
Beziehungen zu anderen Entitäten	X		X	X		X
Synonyme	X		X	X	X	
Homonyme	X					
Ownership, Verantwortung (Rolle bzw. Organisationseinheit)	X	X	X	X	X	X
Nutzer (Rolle bzw. Organisationseinheit mit Rechten)	X					
Verwendung (Geschäftsprozess, Unternehmensfunktion)	X	X	X	X	X	X
Ursprungsapplikation	X	X	X			X
Führendes System (system of record)	X	X	X	X		X
Geschäftsregel		X	X	X		X
Gültige Werte, Beispielinstanzen		X	X	X	X	
Regulatorische Einschränkungen		X				
Vertraulichkeitsstufe	X					
Letzte Änderung, Änderungshistorie	X	X	X			
Derzeitiger Status				X		

Tabelle 5-14: Metadatenattribute zur Beschreibung von Geschäfts-/Datenobjekttypen

Fallstudien und Fokusgruppeninterviews

Schliesslich wurden für die Ableitung des Referenz-Metadatenmodells die aufgenommenen Fallstudien (siehe Kapitel 4.2) bezüglich der dort verwendeten fachlichen Metadaten zur Beschreibung von Geschäftsobjekten ausgewertet sowie die Ergebnisse von Fokusgruppeninterviews mit Fachexperten²⁷ aus dem Bereich MDM und DQM berücksichtigt.

²⁷ Das Fokusgruppeninterview wurde am 16. November 2007 auf dem 5. CC CDQ-Workshop durchgeführt.

Aus den Fallstudien konnten folgende Erkenntnisse gewonnen werden. Die SBB Cargo AG beschreibt ihre Geschäftsobjekttypen im Glossar anhand der in Tabelle 5-15 aufgeführten Metadatenattribute (ein Ausschnitt aus dem Glossar ist in Tabelle 4-2 abgebildet). Zusätzlich werden im UML-Klassendiagramm für jeden Geschäftsobjekttyp die wichtigsten Attribute sowie die Beziehungen zu anderen Geschäftsobjekttypen durch Assoziationen mit ihren Kardinalitäten spezifiziert.

Metadatenattribut	Beschreibung
Name	Eindeutiger Bezeichner des Geschäftsobjekttyps
Definition	Kurze, verbale Beschreibung des Geschäftsobjekttyps zur eindeutigen Abgrenzung gegenüber anderen
Mögliche Synonyme	Auflistung von Begriffen, die denselben Geschäftsobjekttyp bezeichnen
Daten-Owner	Benennung der für den Geschäftsobjekttyp verantwortlichen Rolle bzw. Person
Business (Unternehmensfunktion)	Name der Unternehmensfunktion(en), die den Geschäftsobjekttyp nutzt
Geschäftsprozess	Auflistung sämtlicher Geschäftsprozesse, die den Geschäftsobjekttyp verwenden Unterscheidung zwischen: <ul style="list-style-type: none"> • Master-Prozess, in dem Geschäftsobjekttyp erzeugt wird (genau ein Geschäftsprozess) und • Nutzer-Prozessen, die auf den Geschäftsobjekttyp zugreifen (mehrere möglich)
Applikation	Auflistung sämtlicher Applikationen, die den Geschäftsobjekttyp verwenden Unterscheidung zwischen: <ul style="list-style-type: none"> • Master-System, in dem Geschäftsobjekttyp erzeugt wird (in der Regel eine Applikation) und • Nutzer-Systemen, die auf den Geschäftsobjekttyp zugreifen (mehrere möglich)

Tabelle 5-15: Metadaten zur Beschreibung von Geschäftsobjekttypen bei SBB Cargo

Tabelle 5-16 fasst die Metadatenattribute bzw. -klassen zusammen, welche die DTAG nutzt, um ihre Geschäftsobjekttypen im Business Object Model (BOM) zu definieren.

Metadatenattribut	Beschreibung
Name	Eindeutiger Bezeichner des Geschäftsobjekttyps
Definition	Kurze, verbale Beschreibung des Geschäftsobjekttyps zur eindeutigen Abgrenzung gegenüber anderen
Kommentar	Zusätzliche Anmerkungen zur Verwendung des Geschäftsobjekttyps
Beziehung zu anderen Geschäftsobjekttypen	<ul style="list-style-type: none"> • Assoziationsname • Rolle • Kardinalität • Name des assoziierten Geschäftsobjekttyps
Geschäftsobjektattribute	<ul style="list-style-type: none"> • Attributname • Typ • Kardinalität • zugehöriger Geschäftsobjekttyp Weitere Beschreibung durch: <ul style="list-style-type: none"> • Assoziation zu anderem Geschäftsobjekttyp • Assoziationsname • Beschreibung • BOM-Version
Wertelisten	<ul style="list-style-type: none"> • Beschreibung der Werteliste (inkl. Beispiel) • Aufzählung zulässiger Werte mit Beschreibung jedes Wertes

Metadatenattribut	Beschreibung
Geschäftsregeln	<ul style="list-style-type: none"> • Beschreibung der Geschäftsregel • betroffene Objekte • zusätzlicher Kommentar

Tabelle 5-16: Metadaten zur Beschreibung von Geschäftsobjekttypen bei der DTAG

Als Grundlage für die Fokusgruppeninterviews diente die vorher beschriebene Analyse der Standards und Publikationen. Die Fachexperten sollten die Bedeutung der Metadaten für ihren Anwendungsfall bewerten und – falls nötig – weitere Metadaten hinzufügen. Als Anforderungen an das Metadatenmodell eines BDD wurden folgende Aspekte besonders hoch bewertet:

- Pflege natürlichsprachlicher Definitionen mit dem Vokabular des Fachbereiches,
- Beschreibung der Beziehungen zu und Abhängigkeiten von anderen Geschäftsobjekttypen,
- Dokumentation von Synonymen und alternativen Definitionen (aber Festlegung auf einen verbindlichen Bezeichner und eine einheitliche Definition),
- Benennung der globalen sowie der identifizierenden Attribute, die unternehmensweit Gültigkeit besitzen (müssen),
- natürlichsprachliche Beschreibung der im Unternehmen für einen Geschäftsobjekttyp gültigen Geschäftsregeln.

Als zusätzliche Notwendigkeit gaben die Fachexperten an, dass die Beziehungen zwischen Geschäftsobjekttypen und Datenobjekttypen bzw. Datenobjektattributen als Integrationsanforderung zwischen Prozess- und Systemebene in einem BDD abgebildet werden sollte. Entsprechend muss das BDD über eine Kategorie „Zugehörige Datenobjektattribute“ verfügen, in der die systemseitigen Repräsentationen eines Geschäftsobjekttyps mit dem entsprechenden Namen der Applikation dokumentiert werden kann.

Ableitung des Metadatenmodells

Aus den verschiedenen Quellen werden im Folgenden die fachlichen Anforderungen an Art und Umfang der in einem BDD gespeicherten Metadaten abgeleitet. Daraufhin wird ein allgemeines Metadatenmodell für ein BDD erstellt, das als Referenz für die Implementierung eines unternehmensspezifischen fachlichen Metadatenkatalogs verwendet werden kann. Das Metadatenmodell bildet die Grundlage für eine umfassende Spezifikation der fachlichen Metadaten zur eindeutigen Definition von Geschäftsobjekttypen. Die Anwendung des Metadatenmodells wird im Anschluss anhand eines Aktionsforschungsprojektes im Rahmen des Kompetenzzentrum CDQ beschrieben.

Abbildung 5-8 fasst in einem vereinfachten Klassendiagramm die fachlichen Anforderungen an die in einem BDD verwalteten fachlichen Metadaten zu einem Geschäftsobjekttyp zusammen. Inhaltlich zusammengehörende Metadatenattribute zur Beschreibung eines Geschäftsobjekttyps sind in Metadatenkategorien gruppiert.

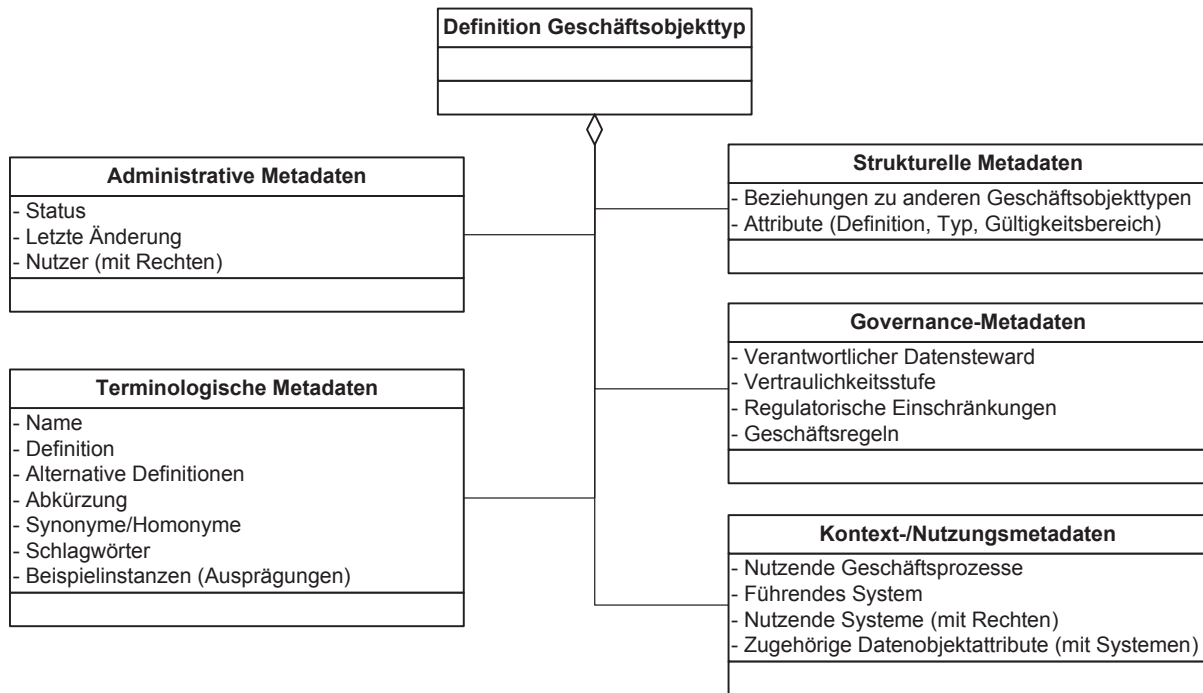


Abbildung 5-8: Fachliche Anforderungen an den Inhalt eines BDD

Die Metadatenkategorien leiten sich aufbauend auf bestehenden Ansätzen zur Strukturierung von Metadaten aus ihrem Verwendungszweck ab [Auth 2003, 47ff.]:

- *Terminologische Metadaten* umfassen diejenigen Metadaten, die der unternehmensweiten Homogenisierung und Vereinheitlichung der Bedeutung (Semantik) von Geschäftsobjekttypen dienen.
- *Strukturelle Metadaten* legen die Struktur der Bestandteile eines Geschäftsobjekttyps (Attribute) sowie die Beziehungen zwischen Geschäftsobjekttypen fest [Müller et al. 2006, 198].
- *Administrative Metadaten* dienen der Verwaltung der Definition von Geschäftsobjekttypen und enthalten Angaben zu Nutzern und ihren Berechtigungen, zum Status sowie zur Änderungshistorie [Shankaranarayan/Even 2004, 254-255].
- *Governance-Metadaten* beinhalten sämtliche Metadaten zu einem Geschäftsobjekttyp, die sich aus regulatorischen Anforderungen, Geschäftsregeln ihrer Verwendung sowie zugeordneter Verantwortung ergeben.
- *Kontextmetadaten* beschreiben in Anlehnung an DEVLIN den *Business Context*, in dem Objekte verwendet werden, und setzen sich aus den drei Dimensionen *Busi-*

ness (Prozessbezug), *Organization* (Organisationsbezug) und *Application Function* (Applikationsbezug) zusammen [Devlin 1997, 277-283].

Aus der Anforderungsanalyse lässt sich im nächsten Entwurfsschritt ein Metdatenmodell als Klassendiagramm ableiten, welches die Grundlage für die Implementierung eines BDD bildet. Ausgangspunkt hierfür war die Synthese der aus den analysierten Quellen (Metadatenstandards, Publikationen, Fallstudien und Experteninterviews) identifizierten Metadatenattribute. Das resultierende Klassendiagramm für die BDD-Implementierung ist in Abbildung 5-9 dargestellt. Eine zusammenfassende Übersicht mit Definitionen sämtlicher Metadatenattribute des Referenz-Metadatenmodells ist in Anhang C.1 beigefügt.

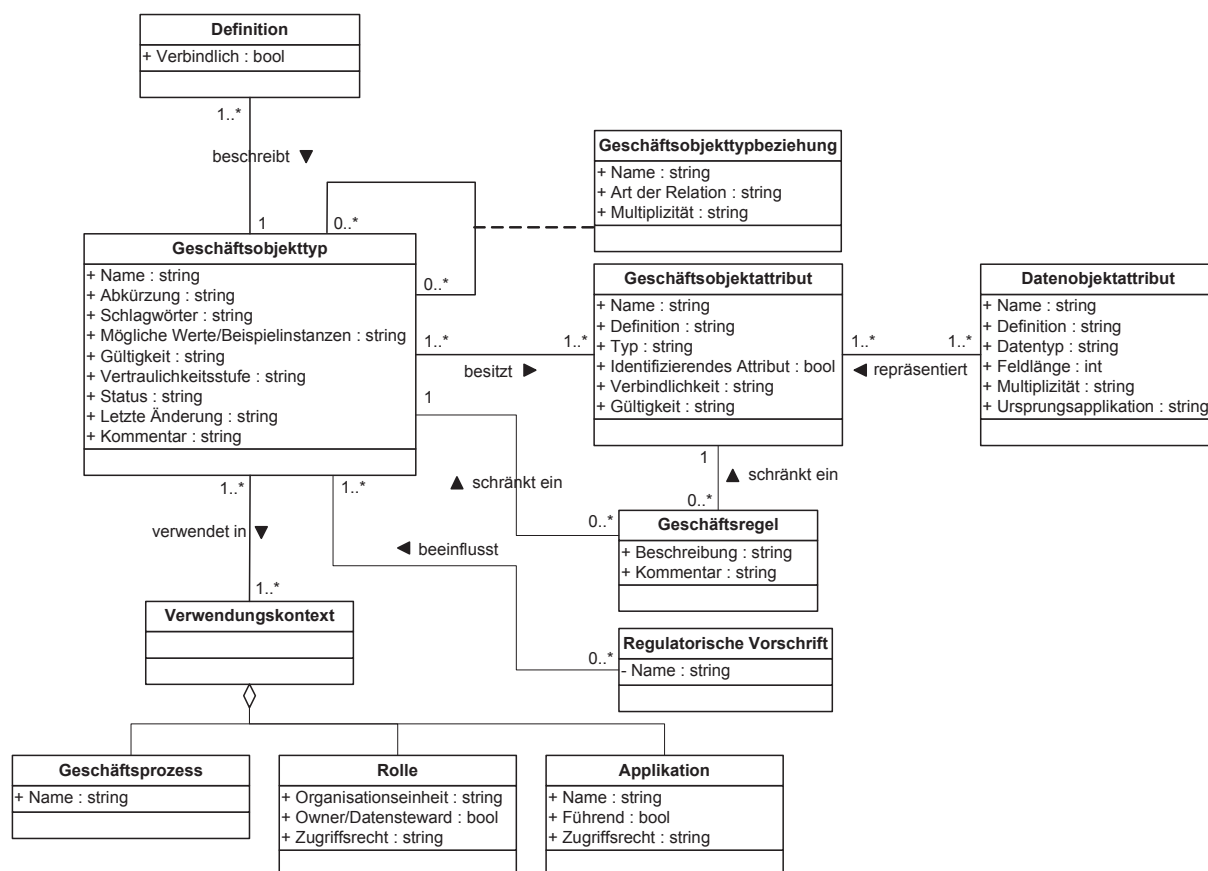


Abbildung 5-9: Klassendiagramm für die Implementierung eines BDD

Die strukturellen Metadaten sind hierbei über die Klassen *Geschäftsobjekttyprelation* (Beziehung zu anderen Geschäftsobjekttypen) sowie über die separate Klasse *Geschäftsobjektattribut* abgebildet. Für die Auslagerung der Geschäftsobjektattribute in eine eigene Klasse sprechen drei Gründe: Erstens können die Geschäftsobjektattribute so detailliert beschrieben werden, zweitens können diese abhängig von ihrem Gültigkeitsbereich als globale oder lokale Attribute charakterisiert werden und drittens können damit *Datenobjektattribute* der Systemebene (siehe Methodenfragment 2) direkt zugeordnet werden. Damit wird die Anforderung erfüllt, die redundante Abbildung

von Geschäftsobjektattributen in den Applikationen zu dokumentieren. Die Art der Geschäftsobjekttyprelation enthält als mögliche Ausprägung den Wert „synonym“, um die synonyme Beziehung zwischen zwei Geschäftsobjekttypen abbilden zu können.

Mit der Abbildung der *Definitionen* in einer gesonderten Klasse wird gewährleistet, dass zu einem Geschäftsobjekttyp mehrere voneinander abweichende Definitionen gepflegt werden können. Dadurch dokumentiert das BDD sowohl unterschiedliche Verständnisse sowie die homonyme Verwendung von Begriffen. Verpflichtend ist jedoch die Angabe einer unternehmensweit verbindlichen Definition. Hervorzuheben ist zudem die Abbildung des *Verwendungskontextes* von Geschäftsobjekttypen. Er setzt sich – wie zuvor beschrieben – aus dem Geschäftsprozess zusammen, in dem die Geschäftsobjekte genutzt werden, aus den Organisationseinheiten und Rollen, die darauf zugreifen bzw. für diese verantwortlich sind, und aus dem Applikationsbezug (führende und nutzende Systeme). Diese fachlichen Metadaten sind insbesondere für die Abbildung im semantischen Stammdatenmodell (siehe Kapitel 5.6) von Bedeutung.

Administrative und Governance-Metadaten sind grösstenteils dem Geschäftsobjekttyp direkt zugeordnet. Aufgrund der Bedeutung der Geschäftsregeln (siehe Anforderungen der Fachexperten) werden diese in einer eigenen Klasse modelliert und können sich sowohl auf Geschäftsobjekttypen als auch auf -attribute beziehen. Ihre Beschreibung im BDD ist in natürlichsprachlicher Form vorgesehen.

Eine detaillierte Beschreibung der einzelnen Klassen und Attribute ist in Anhang C.2 zu finden.

Die ETA S.A., ein schweizerischer Hersteller für Uhrwerke, entwickelte aufbauend auf dem in Abbildung 5-9 dargestellten Referenz-Metadatenmodell eine webbasiertes BDD zur Beschreibung seiner Kerngeschäftsobjekte, insbesondere der Materialien und Artikel. Abbildung 5-10 zeigt einen Screenshot aus dem BDD mit den Metadatenattributen Name, Gültigkeit, Synonyme, Schlagwörter, Vertraulichkeitsstufe (Data Classification), Owner des Geschäftsobjekttyps (Contact), Definition, Kommentar (Rationale) sowie letzte Änderung. Die Metadatenattribute „Descriptive Convention“ und „Coding Convention“ enthalten Geschäftsregeln für die Verwendung und Instanziierung des Geschäftsobjekttyps.

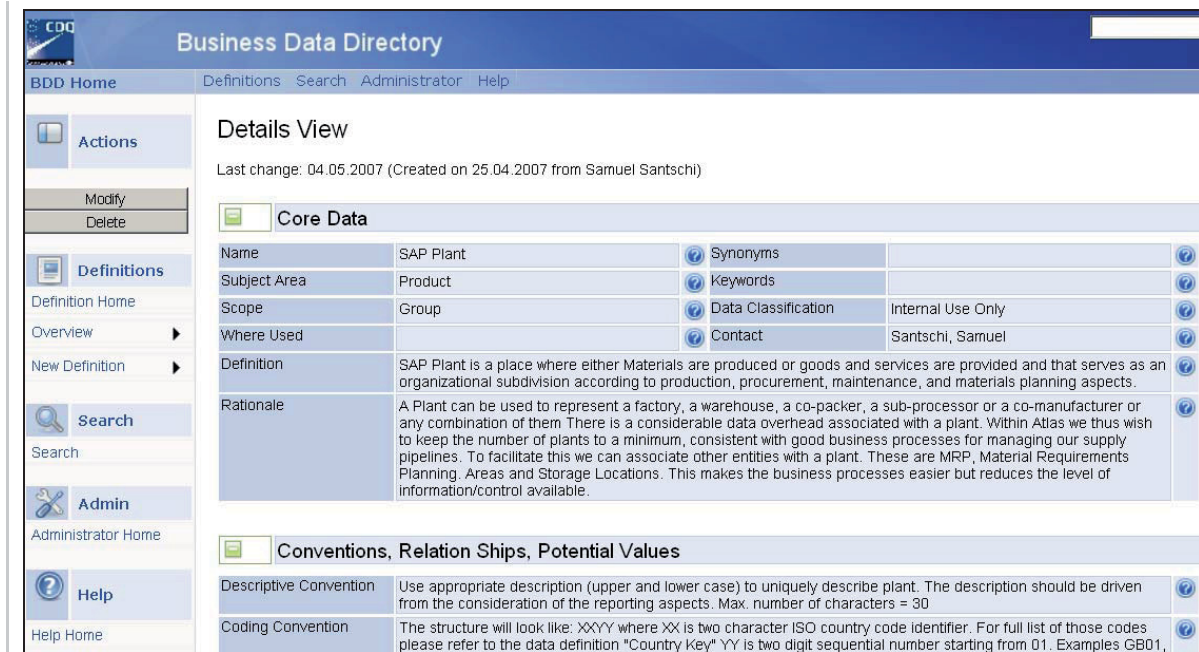


Abbildung 5-10: Screenshot des BDD bei der ETA S.A.

Ein gemäss Klassendiagramm in Abbildung 5-9 vollständig beschriebener Geschäftsobjekttyp ist anhand des ausgefüllten Templates zur Beschreibung von Geschäftsobjekttypen in Anhang D.2 für den Geschäftsobjekttyp „Uhrwerk“ enthalten.

5.5.3 Techniken

Kapitel 5.5.1 beschreibt die Aktivitäten des ersten Methodenfragmentes und verweist auf einzelne Techniken, die als Arbeitsanweisung bei der Ausführung der Aktivität behilflich sind. Hierbei wird teilweise auf Techniken zurückgegriffen, die sich bereits in der Praxis bewährt haben und entsprechend veröffentlicht sind. Tabelle 2-1 zeigt einen Überblick über die in Methodenfragment 1 verwendeten Techniken. Die aus der Literatur entlehnten Techniken sind jeweils referenziert. Das Kapitel erläutert anschliessend die Übertragung und Anwendung der Techniken in der vorliegenden Methode und führt Beispiele an.

Nr.	Name	Aktivität	Referenz
T.I.1	Bewertung der Gestaltungsbereiche	MF1: I.1	[Martin/Leben 1989, 247-249], [IMG 1999, 159-163] ²⁸
T.I.2	Identifikation der Daten- bzw. Prozess-Owner	MF1: I.2	–
T.II.1	Inhaltsanalyse Geschäftsprozessdokumentation	MF1: II.1	[Gläser/Laudel 2009, 197ff.], [Mayring 2008], [Myers 2002]
T.II.2	Einzelinterviews	MF1: II.2, MF1: III.3	[Bortz/Döring 2006, 242-243], [Gläser/Laudel 2009, 111ff.]

²⁸ Ein allgemeines Vorgehen für die Durchführung von Bewertungen ist im PROMET Methodenhandbuch für die System- und Technologieplanung dargestellt (vgl. [IMG 1999, 159-163]). Dieses Vorgehen liegt den Bewertungstechniken dieser Arbeit zugrunde, wird jedoch für den spezifischen Anwendungsfall beschrieben.

Nr.	Name	Aktivität	Referenz
T.II.3	Identifikation von Synonymen und Homonymen / Homogenisierung	MF1: II.2, MF1: III.3	[Brenner 1985, 58ff.], [Strauch 2002, 197-203]
T.II.4	Bewertung der Geschäftsobjekttypen	MF1: II.2	[IMG 1999, 159-163]
T.III.1	Datenmodellierung / Auswahl fachlicher Metadatenklassen	MF1: III.1	–
T.III.2	Bewertung von Software-Lösungen für ein BDD	MF1: III.2	[Hay 2002], [Pohl 2008], [IMG 1999, 159-163]
T.III.3	Workshops / Gruppendiskussion	MF1: III.3	[Bortz/Döring 2006, 319-320], [Cavana et al. 2001, 153-159]
T.IV.1	Ablaufplanung	MF1: IV.1	[IMG 1997, 230-244], [Österle 1995, 61ff.]
T.IV.2	Definition von Change-Request-Prozessen	MF1: IV.2	[OGC 2007, 42ff.]

Tabelle 5-17: Techniken des Methodenfragmentes 1

T.I.1 Bewertung der Gestaltungsbereiche

1. Schritt: Bewertungskriterien festlegen

Um eine systematische Auswahl der potenziellen Gestaltungsbereiche zu ermöglichen, in denen die Methode angewandt werden soll, gilt es in einem ersten Schritt, Kriterien festzulegen. Mögliche Kriterien sind unter anderem [Martin/Leben 1989, 248f.]:

- die Bedeutung des Gestaltungsbereiches (z. B. des Geschäftsprozesses, einer Domäne) für das Unternehmen,
- die Komplexität des Gestaltungsbereiches und der damit verbundene Aufwand für die Methodenanwendung,
- das Nutzenpotenzial sowohl finanziell als auch in Bezug auf die Verbesserung der Datenqualität,
- die abteilungs- und applikationsübergreifende Ausdehnung.

Weitere Bewertungskriterien sowie die Gewichtung der Kriterien sind unternehmensspezifisch festzulegen.

2. Schritt: Bewertung durchführen

Im nächsten Schritt werden die in Frage kommenden Gestaltungsbereiche anhand der definierten Kriterien bewertet bzw. priorisiert. Abhängig von der Anzahl ausgewählter Kriterien sind verschiedene Darstellungsformen möglich. Bei zwei Kriterien zur Bewertung bietet sich eine Matrixdarstellung an (siehe Abbildung 5-11). Ausserdem ist eine tabellarische Auflistung mit Einzelbewertungen für jedes Kriterium möglich, wobei diese zusätzlich gemäss ihrer Bedeutung gewichtet werden können.

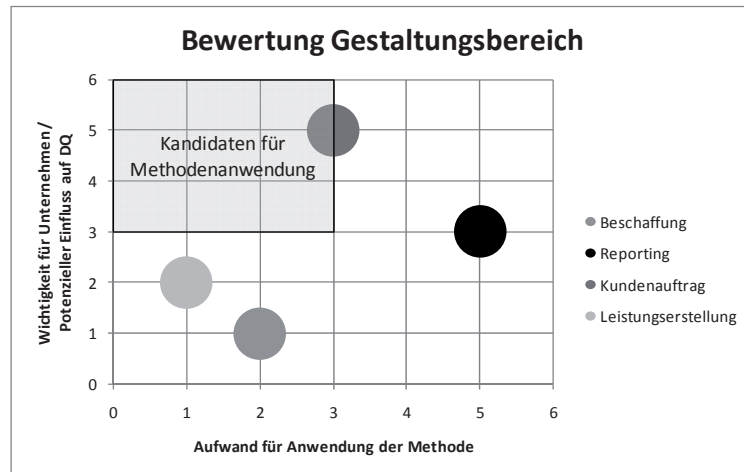


Abbildung 5-11: Matrix zur Bewertung möglicher Gestaltungsbereiche

T.I.2 Identifikation der Daten- bzw. Prozess-Owner

Prozess-Owner können in der Regel (so sie definiert sind) direkt aus einer vorhandenen Geschäftsprozessdokumentation bzw. aus Stellenbeschreibungen identifiziert werden. Besteht keine formal dokumentierte Zuordnung von Rollen zu Geschäftsobjekttypen, so sind potenzielle Owner sowie Nutzer der Geschäftsobjekte am ehesten in Absprache mit dem Prozess-Owner zu ermitteln. Ergebnis der Technik ist eine Auflistung von Rollen oder Personen, die als Akteure in den Geschäftsprozessen Geschäftsobjekte nutzen oder verändern und somit für die Aktivitäten der Identifikation (MF1: II.1) und Beschreibung (MF1: III.3) von Geschäftsobjekttypen von Bedeutung sind. Zudem können sie bei der Geschäftsprozessanalyse (siehe Technik T.II.1 der Aktivität MF1: II.1) als Interviewpartner behilflich sein.

T.II.1 Inhaltsanalyse Geschäftsprozessdokumentation

Die qualitative Inhaltsanalyse beschreibt ein strukturiertes Verfahren zur systematischen Auswertung von Texten und Dokumenten. In der Literatur sind Vorgehensmodelle zur qualitativen Inhaltsanalyse ausführlich beschrieben (vgl. [Mayring 2008; Gläser/Laudel 2009, 197ff.]), weswegen sie an dieser Stelle nicht näher ausgeführt werden. Die Technik sollte zur Auswertung vorhandener Dokumentation zu Geschäftsprozessen genutzt werden. In Aktivität MF1: III.3 unterstützt sie bei der Aufgabe, die Semantik der Geschäftsobjekttypen, d. h. deren Verwendung in Geschäftsprozessen und durch verschiedene Nutzer, aufzudecken. Die Geschäftsprozessdokumentation umfasst u. a. Geschäftsprozessmodelle (Aufgabenkettendiagramme, Aktivitätsdiagramme), Handbücher, Anforderungs- und Rollenspezifikationen, Aufgabenverzeichnisse oder regulatorische Vorschriften.

T.II.2 Einzelinterviews / T.III.4 Workshops / Gruppeninterviews

Sowohl Einzel- als auch Gruppeninterviews zur Identifikation sowie zur Beschreibung von Geschäftsobjekttypen sollten semi-strukturiert durchgeführt werden (vgl. [Cavana

et al. 2001, 153-159]). Semi-strukturierte Interviews setzen die vorherige Entwicklung eines Interviewleitfadens voraus, der als Orientierungshilfe während der Durchführung der Interviews dient [Bortz/Döring 2006, 239]. Dem Ziel der Definition von Geschäftsobjekttypen dienend, sollten die Fragen grösstenteils offen formuliert werden und sich an den Unternehmensanforderungen an fachliche Metadaten orientieren. Folglich können beispielsweise das Metadatenmodell des BDD (siehe Aktivität MF1: III.1), welches diese Anforderungen abbildet, und die darin enthaltenen Metadatenklassen als Hilfe zur Entwicklung des Interviewleitfadens genutzt werden. Dieses Vorgehen gewährleistet die vollständige Beschreibung von Inhalt und Verwendungskontext der Geschäftsobjekttypen und ihrer Attribute.

Die Ergebnisse der Einzelinterviews stellen in der Regel lediglich eine initiale Grundlage für die Diskussion in der Gruppe dar, ohne diese zu ersetzen. Da insbesondere die Festlegung einheitlicher Definitionen für Geschäftsobjekttypen ein kollaborativer Prozess ist, der verschiedene Sichten integrieren muss, ist die Durchführung von Workshops mit Teilnehmern aus verschiedenen Bereichen zwingend notwendig.

In Anhang D.2 ist ein exemplarischer Interviewleitfaden, aufbauend auf dem in Kapitel 5.5.2 hergeleiteten Referenz-Metadatenmodell, enthalten, der für die Befragungen im Rahmen des Aktionsforschungsprojektes bei der Daimler AG (siehe Kapitel 6.1.2) eingesetzt wurde.

T.II.3 Identifikation von Synonymen und Homonymen / Homogenisierung

1. Schritt: Potenzielle Synonyme identifizieren

Synonymie bezeichnet die gleiche Bedeutung verschiedener sprachlicher Ausdrücke. Bei Geschäftsobjekttypen (oder deren Attribute) liegt Synonymie dann vor, wenn sie mit demselben Inhalt mehrere verschiedene Bezeichner besitzen [Brenner 1985, 61; Batini et al. 1986, 344]. Der Inhalt eines Geschäftsobjekttyps setzt sich aus der Summe der Merkmale, also der dem Typ zugeordneten (fachlichen) Metadaten, zusammen. Folglich sind zwei Geschäftsobjekttypen (oder deren Attribute) synonym, wenn sie in ihren fachlichen Metadaten übereinstimmen. Da diese fachlichen Metadaten im BDD natürlichsprachlich gepflegt werden, ist eine eindeutige automatisierte Identifikation von Synonymen nicht möglich [Brenner 1985, 64]. Entsprechend müssen synonyme Geschäftsobjekttypen im Zuge der Definition durch die beteiligten Rollen ermittelt werden. Synonymkandidaten sind in der Regel bereits während der Identifikation bzw. Definition der Geschäftsobjekttypen durch die fachlichen Datenstewards erkennbar. Um definitiv entscheiden zu können, ob es sich bei den Kandidaten tatsächlich um Synonyme handelt, sind die fachlichen Metadaten der Synonymkandidaten miteinander zu vergleichen.

Synonyme müssen nicht zwingend zu Problemen führen, wichtig ist die Kenntnis über ihre Existenz und ihre Kontrolle [Lehmann/Jaszewski 1999, 2]. Werden im Zuge der

Definition Synonyme identifiziert, ist festzulegen, welcher Bezeichner fortan verbindlich zu verwenden ist, um eine homogene Begrifflichkeit zu gewährleisten. Die synonym verwendeten Begriffe werden in dem entsprechenden Metadatenattribut im BDD gepflegt (siehe Referenz-Metadatenmodell in Abbildung 5-9).

2. Schritt: Potenzielle Homonyme identifizieren

Homonymie ist dann gegeben, wenn derselbe Bezeichner für Geschäftsobjekttypen (oder deren Attribute) mit verschiedener Bedeutung verwendet wird, was zwangsläufig zu Inkonsistenzen führt [Brenner 1985, 144-145; Batini et al. 1986, 344]. Analog zur Identifikation von Synonymen können Homonyme nur bestimmt werden, indem im Fall gleichlautender Bezeichner fachliche Datenstewards durch paarweisen Vergleich überprüfen, ob die Beschreibungen der Geschäftsobjekttypen (also die fachlichen Metadaten) übereinstimmen. Homonym verwendete Benennungen müssen beseitigt und durch präzisere, eindeutige Bezeichner differenziert werden.

Ein Beispiel für einen homonym verwendeten Begriff ist der Geschäftsobjekttyp Gleis in Abbildung 2-6. Zwar sprechen die verschiedenen Geschäftsbereiche der DB Netz AG alle von einem Gleis, während sich die Anlagenbuchhaltung jedoch auf die anlagenbuchhalterischen Attribute (wie z. B. die Streckenkostenstelle) beziehen, sind für den Bereich Infrastruktur lediglich Attribute wie der Technische Platz oder die Länge von Bedeutung.

Die Identifikation von Homonymen und Synonymen ist nicht nur bei der Erstellung eines BDD von Bedeutung, sondern auch im bei der Pflege der Metadaten eine fortwährende Aufgabe. So ist bei der Überprüfung der Richtigkeit eines Änderungsantrages für fachliche Metadaten (siehe Aktivität MF1: IV.2) darauf zu achten, dass für homonyme Geschäftsobjekttypen oder -attribute neue, eindeutige Bezeichner gefunden bzw. synonyme Bezeichner in den fachlichen Metadaten des verbindlich vorgegebenen Geschäftsobjekttypen dokumentiert werden.

Die Durchführung der Synonym- und Homonymidentifikation ist analog bei der Ableitung des semantischen Stammdatenmodells in Methodenfragment 2 für Datenobjekttypen durchzuführen. Während die Bestimmung auf Ebene der Geschäftsobjekttypen relativ einfach zu bewerkstelligen ist, empfiehlt sich bei der Identifikation synonyme und homonyme Datenobjektattribute aufgrund der grösseren Anzahl eine werkzeugunterstützte Suche.

T.II.4 Bewertung der Geschäftsobjekttypen

1. Schritt: Bewertungskriterien definieren

Für die Bewertung der zu definierenden Geschäftsobjekttypen sind in einem ersten Schritt Kriterien zu bestimmen. Dies können beispielsweise sein:

- die Wichtigkeit des Geschäftsobjekttyps für das Unternehmen, die z. B. anhand der Kritikalität, der Gültigkeitsbereiches im Unternehmen oder der Anzahl der Instanzen bewertet werden kann,
- das erwartete Potenzial zur Verbesserung der Datenqualität,
- Verfügbarkeit von Dokumentation.

2. Schritt: Bewertung durchführen

Im nächsten Schritt werden die identifizierten Geschäftsobjekttypen anhand der definierten Kriterien bewertet bzw. priorisiert. Analog zu Technik T.I.1 sind abhängig von der Anzahl ausgewählter Kriterien verschiedene Darstellungsformen möglich. Tabelle 5-18 zeigt beispielhaft eine tabellarische Auflistung mit Bewertungen einzelner Geschäftsobjekttypen aus dem Aktionsforschungsprojekt DB Netz AG (siehe Kapitel 6.1.1). Die Kriterien wurden in diesem Fall nicht gewichtet.

Geschäftsobjekttyp	Bewertungskriterien			Summe (Prioritätskennzahl)
	Kritikalität	Anzahl Instanzen (Mengengerüst)	Mögliche DQ- Verbesserung	
Bahnübergang	8	6	6	20
Betriebsstellen	4	3	3	10
Brücken	8	7	5	20
Gleis	9	10	10	29
Hauptsignal	5	4	2	11
Kilometrierungssprünge	5	3	3	11
Stellwerke	7	4	3	14
Strecken	7	7	5	19
Tunnel	9	2	7	18
Weichen	9	10	7	26

Tabelle 5-18: Bewertung von Geschäftsobjekttypen anhand definierter Kriterien

T.III.1 Datenmodellierung bzw. Auswahl fachlicher Metadatenattribute

Da der Modellierungsprozess zur Erstellung des Metadatenmodells sehr aufwendig sein kann, wird in Kapitel 5.5.2 ein Metadatenmodell für ein BDD vorgeschlagen, das als Referenz durch Entfernen oder Hinzufügen einzelner Metadatenklassen für die Ableitung eines unternehmensspezifischen Metadatenmodells verwendet werden kann.

T.III.2 Bewertung von Software-Lösungen für ein BDD

1. Schritt: Anforderungen definieren

Ziel der Anforderungsanalyse ist es, eine Liste der Anforderungen an ein BDD zu erstellen (Anforderungsspezifikation). Das Vorgehen für eine Anforderungsanalyse ist

in der Literatur zur Software- und Systementwicklung in zahlreichen Quellen beschrieben (vgl. [Hay 2002; Pohl 2008]). Aufgrund des auf fachliche Metadaten beschränkten Umfangs der zu implementierenden Applikation können die Schritte zur Erhebung der Anforderungen in vereinfachter Form durchgeführt werden:

- Anwendungsbereich definieren,
- Information sammeln (in Absprache mit vorgesehenen Nutzern),
- Anforderungen beschreiben,
- Einführung der Applikation planen.

Der Aufwand für die Anforderungsanalyse kann zudem durch die Verwendung von Ergebnissen vorheriger bzw. parallel laufender Aktivitäten (z. B. des in Aktivität MF1: III.1 entwickelten Metadatenmodells) reduziert werden.

Neben den funktionalen Anforderungen an ein BDD, wie z. B. effiziente Navigations- und Suchmechanismen oder die Benachrichtigung von Nutzern zur Aktualisierung von Metadaten, sind auch die nicht-funktionalen Anforderungen zu dokumentieren. Hierzu zählen u. a. die Integrationsfähigkeit in die bestehende Applikationsarchitektur, die einfache Erweiterbarkeit des Metadatenmodells sowie geringe Implementierungskosten. Zusätzlich sind die Anforderungen, die sich aus der Definition der Metadatenmanagementprozesse ergeben (Nutzerrollen, Änderungsrechte), zu berücksichtigen. Eine Liste funktionaler und nicht-funktionaler Anforderungen an ein BDD, welche im Rahmen des Kompetenzzentrums CDQ gemeinsam mit den Partnerunternehmen erarbeitet wurde, beschreibt Kapitel 5.5.4 zur Werkzeugunterstützung für dieses Methodenfragment. Die Liste kann als Referenz für die Definition unternehmensspezifischer Anforderungen genutzt werden.

2. Schritt: Bewertung durchführen

Auf Grundlage der definierten Anforderungen bewerten Mitarbeiter der IT gemeinsam mit Vertretern des Fachbereiches potenzielle Software-Lösungen und prüfen Realisierungsoptionen für ein BDD. Grundsätzlich ist hierbei im Sinne einer Make-or-Buy-Entscheidung festzulegen, ob das Unternehmen das BDD intern erstellen kann (Eigenentwicklung) oder eine kommerziell erhältliche Software erwirbt (Fremdbezug). Ausserdem ist zu prüfen, inwiefern bereits im Unternehmen existierende Software-Tools, wie z. B. MDM- oder EAM-Werkzeuge, Funktionalitäten zum Management fachlicher Metadaten für Geschäftsobjekte bieten und somit mit geringen Kosten und wenig Aufwand als BDD verwendet werden können. Einen Überblick über BDD-Funktionalität in MDM-Tools sowie eine Möglichkeit, ein BDD mit geringem Aufwand intern zu erstellen, zeigt Kapitel 5.5.4 zur Werkzeugunterstützung.

T.IV.1 Ablaufplanung

Die Entwurfstechnik der Ablaufplanung der Methode PROMET-BPR für den Entwurf betrieblicher Prozesse (vgl. [Hess 1996, 202-210; IMG 1997, 230-244]) in vereinfachter Form auf den Entwurf von Metadatenmanagementprozessen übertragen. In Analogie zu den beschriebenen Aufgaben bei der Definition des Metadatenmanagementprozesses fordert die Methode ein systematisches Vorgehen in drei Schritten: 1) Identifizieren von Aufgaben, 2) Festlegen der Aufgabenreihenfolge und 3) Zuordnen von Aufgabenträgern (Verantwortlichen) zu den einzelnen Aufgaben. Dementsprechend sind die wesentlichen Ergebnisse der Ablaufplanung das Aufgabenketten- oder Aktivitätsdiagramm zur Dokumentation des Prozessablaufes, das Aufgabenverzeichnis zur verbalen Beschreibung der Aufgaben sowie das Rollenmodell, welches die personellen Aufgabenträger charakterisiert [IMG 1997, 230]. Für die Ablaufplanung der Metadatenmanagementprozesse ist eine Spezifikation der Abläufe auf Makroebene ausreichend. Die Ergebnisse sind insbesondere bei Entwurf und Implementierung des BDD (siehe Aktivität MF1: III.1) zu berücksichtigen.

In Bezug auf das Rollenmodell für das Metadatenmanagement (und somit für die BDD-Nutzung) können im Wesentlichen drei Rollen unterschieden werden [Auth 2003, 181]:

- Der *Metadatennutzer* hat ausschliesslich lesenden Zugriff auf die fachlichen Metadaten, kann aber auch Change Requests für die Veränderung bestimmter Metadaten initiieren. Sämtliche Mitarbeiter (insbesondere Fachbereichsvertreter) des Unternehmens können diese Rolle einnehmen.
- Der *Owner* des Geschäftsobjektyps entscheidet über Änderungsanforderungen (Change Requests) an die im BDD gespeicherten, fachlichen Metadaten. Zusätzlich überprüft er in regelmässigen Abständen die Richtigkeit und Aktualität der fachlichen Metadaten. In der Regel können die strategischen Datenstewards diese Rolle erfüllen, je nachdem, welche Rolle die fachliche Verantwortung für bestimmte Geschäftsobjektypen trägt.
- Der *Metadatenadministrator* sorgt nach Freigabe der Änderungen durch den Owner des Geschäftsobjektyps dafür, dass die fachlichen Metadaten im BDD entsprechend aktualisiert und gepflegt werden. Die fachlichen Datenstewards bzw. der für die Modellierung zuständige Datenarchitekt können diese Rolle übernehmen.

Die SBB Cargo hat für ihre Metadatenpflegeprozesse vier Gruppen definiert (siehe Kapitel 4.2.1.3). Die Rolle der sogenannten Auftraggeber kann durch sämtliche Mitarbeiter, die Metadaten nutzen, ausgefüllt werden. Sie können Änderungsanforderungen für Metadaten erstellen. Die Notwendigkeit zur Änderung wird durch eine

zentrale Prüfstellung in Zusammenarbeit mit dem Owner des Geschäftsobjekttyps geprüft und durch den Owner freigegeben. Die Mutationsstelle pflegt die beschlossenen Änderungen schliesslich als zentrale Organisationseinheit zur Metadatenpflege in Glossar und Geschäftsdatenmodell ein.

Bei der DTAG übernehmen die Data Architects als Verantwortliche für Datenmodelle und fachliche Metadatenkataloge die Rolle der Metadatenadministratoren (siehe Kapitel 4.2.2.3). Sie führen Änderungen an den Metadaten aus, nachdem diese durch die Datenverantwortlichen, die Owner der Geschäftsobjekttypen sind, freigegeben wurden.

T.IV.2 Definition von Change-Request-Prozessen

Das Thema des Managements von Änderungsanforderungen (engl.: Change Requests) wird ausführlich in der IT Infrastructure Library (ITIL) des Office of Government Commerce [OGC] behandelt [OGC 2007, 42ff.]. Ziel ist die Definition eines standardisierten Vorgehens, um Änderungsanfragen an die IT-Infrastruktur (bei ITIL insbesondere Services) schnell und effizient umzusetzen. Im Rahmen der Methode dieser Arbeit müssen Change-Request-Prozesse definiert werden, um ein effizientes Hinzu-fügen, Ändern und Löschen fachlicher Metadaten im BDD zu gewährleisten.

Das in ITIL formulierte Vorgehen für das Management von Änderungsanforderungen kann auf das Metadatenmanagement übertragen werden. ITIL definiert folgende Aktivitäten [OGC 2007, 49]:

- Erstellen eines Change Requests,
- (Formale) Prüfung des Change Requests,
- (Inhaltliche) Bewertung des Change Requests,
- Freigabe des Change Requests,
- Planung und Durchführung der Implementierung des Change Requests,
- Dokumentation und Abschluss des Change Requests.

Ein an dieses Vorgehen angelehnter Prozess für die Bearbeitung von Änderungsanforderungen am Beispiel des Metadatenmanagementprozesses *Anlegen/Ändern fachlicher Metadaten* ist als UML-Aktivitätsdiagramm in Abbildung 5-12 dargestellt.

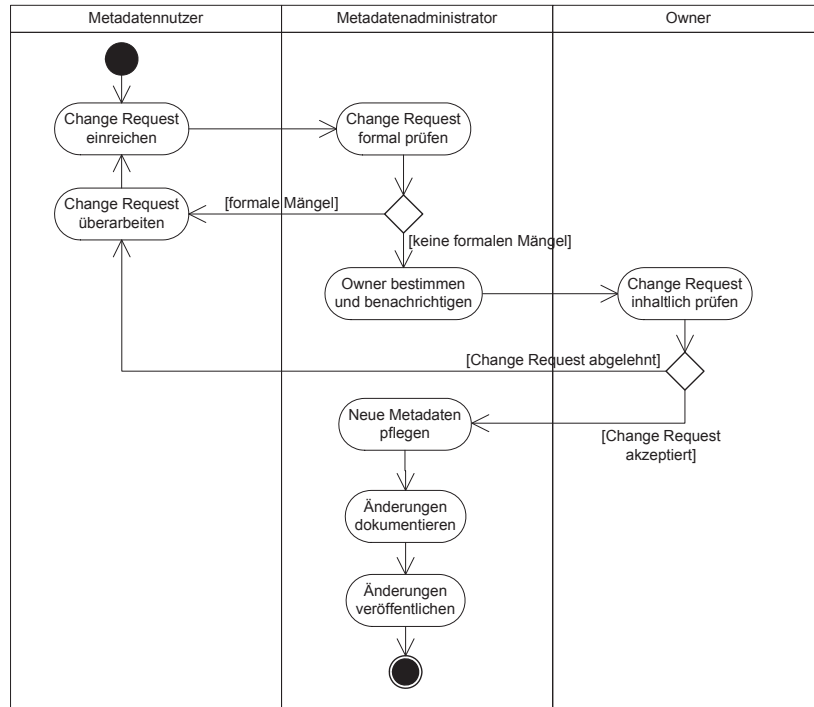


Abbildung 5-12: Aktivitätsdiagramm für den Metadatenmanagementprozess Anlegen/Ändern fachlicher Metadaten

Die einzelnen Aktivitäten des Aktivitätsdiagramms sind in einem Aufgabenverzeichnis in Tabelle 5-19 kurz charakterisiert. Die Ausführung des Prozesses sollte durch einen Workflow automatisiert gesteuert werden, um die Bearbeitung von Änderungsanforderungen zu beschleunigen.

Aufgabe	Beschreibung	Aufgabenträger
Change Request einreichen	Formulieren des Change Requests für neu anzulegende oder zu ändernde fachliche Metadaten entsprechend dem vorgegebenen Template und Übermitteln an den Metadatenverantwortlichen (z. B. per Workflow oder E-Mail)	Metadatenutzer
Change Request formal prüfen	Überprüfen des Change Requests auf formale Richtigkeit (z. B. richtiger Adressat)	Metadatenadministrator
Owner bestimmen und benachrichtigen	Zuweisung des Change Requests an den zuständigen Owner, in der Regel einen fachlichen Datensteward, der über das notwendige fachliche Wissen zur inhaltlichen Beurteilung verfügt	Metadatenadministrator
Change Request inhaltlich prüfen	Überprüfen des Change Requests auf sachlich-logische Richtigkeit und Abhängigkeiten zu anderen fachlichen Metadaten. Dies kann die Abstimmung mit Fachbereichsvertretern oder Prozess-Ownern erforderlich machen.	Owner
Change Request überarbeiten	Beseitigung der formalen oder inhaltlichen Mängel im Falle einer Ablehnung des Change Requests	Metadatenutzer
Neue Metadaten pflegen	Hinzufügen neuer oder Verändern bestehender Inhalte im BDD	Metadatenadministrator
Änderungen dokumentieren	Dauerhafte Dokumentation sowohl des Change Requests als auch seiner Umsetzung, um konsistente Weiterentwicklung in der Zukunft zu ermöglichen	Metadatenadministrator
Änderungen veröffentlichen	Benachrichtigung des Antragstellers sowie des Owners und unternehmensweite Publikation der Änderungen	Metadatenadministrator

Tabelle 5-19: Aufgabenverzeichnis für den Metadatenmanagementprozess „Anlegen/Ändern fachlicher Metadaten“

Zusätzlich zum Vorgehen für das Management von Change Requests sollten entsprechende Antragsformulare standardisiert vorgegeben werden, die dem Metadatenutzer die Beantragung von Änderungen erleichtern. Eine Vorlage für ein solches Formular ist Anhang D.3 beigelegt.

5.5.4 Werkzeugunterstützung

Das nachhaltige Management fachlicher Metadaten in einem BDD ist nur werkzeugu-nterstützt effizient möglich [White et al. 2008, 12]. Aufbauend auf einer Analyse der Anforderungen an ein BDD untersucht dieses Kapitel mögliche Software-Lösungen aus dem MDM-Umfeld, die Aufgaben des Metadatenmanagements unterstützen und somit BDD-Funktionalität bieten.

In einem Fokusgruppeninterview²⁹ im Rahmen des Kompetenzzentrums CDQ identifizierten Fachexperten aus dem Bereich des Daten- und Datenqualitätsmanagements folgende funktionale und nicht-funktionale Anforderungen an ein BDD (siehe Tabelle 5-20). Die funktionalen Anforderungen mit Bezug zum Metadatenmodell eines BDD werden an dieser Stelle nicht erneut aufgezählt. Sie sind in Kapitel 5.5.2 zur Ableitung des Referenz-Metadatenmodells für ein BDD beschrieben.

Anforderung	Erläuterung
Funktionale Anforderungen	
A01. Abbildung von Administrations-Workflows	Dem BDD sollten klar definierte und dokumentierte Arbeitsabläufe zur Nutzung und Pflege zugrunde liegen, beispielsweise Workflows für Fehlermeldungen oder zur Aktualisierung von Metadaten.
A02. Verweise zwischen Metadaten	Das BDD sollte die Verlinkung der Metadaten erlauben. Durch die Verlinkung werden die Metadaten strukturiert, die redundante Speicherung von Metadaten wird vermieden und die effiziente Navigation im fachlichen Metadatenkatalog unterstützt. Dadurch realisiert ein BDD eine über ein einfaches Glossar hinausgehende Funktionalität.
A03. Rollenbasierte Benutzerschnittstelle	Um eine hohe Nutzung des BDD durch Vertreter verschiedenster Bereiche zu gewährleisten, sollte die Benutzerschnittstelle individuell für verschiedene Rollen mit unterschiedlichen Aufgaben und Rechten vorkonfiguriert werden können.
Nicht-funktionale Anforderungen	
A04. Unterstützung des Fachbereiches	Ziel eines BDD ist es, das Wissen des Fachbereiches über Geschäftsobjekttypen zu erfassen. Daher ist eine einfache und intuitive Bedienung wichtig (z. B. durch Dateneingaben über Formulare), um die Nutzung für den Fachbereich zu vereinfachen.
A05. Vorgaben der IT-Strategie	Die Lösung sollte bestehende Anforderungen der IT-Strategie erfüllen, u. a. die Unterstützung vorhandener Authentifizierungsverfahren (z. B. Single Sign-On).
A06. Flexible Metadatenstruktur	Die einfache Anpassung von Metadaten (z. B. neue Attribute) und Funktionen des Metadatenkatalogs (z. B. neue Suchfilter) durch den Nutzer muss möglich sein, da sich Metadaten im Geschäftsverlauf sowohl inhaltlich als auch strukturell ändern.
A07. Integration der Metadaten in operative Applikationen	Um Mitarbeitern einen einfachen Zugang zu den Metadaten zu ermöglichen, ist eine Integration der Metadaten in operative Applikationen wünschenswert. Der Export der Metadaten aus dem BDD sowie der Import in operative Applikationen sind nur durch standardisierte Metadatenformate möglich.
A08. Versionierung und Änderungshistorie	Um Änderungen nachvollziehen und Fehler korrigieren zu können, sollten verschiedene Versionen sowie die Änderungshistorie von Metadaten gespeichert und darauf aufbauend ältere Zustände der Metadaten wiederhergestellt werden können.

²⁹ Das Fokusgruppeninterview wurde am 16. November 2007 auf dem 5. CC CDQ-Workshops durchgeführt.

Anforderung	Erläuterung
A09. Mehrsprachigkeit	Sowohl Benutzerschnittstelle als auch Metadaten sollten in verschiedenen Sprachen angeboten und erfasst werden können.

Tabelle 5-20: Funktionale Anforderungen an einen fachlichen Metadatenkatalog (in Anlehnung an [Hüner et al. 2011, 5f.]

Das Angebot an kommerziellen Software-Produkten zur Unterstützung des Managements fachlicher Metadaten muss derzeit als gering bewertet werden. So zeigt eine aktuelle Marktuntersuchung unter Anbietern von Software für das Stammdatenmanagement, dass von 62 untersuchten MDM-Lösungen lediglich vierzehn Funktionalitäten in der Rubrik *Metadata Modeling/Semantic Modeling Related to MDM* anbieten [White et al. 2009, 5-7]. Von diesen vierzehn wiederum deckt kein Anbieter mit seiner Lösung das gesamte Anforderungsspektrum an einen fachlichen Metadatenkatalog ab. Die meisten Produkte sind erweiterte Modellierungswerkzeuge, die fachliche Metadaten zur Beschreibung von Stammdaten mit aufnehmen.

Die meisten Software-Anbieter haben derzeit zumindest Teilfunktionalitäten für das Metadatenmanagement in ihren MDM-Lösungen integriert [Otto/Hüner 2009, 53]. So enthält die Komponente *InfoSphere Metadata Server*³⁰ des Produktes InfoSphere von IBM Funktionen zur Analyse, Modellierung, Nutzung und Verwaltung von Metadaten [Otto/Hüner 2009, 47]. Mit dem *InfoSphere Business Glossary*³¹ können Glossare erstellt werden, die Beschreibungen von Geschäftsobjekttypen verwalten. Das zugrunde liegende Metamodell ist an unternehmensspezifische Anforderungen anpassbar. Die SAP hat in ihrer NetWeaver MDM-Lösung ähnliche Funktionalitäten in der Komponente *SAP BO Metadata Management* realisiert, die auch die Möglichkeit der Erstellung einfacher Glossare mit Definitionen verwendeter Fachbegriffe (Geschäftsobjekttypen) umfasst. Die Fachbegriffe können anschliessend auch Datenobjekttypen zugeordnet werden. Die Datenintegrationslösung PowerCenter von Informatica enthält innerhalb der Komponente *Metadata Manager* die Funktion *Business Glossary*³². Mit dieser wird das Ziel verfolgt, die organisationsbereichsübergreifende Integration durch Erstellung, Pflege und Verteilung konsistenter, für Fachbereich und IT verständlicher Definitionen von Fachbegriffen und Daten zu unterstützen. Die Allen Systems Group (ASG) bietet als Teil ihrer Metadatenmanagement-Lösung die Komponente *ASG-metaGlossary*³³ für die kollaborative, workflowbasierte Erstellung und Pflege von Definitionen zu Fachbegriffen. Die terminologischen Metadaten können mit sonstigen Metadaten des Unternehmens (über Geschäftsprozesse, Applikationen, Datenobjekttypen usw.) im Rochade-Repository verknüpft werden. Bei allen vier Tools ist es möglich, die Geschäftsobjekttypen den verantwortlichen Datenstewards zuzuordnen sowie der Änderungshistorie von Metadaten zu dokumentieren.

³⁰ Siehe <http://www-142.ibm.com/software/products/gb/en/ibminfometawork>

³¹ Siehe <http://www.sap.com/about/newsroom/press.epx?pressid=10070>

³² Siehe http://www.informatica.com/INFA_Resources/brief_business_glossary_6922.pdf

³³ Siehe http://www.asg.com/products/product_details.asp?code=AMG

Eine weitere, für Unternehmen zunehmend attraktivere Option der Realisierung fachlicher Metadatenkataloge sind wiki-basierte Implementierungen. Mehrere Gründe sprechen für eine derartige Lösung: Erstens sind die Kosten für Implementierung und Wartung eines Wikis gering. Die notwendige Software ist Open Source verfügbar, wodurch keinerlei Lizenzkosten entstehen. Da der Nutzer auf Inhalte eines Wikis webbasiert zugreift, ist keine aufwendige Softwareinstallation auf lokalen Arbeitsplatzrechnern erforderlich. Zudem existieren in vielen Unternehmen bereits interne Wikis, z. B. als Werkzeug für das Wissensmanagement [Müller/Gronau 2008, 13-14; Bughin et al. 2009, 13], sowie das Fachwissen zum Betreiben solcher Wikis. Zweitens verfügen die Mitarbeiter in der Regel über Erfahrung im Umgang mit Wikis, sei es aufgrund der arbeitsbedingten oder der privaten Nutzung. Dies beeinflusst die Bereitschaft der Mitarbeiter, das Tool aktiv zu nutzen, positiv. Dadurch kann auch für das Management fachlicher Metadaten, das häufig durch das Problem der unzureichenden Integration bestehender Metadaten-Repositorys in die Prozesse und Arbeitsabläufe gekennzeichnet ist [Informatica 2006, 4], ein Mehrwert geschaffen werden. Durch die Möglichkeit für jeden Nutzer, Inhalte von Wiki-Seiten nicht nur lesen, sondern mit Hilfe einer einfachen Auszeichnungssprache auch ändern zu können, unterstützen Wikis drittens die gemeinschaftliche Erstellung von Inhalten und den Wissensaustausch zwischen Nutzern [Frankfurth/Schellhase 2007, 118; Krötzsch et al. 2007, 251]. Der kollaborative Ansatz ermöglicht es, das in Unternehmen verteilt vorliegende Fachwissen (beispielsweise über Geschäftsobjekttypen) effizient zu bündeln und allen Nutzern zugänglich zu machen [Inmon et al. 2008, 104]. Durch einfache Verlinkungsmöglichkeiten wird zudem gewährleistet, dass Inhalte redundanzfrei erzeugt werden [Müller/Gronau 2008, 14].

Das Potenzial eines *wiki-basierten Metadatenmanagements* wurde in den letzten Jahren durch die zunehmende Integration semantischer Konzepte zusätzlich erhöht. Bei den daraus resultierenden *semantischen Wikis* wird Wissen über Seiten (d. h. Wiki-Artikel über einen Geschäftsobjekttyp) und ihre Relationen in einer formalsprachlich definierten Wissensbasis in Form von Metadaten gespeichert. Diese Metadaten können aus dem Wiki exportiert und aufgrund ihrer formalsprachlichen Darstellung durch Maschinen verarbeitet werden. Für die strukturierte Abbildung werden formale Beschreibungssprachen wie RDF oder OWL verwendet [Souzis 2005, 88]. Mit Hilfe der annotierten Seiten und Links können erweiterte, schnellere und benutzerfreundlichere Such- und Navigationsfunktionalitäten realisiert und automatisierte Analysen (z. B. Liste aller Attribute eines Geschäftsobjekttypen) erstellt werden [Schaffert et al. 2007, 435; Hüner/Otto 2009, 8]. Für die Anwendung dieser Technologien im Rahmen eines BDD bedeutet dies, dass Metadaten zu einzelnen Entitäten für den BDD-Nutzer einfacher zu finden sind. Die Komponente SAP BO Metadata Management der NetWeaver MDM-Lösung von SAP erlaubt in ihrer neuesten Version 3.0 bereits die Erstellung eines sogenannten Metapedia als Enzyklopädie fachlicher Metadaten, die die technolo-

gischen Möglichkeiten der Verlinkung und semantischen Annotationen von Einträgen nutzt [Business Objects 2008, 2-3].

Wie die Vorgehensbeschreibung in Kapitel 5.5.1 zeigt, bedarf es bei der Definition von Geschäftsobjekttypen der Einbindung einer Vielzahl von Fachexperten aus verschiedenen Unternehmensbereichen. Durch die beschriebenen Eigenschaften, insbesondere das Zusammentragen und Konsolidieren von Informationen in einem kollaborativen Prozess, können semantische Wikis sehr gut als Tool für die Unterstützung des Definitionsprozesses (Identifikation und Spezifikation von Geschäftsobjekttypen) sowie die anschließende Nutzung und Pflege der Metadaten eingesetzt werden [Muljadi et al. 2006, 273; Egloff 2008, 26f.].

Die Erfüllung der in Tabelle 5-20 definierten Anforderungen durch ein semantisches Wiki wurde prototypisch in einem Projekt mit der Bayer CropScience im Rahmen des Kompetenzzentrums CDQ evaluiert. Hierzu wurde mit Hilfe der Software Semantic MediaWiki³⁴, einer durch semantische Konzepte erweiterten Wiki-Software, ein BDD implementiert [Hüner et al. 2011, 6-10] und das bestehende Lotus Notes basierte Master Data Handbook abgelöst. Das Master Data Handbook dokumentiert als fachlicher Metadatenkatalog Informationen zu den Stammdaten des Unternehmens. Die neue Lösung wurde anhand des Szenarios „Metadaten pflegen“ getestet, in dem fehlerhafte Metadaten verändert werden. Um die Benutzerfreundlichkeit der Wiki-Lösung zu erhöhen, wurden zusätzlich Eingabeformulare entwickelt.

In Bezug auf die Anforderungen konnten durch die Anwendung des Prototyps sowie die Einschätzung des Prototyps durch Fachexperten des Kompetenzzentrums CDQ in einem Fokusgruppeninterview³⁵ folgende Erkenntnisse gewonnen werden [Hüner et al. 2011, 10f.]:

- *Die Anforderungen A02, A03, A04, A06, A08 und A09 wurden vollständig erfüllt. Besonders hervorzuheben ist hierbei die Flexibilität in Bezug auf das zugrunde liegende Metadatenmodell des BDD (A06), das den Unternehmensanforderungen mit geringem Aufwand angepasst werden kann. Die Erfüllung von Anforderung A07 ist durch die Möglichkeit, den Wiki-Inhalt über einen CSV- oder XML-Export zu integrieren, gegeben.*
- *Durch die Bereitstellung von Formularen zur Eingabe der Metadaten konnte Anforderung A01 teilweise erfüllt werden. Zwar können allein mit der Software keine Workflows definiert werden, jedoch erlaubt es die Vorgabe von Folgen von Formularen, einfache Arbeitsabläufe zu unterstützen.*

³⁴ Siehe <http://www.semantic-mediawiki.org>

³⁵ Das Fokusgruppeninterview wurde im Rahmen eines Workshops am 03. Dezember 2008 durchgeführt.

- *Nicht bzw. nur unzureichend erfüllt wurde Anforderung A05, da der Prototyp bestehende Authentifizierungsverfahren (z. B. Single Sign-on) nicht unterstützt. Grundsätzlich ist diese durch die Software Semantic MediaWiki jedoch möglich.*

Zudem wurde die Annahme der vereinfachten Nutzung durch Vertreter des Fachbereiches bestätigt, da sämtliche Testnutzer die ihnen gestellten Aufgaben (Suchen, Anlegen, Verändern von Metadaten) ohne Schulung problemlos erfüllen konnten. Benutzerfreundlichkeit und Einfachheit der Navigation stellten wesentliche Verbesserungen gegenüber dem Master Data Handbook dar. Durch die automatisierte Benachrichtigung des verantwortlichen Metadatenpflegers im Falle veralteter, d. h. seit längerem nicht gepflegter Metadaten, verspricht sich das Unternehmen, auch die Aktualität der BDD-Inhalte zu erhöhen.

Abbildung 5-13 zeigt einen Ausschnitt des mittels Semantic MediaWiki realisierten BDD-Prototyps der Bayer CropScience.

Abbildung 5-13: Screenshot des BDD der Bayer CropScience

Da die ausschliessliche Sammlung fachlicher Metadaten ohne deren intensive Verwendung wenig Nutzenpotenzial bietet [Loshin 2008, 53], besteht ein nächster Schritt in der kontextabhängigen Integration der im BDD gespeicherten Metadaten in die Applikationen des Fachbereiches (z. B. SAP ERP oder CRM). Hierdurch erhielt der Fachbereich benötigte Informationen zu Geschäftsobjekttypen und deren Attributen direkt und medienbruchfrei in seiner gewohnten Arbeitsumgebung, ohne nach diesen Informationen in anderen Applikationen suchen zu müssen. Eine derartige Lösung wird derzeit innerhalb des Kompetenzzentrums CDQ auf Grundlage von Technologien der

Texterkennung (engl.: Object Character Recognition) zur Interpretation und Auswertung von Bildschirminhalten erarbeitet.

5.6 Methodenfragment 2: Semantisches Stammdatenmodell ableiten

Die fachlichen Metadaten des BDD beschreiben einerseits Bedeutung und Struktur von Geschäftsobjekttypen (z. B. deren Attribute oder Beziehungen zu anderen Geschäftsobjekttypen). Sie liefern andererseits Informationen zum Kontext, in dem Geschäftsobjekte genutzt werden, sowie zur korrekten Verwendung der Geschäftsobjekte in Geschäftsprozessen. Das BDD als Ergebnis des ersten Methodenfragmentes speichert die fachlichen Metadaten der für ein Unternehmen wesentlichen Geschäftsobjekte in Form natürlichsprachlicher Beschreibungen und wendet sich damit speziell an den Fachbereich. Der nächste Schritt besteht darin, die im BDD gepflegten, fachlichen Metadaten in eine technische Repräsentation zu überführen [Inmon et al. 2008, 209-211]. Für ein effektives Management von Stammdaten sowie deren Integration ist ein formal und präzise definierter Verwendungskontext essenziell [Sciore et al. 1994, 255; Madnick 1999, 3; Kentouris 2009, 10]. Methodenfragment 2 widmet sich dieser Herausforderung und dient der Ableitung eines semantischen Stammdatenmodells.

In der Datenmodellierung fand die Abbildung des Verwendungskontextes in der Vergangenheit wenig Berücksichtigung, da er entweder unzureichend vom Fachbereich definiert wurde oder die Modellierungskonzepte und -notation deren Modellierung nicht vorsahen [Sheth/Larson 1990, 187]. PARENT und SPACCAPIETRA verweisen zudem auf die Unmöglichkeit, den Kontext bzw. die Semantik vollständig in einem konzeptionellen Datenmodell abzubilden [Parent/Spaccapietra 1998, 168]. Jedoch gibt es in der Literatur erste Ansätze zur „semantischen Anreicherung“ [Parent/Spaccapietra 1998, 168], indem einzelnen Datenobjektattributen ein Kontext zugewiesen wird, welcher die Semantik von Datenobjektattributen und Datenobjekttypen erfasst (vgl. [Sciore et al. 1994], [Siegel/Madnick 1991]).

Die Core Component Technical Specification (CCTS) definiert eine Methode zur Standardisierung und Integration von Semantik (insbesondere des Verwendungskontextes) in Daten- und Informationsmodellen [Janiesch 2007, 2051]. Sie wird daher in Kapitel 5.6.1 kurz einführend vorgestellt, bevor in den anschließenden Kapiteln das zweite Methodenfragment spezifiziert wird. Ergebnis des Methodenfragmentes ist ein aus dem BDD abgeleitetes semantisches Stammdatenmodell, das einen Teil der fachlichen Metadaten erfasst und die heterogenen Datenmodelle einzelner Applikationen zueinander in Beziehung setzt.

5.6.1 CCTS als Verfahren zur semantischen Datenmodellierung

Die *Core Component Technical Specification* ist eine von fünf Teilspezifikationen des Standards ISO 15000, welche die Vereinten Nationen (UN) und die Organization for the Advancement of Structured Information Standards (OASIS) im Zuge der gemeinsamen ebXML-Initiative entwickelt haben. Die CCTS spezifiziert eine Methode für die Entwicklung semantisch eindeutiger Datenobjekttypen, auf deren Basis Geschäftsvokabulare erstellt und bereits bestehende Vokabulare integriert werden können [Kramer 2007, 14]. Die CCTS V3.0 (inklusive Core Component Data Type Catalogue) wurde im Oktober 2009 offiziell verabschiedet und ist die seitdem gültige Version der Spezifikation (vgl. [UN/CEFACT 2009b]).

Neben ihrem primären Anwendungsbereich zur Spezifikation unternehmensübergreifend verwendeter Datenmodelle für die Definition von Nachrichten erfährt die Spezifikation mittlerweile auch bei der Entwicklung unternehmensinterner Datentypen zunehmend Beachtung. Unter anderem nutzt die SAP die CCTS zur Definition der SAP Global Data Types. Durch Aggregation der standardisierten Global Data Types gewährleistet die SAP eine unternehmensweit einheitliche Modellierung ihrer Datenobjekttypen. Des Weiteren bilden die Global Data Types die Grundlage für die Ableitung applikationsspezifischer Datentypen in einem bestimmten Kontext [SAP 2009]. Sowohl die Global Data Types als auch die daraus abgeleiteten, applikationsspezifischen Context Data Types werden zentral im Enterprise Service Repository gespeichert und stehen damit zur Wiederverwendung zur Verfügung.

CCTS nutzt für die Modellierung von Datenobjekten hierarchische Aggregations- und Kompositionsbeziehungen und besteht aus vier Gestaltungselementen, die im Metamodell der Core Component Technical Specification definiert sind [UN/CEFACT 2009b, 51-76; Vogel 2009, 106]:

- *Aggregated Business Information Entities (ABIEs)* fassen eine Menge elementarer und komplexer Datenobjektattribute zu einem Datenobjekttyp zusammen.
- *Basic Business Information Entities (BBIEs)* sind elementare Datenobjektattribute, denen bei der Instanziierung ein atomarer Wert zugeordnet wird.
- *Associated Business Information Entities (ASBIEs)* repräsentieren komplexe Datenobjektattribute, d. h. Felder, die auf einen anderen Datenobjekttyp (ABIE) verweisen und diesen somit hierarchisch einbinden.
- *Business Data Types (BDTs)* sind Datenobjektattributtypen wie z. B. Datum, Zeit, Betrag oder Text, die unabhängig von technischen Plattformen oder Programmiersprachen das Format und den Wertebereich von elementaren Attributen definieren.

Durch die aufeinander aufbauenden Komponenten des Datenobjektmodells wird die Wiederverwendung vereinfacht. Hierzu ist es nötig, sämtliche Komponenten durch

Metadaten vollständig zu spezifizieren. Die Datenobjekttypen, d. h. sämtliche einmal verwendete BIEs und BDTs, werden mit ihren Metadaten (vor allem den natürlichsprachlichen Definitionen) in einem zentralen Repository, der Core Component Library (CCL), gespeichert, von wo aus sie andere Nutzer wiederverwenden können.

Für das Ziel, Datenobjekttypen und deren Attribute semantisch eindeutig zu definieren, gibt die UN/CEFACT in ihrer Spezifikation exakte Benennungs- und Definitionsregeln vor [UN/CEFACT 2009b, 31-33]. Diese Namenskonventionen lehnen sich an den ISO-Standard 11179 (siehe Tabelle 2-4) an, insbesondere an den fünften Teil zur Identifikation und Benennung von Datenelementen. Jede Datenobjekttypkomponente wird durch einen sogenannten *Dictionary Entry Name (DEN)* repräsentiert. Dieser setzt sich aus drei Teilen zusammen [ISO/IEC 2005b, 7-10; UN/CEFACT 2009b, 54]:

- Der *Representation Term* ist der Teil einer Datenobjekttypkomponente, der ihren Typ spezifiziert und somit den gültigen Wertebereich festlegt.
- Der *Property Term* beschreibt die Eigenschaft einer Datenobjekttypkomponente, durch die sie sich von anderen unterscheidet.
- Der *Object Class Term* stellt ein Objekt dar, das Datenobjekttypkomponenten logisch gruppiert.

Des Weiteren ermöglichen sogenannte *Qualifier Terms* die Detaillierung der Semantik eines DENs, indem sie als beigefügtes Wort einen Object Class oder Property Term eingrenzen und die Datenobjekttypkomponente damit von anderen abgrenzen [UN/CEFACT 2009b, 54]. Durch eine derart aufgebaute Benennung ist es möglich, DENs vereinfacht aus natürlichsprachlichen Definitionen (z. B. aus einem BDD) abzuleiten und die den Definitionen inhärente Semantik in einem Datenmodell abzubilden.

Die vier Datenobjekttypkomponenten – also sämtliche BIEs und BDTs – können mit Hilfe des *Kontexttreiberprinzips* zusätzlich semantisch annotiert und somit an spezifische Anforderungen des Anwendungsfalls, in dem sie genutzt werden, angepasst werden. Dem Kontexttreiberprinzip liegt die Annahme zugrunde, dass die inhaltliche Bedeutung von Daten immer von der Verwendung im konkreten Anwendungskontext abhängt. Der Ansatz folgt dem Konzept der konfigurativen Referenzmodellierung, bei der spezifische Modellvarianten eines Informationsmodells durch Ausprägung vorgegebener Konfigurationsparameter aus einem integrierten Gesamtmodell erzeugt werden können [Becker et al. 2008a, 841]. Für die Konfigurationen werden Datenobjekttypkomponenten kategorisiert, d. h. ihnen werden für jede der vorgegebenen Kontextkategorien kein Wert, ein Wert oder mehrere Werte zugeordnet. Wird kein Wert zugeordnet, so bedeutet dies, dass keine Einschränkung für einen spezifischen Anwendungsfall vorgenommen wird und die Datenobjekttypkomponente allgemein gültig ist. Der Gesamtkontext einer Datenobjekttypkomponente ergibt sich als logische Konjunktion (sogenannte Kontextvereinigung, engl.: context unit) aus der Menge der Werte aller

Kontextkategorien. Dadurch kann der Verwendungskontext der BIEs und BDTs formal in Form prädikatenlogischer Ausdrücke beschrieben werden [Pentcheva 2007, 62f.; Yu 2007, 29-37]. Das Kontexttreiberprinzip korrespondiert mit früheren Ansätzen zur Beschreibung der betriebswirtschaftlichen Bedeutung von Datenelementen, beispielsweise der Verwendung von Deskriptoren und Deskriptorenklassen. Auch hier wurden vorgegebene Deskriptoren, also Worte, die inhaltliche Eigenschaften darstellen, aus einer Deskriptorenklasse ausgewählt und einem Datenelement zugeordnet [Brenner 1985, 69-70]. Die Deskriptorenklassen sowie die dazugehörigen Deskriptoren waren damals jedoch auf eine geringe Anzahl beschränkt und nicht unternehmensübergreifend standardisiert [Brenner 1985, 90].

Die UN/CEFACT gibt für die Konfiguration u.a. industrie-, prozess- und produktspezifische *Verwendungskontexte* (engl.: business contexts) vor [UN/CEFACT 2009b, 107ff.]. Die Verwendungskontexte entsprechen fachlichen Metadaten. In der aktuellen Spezifikation sind acht Kontextkategorien definiert. Für die Vorgabe von Werten für jede der Kontextkategorien greift die UN/CEFACT auf international standardisierte Codelisten zurück, wie beispielsweise die Universal Standard Product and Service Specification (UNSPSC) zur Klassifikation von Produkten oder den International Standard for Industrial Classification (ISIC) für die Unterscheidung von Industrien. Die Kontextkategorien sind mit den für sie vorgegebenen Codelisten sowie einer kurzen Beschreibung in Tabelle 5-21 zusammengefasst.

Kontextkategorie	Beschreibung	Verwendete Codelisten
Business Process Context (Geschäftsprozess)	Geschäftsprozesse und Aktivitäten eines Unternehmens, in denen Daten genutzt werden	UN/CEFACT Catalogue of Common Business Processes
Product Classification Context (Produktklassifikation)	Aspekte von Gütern und Dienstleistungen, die im Geschäftsprozess erzeugt, verändert oder ausgetauscht werden	Universal Standard Product and Service Specification (UNSPSC), Standard International Trade Classification (SITC)
Industry Classification Context (Industriezugehörigkeit)	Industrien bzw. Industriezweige, in denen der Geschäftsprozess abläuft	International Standard for Industrial Classification (ISIC)
Geopolitical Context (Länderzugehörigkeit)	Geographische Faktoren, welche zu unterschiedlichen Repräsentationen einer Datenobjektkomponente (z. B. die Struktur einer Adresse) führen	ISO 3166.1 Country Code List
Official Constraints (Regularien)	Gesetzliche oder regulatorische Anforderungen, welche die Ausführung von Geschäftsprozessen beeinflussen	Keine Codeliste vorgegeben; verwendetes Klassifikationschema ist anzugeben
Business Process Role Context (Geschäftsprozessrolle)	Akteure, die in einem Geschäftsprozess miteinander agieren	UN/CEFACT Catalogue of Common Business Processes
Supporting Role Context (Unterstützungsrolle)	Akteure, die nicht aktiv in den Geschäftsprozess eingreifen, jedoch an dessen Ausführung und Prozessleistungen interessiert sind	UN/CEFACT Catalogue of Common Business Processes, UN/EDIFACT Code List for DE 3035 Party Roles
System Capabilities Context (Applikationseigenschaften)	Einflussfaktoren, die sich aus Eigenschaften einer Applikation, einer Klasse von Applikationen oder einem bestimmten Standard ergeben	Keine Codeliste vorgegeben; verwendetes Klassifikationschema ist anzugeben

Tabelle 5-21: Kontextkategorien der CCTS 3.0 [UN/CEFACT 2009b, 109-114]

Das Kontexttreiberprinzip ist derzeit lediglich als Teil der CCTS spezifiziert. Eine gesonderte Spezifikation, welche die Verwendung von Kontexten in Datenobjekt- und Prozessmodellen einheitlich und in Form einer Methode (Unified Context Methodology) definiert, wird derzeit durch eine gesonderte Arbeitsgruppe der UN/CEFACT erarbeitet³⁶. Sie soll die systematische Abbildung des Verwendungskontextes durch Kontextinformationen für sämtliche Modellierungsartefakte der UN/CEFACT-Spezifikationen (u. a. Datentypen, Datenobjektattribute, Geschäftsprozessbausteine) gewährleisten.

Die vorliegende Methode übernimmt das Kontexttreiberprinzips der CCTS (siehe Aktivität III.1 von Methodenfragment 2). Das Konzept wird durch den in Kapitel 5.7.4 beschriebenen Modellierungsprototyp unterstützt und ermöglicht durch die Kontextualisierung einzelner Datenobjektattribute und Datenobjekttypen die Abbildung applikationsspezifischer Datenmodelle. Diese können mit Hilfe des Modellierungsprototyps durch Verwendung der Metadaten teilautomatisiert auf ein applikationsübergreifend konsolidiertes Stammdatenmodell abgebildet werden. Das konsolidierte Stammdatenmodell kann als Grundlage für das Datenmodell eines Stammdatenservers als Bestandteil der Integrationsarchitektur (siehe Methodenfragment 3 in Kapitel 5.7) genutzt werden. Aus dem konsolidierten Stammdatenmodell sind dann kontextspezifische Datenmodelle (z. B. für einzelne Applikationen) konfigurierbar, indem bestimmte Kontextwerte für Kontextkategorien ausgewählt werden (siehe Technik T.III.3 sowie das Beispiel aus dem Aktionsforschungsprojekt Daimler AG in Kapitel 6.1.2). Der Abgleich von Datenobjektattributen und Datenobjekttypen mit den BIEs der CCL und die standardisierte Benennung der Datenobjektcomponenten entsprechend der CCTS ist für den innerbetrieblichen Anwendungsfall nicht zwingend. Die sich aus der Standardisierung ergebenden Vorteile sind vor allem im überbetrieblichen Kontext von Bedeutung.

5.6.2 Vorgehensmodell und Aktivitäten

Die Aktivitäten des zweiten Methodenfragmentes werden wiederum aus den konsolidierten Aktivitäten der Fallstudien abgeleitet (siehe Tabelle 5-22).

Nr.	Aktivität	Ergebnisse	Aktivität aus Fallstudie		
			SBB Cargo	DTAG	Bosch Rexroth
I Vorbereitung					
I.1	Stakeholder und Applikationen identifizieren	Identifizierte Ansprechpartner, Applikationsverzeichnis	A.I.2	B.I.3	C.II.1
II Bestandsaufnahme & Definition					
II.1	Applikationsanalyse	Applikationsbeschreibung, Liste we-	A.II.1,	B.II.1b,	–

³⁶ Die Spezifikation ist derzeit noch in einer nicht öffentlich zugänglichen Draft-Version verfügbar (siehe <http://unstandards.org:8080/display/public/UCM+-+Unified+Context+Methodology>).

Nr.	Aktivität	Ergebnisse	Aktivität aus Fallstudie		
			SBB Cargo	DTAG	Bosch Rexroth
	durchführen	sentlicher Datenobjektattribute	A.II.2	B.II.1c	
II.2a	Datenobjektattribute spezifizieren	Vollständig definierte Datenobjektattribute	A.II.5b	B.II.4b	C.II.2
II.2b	Stammdatenlandkarte erstellen	Stammdatenlandkarte	A.IV.1, A.IV.2	B.IV.1	–
II.3	Daten- und Geschäftsobjekttypen konsolidieren	Vollständige Zuordnung von Daten- zu Geschäftsobjekttypen	–	B.II.5	–
II.4	BDD aktualisieren	Aktualisierte Metadaten im BDD	–	B.II.5	–
III Datenmodellierung & Kontextualisierung					
III.1	Semantisches Stammdatenmodell ableiten	Semantisches Stammdatenmodell mit kontextualisierten Datenobjekttypen	–	(B.II.6)	C.II.2

Tabelle 5-22: Aktivitäten des induzierten Vorgehensmodells von Methodenfragment 2

Die Aktivität MF2: III.1 ergibt sich aus dem Vorgehensmodell der CCTS [Kramer 2007, 108-109; Lemm 2007, 29-33] und ist nicht Bestandteil der Fallstudien. Sie dient der CCTS-konformen und somit standardbasierten Benennung und Modellierung von Datenobjekttypen und Datenobjektattributen. Eine entsprechende Aktivität ist in der Fallstudie der DTAG (siehe Kapitel 4.2.2) zu finden. Hier werden die Entitäten mit einem industriespezifischen Standard, dem SID, abgeglichen.

Die sachlogische Abfolge der einzelnen Aktivitäten von Methodenfragment 2 ist in Abbildung 5-14 dargestellt.

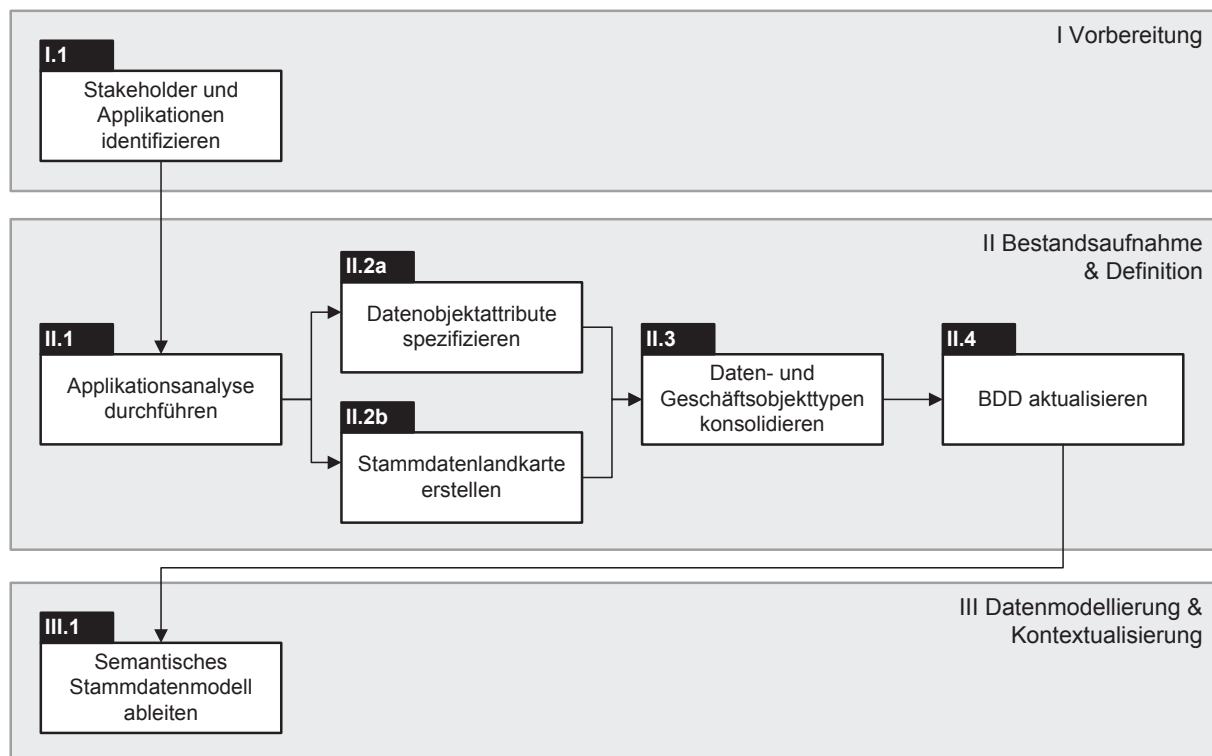


Abbildung 5-14: Vorgehensmodell des Methodenfragmentes 2

MF1 folgte einem Top-Down-Vorgehen, um Geschäftsobjekttypen aus den Geschäftsprozessen, die der Zielerreichung des Unternehmens dienen, abzuleiten und in einem BDD zu beschreiben. Dies ist notwendig, um Bedeutung und Verwendungskontext der Stammdaten aus fachlicher Sicht umfassend dokumentieren zu können. Methodenfragment 2 verfolgt das Ziel, der Fachbereichssicht die systemseitige Repräsentation der Stammdaten gegenüberzustellen. Dadurch soll das semantische Stammdatenmodell unter Kombination aus einem Top-Down- und Bottom-Up-Ansatz eine einheitliche Sicht von Fachbereich und IT auf die Stammdaten des Unternehmens wiedergeben [Martin/Leben 1989, 5; Adelman et al. 2005, 112-115; Dippold et al. 2005, 82ff.]. Das semantische Stammdatenmodell schliesst somit die Lücke zwischen fachseitiger Nutzung der Stammdaten und IT-technischer Implementierung, indem es den Geschäftsobjektattributen Datenobjektattribute zuordnet und Mehrfachrepräsentationen auflöst [Loshin 2008, 48]. Zudem unterstützt der kombinierte Ansatz eine bedarfsgetriebene Modellierung, bei der lediglich die Bereiche bzw. Objekte detailliert beschrieben werden (Bottom-Up), für die als Ergebnis von Methodenfragment 1 eine entsprechende Anforderung besteht.

Analog zum ersten Methodenfragment werden die Aktivitäten des zweiten Methodenfragmentes im Folgenden kurz textuell beschrieben und jeweils durch eine Tabelle, die verantwortliche Rollen, Ergebnisse, notwendige Vorbedingungen und Techniken nennt, charakterisiert. Die Techniken erläutert Kapitel 5.6.3.

5.6.2.1 Stakeholder und Applikationen identifizieren

Ähnlich wie in Methodenfragment 1 besteht die erste Aktivität darin, sich einen Überblick über die wichtigsten Ansprechpartner seitens der IT zu verschaffen, die über Wissen bezüglich der von den Applikationen genutzten Stammdatenobjekte verfügen [DAMA 2009, 180]. Hierfür kann auf der Ist-Analyse der Geschäftsprozesse aus Methodenfragment 1 aufgebaut werden, da diese in der Regel auch für die Geschäftsobjekttypen bedeutsame Applikationen sowie deren Owner aufzeigt. Als zusätzliche Informationsquelle können die Prozess-Owner befragt werden. Ausserdem können IT-Bebauungspläne Aufschluss über bedeutsame Applikationen geben. Sowohl Applikationen (mit Name und ID) als auch System-Owner sowie – falls im Unternehmen definiert – die technischen Datenstewards sind für die anschliessende Analyse durch den Konzern-Datensteward in einem Applikationsverzeichnis zu dokumentieren. Eine Vorlage für ein derartiges Applikationsverzeichnis ist Anhang D.4 zu entnehmen.

MF2: I.1 Stakeholder und Applikationen identifizieren	
Ergebnis	Identifizierte Ansprechpartner, Applikationsverzeichnis (falls nicht bereits vorhanden)
Rollen	Konzern-Datensteward [ausführend, verantwortlich] Prozess-Owner, Technische Datenstewards, System-Owner, Datenbankadministratoren [unterstützend]
Techniken	T.I.1 Identifikation der Daten- und System-Owner
Input/Vorbedingungen	Gestaltungsbereich der Methodenanwendung ist definiert (siehe Aktivität MF1: I.1), Ist-Geschäftsprozessanalyse, IT-Bebauungspläne

Tabelle 5-23: MF2: I.1 Stakeholder und Applikationen identifizieren

5.6.2.2 Applikationsanalyse durchführen

Im Mittelpunkt der Applikationsanalyse steht die Auswertung applikationsübergreifender Datenstrukturen zur Identifikation von Datenobjekttypen bzw. Datenobjektattributen [Spath et al. 2009, 7]. Die Applikationsanalyse baut auf zwei Informationsquellen auf. Einerseits kann die bestehende Dokumentation der in Aktivität MF2: I.1 identifizierten Applikationen ausgewertet werden (Technik T.II.1). Da die ausreichende Dokumentation für die vorhandenen Applikationen oft fehlt [Vogler 2004, 112], müssen andererseits die Vertreter der IT direkt befragt werden (Technik T.II.2). Als Ergebnis sollten einzelne Applikationsbeschreibungen zur Verfügung stehen, in der neben Namen, Schnittstellen der Applikation und Beschreibung der Applikationsfunktionen der zugehörige Datenbank und die von der Applikation benötigten Datenobjektattribute spezifiziert sind [Höning 2009, 201-203].

Die technischen Datenstewards, welche die Aktivität unter Koordination des Konzern-Datenstewards durchführen, erstellen schliesslich eine applikationsübergreifend konsolidierte Liste wesentlicher Datenobjektattribute [Marco/Michael 2004, 57]. Diese werden in der Literatur auch als kritische Datenobjektattribute bezeichnet (vgl. [Loshin 2008, 78f.]). Die Kritikalität der Datenobjektattribute ist hierbei massgeblich durch die im BDD beschriebenen Geschäftsobjekttypen und deren Attribute vorgegeben, da es genau diejenigen Objekte enthält, die für die Geschäftstätigkeit des Unternehmens von Bedeutung sind. Dadurch wird die Menge der zu spezifizierenden Datenobjektattribute auf eine zweckmässige Anzahl eingeschränkt.

MF2: II.1 Applikationsanalyse durchführen	
Ergebnis	Applikationsbeschreibung (falls nicht bereits vorhanden), Liste wesentlicher Datenobjektattribute
Rollen	Konzern-Datensteward [verantwortlich] Technische Datenstewards [ausführend] System-Owner, Datenbankadministratoren, Vertreter IT-Bereich [unterstützend]
Techniken	T.II.1 Inhaltsanalyse Applikationsdokumentation T.II.2 Einzelinterviews
Input/Vorbedingungen	Identifizierte Ansprechpartner (IT), BDD, Applikationsverzeichnis, IT-Bebauungspläne

Tabelle 5-24: MF2: II.1 Applikationsanalyse durchführen

5.6.2.3 Datenobjektattribute spezifizieren

Die Datenobjektattribute sind entsprechend den Vorgaben des Metadatenmodells des BDD zu spezifizieren. Für das in Kapitel 5.5.2 entwickelte Referenz-Metadatenmodell wären folglich Name, Datentyp, Feldlänge, Multiplizität und Ursprungsapplikationen für jedes Datenobjektattribut festzulegen und dieses dem entsprechenden Geschäftsobjektattribut zuzuordnen. Durch die Möglichkeit, einem Geschäftsobjektattribut mehrere Datenobjektattribute zuzuordnen, können redundante Repräsentationen in den Applikationen identifiziert werden. Dies erlaubt es, synonyme sowie homonyme Datenobjektattribute zu erkennen. Über die Definition der Multiplizität der Datenobjektattribute können zudem Rückschlüsse auf die Verbindlichkeit der Attribute gewonnen werden.

MF2: II.2a Datenobjektattribute spezifizieren	
Ergebnis	Vollständig definierte Datenobjektattribute
Rollen	Konzern-Datensteward [verantwortlich] Datenarchitekt, Technische Datenstewards [ausführend] System-Owner, Datenbankadministratoren [unterstützend]
Techniken	T.II.1 Inhaltsanalyse Applikationsdokumentation T.II.2 Einzelinterviews T.II.3 Workshops / Gruppendiskussion
Input/Vorbedingungen	Applikationsverzeichnis, Applikationsbeschreibung (inkl. verfügbarer Datenmodelle), Liste Datenobjektattribute

Tabelle 5-25: MF2: II. 2a Datenobjektattribute spezifizieren

Für die Spezifikation dieser Metadaten sollte auf verfügbare Datenmodelle einzelner Applikationen, in denen Instanzen der Datenobjektattribute gespeichert werden, sowie auf das Fachwissen der Datenarchitekten (die für die Datenmodellierung verantwortlich sind) zurückgegriffen werden.

5.6.2.4 Stammdatenlandkarte erstellen

Parallel zur Identifikation und inhaltlichen Spezifikation der Datenobjektattribute gilt es, diese in aggregierter Form (als Datenobjekttypen) den Applikationen mit den entsprechenden Zugriffsoperationen (CRUD) zuzuordnen, woraus eine sogenannte Stammdatenlandkarte resultiert. Hierfür müssen die technischen Datenstewards in Zusammenarbeit mit den System-Ownern führende Applikationen (in denen Stammdatenobjekte erstellt werden) und nutzende Applikationen bestimmen. Eine Vorlage für die Erstellung einer Stammdatenlandkarte ist in Anhang D.5 der Arbeit enthalten. Die Informationen der ersten Spalten können dem Applikationsverzeichnis (siehe Kapitel 5.6.2.1 bzw. Anhang D.4) entnommen werden.

In einer zusätzlichen Stammdatenlandkarte können die Datenobjekttypen den Organisationseinheiten eines Unternehmens zugeordnet werden. Eine solche organisatorische

Stammdatenlandkarte dokumentiert den Datenzugriff der Fachbereiche auf die Datenobjekttypen und zeigt, inwiefern die Entitäten unternehmensweit verwendet werden.

MF2: II.2b Stammdatenlandkarte erstellen	
Ergebnis	Stammdatenlandkarte (CRUD-Matrix)
Rollen	Konzern-Datensteward [verantwortlich] Datenarchitekt, Technische Datenstewards [ausführend] System-Owner, Datenbankadministratoren [unterstützend]
Techniken	T.II.4 Zuordnung von Datenobjekttypen zu Applikationen
Input/Vorbedingungen	Applikationsverzeichnis, Applikationsbeschreibung, Liste Datenobjektattribute

Tabelle 5-26: MF2: II.2b Stammdatenlandkarte erstellen

Die Stammdatenlandkarte ist ein wichtiger Input für die Gestaltung der Integrationsarchitektur im Rahmen des dritten Methodenfragmentes (siehe Kapitel 5.7). EAM-Werkzeuge (siehe Kapitel 5.7.4), welche eine Abbildung von Datenobjekttypen auf Applikationen erlauben, können die Erstellung der Stammdatenlandkarte unterstützen.

5.6.2.5 Daten- und Geschäftsobjekttypen konsolidieren

Die Aktivität ist Grundlage für ein kohärentes Metadatenmanagement zwischen Fachbereich und IT [Loshin 2008, 48]. Insbesondere gilt es, die Geschäftsverwendung der Datenelemente zu ermitteln, indem Datenobjektattribute anhand der im BDD gespeicherten Definitionen eindeutig den Geschäftsobjektattributen zugeordnet werden, die sie repräsentieren. Zudem sind Abweichungen im Verständnis zwischen Fachbereich und IT zu beseitigen. In seiner übergreifenden Rolle als Mediator zwischen Fachbereich und IT ist der Konzern-Datensteward für die Abstimmung zwischen fachlichen und technischen Datenstewards verantwortlich.

MF2: II.3 Daten- und Geschäftsobjekttypen konsolidieren	
Ergebnis	Vollständige Zuordnung von Daten- zu Geschäftsobjekttypen
Rollen	Konzern-Datensteward [ausführend, verantwortlich] Fachliche Datenstewards, Technische Datenstewards [unterstützend]
Techniken	T.II.3 Workshops / Gruppendiskussion
Input/Vorbedingungen	Spezifizierte Datenobjektattribute, Definierte Geschäftsobjekttypen bzw. Geschäftsobjektattribute (aus dem BDD)

Tabelle 5-27: MF2: II.3 Daten- und Geschäftsobjekttypen konsolidieren

Die Aktivität der Konsolidierung kann die Veränderung bestehender BDD-Inhalte zur Folge haben, wie z. B. das Ergänzen einzelner Geschäftsobjektattribute, das Anpassen der Definitionen von Geschäftsobjekttypen oder -attributen oder das Einfügen zusätzlicher Geschäftsregeln. In bestimmten Fällen kann daher an dieser Stelle ein Rücksprung in Methodenfragment 1 erforderlich sein, um zusätzlich identifizierte Geschäftsobjekttypen zu definieren oder das Metadatenmodell des BDD anzupassen. Notwendige Veränderungen sind durch den Konzern-Datensteward zu dokumentieren, um eine konsistente Aktualisierung des BDDs zu gewährleisten.

5.6.2.6 BDD aktualisieren

Die vorherigen Aktivitäten dienen der Sammlung von Metadaten zu den Datenobjekttypen und -attributen aus Sicht der IT (siehe entsprechende Klasse im BDD-Metadatenmodell in Abbildung 5-9). Diese sollen die bereits im BDD gespeicherten fachlichen Metadaten komplementieren. Für die Aktualisierung der Metadaten im BDD ist den in Methodenfragment 1 definierten Metadatenmanagementprozessen (siehe Aktivitäten MF1: IV.1 und MF1: IV.2) zu folgen. Der Owner des Geschäftsobjekttyps (in der Regel der fachliche Datensteward) sollte die hinzuzufügenden Metadaten bewilligen und der entsprechende Metadatenadministrator diese anschliessend im BDD pflegen. Hierbei ist darauf zu achten, dass wichtige Informationen der Ergebnisdokumente von Methodenfragment 2 (z. B. der Stammdatenlandkarte, die u. a. die für Datenobjekttypen führenden Systeme darstellt) vollständig im BDD gepflegt werden.

MF2: II.4 BDD aktualisieren	
Ergebnis	Aktualisierte Metadaten im BDD
Rollen	Owner des Geschäftsobjekttyps (z. B. fachliche Datenstewards) [verantwortlich] Metadatenadministrator [ausführend] Technische Datenstewards [unterstützend]
Techniken	–
Input/Vorbedingungen	Vollständige Zuordnung von Daten- zu Geschäftsobjekttypen

Tabelle 5-28: MF2: II.4 BDD aktualisieren

5.6.2.7 Semantisches Stammdatenmodell ableiten

Die im BDD gespeicherten Metadaten werden für die Ableitung eines semantischen, unternehmensweit gültigen Stammdatenmodells benötigt, das wiederum logischer Ausgangspunkt für Aktivitäten der Systemgestaltung und Stammdatenverteilung ist [Loshin 2008, 119]. Das semantische Stammdatenmodell bildet die Grundlage für ein einheitliches Verständnis der Stammdatenobjekte auf Typebene über sämtliche operative, stammdatennutzende Applikationen hinweg. Dies ist insbesondere für den konsistenten Austausch von Stammdatenobjekten zwischen führenden bzw. zentralen Stammdatenapplikationen und lokalen Applikationen (siehe Architekturmuster in Kapitel 5.7.1) und somit für die operative Stammdatenintegration von Bedeutung. Bei der DTAG (operatives Anwendungsszenario) besteht die aufwendigste Aktivität in der Erstellung des applikationsübergreifenden, semantischen Datenmodells auf fachlogischer Architekturebene in Form des BOM (siehe Kapitel 4.2.2.3). Für das analytische Anwendungsszenario (siehe Fallstudien SBB Cargo) sowie das Anwendungsszenario der Konsolidierung (siehe Fallstudie Bosch Rexroth) ist diese Aktivität optional.

MF2: III.1 Semantisches Stammdatenmodell ableiten	
Ergebnis	Semantisches Stammdatenmodell mit standardisierten Datenobjekttypen
Rollen	Strategischer Datensteward [verantwortlich] Datenarchitekt [ausführend] Fachliche Datenstewards, Technische Datenstewards, Konzern-Datensteward [unterstützend]
Techniken	T.III.1 Abgleich mit bestehenden Objektdefinitionen der CCL (optional) T.III.2 Definieren neuer BIEs gemäss ISO 11179-5 (optional) T.III.3 Kontextualisierung der Datenobjekttypen T.III.4 Schema Matching
Input/Vorbedingungen	Metadaten des BDD, Datenmodelle der Applikationen

Tabelle 5-29: MF2: III.1 Semantisches Stammdatenmodell ableiten

Nachdem die vorherigen Aktivitäten von Methodenfragment 1 und Methodenfragment 2 zu einer konsistenten Metadatenbasis in Bezug auf die wesentlichen Geschäfts- und Datenobjekttypen eines Unternehmens beitragen und diese eindeutig definieren, gilt es nun, diese Metadaten in die Datenmodelle zu überführen. Während die Bezeichnungen der Objekttypen und ihrer Attribute direkt aus dem BDD übernommen werden können, ist die Abbildung der Semantik (insbesondere des Verwendungskontextes) weniger trivial. Wie in Kapitel 5.6.1 beschrieben, bietet die CCTS-Spezifikation einen systematischen Ansatz ein Verfahren, durch das Datenobjekttypen mit Metadaten annotiert werden können (Kontextualisierung). Dadurch bildet das resultierende semantische Datenmodell einen Teil der Semantik sowie des Verwendungskontextes von Datenobjekttypen ab. Das Hauptaugenmerk im Rahmen der vorliegenden Methode liegt daher auf Technik T.III.3, in der das Kontexttreiberprinzip der CCTS angewendet wird.

Ansonsten sind die Techniken der Aktivität an die CCTS angelehnt (vgl. [Lemm 2007, 29-33]). Die Durchführung der Techniken T.III.1 bis T.III.3 kann durch den in Kapitel 5.6.4 beschriebenen Modellierungsprototyp teilautomatisiert werden. Zudem müssen die verschiedenen Datenmodelle der Applikationen (bzw. Datenbankschemata der zugehörigen Datenbanken) auf das semantische Stammdatenmodell abgebildet werden [Spath et al. 2009, 7], wofür Schema-Matching-Verfahren genutzt werden können.

5.6.3 Techniken

Die vorherigen Kapitel beschreiben die Aktivitäten des zweiten Methodenfragmentes und verweisen auf einzelne Techniken, die als Arbeitsanweisung bei der Ausführung der Aktivität behilflich sind. Hierbei wird teilweise auf Techniken zurückgegriffen, die sich in der Praxis bewährt haben oder bereits in Methodenfragment 1 verwendet wurden (siehe Kapitel 5.5.3). Tabelle 5-30 zeigt einen Überblick über die in Methodenfragment 2 verwendeten Techniken mit möglichen Referenzen auf Methodenfragment 1 bzw. Publikationen, welche die Techniken erläutern. Das Kapitel beschreibt abschliessend die Übertragung und Anwendung der Techniken in der vorliegenden Me-

thode und führt Beispiele an. Auf die in Methodenfragment 1 spezifizierten Techniken wird an dieser Stelle nicht noch einmal im Detail eingegangen.

Nr.	Name	Aktivität	Referenz
T.I.1	Identifikation der Daten- und System-Owner	MF2: I.1	Siehe T.I.2 von Methodenfragment 1 (Kapitel 5.5.3)
T.II.1	Inhaltsanalyse Applikationsdokumentation	MF2: II.1, MF2: II.2a	Zur qualitativen Inhaltsanalyse allgemein siehe T.II.1 von Methodenfragment 1 (Kapitel 5.5.3)
T.II.2	Einzelinterviews	MF2: II.1, MF2: II.2a	Siehe T.II.2 von Methodenfragment 1 (Kapitel 5.5.3)
T.II.3	Workshops / Gruppendiskussion	MF2: II.2a, MF2: II.3	Siehe T.III.4 von Methodenfragment 1 (Kapitel 5.5.3)
T.II.4	Zuordnung von Datenobjekttypen zu Applikationen	MF2: II.2b	[Martin/Leben 1989, 225-226], [IMG 1996, 179-196]
T.III.1	Abgleich mit bestehenden Objektdefinitionen der CCL	MF2: III.1	[Lemm 2007, 29-33]
T.III.2	Definition von BIEs gemäss ISO 11179-5	MF2: III.1	[ISO/IEC 2005b, 7-11], [UN/CEFACT 2009b, 54ff.]
T.III.3	Kontextualisierung der Datenobjekttypen	MF2: III.1	[UN/CEFACT 2009b, 107-114], [Vogel et al. 2008, 24-25]
T.III.4	Schema Matching	MF2: III.1	[Rahm/Bernstein 2001], [Conrad 2002], [Leser/Naumann 2007], [Batini et al. 1986]

Tabelle 5-30: Techniken des Methodenfragmentes 2

T.II.1 Inhaltsanalyse Applikationsdokumentation

Die Technik T.II.1 gilt der strukturierten Auswertung vorhandener Dokumentation zu Applikationen eines Unternehmens, um Datenobjekttypen und -attribute zu identifizieren (Aktivität MF2: II.1) und zu beschreiben (Aktivität MF2: II.2a). Wichtige Dokumente für die Inhaltsanalyse sind bestehende Applikationsbeschreibungen bzw. Applikationsverzeichnisse, IT-Bebauungspläne, die Datenmodelle der Applikationen sowie vorhandene Datenflussdiagramme und Schnittstellenbeschreibungen [McComb 2003, 176].

T.II.2 Einzelinterviews; T.II.3 Workshops / Gruppendiskussion

Analog zu den Interviews mit fachlichen Datenstewards sowie Vertretern des Fachbereiches in Methodenfragment 1 sollten die Befragungen semi-strukturiert auf Grundlage eines Interviewleitfadens durchgeführt werden. Als Ansprechpartner sind insbesondere Applikationsentwickler, System-Owner und Datenarchitekten zu berücksichtigen.

Bei den Workshops liegt der Fokus auf der Abstimmung unterschiedlicher Definitionen von Datenobjektattributen über mehrere Applikationen hinweg sowie mit den im BDD gespeicherten Definitionen aus Sicht des Fachbereiches. Grundlage hierfür ist die eindeutige Zuordnung von Datenobjektattributen zu Geschäftsobjektattributen.

T.II.4 Zuordnung von Datenobjekttypen zu Applikationen

Die Zuordnung von Datenobjekttypen zu einzelnen Applikationen bzw. Applikationsfunktionen kann bei geringer Anzahl von Applikationen bzw. Datenobjekttypen graphisch dargestellt werden [IMG 1996, 182], häufiger wird sie jedoch in einer CRUD-Matrix dokumentiert [Martin/Leben 1989, 225-226]. Letztere spezifiziert zusätzlich die Art des Zugriffs. Dem Akronym entsprechend kann der Zugriff erzeugend (Create), lesend (Read), schreibend bzw. verändernd (Update) oder löschend (Delete) erfolgen. Eine beispielhafte Vorlage für die Erstellung einer CRUD-Matrix als Stammdatenlandkarte ist Anhang D.5 zu entnehmen.

T.III.1 Abgleich mit bestehenden Objektdefinitionen der CCL (optional)

Aufbauend auf der Argumentation in Kapitel 3.2, dass eine effiziente Integration, u. a. von Stammdaten, durch die Verwendung von Standards unterstützt wird [Chari/Seshadri 2004, 59; Schemm 2008, 37], besteht ein optionaler Schritt bei der Datenobjektmodellierung in der Standardisierung der Namen von Datenobjekttypen. Mit der CCL liefert die UN/CEFACT ein Verzeichnis für Datenobjekttypen und für Datenobjektattribute, deren Benennung (der DEN) gemäss den Namenskonventionen der ISO 11179-5 standardisiert ist. Die CCL enthält für jede der Datenobjekttypen eine kurze Definition, welche die Komponente genauer spezifiziert (vgl. [UN/CEFACT 2009a]). Dadurch wird ein einheitliches Verständnis der Datenobjekttypen gefördert. Durch Vergleich der im BDD enthaltenen Definitionen mit denen der CCL können die bereits in der CCL enthaltenen Datenobjekttypen identifiziert und für das eigene Datenobjektmodell verwendet werden. Name und Definition können direkt aus dem BDD übernommen werden.

Datenobjekttypen, für die kein BIE in der CCL gefunden wird, sind separat auf einer Liste zu vermerken. Für sie sind neue DENs zu definieren (siehe Technik T.III.2). Die Identifikation von korrespondierenden BIEs kann durch Software-Werkzeuge, welche die Inhalte der CCL enthalten, unterstützt werden. Sie ermöglichen durch einen automatisierten Abgleich der Definitionen die Bestimmung potenziell passender Datenobjekttypen, aus denen der Modellierer auswählen kann. Ein entsprechendes Beispiel wird in Kapitel 5.6.4 beschrieben.

T.III.2 Definition neuer BIEs gemäss ISO 11179-5 (optional)

Für Datenobjekttypen, für die keine Entsprechungen in der CCL gefunden wurden, sind neue BIEs zu definieren. Dies umfasst die Spezifikation sämtlicher durch die CCL vorgegebenen Eigenschaften, insbesondere des standardisierten DENs sowie einer dazugehörigen Kurzdefinition. Gemäss ISO 11179-5 sind für jeden DEN der Object Class Term, der Property Term, der Representation Term sowie der Qualifier Term festzulegen [ISO/IEC 2005b, 7-10]. Zusätzlich sind zu jeder Datenobjekttypen eine kurze Definition sowie ein oder mehrere sogenannte Business Terms an-

zugeben [UN/CEFACT 2009b, 54ff.]. Die Definition gibt die genaue Bedeutung der Datenobjektkomponente wieder, während Business Terms synonyme Begriffe bezeichnen, die für die Datenobjektkomponente gebräuchlich sind. Das BDD sollte hierbei als Informationsquelle genutzt werden.

Analog zu Technik T.III.1 kann die Technik durch CCTS-konforme Modellierungswerkzeuge, welche aus Definitionen mögliche DENs generieren, unterstützt werden (siehe Kapitel 5.6.4).

T.III.3 Kontextualisierung der Datenobjekttypen

Für das Ziel der Ableitung eines semantischen Stammdatenmodells ist die Technik zur Kontextualisierung wesentlich. Bei der Kontextualisierung der Datenobjekttypen und der Datenobjektattribute geht es vordergründig um die Integration der fachlichen Metadaten zum Verwendungskontext in die Datenmodelle. Wie bereits in Kapitel 5.6.1 angedeutet, bietet die CCTS-Methode Konzepte, mit denen Datenobjekttypen (ABIEs) und Datenobjektattribute (BBIEs bzw. ASBIEs) mit ihrem Verwendungskontext semantisch annotiert werden können (Kontexttreiberprinzip). Die Technik adaptiert das Kontexttreiberprinzip für die Methode zur Stammdatenintegration, um Metadaten des BDD in die Datenmodelle zu übernehmen und dadurch applikationsspezifische Datenobjekttypen und Datenobjektattribute als Varianten im konsolidierten Stammdatenmodell abzubilden.

Zur Kontextualisierung können zum einen die vordefinierten Kontextkategorien (siehe Tabelle 5-21) genutzt werden. Folgende Zuordnung der im BDD gespeicherten Kontextmetadaten bezüglich (siehe Abbildung 5-9) zu den Kontextkategorien ist möglich:

- Die BDD-Metadatenklasse Geschäftsprozess kann durch die Kontextkategorie Business Process abgebildet werden.
- Die BDD-Metadatenklasse Applikation wird teilweise durch die Kontextkategorie System Capabilities abgedeckt.
- Die Kontextkategorien Business Process Role und Supporting Role geben die BDD-Metadatenklasse Rolle wieder.

Des Weiteren ist es möglich, eigene Kontextkategorien zu definieren sowie für einzelne Kontextkategorien unternehmensspezifische Wertelisten zu hinterlegen. Im Aktionsforschungsprojekt Daimler AG (siehe Kapitel 6.1.2) korrespondieren die Werte der Kontextkategorie System Capabilities Context mit den Codes der Applikationen, welche die identifizierten Datenobjektattribute nutzen. Dadurch bildet das applikationsübergreifend gültige, semantische Stammdatenmodell (siehe Abbildung E-2 in Anhang E) ab, in welchen Applikationen die entsprechenden Datenobjektattribute verwendet werden. Beispielsweise entsprechen die mit dem Kontextwert CRM annotierten Datenobjektattribute und Datenobjekttypen den Datenobjektkomponenten des CRM-

Systems. Durch Einschränkung der Kontextkategorie auf bestimmte Kontextwerte können applikationsspezifische Datenmodelle generiert werden.

T.III.4 Schema Matching

Schema Matching ist eine essenzielle Aufgabe bei der Integration von Daten aus verschiedenen Datenquellen [Rahm/Bernstein 2001, 348; Gal 2006, 2]. Es integriert unabhängig voneinander entwickelte und dadurch heterogene Datenschemata³⁷ mittels Korrespondenzen (auch Mapping genannt) zwischen Schemaelementen [Batini et al. 1986, 323; Rahm/Bernstein 2001, 334]. Das Schema Matching versucht, Korrespondenzen zwischen semantisch äquivalenten Elementen automatisch zu erkennen [Leser/Naumann 2007, 144]. Die Arbeit verwendet im Folgenden den Begriff des Schema Matching, um die Notwendigkeit (semi-)automatischer Verfahren zur Integration von Datenschemata in grossen Unternehmen zu betonen und es vom manuellen Schema Mapping abzugrenzen. Ein rein manuelles Mapping der für diese Unternehmen typischen, grossen Anzahl und Komplexität (zahlreiche Attribute, Fremdschlüsselbeziehungen und Relationen) der Schemata ist mit angemessenem Aufwand und geringer Fehlerrate nicht zu bewerkstelligen [Rahm/Bernstein 2001, 334; Melchert 2006, 232]. Grundlage des Schema Matching sind Verfahren des Schema Mapping, also der gerichteten Abbildung eines Schemas (im vorliegenden Fall z. B. das Datenschema einer Datenbank) auf ein anderes (das semantische Stammdatenmodell). Dabei werden Korrespondenzen zwischen Schemaattributen (sogenannte logische Mappings) in der Regel aus Wertkorrespondenzen (auf Instanzebene) zwischen Quell- und Zielschema abgeleitet [Leser/Naumann 2007, 125f.]. Das logische Mapping wird anschliessend in eine Transformationsvorschrift in Form einer Anfragesprache (z. B. SQL) übersetzt. Durch Ausführung der Transformationsvorschrift können die Instanzdaten in das gewünschte Zielschema transformiert werden.

Das Vorgehen beim Schema Matching besteht im Wesentlichen aus zwei Schritten: dem Schemavergleich zur Ermittlung von Korrespondenzen und Konflikten zwischen Quell- und Zielschema und der Schematransformation, bei der die Konflikte behoben und das Quell- in das Zielschema überführt wird [Spaccapietra et al. 1992, 84].³⁸ Konflikte umfassen die in Kapitel 2.1.3 genannten Heterogenitäten auf logischer Ebene, wobei die grösste Herausforderung in der Beseitigung der strukturellen (Schlüssel, Schachtelung) und semantischen (homonyme und synonyme Benennungen) Heterogenität liegt [Schmitt/Saake 2005, 478f.]. Die Behebung der Konflikte ist durch Umbenennung oder Normalisierung einzelner Attribute möglich.

³⁷ Der Begriff des (konzeptionellen) Datenschemas bezeichnet das für ein konkretes Datenbankmanagementsystem formulierte Datenmodell [Jarosch 2010, 24f.]. Es spezifiziert die Datenobjekttypen, Datenobjektattribute und Datenobjekttypbeziehungen des Datenmodells in einer formalen Sprache wie z. B. XML oder SQL.

³⁸ Andere Autoren differenzieren beim Vorgehen noch weitere Phasen der Vorintegration oder Schemaangleichung (vgl. [Batini et al. 1986, 336f.], [Parent/Spaccapietra 1998, 166]), deren Aktivitäten jedoch als Bestandteil der beiden beschriebenen Phasen betrachtet werden.

Beim Schema Matching kommen sogenannte Matcher zum Einsatz, welche die Struktur, Integrationsbedingungen und Beispieldaten von Schemata analysieren, um Korrespondenzen zu identifizieren. Die Matcher vergleichen Attribute von Quell- und Zielschema paarweise und verwenden Ähnlichkeitsmasse zur Bestimmung von Korrespondenzen. Bei Überschreiten eines vorher definierten Schwellwertes beim Vergleich zweier Attribute werden diese als Korrespondenz vorgeschlagen. Matchern liegen verschiedene Matching-Verfahren zugrunde, die abhängig von den Informationen, die sie zum Vergleich nutzen, in vier Kategorien eingeteilt werden können [Rahm/Bernstein 2001, 338-343; Leser/Naumann 2007, 145-152].

- *Schemabasiertes Schema Matching* verwendet für den Vergleich ausschliesslich Schemainformationen, wie z. B. Namen, Struktur (Beziehungstypen) oder Datentypen, ohne Instanzen einzubeziehen. Schemabasierte Schema-Matching-Verfahren können zusätzlich noch dahingehend unterschieden werden, ob sie lediglich einzelne Schemaelemente vergleichen (Element Matching) oder auch Kombinationen aus Schemaelementen, also komplexe Strukturen (Structure Matching).
- *Instanzbasiertes Schema Matching* hingegen betrachtet nur die aus den Schemata instanziierten Daten und versucht aus den Attributwerten kennzeichnende Eigenschaften abzuleiten, die dann verglichen werden. Voraussetzung hierfür ist die Existenz von Beispieldaten für beide Schemata. Instanzbasiertes Schema Matching kommt häufig bei semistrukturierten Daten zur Anwendung, bei denen kein Schema vorhanden ist.
- *Linguistisches Schema Matching* nutzt Namen von Schemaelementen (Name Matching) sowie textuelle Beschreibungen (Description Matching) für den Vergleich. Korrespondenzen werden bei diesem Verfahren aufgrund von Namensgleichheit bzw. Namensähnlichkeit identifiziert.
- *Constraint-basiertes Schema Matching* wertet in Datenschemata enthaltene Bedingungen (constraints) aus, die Datentypen, Wertebereiche, Schlüsselbeziehungen oder Kardinalitäten einschränken. Die alleinige Verwendung dieses Verfahrens liefert häufig qualitativ unzureichende Ergebnisse, weshalb constraint-basiertes Schema Matching in Kombination mit anderen Verfahren angewendet wird.

Heutige Matcher kombinieren in der Regel mehrere dieser Verfahren, indem sie diese entweder verschachteln (sogenannte hybride Matcher) oder unabhängig voneinander ausführen und die Ergebnisse anschliessend verknüpfen (zusammengesetzte Matcher). Derartige Matcher bieten den Vorteil, dass sie die Nachteile individuell angewandter Verfahren ausgleichen und somit die Qualität der Matching-Ergebnisse verbessern. Daher kombiniert auch der im folgenden Kapitel beschriebene Modellierungsprototyp mehrere Matching-Verfahren. Des Weiteren nutzen Matching-Tools zusätzliche Informationen aus (Data) Dictionarys, aus Thesauri (z. B. zur Bestimmung synonymmer

Attributnamen) oder aus der Historie (bereits in der Vergangenheit identifizierte Korrespondenzen). Prototypische Implementierungen von Matchern, die jeweils eines oder mehrere der vorgestellten Verfahren nutzen, werden in der Literatur durch verschiedene Autoren beschrieben (vgl. [Rahm/Bernstein 2001, 343-348; Leser/Naumann 2007, 148-155]). Einschränkend ist anzumerken, dass die bestehenden Forschungsprototypen bisher kaum Eingang in kommerzielle Produkte gefunden haben, weswegen der Transfer der Erkenntnisse in die Unternehmen noch unzureichend ist [Melchert 2006, 231f.].

Eine vollständig automatisierte Identifikation und richtige Interpretation der Korrespondenzen ist aufgrund der oftmals in den Schemata nur implizit vorhandenen Semantik nicht möglich [Rahm/Bernstein 2001, 337; Gal 2006, 2]. So zeigt die Arbeit von LEGLER und NAUMANN, dass für diese beiden Schritte des Schema Matching die kommerziell verfügbaren Tools lediglich Vorschläge für den Anwender generieren können und teilweise falsche Ergebnisse liefern [Legler/Naumann 2007, 457-461]. Daher sind für das Schema Matching Experten mit umfangreichem Fachwissen über die Anwendungsdomäne erforderlich [Melchert 2006, 232], die Matching-Ergebnisse über graphische Benutzerschnittstellen überprüfen. Matcher sollten diesen Experten Matching-Kandidaten mit möglichst hoher Wahrscheinlichkeit vorschlagen.

5.6.4 Werkzeugunterstützung

Das grundsätzliche Defizit des Ansatzes, standardisierte, semantische Datenmodelle auf Grundlage der CCTS zu erstellen, war der Mangel an einer Toolunterstützung zur Modellierung [Vogel 2009, 205]. Zwar existieren mit dem H2-Toolset (vgl. [Becker et al. 2007a, 143ff.; Becker et al. 2008b]) des European Research Center for Information Systems der Westfälischen Wilhelms-Universität sowie mit dem Tool GEFEG.FX³⁹ erste kommerzielle Modellierungstools. Jedoch fokussieren diese sehr stark auf die Modellierung von Geschäftsdokumenten in B2B-Szenarien und setzen CCTS nicht vollumfänglich um (z. B. in Bezug auf das Kontexttreiberprinzip). Aufgrund dieses Mankos sowie der zunehmenden Reife der CCTS arbeitet die SAP Global Ecosystem and Partner Group zusammen mit der SAP Research seit 2007 an der Entwicklung eines entsprechenden Modellierungstools.⁴⁰

Das Modellierungstool soll die Integration von (Geschäfts-)Daten auf Typebene vereinfachen. Der realisierte Forschungsprototyp setzt die CCTS und das darin definierte Kontexttreiberprinzip als Konzept zur semantischen Datenmodellierung vollständig um (siehe Kapitel 5.6.1). Es unterstützt den Modellierer, CCTS-konforme Datenobjekt-komponenten zu definieren [Stuhec 2007, 3]. Darüber hinaus lassen sich Metada-

³⁹ Siehe http://www.gefeg.com/de/gefeg.fx/fx_kurz.htm.

⁴⁰ Der Autor der Dissertation war über das Forschungsprojekt ITAIDE (siehe Kapitel 1.3) in die Entwicklung des Prototyps involviert und trug zur Anforderungsanalyse sowie der Modellierung von Anwendungsfällen zu Testzwecken bei.

ten, wie z. B. Bezeichner, Multiplizitäten, Verwendungsregeln, aber auch natürlichsprachliche Definitionen und der Verwendungskontext der Datenobjektcomponenten, erfassen. Folglich können auch relevante, fachliche Metadaten aus dem BDD (oder anderen Repositorys) in das Datenmodell übernommen werden. Ziel dieser Übernahme ist es, die Semantik der Datenobjektcomponenten genauer zu spezifizieren und somit für notwendige Abgleiche mit anderen Datenmodellen (Schema Matching) verfügbar zu machen. Die im Tool vorgegebenen Verwendungskontexte, durch die Datenobjekttypen sowie Datenobjektattribute semantisch annotiert werden können, entsprechen denen der CCTS (siehe Tabelle 5-21). Sie können jedoch für unternehmensspezifische Anforderungen erweitert werden, indem neue Kontextkategorien hinzugefügt oder weitere Codelisten zu einer Kategorie angefügt werden.

Das Tool erfüllt somit zwei wesentliche Anforderungen für die Integration bzw. Integrierbarkeit von Stammdatenobjekten. Es baut auf einem internationalen Standard (CCTS) auf, der die Semantik von Datenobjektcomponenten bis auf Attributebene vereinheitlicht. Zusätzlich basiert das Tool auf einem semantischen Konzept, durch welches die Verwendung gleicher Datenobjektcomponenten in verschiedenen Verwendungskontexten (Unternehmensfunktion, Prozesse, Applikationen) abgebildet werden kann. Hierdurch können unabhängig voneinander entwickelte Datenmodelle und Schnittstellen der Applikationen sowie darin voneinander abweichend definierte Datenobjektattribute in einem kanonischen Datenmodell integriert bzw. föderiert werden. Das Modellierungstool generiert dieses Datenmodell halbautomatisiert, indem es Datenobjekttypen und -attribute auf CCTS-konforme Datenobjektcomponenten abbildet. Elemente weiterer Datenmodelle bzw. Schnittstellen werden jeweils in das kanonische Datenmodell überführt und müssen nicht mehr zwischen einzelnen Datenmodellen abgebildet werden, wodurch sich der Mapping-Aufwand mit steigender Anzahl der zu integrierenden Datenmodelle merklich reduziert.

Die beschriebenen Eigenschaften des Modellierungstools werden technologisch durch folgende Konzepte bzw. Komponenten in einer integrierten Architektur realisiert:

- Im Mittelpunkt der Architektur des Modellierungstools steht eine auf *Frame Logic* (F-Logic) basierende *Ontologie*, die einerseits die Speicherung und Integration von Datenmodellen (mittels semi-automatischen Schema Matching) erlaubt, andererseits das bereits beschriebene Kontexttreiberprinzip umsetzt. F-Logic ist eine formale Sprache zur Wissensrepräsentation (und damit zur Formulierung von Ontologien), die ihren Ursprung im Bereich der deduktiven Datenbanken hat. Sie vereint die Vorteile objektorientierter, frame-basierter Modellierungssprachen mit der wohldefinierten Semantik einer logikbasierten Sprache [Kifer et al. 1995, 742-746]. Zur Beschreibung von Informationen nutzt F-Logic die drei Komponenten Schemafakten (entsprechen Klassen in der objektorientierten Modellierung), Fakten (entsprechen Objekten bzw. Instanzen) und Regeln [Yu 2007, 20]. Durch die

Anwendung der Inferenzregeln können neue Informationen abgeleitet werden, wodurch F-Logic als regelbasierte, deduktive Ontologiesprache ähnlich RDF oder OWL verstanden werden kann. Als Kern der Ontologie ist auch das Metamodell der CCTS [UN/CEFACT 2009b, 52] in F-Logic modelliert. Hierbei wird die Eigenschaft von Ontologien genutzt, Modellentitäten (in diesem Fall Datenobjekt-komponenten der CCTS) und deren Beziehungen formal zu beschreiben und mit Metadaten annotieren zu können [Hahn 2005, 279f.].

- Die auf F-Logic sowie der Graphentheorie aufbauende *Kontextlogik* ist Bestandteil der *Kontexttreiberkomponente* [Pentcheva 2007, 42; Yu 2007, 38ff.]. Mit deren Hilfe sind eine kontextsensitive (und damit semantische) Suche in den Datenmodellen sowie die Bildung kontextspezifischer Teilmodelle möglich.
- Die auf F-Logic basierende Ontologie ist Voraussetzung für die Integration von nicht CCTS-konformen Datenmodellen durch *semi-automatisches Mapping*. Hierdurch können insbesondere bereits im Unternehmen vorliegende, applikationsspezifische Datenmodelle auf das generische Datenmodell abgebildet werden. Für das Schema Matching zur Integration von Datenmodellen kombiniert der Prototyp schema-, instanz- und constraint-basierte Matching-Verfahren (siehe Kapitel 5.6.3). Die Komponente ist mit Hilfe einer Menge von Matchern realisiert, die jeweils die im Tool-Repository gespeicherten, CCTS-konformen Entitätsnamen, Datentypen, dazugehörige Definitionen sowie Kontextkategorien und -werte durchsuchen und mit einer potenziell neuen Datenobjektkomponente vergleichen. Auf Grundlage dieser zahlreichen, gewichteten Vergleiche wird eine Korrespondenzwahrscheinlichkeit zweier Datenobjektkomponenten berechnet, die den Modellierer bei der Entscheidung unterstützt, ob zwei Elemente gleich (synonym) sind. Dadurch wird er u. a. beim Versuch, einen neuen Datenobjekttyp oder ein neues Datenobjektattribut anzulegen, auf ein gleiches oder ähnliches Element verwiesen, das z. B. in einem anderen Kontext bereits definiert ist. Dieses Element kann der Modellierer dann wiederverwenden und – falls nötig – an seine spezifischen Anforderungen anpassen. Die finale Entscheidung liegt dabei beim Modellierer.
- Werden keine Entsprechungen gefunden, unterstützt das Modellierungstool bei der natürlichsprachlichen Erzeugung von Bezeichnungen der Datenobjekt-komponenten gemäss den in Kapitel 5.6.1 beschriebenen DENs. Hierzu schlägt eine *Natural Language Processing-Komponente* des Modellierungstools Bezeichnungen vor, die mit den Namensregeln der ISO 11179 konform sind, und gewährleistet, dass neu angelegte Elemente den Namensregeln entsprechen. Die zugehörigen Definitionen und Fachbegriffe (die der Fachbereich verwendet) werden in einer Referenzbibliothek gespeichert.

- Auf Grundlage eines Vergleichs semantischer Nähe zu existierenden Datenobjekt-komponenten in der Ontologie werden neu eingefügte oder aus anderen Modellen importierte BIEs oder BDTs durch das Modellierungstool im Datenmodell positioniert. Dies ist Aufgabe der *Konsolidierungs- und Normalisierungskomponente*, die für eine automatisierte Transformation des Datenmodells in eine normalisierte Repräsentation sorgt.
- Die *webbasierte Benutzeroberfläche* unterstützt die Modellierung, indem sie dem Anwender die Definition von Datenobjekt-komponenten sowie deren Kontextualisierung durch einfache Dropdown-Listen erleichtert. Die Navigation durch das Datenmodell ist durch die hierarchische Anordnung von ABIEs (existenzabhängige ABIEs stehen rechts) intuitiv möglich. Dies ermöglicht es, Vertreter des Fachbereiches einfacher in den Modellierungsprozess zu involvieren.

Der implementierungsunabhängigen Definition des CCTS-Standard folgend, der keine bestimmte Syntax vorgibt, kann das Modellierungstool automatisiert verschiedene Syntaxrepräsentationen der Datenmodelle generieren, wie beispielsweise XML oder SQL-basierte, relationale Datenbankschemata.

5.7 Methodenfragment 3: Stammdatenintegrationsarchitektur gestalten

Die Stammdatenintegrationsarchitektur steuert den verteilten Zugriff auf Stammdaten sowie die Replikation und die Verteilung von Datenelementen zwischen Applikationen mit dem Ziel, eine applikationsübergreifende Konsistenz zu gewährleisten [DAMA 2009, 181]. In global agierenden Unternehmen, in denen die Speicherung von Stammdaten in einer einzelnen Applikation nicht realisierbar ist, ergibt sich zwangsläufig die Notwendigkeit der Datenreplikation und -verteilung und somit Probleme der Datenfragmentierung und der redundanten Datenhaltung [Priglinger/Friedrich 2008, 15]. Daher stellt sich die Frage, mittels welcher Architekturansätze und Verteilungsmechanismen sich das aus dem BDD abgeleitete, applikationsübergreifende semantische Stammdatenmodell im Unternehmen implementieren und somit ein transparenter, verteilter Zugriff auf eindeutige Stammdatenobjekte (auf Instanzebene) realisieren lässt. Diese Frage behandelt das dritte Methodenfragment zum Entwurf und Gestaltung einer Stammdatenintegrationsarchitektur. Voraussetzung hierfür ist einerseits die Kenntnis der fachlichen Anforderungen an eine integrierte Stammdatenbasis, andererseits die Integration auf Typebene. Die beiden Aspekte waren Bestandteil der vorherigen Methodenfragmente. Das Methodenfragment nutzt daher das BDD und das semantische Stammdatenmodell als Grundlage für die Architekturgestaltung [Vetere/Lenzerini 2005, 893] und die Integration von Stammdatenobjekten und Stammdatenelementen.

Das Kapitel zu Methodenfragment 3 ist wie folgt aufgebaut: Abschnitt 5.7.1 stellt einleitend konzeptionelle Architekturmuster vor, die sich für die Speicherung und Distribution von Stammdaten in verteilten Applikationsarchitekturen in der Praxis bewährt haben, und charakterisiert diese anhand einheitlicher Merkmale. Daraufhin beschreibt Abschnitt 5.7.2 das Vorgehensmodell des Methodenfragmentes, dessen Ziel die Entwicklung und Implementierung einer Stammdatenintegrationsarchitektur ist. Analog zu den vorherigen Methodenfragmenten werden zusätzlich Techniken (Abschnitt 5.7.3) und Werkzeugunterstützung (Abschnitt 5.7.4) beschrieben.

5.7.1 Architekturmuster für das Stammdatenmanagement

Sowohl Wissenschaft als auch Praxis diskutieren verschiedene konzeptionelle Architekturmuster für die Verwaltung von Stammdaten in verteilten Applikationsarchitekturen (vgl. [Loser et al. 2004], [Schemm 2008, 157-159], [DAMA 2009, 181-184], [Berson/Dubov 2007, 76-78], [Spath et al. 2009, 2-6], [Dreibelbis et al. 2008, 226ff.], [Loshin 2008, 165-171]). *Architekturmuster* bündeln Erfahrungen und Wissen für die Lösung immer wiederkehrender Probleme bei der Gestaltung von Applikations- und Integrationsarchitekturen [Lutz 2000, 66; Dreibelbis et al. 2008, 224]. Sie enthalten Vorlagen zur strukturellen Organisation von Applikationen (und ihrer Beziehungen zueinander) in einem IS auf einer konzeptionellen Ebene [Buschmann et al. 1996, 8].

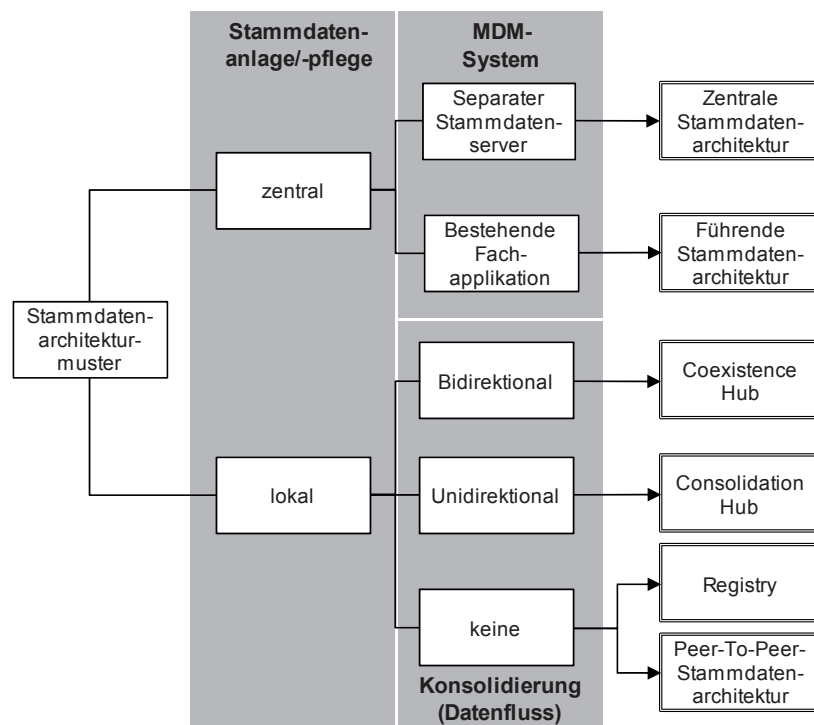


Abbildung 5-15: Grundlegende Architekturmuster für das Stammdatenmanagement

Zur Strukturierung und Abgrenzung der in der Literatur diskutierten *Stammdatenarchitekturmuster* können diese in einem ersten Schritt anhand des Ortes der Anlage

bzw. der Pflege der Stammdatenobjekte unterschieden werden (linke, grau hinterlegte Spalte in Abbildung 5-15).

Datenelemente von Stammdatenobjekten können entweder zentral in einer einzigen Applikation angelegt und dauerhaft gespeichert werden oder dezentral in den Fachapplikationen eines Unternehmens. Bei der zentralen Stammdatenanlage und -pflege kann zudem eine Unterscheidung dahingehend getroffen werden, ob die Stammdaten in einer speziell für das Stammdatenmanagement vorgesehenen Applikation, einem Stammdatenserver, oder in einer Fachapplikation (z. B. einem ERP- oder einem CRM-System) gehalten werden [Schemm 2008, 157f.; Spath et al. 2009, 2f.].

Bei einer *zentralen Stammdatenarchitektur* hält eine separate Applikation Stammdaten vor, von wo aus die Stammdaten entweder in die angebundenen Applikationen verteilt werden (Push-Verfahren) oder die Applikationen die Stammdaten auf dem zentralen System abrufen (Pull-Verfahren) [Pohland 2000, 188f.; Legner/Otto 2007, 9]. Diesem Architekturmuster liegt ein harmonisiertes Datenmodell im zentralen Stammdatensystem zugrunde, weswegen es den höchsten Grad an Integration und Aktualität der Daten verspricht [Spath et al. 2009, 3]. Die meisten kommerziell erhältlichen MDM-Tools, wie z. B. SAP NetWeaver MDM oder IBM InfoSphere MDM, realisieren eine zentrale Stammdatenarchitektur [Liebhart 2007, 290; Spath et al. 2009, 21]. Bei der *führenden Stammdatenarchitektur* übernimmt eine Fachapplikation die Rolle des zentralen Stammdatenservers [Legner/Otto 2007, 9; Spath et al. 2009, 2f.]. In der Regel ist die Fachapplikation für eine bestimmte Stammdatenklasse – beispielsweise ein PLM-System für Produktstammdaten oder ein CRM-System für Kundenstammdaten – und nicht für sämtliche Stammdaten führend. Der Umfang der im zentralen bzw. führenden System gepflegten Datenobjektattribute kann differieren. Sowohl die zentrale Pflege sämtlicher Attribute ist möglich als auch die Festlegung einer Teilmenge global gültiger Attribute, die zentral gepflegt und durch lokale Attribute in den angebundenen Applikationen angereichert werden.

Die zentrale Stammdatenanlage und -pflege ist eine in der Praxis häufig präferierte Lösung. Das schweizerische Maschinenbauunternehmen Oerlikon nutzt SAP NetWeaver MDM als zentrales System zur Anlage und Pflege seiner Debitorenstammdaten [Siebertz 2009, 13]. Diese Lösung ersetzte die vorherige Architektur, bestehend aus sieben SAP-ERP-Systemen der einzelnen Geschäftsbereiche, in denen Kundenstammdaten unabhängig voneinander angelegt wurden. Die Folge waren isolierte, inkonsistente Kundendatenbestände (u. a. synonyme und homonyme Kundennummern), die zu wachsenden Problemen bei geschäftsbereichsübergreifenden Marktaktivitäten führten. Auf Basis der im MDM einheitlich definierten Datenobjektattribute (u. a. global genutzte Kontengruppen) wurden die Kundenstammdaten aus den sieben ERP-Systemen konsolidiert und in bereinigter Form wieder an diese

verteilt. Damit verbunden ist die Anpassung und Zentralisierung des Stammdatenanlageprozesses, der basierend auf einem systemgestützten Workflow die redundante Speicherung von Debitoren im zentralen MDM-System verhindert.

Der Konsumgüterhersteller Beiersdorf verwaltet globale Stammdaten wie beispielsweise Artikelnummern, Produkthierarchien und Stücklisten in einer 2004 eingeführten zentralen Applikation [Rupprecht 2007, 13]. Die Applikation (SAP PLM) fungiert als zentraler Stammdatenserver und verteilt geänderte globale Stammdatenattribute periodisch alle drei Stunden über einen ALE-Verteilungsmechanismus an fünf lokale SAP R/3-Systeme sowie eine Menge von globalen Applikationen wie z. B. das SCM-System SAP APO. Die Daten werden in den empfangenden Applikationen verbucht und überschreiben veraltete Inhalte. Anschliessend können die Mitarbeiter in den Tochtergesellschaften weitere Attribute ergänzen.

Bei der SBB Cargo werden sämtliche Kundenstammdaten (inklusive Angebote und Verträge) führend in einem CRM-System angelegt und verwaltet (siehe Kapitel 4.2.1). Das CRM-System versorgt das ERP-System sowie angebundene Reporting-Systeme mit einem harmonisierten Kundenstamm für die operativen und analytischen Prozesse des Unternehmens.

Die Architekturmuster, die auf einer lokalen Stammdatenanlage und -pflege basieren, lassen sich am besten anhand Art und Grad der Konsolidierung der Stammdaten differenzieren. Den meisten Architekturmustern gemein ist die Konsolidierung der dezentral angelegten Stammdaten über ein *Hub*, d. h. eine zentrale Plattform, der ein harmonisiertes Stammdatenmodell zugrunde liegt und in der die Stammdaten virtuell oder physisch integriert werden [Berson/Dubov 2007, 78; Dreibelbis et al. 2008, 235]. Von da aus werden sie anschliessend wieder an die Fachapplikationen verteilt. Die verschiedenen Hub-Architekturen unterscheiden sich primär aufgrund der Art der Interaktion zwischen dem zentralen System und den angebotenen Fachapplikationen [Radcliffe et al. 2006, 2] sowie des Umfangs der in dem Hub gehaltenen Datenobjektattribute [Loshin 2008, 166].

Das *Consolidation Hub* dient ausschliesslich der Zusammenführung von Stammdaten in einer zentralen Applikation mit dem Ziel, Berichts- und Kontrollsysteme mit konsolidierten Stammdaten zu versorgen [Radcliffe et al. 2006, 3f.; Dreibelbis et al. 2008, 26f.]. Ein Rückfluss der Stammdaten in die Fachapplikationen ist nicht vorgesehen, womit das Architekturmuster im Wesentlichen den Konzepten der Data Warehouses bzw. Operational Data Stores (vgl. [Inmon 1999]) entspricht. Eine derartig realisierte Stammdatenarchitektur ist daher für ein analytisches Anwendungsszenario von Interesse (siehe Kapitel 4.4). Dahingegen extrahiert das *Coexistence Hub* (auch Reconciliation Hub) die Stammdaten (bzw. eine Teilmenge an Attributen) aus den angebotenen Applikationen, harmonisiert diese auf Grundlage des Datenmodells des

Hubs und spielt sie wieder in die Applikationen zurück [Radcliffe et al. 2006, 5f.; Dreibelbis et al. 2008, 242-249]. Auch in dieser Architekturvariante verfügen die Fachapplikationen lediglich über lesenden Zugriff, jedoch werden die Stammdaten in sämtlichen angebotenen Applikationen über das Hub harmonisiert.

Als eine Spezialform der Hub-Architekturen ist das *Registry* (Verzeichnis) zu betrachten, bei der Datenelemente nicht zentral harmonisiert und verteilt werden. Vielmehr werden die Stammdaten dezentral und bilateral zwischen Applikationen ausgetauscht [Loser et al. 2004, 4]. Das Verzeichnis enthält nicht die eigentlichen Datensätze, sondern lediglich eine minimale Anzahl der zur eindeutigen Identifikation der Stammdatenobjekte benötigten Attribute sowie einen Verweis auf den Speicherort des jeweiligen Objektes in Form von Metadaten [Dreibelbis et al. 2008, 240f.]. Somit werden die Stammdaten hier nicht physisch, sondern rein virtuell bzw. föderiert integriert, weswegen für dieses Architekturmuster auch der Begriff *Virtual Hub* gebräuchlich ist [Loshin 2008, 166-168]. Der Umfang der zur eindeutigen Identifikation benötigten Datenobjektattribute kann zwischen verschiedenen Stammdatenklassen variieren. Architekturen nach dem Registry-Prinzip finden häufig im überbetrieblichen Kontext Anwendung (z. B. beim Global Data Synchronisation Network). Ohne jegliche zentrale Koordination werden Stammdaten in *Peer-To-Peer Architekturen* gehalten [Spath et al. 2009, 5f.]. Grundlage für eine Integration der in diesem Muster voneinander unabhängigen Applikationen mit jeweils eigener Stammdatenhaltung ist eine semantische Standardisierung durch eine zusätzliche Abstraktionsschicht (Adapter bzw. Wrapper) und die bilaterale Replikation der Daten.

Das schweizerische Biotechnologie- und Pharmaunternehmen Novartis hat einen schrittweisen Übergang von den vormals dezentral in mehreren Applikationen (u. a. SAP R/3, SAP SRM, proprietäre Applikationen) gehaltenen Lieferantenstammdaten zu einer zentralen Architektur für Lieferantenstammdaten definiert. Hierdurch will Novartis das bestehende Manko fehlender Festlegung, welche Applikation für bestimmte Stammdatenobjekte und deren Konsolidierung verantwortlich ist, beseitigen. Zur Erreichung dieses Ziels wird das neu eingeführte SAP NetWeaver MDM in einem ersten Schritt dazu genutzt, die Lieferantenstammdaten aus den einzelnen Applikationen zu extrahieren und für das Reporting zu konsolidieren (Consolidation Hub). Der initiale Aufwand wird bewusst gering gehalten, indem aufwendige Anpassungen in den dezentralen Applikationen vermieden werden und lediglich eine harmonisierte Kopie zu Analysezwecken (nur lesender Zugriff) erstellt wird. Stufenweise sollen dann schreibende Zugriffe auf das MDM erlaubt und eine zunehmende Anzahl von Stammdatenattributen im MDM gepflegt werden (Coexistence Hub) [Braun 2009, 18-20]. Das langfristige Ziel ist es, das MDM als zentrale Applikation für die Anlage und Pflege sämtlicher Lieferantendaten zu etablieren (zentrales System).

Die derzeitige MDM-Lösung bei der ZF Friedrichshafen AG, einem deutschen Automobilzulieferer, trägt der dezentralen Organisationsstruktur des Konzerns Rechnung. Entsprechend werden die Kundenstammdaten dezentral in relativ unabhängig voneinander laufenden ERP-Systemen gepflegt. Die Kundenstammdaten werden im zentralen SAP NetWeaver MDM, das als Consolidation Hub dient, zusammengeführt (Key Mapping mittels ZF-eigener führender Konzern-ID) [Jakic 2009, pp. 30-33]. Die zentralen Kundenstammdaten werden führend für die übergreifende Generierung konsistenter Reports in den dezentralen Einheiten und auf Konzernebene sowie für die Steuerung von zentralisierten Geschäftsprozessen genutzt. Daraus ergibt sich der Vorteil, dass die einzelnen Standorte sehr flexibel auf das operative Tagesgeschäft der weltweiten Kundenbearbeitung reagieren können und dass dem wachsenden Bedarf der Zentraleinheiten an tagesaktuellen Auswertungen entsprochen wird. Der Aufwand für den Aufbau einer zentralen Stammdatenorganisation wird hierbei gleichzeitig gering gehalten. Aufgrund der wachsenden, unkontrollierten Redundanz der Kundenstammdaten, die den Mapping-Aufwand kontinuierlich erhöhen, und um das Unternehmensziel der Prozessstandardisierung zu unterstützen, wird für die Zukunft eine zentralere Architekturlösung angestrebt. Diese soll in einer Transformation in den drei Phasen Konsolidierung, Harmonisierung (Coexistence Hub mit zentraler Kundennummer und einheitlich definierten Kundenstammattributen) und zentrales Stammdatenmanagement die Anzahl der lokalen ERP-Systeme reduzieren und eine schrittweise Harmonisierung der Kundenstammdaten ermöglichen [Jakic 2009, pp. 17-19]. Die Möglichkeit, die dritte Phase zu erreichen, wird für die ZF aufgrund des Zukaufs neuer Unternehmensteile mit eigenen ERP-Systemen jedoch kritisch bewertet.

Anhand der Merkmale Redundanz, Datenverteilung und zeitliche Verzögerung der Datenbereitstellung (Datenverfügbarkeit bzw. -aktualisierung), die im Wesentlichen von der Art des Stammdatenaustausches zwischen den Applikationen abhängen, können die Stammdatenarchitekturmuster verschieden implementiert werden [Heine 1999, 173f.]. Sowohl bei einer zentralen als auch bei einer führenden Architektur können die Fachapplikationen entweder direkt und somit in Echtzeit auf das zentrale bzw. führende System zugreifen oder die zu verteilenden Stammdaten werden in regelmässigen Abständen auf die Fachapplikationen übertragen. Während die erste Variante eine redundanzfreie Haltung stets aktueller Stammdaten ermöglicht, werden bei der zweiten Variante in den Fachapplikationen (kontrolliert) redundante Kopien gehalten, die in regelmässigen Abständen (Batch) aktualisiert werden. Eine zentrale Stammdatenarchitektur mit lesendem und schreibendem Zugriff der angeschlossenen Applikationen in Echtzeit (und somit ohne lokale Kopien) entspricht dabei dem in der Literatur häufig

beschriebenen *Transaction Hub*⁴¹ [Radcliffe et al. 2006, 6-7; Berson/Dubov 2007, 78; Dreibelbis et al. 2008, 249-253].

Die erwähnten Stammdatenarchitekturmuster sind in Tabelle 5-31 mit ihren charakterisierenden Merkmalen zusammengefasst.

	Zentrales System	Führendes System	Coexistence Hub	Consolidation Hub	Registry	Peer-To-Peer
Datenanlage/ Datenpflege	Zentral	Zentral	Lokal	Lokal	Lokal	Lokal
Speicherung der Datenobjektattribute	Vollständig im zentralen System verwaltet; lokale Ergänzungen möglich	Sämtliche Attribute einer Stammdatenklasse im führenden System; lokale Ergänzungen	Bereinigter Referenzdatensatz („Golden Record“) im Hub	Sämtliche Attribute lokal gehalten	Identifizierende Attribute zentral im Verzeichnis gehalten	Sämtliche Attribute lokal gehalten
Separates MDM-System	Ja	Nein (Fachapplikation)	Ja (Hub)	Ja (Hub)	Ja (Verzeichnis)	Nein
Grad und Art der Harmonisierung	Hoch; durch Datenmodell des zentralen Systems	Hoch; durch Datenmodell des führenden Systems	Relativ hoch; durch Datenmodell des Hubs	Gering; durch Datenmodell des Hubs (keine Harmonisierung in Fachapplikationen)	Gering; lediglich Harmonisierung der identifizierenden Attribute	Keine Harmonisierung der Stammdaten
Redundanz	Kontrolliert, im Fall von lokalen Kopien	Kontrolliert, im Fall von lokalen Kopien	Redundante Datenhaltung in lokalen Applikationen durch Hub kontrolliert	Statische Kopie für Konsolidierung	Ja (teilweise kontrolliert über identifizierende Attribute)	Ja
Mapping	Nicht notwendig	Nicht notwendig	Notwendig (durch das Hub)	Notwendig (durch das Hub)	Bilateral zwischen Applikationen, die Daten austauschen	Bilateral zwischen Applikationen, die Daten austauschen
Verfügbarkeit der Daten	Echtzeitaktualisierung der Stammdaten möglich (aber aufwendig)	Echtzeitaktualisierung der Stammdaten möglich (aber aufwendig)	Mit zeitlicher Verzögerung (Batch)	Keine Aktualität der Daten im Hub (nur statische Kopie)	Aktualität der Stammdaten, da föderierte Sicht jeweils zum Zeitpunkt der Anfrage erstellt wird	Aktualität der Stammdaten nur in lokalen Applikationen
Datenverteilung	Push und Pull möglich	Push und Pull möglich	erst Pull, dann Push (jeweils das durch Hub)	Pull (durch das Hub)	Pull (durch anfragende Applikation)	Pull (durch anfragende Applikation)

Tabelle 5-31: Charakterisierung von Stammdatenarchitekturmustern

5.7.2 Vorgehensmodell und Aktivitäten

Ziel des Methodenfragmentes ist die Etablierung einer verbindlichen Stammdatenintegrationsarchitektur, welche die Grundlage für die Integration der Instanzen der Stammdatenobjekte bildet und die konsistente Verteilung an sämtliche stammdaten-nutzenden Applikationen des Unternehmens gewährleistet. Für die Ausführung der

⁴¹ Ein Beispiel hierfür ist der in den Fallstudien beschriebene Global Data Client von Bosch Rexroth.

Aktivitäten dieses Methodenfragmentes ist auf die Ergebnisse der vorherigen Methodenfragmente zurückzugreifen, insbesondere auf die Informationen aus dem BDD (wie beschrieben in [Thangarathinam et al. 2004, 340-341]), auf die Inhalte der Stammdatenlandkarte sowie des semantischen Stammdatenmodells. Die zentral verwalteten Metadaten sowie konsolidierte, semantische Datenmodelle sind Ausgangspunkt für Aktivitäten der Neugestaltung und Optimierung von Datenintegrationsarchitekturen [Marco 2005, 10; Loshin 2008, 119].

Analog zu den ersten beiden Methodenfragmenten werden die Aktivitäten des dritten Methodenfragmentes aus den Aktivitäten der Fallstudien abgeleitet (siehe Tabelle 5-32).

Nr.	Aktivität	Ergebnisse	Aktivität aus Fallstudie		
			SBB Cargo	DTAG	Bosch Rexroth
I Ist-Architektur					
I.1	Stammdatenflüsse und Schnittstellen modellieren	Stammdatenflussdiagramm	A.IV.1, A.IV.2	B.IV.2	–
I.2	Priorität der Applikationen festlegen	Liste priorisierter Applikationen	–	B.IV.3	–
II Ziel-Architektur					
II.1	Architekturmuster evaluieren	Ausgewähltes Architekturmuster	–	–	C.I.2
II.2	Stammdatenintegrationsarchitektur definieren	Integrationskontextdiagramm, Implementierungsroadmap	A.IV.4	B.IV.4	C.I.2
III Umsetzung					
III.1	Stammdatenintegrationsarchitektur implementieren	Stammdatenintegrationsarchitektur	A.V.1	–	C.III.1, C.III.2, C.III.3
III.2	Datenpflegeprozesse spezifizieren	Unternehmensweite Stammdatenmanagementprozesse	A.III.1	B.III.1	C.IV.1, C.IV.2

Tabelle 5-32: Aktivitäten des induzierten Vorgehensmodells von Methodenfragment 3

Die zeitlogische Abfolge der konsolidierten Aktivitäten des dritten Methodenfragmentes ist in Abbildung 5-16 dargestellt.

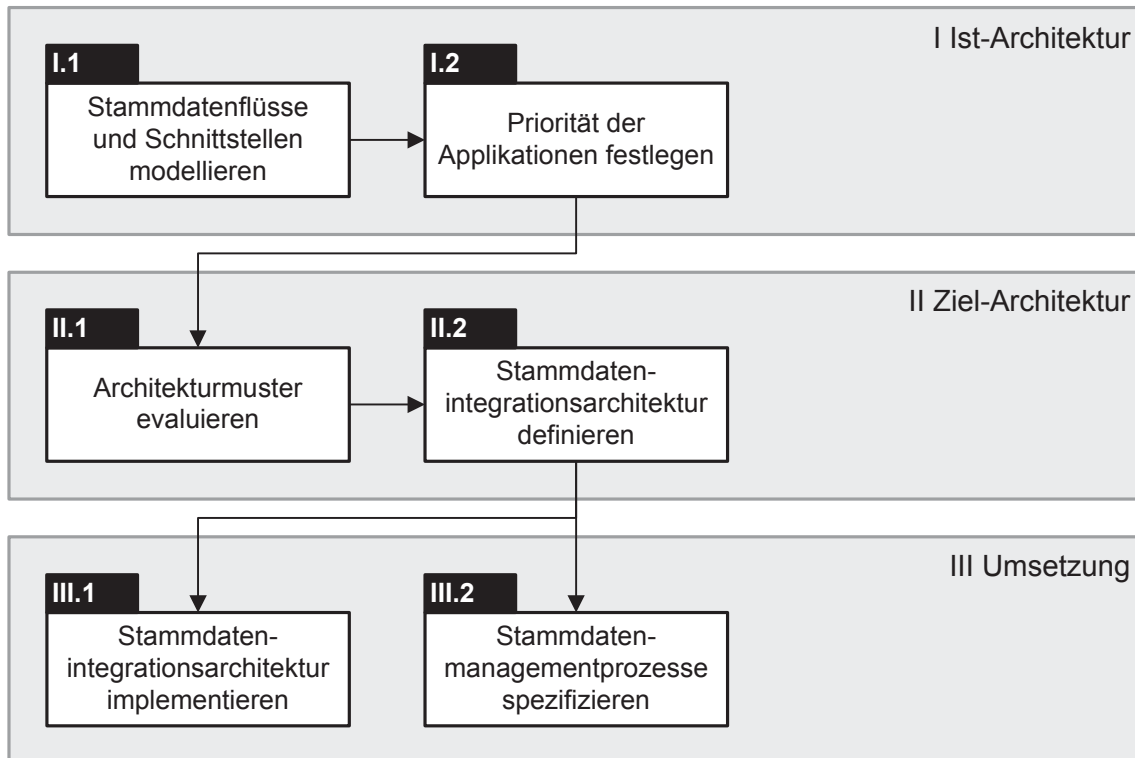


Abbildung 5-16: Vorgehensmodell des Methodenfragmentes 3

Auch die einzelnen Aktivitäten des dritten Methodenfragmentes werden im Folgenden kurz textuell beschrieben und jeweils durch eine Tabelle, die verantwortliche Rollen, Ergebnisse, notwendige Vorbedingungen und Techniken nennt, charakterisiert. Die Techniken erläutert Kapitel 5.7.3.

5.7.2.1 Stammdatenflüsse und Schnittstellen modellieren

Für das Ziel einer Applikations- und Integrationsarchitektur mit einer integrierten Stammdatenbasis ist eine ausreichende Kenntnis der vorhandenen Applikationen und ihrer Beziehungen notwendig [Höning 2009, 193]. Die Datenflussmodellierung ermöglicht eine kombinierte Betrachtung von Applikationsfunktionen und Daten, in der Regel abstrahiert von den physikalischen Gegebenheiten der Applikation. Sie ist wesentlicher Bestandteil der Analyse und Gestaltung von Integrationsarchitekturen. Die Modellierung ist Aufgabe des Datenarchitekten, als Verantwortlicher für die Informations- und Datenmodelle des Unternehmens, sowie des Konzern-Datenstewards, der die Sicht des qualitätsorientierten Stammdatenmanagements vertritt.

MF3: I.1 Stammdatenflüsse und Schnittstellen modellieren	
Ergebnis	Stammdatenflussdiagramm, Schnittstellenbeschreibungen
Rollen	Strategischer Datensteward [verantwortlich] Datenarchitekt, Konzern-Datensteward [ausführend] Technische Datenstewards, System-Owner, IS-Architekt ⁴² [unterstützend]
Techniken	T.I.1 Datenflussmodellierung T.I.2 Entwurf einer Integrationsarchitektur
Input/Vorbedingungen	Applikationsverzeichnis, Applikationsbeschreibung, Stammdatenlandkarte, Semantisches Stammdatenmodell

Tabelle 5-33: MF3: I.1 Stammdatenflüsse und Schnittstellen modellieren

Für die Modellierung der Stammdatenflüsse wird im Wesentlichen auf das Applikationsverzeichnis und die Applikationsbeschreibungen (siehe Aktivität MF2: II.1) sowie auf die Stammdatenlandkarte zurückgegriffen, welche Informationen über die statische Verteilung von Datenobjekttypen auf Applikationen enthalten (siehe Kapitel 5.6.2.4). Ihnen kann entnommen werden, welche Applikationen welche Stammdatenobjekte erzeugen, verteilen oder erhalten. Das Stammdatenflussdiagramm ergänzt diese Sicht um eine graphische Repräsentation der dynamischen Aspekte (Datenaustausch), wobei es sich auf die Datenobjekttypen, die in der Stammdatenlandkarte und im BDD enthalten sind, beschränkt. Das Stammdatenflussdiagramm weist Ähnlichkeiten zum Ergebnisdokument Applikationsübersicht des BE-Methodenkerns auf, indem es die Beziehungen zwischen Applikationen abbildet. Der Schwerpunkt der Modellierung liegt hierbei auf den ausgetauschten Datenobjektattributen, wodurch der Detaillierungsgrad höher ist. Die zwischen den Applikationen ausgetauschten Datenobjektattribute sind zusätzlich in der Schnittstellenbeschreibung zu dokumentieren (siehe Beschreibung der Technik T.I.2).

5.7.2.2 Priorität der Applikationen festlegen

Aufbauend auf Informationen des BDD sowie der Stammdatenlandkarte sind die bestehenden Applikationen, welche im BDD definierte Stammdatenobjekte erzeugen bzw. nutzen, durch den IS-Architekten in Zusammenarbeit mit dem strategischen Datensteward zu priorisieren. Die Priorisierung soll die Rolle für jede Applikation innerhalb der zukünftigen Integrationsarchitektur definieren. Hierbei sind vor allem diejenigen Applikationen zu identifizieren, die in einer zentralen Architektur für eine oder mehrere Stammdatenklassen potenziell führend sind oder die in der Ziel-Architektur abgelöst werden können (z. B. wenn die Neugestaltung mit dem Ziel verbunden ist, Altapplikationen abzulösen bzw. die Komplexität der Applikationsarchitektur zu reduzieren, wie im Fall der Daimler AG). Für Letzteres bietet die Stammdatenlandkarte Hinweise, da sie Applikationen aufzeigt, die auf keine Datenobjekt-

⁴² Der IS-Architekt hat einen Überblick über die gesamte Informations- und Applikationsarchitektur des Unternehmens und ist massgeblich für deren langfristige Planung und Entwicklung verantwortlich [Schwinn 2005, 235].

attribute zugreifen oder die Datenobjektattribute duplizieren, indem sie diese neu anlegen, anstatt bestehende zu nutzen. Mit Hilfe der Priorisierung der Applikationen ist es bei der Implementierung einer zentralen oder einer Hub-Architektur möglich, die Reihenfolge zu bestimmen, in der die Applikationen an das Hub oder den MDM-Server angebunden werden sollten. Das Ergebnis der Priorisierung sollte als zusätzliche Information in der Stammdatenlandkarte dokumentiert werden (siehe Anhang D.5).

MF3: I.2 Priorität der Applikationen festlegen	
Ergebnis	Liste priorisierter Applikationen (bzw. Priorisierung in Stammdatenlandkarte eingefügt)
Rollen	IS-Architekt, Strategischer Datensteward [ausführend, verantwortlich] Fachliche Datenstewards, Technische Datenstewards, System-Owner [unterstützend]
Techniken	T.I.3 Bewertung der Applikationen
Input/Vorbedingungen	BDD, Stammdatenlandkarte, Applikationsverzeichnis, Stammdatenflussdiagramm

Tabelle 5-34: MF3: I.2 Priorität der Applikationen festlegen

5.7.2.3 Architekturmuster evaluieren

Als Grundlage für die Entwicklung der zukünftigen Stammdatenintegrationsarchitektur kann die Erfahrung aus Wissenschaft und Praxis genutzt werden, die in Form von Architekturmustern dokumentiert ist (siehe Kapitel 5.7.1). Diese zeigen auf einer konzeptionellen, implementierungsunabhängigen Ebene, welche Varianten des Managements von Stammdaten in verteilten Applikationsarchitekturen existieren, indem sie für die einzelnen Muster unter anderem Aussagen zum Ort der Stammdatenhaltung und -pflege, zum Umfang der harmonisierten Datenobjektattribute oder zur Art der Stammdatenverteilung treffen. Den Architekturmustern gemein ist das Ziel, einen transparenten Zugriff auf eindeutig identifizierbare Stammdatenobjekte zu gewährleisten [Loshin 2008, 165].

MF3: II.1 Architekturmuster evaluieren	
Ergebnis	Ausgewähltes Architekturmuster
Rollen	IS-Architekt, Strategischer Datensteward [verantwortlich] Konzern-Datensteward [ausführend]
Techniken	T.II.1 Bewertung der Architekturmuster
Input/Vorbedingungen	Architekturprinzipien, Stammdatenlandkarte, Stammdatenflussdiagramm, Schnittstellenbeschreibung

Tabelle 5-35: MF3: II.1 Architekturmuster evaluieren

Im Rahmen der Beschreibung der Technik zur Bewertung (siehe Kapitel 5.7.3) werden Vor- und Nachteile der einzelnen Architekturmuster diskutiert sowie konkrete Bewertungskriterien vorgestellt. Hierbei wird explizit auf die Eignung der Architekturmuster für die verschiedenen Anwendungsszenarien (siehe Kapitel 4.4) eingegangen. In einem analytischen Anwendungsszenario der Stammdatenintegration ist diese Aktivität optional, da das Architekturmuster des Consolidation Hub per se auf die Anforderungen dieses Szenarios zugeschnitten ist.

5.7.2.4 Stammdatenintegrationsarchitektur definieren

Massgeblich für den Entwurf der Stammdatenintegrationsarchitektur ist das in der vorherigen Aktivität gewählte Architekturmuster. Dieses ist nun mit konkreten Applikationen und Schnittstellen zum Datenaustausch zwischen den Applikationen auszuprägen. Insbesondere ist die zukünftig als Hub oder führendes System zu nutzende Applikation als wichtigste Komponente der neuen Architektur zu spezifizieren. Grundlage für die Datenarchitektur des MDM-Systems sowie für die Spezifikation der Schnittstellen ist das aus dem BDD abgeleitete semantische Stammdatenmodell (siehe Kapitel 5.6). Die Technik für den Entwurf der Integrationsarchitektur gewährleistet die Ausrichtung der Applikationen auf die im BDD definierten Stammdatenanforderungen der Geschäftsprozesse bzw. Fachbereiche (vgl. [Höning 2009, 192]). Die konzeptionelle Stammdatenintegrationsarchitektur kann mit Hilfe eines Integrationskontextdiagramms (vgl. [Schwinn 2005, 205]) graphisch dargestellt werden. Im Gegensatz zum Stammdatenflussdiagramm aus MF3: I.1 bildet das Integrationskontextdiagramm die Ziel-Integrationsarchitektur mit den Applikationen und ihren zukünftigen Integrationsbeziehungen (Datenflüsse mit Datenobjekttypen) ab.

Die Ziel-Architektur kann als langfristige Lösung definiert werden, deren Umsetzung aufgrund organisatorischer bzw. technologischer Einschränkungen (z. B. Einschränkung des operativen Betriebes) nicht direkt, sondern über Zwischenlösungen möglich ist. Der Fall von Novartis (siehe Kapitel 5.7.1) beispielsweise zeigt, dass Zwischenschritte auf dem Weg zu einer zentralen Stammdatenarchitektur (mit zunehmender Funktionalität, aber auch erhöhtem Implementierungsaufwand) notwendig sein können. Die Implementierung eines Consolidation Hub ist häufig ein erster Schritt für eine weitergehende Integration der Stammdaten in Form eines Coexistence Hub oder einer zentralen Stammdatenarchitektur [Dreibelbis et al. 2008, 26]. Ein solcher Entwicklungspfad ist in der Implementierungsroadmap zu dokumentieren.

MF3: II.2 Stammdatenintegrationsarchitektur definieren	
Ergebnis	Integrationskontextdiagramm (mit aktualisierten Applikations- und Schnittstellenbeschreibungen) Implementierungsroadmap
Rollen	IS-Architekt, Strategischer Datensteward [verantwortlich] Datenarchitekt, Konzern-Datensteward [ausführend] Fachliche Datenstewards, Technische Datenstewards, System-Owner, Datenbankadministratoren [unterstützend]
Techniken	T.I.2 Entwurf einer Integrationsarchitektur
Input/Vorbedingungen	Ist-Analyse der Geschäftsprozesse, Architekturprinzipien, Stammdatenlandkarte, Semantisches Stammdatenmodell, Stammdatenflussdiagramm, Schnittstellenbeschreibung, ausgewähltes Architekturmuster

Tabelle 5-36: MF3: II.2 Stammdatenintegrationsarchitektur definieren

5.7.2.5 Stammdatenintegrationsarchitektur implementieren

Die Implementierung der Ziel-Architektur bedingt mehrere Teilprojekte, die aufgrund ihrer Komplexität und benötigten Ressourcen inkrementell und iterativ durchzuführen sind [Riebisch 2006, 65; DAMA 2009, 185]. Sowohl für die zentralen Architekturmuster als auch für die Hub-Architekturen sind die zu integrierenden Fachapplikationen schrittweise an das MDM-System anzubinden. Die als Ergebnis der Aktivität MF3: I.2 vorliegende Priorisierung der Applikationen (siehe Kapitel 5.7.2.2) liefert eine Hilfestellung, die Reihenfolge der Anbindung zu bestimmen. Die Anbindung verläuft für jede der Fachapplikationen analog: Identifikation und Bereinigung der Dubletten in der Fachapplikation (sowie zwischen Fachapplikation und MDM-System im Fall, dass das MDM-System bereits Daten enthält), Migration der zentral zu verwaltenden Datenelemente und Festlegung der Verteilmechanismen zwischen den Applikationen.

MF3: III.1 Stammdatenintegrationsarchitektur implementieren	
Ergebnis	Stammdatenintegrationsarchitektur
Rollen	IS-Architekt, Strategischer Datensteward [verantwortlich] Datenarchitekt, System-Owner, Datenbankadministratoren, Konzern-Datensteward [ausführend]
Techniken	T.III.1 Dublettenbereinigung (Instanzintegration) T.III.2 Datenmigration
Input/Vorbedingungen	Integrationskontextdiagramm, Schnittstellenbeschreibungen, ausgewähltes Architekturmuster, Implementierungsroadmap

Tabelle 5-37: MF3: III.1 Stammdatenintegrationsarchitektur implementieren

Während die Dublettenbereinigung in sämtlichen Anwendungsszenarien der Stammdatenintegration wichtig ist, sind die Datenmigration sowie die Festlegung der Verteilmechanismen für die operative Stammdatenintegration und die Konsolidierung der Applikationsarchitektur von Bedeutung. In den letztgenannten Anwendungsszenarien müssen Stammdatenbestände aus Altapplikationen in die neuen (führenden) Stammdatensysteme migriert und notwendige Synchronisationsbedarfe zwischen stammdatenführenden und -nutzenden Applikationen definiert werden [Loshin 2008, 163]. Bei Bosch Rexroth (siehe Fallstudie in Kapitel 4.2.3) besteht das zentrale Ziel der Stammdatenintegration in der Zentralisierung der Stammdatenintegrationsarchitektur (Anwendungsszenario Konsolidierung). Die schrittweise Migration und die Synchronisation der Datenelemente zwischen lokalen ERP-Systemen und zentralem Stammdatenserver stellten den zeit- und ressourcenaufwendigsten Teil der Kunden- und Materialstammdatenintegration dar. Im analytischen Anwendungsszenario müssen lediglich die unidirektionalen Beziehungen zwischen stammdatenführenden Applikationen und dem Consolidation Hub implementiert werden. Mögliche Dubletten werden in der Regel nach Extraktion der Datenkopien aus den operativen Applikationen im Consolidation Hub bereinigt.

5.7.2.6 Stammdatenmanagementprozesse spezifizieren

Die Änderung der Stammdatenintegrationsarchitektur bedingt die Neugestaltung von Anlage- und Pflegeprozessen für Stammdaten. Diese Aktivität ist vor allem für die operative Stammdatenintegration wesentlich, da Änderungen der Stammdaten möglichst in Echtzeit in den Applikationen verfügbar sein müssen. Die Spezifikation der Stammdatenmanagementprozesse umfasst die Definition der einzelnen Aufgaben (u. a. Anlage bzw. Erzeugung, Änderung, Löschung bzw. Bereinigung von Stammdaten) und ihrer Reihenfolge sowie die langfristige Festlegung der Verantwortlichkeiten für diese Aufgaben. Die Aufgaben des Prozessentwurfs können sich an etablierten Techniken des Business Process (Re-)Engineering orientieren. So lässt sich beispielsweise die Entwurfstechnik der Ablaufplanung der Methode PROMET-BPR für den Entwurf betrieblicher Prozesse [Hess 1996, 158; IMG 1997, AKTI 4ff.] in vereinfachter Form auf den Entwurf von Stammdatenmanagementprozessen übertragen. In Analogie zu den beschriebenen Aufgaben bei der Definition des Stammdatenmanagementprozesses fordert die Methode ein systematisches Vorgehen in drei Schritten: 1) Identifizieren von Aufgaben, 2) Festlegen der Aufgabenreihenfolge und 3) Zuordnen von Aufgabenträgern (Verantwortlichen) zu den einzelnen Aufgaben.

Langfristig sind zudem Kennzahlen und Aktivitäten zur Messung und Steuerung der Stammdatenqualität zu spezifizieren und in die Stammdatenmanagementprozesse zu integrieren, die jedoch nicht im Mittelpunkt dieser Arbeit stehen. Hierfür wird auf entsprechende Arbeiten des Kompetenzzentrums (vgl. [Otto et al. 2009]) sowie weitere Publikationen zum Thema Datenqualitätsmessung verwiesen (vgl. [Batini et al. 2007], [Caballero et al. 2008]).

MF3: III.2 Stammdatenmanagementprozesse spezifizieren	
Ergebnis	Unternehmensweite Stammdatenmanagementprozesse
Rollen	Strategischer Datensteward [verantwortlich] Konzern-Datensteward, Fachliche Datenstewards, Technische Datenstewards [ausführend] Prozess-Owner, System-Owner [unterstützend]
Techniken	T.III.3 Ablaufplanung
Input/Vorbedingungen	Integrationskontextdiagramm

Tabelle 5-38: MF3: III.2 Stammdatenmanagementprozesse spezifizieren

5.7.3 Techniken

Die vorherigen Kapitel beschreiben die Aktivitäten des dritten Methodenfragmentes und nennen einzelne Techniken, die als Arbeitsanweisung bei der Ausführung der Aktivität behilflich sind. Hierbei wird teilweise auf Techniken zurückgegriffen, die sich bereits in der Praxis bewährt haben oder in den ersten beiden Methodenfragmenten verwendet wurden. Tabelle 5-39 zeigt einen Überblick über die in Methodenfragment

3 verwendeten Techniken mit möglichen Referenzen auf Methodenfragment 1 bzw. Methodenfragment 2 respektive Publikationen, in denen die jeweiligen Techniken detailliert erläutert werden.

Nr.	Name	Aktivität	Referenz
T.I.1	Datenflussmodellierung	MF3: I.1	[DeMarco 1979], [Balzert 2000, 141-146]
T.I.2	Entwurf einer Integrationsarchitektur	MF3: I.1, MF3: II.3	[Schwinn 2005, 157-185], [Höning 2009, 192-203], [IMG 1996, 178-194]
T.I.3	Bewertung der Applikationen	MF3: I.2	Siehe T.I.1 und T.II.4 von MF1, [IMG 1999, 159-163]
T.II.1	Bewertung der Architekturmuster	MF3: II.1	Siehe T.I.1 und T.II.4 von MF1 sowie Kapitel 5.7.1, [IMG 1999, 159-163]
T.III.1	Dublettenbereinigung (Instanzintegration)	MF3: III.1	[Leser/Naumann 2007, 317ff.], [Hernandez/Stolfo 1998], [Rahm/Do 2000, 7ff.], [Batini/Scannapieco 2006, 97ff.]
T.III.2	Datenmigration	MF3: III.1	[IMG 1999, 146-152], [Endava 2006], [Haller 2009], [Morris 2006]
T.III.3	Ablaufplanung	MF3: III.2	Siehe T.IV.1 von MF1, [IMG 1997, 230-244], [Österle 1995, 61ff.]

Tabelle 5-39: Techniken des Methodenfragmentes 3

Das Kapitel beschreibt im Folgenden die Übertragung und Anwendung der Techniken in der vorliegenden Methode und führt Beispiele an. Auf die bereits in den vorherigen Methodenfragmenten spezifizierten Techniken wird an dieser Stelle nicht noch einmal im Detail eingegangen.

T.I.1 Datenflussmodellierung

Der Fokus der Datenflussmodellierung liegt auf den Datenflüssen zwischen einzelnen Applikationen eines Unternehmens. Für die Erstellung von Datenflussdiagrammen stehen verschiedene Vorgehen und Notationen zur Verfügung. Das am weitesten verbreitete Vorgehen ist die Datenflussmodellierung im Rahmen der Strukturierten Analyse als Methode zur Systementwicklung (vgl. [DeMarco 1979]). Ausgehend von einem Kontextdiagramm, das die Datenflüsse des Systems von und zu seiner Umwelt (Aussenbeziehungen) abbildet, werden über ein Top-Down-Vorgehen zunehmend detailliertere Datenflussmodelle erstellt [Balzert 2000, 432ff.]. Für die Anwendung innerhalb der beschriebenen Methode ist eine Modellierung auf der ersten Detaillierungsebene unterhalb des Kontextdiagrammes ausreichend. Für die Spezifikation der Datenflüsse sollten die im BDD spezifizierten Datenobjektattribute verwendet werden.

Die UML bietet mit ihrer Version 2.0 (Superstructure Specification) die Möglichkeit der Modellierung von Informationsflüssen als Austausch von Informationseinheiten in Form von gerichteten Beziehungen (vgl. [OMG 2007, 605-610]). Als Start- und Endpunkt können zahlreiche Modellelemente der UML verwendet werden, u. a. Anwendungsfälle, Akteure, Komponenten, Pakete und Schnittstellen (vgl. [OMG 2007, 607]).

Im Gegensatz zur Datenflussmodellierung der Strukturierten Analyse erlaubt die UML eine Modellierung der Datenflüsse auf relativ hoher Abstraktionsstufe.

Die verbindliche Vorgabe einer bestimmten Notation erscheint an dieser Stelle wenig sinnvoll, da in den Unternehmen gerade zur Modellierung der Datenflüsse häufig proprietäre Darstellungen verwendet werden (siehe Abbildung 6-1 in Kapitel 6.1.1, die das Stammdatenflussdiagramm aus dem Aktionsforschungsprojekt DB Netz AG darstellt).

T.I.2 Entwurf der Integrationsarchitektur

Techniken zum Entwurf einer Applikations- und Integrationsarchitektur sind bereits Bestandteil verschiedener Methoden des BE (vgl. [Schwinn 2005, 157-185]) sowie des Methodenkerns des BE (vgl. [Höning 2009, 192-203]). Die entsprechende Technik des Methodenkerns, der als Synthese der einzelnen Methoden des BE betrachtet werden kann, definiert ein Vorgehen in den drei Schritten *IS-Ist-Analyse erstellen*, *Integrationsbedarfe identifizieren* und *IS-Soll-Architektur definieren*, in denen die drei Ergebnisdokumente Applikationsübersicht, Applikationsbeschreibung und Schnittstellenbeschreibung entwickelt werden [Höning 2009, 192]. Im Sinne einer konsequenten Trennung zwischen Ist-Analyse und Definition der Soll-Architektur sowie aufgrund des Aufwandes für die Durchführung wurden im vorliegenden Vorgehensmodell zwei separate Aktivitäten definiert (siehe Abbildung 5-16). Die Modellierung der Stammdatenflüsse bzw. die Beschreibung der bestehenden Schnittstellen (Aktivität MF3: I.1) entspricht gemeinsam mit der Erstellung der Stammdatenlandkarte (Aktivitäten MF2: II.1 und MF2: II.2b) der IS-Ist-Analyse mit Fokus auf die gespeicherten bzw. ausgetauschten Stammdaten. Entsprechend entsteht als Ergebnis der Aktivität MF3: I.1 das Stammdatenflussdiagramm sowie eine Beschreibung der Schnittstellen zwischen den Applikationen, die gemeinsam die Datentransferbeziehungen zwischen den einzelnen Applikationen dokumentieren [IMG 1996, 179]. Vorlagen für die Beschreibung von Schnittstellen sind in mehreren Methoden des BE zu finden (vgl. [Schwinn 2005, 166f.; Höning 2009, 197-200]).

Die beiden Schritte Integrationsbedarfe identifizieren und Soll-Architektur definieren werden im Zuge der Aktivität MF3: II.2 (Stammdatenintegrationsarchitektur definieren) ausgeführt. Der Integrationsbedarf ergibt sich aus der Notwendigkeit, Applikationen im Zuge der Umgestaltung der Integrationsarchitektur neu miteinander verbinden zu müssen, um den Bedarf an Stammdaten seitens der Fachbereiche abzudecken [Vogler 2004, 230]. Zur Identifikation des (fachlichen) Integrationsbedarfs ist daher der im BDD vom Fachbereich definierte Bedarf mit der Stammdatenbasis der bestehenden Applikationsarchitektur abzugleichen [DAMA 2009, 180]. So lässt sich feststellen, welche Applikationen (bzw. welche Datenbanken) zwingend in die Ziel-Architektur zu übernehmen und miteinander oder einem MDM-System zu verbinden sind [Höning 2009, 196]. Ein wichtiges Hilfsmittel für die Bestimmung notwendiger

Integrationsbedarfe zwischen Applikationen ist die Stammdatenlandkarte (siehe Aktivität MF2: II.2b). Ihr kann entnommen werden,

- welche Applikationen ein Datenobjekt produzieren (create) und welche Applikationen dasselbe Datenobjekt konsumieren (read) und
- welche Applikationen in ein und derselben Spalte ein „C“ stehen haben und folglich ein Datenobjekt redundant erstellen.

Für beide Fälle ist die Integrationsbeziehung zwischen den jeweiligen Applikationen auf Attributebene näher zu untersuchen. Im ersten Fall sind fehlende, aber notwendige Beziehungen zwischen produzierenden und konsumierenden Applikationen als Integrationsbedarf festzuhalten. Im zweiten Fall ist die für einen Datenobjekttyp massgebliche Applikation zu bestimmen und – falls nötig – die Integrationsbeziehung zu konsumierenden Applikationen neu zu spezifizieren, um unkontrollierte Redundanz zu vermeiden.

Die sich daraus ergebenden neuen Applikationsbeziehungen sind den bestehenden Applikationsbeziehungen gegenüberzustellen, die als Ergebnis von Aktivität MF3: I.1 in Form der Stammdatenflussdiagramme bzw. Schnittstellenbeschreibungen dokumentiert sind. Daraus lassen sich die nicht mehr benötigten, die zu modifizierenden sowie die neu zu realisierenden Schnittstellen ableiten und somit der Aufwand für die Implementierung der neuen Integrationsarchitektur abschätzen.

Das im zweiten Methodenfragment entwickelte semantische Stammdatenmodell vereinfacht und beschleunigt die Spezifikation von Schnittstellen. Durch Einschränkung des Verwendungskontextes (z. B. auf spezifische Applikationen oder Geschäftsprozesse) werden nur diejenigen Datenobjekttypen und die dazugehörigen Datenobjektattribute ausgewählt, die in dem Kontext benötigt werden. Somit lassen sich aus dem generischen Stammdatenmodell applikations- bzw. schnittstellenspezifische Datenmodelle ableiten. Die Verwendung des in Kapitel 5.6.4 vorgestellten Modellierungsprototyps ermöglicht es, diesen Schritt halbautomatisiert durchzuführen. Der Anwender wählt hierzu den passenden Verwendungskontext aus. Die durch das Tool generierte Menge von Datenobjektattributen kann anschliessend manuell erweitert oder verändert werden. Manuell neu definierte Datenobjektattribute werden durch den Modellierungsprototyp auf Konsistenz mit dem semantischen Stammdatenmodell überprüft. Das Modellierungstool prüft, ob in dem Datenmodell bereits gleichbedeutende Datenobjektattribute existieren, um Duplikate zu vermeiden. Ist das neu definierte Datenobjektattribut nicht vorhanden und konform mit den Regeln der CCTS spezifiziert, wird es in das Datenmodell eingefügt. Dadurch sind sie für zukünftige Veränderungen gespeichert und können wiederverwendet werden. Die Eigenschaften und Anforderungen der neu zu implementierenden Applikationsbeziehungen als Teil der Soll-Architektur sind mittels Beschreibungen der Soll-Schnittstellen zu dokumentieren.

Aufbauend auf den analysierten Integrationsbedarfen und den daraus resultierenden neuen Applikationsbeziehungen ist die zukünftige Stammdatenintegrationsarchitektur zu definieren [Höning 2009, 201]. Unterschiede zwischen Soll und Ist sollten als Aktualisierungen in den Stammdatenflussdiagrammen, der Stammdatenlandkarte (durch Veränderung der systemtechnischen Verantwortung für Stammdatenobjekte) und den Applikationsbeschreibungen dokumentiert werden.

T.I.3 Bewertung der Applikationen

Die Technik folgt demselben Vorgehen wie die bisherigen Priorisierungen (siehe u. a. die Techniken T.I.1 bzw. T.II.5 des ersten Methodenfragments), beschränkt sich jedoch auf die Definition der Bewertungskriterien und die eigentliche Bewertung. In die Bewertung sind sämtliche Applikationen einzubeziehen, die Stammdatenobjekte zu den im BDD definierten Geschäftsobjekttypen erzeugen bzw. nutzen.

1. Schritt: Bewertungskriterien definieren

Ziel der Priorisierung ist die Bestimmung einer Reihenfolge der Applikationen gemäss ihrer Bedeutung für die Stammdatenintegration im Unternehmen. Mögliche Kriterien sind:

- Ist die Applikation für ein Stammdatenobjekt führend („System of Record“) bzw. werden in der Applikation Stammdatenobjekte initial erzeugt?
- Erzeugt die Applikation Duplikate, also Datenobjekte bzw. Datenelemente, die im Unternehmen bereits in anderen Applikationen existieren und dasselbe Geschäftsobjekt repräsentieren?
- Wie viele Stammdatenobjekte enthält die Applikation?
- Wie viele Datenelemente werden pro Stammdatenobjekt in der Applikation gepflegt?
- Sind die Stammdatenobjekte eindeutig identifizierbar?
- Integriert die Applikation Datenelemente aus mehreren Applikationen?

2. Schritt: Bewertung durchführen

Zur Bewertung anhand der Kriterien sollten bereits im Zuge der vorherigen Methodenfragmente erarbeitete Ergebnisdokumente genutzt werden. So gibt das BDD eine Orientierung dafür, welche Stammdaten für das Unternehmen von Bedeutung sind, während der Stammdatenlandkarte entnommen werden kann, ob Applikationen für die initiale Anlage von Stammdatenobjekten verantwortlich sind. Für die einzelnen Fragen können Antwortkategorien vorgegeben werden, die jeweils mit einer Punktzahl versehen werden. Anhand der Summe ist eine Reihenfolge der Applikationen festzulegen

bzw. eine Kategorisierung (hohe Bedeutung, mittlere Bedeutung, niedrige Bedeutung) vorzunehmen.

T.II.1 Bewertung der Architekturmuster

Die Bewertung und Priorisierung von Architekturmustern lehnt sich in Bezug auf das Vorgehen an die bisherigen Bewertungstechniken der Methode und die dort definierten drei Schritte an (vgl. [IMG 1999, 159-163]). Sie werden an dieser Stelle anhand konkreter Kriterien und einer unternehmensunabhängigen Evaluation erläutert.

1. Schritt: Bewertungskriterien definieren

Die einzelnen, in Kapitel 5.7.1 beschriebenen Architekturmuster bieten unterschiedliche Vor- und Nachteile, die abhängig von der konkreten Situation im Unternehmen abzuwägen sind. Bewertungskriterien erleichtern Praktikern eine objektive Beurteilung der Angemessenheit eines Architekturmusters im Unternehmenskontext. Die folgenden Bewertungskriterien wurden basierend auf einer Literaturanalyse identifiziert und in Gruppeninterviews⁴³ mit Fachexperten validiert respektive ergänzt:

- *Eignung für operatives Anwendungsszenario.* Nicht sämtliche Architekturmuster sind für ein operatives Anwendungsszenario (siehe Kapitel 4.4) geeignet. So generiert das Consolidation Hub lediglich eine konsistente Sicht auf operative Stammdaten durch das Erstellen von Kopien mit nur lesendem Zugriff, ohne die Daten zu harmonisieren oder deren Aktualität in den Fachapplikationen zu sichern. Es ist daher nur für ein analytisches Anwendungsszenario geeignet.
- *Reifegrad des Stammdatenmanagements.* Die Hub-Architekturmuster sind häufig Kompromisslösungen für den Fall, dass sich keine führenden Systeme definieren lassen bzw. zentrale Umsetzungsvarianten aufgrund organisatorischer oder technologischer Hindernisse nicht möglich sind [Schemm 2008, 158]. Zudem werden sie als Zwischenschritte auf dem Weg zu einer zentralen Stammdatenarchitektur (mit zunehmender Funktionalität, aber auch erhöhtem Implementierungsaufwand) genutzt (siehe Kapitel 5.7.2.4). Insofern können verschiedene Architekturmuster für Unternehmen mit unterschiedlichem Reifegrad des Stammdatenmanagements angemessen sein.
- *Implementierungs- und Wartungsaufwand.* Erhebliche Unterschiede bestehen zwischen den einzelnen Architekturmustern in Bezug auf Kosten und Zeit für Implementierung und Wartung der jeweiligen Architekturlösungen [Schemm 2008, 159].

⁴³ Die Gruppeninterviews fanden im Rahmen eines Workshops des Kompetenzzentrums am 03. und 04. September 2008 in Biel (Schweiz) statt.

- *Konsistenz / Kontrollierte Redundanz.* Wichtigstes Ziel der Stammdatenintegrationsarchitektur ist die Gewährleistung konsistenter, nicht redundanter Stammdaten [Schemm 2008, 159]. Redundanz bzw. Inkonsistenz liegt vor, wenn mehrere Datenelemente, die ein Merkmal eines Geschäftsobjektes repräsentieren, sich nicht entsprechen [Wand/Wang 1996, 93f.; Jung 2006, 217].
- *Eignung für Verbesserung der Stammdatenqualität.* Das letzte Kriterium beschreibt die Eignung der Architekturmuster, die Qualität der Stammdaten unternehmensweit zu verbessern [Dreibelbis et al. 2008, 29].
- *Aktualität der Stammdaten.* Das Kriterium der Aktualität beschreibt die Fähigkeit der implementierten Architekturlösung, Nutzern unternehmensweit aktuelle Stammdaten zur Verfügung zu stellen [Loser et al. 2004, 3ff.; Dreibelbis et al. 2008, 27-32]. Dieses Kriterium hängt entscheidend von der Möglichkeit der Echtzeitübertragung von Aktualisierungen ab.
- *Skalierbarkeit / Performanz.* Als Skalierbarkeit der Architekturlösung wird die Fähigkeit bezeichnet, ein erhöhtes Stammdatenvolumen ohne Performanzverluste handhaben zu können [Ariyachandra/Watson 2005, 19], wobei die Performanz der Zeit entspricht, die für die Bereitstellung von Datenelementen durch die Applikation notwendig ist [Jung 2006, 216]. Skalierbarkeit bezieht sich auf die Möglichkeit, weitere Applikationen und Nutzer mit geringem Aufwand an die Architektur anzubinden [Medjahed et al. 2003, 62].

Die Liste der Bewertungskriterien erhebt keinen Anspruch auf Vollständigkeit, sondern spiegelt eine Synthese der Aussagen von Fachexperten wider. Sie kann durch unternehmensspezifische Kriterien ergänzt werden.

2. Schritt: Bewertung durchführen

Die Architekturmuster können anhand der identifizierten Kriterien bewertet werden. Tabelle 5-40 gibt eine Bewertung unabhängig vom spezifischen Unternehmenskontext wieder. Das Kriterium Reifegrad des Stammdatenmanagements wurde an dieser Stelle nicht in die Übersicht aufgenommen, da die Bewertung abhängig vom Reifegrad eines Unternehmens unterschiedlich ausfällt. Erklärungen zu einzelnen Bewertungen schliessen sich an die Tabelle an. Durch die unterschiedliche Gewichtung der Kriterien je nach der konkreten Unternehmenssituation ist es möglich, konzeptionelle Architekturmuster für das eigene Unternehmen zu priorisieren.

	Zentrales System	Führendes System	Coexistence Hub	Consolidation Hub	Registry	Peer-To-Peer
Eignung für operatives Szenario	●	●	●	○	◐	◐
Geringer Implementierungs-/Wartungsaufwand	○	○	◐	●	◐	◐
Konsistenz / Kontrollierte Redundanz	●	●	◐	○	◐	○
Eignung für Verbesserung der Datenqualität	●	●	◐	◐	◐	○
Aktualität der Stammdaten	●	●	◐	○	◐	◐
Skalierbarkeit/ Performanz	◐	◐	◐	●	●	●

● Kriterium vollständig erfüllt ◐ Kriterium teilweise erfüllt ○ Kriterium nicht erfüllt

Tabelle 5-40: Bewertung von Stammdatenarchitekturmustern

Wie bereits bei der Beschreibung der Bewertungskriterien erläutert, sind die Architekturmuster nicht gleichermassen für das operative Stammdatenmanagement geeignet. So generiert das Consolidation Hub lediglich eine konsistente Sicht auf operative Stammdaten durch das Erstellen von Kopien mit nur lesendem Zugriff, ohne die Daten zu harmonisieren oder deren Aktualität in den Fachapplikationen zu sichern. Auch Registry- und Peer-To-Peer-Architekturen erscheinen für dieses Anwendungsszenario nur bedingt zweckmässig, da sie keine umfassende Harmonisierung von Stammdatenattributen gewährleisten.

Grosse Unterschiede bestehen in Bezug auf den Implementierungs- und Wartungsaufwand. Ein Registry mit einer geringen Anzahl zentral zu speichernder und zu pflegender Attribute ist einfacher und schneller zu realisieren als andere Lösungen, bei denen viele Daten zentral gehalten und entsprechend harmonisiert werden müssen. Zusätzlich sind bei den Architekturmustern mit lokaler Stammdatenanlage und -pflege geringere Eingriffe in die bestehende Applikations- und Prozessarchitektur (inklusive Verantwortlichkeiten) erforderlich [Radcliffe et al. 2006, 6; Berson/Dubov 2007, 78], da keine zentrale Stammdatenorganisation aufgebaut und die lokalen Prozesse nicht verändert werden müssen. Das Kriterium wird zudem massgeblich durch den Grad der Harmonisierung (siehe Tabelle 5-31) beeinflusst. Bei zentralen Lösungen sowie Hub-Lösungen mit einer hohen Anzahl im Hub gehaltener Attribute ergeben sich beträchtliche Aufwände für die (initiale) Harmonisierung und Migration (inklusive Extraktion, Bereinigung und Transformation) der Stammdaten sowie für den Entwurf und die Implementierung neuer Pflegeprozesse [Dreibelbis et al. 2008, 237ff.]. Zudem sind in diesen Fällen bei wachsenden Datenvolumina und beim Zugriff vieler Applikationen, die das führende bzw. zentrale MDM-System serialisieren muss, Performanzverluste in Kauf zu nehmen [Loshin 2008, 175f.]. Bei der Peer-To-Peer-Architektur ist der

Implementierungsaufwand zwar sehr gering, jedoch ist ein erhöhter Wartungsaufwand für die zahlreichen Schnittstellen zwischen den Applikationen notwendig (insbesondere bei einer hohen Anzahl von Applikationen im Unternehmen).

Die Speicherung der Stammdaten in einheitlicher Form in einer Applikation (auf Grundlage des harmonisierten Datenmodells) und die zentrale Steuerung der Stammdatenverteilung gewährleistet die applikationsübergreifende Konsistenz bei zentralen Architekturvarianten. Dies hat direkten Einfluss auf die Möglichkeit, Datenredundanz zu kontrollieren und die Stammdatenqualität im Unternehmen zu verbessern, die mit steigender Anzahl zentral verwalteter Attribute wächst. Je höher die Autonomie der einzelnen Fachapplikationen im Unternehmen, desto schwieriger ist die Kontrolle der Redundanz von Stammdaten im Unternehmen. Hub-Architekturen, die keinen oder nur einen geringen Umfang an Stammdatenattributen zentral verwalten (z. B. ein Registry), erlauben nur bedingt die Harmonisierung und Bereinigung redundanter Stammdatenattribute.

Das Kriterium der Aktualität hängt entscheidend von der Möglichkeit der Echtzeitübertragung von Aktualisierungen ab. Hierbei weisen Hub-Architekturen (Coexistence, Consolidation) Defizite auf, da die Stammdaten nach der Extraktion und Harmonisierung, wenn überhaupt, nur mit zeitlicher Verzögerung (per Batch-Job) wieder in die angebundene Fachapplikationen verteilt werden können.

T.III.1 Dublettenbereinigung (Instanzintegration)

Nach der Überwindung von Heterogenitäten auf Schemaebene (siehe Methodenfragment 2) müssen im Zuge der Gestaltung der Integrationsarchitektur auch die Instanzen integriert und Duplikate beseitigt werden (engl.: data cleansing). Auf Instanzebene können grundsätzlich Fehler innerhalb eines Datenelementes (Tupel) von tupelübergreifenden Datenfehlern unterschieden werden [Leser/Naumann 2007, 326]. Die erste Kategorie umfasst Probleme aufgrund [Rahm/Do 2000, 3f.; Leser/Naumann 2007, 318-322]:

- Falscher Werte (Schreibfehler, falsch verwendete Felder),
- Fehlender Werte (Nullwerte),
- Kryptischer oder eingebetteter Werte,
- Unterschiedlicher Wertrepräsentationen (aufgrund abweichender Wertebereiche).

Die Beseitigung dieser Fehler ist durch Normalisierungsregeln zur Transformation in ein Standardformat (automatische Rechtschreibprüfung, Parsing bzw. Attribut-Splitting), durch Konvertierungsregeln (zur Umwandlung von Datenwerten von einer Einheit in eine andere) oder die Verwendung von Referenztabellen (z. B. von Orts-, Strassen- oder Namensverzeichnissen) möglich. Zusätzlich können die Metadaten ei-

nes BDD als Referenz genutzt werden, um falsche und unklare Werte zu korrigieren [Rahm/Do 2000, 7]. Entsprechende Techniken und Funktionen sind in kommerziellen Produkten enthalten und unterstützen die automatisierte Bereinigung derartiger Datenfehler.

Die grösste Herausforderung besteht jedoch in der Beseitigung von Duplikaten [Wang/Madnick 1989, 46] und somit in der Bereinigung tupelübergreifender Datenfehler. Für eine Bereinigung ist schrittweise die Identifikation von Duplikaten und anschliessend deren Beseitigung (auch Datenfusion) notwendig. Idealerweise wird das Vorgehen auf bereits vorbereinigten Daten durchgeführt, bei denen die beschriebenen Datenfehler innerhalb einzelner Datenelemente behoben wurden (Phase der Vorverarbeitung [Batini/Scannapieco 2006, 103f.]). Die dadurch erreichte Standardisierung der Datensätze erhöht deren Vergleichbarkeit und verbessert die Ergebnisse der Duplikaterkennung [Low et al. 2001, 588; Müller/Freytag 2003, 12].

Die Identifikation von Duplikaten beruht auf dem paarweisen Vergleich sämtlicher Datenelemente. Im Idealfall gibt es identifizierende Attribute, welche die eindeutige Duplikaterkennung (engl.: record linkage, object/instance identification oder entity resolution) ermöglichen und über die daraufhin die Datensätze zusammengeführt werden können. Das BDD bietet durch die als identifizierend festgelegten Geschäftsobjektattribute Anhaltspunkte zur Bestimmung eindeutiger Schlüssel. Sind keine identifizierenden Attribute vorhanden, so müssen Ähnlichkeitsmasse für Attributwerte definiert werden, auf deren Grundlage eine Wahrscheinlichkeit der Übereinstimmung von Tupeln berechnet werden kann [Rahm/Do 2000, 7]. Ähnlichkeitsmasse hängen stark vom Anwendungskontext bzw. von der Art der Daten ab [Leser/Naumann 2007, 335]. Ein ausführlicher Überblick über eine Vielzahl von Ähnlichkeitsmassen sowie über deren Vor- und Nachteile ist u. a. bei ELMAGARMID ET AL. (vgl. [Elmagarmid et al. 2007, 3-6]) sowie bei LESER und NAUMANN (vgl. [Leser/Naumann 2007, 334-340]) zu finden.

Die Ähnlichkeitsmasse beziehen sich lediglich auf den Vergleich einzelner Attribute. In der Praxis müssen jedoch komplexere Datensätze bestehend aus mehreren Attributen überprüft werden, wofür erheblich kompliziertere Techniken und Entscheidungsmodelle notwendig sind. Bei der Duplikaterkennung (Objektidentifikation) können empirische, probabilistische und wissensbasierte Techniken unterschieden werden [Batini/Scannapieco 2006, 106]. Sie sind ebenfalls in der Literatur ausführlich beschrieben (vgl. [Batini/Scannapieco 2006, 106-132], [Elmagarmid et al. 2007, 6-13], [Hernandez/Stolfo 1998, 9-18]).

Mit den bestehenden Verfahren bzw. Techniken ist keine vollständig automatisierte Duplikaterkennung möglich. In der Regel können mehrere Datensätze nur mit einer bestimmten Wahrscheinlichkeit als Duplikate identifiziert werden. Diese potenziellen Duplikate müssen in einem letzten Schritt manuell durch den Anwender verifiziert werden [Low et al. 2001, 593; Batini/Scannapieco 2006, 102]. Eine allgemeine Aussa-

ge darüber, welche der existierenden Techniken zur Duplikaterkennung die qualitativ besten Ergebnisse liefert bzw. am effizientesten arbeitet, ist nicht möglich, da die Verfahren sehr stark vom Anwendungskontext und der Art der Daten abhängen [Elmagarmid et al. 2007, 11]. Zwar sind in der Literatur Experimente mit Beispieldatensätzen zu finden, mittels derer die Präzision der einzelnen Verfahren vergleichend getestet werden (vgl. [Batini/Scannapieco 2006, 125-132], [Low et al. 2001, 597-604]). Jedoch weichen die einzelnen Experimente in ihren Bedingungen und Annahmen teilweise voneinander ab, wodurch die Vergleichbarkeit eingeschränkt wird.

An die Duplikaterkennung schliesst sich die Beseitigung der identifizierten Dubletten und somit die eigentliche Integration von Datenelementen (Datenfusion, engl.: record merging) an, sodass als Ergebnis ein Geschäftsobjekt nur noch durch ein Datenobjekt repräsentiert wird. Problematisch bei der Datenfusion sind widersprüchliche Werte. Derartige Datenkonflikte treten auf, wenn zwei identifizierte Duplikate in mindestens einem Attribut voneinander abweichende Werte enthalten [Leser/Naumann 2007, 345]. In diesem Fall ist zu entscheiden, welcher Wert in den integrierten Datensatz übernommen wird. Mittels Vereinigungs- und Join-Operatoren sowie Kombinationen aus beiden können diese Datenkonflikte nur bedingt behoben werden. LESER und NAUMANN diskutieren ausführlich Vor- und Nachteile der Datenfusion mit Hilfe von Vereinigungs- und Join-Operatoren (vgl. [Leser/Naumann 2007, 347-353]).

T.III.1 Datenmigration

Die Datenmigration bezeichnet die Übernahme von Daten aus abzulösenden Altapplikationen in definierte Zielapplikationen. Die Datenmigration setzt sich aus den drei Schritten Extraktion (aus der Altapplikation), Transformation (gemäß dem Datenmodell der Zielapplikation) und Ladung der Daten (in die Zielapplikation) zusammen [Haller 2009, 65f.]. Entsprechend werden ETL-Tools (Extract, Transform, Load) zur Unterstützung der Migration eingesetzt. Das Vorgehen und das Projektmanagement bei der Migration (vgl. [IMG 1999, 146-152; Morris 2006]) sowie Migrationsarchitekturen und Implementierungsansätze für Extraktion, Transformation und Ladung der Daten (vgl. [Haller 2009]) sind in der Literatur ausführlich beschrieben.

In der vorliegenden Methode beschränkt sich die Datenmigration als Technik auf die technologische Übertragung der Daten. Vorgelagerte Aktivitäten der Datenanalyse (vgl. [Wu et al. 1997; Endava 2006]) wurden in den vorherigen Methodenfragmenten behandelt. Die daraus resultierenden Ergebnisse und Ergebnisdokumente fließen in die Migrationsplanung sowie die eigentliche Migration ein. Im Zuge der Transformation bzw. Datenladung wird oftmals eine Überprüfung des Datenbestandes auf Duplikate und deren Beseitigung (siehe vorherige Technik) durchgeführt. Diese kann durch eine vorgelagerte Dublettenprüfung und -bereinigung in der Altapplikation ergänzt werden, um eine hohe Qualität der zu migrierenden Daten zu gewährleisten.

5.7.4 Werkzeugunterstützung

Für die Dokumentation, Analyse und Gestaltung der Stammdatenintegrationsarchitektur bieten sich Software-Tools zur Modellierung sowie zum Management der Unternehmensarchitektur, sogenannte *Enterprise-Architecture-Management-Werkzeuge (EAM-Werkzeuge)*, an. Diese ermöglichen die Abbildung verschiedener Unternehmensarchitekturkomponenten (Daten, Applikationen, Geschäftsprozesse) sowie die Analyse über mehrere Architekturebenen hinweg. Folglich können Daten-, Applikations- und Integrationsarchitekturen mit Hilfe desselben Werkzeuges modelliert und wichtige Querbezüge (Zuordnung der modellierten Datenobjekttypen zu Applikationen, Darstellung der Datenflüsse zwischen Applikationen, Verknüpfung von Geschäftsobjekttypen mit der Geschäftsprozessarchitektur) veranschaulicht werden.

Ein EAM-Werkzeug, das unter anderem im Aktionsforschungsprojekt bei der Daimler AG zur Anwendung kam, ist die Software-Suite *planningIT* von Alfabet (vgl. [Alfabet 2009]). Die darin enthaltenen Module *Enterprise Architecture Management* und *Application Architecture Management* stellen verschiedene Modelltypen zur Beschreibung von IS auf Geschäfts-, Applikations-, Informations- und technischer Ebene zur Verfügung. Schwerpunkt von *planningIT* ist die Applikationsebene, auf der sämtliche Applikationen mit dem primären Ziel der IT-Planung inventarisiert bzw. dokumentiert werden können (Applikationsarchitekturdiagramm). Auf Informationsebene können Geschäfts- und Datenobjekttypen dokumentiert und mit wenigen (jedoch erweiterbaren) Metadatenattributen beschrieben werden. Das Werkzeug bietet somit auch vereinfachte BDD-Funktionalitäten an. Die Abbildung der Integrationsarchitektur ist durch automatisch erstellbare Datenflussdiagramme (Austausch der Datenobjekte zwischen verschiedenen Applikationen) und Informationsflüsse (Austausch der Geschäftsobjekte in den Prozessen) möglich. Auf Grundlage dieser Dokumentation sind umfassende Analyseberichte zur Identifikation von Redundanzen sowie Abweichungen zu einem definierten Soll-Szenario sowohl ebenenspezifisch als auch ebenenübergreifend möglich.

Einen ähnlichen Ansatz, allerdings mit ausschliesslichem Fokus auf EAM, verfolgt das vom Institut für Wirtschaftsinformatik der Universität St. Gallen gemeinsam mit der BOC Information Technologies Consulting AG entwickelte Software-Werkzeug *ADOben* [ADOben 2009; Kurpjuweit 2009, 208-211]. Darin sind 18 Modelltypen definiert, die jeweils eine Sicht auf die Unternehmensarchitektur (einen sogenannten Viewpoint) repräsentieren und einer Ebene des Business Engineering zugeordnet werden können [Kurpjuweit 2009, 208]. In Bezug auf die für diese Arbeit bedeutenden Teilarchitekturen wird:

- die (statische) Datenarchitektur im Modelltyp Datenmodell abgebildet,

- die Applikationsarchitektur durch die Modelltypen Applikationen (Bestandsführung) sowie Softwarekomponenten (Bestandsführung) abgedeckt, und
- die Integrationsarchitektur durch die Applikationslandkarte (Datenflüsse) sowie die Informationslandkarte (Informationsflüsse) repräsentiert.

Bei der Implementierung neuer Integrationsarchitekturen (siehe Aktivität MF3: III.1) und den damit verbundenen Aufgaben der Extraktion, der Transformation und des Imports von Stammdaten sowie der Dublettenprüfung und -bereinigung bieten eine Vielzahl von *MDM-Tools* weitreichende Funktionalitäten (vgl. [Otto/Hüner 2009, 35ff.], [Spath et al. 2009, 26ff.]).

5.8 Dokumentationsmodell

Die Entwurfsergebnisse der Methode wurden bereits im Rahmen der einzelnen Aktivitäten und Techniken beschrieben bzw. werden in Kapitel 6.1 anhand konkreter Ergebnisse aus den Aktionsforschungsprojekten illustriert. Daher beschränkt sich das Dokumentationsmodell in Abbildung 5-17 auf eine überblicksartige Darstellung sämtlicher Entwurfsergebnisse der Methode sowie die Abhängigkeiten zwischen ihnen in graphischer Form. Die Nummerierung der einzelnen Ergebnisse bzw. Ergebnisdokumente verweist auf die Aktivitäten des Vorgehensmodells, in denen sie erstellt werden. Die dargestellten Beziehungen implizieren primär inhaltliche, teilweise auch zeitliche Abhängigkeiten. Die Abbildung verdeutlicht, dass Ergebnisse aufeinander aufbauen, d. h. Ergebnisse früherer Aktivitäten in die Nachfolgeaktivitäten einfließen.

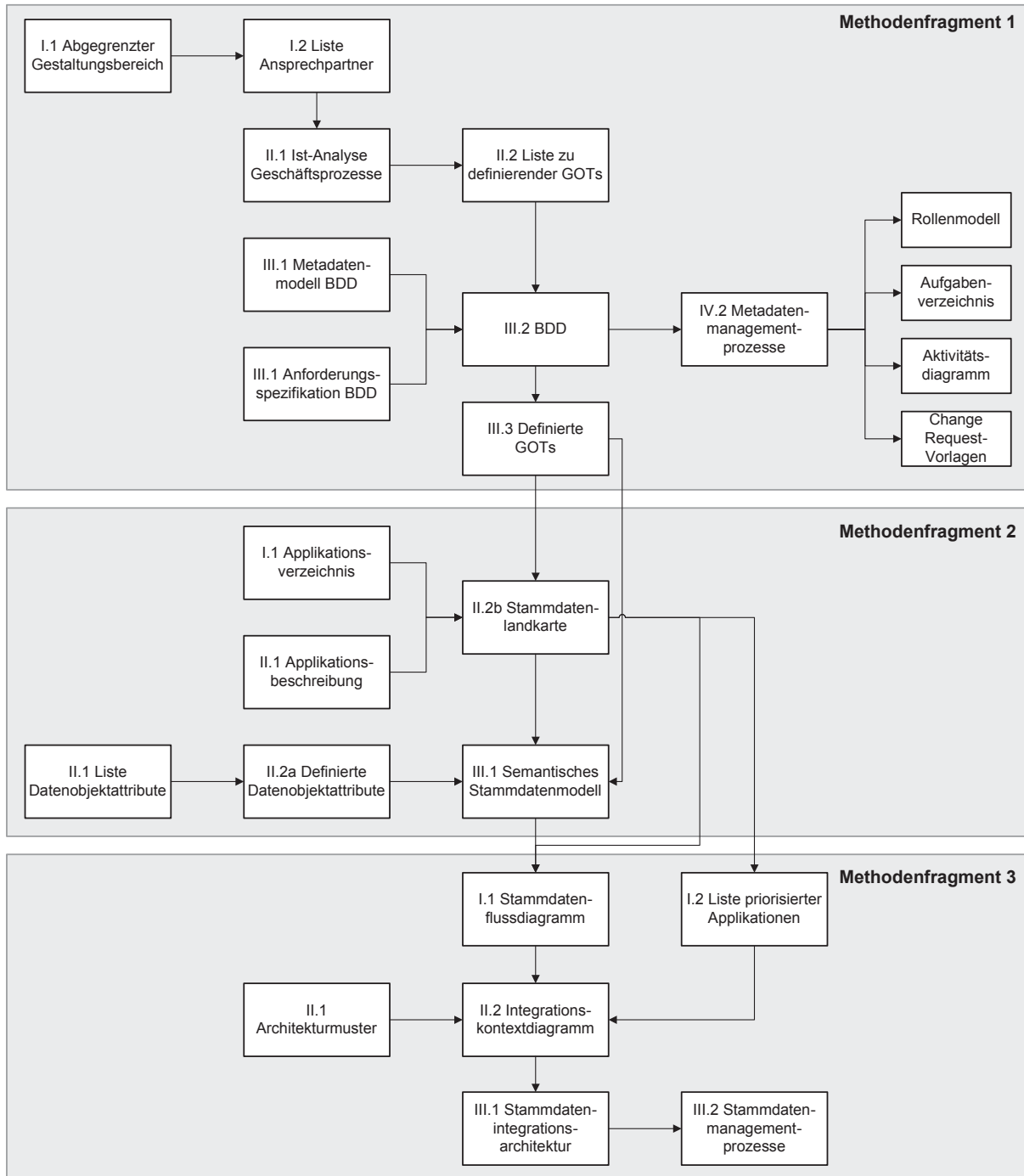


Abbildung 5-17: Entwurfsergebnisse der Methode zur Stammdatenintegration

Anhang B.3 enthält eine Gegenüberstellung der Entwurfsergebnisse und der Metaentitätstypen des Metamodells, auf die sich jeweils ein Ergebnis bezieht. Die Gegenüberstellung lässt zwei Rückschlüsse in Bezug auf das Metamodell zu: Zum einen verdeutlicht sie, dass sämtliche Metaentitätstypen notwendig sind und das Metamodell somit die minimale Anzahl an Gestaltungsobjekten enthält. Zum anderen zeigt die Gegenüberstellung, dass keine weiteren Metaentitätstypen erforderlich sind und das Metamodell folglich vollständig ist.

6 Anwendung und Bewertung der Methode

In den vorhergehenden Kapiteln wurde eine Methode zur Stammdatenintegration entwickelt. Dieses Kapitel beschreibt Anwendung und Bewertung der Methode. Kapitel 6.1 dokumentiert die Verwendung der Methode bzw. einzelner Methodenfragmente in den beiden Aktionsforschungsprojekten bei der Daimler AG und der DB Netz AG. Anschliessend bewertet Kapitel 6.2 die Methode einerseits merkmalsbasiert, andererseits anhand von Fokusgruppeninterviews. Eine Aufwand-Nutzen-Betrachtung in Abschnitt 6.3, die wesentlich auf Erfahrungen aus den Aktionsforschungsprojekten und Erkenntnissen der Fallstudien aufbaut, schliesst das Kapitel ab.

6.1 Anwendung der Methode in Aktionsforschungsprojekten

Die beiden Aktionsforschungsprojekte folgen den fünf Phasen der (kanonischen) Aktionsforschung Diagnose, Aktionsplanung, Aktionsumsetzung, Evaluation und Lernen (vgl. [Baskerville/Wood-Harper 1998, 97; Davison et al. 2004, 72]). Die fünf Phasen sind jeweils mit ihren Ergebnissen in Tabelle 6-1 beschrieben.

Phase	Beschreibung	Ergebnis
Diagnose	Probleme als Ursachen des organisatorischen Wandels bestimmen und gewünschtes Ergebnis beschreiben	Problembeschreibung und mögliche Lösungen
Aktionsplanung	Massnahmen zur Behebung der Probleme festlegen	Aktionsplan mit Zielen und Vorgehen
Aktionsumsetzung	Durch Umsetzung des definierten Aktionsplans aktiv in Organisation eingreifen	Veränderte Organisation
Evaluation	Ergebnisse und Konsequenzen des Eingriffs analysieren und durchgeführte Lösung bewerten	Veränderte theoretische Annahmen und Massnahmen (bei Misserfolg)
Lernen	Generelle Erkenntnisse ableiten und dokumentieren	Erkenntnisfortschritt für Wissenschaft und Praxis

Tabelle 6-1: Phasen der kanonischen Aktionsforschung [Susman/Evered 1978, 588; Baskerville/Wood-Harper 1998, 97]

Das Kapitel konzentriert sich dabei auf die Ergebnisse der Phase der Aktionsplanung und -umsetzung. Die Aktionsumsetzung wird anhand der Phasen der Methodenfragmente erläutert. Die Phase Diagnose wurde im Zuge der Vorstellung der Projekte in Kapitel 4.3 beschrieben. Die Phasen Evaluation und Lernen werden im anschliessenden Kapitel 6.2 behandelt.

6.1.1 Integration von Infrastrukturstammdaten bei der DB Netz AG

Die bereits in Kapitel 4.3.2 einleitend beschriebenen Geschäftstreiber und die damit verbundenen Fragen führten zur Initiierung von acht DQM-Arbeitspaketen⁴⁴. Eines dieser Arbeitspakete diente der Etablierung eines Datenkatalogs (BDD) mit unterneh-

⁴⁴ Die acht DQM-Arbeitspakete sind als Ergebnisse der Aktionsplanung für das Aktionsforschungsprojekt in Anhang A.2 gemeinsam mit der geplanten Projektdauer sowie dem Arbeitsmodus zusammengefasst.

mensweit einheitlichen Definitionen von Infrastrukturstammdaten, ein zweites der Analyse der Informationsflüsse. Die Definitionen sollten insbesondere zwischen dem Fachbereich und der IT sowie den einzelnen in den Geschäftsprozessen beteiligten Abteilungen eine gemeinsame Sprache etablieren. Verantwortlich hierfür war die 2007 geschaffene Abteilung Infrastrukturdatenmanagement (IDM). Sie verfolgt das Ziel einer ganzheitlichen Unterstützung der Prozesse der DB Netz (Fahrplan, Anlagenbau, Vertrieb, Betrieb und Instandhaltung) mit qualitativ hochwertigen Infrastrukturdaten.

Mit dem BDD verbundene Ziele des Projektes waren:

- Ein generisches Metadatenmodell zur Beschreibung der Infrastrukturdaten,
- eine „single source of truth“ mit Definitionen der wesentlichen Infrastrukturstammdaten und ihrer Attribute aus einer fachlichen Perspektive, die als Referenz für sämtliche Mitarbeiter dient,
- definierte Verantwortlichkeiten für Stammdatenobjekte,
- Analyse und Verbesserung der Applikationsschnittstellen.

Methodenfragment 1 – Vorbereitung

Um den initialen Aufwand zu beschränken, wurde zu Beginn des Projektes der Geltungsbereich eingeschränkt und ein Geschäftsprozess identifiziert, für dessen Geschäftsobjekttypen ein erstes BDD entwickelt werden sollte. Für den Geschäftsprozess wurden folgende Kriterien definiert:

- An der Ausführung des Geschäftsprozesses sollten mehrere Geschäftseinheiten (DB Netz, DB Energie, DB Projektbau usw.) beteiligt und der Zugriff auf mehrere Applikationen erforderlich sein.
- Der Geschäftsprozess sollte regelmässig durchlaufen werden und daher für die DB Netz von Bedeutung sein.
- Es sollte eine gute Dokumentation des Geschäftsprozesses (detailliertes Prozessmodell, Handbuch) vorhanden sein.

Für die Anwendung des Vorgehens zur Erstellung des BDD wählten die Unternehmensvertreter schliesslich den Prozess *Umbau eines Bahnsteiggleises*, da dieser sämtliche Kriterien erfüllte.

Methodenfragment 1 - Bestandsaufnahme

Ausgangspunkt der Bestandsaufnahme waren die in einer Excel-Tabelle beschriebenen einzelnen Phasen, Aufgaben und zugehörigen Aktivitäten des Geschäftsprozesses (in Abbildung 6-1 ganz oben dargestellt).

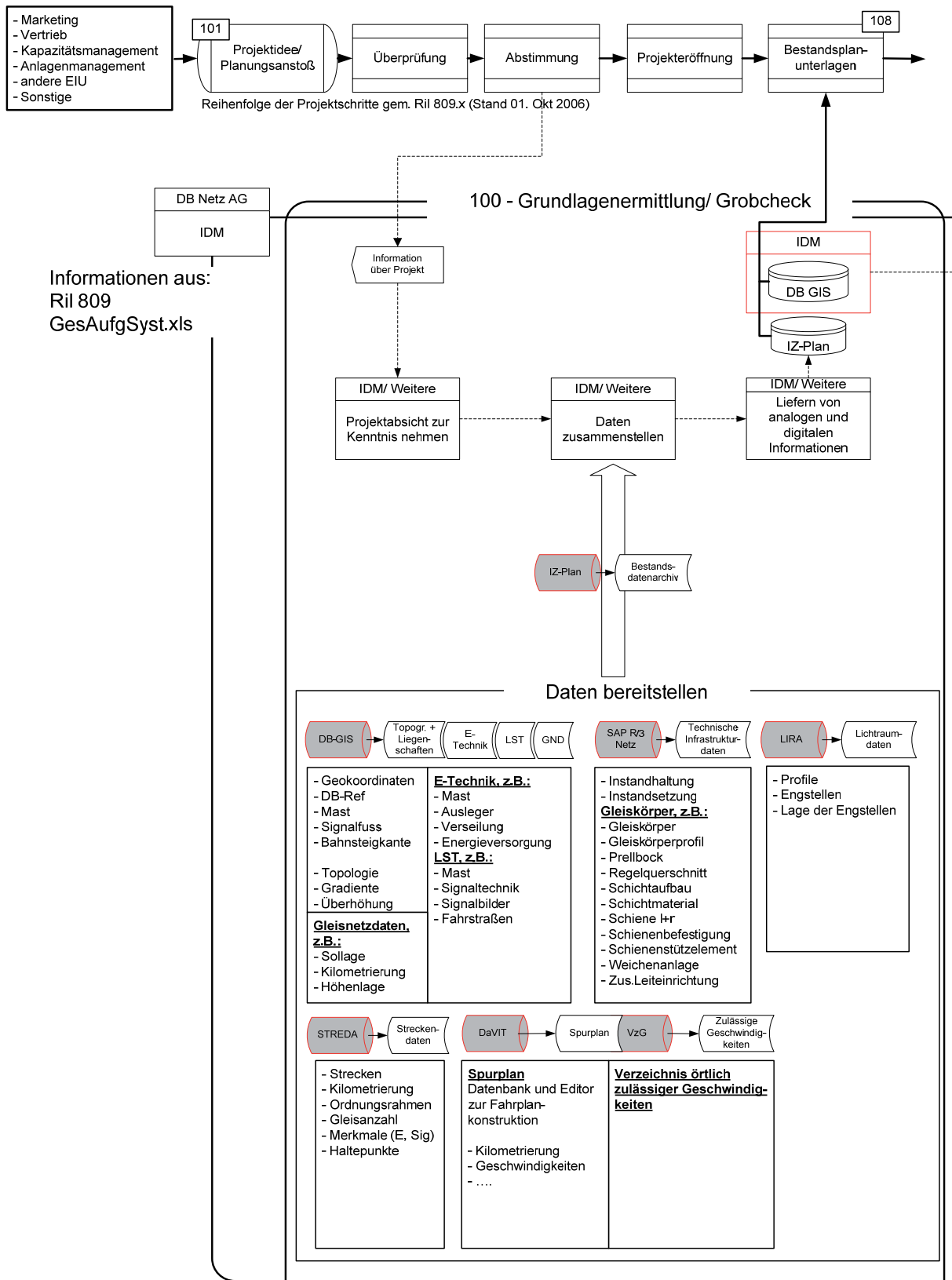


Abbildung 6-1: Auszug aus dem Informationsflussdiagramm für den Geschäftsprozess Umbau eines Bahnsteiggleises

Aus der Prozessbeschreibung konnten die involvierten Konzernbereiche (DB Netze Fahrweg, DB Netze Personenbahnhöfe, DB Netze Energie, DB Netze Projektbau, DB

System) sowie Fachbereiche der DB Netz AG (IDM, Anlagenbuchhaltung, Instandhaltung, Fahrplan) ermittelt werden. Über die Fachbereiche wurden weitere Informationen zum Geschäftsprozess in die Analyse einbezogen, unter anderem in Microsoft PowerPoint und Visio dokumentierte Datenflussdiagramme. Zusätzlich konnten die in den Fachbereichen zuständigen Prozess-Owner und fachlichen Datenmanager, die den fachlichen Datenstewards im Rollenmodell entsprechen⁴⁵, identifiziert werden, um von ihnen verwendete bzw. verantwortete Geschäftsobjekte sowie Applikationen zu bestimmen. Als weitere Informationsquelle konnte das Infrastrukturkataster genutzt werden, in dem sämtliche Infrastrukturobjekte des DB-Konzerns inventarisiert werden.

Auf Grundlage der bestehenden Dokumentation sowie der mit den Prozess-Ownern durchgeführten Interviews wurde der Geschäftsprozess Umbau eines Bahnsteiggleises mit Fokus auf die im Geschäftsprozess verwendeten Entitäten modelliert. Ein Auszug des resultierenden Informationsflussdiagramms ist in Abbildung 6-1 dargestellt. Das Informationsflussdiagramm visualisiert für jede der Prozessphasen (im Diagramm mit gestrichelter Linie umrandete Spalten) die für eine Organisationseinheit (in Abbildung 6-1 ist aus Gründen der Übersichtlichkeit nur der Bereich IDM abgebildet) durchzuführenden Aktivitäten. Zusätzlich sind die jeweils für die Ausführung einer Aktivität benötigten Stammdaten (zugeordnet zu den Applikationen) als einfache Aufzählung enthalten. Hierbei wurde noch keine Unterscheidung zwischen Objekt und Attribut getroffen. Die Informationsflussdiagramme bildeten den Ausgangspunkt für die Identifikation und Priorisierung der in einem fachlichen Metadatenkatalog zu beschreibenden Geschäftsobjekttypen (die Ergebnisse der Priorisierung enthält Tabelle 5-18).

Methodenfragment 1 - Definition

Zur Beschreibung der Geschäftsobjekttypen wurde initial ein Metadatenmodell mit möglichst geringer Komplexität definiert. Hierzu wurden auf Grundlage des Referenz-Metadatenmodells (siehe Abbildung 5-9) für die DB Netz AG relevante Metadaten ausgewählt. Das BDD umfasst folgende Metadaten zur Definition von Geschäftsobjekttypen:

- Name
- Unternehmensweit gültige Definition
- Globale Attribute (unternehmensweit gültig)
- Synonyme / Verwandte bzw. weiterführende Begriffe
- Beziehungen zu anderen Geschäftsobjekttypen (inklusive Hierarchiebeziehungen)
- Fachbereiche, in denen Geschäftsobjekte genutzt werden

⁴⁵ Das vollständige Rollenmodell der DB Netz AG ist bei [Weber 2009, 187-191] dokumentiert und zeigt die Ausprägungen der in Kapitel 5.4 definierten Rollen für das Unternehmen.

- Zugeordnete Datenobjekttypen / Abweichende Objektdefinitionen
- Applikationen, in denen Datenobjekte gespeichert werden
- Weitere Anmerkungen

Hierbei ist zu beachten, dass das BDD bewusst voneinander abweichende bzw. ergänzende Definitionen verwaltet, die sich aus prozessspezifischen Sichten oder unterschiedlichen Datenmodellen der Applikationen (siehe Abbildung 2-6) ergeben. Diese Definitionen sowie applikationsspezifische, lokale Attribute wurden u. a. als Ergebnis der Analyse der Datenmodelle (siehe Methodenfragment 2 – Bestandsaufnahme & Definition) schrittweise hinzugefügt. Aus den verschiedenen Definitionen wird eine konsolidierte, unternehmensweit gültige Definition des Geschäftsobjekttyps abgeleitet. Die globalen Attribute werden durch einen Attributnamen, eine Attributbeschreibung, die Zugehörigkeit zu einem Geschäftsobjekttyp, eine Wertetabelle sowie die Angabe, ob es sich um ein Schlüsselattribut handelt, spezifiziert.

Für den Definitionsprozess wurde ein dezentraler, kollaborativer Ansatz unter Verwendung der im Unternehmen vorhandenen Wiki-Software (allerdings ohne erweiterte semantische Konzepte) gewählt, wobei einzelne Metadaten als Ergebnis der Bestandsaufnahme bereits dokumentiert wurden. In dem sogenannten IDMwiki sind Mitarbeiter berechtigt, Anmerkungen und Änderungsvorschläge zu Objektbeschreibungen zu verfassen. Die Aufgabe der Konsolidierung obliegt dem fachlichen Datenmanager. Abbildung 6-2 zeigt einen Auszug aus dem wiki-basierten BDD der DB Netz AG mit der zugrunde liegenden Struktur der Beschreibung sowie der Definition des Geschäftsobjekttyps Weiche.

The screenshot shows a wiki page for 'Weiche'. At the top, there are navigation buttons: 'Seite', 'Diskussion', 'bearbeiten', 'Versionen/Autoren', 'verschieben', and 'beobachten'. Below the title 'Weiche', there is a table of contents on the left and a table of systems on the right.

Weiche	
Systeme	SAP R/3 Netz, SAP R/3 K, STREDA, Strecken-GIS, DB-GIS, GFD-I

Inhaltsverzeichnis [Verbergen]

- Objektdefinition
- Attribute für das Infrastrukturdatenmanagement
- Ergänzende Objektbeschreibungen und Attribute in den IT-Systemen der DB Netz
 - SAP R/3 Netz
 - Ergänzende Objektbeschreibung
 - STREDA und Strecken-GIS
 - Ergänzende Objektbeschreibung
 - DB-GIS
 - Ergänzende Objektbeschreibung
 - GFD-I
 - Ergänzende Objektbeschreibung
 - SAP R/3 K
 - Ergänzende Objektbeschreibung
 - Attribute
 - Attribute
- Begriffseingrenzung
 - Synonyme
 - Themenbereich / Oberbegriff
 - Verwandte / Weiterführende Begriffe
 - Beziehungen zu anderen Objekten / Attributen
- Weitere Erläuterungen

Objektdefinition [bearbeiten]

Eine **Weiche** ist eine Gleiskonstruktion, die Schienenfahrzeugen den Übergang von einem Gleis in ein anderes ohne Unterbrechung der Fahrt ermöglicht. Zu unterscheiden sind die Weichen nach der Art in einfache Weichen, Doppelweichen und Kreuzungsweichen. Alle Weichenarten können als gerade Weiche oder Bogenweiche ausgeführt sein.

Abbildung 6-2: Auszug aus dem wiki-basierten BDD zum Geschäftsobjekttyp Weiche

Eine Erweiterung des Metadatenmodells mit dem Ziel einer umfangreicheren Sammlung von Metadaten ist für die Zukunft vor allem im Hinblick auf die Gestaltung der Integrationsarchitektur und der Stammdatenverteilung geplant. Ausserdem sollen Metadaten zu Objektverantwortlichkeiten (fachlichen Datenmanagern) sowie Geschäftsregeln im BDD hinterlegt werden.

Methodenfragment 1 - Metadatenpflege

Aufgrund des Einsatzes eines Wiki als fachlicher Metadatenkatalog ist der Pflegeprozess für die Metadaten derzeit nicht explizit definiert bzw. modelliert, sondern durch das Rollen- und Berechtigungskonzept des IDMwiki vorgegeben. Grundsätzlich darf jeder Mitarbeiter über die Diskussionsseiten jedes Geschäftsobjektyps Anmerkungen und Änderungswünsche in das IDMwiki einfügen. Diese werden anschliessend durch den fachlichen Datenmanager des jeweiligen Fachbereiches konsolidiert und im IDM-Arbeitskreis Netz, der sich aus fachlichen und technischen Datenmanagern zusammensetzt, diskutiert. Der IDM-Arbeitskreis Netz trifft schliesslich auf Basis des Vorschlages des zuständigen fachlichen Datenmanagers die finale Entscheidung über die Änderung und Veröffentlichung von Metadaten.

Eine mögliche Ablösung des wiki-basierten BDD (IDMwiki) durch ein kommerzielles Werkzeug für das Metadatenmanagement oder ein semantisches Wiki, welches weitergehende Funktionalitäten bietet, ist derzeit in der Evaluationsphase. Hierzu soll jedoch als erstes die Definition der wichtigsten Geschäftsobjektypen abgeschlossen werden.

Methodenfragment 2 – Bestandsaufnahme & Definition

Bereits parallel zur Auswertung der Geschäftsprozesse wurde mit einer ausführlichen Analyse der bestehenden Applikationen und der ihnen zugrunde liegenden Datenmodelle begonnen. In den Analyseprozess wurden auch die jeweiligen System-Owner involviert. Die Datenobjektypen wurden anhand der Datenobjektattribute, die einem Geschäftsobjekt zugeordnet werden konnten, identifiziert und anschliessend aufgrund ihrer Wichtigkeit für das Unternehmen (Anzahl der Instanzen bzw. Datensätze, die bei der DB Netz zu einem Datenobjektyp geführt werden) priorisiert.

Tabelle 6-2 zeigt eine Zuordnung der zehn priorisierten Datenobjektypen zu den wichtigsten Applikationen der DB Netz AG. Hierbei wird ein Datenobjektyp immer dann einer Applikation zugeordnet, wenn sie mindestens ein Datenobjektattribut enthält.

Objekt:	DB-GIS	R/3-K	R/3-Netz	GFD-I	DaViT Spurplan	STREDA S-GIS	AG 850	APS	TPS
Bahnübergang	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			
Betriebsstellen	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Brücken	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>			
Gleis	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	
Hauptsignal	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>				
KM-Sprünge	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			
Stellwerke	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>						
Strecken				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>
Tunnel	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>			
Weichen	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	

Legende: DB-GIS.... Geoinformationssystem (Geodaten)
R/3-K..... SAP R/3 Konzern
R/3-Netz...Status der Infrastruktur (nach technischen Plätzen)
GFD-I.....Gemeinsame Fahrplandaten Infrastruktur
AG 850.....Konzernweite Bahnstellen und Organisationsdatenbank

DaViT Spurplan...Betriebssteuerung und Fahrplanung
STREDA S-GIS...Streckendaten, Strecken-Geo-IS
APS.....Anlagenpreissystem
TPS.....Trassenpreissystem

Tabelle 6-2: Liste von Datenobjekttypen mit ihrer Zuordnung zu Applikationen

Die in dieser Phase identifizierten Datenobjekttypen werden im IDMwiki den Geschäftsobjekttypen, die sie auf Systemebene repräsentieren, zugeordnet. In dem Abschnitt „Ergänzende Objektbeschreibungen und Attribute in den IT-Systemen der DB Netze“ werden die Datenobjekttypen mit ihren applikationsspezifischen Bezeichnungen und Attributen dokumentiert. Das BDD der DB Netz AG enthält somit nicht nur Definitionen der Geschäftsobjekttypen, sondern zusätzlich auch Metadaten der dazugehörigen Datenobjekttypen. Während erstere jedoch durch Angabe einer verbalen Definition, wesentlicher Eigenschaften (globale Attribute) sowie von Synonymen primär für die Zielgruppe des Fachbereiches beschrieben werden, umfasst die Charakterisierung der Datenobjekttypen eine tabellarische Auflistung sämtlicher in einer Applikation geführten Datenobjektattribute mit Name, Feldnummer, Typ, Datenformat, (Datenbank-)Tabelle und Kurzbeschreibung. Mit diesem integrierten Ansatz verfolgt die DB Netz AG das Ziel, das BDD als Übersetzungsschicht zwischen Fachbereich und IT zu etablieren.

Methodenfragment 3 – Ist-Architektur

In Bezug auf die Gestaltung der Integrationsarchitektur befindet sich die DB Netz AG derzeit noch in einem frühen Stadium. Dies liegt zum einen daran, dass die ersten beiden Methodenfragmente nach den positiven Erfahrungen mit dem Beispielgeschäftsprozess Umbau eines Bahngleises derzeit in anderen Geschäftsprozessen Anwendung finden, um weitere Geschäfts- und Datenobjekttypen zu identifizieren. Die Sammlung der dazugehörigen Metadaten ist zum jetzigen Zeitpunkt noch nicht abgeschlossen. Zum anderen besteht noch keine vollständige Einigkeit darüber, welche Stammdatenattribute global vorgegeben und welche lokal verändert werden können. Dies ist je-

doch wesentliche Voraussetzung für die zukünftige Datenverteilung auf die einzelnen Applikationen.

Aufbauend auf den Ergebnissen der Analyse der Datenobjekttypen (siehe Kapitel 5.6.2.3) wurden einzelne Aktivitäten der ersten Phase (Vorbereitung) bereits durchgeführt. Insbesondere konnten die stammdatenführenden bzw. -nutzenden Applikationen ermittelt werden und grobe Datenflüsse zwischen den neun in Tabelle 6-2 genannten Applikationen modelliert werden. Die Modellierung der Datenflüsse soll in einem nächsten Schritt detailliert werden, um eine umfassende Analyse und Verbesserung der Schnittstellen zu ermöglichen.

6.1.2 Konsolidierung der Applikationsarchitektur bei der Daimler AG

Als Antwort auf die in Kapitel 4.3.1 erläuterten Herausforderungen plante der Bereich Mercedes-Benz Cars (MBC, früher Mercedes Car Group) die Einführung eines unternehmensweiten DQM. Die Zielsetzungen und Massnahmen des DQM wurden in einem Vorprojekt in Form einer Entscheidungsvorlage erarbeitet⁴⁶:

- Konzeption und Einführung eines Metadaten- und Stammdatenmanagements,
- Erstellung eines Kerndatenmodells MBC,
- Evaluation und Auswahl eines DQM-Tools,
- Etablierung des DQM in einem Data Service Center und Konzeption von DQM-Arbeitspaketen,
- Konzeption Erfolgskontrolle des Data Service Center.

Die Erstellung des Kerndatenmodells in Form eines BDD ist ein wesentliches Zwischenergebnis für die Erreichung des übergeordneten Ziels, die Anzahl der Applikationen zu verringern und somit die Kosten der Applikationsarchitektur zu reduzieren. Für eine systematische Konsolidierung der Applikationsarchitektur ist die Transparenz notwendig, welche Applikation welche Stammdatenobjekte nutzt und speichert. Zur Verwirklichung dieser Transparenz sowie zur Erstellung des Kerndatenmodells wurde die Methode zur Stammdatenintegration angewendet.

Methodenfragment 1 – Vorbereitung

Analog zum Vorgehen bei der DB Netz AG wurde zu Beginn des Projektes der Geltungsbereich für die Stammdatenintegration eingeschränkt und ein Geschäftsprozess identifiziert, für dessen Geschäftsobjekttypen ein erstes BDD entwickelt wurde. Auf Grundlage derselben Kriterien wie im Aktionsforschungsprojekt DB Netz AG (Betei-

⁴⁶ Die Ergebnisse der Aktionsplanung für das Aktionsforschungsprojekt sind in Anhang A.2 anhand der Arbeitspakete, der geplanten Projektdauer sowie des Arbeitsmodus zusammengefasst.

ligung mehrerer Abteilungen und Applikationen am Geschäftsprozess, hohe Bedeutung und regelmässig wiederkehrende Ausführung des Geschäftsprozesses, vorhandenes Geschäftsprozessmodell) wurde der Geschäftsprozess „Kundenauftrag managen“ für die initiale Anwendung der Methode ausgewählt.

Methodenfragment 1 – Bestandsaufnahme

Die bestehende Dokumentation des Geschäftsprozesses bestand in einer veralteten EPK sowie vereinzelt Aufgabenbeschreibungen. Darauf aufbauend wurde der Geschäftsprozess als (erweiterte) EPK modelliert [Rosemann 1996, 122f.; Scheer 2001, 125-128], wie er in Abbildung E-1 (siehe Anhang E) in vereinfachter Form (ohne sämtliche Ereignisse) dargestellt ist.

Zu jeder der in Abbildung E-1 modellierten Aktivitäten wurden dabei die verwendeten Geschäftsobjekttypen identifiziert, in der EPK modelliert und in einer Liste gesammelt, wobei den einzelnen Geschäftsobjekttypen eine Priorität zugeordnet wurde (siehe Tabelle 6-3). Für die Festlegung der Priorität schätzten die Workshop-Teilnehmer die Kritikalität (Bedeutung für den Geschäftsprozess, Anzahl Instanzen) der Geschäftsobjekttypen auf einer zehnwertigen Skala, wobei ein hoher Wert einem Geschäftsobjekttyp eine hohe Priorität zumisst.

Geschäftsobjekttyp	Priorität	Erstmalige Verwendung eines Geschäftsobjektes
Aggregat (Motor, Getriebe, Achse)	8	B&D Aggregate
Angebot	7	Aufträge konfigurieren
Einzelhändler (Retailer)	7	Aufträge konfigurieren
Endkunde (Kunde)	9	Aufträge konfigurieren
Fahrzeug	9	Aufträge konfigurieren
Fahrzeugauftrag (Vertrieb)	9	Aufträge konfigurieren
Fahrzeugdokumente	5	Aufträge einbuchen
Grosshändler (Wholesaler)	7	Aufträge konfigurieren
Komponente	8	Powertrainprogramm planen
Kontingentsplan	7	Kontingente planen
Lieferant	7	Aufträge einbuchen
Produktionsauftrag Aggregat	8	B&D Aggregate
Produktionsauftrag Fahrzeug	8	B&D Fahrzeuge
Produktionsprogrammplan Aggregat	7	Produktionsprogramm planen, Powertrainprogramm planen
Produktionsprogrammplan Fahrzeug	7	Produktionsprogramm planen, Fahrzeugprogramm planen
Rechnung	6	Rechnung stellen
Stückliste	8	Aufträge einbuchen
Teil (Rohteil, Kaufteil, Ersatzteil)	8	Powertrainprogramm planen
Werk (Aggregatwerk, Fahrzeugwerk)	5	Aufträge einbuchen

Tabelle 6-3: Geschäftsobjekttypen des Geschäftsprozesses „Kundenauftrag managen“

Methodenfragment 1 – Definition

Die Geschäftsobjekttypen wurden in Reihenfolge der Priorisierung definiert. Das zugrundeliegende Metadatenmodell für das BDD wurde – auch vor dem Hintergrund der Stammdatenintegration mit dem Ziel der Konsolidierung der Applikationsarchitektur (siehe Kapitel 4.4) – bewusst schlank gehalten. Für jeden Geschäftsobjekttyp wurden in Interviews und Workshops auf Basis des in Anhang D.2 abgebildeten Interviewleitfadens sogenannte Steckbriefe erstellt. Interviewpartner waren Vertreter des Fachbereiches, u. a. Werkleiter und Mitarbeiter des Bereiches Vertrieb. Die Geschäftsobjekttypdefinitionen und die fachlichen Metadaten wurden im EAM-Werkzeug planningIT (siehe Kapitel 5.7.4) gepflegt, um Aufwand und Kosten für eine BDD-Implementierung gering zu halten. Die Daimler AG nutzt das Tool u. a. für seine IT-Bebauungsplanung und die Dokumentation der Applikations- und Integrationsarchitektur. Das Tool unterscheidet Business Objects (Geschäftsobjekte) und Business Data (Datenobjekte), zu denen standardmässig folgende Metadaten gepflegt werden können:

- ID
- Name
- Definition / Beschreibung
- Attribute
- Verantwortlicher für Business Object bzw. Business Data und Nutzergruppen
- Nutzende Geschäftsprozesse (über Zuordnung von Business Objects zu Geschäftsprozessen in CRUD-Matrix)
- Nutzende Applikationen (über Zuordnung von Business Data zu Applikationen in CRUD-Matrix)
- Informationsflüsse (Business Objects) zwischen Geschäftsprozessen und Datenflüsse (Business Data) zwischen Applikationen
- Zuordnung von Business Data zu Business Objects
- Änderungshistorie / Version

Die Metadatenattribute stellen eine Teilmenge des Referenz-Metadatenmodells (siehe Abbildung 5-9) dar. Das Metadatenmodell in planningIT ist flexibel erweiterbar, so dass weitere fachliche Metadaten im Tool dokumentiert werden können. Durch die Pflege in planningIT können die Business Objects den einzelnen Aktivitäten bzw. Geschäftsprozessen sowie die Business Daten den Applikationen, in denen sie geführt werden, zugeordnet werden. Dadurch pflegt MBC automatisch eine Stammdatenland-

karte ihrer Kerndaten und erzielt Transparenz über Entstehungs- und Speicherort von Geschäfts- und Datenobjekten. Ausserdem ist die Zuordnung von Business Data zu Business Objects möglich.

Methodenfragment 1 – Metadatenpflege

Die Definition und Implementierung von Metadatenmanagementprozessen inklusive der Zuordnung von Verantwortlichkeiten zu Aufgaben des Metadatenmanagements steht zum derzeitigen Zeitpunkt noch aus. Sie sollen auf Grundlage eines noch zu erarbeitenden Data-Governance-Modells festgelegt werden. Damit verbunden ist das langfristige Ziel, ein Data Service Center als zentrale Organisationseinheit für das Daten- und Metadatenmanagement zu etablieren.

Methodenfragment 2 – Vorbereitung

Für die Analyse der in den Applikationen gehaltenen Datenobjekttypen werden die Leiter der Center of Competence (die für die Applikationsarchitektur eines bestimmten Geschäftsbereiches verantwortlich sind) sowie die Owner der Applikationen kontaktiert. Diese stellten Datenmodelle ihrer Applikationen zur Verfügung und konkretisierten in Interviews die Definitionen von Datenobjektattributen aus technischer Sicht.

Methodenfragment 2 – Bestandsaufnahme & Definition

An den Interviews zur Definition von Datenobjekttypen (Business Data) nahmen ausserdem die IT-Projektleiter teil. Die Datenobjekttypen wurden in planningIT mit Definitionen und Zuordnungen zu den entsprechenden Business Objects gepflegt. Mögliche Widersprüche zu den in Methodenfragment 1 erstellten Steckbriefen respektive darin enthaltenen Metadaten wurden dokumentiert und anschliessend erneut mit den Vertretern der Fachbereiche diskutiert. Sämtliche Mitarbeiter konnten die finalen Metadaten der Stammdatenobjekte in planningIT lesen und überprüfen, bevor sie vom DQM-Team freigegeben wurden.

Der Schwerpunkt der Phase Bestandsaufnahme & Definition lag neben der Konsolidierung der Definitionen auf der Erstellung der Stammdatenlandkarte. Diese Aktivität wurde parallel zur Modellierung der Datenflüsse (siehe Methodenfragment 3 – Ist-Architektur) durchgeführt. Als Modellierungstool wurde wiederum planningIT verwendet, welches mittels CRUD-Matrizen (siehe Kapitel 5.6.2.4) die Zuordnung von Datenobjekttypen (Business Data) zu Elementen der Applikationsarchitektur (Applikationen) ermöglicht.

Methodenfragment 2 – Datenmodellierung & Kontextualisierung

Auf Basis der im planningIT gepflegten Metadaten wurde vom DQM-Team ein semantisches Stammdatenmodell gemäss Aktivität MF2: III.1 (siehe Kapitel 5.6.2.7) erstellt. Die Modellierung folgte dem in der CCTS definierten Vorgehen (siehe Kapitel

5.6.1) und nutzte dabei die standardisierten Datenobjektattribute der CCL Version 09A (vgl. [UN/CEFACT 2009a]). Bei der Modellierung kam der in Kapitel 5.6.4 beschriebene Modellierungsprototyp zur Anwendung. Durch ihn konnten Datenobjektattribute teilweise auf in der CCL enthaltene BBIEs abgebildet und diese somit wiederverwendet werden. Einzelne Datenobjektattribute (insbesondere solche, die automobilspezifische Geschäftsobjekte wie Aggregate oder Komponenten abbilden) mussten manuell gemäss den Namensregeln der CCTS neu definiert werden. Die Kontextwerte für die einzelnen Kontextkategorien, durch die einzelne Datenobjektattribute (BBIEs in den Zeilen) mit Metadaten semantisch annotiert sind, wurden manuell hinzugefügt. Ein Ausschnitt aus dem Stammdatenmodell ist in Anhang E (siehe Abbildung E-2) zu finden. Aufgrund des übergeordneten Zieles der Konsolidierung der Applikationsarchitektur lag ein besonderes Augenmerk auf der Annotation durch den Systemkontext. Daher umfassen die Kontextwerte der Kontextkategorie System Capabilities Context (Sys) Codes für die Applikationen, welche die identifizierten Datenobjektattribute nutzen. Durch Auswahl eines bestimmten Kontextwertes dieser Kontextkategorie ist die Ableitung applikationsspezifischer Stammdatenmodelle aus dem applikationsübergreifenden Stammdatenmodell möglich.

Methodenfragment 3 – Ist-Architektur

Für das Ziel der Komplexitätsreduktion der Applikationsarchitektur ist die Transparenz über die Beziehungen der Applikationen zueinander notwendig, um erkennen zu können, welche Schnittstellen überflüssig werden bzw. welche neuen Schnittstellenanforderungen entstehen. Daher wurde auf die Modellierung der Datenflüsse und Schnittstellen besonderer Wert gelegt. An der Spezifikation waren dieselben Rollen beteiligt wie bei der Definition der Stammdatenlandkarte (siehe Abschnitt Methodenfragment 2 – Bestandsaufnahme & Definition). Datenflüsse und Schnittstellen werden als „Information Flows“ zwischen Applikationen bzw. Applikationskomponenten im EAM-Werkzeug planningIT modelliert. Ein Ausschnitt aus dem Datenflussdiagramm für den Datenobjekttyp Fahrzeugauftrag ist Anhang E (siehe Abbildung E-3) beigelegt. Zusätzlich können dem Business Data Usage Diagram (Stammdatenlandkarte) in planningIT die Zugriffsoperationen jeder Applikation entnommen werden. Jeder Datenfluss wird mit den in Tabelle 6-4 genannten Attributen gepflegt.

Attribut des Datenflusses	Beschreibung
ID	Eindeutige Kennnummer des Datenflusses (IF-xxxx)
Name	Bezeichnung, die beim Anlegen automatisch aus den Namen von Quell- und Zielsystem des Datenflusses generiert wird
Version	Versionsnummer, die bei Anlage vom System automatisch vergeben wird
Status (inklusive Start- / Enddatum)	Information zum Freigabestatus (Planned, Approved, Operation) des Datenflusses
Quellsystem	Name der Applikation, von welcher der Datenfluss ausgeht
Zielsystem	Name der Applikation, bei welcher der Datenfluss eingeht
Schnittstellentyp (Connection Type)	Art der Schnittstelle mit vorgegebenen Werten (Batch, Manual, Online)
Häufigkeit (Connection Frequency)	Regelmässigkeit des Datenflusses mit vorgegebenen Werten (Daily, Weekly, Monthly, Yearly, On Demand, Realtime)
Beschreibung (Description)	Ausführliche Charakterisierung des Datenflusses und seines Zwecks, inklusive einer Angabe der übertragenen Datenelemente bzw. -objekte

Tabelle 6-4: Attribute von Datenflüssen in planningIT

Im Hinblick auf die zukünftige Gestaltung der Applikations- und Integrationsarchitektur wurden die Applikationen mit ihren Datenflüssen bewertet. Dadurch sollen diese priorisiert und abzulösende Applikationen identifiziert werden. Die Kriterien, die jeweils mit einer zehnwertigen Ordinalskala bewertet werden, umfassen den sogenannten Technical Value⁴⁷ der Applikation sowie des Datenflusses, die Ausfallwahrscheinlichkeit des Datenflusses (hoher Wert bedeutet hohe Ausfallwahrscheinlichkeit) und die Komplexität der Implementierung des Datenflusses (hoher Wert bedeutet hohe Ausfallwahrscheinlichkeit). Ein hoher Gesamtwert (Kritikalität) verweist auf zu ändernde Datenflüsse bzw. potenziell abzulösende Applikationen.

Eine Entscheidung über die zukünftige Applikations- und Integrationsarchitektur und somit über deren Entwurf und Implementierung steht derzeit noch aus, auch weil die Bewertung noch nicht abgeschlossen ist. Damit verbunden ist die Frage, in welcher Form die Methode breitere Anwendung finden soll (z. B. für weitere Geschäftsprozesse).

6.2 Bewertung der Methode

Die Evaluation von konstruierten Artefakten (wie z. B. einer Methode) ist wesentlicher Bestandteil des Forschungsprozesses der konstruktionsorientierten Wirtschaftsinformatik [March/Storey 2008, 726; Österle/Otto 2009, 15]. Das Ziel der Bewertung der Methode ist es, deren Eignung und Nutzen systematisch und zielgerichtet zu bestimmen [Fettke/Loos 2003, 81]. FETTKE und LOOS strukturieren und beschreiben Ansätze zur multiperspektivischen Evaluation, wobei sie im Wesentlichen analytische (z. B. merkmals-, metrik- oder metamodellbasierte) und empirische Bewertungsansätze (z. B. Fallstudien, Aktionsforschung) unterscheiden [Fettke/Loos 2003, 83-89]. Aus

⁴⁷ Der Technical Value steht für das Alter der Technologie, wobei eine hohe Bewertung für eine alte Technologie steht.

diesem Rahmenwerk werden in der vorliegenden Arbeit drei Perspektiven ausgewählt und auf die Methodenevaluation übertragen – die merkmalsbasierte Evaluation als analytischer Ansatz, Evaluation durch Befragung und Evaluation durch Aktionsforschungsprojekte als empirische Ansätze. Die Kombination der verschiedenen Evaluationsansätze erlaubt es, Schwächen einzelner Ansätze auszugleichen und die Methode differenziert – also in Bezug sowohl auf das Schliessen der identifizierten Forschungslücke als auch auf die tatsächliche Lösung eines realen Problems – zu beurteilen [Bucher et al. 2008, 82f.]. Die Bewertung der Methode beruht folglich auf zwei Komponenten:

- auf den in Kapitel 3.3 hergeleiteten Anforderungen an eine Methode zur Stammdatenintegration (merkmalsbasierte Evaluation in Kapitel 6.2.1) und
- auf der Auswertung von Fokusgruppeninterviews mit Fachexperten im Bereich DQM, MDM und Unternehmensarchitektur (Evaluation durch Befragung in Kapitel 6.2.2).

Anschliessend wertet Kapitel 6.3 die Erfahrungen aus der Anwendung der Methode in den Projekten im Sinne der Evaluation als Aufgabe der Aktions- und Konsortialforschung [Österle/Otto 2009, 15] aus und führt darauf aufbauend eine Aufwand-Nutzen-Betrachtung durch (Evaluation durch Aktionsforschung). Hierin fliessen zusätzlich Erkenntnisse aus den Fallstudien ein.

6.2.1 Merkmalsbasierte Evaluation

Eine rein formale, analytische Bewertung der in dieser Arbeit konzipierten Methode ist anhand der in Kapitel 3.3 hergeleiteten Anforderungen an die Methode merkmalsbasiert möglich. Sie dient der Evaluation der entwickelten Methode hinsichtlich ihres Potenzials zum Schliessen der identifizierten Forschungslücke. Die strukturellen Anforderungen an eine Methode sind durch die Spezifikation der Methodenbestandteile vollständig erfüllt: Definition des Metamodells in Kapitel 5.3, Spezifikation des Rollenmodells in Kapitel 5.4, Herleitung eines Vorgehensmodells mit Aktivitäten und zugehörigen Techniken in den Kapiteln 5.5 bis 5.7 sowie Beschreibung des Dokumentationsmodells in Kapitel 5.8.

Die inhaltlichen Charakteristika wurden innerhalb der Methodenfragmente behandelt und entsprechende Lösungsvorschläge erarbeitet. Dies umfasst auch die Diskussion von Software-Tools zur Unterstützung einzelner Aktivitäten. Die inhaltlichen Kriterien und deren Erfüllung im Rahmen dieser Arbeit sind in Tabelle 6-5 zusammengefasst.

Bewertungskriterium	Erfüllung der inhaltlichen Kriterien
Bezug zu Stammdaten	Beschränkung der gesamten Methode auf Stammdatenobjekte und deren Eigenschaften, wodurch der Anwendungsbereich der Methode eingegrenzt werden konnte
Eindeutige Definition von Geschäftsobjekttypen	Vorgehen zur Etablierung einheitlicher Definitionen durch die systematische Sammlung (Explikation), Abstimmung und unternehmensweite Publikation fachlicher Metadaten in einem BDD (siehe Methodenfragment 1)
Management fachlicher Metadaten	Ausführliche Diskussion des Managements fachlicher Metadaten: <ul style="list-style-type: none"> Vorgehen für die systematische Sammlung und Speicherung fachlicher Metadaten in einem BDD im Zuge der Definition (siehe Methodenfragment 1 und Methodenfragment 2) Entwurf und Implementierung von Prozessen und Rollen zur dauerhaften und konsistenten Metadatenpflege und -änderung (als dedizierte Phase in Methodenfragment 1)
Ableitung eines semantischen Stammdatenmodells	Vorschlag eines Vorgehens (inklusive Technologie) zur Integration semantischer Informationen (insbesondere des Verwendungskontextes) in ein Datenmodell, durch die sowohl ein übergreifendes Stammdatenmodell erstellt als auch applikationsspezifische Stammdatenmodelle generiert werden können
Entwicklung der Stammdatenintegrationsarchitektur	Vorgehen und Handlungsoptionen zur Integration und konsistenten Verteilung von Stammdaten in verteilten Applikationsarchitekturen (siehe Methodenfragment 3)
Werkzeugunterstützung	Überblick über mögliche Tools (u. a. kommerziell erhältliche Softwareprodukte) zur Unterstützung der einzelnen Methodenfragmente sowie detaillierte Beschreibung eigenentwickelter Tools für Methodenfragment 1 (Kapitel 5.5.4) und Methodenfragment 2 (Kapitel 5.6.4)

Tabelle 6-5: Erfüllung der inhaltlichen Kriterien durch die entwickelte Methode

6.2.2 Auswertung von Fokusgruppeninterviews

Mit dem Ziel, die Objektivität der Bewertung zu erhöhen, wurden zusätzlich mehrere Fokusgruppeninterviews (vgl. [Morgan 1997, 7ff.; Rosemann/Vessey 2008, 12]) mit Fachexperten aus dem Bereich DQM und MDM durchgeführt, in denen die Methode bzw. einzelne Methodenfragmente evaluiert wurden. Tabelle 6-6 fasst diese Fokusgruppeninterviews mit den jeweils behandelten Methodenbestandteilen (Inhalt) sowie den Teilnehmern zusammen. Die Teilnehmer sind anhand der Unternehmen benannt, die sie vertreten, wobei die Zahl in Klammern für die jeweilige Anzahl der Unternehmensvertreter steht.

Inhalt	Fokusgruppenteilnehmer	Ort, Datum
Diskussion von Methodenfragment 1	Bayer CropScience (1), Cablecom (1), DaimlerChrysler (2), Deutsche Telekom (2), ETA (3), IBM (2), Masterfoods (1)	Esslingen, 25. April 2007
Diskussion und Bewertung von Methodenfragment 1 und 2	alfabet (1), Bayer CropScience (2), Bosch (1), DaimlerChrysler (2), DB Netz (1), Deutsche Telekom (2), E.ON (1), ETA (2), IBM (1), ZF (2)	Berlin, 16. November 2007
Diskussion und Bewertung von Methodenfragment 3	Bayer CropScience (2), Bosch (1), British Telecom (2), DaimlerChrysler (2), DB Netz (2), Deutsche Börse (2), Deutsche Telekom (2), ETA (2), IBM (2), Nestlé (1), ZF (1)	St. Gallen, 16. Januar 2008
Diskussion der Methodenfragmente 2 und 3	Adidas (1), Alstom (2), BearingPoint (1), Deutsche Bahn (1), Helsana (1), Kraft Foods (1), Lidl (1), Metro Group (1), MioSoft (2), Philips (2), Swisscom (1), Teradata (1), Voestalpine (1)	Köln, 09. Februar 2009
Diskussion und Bewertung der gesamten Methode	Credit Suisse (21 Fachexperten, u. a. aus den Bereichen Domänen-, Daten- und Applikationsarchitektur)	Zürich, 08. Februar 2010

Tabelle 6-6: Inhalte und Teilnehmer der Fokusgruppeninterviews

Die Fokusgruppeninterviews umfassten jeweils eine Einführung des Methodenfragmentes bzw. der Methode mit Hintergrundinformationen zur Herleitung und Anwendung, gefolgt von einem Hauptteil mit Fragen zur Evaluation. Grundlage der Bewertung waren in erster Linie etablierte Ansätze zur Evaluation von Referenzmodellen aus der Wissenschaft, die für die Bewertung der Methode übernommen wurden (vgl. [Moody/Shanks 1994, 101-106; Rosemann/Schütte 1997, 17-25; Frank 2006, 123-136]). Die in diesen Ansätzen genannten Kriterien, insbesondere die Grundsätze ordnungsmässiger Referenzmodellierung, stellten den Rahmen für eine strukturierte, merkmalsbasierte Bewertung dar, jedoch war die Evaluation nicht auf diese beschränkt. Die Grundsätze der *Richtigkeit* [Rosemann/Schütte 1997, 19-20] und des *systematischen Aufbaus* [Rosemann/Schütte 1997, 25] wurden in den Fokusgruppeninterviews nicht diskutiert, da diese durch die Spezifikation sämtlicher Methodenelemente respektive die Definition eines methodenfragmentübergreifenden Metamodells formal erfüllt sind. Die *Umsetz- bzw. Anwendbarkeit* der Methode als Bewertungskriterium [Moody/Shanks 1994, 106] wird anhand der Methodenanwendung in den Aktionsforschungsprojekten gezeigt und in Kapitel 6.3 diskutiert.

Das Kriterium der *Klarheit und Verständlichkeit* [Moody/Shanks 1994, 105; Rosemann/Schütte 1997, 22-24] sahen sämtliche Fachexperten erfüllt. Vorgehen und Techniken konnten durch kurze Einführungsvorträge nachvollziehbar dargelegt werden. Die Interviewteilnehmer befanden die Aktivitäten und Techniken als hinreichend detailliert beschrieben, um in der Praxis angewendet werden zu können. Die Aufteilung der Methode sowohl in verschiedene Methodenelemente als auch in Methodenfragmente wurde als sehr theoretisch beurteilt, jedoch stimmten die Interviewteilnehmer überein, dass die Methode dadurch übersichtlicher und strukturierter wird. Die durch das Vorgehensmodell definierte Reihenfolge der Aktivitäten wurde als zielführend bewertet mit der Anmerkung, dass Aktivitäten der Methodenfragmente 1 und 2 teilweise parallel durchgeführt werden können.

Das Kriterium der *Relevanz* [Rosemann/Schütte 1997, 20-21] bezieht sich bei der Bewertung der Methode auf zwei Aspekte: Zum einen auf die Bedeutung der mit der Methode adressierten Problemstellung in der Praxis (Problemrelevanz), zum anderen auf die Relevanz der Methodenbestandteile, insbesondere der Aktivitäten des Vorgehensmodells. Die Problemrelevanz der in den Methodenfragmenten adressierten Themen (Standardisierung der Semantik durch einheitliche Definitionen, Metadatenmanagement, Stammdatenintegrationsarchitektur) wurde aus praktischer Sicht bestätigt. Ein systematisches, schrittweises Vorgehen, welches in Form von Aktivitäten und zugehörigen Ergebnisdokumenten eine Orientierungshilfe für die komplexe Gesamtaufgabe Stammdatenintegration gibt, wurde ausdrücklich begrüsst und für zweckmässig befunden. Ein wesentlicher Diskussionspunkt in mehreren der Fokusgruppeninterviews war die Komplexität der Methode und die Notwendigkeit der einzelnen Aktivitäten des Vorgehensmodells. Hierbei wurde einerseits die Skalierbarkeit der Methode bei einer

hohen Anzahl von Stammdatenobjekten und Applikationen in Frage gestellt. Andererseits wurde die Findung einheitlicher Definitionen gerade bei Stammdatenobjekten, die von einer Vielzahl von Geschäftsbereichen und Nutzern verwendet werden, kritisch beurteilt und als sehr langwieriger Prozess gesehen. In diesem Zusammenhang wird die Unterstützung durch ein wiki-basiertes Werkzeug als Möglichkeit betrachtet, diesen Prozess effizienter zu gestalten. Die Fachexperten konnten allerdings keine alternativen Herangehensweisen – in Bezug sowohl auf die Definitionsfindung als auch auf das gesamte Vorgehensmodell – benennen. Für eine objektive Beurteilung äusseren die Fachexperten daher den Bedarf an Erfahrungswerten aus der Anwendung der Methode sowie an einer Aufwand-Nutzen-Betrachtung.

Bei den letzten beiden Fokusgruppeninterviews wurde der unterschiedliche Detaillierungsgrad der Aktivitäten in den verschiedenen Methodenfragmenten hinterfragt. Während die Aktivitäten in Methodenfragment 2 sehr konkret benannt seien, schätzten die Fachexperten die Aktivitäten von Methodenfragment 2 als relativ generisch und dementsprechend umfangreicher bzw. komplexer ein.

Ausdrücklich positiv und als notwendig bewertet wurde die Kombination aus Top-Down- und Bottom-Up-Vorgehen (Methodenfragment 1 und 2), wobei darauf verwiesen wurde, dass die Anforderungen aus Prozesssicht massgeblich und deren Analyse der Ausgangspunkt sein sollten. Der Top-Down-Ansatz unterstützt die Beschränkung des Anwendungsbereiches auf für ein Unternehmen wesentliche Stammdatenobjekte und gibt die notwendige einheitliche Strukturierung der Modellierung vor, die begleitende Bottom-Up-Modellierung gewährleistet den gewünschten Konkretisierungsgrad. Diese Auffassung deckt sich mit den Erkenntnissen aus den Fallstudien der SBB Cargo und der DTAG.

6.3 Aufwand-Nutzen-Betrachtung

Dieses Kapitel nimmt Bezug auf die Bewertungskriterien *Umsetz- und Anwendbarkeit* [Moody/Shanks 1994, 106] sowie *Wirtschaftlichkeit* [Rosemann/Schütte 1997, 21-22] der Methode auf Basis der Erfahrungen aus den Aktionsforschungsprojekten. Zur Bewertung der Wirtschaftlichkeit werden zudem Erkenntnisse aus den Fallstudien genutzt. Die Betrachtung der Wirtschaftlichkeit wird auf die Phase der Methodennutzung beschränkt, während der Aufwand für die Methodenentwicklung vernachlässigt wird. Der Umsetzungsaufwand der Methode in einem bestimmten Kontext (also einem Unternehmen) sowie die Möglichkeit, diese flexibel an die Anforderungen anpassen zu können (Flexibilität bzw. Adaptierbarkeit), beeinflusst die Wirtschaftlichkeit der Methode [Moody/Shanks 1994, 103-104]. Diesem Aufwand wird der durch die Methode generierte Nutzen (z. B. in Form von verbesserter Datenqualität) gegenübergestellt (Aufwand-Nutzen-Betrachtung).

Die erste Erkenntnis der Aktionsforschungsprojekte für die Bewertung der Methode war die Bestätigung der Umsetzbarkeit der Methodenfragmente 1 und 2 in beiden Fallstudien sowie teilweise von Methodenfragment 3 im Projekt bei der Daimler AG. Daraus resultierte gleichzeitig die Einschränkung, dass die Methode nicht vollständig, d. h. mit sämtlichen Methodenfragmenten und Aktivitäten, implementiert werden konnte. Diese Einschränkung war auf zeitliche (beschränkte Projektlaufzeit) sowie unternehmenspolitische Gründe zurückzuführen. Die unvollständige Anwendung wirft die Frage nach dem Aufwand der Methodenumsetzung auf und inwiefern eine zu hohe Komplexität der Methode hierfür verantwortlich war.

Aufwand

Der Projektzeitraum beider Aktionsforschungsprojekte belief sich auf 1,5 bis 2 Jahre. In dieser Zeit konnten bei der Daimler AG die ersten beiden Methodenfragmente sowie die ersten beiden Aktivitäten des dritten Methodenfragmentes für einen Geschäftsprozess durchgeführt werden. Im Fall der DB Netz konnten lediglich Teile der ersten beiden Methodenfragmente umgesetzt werden. Dies zeigt, dass trotz der Einschränkung des Anwendungsbereiches auf einen Geschäftsprozess die Analyse, einheitliche Definition und Modellierung von Stammdatenobjekten als Voraussetzung zur Integration auf Schemaebene bereits ein sehr langwieriger Prozess ist. Der Aufwand deckt sich jedoch mit vergleichbaren Vorgehensweisen in anderen Unternehmen und den Erkenntnissen aus den in dieser Arbeit vorgestellten Fallstudien (siehe Kapitel 4.2). Diese sind in Tabelle 6-7 zusammengefasst.

	SBB Cargo AG	DTAG	Bosch Rexroth
Umfang	47 Stammdatenobjekte, insgesamt 303 Geschäftsobjekttypen (inkl. Bewegungsdaten)	Sämtliche Stammdatenobjekte	a) Kundenstammdaten b) Materialstammdaten
Projektlaufzeit	06/2008 bis 12/2009	01/2007 bis 12/2009	a) 2000 bis Ende 2005 b) 2000 bis Ende 2006
Personelle Ressourcen	MDM-Projektteam mit 3 Vollzeitstellen	Zwei Abteilungen mit mehreren Modellierungsteams	MDM-Abteilung mit 7 Vollzeitstellen
Bewältigte Aufgaben	Aktivitäten der Methodenfragmente 1 und 2 (bis 12/2010 abzuschliessen) Abschluss der Aktivitäten zur Optimierung der Stammdatenintegrationsarchitektur bis 12/2011 (geplant)	Aktivitäten der Methodenfragmente 1 und 2 (noch nicht vollständig abgeschlossen)	Aktivitäten der Methodenfragmente 2 (teilweise) und 3

Tabelle 6-7: Aufwand und Laufzeit der Projekte zur Stammdatenintegration in den Fallstudien

Insbesondere die beiden Fallstudien SBB Cargo AG und DTAG haben in Bezug auf die bereits bewältigten Aufgaben für die Stammdatenintegration einen ähnlichen Umfang wie die beiden Aktionsforschungsprojekte bei der Daimler AG und der DB Netz AG. Zwar ist die Anzahl der zu definierenden Objekte im Vergleich zu den zehn Stammdatenobjekten bei der DB Netz und den 19 Stammdatenobjekten bei Daimler

grösser, jedoch standen hierfür auch erheblich mehr personelle Ressourcen zur Verfügung. Während bei Daimler lediglich eine Person in Vollzeit für die Durchführung des Projektes zur Stammdatenintegration verantwortlich war, bestand bei der DB Netz die Einschränkung, dass das IDM-Team (bestehend aus 2,5 bis 3,5 Vollzeitstellen) parallel zum Projekt zur Stammdatenintegration weitere Arbeitspakete (u. a. die Erarbeitung eines Data Governance-Modells) zu erfüllen hatten.

Zusätzlich kam in beiden Aktionsforschungsprojekten die fehlende Geschäftsprozessdokumentation als wesentlicher Aufwandstreiber erschwerend hinzu. Dadurch musste viel Zeit in die ersten Aktivitäten des Methodenfragments 1 (u. a. Geschäftsprozessmodellierung und -analyse, Bestimmung der relevanten Ansprechpartner) investiert werden. In Unternehmen, in denen eine umfassendere Dokumentation besteht, fällt dieser Aufwand weg.

Die Möglichkeit, die Methode abhängig vom Anwendungsszenario zu konfigurieren, gestattet eine Verwendung und reduzierte damit den Umsetzungsaufwand. Im Zuge der Methodenanwendung in den beiden Aktionsforschungsprojekten konnte beispielsweise der Definitionsaufwand für die Geschäftsobjekttypen dadurch begrenzt werden, dass das Metadatenmodell vor dem Hintergrund der verfolgten Ziele (Anwendungsszenarien Reporting und Konsolidierung) bewusst eingeschränkt wurde. Dadurch konnten Umfang des BDD sowie des semantischen Stammdatenmodells eingeschränkt werden.

Nutzen

In Bezug auf den Nutzen der Methode können qualitative und quantitative Nutzenpotenziale unterschieden werden. Während erstere Verbesserungen ermöglichen, die monetär nicht bewertbar sind, führen quantitative Nutzenpotenziale zu konkreten Einsparungen. Seitens der Unternehmensvertreter von Daimler und DB Netz wurden u. a. folgende qualitative Nutzenpotenziale der Methode zur Stammdatenintegration genannt:

- Die ausführliche Analyse und Dokumentation der Eigenschaften und Verwendung von Stammdatenobjekten erhöht die *Transparenz* über Stammdaten und deren Herkunft. Sie ist Grundlage für eine Verbesserung der Stammdatenqualität.
- Die Anstrengungen, Stammdatenobjekte projekt- und geschäftsbereichsübergreifend zu definieren und zu standardisieren, *reduziert Mehrarbeit* aufgrund sich wiederholender Neudefinitionen gleicher Stammdatenobjekte. Die einmal festgelegten Eigenschaften von Stammdatenobjekten können in verschiedenen Kontexten und Projekten *wiederverwendet* und falls nötig durch neue Metadaten ergänzt werden.
- Die Vorgabe von unternehmensweit gültigen Referenzdefinitionen und die durch das BDD bereitgestellten Metadaten erleichtern das Verständnis der Stammdatenobjekte und *reduzieren Fehler* durch unsachgemässe Verwendung (z. B. Eintragen

falscher Werte in Datenfelder). In der Folge lassen sich auch manuelle Aufwände zur Fehlerkorrektur und Datenbereinigung verringern.

- Bereits die Diskussion über ein einheitliches Stammdatenmodell wurde als wichtiger Schritt gesehen, die *Ausrichtung der IT an den Geschäftsanforderungen* und die Kommunikation über Abteilungsgrenzen zu fördern, indem eine gemeinsame und konsistente Terminologie definiert wird.

Von grosser Bedeutung für die Nutzenbewertung der Methode ist die Messung der Auswirkungen auf die Stammdatenqualität, insbesondere auf die in Tabelle 2-2 beschriebenen Datenqualitätsdimensionen. Die Aufgaben der Identifikation und Spezifikation von zugehörigen Datenqualitätskennzahlen, der Festlegung von Messverfahren sowie der eigentlichen Messung bilden einen längeren und zeitaufwendigen Prozess, der in zahlreichen Publikationen thematisiert wird (vgl. [Batini/Scannapieco 2006, 181-188], [Caballero et al. 2008, 330-338], [Heinrich et al. 2009, 5:10-5:17], [Otto et al. 2009, 129-132]). Hierfür wurden in beiden Aktionsforschungsprojekten gesonderte DQM-Arbeitspakete definiert (siehe Anhang A.2), mit deren Umsetzung jedoch noch nicht begonnen wurde. Daher können in Bezug auf die Verbesserung der Stammdatenqualität keine quantitativen Aussagen getroffen werden.

Jedoch lassen sich aus den Erfahrungen sowie Gesprächen mit den Unternehmensvertretern in den Aktionsforschungsprojekten erste Rückschlüsse ziehen, welche Auswirkungen die Methodenergebnisse auf einzelne Qualitätsdimensionen haben. Die grössten Verbesserungen wurden bei den Dimensionen Konsistenz, Eindeutigkeit und Zugänglichkeit gesehen (die Definitionen der Dimensionen sind Tabelle 2-2 zu entnehmen). Einen positiven Einfluss auf die Konsistenz der Stammdatenbasis sehen Unternehmensvertreter von Daimler und der DB Netz bereits durch das Vorhandensein einheitlicher Definitionen und ihrer systematischen Pflege in einem BDD gegeben. Während bisher keinerlei einheitliches Verständnis über die Geschäftsobjekttypen und deren Attribute bestand und folglich nicht einmal bestimmt werden konnte, inwiefern die Datenobjekte mit den Definitionen im BDD übereinstimmen, ist eine Überprüfung nunmehr möglich. Die im BDD gepflegten Metadaten bilden damit die Grundlage für eine zukünftige Datenqualitätsmessung in Bezug auf die Dimension Konsistenz. Gleiches gilt für die Überprüfung, ob sämtliche notwendigen Werte und Attribute eines Datensatzes vorhanden sind (Dimension Vollständigkeit).

Des Weiteren fördert die Festlegung eindeutiger Attribute im Rahmen der Geschäfts- und Datenobjekttypdefinition (siehe Methodenfragment 1 und 2), die in das semantische Stammdatenmodell übernommen wird, die eindeutige Repräsentation von Realweltobjekten in den Applikationen (Dimension Eindeutigkeit).

Schliesslich wurde in den Aktionsforschungsprojekten der positive Einfluss der Ergebnisdokumente Stammdatenlandkarte (Aktivität MF2: II.2b) und Stammdatenfluss-

diagramm (Aktivität MF3: I.1) auf die Dimension Zugänglichkeit deutlich. Die Transparenz über Ursprung und Speicherort der Datenobjekte erleichtert es den Datennutzern, Zugriff auf von ihnen benötigte Datenelemente zu erhalten, respektive den IS-Architekten, Änderungen an der Stammdatenintegrationsarchitektur entsprechend den Bedürfnissen der Datennutzer vorzunehmen.

Valide monetäre Nutzenpotenziale der Methode sind derzeit nicht möglich. Dies liegt zum einen daran, dass sich der monetäre Nutzen integrierter Stammdaten eher langfristig und indirekt über die Geschäftsprozesse, die auf die Stammdaten zugreifen, einstellt [Kagermann/Österle 2006, 231]. Zum anderen fehlen etablierte Nutzwertrechnungen für integrierte Stammdaten⁴⁸. Näherungsweise kann das monetäre Nutzenpotenzial mit Hilfe von Expertenschätzungen ermittelt werden. Ein solches Vorgehen zur Erstellung einer Aufwand-Nutzen-Schätzung für die Stammdatenintegration auf Basis eines durchgängigen Vorgehens zur Datenmodellierung⁴⁹ wurde im Fall der Deutschen Telekom AG gewählt. Es soll an dieser Stelle exemplarisch beschrieben werden.

Für die Berechnung monetärer Nutzenpotenziale einer durchgängigen Datenmodellierung als Grundlage für eine effizientere Stammdatenintegration nutzt die DTAG die Phasen des IT-Lebenszyklus – Plan, Build (inklusive Test), Run, Retire – als Strukturierungsrahmen. Auf diese Phasen wird das Gesamt-IT-Budget aufgeteilt. Für jede der Phasen schätzen Experten das Einsparpotenzial des qualitätsbezogenen Anteils. Die vier Schritte des Vorgehens für die Ermittlung des quantitativen Nutzens einheitlicher Datenmodelle sind in Abbildung 6-3 dargestellt.

⁴⁸ Die Problematik der Quantifizierung von Nutzenpotenzialen integrierter Stammdaten zeigt viele Parallelen zur Diskussion des monetären Nutzens von IT (siehe hierzu u. a. [Brugger 2005]).

⁴⁹ Der durchgängige Datenmodellierungsansatz umfasst wie in Kapitel 4.2.2.3 beschrieben die natürlichsprachliche Beschreibung von Geschäftsobjekttypen im Business Object Glossary, das fachlogische Datenmodell (Business Object Model) sowie die applikationsspezifischen Datenmodelle (siehe Tabelle 4-5).

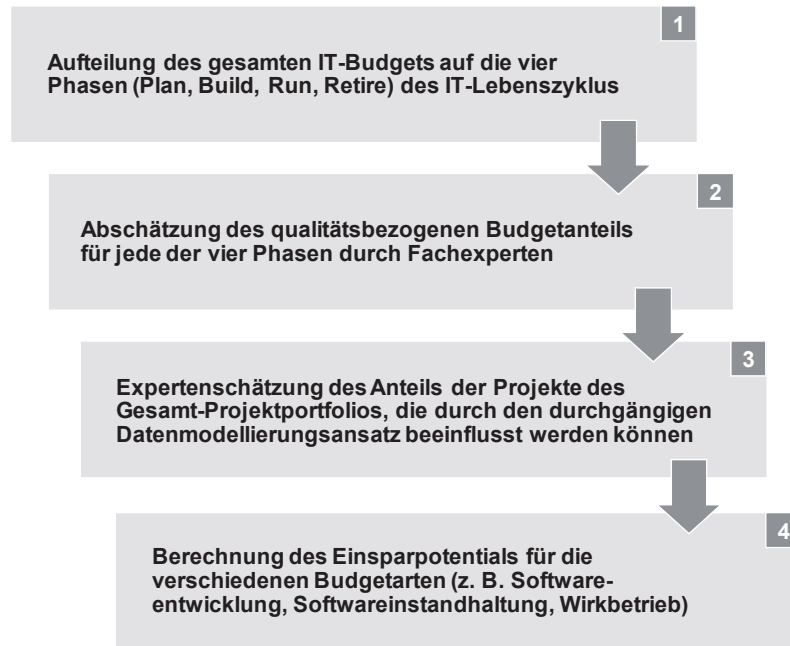


Abbildung 6-3: Vorgehensweise zur Ermittlung des quantitativen Nutzens einheitlicher Datenmodelle (in Anlehnung an [Grewe 2009, 31])

Der qualitätsbezogene Budgetanteil, der im zweiten Schritt von Fachexperten als Prozentsatz geschätzt wird, setzt sich aus den Kosten zusammen, die für Aufgaben des DQM verwendet werden (z. B. Datenmodellierung, Datenbereinigung, Migration). Ausgehend von diesem Budgetanteil wird daraufhin in Schritt 3 der Prozentsatz geschätzt, der durch das durchgängige Modellierungsvorgehen beeinflusst werden kann (sogenannter beeinflussbarer datenqualitätsbezogener Anteil). Der Schätzung liegen die in Tabelle 6-8 genannten Nutzenpotenziale zugrunde [Grewe 2009, 30].

IT-Lebenszyklus	Nutzenpotenziale
Plan	<ul style="list-style-type: none"> • Vereinfachung der Schnittstellen-Vereinbarungen • Verbesserte Wiederverwendbarkeit von Geschäfts- und Datenobjekttypen • Höhere Vollständigkeit der Datenmodelle • Stärkere Standardisierung von Geschäfts- und Datenobjekttypen
Build	<ul style="list-style-type: none"> • Wiederverwendung vorhandener Services/Module • Bessere Verfügbarkeit der Daten • Höherer Standardisierungsgrad • Weniger komplexe Datenmigrationen
Run	<ul style="list-style-type: none"> • Weniger Aufwand für das DQM (z. B. Datenbereinigungen, Konsistenzsicherung) • Kürzere Prozesslaufzeiten • Weniger redundante/überflüssige Daten • Einfachere Datenlieferungen an Data Warehouse, weniger Nacharbeit
Retire	<ul style="list-style-type: none"> • Weniger Abhängigkeiten zwischen Applikationen • Verminderung des Migrationsaufwands • Einfacherer Austausch von Services/Modulen

Tabelle 6-8: Nutzenpotenziale des Modellierungsvorgehens bei der DTAG

Eine beispielhafte Nutzenbetrachtung wurde mit einer Kalkulation basierend auf normalisierten Budgetwerten und über einen Betrachtungszeitraum von drei Jahren erstellt. Sie ist in Tabelle 6-9 abgebildet.

Jahr		Budget	Plan	Build	Test	Run	Retire	Budgetanteil	Savings (%)	Savings (€)	Savings p.a.	% Budget
2008	beeinflussbarer Anteil		20.00%	7.00%								
	DQ bezogener Anteil		12.50%	22.50%								
	ORP gesamt	2'000.00 €	50.00 €	31.50 €	0.00 €	0.00 €	0.00 €	81.50 €	5.00%	4.08 €		
	beeinflussbarer Anteil		10.00%	10.00%	10.00%							
	DQ bezogener Anteil		7.50%	7.50%	10.00%							
	SWI gesamt	0.00 €	0.00 €	0.00 €	0.00 €	0.00 €	0.00 €	0.00 €	5.00%	0.00 €		
	beeinflussbarer Anteil					7.50%						
DQ bezogener Anteil					7.50%							
WIB gesamt	0.00 €	0.00 €	0.00 €	0.00 €	0.00 €	0.00 €	0.00 €	0.00 €	5.00%	0.00 €	4.08 €	0.20%
2009	beeinflussbarer Anteil		50.00%	30.00%	20.00%							
	DQ bezogener Anteil		15.00%	22.50%	12.50%							
	ORP gesamt	1'000.00 €	75.00 €	67.50 €	25.00 €	0.00 €	0.00 €	167.50 €	7.50%	12.56 €		
	beeinflussbarer Anteil		10.00%	10.00%	10.00%							
	DQ bezogener Anteil		7.50%	7.50%	10.00%							
	SWI gesamt	100.00 €	0.75 €	0.75 €	1.00 €	0.00 €	0.00 €	2.50 €	7.50%	0.19 €		
	beeinflussbarer Anteil					10.00%						
DQ bezogener Anteil					10.00%							
WIB gesamt	900.00 €	0.00 €	0.00 €	0.00 €	9.00 €	0.00 €	9.00 €	7.50%	0.68 €		13.43 €	0.67%
2010	beeinflussbarer Anteil		70.00%	50.00%	40.00%							
	DQ bezogener Anteil		17.50%	22.50%	10.00%							
	ORP gesamt	1'000.00 €	122.50 €	112.50 €	40.00 €	0.00 €	0.00 €	275.00 €	10.00%	27.50 €		
	beeinflussbarer Anteil		10.00%	10.00%	10.00%							
	DQ bezogener Anteil		7.50%	7.50%	10.00%							
	SWI gesamt	100.00 €	0.75 €	0.75 €	1.00 €	0.00 €	0.00 €	2.50 €	10.00%	0.25 €		
	beeinflussbarer Anteil					12.50%						
DQ bezogener Anteil					12.50%							
WIB gesamt	900.00 €	0.00 €	0.00 €	0.00 €	14.06 €	0.00 €	14.06 €	10.00%	1.41 €		29.16 €	1.46%
Budget gesamt												
(über 3 Jahre)		6'000.00 €							Savings			
									(über 3 Jahre)	46.66 €		0.78%

Legende: ORP: Neuentwicklungsbudget SWI: SW-Instandhaltung WIB: Wirkbetrieb

Tabelle 6-9: Nutzenbetrachtung für einheitliche Datenmodelle

Die Berechnung zeigt positive Nutzeneffekte, die über die Zeit durch Lerneffekte und Wiederverwendung zunehmen. Das Gesamteinsparpotenzial liegt bei knapp 0,8 Prozent des IT-Budgets. Bei einem geschätzten, durchschnittlichen Anteil des IT-Budgets von 4,6 Prozent des Umsatzes bei Unternehmen der Telekommunikationsbranche [Smith/Potter 2009, 31], liesse sich für die DTAG folglich ein Einsparpotenzial in Höhe von mehr als 23 Mio. EUR abschätzen.

Auch die SBB Cargo hat – wie in der Fallstudie beschrieben – den Nutzen der bisherigen Massnahmen zur einheitlichen Definition und Modellierung ihrer Stammdatenobjekte geschätzt (siehe Kapitel 4.2.1.4). Die Massnahmen entsprechen den Aktivitäten von Methodenfragment 1 und teilweise Methodenfragment 2. Die durch die Massnahmen geschätzte Reduktion administrativer Kosten beläuft sich auf 7,3 Vollzeitmitarbeiter. Gleichzeitig verringert sich die Zahl der nicht abrechenbaren Sendungen, die durch fehlerhafte Daten verursacht wurden. Daraus resultiert ein Einsparpotenzial (z. B. infolge eines geringeren Zinsverlustes) aus mehr als CHF 0,5 Mio.

Zusammenfassend kann geschlussfolgert werden, dass die Ergebnisse der Fokusgruppeninterviews und der Aufwand-Nutzen-Betrachtung für eine Anwendung der Methode sprechen. Durch die hinreichend detaillierten Techniken erlaubt die Methode die Erarbeitung notwendiger (Teil-)Ergebnisse mit einem verhältnismässig geringen Zeit-

aufwand. Die sich bereits aus den Teilergebnissen ergebenden positiven Auswirkungen auf die Konsistenz, Vollständigkeit und Eindeutigkeit der Stammdaten konnten in den Anwendungsprojekten qualitativ und in den Fallstudien auch quantitativ nachgewiesen werden. Weiterentwicklungsbedarf besteht in Bezug auf die Flexibilität der Methodenanwendung. Die Notwendigkeit, einzelne Aktivitäten nur in Abhängigkeit des Anwendungsszenarios durchzuführen, ist ein erster Schritt in Richtung einer höheren Flexibilität, die Identifikation weiterer Einflussfaktoren ist jedoch erforderlich.

Des Weiteren fokussiert die Methode auf global agierende Unternehmen, die sich aufgrund ihrer Grösse und tendenziell dezentraler Organisationsstrukturen durch heterogene Applikationen und komplexe Datenarchitekturen auszeichnen. Eine Anwendung der Methode in kleineren und mittleren Unternehmen ist nicht vorgesehen. In diesen Fällen erscheint eine erheblich reduzierte Methode ausreichend. Da sich die Methode zudem auf Planung und Entwurf von Architekturen beschränkt, wird der gesamte Betrieb einmal realisierter Stammdatenintegrationsarchitekturen nicht abgedeckt. Eine Erweiterung der Methode in diese Richtung erfordert neben Prozessen zur stetigen Aktualisierung des BDD sowie von Daten- und Architekturmodellen eine enge Verknüpfung mit Themen der Data Governance (vgl. [Weber 2009]) und zur Datenqualitätsmessung (vgl. [Otto et al. 2009]).

7 Zusammenfassung und Ausblick

Das folgende Kapitel fasst Ergebnisse und Erkenntnisse der Arbeit zusammen. Anschliessend diskutiert Kapitel 7.2 Einschränkungen der Arbeit, die sich u. a. aus der Forschungsmethode ergeben. Kapitel 7.3 schliesst die Arbeit ab, indem es ausgehend von den gewonnenen Erkenntnissen weitergehenden Forschungsbedarf identifiziert.

7.1 Ergebnisse und Erkenntnisse der Arbeit

Die Arbeit entwickelt eine Methode zur Integration von Stammdaten in global agierenden Unternehmen, die durch dezentrale Organisationsstrukturen und dadurch bedingte heterogene Applikationsarchitekturen gekennzeichnet sind. Sie nutzt das Methoden-Engineering als Bezugsrahmen zur Methodenentwicklung und spezifiziert sämtliche im Methoden-Engineering definierten Methodenelemente: ein Metamodell als Erweiterung des Metamodells des BE-Methodenkerns, ein Vorgehensmodell sowie konkrete Techniken zur systematischen Erarbeitung der Ergebnisse, ein Dokumentationsmodell, das die Ergebnisse miteinander in Beziehung setzt, und ein Rollenmodell, welches organisatorische Rollen für die Stammdatenintegration definiert und sie den Aktivitäten der Methode zuweist.

Mit dem Ziel, den Umsetzungsaufwand einzuschränken und die Flexibilität der Methodenanwendung zu erhöhen, ist das Vorgehensmodell in drei Methodenfragmente unterteilt, die anhand wesentlicher Teilergebnisse inhaltlich voneinander abgegrenzt sind. Methodenfragment 1 zeigt ein Vorgehen für die Etablierung einheitlicher Definitionen von Geschäftsobjekttypen durch die systematische Sammlung fachlicher Metadaten und für den Aufbau einer konsistenten Pflege dieser Metadaten. Dies bildet die Grundlage für eine integrierte Schemadefinition in Methodenfragment 2 in Form eines semantischen Stammdatenmodells, welches die Metadaten – insbesondere den Verwendungskontext der Stammdaten – auf Basis eines standardisierten Vorgehens zur semantischen Datenmodellierung (CCTS) in das Datenmodell einbezieht. Darauf aufbauend beinhaltet das dritte Methodenfragment die physische Integration der Stammdatenobjekte in einer Integrationsarchitektur. Konkrete Architekturmuster für den Entwurf und die Implementierung von Stammdatenintegrationsarchitekturen unterstützen diese Gestaltungsaufgabe. Des Weiteren liefert die Arbeit für jedes Methodenfragment einen Überblick über mögliche Tools zur Unterstützung der Aktivitäten. Dabei stellt die Arbeit kommerziell erhältliche Softwareprodukte kurz vor, während sie eigenentwickelte Werkzeuge mit ihrem Zweck und ihrer Funktionsweise ausführlicher beschreibt. Die Identifikation von Anwendungsszenarien auf Basis der Praxisfälle (Fallstudien und Aktionsforschungsprojekte) dient dem Zweck, die Methode flexibel umsetzen zu können. Je nach Anwendungsfall können einzelne Aktivitäten des Vorgehensmodells optional oder in geringerem Umfang durchgeführt werden.

Zusätzlich zu der Methode als wichtigstes Ergebnis liefert die Arbeit ein Integrationsmodell, welches aktuelle Integrationsprobleme in der Unternehmenspraxis zusammenfasst, nach Integrationsobjekten strukturiert und offene Handlungsfelder ableitet. Als Synthese praktischer Integrationsprobleme verdeutlicht das Integrationsmodell, dass vor allem die semantische Heterogenität auf Datenebene ein in der Praxis noch unzureichend gelöstes Problem ist. Diese Herausforderung wird daher in der anschließenden Methodenentwicklung in den ersten beiden Methodenfragmenten adressiert, indem die Metadaten der Geschäfts- und Datenobjekttypen systematisch erfasst und strukturiert gespeichert werden.

Die zwei Aktionsforschungsprojekte dienen dazu, die Anwendbarkeit der vorgeschlagenen Methode nachzuweisen und sie anhand der Anforderungen aus der Praxis zu evaluieren. Die Ergebnisse der Aktionsforschungsprojekte zeigen, dass durch einen kombinierten Ansatz aus Top-Down- und Bottom-Up-Vorgehen eine umfassende und zwischen Fachbereich und IT konsolidierte Spezifikation von Metadaten zu wesentlichen Stammdatenobjekten möglich ist. Die Metadaten können nicht nur als Grundlage für eine Vereinheitlichung der Terminologie im Unternehmen, sondern auch für die konzeptionelle Datenmodellierung zur applikationsübergreifenden, semantisch korrekten Integration von Datenobjekttypen genutzt werden. Die Anwendung der Methode offenbarte den hohen zeitlichen Aufwand, den eine unternehmensweite Integration von Stammdaten benötigt. Aufgrund der beschränkten Projektlaufzeit konnte die Methode in beiden Fällen nicht vollständig umgesetzt und insbesondere die Integration auf Instanzebene nicht begleitet werden.

7.2 Einschränkungen

Neben Einschränkungen, die aus der zu Beginn der Arbeit vorgenommenen Eingrenzung der Themenstellung resultieren (u. a. die Beschränkung des Untersuchungsbereiches auf strukturierte Stammdaten), unterliegen die Ergebnisse der Arbeit Einschränkungen aufgrund des gewählten Forschungsansatzes (siehe Kapitel 1.3). Die Anwendung der Methode beschränkt sich auf zwei Aktionsforschungsprojekte mit dem Zweck, die Umsetzbarkeit der Methode zu demonstrieren und Erkenntnisse für deren Verbesserung abzuleiten. Der begrenzte zeitliche Rahmen erschwerte es, die Wirkung der durch die Methode ergriffenen Massnahmen, die nicht selten erst nach mehreren Monaten deutlich wird, nachzuweisen. Folglich beschränkt sich die Evaluation der Methode als wesentliche Aufgabe der Aktions- und Konsortialforschung auf eine qualitative Nutzenbetrachtung. Für eine quantitative Bewertung des Nutzens sind längerfristige Projekte erforderlich, die u. a. den Einfluss integrierter Stammdaten auf die Datenqualität messen. Des Weiteren ist eine vollständige Anwendung der Methode in weiteren Unternehmen notwendig, um die Praxistauglichkeit und die Nutzenpotenziale zu beweisen.

Ausserdem besteht eine mögliche Einschränkung der Methodenanwendung in ihrer Abhängigkeit von der weiteren Entwicklung des UN/CEFACT-Standards CCTS (siehe Methodenfragment 2), der zum derzeitigen Zeitpunkt in der Praxis noch wenig verbreitet ist. Mehrere Tendenzen deuten jedoch auf eine zunehmende Verbreitung von Methode (CCTS und Kontexttreiberprinzip) und Inhalt (semantische Datenobjekt-komponenten der CCL) hin. Zum einen spricht die Konvergenz zwischen zahlreichen Standards sowohl aus dem Bereich B2B (UBL, OASIS) als auch aus dem Bereich der Stammdaten (GS1) mit dem UN/CEFACT-Rahmenwerk für einen Konsens zwischen den bisher voneinander abweichenden Standardisierungsinitiativen. Zum anderen sind die Entscheidungen der SAP, ihre Global Data Types auf Basis der CCTS und CCL zu definieren (vgl. [Stuhec 2005]), sowie von Oracle, die CCTS als wesentlichen Baustein in ihrer Standardisierungsstrategie nennt (vgl. [Rowell 2008]), deutliche Zeichen einer zunehmenden Akzeptanz bei Software-Anbietern. In diesem Zusammenhang ist auch die Integration des in der Arbeit beschriebenen Modellierungsprototyps in die SAP NetWeaver-Plattform (und in der Folge auch in SAP-Stammdatenlösung NetWeaver MDM) positiv zu bewerten, da eine toolgestützte Modellierung den Aufwand reduziert und die Akzeptanz des Standards zusätzlich fördert [Vogel 2009, 205].

7.3 Weiterer Forschungsbedarf

Aufbauend auf den Erkenntnissen dieser Arbeit sowie auf laufenden Forschungsaktivitäten, leitet der folgende Abschnitt weitergehenden Forschungsbedarf mit Bezug zur Stammdatenintegration ab und stellt kurz innovative Ansätze zur Lösung des Datenintegrationsproblems vor.

7.3.1 Linked Data

Die Idee von Linked Data hat ihren Ursprung im World Wide Web bzw. im Semantic Web, dessen Architektur – bestehend aus einer Menge lose miteinander verknüpfter Informationen, Internetseiten und Dokumente – auf geschäftsrelevante, unternehmensinterne Daten (z. B. Stammdaten) übertragen werden soll. Im Semantic Web sind sämtliche Informationsressourcen über einen Uniform Resource Identifier (URI) eindeutig bezeichnet und identifizierbar. Mit Hilfe der RDF-Standardfamilie werden diese Ressourcen typisiert, d. h. es werden ihnen beschreibende Informationen (Metadaten) hinzugefügt [Bizer et al. 2009, 2]. Durch die formale Repräsentation der Informationen in RDF können diese durch Maschinen ausgewertet werden.

Wenn diese Technologien auf das Management und die Integration von Daten im betrieblichen Kontext übertragen werden, müssen sämtliche Datenobjekte (auf Instanzebene) zu verknüpften bzw. verknüpfbaren Daten (Linked Data) transformiert werden. Hierzu werden sie mit einem unternehmensweiten eindeutigen Bezeichner (einem URI) versehen. Dadurch können die Datenobjekte z. B. im Intranet eines Unterneh-

mens veröffentlicht werden. Die Datenstrukturen (Datenelemente eines Datenobjektes) werden standardisiert durch RDF beschrieben, die Beziehungen zu anderen Datenobjekten (URIs) mittels RDF Links abgebildet. Das Vorgehen zur Transformation von Datenobjekten zu Linked Data stellt Abbildung 7-1 beispielhaft dar.

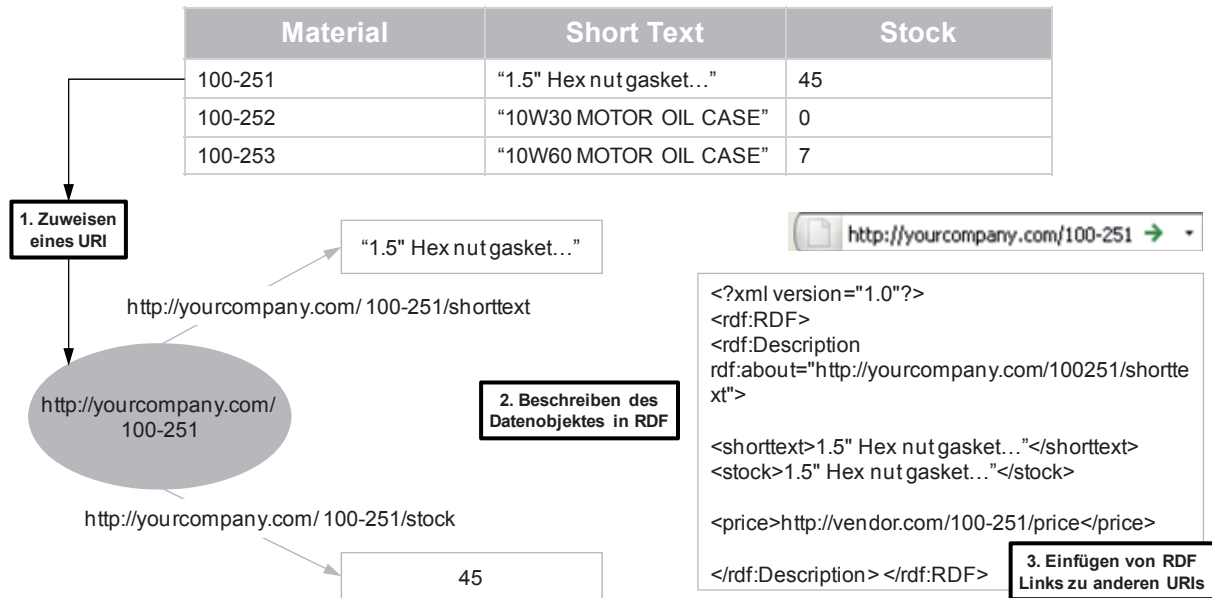


Abbildung 7-1: Übertragung des Linked Data-Konzeptes auf Datenobjekte eines Unternehmens (in Anlehnung an [Hüner/Ofner 2009, 16])

Dadurch entsteht ein datenquellenübergreifendes semantisches Datennetz. Dieses „Web of Data“ bietet mehrere Vorteile:

- es kann Daten beliebiger Formate und dementsprechend auch unstrukturierte Daten integrieren [Bizer et al. 2009, 4],
- es ist auf grosse Datenmengen und auf eine hohe Anzahl von Datenquellen skalierbar [Halb et al. 2008, 2],
- es ermöglicht die Integration von Daten aus Legacy-Systemen,
- es ist nicht auf unternehmenseigene Daten beschränkt, sondern auch auf unternehmensexterne Daten übertragbar.

Da sich die Übertragung des Linked Data-Konzeptes auf die Datenintegration in Unternehmen noch in einem sehr frühen Stadium befindet, gibt es derzeit noch keine Erfahrungswerte aus Pilotimplementierungen. Insbesondere die Aufwände für die Transformation von Datenobjekten in RDF sowie für die Einrichtung der Infrastruktur sind schwer einzuschätzen. So bestehen ungelöste Herausforderungen in Bezug auf die Einbindung komplexer Unternehmensapplikationen (z. B. ERP-Systeme) durch die notwendige Transformation der Daten in RDF. Da eine manuelle Transformation aufgrund der in diesen Applikationen gehaltenen Datenmengen unrealistisch ist, sind

Software-Tools erforderlich, um diese Aufgabe zu automatisieren. Für die Infrastruktur bedarf es spezieller Browser sowie semantischer Suchmaschinen, für die es erste Forschungsprototypen gibt, deren Entwicklung aber noch weit von einer kommerziellen Nutzung entfernt ist.

Die derzeitige Forschung im Bereich Linked Data konzentriert sich daher auf die Weiterentwicklung bestehender Technologien sowie auf die Entwicklung von Gestaltungsprinzipien und Sicherheitskonzepten, die eine erfolgreiche Implementierung des Konzeptes im Unternehmenskontext gewährleisten.

Die Ergebnisse der Methode zur Stammdatenintegration, u. a. die im BDD gepflegten Metadaten, können als hilfreicher Input genutzt werden. Insbesondere die in Kapitel 5.5.4 beschriebene, auf Technologien des semantischen Web beruhende BDD-Implementierung, in denen die Metadaten in RDF-Tripeln abgebildet sind, können den Aufwand für die Umsetzung des Linked Data-Konzeptes erheblich verringern.

7.3.2 Master Data as a Service

Ein in wissenschaftlichen und praxisorientierten Veröffentlichungen zunehmend behandeltes Thema ist der Zusammenhang von Stammdatenintegration und serviceorientierten Konzepten. Hierbei verweisen zahlreiche Publikationen auf die Rolle integrierter Stammdaten als Voraussetzung für die Implementierung einer SOA (vgl. [Herwig 2006, 156; Dreibelbis et al. 2008, 86]). Seltener wird die Anwendung von SOA- bzw. Servicedesignprinzipien thematisiert, um Stammdaten als Service zur Verfügung zu stellen (vgl. [Kagermann/Österle 2006, 189f.; Dan et al. 2007; Liebhart 2007, 288-299]).

Basierend auf einer Analyse von Beiträgen von Wissenschaftlern und Standardisierungsgremien definiert HEUTSCHI einen Service als „ein abstraktes Software-Element bzw. eine Schnittstelle, die anderen Applikationen über ein Netzwerk einen standardisierten Zugriff auf Anwendungsfunktionen anbietet“ [Heutschi 2007, 27]. Stammdatenservices kapseln Funktionalität für den Zugriff und die Manipulation von Daten. Die für die Serviceorientierung als konstituierend geltenden Designprinzipien umfassen u. a. die Schnittstellenorientierung (z. B. durch eine einheitliche Servicespezifikation und Abstraktion von der Serviceimplementierung) sowie die Autonomie bzw. Modularität (z. B. durch schwache logische Kopplung) [Legner/Heutschi 2007, 1644; Vogel 2009, 37]. Die Anwendung dieser Prinzipien auf Stammdaten ermöglicht es, Services mit standardisierten Schnittstellen zu definieren und zu implementieren, über die Daten einheitlich zugänglich gemacht und gepflegt werden können [Truong/Dustdar 2009, 88]. Web Service-Technologien und entsprechende technische Standards, u. a. für die Schnittstellenbeschreibung, für das Nachrichtenformat und für

Verzeichnisdienste für Web Services⁵⁰, bieten eine Möglichkeit zur technischen Umsetzung serviceorientierter Architekturen [Heutschi 2007, 59; Vogel 2009, 43f.].

Angelehnt an die Phasen des Stammdatenlebenszyklus können grundsätzlich zwei Datenservicetypen unterschieden werden:

- Read-only Services stellen lediglich einen Zugang zu bestehenden Quellen her.
- CRUD-Services erlauben es dem Nutzer zusätzlich, Stammdaten zu erstellen, zu verändern und zu löschen, wobei eine weitere Unterscheidung zwischen Services zum Anlegen (Create), Aktualisieren (Update) und Löschen (Delete) möglich ist.

Des Weiteren existieren spezielle Retrieve-Services, welche die Suche nach Stammdatenobjekten unterstützen. Derzeit werden lediglich read-only Datenservices spezialisierter Anbieter von Data as a Service bereitgestellt, die als Online-Services flexibel in Applikationen und Geschäftsprozesse eines Unternehmens eingebunden werden können [Kagermann/Österle 2006, 189f.]. Einen Überblick über ausgewählte Anbieter solcher externen Stammdatenservices gibt Tabelle 7-1.

	Dun & Bradstreet ⁵¹	GS1	SINFOS ⁵²	Strikelron ⁵³ / ServiceObjects ⁵⁴
Stammdatenklasse	Geschäftspartnerstammdaten	Materialstammdaten	primär Materialstammdaten	Geschäftspartnerstammdaten
Funktionalität	Validierung Anreicherung	Identifikation (u. a. GTIN, Global Location Number)	Validierung Anreicherung	Validierung
Einschränkungen	Keine	Keine	Fokus auf Einzelhandel	Nur amerikanische Adressen

Tabelle 7-1: Anbieter externer Stammdatenservices

Der potenzielle Nutzen einer Implementierung von Stammdatenservices innerhalb eines Unternehmens besteht analog zu den Zielen einer SOA in der vereinfachten Prozess- und Applikationsintegration (vgl. [Curbera et al. 2003, 30; Heutschi 2007, 3]). Applikationen können im Prozessverlauf über einheitliche Schnittstellen der Services, die pro Datenobjekttyp oder Datenobjektattribut definiert sein können, zum Bedarfszeitpunkt auf ein Stammdatenobjekt zugreifen. Dadurch wird die durchgängige und medienbruchfreie Aufgabenbearbeitung (siehe Kapitel 2.1.2 zur Prozessintegration) unterstützt. Zudem verspricht der serviceorientierte Ansatz eine höhere Flexibilität und Wiederverwendung (vgl. [Dreibelbis et al. 2008, 55; Vogel 2009, 40]). So hat die Trennung von Applikationslogik und Datenzugriff den Vorteil, dass aufrufende Applikationen kein Wissen über die konkrete Implementierung des Services benötigen, sondern lediglich die Zugriffsparameter kennen müssen. Die zugrunde liegenden

⁵⁰ Eine umfassende Übersicht über den gesamten Web Service Standards Stack geben VOGEL [Vogel 2009, 42] sowie HEUTSCHI [Heutschi 2007, 62].

⁵¹ Siehe <http://www.dnb.com/>

⁵² Siehe <http://www.sinfosweb.de/serviceDE/>

⁵³ Siehe <http://www.strikeiron.com/strikeironservices.aspx>

⁵⁴ Siehe <http://www.serviceobjects.com/products/product-directory>

Datenquellen und Datenmodelle können – bei sich ändernden Anforderungen – flexibel ohne Auswirkungen auf den Datenzugriff verändert werden [Dan et al. 2007], wodurch die Wiederverwendung der Stammdatenservices vereinfacht wird.

Ferner unterstützt das Master Data as a Service-Konzept den Erhalt einer konsistenten Stammdatenbasis. Die Kapselung von Funktionalität zur Datenmanipulation in einem Stammdatenservice, der als „single point of entry“ in Bezug auf ein Stammdatenobjekt dient, fördert die Vermeidung von Duplikaten und die Standardisierung von Stammdaten [Dreibelbis et al. 2008, 86]. Hierfür müssen Datenobjekte auf Instanzebene eine applikationsübergreifend gültige, eindeutige Identifikation besitzen, die als Eingabeparameter bei Aufruf eines Stammdatenservices (z. B. zur Änderung) übergeben wird.

LIEBHART diskutiert servicebasierte Stammdatenintegrationsarchitekturen auf Grundlage verschiedener Architekturmuster, wie sie in Kapitel 5.7.1 dieser Arbeit vorgestellt wurden [Liebhart 2007, 293-299]. Er verdeutlicht, dass die Implementierung von Stammdatenservices grundsätzlich unabhängig von der zugrunde liegenden Architektur ist, seine Ausführungen bleiben jedoch auf einer sehr groben, konzeptionellen Ebene, die keine technische Umsetzung betrachtet oder konkrete Anwendungsfälle beschreibt. Das Fehlen von prototypischen Implementierungen in der Unternehmenspraxis, anhand derer nicht nur die Umsetzbarkeit des Konzeptes gezeigt, sondern der reale Nutzen bewertet werden kann, ist ein allgemeines Defizit von Veröffentlichungen zu diesem Thema. Weitere, aktuell diskutierte Fragen in Bezug auf das Master Data as a Service-Konzept betreffen [Dan et al. 2007; Truong/Dustdar 2009]:

- Welche Eigenschaften unterscheiden Master Data Services von herkömmlichen Business Services? Welche (Meta-)Informationen müssen diese Services folglich beinhalten?
- Welche Granularität sollte ein Stammdatenservice besitzen?
- Wie sieht ein methodisches Vorgehen zur Identifikation, Modellierung und Implementierung von Stammdatenservices aus?
- Wie kann der Nutzen von Stammdatenservices gemessen werden?

Voraussetzung für die Umsetzung des Konzeptes – insbesondere für die Suche, Auswahl und Nutzung von Stammdatenservices – ist eine umfassende Beschreibung der Services sowie der Identität, Semantik, Herkunft (Datenquelle), Verwendung und Qualität der Stammdatenobjekte [Truong/Dustdar 2009, 88-90]. Da die Definition von Stammdatenobjekten und die strukturierte Pflege ihrer Metadaten Gegenstand der vorliegenden Arbeit sind, liefert die Arbeit wichtige Vorarbeiten für die Umsetzung von Master Data as a Service.

Anhang A Dokumentation zur Forschungsmethodik

Anhang A.1 Fallstudieninterviews

Nr.	Datum, Uhrzeit, Ort	Interviewpartner
SBB Cargo AG		
1	16. Juli 2009, 10:00 – 12:00, Basel (Schweiz)	Beat Schwaar, Leiter des Projektes zum Stamm- und Geschäftsdatenmanagement, Qualitäts- und Informationsmanagement (QI), Prozessmanagement & IT, SBB Cargo AG
2	16. Juli 2009, 13:00 – 14:30, Basel (Schweiz)	Mathias Hirschi, Teilprojektleiter Analysesysteme, Management & Accounting (innerbetriebliche Controllingsysteme), Finanzen & Controlling, SBB Cargo AG
3	16. Juli 2009, 14:45 – 16:15, Basel (Schweiz)	Karl-Heinz Diekmann, Bereichsleiter Stammdatenmanagement, gicom GmbH (seit November 2008 als externer Berater aktiv in das Projekt zum Stamm- und Geschäftsdatenmanagement bei der SBB Cargo involviert)
Deutsche Telekom AG		
1	06. November 2010, 09:00 – 10:30, Darmstadt (Deutschland)	Dr. Axel Grewe, Leiter Data Governance, Zentrum für Informationstechnik, Deutsche Telekom AG
2	06. November 2010, 11:00 – 12:30, Darmstadt (Deutschland)	Bernd Hofner, Geschäftsführer Ingenieurbüro technische Informatik, Externer Dienstleister bei der Deutschen Telekom AG
3	06. November 2010, 13:00 – 14:30, Darmstadt (Deutschland)	Dr. Ina Pitschke, Data Architect, Zentrum für Informationstechnik, Deutsche Telekom AG
Bosch Rexroth AG		
1	02. Oktober 2008, 14:00 – 16:00 Gemünden-Langenprozelten (Deutschland)	Werner Schiff, Projektleiter Global Data Client and Master Data Management, Information Technology – Competence Center Global Services, Bosch Rexroth AG
2	21. November 2008, 14:00 – 16:00 St. Gallen (Schweiz) Telefoninterview	Werner Schiff, Projektleiter Global Data Client and Master Data Management, Information Technology – Competence Center Global Services, Bosch Rexroth AG

Tabelle A-1: Übersicht Fallstudieninterviews

Anhang A.2 Charakterisierung der Aktionsforschungsprojekte

Tabelle A-2 charakterisiert die Aktionsforschungsprojekte der Arbeit anhand der Ansprechpartner in den Unternehmen, der definierten Arbeitspakete (als Ergebnis der Aktionsplanung), der Projektdauer sowie dem Arbeitsmodus. Die fett markierten Arbeitspakete verfügen über direkten Bezug zur Arbeit.

Merkmal	DB Netz AG	Daimler AG
Projektpartner	Abteilung Infrastrukturdatenmanagement (IDM)	IT Management (ITM), Methoden Bereich Bebauungsplanung
Arbeitspakete (mit Bezug zur Arbeit)	<ul style="list-style-type: none"> • Entwicklung der IDM-Strategie • Entwurf einer Wirtschaftlichkeitsbetrachtung (Business Case) • Aufbau eines Datenkatalogs (BDD) • Aufbau eines Kennzahlensystems • Entwicklung eines Data-Governance-Modells • Entwurf von Datenmanagementprozessen • Entwurf eines Kommunikationskonzepts • Analyse der Informationsflüsse und Entwurf des IDM-Zieldatenmodell 	<ul style="list-style-type: none"> • Analyse Stammdatenarchitektur • Erstellung eines Kerndatenmodells inklusive fachlichen Metadatenkatalogs (BDD) für Stammdaten • Bestimmung führender Stammdatensysteme als Grundlage für die Konsolidierung der Applikationsarchitektur • Definition von Datenmanagementprozessen (inklusive Rollen und Verantwortlichkeiten) • Etablierung des DQM in einem Data Service Center • Evaluation und Auswahl eines DQM-Tools
Projektdauer	1,5 Jahre (Juni 2007-Dezember 2008)	1,5 Jahre (März 2007-September 2008)
Arbeitsmodus	Zweiwöchentlicher Jour Fixe des Projektteams, ergänzt durch Workshops mit Vertretern des Fachbereiches	Eintägige Workshops mit Unternehmensvertretern alle 1-2 Monate

Tabelle A-2: Übersicht Aktionsforschungsprojekte

Anhang A.3 Workshops und Präsentationen

Anlass	Thema	Ort, Datum
1. CC CDQ-Workshop	Präsentation: Innovative Approaches for Information Modelling – The Example of the UN/CEFACT CCTS	Darmstadt, 01. Februar 2007
2. CC CDQ-Workshop	Präsentation: How to Achieve a Common Understanding of Enterprise Information Objects? – Objectives, Best Practices and Possible Approaches Gruppendiskussion: Vorstellung des BDD-Prototypen der ETA S.A.	Esslingen, 25. April 2007
5. CC CDQ-Workshop	Präsentation: Method for Establishing Transparency on Information Objects (METIO) Gruppenarbeit: Tools and Methodologies for Metadata Management (Anforderungen an ein BDD)	Berlin, 16. November 2007
6. CC CDQ-Workshop	Präsentation und Gruppenarbeit: Corporate Information Architecture	St. Gallen, 16. Januar 2008
Meeting Deutsche Telekom	Präsentation und Gruppendiskussion: Method for Establishing Transparency on Information Objects	Darmstadt, 23. Januar 2008
IIR Data Management Kongress 2008	Präsentation: Corporate Information Architecture	Köln, 25. Februar 2008
8. CC CDQ-Workshop	Präsentation: Information Architecture for Corporate Data Quality	Hamburg, 12. Juni 2008
9. CC CDQ-Workshop	Präsentation: Information Architecture for Corporate Data Quality Gruppenarbeit: Criteria for the Evaluation of a Corporate Information Architecture	Biel/Bienne, 03. September 2008
DW 2008	Präsentation: Harmonizing Company-Wide Information Objects	St. Gallen, 27. Oktober 2008
International Conference on Information Quality 2008	Präsentation: A Method for the Identification and Definition of Information Objects	Cambridge, 15. November 2008
IIR Data Management Kongress 2009	Präsentation: Stammdatenarchitektur und Lebenszyklusmanagement von Stammdaten Gruppenarbeit: Strukturierte Analyse von Stammdatenklassen	Köln, 09. Februar 2009
2. CC CDQ2-Workshop	Präsentation: Corporate Information Architecture – Business Drivers and Fundamental Concepts	St. Gallen, 17. Februar 2009
Meeting Credit Suisse	Präsentation: Company-Wide Integration of Master Data	Zürich, 08. Februar 2010

Tabelle A-3: Workshops und Präsentationen mit Bezug zur Arbeit

Anhang B Dokumentation der Methode zur Stammdatenintegration

Anhang B.1 Definition der Metaentitätstypen der Methode

Dieser Anhang beschreibt sämtliche Metaentitätstypen der Methode zur Stammdatenintegration aus Kapitel 5.3 (siehe Abbildung 5-4) in alphabetischer Reihenfolge. Die Definitionen der Metaentitätstypen, welche zusätzlich mit dem Bezeichnung „(BE)“ versehen sind, wurden direkt dem Metamodell des Methodenkerns des BE entnommen (vgl.[Österle et al. 2007]).

Metaentitätstyp	Beschreibung und Beziehungen
Aktivität (BE)	<p>Aktivitäten sind in sich geschlossene Verrichtungseinheiten im Arbeitsablauf. Sie fassen Arbeitsschritte zusammen, die der Benutzer im gleichen fachlichen und zeitlichen Zusammenhang anwendet und in einem Zug ausführt (vgl. [Morschheuser et al. 1996]). Ein Arbeitsschritt ist eine elementare, nicht weiter zerlegbare Verrichtungseinheit. Aktivitäten können vollständig automatisch (automatisierte Aktivität) oder im Dialog mit dem Benutzer (manuelle Aktivität) ablaufen [Vogler 2004, 40f.]</p> <p>Beziehungen: Eine Aktivität ist Bestandteil einer Aufgabe. Eine Aktivität erzeugt bzw. konsumiert mehrere Geschäftsobjekte.</p>
Applikation (BE)	<p>Eine Applikation ist Software, die Funktionen und Daten zur Unterstützung betrieblicher Aufgaben zur Verfügung stellt [Mertens 2001, 37; Vogler 2004, 41].</p> <p>Beziehungen: Eine Applikation ist mit einer oder mehreren anderen Applikationen über einen Datenfluss verbunden. Eine Applikation besteht u. a. aus einer oder mehreren Applikationsfunktionen.</p>
Applikationsfunktion (BE)	<p>Eine Applikationsfunktion implementiert die fachliche Logik der Applikation in Form von Programmen bzw. Operationen [Heutschi 2007, 14] und ist Bestandteil einer Applikation [Schwinn 2005, 34].</p> <p>Beziehungen: Eine Applikationsfunktion ist Bestandteil einer Applikation. Eine Applikationskomponente operiert auf mehreren Datenelementen.</p>
Aufgabe (BE)	<p>Eine Aufgabe formuliert ein klar definiertes Handlungsziel einer betrieblichen Tätigkeit, das von einem Aufgabenträger zu realisieren ist [Bleicher 1991, 35]. Aufgabenträger kann ein Mensch und/oder eine Maschine sein [Österle 1995, 50; Rüegg-Stürm 2002, 67].</p> <p>Beziehungen: Eine Aufgabe ist Bestandteil eines Geschäftsprozesses. Eine Aufgabe besteht selber aus mehreren Aktivitäten. Eine Aufgabe kann von mehreren Rollen ausgeführt werden.</p>
Business Data Dictionary	<p>Ein Business Data Dictionary (BDD) bezeichnen fachliche Metadatenkataloge [Schemm 2007, 4; Schmidt/Otto 2008, 215ff.]. BDDs unterstützen das Management von Metadaten, indem sie fachliche Beschreibungen (Metadaten) der Kernentitäten eines Unternehmens speichern, insbesondere von Stammdatenobjekten, mit dem Ziel, ein einheitliches Verständnis im Unternehmen zu etablieren.</p> <p>Beziehungen: Ein BDD ist eine spezielle Applikation (Spezialisierung). Das BDD speichert (primär fachliche) Metadaten. Ein BDD ist ein wichtiges Werkzeug für das Metadatenmanagement und Grundlage für die Ableitung eines semantischen Datenmodells.</p>

Metaentitätstyp	Beschreibung und Beziehungen
Datenelement (BE)	<p>Datenelemente repräsentieren Geschäftsobjekte auf Ebene der Informationssysteme [Martin 1977, 50; Brenner 1995, 177; Höning 2009, 111]. Sie sind in einer Datenbasis persistent gespeicherte Ausprägungen eines Datenelementtyps bzw. Datenobjektattributs [Jung 2006, 17]. Unter einem Datenelement ist jener Teil der Datenstruktur zu verstehen, der sich in einem gegebenen oder unterstellten Zusammenhang logisch nicht mehr sinnvoll unterteilen lässt [Stahlknecht/Hasenkamp 2002, 139]. Im relationalen Datenmodell entspricht das Datenelement einem Attribut [Mayr et al. 1987, 518].</p> <p>Beziehungen: Ein Datenelement ist Bestandteil eines Datenobjektes. Ein Datenelement ist Instanz genau eines Datenobjektattributs. Datenelemente sind Bestandteil eines Datenflusses. Auf Datenelementen operieren eine oder mehrere Applikationsfunktion.</p>
Datenfluss	<p>Ein Datenfluss bezeichnet den elektronischen Austausch von Datenelementen zwischen zwei Applikationen.</p> <p>Beziehungen: Ein Datenfluss verbindet zwei Applikationen miteinander. Ein Datenfluss besteht aus mehreren Datenelementen, die zwischen zwei Applikationen ausgetauscht werden.</p>
Datenmodell	<p>Ein Datenmodell stellt Datenobjekttypen und ihre Beziehungen zueinander ganzheitlich dar [Krcmar 2005, 112]. Datenmodelle bilden die Grundlage für eine Datenintegration auf konzeptioneller Ebene sowie die Verbesserung von Datenqualität und die Verringerung von Datenredundanz [Schüngel 1995, 1; Shanks/Darke 1999, 20].</p> <p>Beziehungen: Ein Datenmodell besteht aus mehreren Datenobjekttypen (und deren Beziehungen zueinander).</p>
Datenobjekt	<p>Ein Datenobjekt ist eine Menge von mehreren Datenelementen, die zusammen die Ausprägung eines entsprechenden Datenobjekttyps bilden [Jung 2006, 17].</p> <p>Beziehungen: Ein Datenobjekt ist Instanz eines Datenobjekttyps. Ein Datenobjekt repräsentiert ein Geschäftsobjekt auf Systemebene. Ein Datenobjekt setzt sich aus mehreren Datenelementen zusammen.</p>
Datenobjektattribut, atomar	<p>Datenelementtypen bzw. Datenobjektattribute entsprechen Attributen in relationalen Datenmodellen [Martin 1977, 49f.]. Atomaren Datenobjektattributen wird bei der Instanziierung ein atomarer Wert zugeordnet [UN/CEFACT 2009b, 51-76; Vogel 2009, 106].</p> <p>Beziehungen: Ein atomares Datenobjektattribut ist Bestandteil eines komplexen Datenobjektattributs sowie eines Datenobjekttyps. Ein atomares Datenobjektattribut wird durch einen Datenobjektattributtyp typisiert. Aus einem atomaren Datenobjektattribut können mehrere Datenelemente als Instanzen erzeugt werden.</p>
Datenobjektattribut, komplex	<p>Datenelementtypen bzw. Datenobjektattribute entsprechen Attributen in relationalen Datenmodellen. Komplexe Datenobjektattribute sind Felder, die auf einen anderen Datenobjekttyp verweisen und diese somit hierarchisch einbinden [UN/CEFACT 2009b, 51-76; Vogel 2009, 106].</p> <p>Beziehungen: Komplexe Datenobjektattribute sind Bestandteil von Datenobjekttypen. Komplexe Datenobjektattribute setzen sich aus mehreren atomaren Datenobjektattributen zusammen.</p>
Datenobjektattributtyp	<p>Datenobjektattributtypen sind fachlichen Attributtypen, welche mögliche Wertzuweisungen von Datenobjektattributen einschränken und diese dadurch näher spezifizieren.</p> <p>Beziehungen: Ein Datenobjektattributtyp typisiert mehrere atomare Datenobjektattribute.</p>

Metaentitätstyp	Beschreibung und Beziehungen
Datenobjektbeziehung	<p>Eine Datenobjektbeziehung bezeichnet die Beziehung zwischen zwei Datenobjekten.</p> <p>Beziehungen: Eine Datenobjektbeziehung ist die Ausprägung genau einer Datenobjekttypbeziehung.</p>
Datenobjekttyp	<p>Der Datenobjekttyp beschreibt ein Datenobjekt z. B. anhand seines Wertebereiches und seiner Klassifizierung [Ferstl/Sinz 2001, 291]. Er vereinigt eine Menge von zusammengehörigen Datenobjektattributen, mit deren Hilfe Datenobjekte (Instanzebene) abstrakt (z. B. in einem Modell) beschrieben werden können [Jung 2006, 17].</p> <p>Beziehungen: Ein Datenobjekttyp besteht aus mehreren atomaren und komplexen Datenobjektattributen. Aus einem Datenobjekttypen können mehrere Datenobjekte instanziiert werden. Ein Datenobjekttyp realisiert genau einen Geschäftsobjekttyp durch Auswahl bestimmter Objekteigenschaften und deren Abbildung auf Datenobjektattribute. Datenobjekttypen sind Bestandteile eines Datenmodells.</p>
Datenobjekttypbeziehung	<p>Eine Datenobjekttypbeziehung modelliert die Beziehung zwischen zwei Datenobjekttypen in einem Datenmodell. Sie kann mit Multiplizitäten und Typisierungsangaben versehen werden.</p> <p>Beziehungen: Eine Datenobjekttypbeziehung wird durch eine oder mehrere Datenobjektbeziehung instanziiert.</p>
Geschäftsobjekt (BE)	<p>Ein Geschäftsobjekt ist ein realer oder gedachter Gegenstand (eine Instanz) der Leistungserstellung, z.B. Geschäftspartner, Anlage, Material, Auftrag, Vertrag [Österle 1995, 87].</p> <p>Beziehungen: Ein Geschäftsobjekt ist die Instanz eines Geschäftsobjekttypen. Ein Geschäftsobjekt wird auf Systemebene durch ein oder mehrere Datenobjekte repräsentiert. Ein Geschäftsobjekt wird durch eine oder mehrere Aktivitäten erzeugt und konsumiert.</p>
Geschäftsobjektattribut	<p>Geschäftsobjektattribute kennzeichnen gemeinsame Eigenschaften von Geschäftsobjekttypen. Sie werden durch fachliche Attributtypen und Multiplizitäten charakterisiert [Vogel 2009, 69].</p> <p>Beziehungen: Ein Geschäftsobjektattribut ist Bestandteil eines Geschäftsobjekttyps.</p>
Geschäftsobjektbeziehung	<p>Eine Geschäftsobjektbeziehung bezeichnet die Beziehung zwischen zwei Geschäftsobjekten.</p> <p>Beziehungen: Eine Geschäftsobjektbeziehung ist die Ausprägung genau einer Geschäftsobjekttypbeziehung.</p>
Geschäftsobjekttyp	<p>Geschäftsobjekttypen vereinigen eine Menge von zusammengehörigen Geschäftsobjektattributen, mit deren Hilfe Geschäftsobjekte der Realwelt (Instanzebene) abstrakt (z. B. in einem Modell) beschrieben werden können [Jung 2006, 17].</p> <p>Beziehungen: Ein Geschäftsobjekttyp ist eine Sammlung von Metadaten zu einem Geschäftsobjekt (Instanz). Aus einem Geschäftsobjekttyp können mehrere Geschäftsobjekte instanziiert werden. Ein Geschäftsobjekttyp besteht aus mehreren Geschäftsobjektattributen. Ein Geschäftsobjekttyp wird durch genau einen Datenobjekttypen realisiert.</p>
Geschäftsobjekttypbeziehung	<p>Eine Geschäftsobjekttypbeziehung beschreibt die Beziehung zwischen zwei Geschäftsobjekttypen. Sie kann mit Multiplizitäten und Typisierungsangaben versehen werden [Vogel 2009, 69].</p> <p>Beziehungen: Eine Geschäftsobjekttypbeziehung wird durch eine oder mehrere Geschäftsobjektbeziehung instanziiert.</p>

Metaentitätstyp	Beschreibung und Beziehungen
Geschäftsprozess (BE)	<p>Ein Geschäftsprozess fasst eine Menge von Aufgaben, die in einer vorgegebenen zeitlichen und sachlogischen Ablauffolge zu erledigen sind, und ggf. durch Applikationen unterstützt werden, als Einheit zusammen mit dem Ziel, eine klar definierte Prozessleistung zur Erzeugung von Kundennutzen zu erstellen [Hess 1996, 124ff.; Vogler 2004, 96]. Ein Geschäftsprozess tauscht mit unternehmensinternen oder -externen Kunden (Prozess-)Leistungen in bestimmtem Umfang und in bestimmter Qualität aus (als Produzent und Konsument). Er transformiert bestimmte Einsatzgüter (Input) unter Beachtung bestimmter Regeln und unter Verwendung verschiedener Ressourcen und Informationen zu Arbeitsergebnissen (Output).</p> <p>Beziehungen: Ein Geschäftsprozess besteht aus mehreren Aufgaben und kann sich aus mehreren Teilprozessen zusammensetzen.</p>
Metadaten	<p>Metadaten bezeichnen Daten, welche die Bedeutung sowie fachliche und technische Eigenschaften anderer Daten beschreiben [Burnett et al. 1999, 184-194; Tozer 1999, xix]. Metadaten werden im Rahmen der Arbeit genutzt, um Geschäfts- und Datenobjekte eines Informationssystems standardisiert zu definieren und zu beschreiben. Sie stellen somit ein wichtiges Werkzeug für die Integration von Informationssystemen sowie das Stammdatenmanagement dar.</p> <p>Beziehungen: Metadaten werden in BDDs gespeichert. Geschäftsobjekttypen und Datenobjekttypen enthalten Metadaten zu Geschäfts- und Datenobjekten. Metadaten werden durch ein oder mehrere Metadatenmodell beschrieben. Metadaten werden durch einen oder mehrere Metadatenmanagementprozess verwaltet bzw. gesteuert.</p>
Metadatenmanagement	<p>Das Metadatenmanagement hat die Aufgabe, Metadaten für alle relevanten Nutzer und über den gesamten Lebenszyklus der Entitäten zur Verfügung zu stellen [Dippold et al. 2005, 97]. Durch die Umsetzung eines zentralen Metadatenmanagements erreichen Unternehmen einen höheren Reifegrad des qualitätsorientierten Stammdatenmanagements [Ryu et al. 2006, 194f.; Loshin 2008, 56ff.].</p> <p>Beziehungen: Das Metadatenmanagement nutzt Werkzeuge, u. a. ein BDD.</p>
Metadatenmanagementprozess	<p>Ein Metadatenmanagementprozess ist ein Prozess, der einen integrierten Umgang mit Metadaten einer oder mehrerer Kategorien gewährleistet [Melchert 2006, 146]. Er umfasst Aktivitäten der Entstehung, Verwaltung bzw. Pflege, Bereitstellung und Nutzung von Metadaten.</p> <p>Beziehungen: Ein Metadatenmanagementprozess ist ein Geschäftsprozess. Ein Metadatenmanagementprozess steuert bzw. verwaltet Metadaten.</p>
Metadatenmodell	<p>Ein Metadatenmodell ist ein Modell der Metadaten, die in einem bestimmten Gegenstandsbereich verwaltet werden [Melchert 2006, 144]. Als spezielles Datenmodell beschreibt es eine bestimmte Art von Daten – nämlich Metadaten –, die in einem speziellen Datenbestand – nämlich einem Metadatenbestand – verwaltet werden⁵⁵.</p> <p>Beziehungen: Ein Metadatenmodell ist ein spezielles Datenmodell. Ein Metadatenmodell ist das konzeptionelle Datenmodell, das einem BDD zugrunde liegt. Ein Metadatenmodell beschreibt (mehrere) Metadaten.</p>
Rolle (BE)	<p>Rollen sind Klassifikationen von Mitarbeitern, die sich aus den Fähigkeiten, den Aufgaben und dem Verantwortungsbereich ableiten lassen [Bartölke/Grieger 2004, 468; Puschmann 2004, 53].</p> <p>Beziehungen: Eine Rolle kann mehrere Aufgaben ausführen.</p>

⁵⁵ Die Vorsilbe Meta bezieht sich in diesem Zusammenhang nicht auf den Begriff des Datenmodells, sondern lediglich auf den Begriff der Daten, sodass es sich bei einem Metadatenmodell keineswegs um ein Metamodell handelt.

Metaentitätstyp	Beschreibung und Beziehungen
Semantisches Datenmodell	<p>Ein semantisches Datenmodell ist eine implementierungsunabhängige und semantisch eindeutige Beschreibung von Entitäten aus einer fachlichen Sicht und somit die gemeinsame Grundlage für die Kommunikation zwischen Vertretern verschiedener Fachbereiche sowie für die systemseitige Implementierung von Datenstrukturen und Datenverteilung [Gemünden/Schmitt 1991, 25; Wand/Weber 2002, 363; Shanks et al. 2003, 85]. Der Schwerpunkt liegt auf der Datenstandardisierung von Datenobjekttypen und Datenobjektattributen über mehrere Applikationen hinweg [Ortner 1991, 319].</p> <p>Beziehungen: Ein semantisches Datenmodell wird aus einem BDD abgeleitet. Ein semantisches Datenmodell konsolidiert mehrere, applikationsspezifische Datenmodelle.</p>
Stammdatenobjekt	<p>Ein Stammdatenobjekt ist ein spezielles Datenobjekt, dessen Attributwerte (Datenelemente) sich im Vergleich zu anderen Datenobjekten (z. B. Bewegungsdatenobjekten) seltener ändern [Mertens 2007, 20; Loshin 2008, 8].</p> <p>Beziehungen: Ein Stammdatenobjekt ist eine Spezialisierung eines Datenobjektes.</p>

Tabelle B-1: Metaentitätstypen und ihre Definition

Anhang B.2 Funktionendiagramm

Tabelle B-2 fasst sämtliche Aktivitäten der Methode zur Stammdatenintegration zusammen und ordnet ihnen Rollen bzw. Organisationseinheiten zu. Das Funktionendiagramm zeigt, welche Rolle für eine Aktivität verantwortlich ist, welche Rollen die Aktivität ausführen bzw. bei der Ausführung unterstützen und wer – falls nötig – das Ergebnis einer Aktivität genehmigen muss.

- a ausführend
- g genehmigend
- u unterstützend
- v verantwortlich

Aktivitäten		Strategischer Datensteward	Fachlicher Datenstewards	Technische Datenstewards	Konzern-Datensteward	Datenqualitäts-Komitee	Prozess-Owner	System-Owner, Datenbankadmin.	IS-Architekt	Datenarchitekt	Metadaten-administrator
MF1: I.1	Gestaltungsbereich definieren	a, v				g					
MF1: I.2	Relevante Stakeholder und Prozesse identifizieren		u		a, v		u				
MF1: II.1	Geschäftsprozesse auswerten		a		v		u				
MF1: II.2	Geschäftsobjekttypen konsolidieren	v			a	g					
MF1: III.1	Metadatenmodell des BDD spezifizieren	v	a		a	g				u	
MF1: III.2	BDD implementieren	v	u		u						
MF1: III.3	Geschäftsobjekttypen definieren	a, v	a		a	g	a				a
MF1: IV.1	Metadatenmanagementprozesse entwerfen	v			a	g					
MF1: IV.2	Metadatenmanagementprozesse implementieren	v			a	g					a
MF2: I.1	Relevante Stakeholder und Systeme identifizieren			u	a, v		u	u			
MF2: II.1	Systemanalyse durchführen			a	v			u			
MF2: II.2a	Datenobjektattribute spezifizieren			a	v			u		a	
MF2: II.2b	Stammdatenlandkarte erstellen			a	v			u		a	
MF2: II.3	Daten-/Geschäftsobjekttypen konsolidieren		u	u	a, v						
MF2: II.4	BDD aktualisieren		v	u							a
MF2: III.1	Semantisches Stammdatenmodell ableiten	v	u	u	u					a	
MF3: I.1	Stammdatenflüsse und Schnittstellen modellieren	v		u	a			u	u	a	
MF3: I.2	Priorität der Applikationen festlegen	a, v	u	u				u	a, v		
MF3: II.1	Architekturmuster evaluieren	v			a				v		
MF3: II.2	Stammdatenintegrationsarchitektur definieren	v	u	u	a			u	v	a	
MF3: III.1	Stammdatenintegrationsarchitektur implementieren	v			a			a	v	a	
MF3: III.2	Stammdatenmanagementprozesse spezifizieren	v	a	a	a		u	u			

Tabelle B-2: Übersicht über die Zuordnung der Rollen zu Aktivitäten der Methode

Anhang B.3 Bezug von Entwurfsergebnissen zu Metaentitätstypen

Tabelle B-3 listet sämtliche Entwurfsergebnisse der Methode gemäss Dokumentationsmodell auf (siehe Abbildung 5-17 in Kapitel 5.8) und ordnet ihnen die Metaentitätstypen des Metamodells (siehe Abbildung 5-4 in Kapitel 5.3) zu, auf die sich das jeweilige Ergebnis bezieht.

Entwurfsergebnisse	Metaentitätstypen
Abgegrenzter Gestaltungsbereich	Geschäftsprozess, Aufgabe
Liste Ansprechpartner	Rolle
Geschäftsprozessarchitektur	Geschäftsprozess, Aufgabe, Aktivität, Geschäftsobjekt, Geschäftsobjektbeziehung
Liste zu definierender Geschäftsobjekttypen	Geschäftsobjekttyp
Metadatenmodell BDD	Metadatenmodell, Metadaten
Anforderungsspezifikation BDD	BDD
BDD	Business Data Dictionary, Metadaten
Definierte Geschäftsobjekttypen (ausgefüllter Interviewleitfaden)	Metadaten, Geschäftsobjekttyp, Geschäftsobjektattribut, Geschäftsobjekttypbeziehung
Metadatenmanagementprozesse	Metadatenmanagement, Metadatenmanagementprozess, Aufgabe, Aktivität, Rolle
Applikationsverzeichnis	Applikation
Applikationsbeschreibung	Applikation, Applikationsfunktion
Stammdatenlandkarte	Applikation, Datenobjekttyp, Datenobjektattribut
Liste Datenobjektattribute	Datenobjektattribut
Definierte Datenobjektattribute	Datenobjekttyp, Datenobjektattribut, Datenobjektattributtyp, Datenobjekttypbeziehung
Semantisches Stammdatenmodell	Semantisches Datenmodell
Stammdatenflussdiagramm	Applikation, Datenfluss, Datenobjektattribut
Liste priorisierter Applikationen	Applikation
Architekturmuster	Applikation, Datenfluss
Integrationskontextdiagramm	Applikation, Datenobjekttyp
Stammdatenintegrationsarchitektur	Applikation, Datenobjekt bzw. Stammdatenobjekt, Datenelement, Datenobjektbeziehung, Datenfluss
Stammdatenmanagementprozesse	Geschäftsprozess, Aufgabe, Aktivität, Rolle

Tabelle B-3: Entwurfsergebnisse der Methode und ihr Bezug zu den Metaentitätstypen

Anhang C Referenz-Metadatenmodell für ein BDD

Anhang C.1 Ableitung der Attribute des Referenz-Metadatenmodells

Im Folgenden werden die Metadatenattribute aus den verschiedenen Quellen – Metadatenstandards, Publikation sowie Fallstudien und Experteninterviews (siehe Kapitel 5.5.2) – zusammenfassend dargestellt. Hierzu zeigt Tabelle C-1, welchen Quellen die jeweiligen Metadatenattribute entnommen wurden. Die Tabelle ist die Grundlage für die Modellierung des Referenz-Metadatenmodells des BDD (siehe Abbildung 5-9).

Metadatenattribut	Metadatenstandards	Publikationen	Fallstudien/Interviews
Geschäftsobjekttyp			
Name	X	X	X
Verbindliche Definition	X	X	X
Alternative Definitionen		X	X
Abkürzung		X	
Schlagwort	X		
Beispielinstanz	X	X	X
Gültigkeit		X	X
Vertraulichkeitsstufe		X	
Status		X	
Letzte Änderung	X	X	X
Regulatorische Vorschrift		X	X
Kommentar	X		X
Geschäftsobjektattribut			
Name	X	X	X
Definition	X	X	X
Identifizierendes Attribut		X	X
Verbindlichkeit	X	X	X
Gültigkeit		X	X
Datenobjektattribut			
Name	X	X	X
Definition	X	X	X
Multiplizität	X	X	X
Datentyp	X	X	
Ursprungsapplikation	X	X	
Geschäftsobjektbeziehung			
Name	X	X	X
Art der Relation	X	X	X
Multiplizität		X	X
Verwendungskontext			
Geschäftsprozess		X	X
Rolle / Owner	X	X	X

Metadatenattribut	Metadatenstandards	Publikationen	Fallstudien/Interviews
Applikation (führend, nutzend)		X	X
Geschäftsregel			
Name		X	X
Beschreibung		X	X

Tabelle C-1: Metadatenattribute mit Zuordnung zu ihren Quellen

Anhang C.2 Klassen und Attribute des Referenz-Metadatenmodells

Tabelle C-2 beschreibt die Klassen (fett hervorgehoben) und Attribute des in Kapitel 5.5.2 dargestellten Referenz-Metadatenmodell als Grundlage für die Implementierung eines BDD, in dem Geschäftsobjekttypen eines Unternehmens eindeutig definiert werden. Bei der Definition der Attribute jeder Klasse wird jeweils angegeben, ob diese zwingend sind oder lediglich optional.

Klasse / Attribut	Beschreibung
Geschäftsobjekttyp	<p>Ausser den beschriebenen Attributen verfügt jeder im BDD gepflegte Geschäftsobjekttyp über:</p> <ul style="list-style-type: none"> • (mindestens) eine oder mehrere Definitionen (Klasse: Definition) • eine oder mehrere Beziehungen zu anderen Geschäftsobjekttypen (Klasse: Geschäftsobjekttyprelation) • ein oder mehrere Geschäftsobjektattribute (Klasse: Geschäftsobjektattribut) • eine oder mehrere Geschäftsregeln (Klasse: Geschäftsregel; optional) • eine oder mehrere regulatorische Vorschriften (Klasse: Regulatorische Vorschrift) • einen Verwendungskontext (Klasse: Verwendungskontext)
Name	Eindeutiger Bezeichner des Geschäftsobjekttyps, um den Geschäftsobjekttyp von einem anderen unterscheiden zu können (zwingendes Attribut)
Abkürzung	Angabe eines im Unternehmen oder in der Branche üblichen Akronyms für den Geschäftsobjekttyp (optionales Attribut)
Schlagwörter	Begriffe, die im Zusammenhang mit dem Geschäftsobjekttypen häufig verwendet werden, und anhand derer die Suche im BDD vereinfacht möglich ist (optionales Attribut) Aufzählung, verschiedene Begriffe sind durch Kommata getrennt
Beispielinstanzen	Beispiele für einen Geschäftsobjekttyp, um Nutzer des BDD das Verständnis zu erleichtern (zwingendes Attribut) Aufzählung, verschiedene Begriffe sind durch Kommata getrennt
Gültigkeit	Definiert die Reichweite innerhalb des Unternehmens bzw. des Konzerns, in dem der Geschäftsobjekttyp verwendet wird (zwingendes Attribut) Mögliche Kategorien: Konzern, Unternehmen, Geschäftsbereich, Projekt Die Angabe des Gültigkeitsbereiches kann im Fall der Kategorisierung Unternehmen, Geschäftsbereich oder Projekt durch eine Aufzählung der betroffenen Organisationseinheiten ergänzt werden.
Vertraulichkeitsstufe	Kategorisierung des Geschäftsobjekttyps gemäss den im Unternehmen gebräuchlichen Sicherheitsstandards für Unternehmensdaten (optionales Attribut) Alternativ kann die Kategorisierung auch anhand der Bedeutung des Geschäftsobjekttyps (Prioritätskennzahl) für das Unternehmen durchgeführt (siehe beispielsweise Kapitel 5.5.1.4 bzw. Technik MF1: T.II.4) und somit ein Attribut „Priorität“ definiert werden.
Status	Derzeitiger Status des Geschäftsobjekttyps (zwingendes Attribut) Mögliche Kategorien: in Bearbeitung, Draft (Freigabe ausstehend), Freigegeben, Ersetzt/Veraltet Im Falle des Status „Ersetzt/Veraltet“ ist eine entsprechende Geschäftsobjekttyprelation zwischen zwei Geschäftsobjekttypen zu definieren (siehe Art der Relation).
Letzte Änderung	Datum, an dem fachliche Metadaten zu einem Geschäftsobjekttyp letztmalig verändert wurden (zwingendes Attribut)
Kommentar	Textfeld für weiterführende Bemerkungen zu einem Geschäftsobjekttyp (optionales Attribut) Hier können u. a. Gründe für die Festlegung auf eine bestimmte Definition angegeben werden
Definition	<p>Definitionen sind textuelle Beschreibungen, mit denen die Bedeutung von Geschäftsobjekttypen expliziert wird. Die Pflege mehrerer Definitionen zu einem Geschäftsobjekttyp sollte grundsätzlich möglich sein, jedoch sollte eine als unternehmensweit verbindlich festgelegt werden</p>

Klasse / Attribut	Beschreibung
Verbindlich	Feld, durch das die unternehmensweite Verbindlichkeit einer Definition zu einem Geschäftsobjekttyp festgelegt wird (zwingendes Attribut)
Geschäftsobjekttyprelation	
Jeder Geschäftsobjekttyp steht mit einem oder mehreren anderen Geschäftsobjekttypen in Beziehung. Diese Relationen sind im BDD zwingend abzubilden.	
Name	Name der Relation zwischen zwei Geschäftsobjekttypen (zwingendes Attribut)
Art der Relation	Typ der Beziehung zwischen zwei Geschäftsobjekttypen (zwingendes Attribut) Art der Relation wird als bidirektionale Beziehung aus einer vordefinierten Liste von Beziehungstypen angegeben Mögliche Kategorien: Generalisierung/Spezialisierung, Aggregation (Teil/Ganzes), Synonymbeziehung, Ersetzt/Ersetzt durch
Multiplizität	Mengenmässige Beziehung zwischen zwei Geschäftsobjekttypen (optionales Attribut)
Geschäftsobjektattribut	
Name	Bezeichner des Geschäftsobjektattributs muss innerhalb eines Geschäftsobjekttyps eindeutig sein, um es von anderen unterscheiden zu können (zwingendes Attribut)
Definition	textuelle Beschreibung, mit der die Bedeutung des Geschäftsobjektattributs expliziert wird (zwingendes Attribut) pro Geschäftsobjektattribut sollte es nur eine eindeutige Definition geben
Identifizierendes Attribut	Feld, welches angibt, ob das Geschäftsobjektattribut zur eindeutigen Identifikation des Geschäftsobjekttyps geeignet ist (zwingendes Attribut)
Verbindlichkeit	Feld, welches angibt, ob das Geschäftsobjektattribut verpflichtend für einen Geschäftsobjekttyp werden muss (mandatory), kann (optional) oder abhängig von der Angabe eines anderen Attributs gefüllt werden muss (conditional) (zwingendes Attribut)
Gültigkeit	Feld, welches angibt, ob das Geschäftsobjektattribut in mehreren Fachbereichen bzw. Applikationen (global bzw. unternehmensweit) verwendet wird oder nur lokal von Bedeutung ist (zwingendes Attribut)
Datenobjektattribut	
Name	Bezeichner des Datenobjektattributs muss eindeutig sein, um es von anderen unterscheiden zu können (zwingendes Attribut)
Definition	textuelle Beschreibung, mit der die Bedeutung des Datenobjektattributs expliziert wird (zwingendes Attribut) pro Datenobjektattribut sollte es nur eine eindeutige Definition geben
Multiplizität	Anzahl der Werte, die ein Datenobjektattribut annehmen kann (zwingendes Attribut) Als Syntax für die Angabe von Multiplizitäten kann beispielsweise auf die UML zurückgegriffen werden. Ein Datenobjektattribut mit der Multiplizität [0..1] wird als optional bezeichnet.
Datentyp	Wertebereich des Datenobjekttyps (optionales Attribut)
Feldlänge	Zulässige Länge eines Datenobjektattributes (optionales Attribut)
Ursprungsapplikation	Applikation, in welcher die zugehörigen Instanzen (Datenelemente) initial angelegt wurden (zwingendes Attribut) Die Pflege dieses Attributes ist substanziell, um bei redundanten Datenelementen das glaubwürdigste und für die Integration massgebliche Datenelement identifizieren zu können [Adelman et al. 2005, 37].
Geschäftsregel	
Die Verwendung von Geschäftsobjekttypen kann durch Geschäftsregeln eingeschränkt werden. Sie geben daher zusätzlichen Aufschluss über die Bedeutung der Geschäftsobjekttypen, auf die sie sich beziehen. Geschäftsregeln können sich sowohl auf einen Objekttypen als auch auf einzelne Attribute beziehen.	
Beschreibung	Jede Geschäftsregel wird textuell in einem Satz beschrieben (zwingendes Attribut)
Kommentar	Textfeld für weiterführende Bemerkungen zu einer Geschäftsregel (optionales Attribut)
Regulatorische Vorschrift	
Regulatorische Vorschriften können die Verwendung von Geschäftsobjekten einschränken. Die für einen Geschäftsobjekttyp relevanten regulatorischen Vorschriften sind daher mit ihrem Namen im BDD zu pflegen.	
Name	Bezeichner bzw. Titel der zu beachtenden regulatorischen Vorschrift (zwingendes Attribut)

Klasse / Attribut	Beschreibung
Verwendungskontext	
Der Verwendungskontext beschreibt das Umfeld, in dem der Geschäftsobjekttyp verwendet wird. Er setzt sich aus den drei Klassen Geschäftsprozess, Rolle und Applikation zusammen.	
Geschäftsprozess	
Name	Im Unternehmen gebräuchlicher Name des Geschäftsprozesses, in denen Geschäftsobjekttyp verwendet wird (zwingendes Attribut)
Rolle	
Organisationseinheit	Abteilung bzw. Unternehmensfunktion, welcher die Rolle im Unternehmen zugeordnet ist (zwingendes Attribut)
Owner/Datensteward	Feld, welches angibt, ob die Rolle für den Geschäftsobjekttyp und dessen Pflege (Bearbeitung von Änderungsanträgen) verantwortlich und damit deren Owner ist (zwingendes Attribut)
Zugriffsrecht	Berechtigungen, über welche die jeweilige Rolle für den Geschäftsobjekttyp verfügt (zwingendes Attribut) Mögliche Kategorien: Create, Read, Update, Delete
Applikation	
Name	Im Unternehmen gebräuchlicher Name der Applikation (zwingendes Attribut)
Führend	Feld, welches angibt, ob die Applikation für den Geschäftsobjekttyp führend ist (zwingendes Attribut)
Zugriffsrecht	Berechtigungen, über welche die jeweilige Rolle für den zugehörigen Datenobjekttyp verfügt (zwingendes Attribut) Das Zugriffsrecht der Applikationen wird im BDD aggregiert für einen Geschäftsobjekttypen definiert, obwohl eine Applikation genau genommen auf ein Datenelement zugreift (siehe Metamodell in Abbildung 5-4). Die dazu notwendigen Informationen können der Stammdatenlandkarte (siehe Aktivität MF2: II.2.b). Eine Zuordnung der Zugriffsrechte zu Datenelementen bzw. Datenobjektattributen ist aufgrund ihrer Vielzahl im BDD jedoch nicht mehr sinnvoll abzubilden. Mögliche Kategorien: Create, Read, Update, Delete

Tabelle C-2: Klassen des Referenzmetamodells für ein BDD

Anhang D.2 Vorlage zur Beschreibung von Geschäftsobjekttypen

Interviewpartner	
Name:	Fritz Bütikofer
Rolle:	Leiter Reference Data Service (Normung, Klassifizierung von Stammdaten)
Abteilung:	Service Unit Organisation & Informatik
Telefon:	E-Mail:
Datum:	19. August 2008

1.) Name des Geschäftsobjekttypen:	Uhrwerk		
Synonyme:	MVT, Kaliber, Werk		
Schlagwörter:	Komponente, Baugruppe		
Gängige Abkürzung:	MVT Kalibernummer Oberflächenbehandlung etc. (Bsp.: MVT 2892A2 DO 3HZ H3 QG3H RO COL)		
2.) Für welche Fachbereiche ist der Geschäftsobjekttyp relevant? (Mehrfachnennungen möglich)			
<input type="checkbox"/> Einkauf	<input checked="" type="checkbox"/> Produktion	<input checked="" type="checkbox"/> Vertrieb	<input type="checkbox"/> Personal (Zutreffendes ankreuzen)
<input type="checkbox"/> Finanz	<input type="checkbox"/> Sonstige:		
3.) Wie gross ist der organisatorische Gültigkeitsbereich des zu definierenden Geschäftsobjekttyps?			
<input type="checkbox"/> Projekt	<input type="checkbox"/> Business Unit	<input type="checkbox"/> Unternehmen	<input checked="" type="checkbox"/> Konzern (Zutreffendes ankreuzen)
4.) In welchen Geschäftsprozessen wird der Geschäftsobjekttyp verwendet?			
	Manufacturing		
	Sales and Distribution		
5.) Gibt es für den Geschäftsobjekttyp eine unternehmensübergreifend anerkannte Definition?			
<input type="checkbox"/> Ja	<input checked="" type="checkbox"/> Nein	(Zutreffendes ankreuzen)	
Falls ja, wie lautet die Definition? (Angabe mit Quelle)			
Falls nein, wie würden Sie den Geschäftsobjekttyp beschreiben bzw. definieren?			
Das Uhrwerk bezeichnet ein komplexes Bauelement der Uhr. Innerhalb der ETA werden mechanische Uhrwerke und Quarzwerke unterschieden. Das mechanische Uhrwerk besteht aus Antrieb, Zahnradgetriebe, Anzeigevorrichtung und Gangregelung. Jedes Uhrwerk muss zwingend über eine eindeutige Artikelnummer, eine zugehörige Zeichnung (Konstruktionsplan) und ein Stückliste verfügen.			

Sind Ihnen im Unternehmen gebräuchliche, abweichende Definitionen bekannt? (jeweils mit Bereich)

Produktverständnis des Uhrwerks im Vertrieb (Artikel für Uhrenhersteller)

6.) Durch welche Attribute wird der Geschäftsobjekttyp beschrieben? (Auszug)

Artikelnummer	(Text, 40)	Identifizierendes Attribut:	<input checked="" type="checkbox"/>
Status	(vorgegebene Werte)	Identifizierendes Attribut:	<input type="checkbox"/>
Variante	(Text, 40)	Identifizierendes Attribut:	<input type="checkbox"/>
Änderungen	(Text)	Identifizierendes Attribut:	<input type="checkbox"/>
ID Stückliste	(Text, 20)	Identifizierendes Attribut:	<input type="checkbox"/>
Zeichnung	(CAD-Datei)	Identifizierendes Attribut:	<input type="checkbox"/>

7.) Ausprägungen / Beispiele für den Geschäftsobjekttyp (mit Geschäftsobjektattributen)

8.) Mit welchen anderen Entitäten steht der Geschäftsobjekttyp in Beziehung?

Geschäftsobjekttyp	Art der Beziehung
Quarzwerk	Generalisierung/Spezialisierung
Mechanisches Uhrwerk	Generalisierung/Spezialisierung
Werkplatte	Teile-Ganzes-Beziehung
Uhr	Teile-Ganzes-Beziehung
Artikel, Produkt	Generalisierung/Spezialisierung

9.) Ist Verantwortlichkeit für den Geschäftsobjekttyp definiert?

Ja Nein (Zutreffendes ankreuzen)

Falls ja, wer ist Owner des Geschäftsobjekttyps?

Name:

Abteilung:

Telefon: E-Mail:

Wer ist für die Pflege des Geschäftsobjekttyps verantwortlich?

Name:

Abteilung:

Telefon: E-Mail:

Gibt es einen Verantwortlichen seitens der IT?

Name:

Abteilung:

Telefon: E-Mail:

10.) Welche regulatorischen Bestimmungen sind für den Geschäftsobjekttyp von Bedeutung?

11.) Wird der Geschäftsobjekttyp von einer Applikation in Ihrem Verantwortungsbereich erzeugt?

Ja Nein (Zutreffendes ankreuzen)

Falls ja, wie ist der Name der Applikation?

Falls nein, woher erhalten Ihre Applikationen das Objekt?

12.) In welchen Applikationen wird dieser Geschäftsobjekttyp in ihrem Verantwortungsbereich verwaltet und in welcher Form greifen sie darauf zu?

Name der Applikation: Art des Zugriffs (Create, Read, Update, Delete):

13.) Welche abweichenden Bezeichnungen für den Geschäftsobjekttyp gibt es in den Applikationen?

14.) Geschäftsregeln

- a.) Die Artikelnummer jedes Uhrwerks beginnt mit dem Kürzel MVT.
b.) Jedes Uhrwerk muss über eine Stückliste und eine Zeichnung (Konstruktionsplan) verfügen.

15.) Sonstige Anmerkungen

Anhang D.3 Vorlage für einen BDD Change Request

Antragsteller	
Name:	<input type="text"/>
Rolle:	<input type="text"/>
Abteilung:	<input type="text"/>
Telefon:	<input type="text"/>
E-Mail:	<input type="text"/>

Änderungsantrag		Datum:	<input type="text"/>
Nummer:	<input type="text"/>		
Auslöser für Change Request:	<input type="text"/>		
Name des Geschäftsobjekttyps:	<input type="text"/>		
Name des Geschäftsobjektattributs:	<input type="text"/>		
Beschreibung des Change Requests:	<input type="text"/>		
Begründung des Change Requests:	<input type="text"/>		
Dringlichkeit:	<input type="checkbox"/> Hoch	<input type="checkbox"/> Mittel	<input type="checkbox"/> Tief (Zutreffendes ankreuzen)
Zuständiger Datensteu-ward:	<input type="text"/>		

Ausführung Änderungsantrag		Datum:	<input type="text"/>
Bearbeitet durch:	<input type="text"/>		
Review-Ergebnis:	<input type="checkbox"/> Antrag angenommen	<input type="checkbox"/> Antrag abgelehnt	
Begründung des Review-Ergebnisses:	<input type="text"/>		
Änderungskommentar:	<input type="text"/>		

Anhang E Dokumentation zum Aktionsforschungsprojekt Daimler AG

Anhang E führt ausgewählte Entwurfsergebnisse des Aktionsforschungsprojektes Daimler AG an, auf die in Kapitel 6.1.2 Bezug genommen wurde.

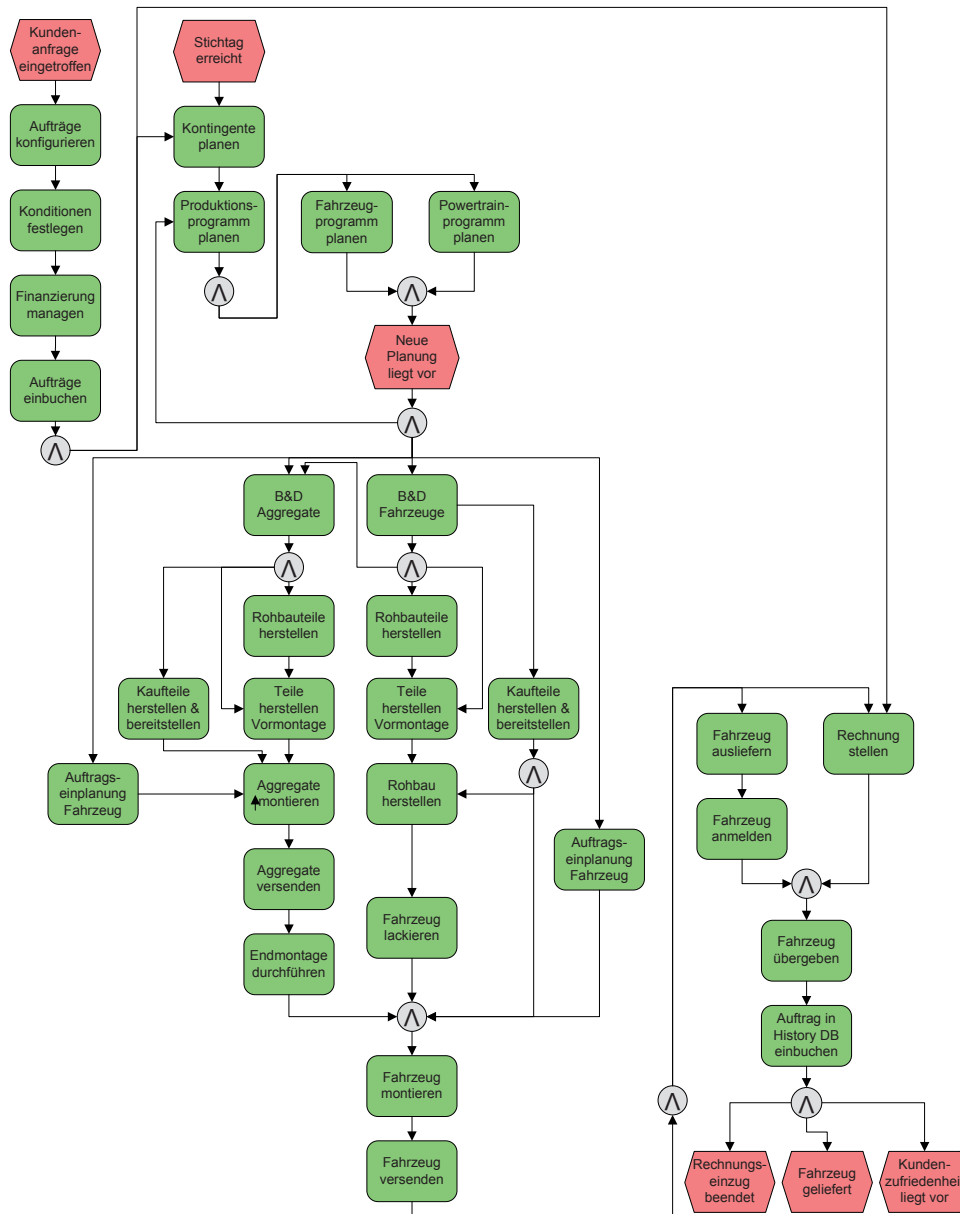


Abbildung E-1: Vereinfachte EPK des Geschäftsprozesse „Kundenauftrag managen“ (Aktionsforschungsprojekt Daimler)

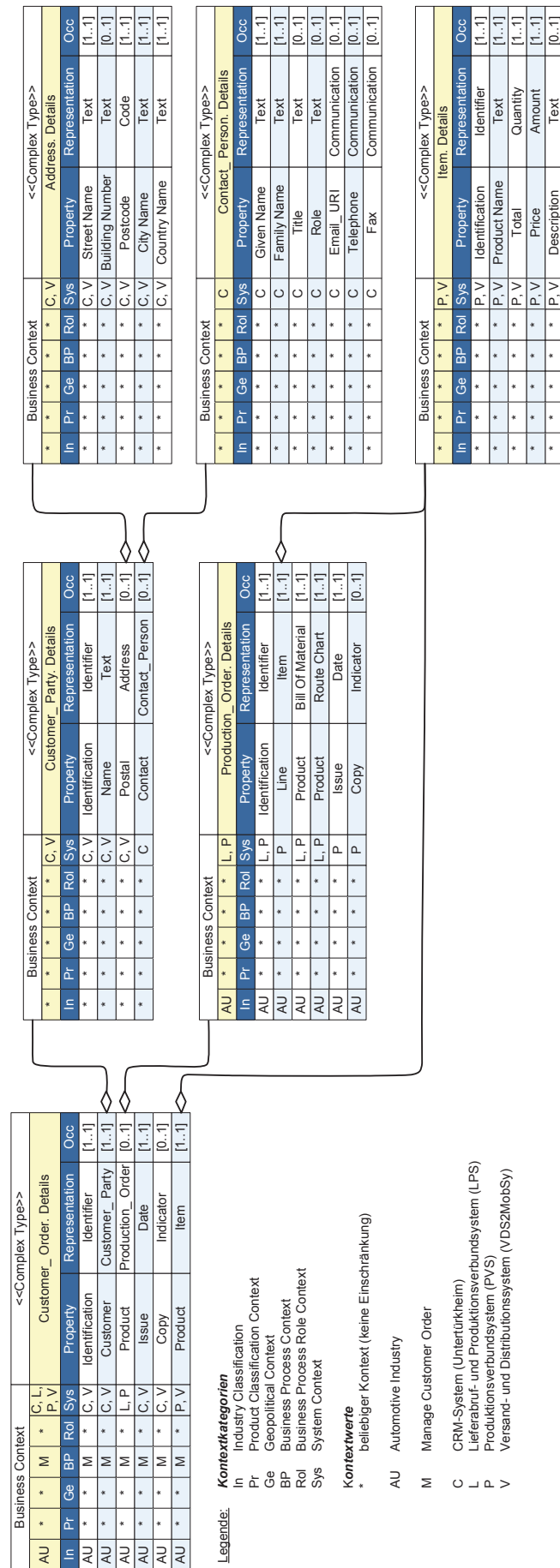


Abbildung E-2: Semantisches Stammdatenmodell (Ausschnitt)

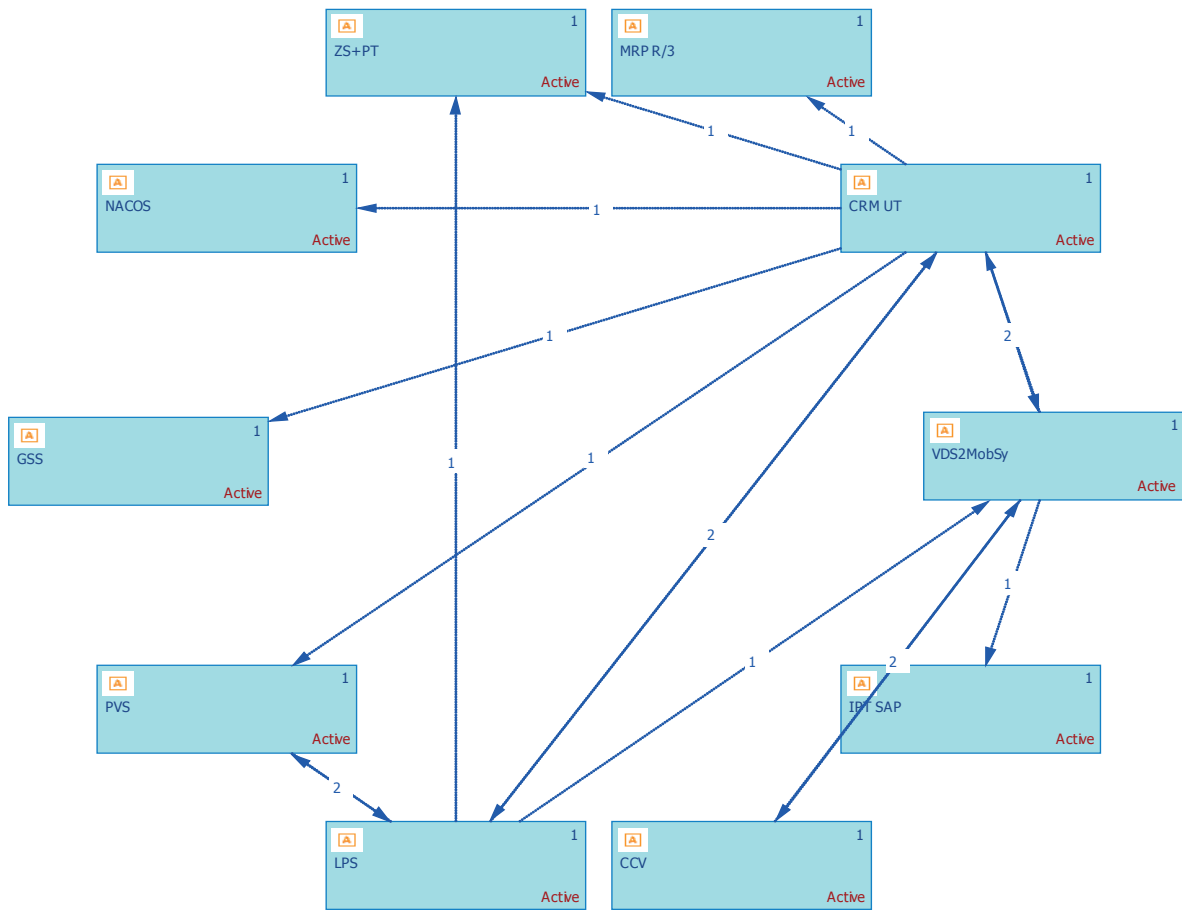


Abbildung E-3: Datenflussdiagramm für den Datenobjekttyp Fahrzeugauftrag (Ausschnitt)

Literaturverzeichnis

[Adelman et al. 2005]

Adelman, S., Moss, L. T., Abai, M., Data Strategy, Addison-Wesley Professional, Boston (MA) 2005

[ADOben 2009]

ADOben, Institut für Wirtschaftsinformatik der Universität St. Gallen (IWI-HSG): ADOben Produktwebseite, <http://adoben.iwi.unisg.ch/de/adoben>, 02.11.2009

[AGLS 2002]

AGLS, AGLS Metadata Element Set – Part1: Reference Description. Version 1.3, Australian Government Locator Service, National Archives of Australia, Canberra 2002

[Aier et al. 2009]

Aier, S., Gleichauf, B., Riege, C., Saat, J., Empirische Validierung von Integrationstypen am Beispiel unternehmensübergreifender Integration, in: Hansen, H.R., Karagiannis, D., Fill, H.-G. (Hrsg.), Business Services: Konzepte, Technologien, Anwendungen; Proceedings der 9. Internationalen Tagung Wirtschaftsinformatik, Band 1, 246, Österreichische Computer Gesellschaft, Wien 2009, S. 99-108

[Aier et al. 2008]

Aier, S., Riege, C., Winter, R., Unternehmensarchitektur - Literaturüberblick und Stand der Praxis, in: Wirtschaftsinformatik, 50, 2008, Nr. 4, S. 292-304

[Alfabet 2009]

Alfabet, planningIT - Produktinformationen, <http://www.alfabet.de/ressourcen/produktinformationen>, 02.11.2009

[Alonso et al. 2003]

Alonso, G., Casati, F., Kuno, H., Machiraju, V., Web Services: Concepts, Architectures and Applications, Springer, Berlin 2003

[Alt et al. 2001]

Alt, R., Cäsar, M. A., Leser, F., Österle, H., Puschmann, T., Reichmayr, C., Architektur für das Business Networking, BE HSG/ CC BN/ 2, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen 2001

[Alt et al. 2004]

Alt, R., Cäsar, M. A., Leser, F., Österle, H., Puschmann, T., Reichmayr, C., Architektur des Echtzeit-Unternehmens, in: Alt, R., Österle, H. (Hrsg.), Real-Time Business: Lösungen, Bausteine und Potentiale des Business Networking, Springer, Berlin 2004, S. 19-52

[Ariyachandra/Watson 2005]

Ariyachandra, T. R., Watson, H. J., Data Warehouse Architectures: Factors in the Selection Decision and the Success of the Architectures, Working Paper, 2005

[Arnett 2005]

Arnett, S., The ECCMA Open Technical Dictionary: A Catalog Standard for eCommerce, http://www.eccma.org/Presentations/Day1/1630_Benson_Arnett.zip, 02.12.2009

- [Auth 2003]
Auth, G., Prozessorientierte Organisation des Metadatenmanagements für Data-Warehouse-Systeme, Dissertation, Universität St. Gallen, Difo-Druck, Bamberg 2003
- [Back et al. 2007]
Back, A., von Krogh, G., Enkel, E., The CC Model as Organizational Design Striving to Combine Relevance and Rigor, in: Systemic Practice and Action Research, 20, 2007, Nr. 1, S. 91-103
- [Balzert 2000]
Balzert, H., Lehrbuch der Software-Technik - Software-Entwicklung, Spektrum Akademischer Verlag, Heidelberg 2000
- [Bartölke/Grieger 2004]
Bartölke, K., Grieger, J., Individuum und Organisation, in: Schreyögg, G., von Werder, A. (Hrsg.), Handwörterbuch Unternehmensführung und Organisation, Poeschel, Stuttgart 2004, S. 464-472
- [Baskerville/Pries-Heje 1999]
Baskerville, R. L., Pries-Heje, J., Grounded action research: a method for understanding IT in practice, in: Accounting Management And Information Technologies, 9, 1999, S. 1-23
- [Baskerville/Wood-Harper 1998]
Baskerville, R. L., Wood-Harper, A. T., Diversity in information systems action research methods, in: European Journal of Information Systems, 7, 1998, S. 90-107
- [Batini et al. 2007]
Batini, C., Barone, D., Mastrella, M., Maurino, A., Ruffini, C., A Framework and a Methodology for Data Quality Assessment and Monitoring, in: Robert, M.A., O'Hare, R., Markus, M.L., Klein, B. (Hrsg.), Proceedings of 12th International Conference on Information Quality, Cambridge (MA) 2007, S. 333-346
- [Batini et al. 2009]
Batini, C., Cappiello, C., Francalanci, C., Maurino, A., Methodologies for Data Quality Assessment and Improvement, in: ACM Computing Survey, 41, 2009, Nr. 3, Article 16, S. 16:1-16:52
- [Batini et al. 1986]
Batini, C., Lenzerini, M., Navathe, S. B., A Comparative Analysis of Methodologies for Database Schema Integration, in: ACM Computer Surveys, 18, 1986, Nr. 4, S. 323-364
- [Batini/Scannapieco 2006]
Batini, C., Scannapieco, M., Data Quality. Concepts, Methodologies and Techniques, Springer, Berlin 2006
- [Baumöl/Meschke 2009]
Baumöl, U., Meschke, M., Das Management von Datenqualität, in: Zeitschrift für Controlling & Management (ZfCM), 53, 2009, Nr. 1, S. 62-65
- [Becker et al. 2002]
Becker, J., Delfmann, P., Knackstedt, R., Kuropka, D., Konfigurative Referenzmodellierung, in: Becker, J., Knackstedt, R. (Hrsg.), Wissensmanagement mit Referenzmodellen, Physica, Heidelberg 2002, S. 25-144

[Becker et al. 2007a]

Becker, J., Janiesch, C., Pfeiffer, D., Context-based Modeling – Conceptualization of a Novel Modeling Approach and Application for the Design of Business Documents, Proceedings of the 11th Pacific Asia Conference on Information Systems, 2007a, S. 143ff

[Becker et al. 2008a]

Becker, J., Janiesch, C., Pöppelbuß, J., Konfiguration kollaborativer Informationsmodelle, in: Bichler, M., Hess, T., Krcmar, H., Lechner, U., Matthes, F.P.A., Wolf, P. (Hrsg.), Proceedings der Multikonferenz Wirtschaftsinformatik 2008, GITO, Berlin 2008a, S. 813-824

[Becker et al. 2007b]

Becker, J., Knackstedt, R., Pfeiffer, D., Janiesch, C., Configurative Method Engineering - On the Applicability of Reference Modeling Mechanisms in Method Engineering, Proceedings of the 13th Americas Conference on Information Systems (AMCIS 2007), Keystone (CO) 2007b, S. 1-12

[Becker et al. 2008b]

Becker, J., Kramer, S., Janiesch, C., Ein Ansatz für das Variantenmanagement elektronischer Geschäftsdokumente, in: Bichler, M., Hess, T., Krcmar, H., Lechner, U., Matthes, F.P.A., Wolf, P. (Hrsg.), Proceedings der Multikonferenz Wirtschaftsinformatik 2008, GITO, Berlin 2008b, S. 837-848

[Benz 2001]

Benz, R., Entwurf überbetrieblicher Prozessnetzwerke, Dissertation, Universität St. Gallen, Bamberg 2001

[Berson/Dubov 2007]

Berson, A., Dubov, L., Master Data Management and Customer Data Integration for a Global Enterprise, McGraw-Hill, New York (NY) 2007

[Biazzo 2000]

Biazzo, S., Approaches to business process analysis: a review, in: Business Process Management, 6, 2000, Nr. 2, S. 99-112

[Birkenbihl 2006]

Birkenbihl, K., Standards für das Semantic Web, in: Pellegrini, T., Blumauer, A. (Hrsg.), Semantic Web - Wege zur vernetzten Wissensgesellschaft, Springer, Berlin/Heidelberg 2006, S. 73-88

[Bizer et al. 2009]

Bizer, C., Health, T., Berners-Lee, T., Linked Data - The Story So Far, in: International Journal on Semantic Web and Information Systems, 5, 2009, Nr. 3, S. 1-22

[Blechar/White 2009]

Blechar, M., White, A., Metadata Management Is Important for MDM's Long-Term Success, Gartner Group, Stamford (CT) 2009

[Bleicher 1991]

Bleicher, K., Organisation: Strategien-Strukturen-Kulturen, 2. Aufl., Gabler, Wiesbaden 1991

- [Boles et al. 2004]
Boles, C., Friebe, J., Luhmann, T., Typische Integrationsszenarien und deren Unterstützung durch Web Services und andere Technologien, in: Hasselbring, W., Reichert, M. (Hrsg.), EAI-Workshop 2004 - Enterprise Application Integration, Tagungsband, Tagungsband des GI-/GMDS-Workshops EAI'04, OFFIS, 2004, S. 57-67
- [Bortz/Döring 2006]
Bortz, J., Döring, N., Forschungsmethoden und Evaluation - für Human - und Sozialwissenschaftler, 4. Aufl., Springer, Berlin 2006
- [Brancheau et al. 1989]
Brancheau, J. C., Schuster, L., March, S. T., Building and Implementing an Information Architecture, in: Data Base, 20, 1989, Nr. 2, S. 9-17
- [Braun 2007]
Braun, C., Modellierung der Unternehmensarchitektur - Weiterentwicklung einer bestehenden Methode und deren Abbildung in einem Meta-Modellierungswerkzeug, Dissertation, Universität St. Gallen, Logos, Berlin 2007
- [Braun 2009]
Braun, K., MDM @ Novartis Pharma AG, Präsentation auf dem Treffen des Verbandes der Chemischen Industrie (VCI), Frankfurt am Main 2009
- [Brenner 1995]
Brenner, C., Techniken und Metamodell des Business Engineering, Dissertation, Universität St. Gallen, St. Gallen 1995
- [Brenner 1985]
Brenner, W., Entwurf betrieblicher Datenelemente - Ein Weg zur Integration von Informationssystemen, Dissertation, Universität St. Gallen, St. Gallen 1985
- [Brinkkemper 1996]
Brinkkemper, S., Method Engineering: Engineering of Information Systems Development Methods and Tools, in: Information and Software Technology, 38, 1996, Nr. 4, S. 275-280
- [Brugger 2005]
Brugger, R., Der IT Business Case, Springer, Berlin 2005
- [Bucher 2009]
Bucher, T., Ausrichtung der Informationslogistik auf operative Prozesse - Entwicklung und Evaluation einer situativen Methode, Dissertation, Universität St. Gallen, St. Gallen 2009
- [Bucher et al. 2007]
Bucher, T., Klesse, M., Kurpjuweit, S., Winter, R., Situational Method Engineering - On the Differentiation of "Context" and "Project Type", in: Ralytė, J., Brinkkemper, S., Henderson-Sellers, B. (Hrsg.), Situational Method Engineering - Fundamentals and Experiences, 244, Springer 2007, S. 33-48
- [Bucher et al. 2008]
Bucher, T., Riege, C., Saat, J., Evaluation in der gestaltungsorientierten Wirtschaftsinformatik - Systematisierung nach Erkenntnisziel und Gestaltungsziel, in: Becker, J., Krcmar, H., Niehaves, B. (Hrsg.), Wissenschaftstheorie und gestaltungsorientierte Wirtschaftsinformatik. Arbeitsbericht Nr. 120 des Instituts für Wirtschaftsinformatik, Universität Münster, 2008, S. 69-86

- [Bughin et al. 2009]
Bughin, J., Chui, M., Miller, A., How companies are benefiting from Web 2.0: McKinsey Global Survey Results, in: McKinsey Quarterly, September 2009, 2009, S. 10-17
- [Bunjes et al. 2002]
Bunjes, B., Friebe, J., Götze, R., Harren, A., Integration von Daten, Anwendungen und Prozessen am Beispiel des Telekommunikationsunternehmens EWE TEL, in: Wirtschaftsinformatik, 44, 2002, Nr. 5, S. 415-423
- [Burnett et al. 1999]
Burnett, K., Ng, K. B., Park, S., A Comparison of the Two Traditions of Metadata Development, in: Journal of the American Society for Information Science, 50, 1999, Nr. 13, S. 1209-1217
- [Buschmann et al. 1996]
Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P., Stal, M., A System of Patterns - Pattern-Oriented Software Architecture, John Wiley & Sons, West Sussex 1996
- [Business Objects 2008]
Business Objects, BusinessObjects Metadata Management XI 3.0 - What's New, Business Objects, San Jose (CA) 2008
- [Busse 2002]
Busse, S., Modellkorrespondenzen für die kontinuierliche Entwicklung mediatorbasierter Informationssysteme, Logos, Berlin 2002
- [Buxmann 1996]
Buxmann, P., Standardisierung betrieblicher Informationssysteme, Deutscher Universitätsverlag, Wiesbaden 1996
- [Caballero et al. 2008]
Caballero, I., Calero, C., Piattini, M., Verbo, E., MMPro: A Methodology based on ISO/IEC 15939 to draw up data quality measurement processes, Proceedings of the Thirteenth International Conference on Information Quality (ICIQ-08), Cambridge (MA) 2008, S. 326-340
- [Cavana et al. 2001]
Cavana, R. Y., Delahaye, B. L., Sekaran, U., Applied Business Research - Qualitative and Quantitative Methods, 3. Aufl., John Wiley & Sons, Milton 2001
- [Chari/Seshadri 2004]
Chari, K., Seshadri, S., Demystifying Integration, in: Communications of the ACM (CACM), 47, 2004, Nr. 7, S. 59-63
- [Checkland/Holwell 1998]
Checkland, P., Holwell, S., Action Research: Its Nature and Validity, in: Systemic Practice and Action Research, 11, 1998, Nr. 1, S. 9-21
- [Chisholm 2000]
Chisholm, M., Managing Reference Data in Enterprise Databases, Morgan Kaufmann, San Francisco (CA) 2000
- [Cleven/Wortmann 2010]
Cleven, A., Wortmann, F., Uncovering four strategies to approach master data management, in: Sprague, R.H. (Hrsg.), Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS 43), IEEE Computer Society, Los Alamitos (CA) 2010, S. 1-10

- [Conrad 1997]
Conrad, S., *Föderierte Datenbanksysteme - Konzepte der Datenintegration*, Springer, Berlin 1997
- [Conrad 2002]
Conrad, S., *Schemaintegration – Integrationskonflikte, Lösungsansätze, aktuelle Herausforderungen*, in: *Informatik - Forschung und Entwicklung*, 17, 2002, Nr. 3, S. 101-111
- [Cudré-Mauroux 2006]
Cudré-Mauroux, P., *Emergent semantics: Rethinking Interoperability for Large Scale Decentralized Information Systems*, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne 2006
- [Curbera et al. 2003]
Curbera, F., Khalaf, R., Mukhi, N., Tai, S., Weerawarana, S., *The next step in Web Services*, in: *Communications of the ACM (CACM)*, 46, 2003, Nr. 10, S. 29-34
- [Daconta et al. 2003]
Daconta, M., Obrst, L., Smith, K., *The Semantic Web*, John Wiley & Sons, Indianapolis (IN) 2003
- [Daimler 2009]
Daimler, *Innovationen für nachhaltige Mobilität - Geschäftsbericht 2008*, http://www.daimler.com/Projects/c2c/channel/documents/1677322_DAI_2008_Geschaeftsbericht.pdf, 05.10.2009
- [DAMA 2008]
DAMA, *The DAMA Dictionary of Data Management*, Technics Publications, Bradley Beach (NJ) 2008
- [DAMA 2009]
DAMA, *The DAMA Guide to the Data Management of Knowledge*, 1. Aufl., Technics Publications, Bradley Beach (NJ) 2009
- [Dan et al. 2007]
Dan, A., Johnson, R., Arsanjani, A., *Information as a Service: Modeling and Realization*, in: *Ieee (Hrsg.), Proceedings of the International Workshop on Systems Development in SOA Environments 2007 (SDSOA'07)*, IEEE Computer Society, Minneapolis (MN) 2007
- [Davenport/Prusak 1998]
Davenport, T. H., Prusak, L., *Working Knowledge - How Organizations Manage What They Know*, Harvard Business School Press, Boston (MA) 1998
- [Davison et al. 2004]
Davison, R. M., Martinsons, M. G., Kock, N., *Principles of canonical action research*, in: *Information Systems Journal*, 14, 2004, Nr. 1, S. 65-86
- [Dayton 2007]
Dayton, M., *Strategic MDM: The Foundation of Enterprise Performance Management*, in: *Cutter IT Journal*, 20, 2007, Nr. 9, S. 13-17
- [DCMI 2008]
DCMI, *Dublin Core Metadata Element Set, Version 1.1*, <http://dublincore.org/documents/dces/>, 10.03.2008

- [DeMarco 1979]
DeMarco, T., Structured Analysis and System Specification, Prentice Hall, Upper Saddle River (NJ) 1979
- [Dessloch et al. 2003]
Dessloch, S., Maier, A., Mattos, N., Wolfson, D., Information Integration – Goals and Challenges, in: Datenbank-Spektrum, 3, 2003, Nr. 6, S. 7-13
- [Deutsche Bahn 2009]
Deutsche Bahn, Deutsche Bahn - Geschäftsbericht 2008, http://www.deutschebahn.com/site/shared/de/dateianhaenge/berichte/geschaeftsbericht_2008_konzern.pdf, 05.10.2009
- [Deutsche Telekom 2009]
Deutsche Telekom, Connected life and work - Geschäftsbericht 2008, http://www.download-telekom.de/dt/StaticPage/62/37/06/090227_dtag_gb_2008.pdf_623706.pdf, 29.10.2009
- [Devlin 1997]
Devlin, B. A., Data Warehouse: from Architecture to Implementation, Addison Wesley, Reading et al. 1997
- [Devlin/Murphy 1988]
Devlin, B. A., Murphy, P. T., An architecture for a business and information system, in: IBM Systems Journal, 27, 1988, Nr. 1, S. 22
- [Dippold et al. 2005]
Dippold, R., Meier, A., Schnider, W., Unternehmensweites Datenmanagement, Vieweg, Wiesbaden 2005
- [Dorloff et al. 2001]
Dorloff, F.-D., Leukel, J., Schmitz, V., Standards für den Austausch von elektronischen Produktkatalogen, in: WISU - Das Wirtschaftsstudium, 30, 2001, Nr. 11, S. 1528-1536
- [Dreibelbis et al. 2008]
Dreibelbis, A., Hechler, E., Milman, I., Oberhofer, M., van Run, P., Wolfson, D., Enterprise Master Data Management: An SOA Approach to Managing Core Information, Pearson Education, Boston (MA) 2008
- [Dyché/Levy 2006]
Dyché, J., Levy, E., Customer Data Integration, John Wiley & Sons, Hoboken (NJ) 2006
- [EFQM 2009]
EFQM, EFQM Framework for Corporate Data Quality Management - Assessing the Organization's Data Quality Management Capabilities, European Foundation for Quality Management (EFQM), Brüssel 2009
- [Egloff 2008]
Egloff, M. T., Applying Semantic Wiki Technology to Corporate Data Metadata Management - An Implementation Project at Bayer CropScience, Diploma, University of St. Gallen, Institute of Information Management, St. Gallen 2008
- [Eisenhardt 1989]
Eisenhardt, K. M., Building Theories from Case Study Research, in: Academy Of Management Review, 14, 1989, Nr. 4, S. 532-550

- [Elmagarmid et al. 2007]
Elmagarmid, A. K., Ipeirotis, P. G., Verykios, V. S., Duplicate Record Detection: A Survey, in: IEEE Transactions on Knowledge and Data Engineering, 19, 2007, Nr. 1, S. 1-16
- [Elmasri/Navathe 2000]
Elmasri, R., Navathe, S. B., Fundamentals of Database Systems, 3. Aufl., Addison Wesley, Reading et al. 2000
- [Endava 2006]
Endava, Data Migration - The Endava Approach, London 2006
- [English 1999]
English, L. P., Improving Data Warehouse and Business Information Quality, 1. Aufl., John Wiley & Sons, Inc., New York (NY) 1999
- [English 2006]
English, L. P., Information Stewardship: Accountability for Information Quality, Information Impact International, Brentwood (TN) 2006
- [Eppler 2006]
Eppler, M. J., Managing Information Quality, 2. Aufl., Springer, Berlin/Heidelberg 2006
- [Ferstl/Sinz 2001]
Ferstl, O. K., Sinz, E. J., Grundlagen der Wirtschaftsinformatik, 4. Aufl., Oldenbourg, München 2001
- [Fettke 2009]
Fettke, P., How Conceptual Modeling Is Used, in: Communications of the AIS (CAIS), 25, 2009, Nr. 43, S. 571-592
- [Fettke/Loos 2003]
Fettke, P., Loos, P., Multiperspective Evaluation of Reference Models - Towards a Framework, in: Jeusfeld, M., Pastor, O. (Hrsg.), Conceptual Modelling for Novel Application Domains, Lecture Notes in Computer Science (LNCS), Vol. 2814, Springer, Berlin et al. 2003, S. 80-91
- [Foshay et al. 2007]
Foshay, N., Mukherjee, A., Taylor, A., Does Data Warehouse End-User Metadata Add Value?, in: Communications of the ACM (CACM), 50, 2007, Nr. 11, S. 70-77
- [Frank 1994]
Frank, U., Multiperspektivische Unternehmensmodellierung: Theoretischer Hintergrund und Entwurf einer objektorientierten Entwicklungsumgebung, Oldenbourg, München 1994
- [Frank 2003]
Frank, U., Ebenen der Abstraktion und ihre Abbildung auf konzeptionelle Modelle - oder: Anmerkungen zur Semantik von Spezialisierungs- und Instanzierungsbeziehungen, in: EMISA Forum, 23, 2003, Nr. 2, S. 14-18
- [Frank 2006]
Frank, U., Evaluation of Reference Models, in: Fettke, P., Loos, P. (Hrsg.), Reference Modeling for Business Systems Analysis, IGI Publishing, 2006, S. 118--139

[Frankfurth/Schellhase 2007]

Frankfurth, A., Schellhase, J., Content-Migration beim Wechsel zwischen verschiedenen Systemkategorien zur Content-Erstellung und -Pflege, in: Eibl, C.J., Magenheim, J., Schubert, S.E., Wessner, M. (Hrsg.), Proceedings der 5. e-Learning Fachtagung Informatik (DeLFI 2007), 111, GI, Siegen 2007, S. 115-126

[Friedman et al. 2008]

Friedman, T., Beyer, M. A., Bitterer, A., Survey Highlights Trends in Organization, Procurement and Standardization Practices for Data Integration, Gartner Group, Stamford 2008

[Gabriel/Röhrs 1996]

Gabriel, R., Röhrs, H.-P., Datenbanksysteme, 2. Aufl., Springer, Berlin 1996

[Gal 2006]

Gal, A., Why is Schema Matching Tough and What Can We Do About It?, in: SIGMOD Record, 35, 2006, Nr. 4, S. 2-5

[Gemünden/Schmitt 1991]

Gemünden, H. G., Schmitt, M., Datenmanagement in deutschen Grossunternehmen - Theoretischer Ansatz und empirische Untersuchung, in: Information Management, 6, 1991, Nr. 4, S. 22-34

[Gibson et al. 2007]

Gibson, A., Wolstencroft, K., Stevens, R., Promotion of Ontological Comprehension: Exposing Terms and Metadata with Web 2.0, in: Noy, N., Alani, H., Stumme, G., Mika, P., Sure, Y., Vrandecic, D. (Hrsg.), Proceedings of the CKC 2007, Workshop on Social and Collaborative Construction of Structured Knowledge at the 16th International World Wide Web Conference (WWW2007), Banff 2007

[Gläser/Laudel 2009]

Gläser, J., Laudel, G., Experteninterviews und qualitative Inhaltsanalyse, 2. Aufl., VS Verlag für Sozialwissenschaften, Wiesbaden 2009

[Grewe 2009]

Grewe, A., Datenarchitekturen - Welche Datenmodelle braucht das Unternehmen?, Präsentation auf dem IIR Data Management Kongress 2009, Köln 2009

[GS1 2005]

GS1, Business Requirements Document for Data Synchronization Data Model for Trade Item, 2005

[GS1 2009]

GS1, Global Data Dictionary, <http://gdd.gs1.org/gdd/public/default.asp>, 01.11.2009

[Gutzwiller 1994]

Gutzwiller, T., Das CC RIM-Referenzmodell für den Entwurf von betrieblichen, transaktionsorientierten Informationssystemen, Physica, Heidelberg 1994

[Hahn 2005]

Hahn, A., Integration verteilter Produktmodelle durch Semantic-Web-Technologien, in: Wirtschaftsinformatik, 47, 2005, Nr. 4, S. 278-284

[Halb et al. 2008]

Halb, W., Raimond, Y., Hausenblas, M., Building Linked Data For Both Humans and Machines, Proceedings of the Linked Data on the Web Workshop at the 17th International World Wide Web Conference 2008 (WWW2008), Beijing 2008, S. 1-8

[Haller 2009]

Haller, K., Towards the Industrialization of Data Migration: Concepts and Patterns for Standard Software Implementation Projects, in: van Eck, P., Gordijn, J., Wieringa, R. (Hrsg.), Proceedings of the 22nd International Conference on Advanced Information Systems Engineering (CAiSE'09), Lecture Notes in Computer Science (LNCS), Vol. 5565, Springer, Berlin/Heidelberg 2009, S. 63–78

[Hansen/Neumann 2005]

Hansen, H. R., Neumann, G., Wirtschaftsinformatik 1, 9. Aufl., Lucius & Lucius, Stuttgart 2005

[Hartmann/Weber 2008]

Hartmann, S., Weber, M., Semantisch homogene Beschreibung von Data-Warehouse-Metadaten mit RDF, in: Bichler, M., Hess, T., Krcmar, H., Lechner, U., Matthes, F., Picot, A., Speitkamp, B., Wolf, P. (Hrsg.), Proceedings der Multikonferenz Wirtschaftsinformatik 2008, GITO, Berlin 2008, S. 117-128

[Hay 2002]

Hay, D. C., Requirements Analysis: From Business Views to Architecture, Prentice Hall, Upper Saddle River (NJ) 2002

[Heine 1999]

Heine, P., Unternehmensweite Datenintegration: Modular-integrierte Datenlogistik in betrieblichen Informationssystemen, B.G. Teubner, Stuttgart 1999

[Heinrich et al. 2009]

Heinrich, B., Klier, M., Kaiser, M., A Procedure to Develop Metrics for Currency and its Application in CRM, in: ACM Journal of Data and Information Quality (JDIQ), 1, 2009, Nr. 1, S. Article 5

[Heinrich et al. 2007]

Heinrich, L. J., Heinzl, A., Roithmayr, F., Wirtschaftsinformatik - Einführung und Grundlegung, 3. Aufl., R. Oldenbourg Verlag, München 2007

[Helfert 2002]

Helfert, M., Planung und Messung der Datenqualität in Data-Warehouse-Systemen, Dissertation, Universität St. Gallen, Difo-Druck, Bamberg 2002

[Hepp 2003]

Hepp, M., Güterklassifikation als semantisches Standardisierungsproblem, Deutscher Universitätsverlag, Wiesbaden 2003

[Hepp et al. 2006]

Hepp, M., Leukel, J., Schmitz, V., A quantitative analysis of product categorization standards: content, coverage, and maintenance of eCl@ss, UNSPSC, eOTD, and the RosettaNet Technical Dictionary, in: Knowledge and Information Systems, 13, 2006, Nr. 1, S. 77-114

[Hernandez/Stolfo 1998]

Hernandez, M. A., Stolfo, S. J., Real-World Data is Dirty: Data Cleansing and the Merge/Purge Problem, in: *Data Mining and Knowledge Discovery*, 2, 1998, Nr. 1, S. 9-37

[Herwig 2006]

Herwig, V., Data Standardization as the Basis for an Service-oriented Architecture, *Proceedings of the IEEE Services Computing Workshops (SCW'06)*, IEEE Computer Society, New York (NY) 2006, S. 156-164

[Hess 1996]

Hess, T., Entwurf betrieblicher Prozesse, Dissertation, Universität St. Gallen, Deutscher Universitätsverlag, Wiesbaden 1996

[Heutschi 2007]

Heutschi, R., Serviceorientierte Architektur: Architekturprinzipien und Umsetzung in die Praxis, Springer, Berlin/Heidelberg 2007

[Heutschi et al. 2004]

Heutschi, R., Leser, F., Erni, F., Alt, R., Österle, H., Webservice-Technologien als Enabler des Real-time Business, in: Alt, R., Österle, H. (Hrsg.), *Real-time Business: Lösungen, Bausteine und Potentiale des Business Networking*, Springer, Berlin 2004, S. 133-155

[Hevner et al. 2004]

Hevner, A. R., March, S. T., Park, J., Ram, S., Design Science in Information Systems Research, in: *Management Information Systems Quarterly*, 28, 2004, Nr. 1, S. 75-105

[Heym 1993]

Heym, M., Methoden-Engineering - Spezifikation und Integration von Entwicklungsmethoden für Informationssysteme, Dissertation, Universität St. Gallen, Rosch-Buch, Hallstadt 1993

[Heym/Österle 1993]

Heym, M., Österle, H., Computer-Aided Methodology Engineering, in: *Information and Software Technology*, 35, 1993, Nr. 6/7, S. 345-354

[Hitzler et al. 2008]

Hitzler, P., Krötzsch, M., Rudolph, S., Sure, Y., *Semantic Web - Grundlagen*, Springer, Berlin/Heidelberg 2008

[Hoberman 2008]

Hoberman, S., Canonical Data Model, in: *DM Review*, 18, 2008, Nr. 8, S. 8

[Hoberman et al. 2009]

Hoberman, S., Burbank, D., Bradley, C., *Data Modeling for the Business - A Handbook for Aligning the Business with IT using High-Level Data Models*, 1. Aufl., Technics Publications, Bradley Beach (NJ) 2009

[Höning 2009]

Höning, F., Methodenkernel des Business Engineering - Metamodell, Vorgehensmodell, Techniken, Ergebnisdokumente und Rollen, Dissertation, Universität St. Gallen, St. Gallen 2009

[Hsu et al. 1992]

Hsu, C., Babin, G., Yee, L., Bouziane, M. h., Cheung, W., Rattner, L., Metadata Modeling for Enterprise Information Integration, in: *Journal of Systems Integration*, 2, 1992, Nr. 1, S. 5-37

- [Hüner et al. 2011]
Hüner, K., Brauer, B., Otto, B., Österle, H., Fachliches Metadatenmanagement mit einem semantischen Wiki, in: HMD -- Praxis der Wirtschaftsinformatik (angenommen), 279, 2011
- [Hüner/Ofner 2009]
Hüner, K., Ofner, M., Linked Data - The Future of the Internet or Old Wine in New Skins?, Presentation held at the 5th CC CDQ2 Workshop, Vevey, Switzerland 2009
- [Hüner et al. 2009]
Hüner, K., Ofner, M., Otto, B., Towards a Maturity Model for Corporate Data Quality Management, in: Shin, D. (Hrsg.), Proceedings of the 2009 ACM Symposium on Applied Computing, Manoa (HI) 2009, S. 231-238
- [Hüner/Otto 2009]
Hüner, K., Otto, B., The Effect of Using a Semantic Wiki for Metadata Management: A Controlled Experiment, in: Sprague, R.H. (Hrsg.), Proceedings of the 42nd Hawaii International Conference on System Sciences, IEEE Computer Society, Los Alamitos 2009, S. 1-9
- [IMG 1996]
IMG, PROMET PSI: Methodenhandbuch für Prozess- und Systemintegration, Version 1.1, Information Management Group (IMG AG), St. Gallen 1996
- [IMG 1997]
IMG, PROMET BPR, Methodenhandbuch für den Entwurf von Geschäftsprozessen, Version 2.0, Information Management Group (IMG AG), St. Gallen 1997
- [IMG 1999]
IMG, PROMET STP: Methodenhandbuch für die System- und Technologieplanung, Release 1.0, Information Management Group (IMG AG), St. Gallen 1999
- [Informatica 2006]
Informatica, Putting Metadata to Work to Achieve the Goals of Data Governance, Informatica Corporation, Redwood City (CA) 2006
- [Inmon 1999]
Inmon, W. H., Building the Operational Data Store, 2. Aufl., John Wiley & Sons, Inc., New York (NY) 1999
- [Inmon et al. 2008]
Inmon, W. H., O'Neil, B. K., Fryman, L., Business Metadata: Capturing Enterprise Knowledge, Morgan Kaufmann, New Jersey, USA 2008
- [ISO/IEC 2003]
ISO/IEC, Information Technology - Metadata Registries (MDR) - Part 3: Registry metamodel and basic attributes (ISO/IEC 11179-3), International Organization for Standardization (ISO), Geneva 2003
- [ISO/IEC 2004a]
ISO/IEC, Information Technology - Metadata Registries (MDR) - Part 1: Framework (ISO/IEC 11179-1), International Organization for Standardization (ISO), Geneva 2004a

[ISO/IEC 2004b]

ISO/IEC, Information Technology - Metadata Registries (MDR) - Part 4: Formulation of data definitions (ISO/IEC 11179-4), International Organization for Standardization (ISO), Geneva 2004b

[ISO/IEC 2004c]

ISO/IEC, Information Technology - Metadata Registries (MDR) - Part 6: Registration (ISO/IEC 11179-6), International Organization for Standardization (ISO), Geneva 2004c

[ISO/IEC 2005a]

ISO/IEC, Information Technology - Metadata Registries (MDR) - Part 2: Classification (ISO/IEC 11179-2), International Organization for Standardization (ISO), Geneva 2005a

[ISO/IEC 2005b]

ISO/IEC, Information Technology - Metadata Registries (MDR) - Part 5: Naming and identification principles (ISO/IEC 11179-5), International Organization for Standardization (ISO), Geneva 2005b

[ITAIDE 2006]

ITAIDE, Information Technology for Adoption and Intelligent Design for E-Government (project nr. 027829), <http://www.itaide.org/>, 15.08.2008

[Jain/Zhao 2003]

Jain, H., Zhao, H., A Conceptual Model for Comparative Analysis of Standardization of Vertical Industry Languages, in: King, J.L., Lyytinen, K. (Hrsg.), Proceedings of the Workshop on Standard Making: A Critical Research Frontier for Information Systems, Seattle (WA) 2003, S. 210-221

[Jakic 2009]

Jakic, K., Strategie-Workshop „Roadmap 2012“ - Zentrales Kundenstammdatenmanagement, Interne Präsentation, Friedrichshafen 2009

[Janiesch 2007]

Janiesch, C., Implementing Views on Business Semantics: Model-based Configuration of Business Document, in: Österle, H., Schelp, J., Winter, R. (Hrsg.), Proceedings of the 15th European Conference on Information Systems, St. Gallen 2007, S. 2050 - 2061

[Jarosch 2010]

Jarosch, H., Grundkurs Datenbankentwurf - Eine beispielorientierte Einführung für Studenten und Praktiker, 2. Aufl., Vieweg+Teubner, Wiesbaden 2010

[Jarvinen 2000]

Jarvinen, P. H., Research Questions Guiding Selection of an Appropriate Research Method, in: Hansen, H.R., Bichler, M., Mahrer, H. (Hrsg.), Proceedings of the 8th European Conference on Information Systems, Vienna 2000, S. 124-131

[Josuttis 2008]

Josuttis, N., SOA in der Praxis, dpunkt, Heidelberg 2008

[Jung 2006]

Jung, R., Architekturen zur Datenintegration. Gestaltungsempfehlungen auf der Basis fachkonzeptueller Anforderungen, Deutscher Universitätsverlag, Wiesbaden 2006

- [Jung/Winter 2006]
Jung, R., Winter, R., Gestaltung von Datenintegrationsarchitekturen, Arbeitsbericht 188, Institut für Wirtschaftsinformatik, Universität Bern, Bern 2006
- [Kagermann/Österle 2006]
Kagermann, H., Österle, H., Geschäftsmodelle 2010. Wie CEOs Unternehmen transformieren, F.A.Z. Buch, Frankfurt am Main 2006
- [Kaib 2002]
Kaib, M., Enterprise Application Integration. Grundlagen, Integrationsprodukte, Anwendungsbeispiele, DUV Wirtschaftsinformatik, Wiesbaden 2002
- [Kaiser 2000]
Kaiser, T. M., Methode zur Konzeption des Intranets, Dissertation, Universität St. Gallen, Difo-Druck, Bamberg 2000
- [Karlsson/Wistrand 2006]
Karlsson, F., Wistrand, K., Combining Method Engineering with Activity Theory - Theoretical Grounding of the Method Component Concept, in: European Journal of Information Systems, 15, 2006, Nr. 1, S. 82-90
- [Kentouris 2009]
Kentouris, C., Business Semantics Repository Aims for Data Precision, in: Securities Industry News, 21, 2009, Nr. 6, S. 10-14
- [Khathri/Brown 2010]
Khathri, V., Brown, C. V., Designing Data Governance, in: Communications of the ACM (CACM), 53, 2010, Nr. 1, S. 148-153
- [Kifer et al. 1995]
Kifer, M., Lausen, G., Wu, J., Logical Foundations of Object-Oriented and Frame-Based Languages, in: Journal of the ACM, 42, 1995, Nr. 4, S. 741-843
- [Kim/Seo 1991]
Kim, W., Seo, J., Classifying schematic and data heterogeneity in multidatabase systems, in: Computer, 24, 1991, Nr. 12, S. 12-18
- [Kramer 2007]
Kramer, S., Modellierung und Konfiguration elektronischer Geschäftsdokumente - Fachkonzeptionelle Spezifikation und exemplarische Umsetzung, Diploma, Lehrstuhl für Wirtschaftsinformatik und Informationsmanagement, Westfälische Wilhelms-Universität Münster, Münster 2007
- [Krcmar 1990]
Krcmar, H., Bedeutung und Ziele von Informationssystemarchitekturen, in: Wirtschaftsinformatik, 32, 1990, Nr. 5, S. 395-402
- [Krcmar 1992]
Krcmar, H., Informationslogistik der Unternehmung: Konzept und Perspektiven, in: Stroetmann, K.A. (Hrsg.), Informationslogistik, Deutsche Gesellschaft für Dokumentation 1992, S. 67-90
- [Krcmar 2005]
Krcmar, H., Informationsmanagement, 4. Aufl., Springer, Berlin 2005
- [Kröttsch et al. 2007]
Kröttsch, M., Vrandečić, D., Völkel, M., Haller, H., Studer, R., Semantic Wikipedia, in: Journal of Web Semantics, 5, 2007, Nr. 4, S. 251-261

[Kugeler/Rosemann 1998]

Kugeler, M., Rosemann, M., Fachbegriffsmodellierung für betriebliche Informationssysteme zur Unterstützung der Unternehmenskommunikation, in: Informationssystem-Architekturen, 5, 1998, Nr. 2, S. 8-15

[Kumar/Welke 1992]

Kumar, K., Welke, R. J., Methodology Engineering - A Proposal for Situation-Specific Methodology Construction, in: Cotterman, W.W., Senn, J.A. (Hrsg.), Challenges and Strategies for Research in Systems Development, John Wiley & Sons, Chichester 1992, S. 257-269

[Kurbel/Rautenstrauch 1996]

Kurbel, K., Rautenstrauch, C., Integration Engineering: Konkurrenz oder Komplement zum Information Engineering? - Methodische Ansätze zur Integration von Informationssystemen, in: Heilmann, H., Heinrich, L.J., Roithmayr, F. (Hrsg.), Information Engineering - Wirtschaftsinformatik im Schnittpunkt von Wirtschafts-, Sozial- und Ingenieurwissenschaften, Oldenbourg, München, Wien 1996, S. 167-191

[Kurpjuweit 2009]

Kurpjuweit, S., Stakeholder-orientierte Modellierung und Analyse der Unternehmensarchitektur : unter besonderer Berücksichtigung der Geschäfts- und IT-Architektur, Universität St.Gallen, St. Gallen 2009

[Lee et al. 2006]

Lee, Y. W., Pipino, L. L., Funk, J., Wang, R. Y., Journey to Data Quality, MIT Press, Cambridge (MA) 2006

[Legler/Naumann 2007]

Legler, F., Naumann, F., A Classification of Schema Mappings and Analysis of Mapping Tools, Proceedings der 12. GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web (BTW 2007), Aachen 2007, S. 449-464

[Legner/Heutschi 2007]

Legner, C., Heutschi, R., SOA Adoption in Practice - Findings from Early SOA Implementations, in: Österle, H., Schelp, J., Winter, R. (Hrsg.), Proceedings of the 15th European Conference on Information Systems, St. Gallen 2007, S. 1643-1654

[Legner/Otto 2007]

Legner, C., Otto, B., Stammdatenmanagement, in: WISU - Das Wirtschaftsstudium, 236, 2007, Nr. 4, S. 562-568

[Legner/Vogel 2008]

Legner, C., Vogel, T., Leveraging Web Services for Implementing Vertical Industry Standards - A Model for Service-Based Interoperability, in: Electronic Markets, 18, 2008, Nr. 1, S. 39-52

[Lehmann 2001]

Lehmann, P., Meta-Datenmanagement in Data-Warehouse-Systemen. Rekonstruierte Fachbegriffe als Grundlage einer konstruktiven, konzeptionellen Modellierung, Dissertation, Universität Magdeburg, Magdeburg 2001

[Lehmann/Jaszewski 1999]

Lehmann, P., Jaszewski, J., Business Terms as a Critical Success Factor for Data Warehousing, Proceedings of the International Workshop on Design and Management of Data Warehouse, Heidelberg 1999, S. 1-5

- [Lemm 2007]
Lemm, A., XML-basierte E-Business-Frameworks - Analyse, Konvergenztendenzen und Vergleich, Studienarbeit, Technische Universität Darmstadt / Universität St. Gallen, Darmstadt 2007
- [Leser/Naumann 2007]
Leser, U., Naumann, F., Informationsintegration - Architekturen und Methoden zur Integration verteilter und heterogener Datenquellen, dpunkt.verlag, Heidelberg 2007
- [Liebhart 2007]
Liebhart, D., SOA goes real: Serviceorientierte Architekturen erfolgreich planen und einführen, Carl Hanser, München 2007
- [Linthicum 2000]
Linthicum, D. S., Enterprise Application Integration, Addison-Wesley, Upper Saddle River (NJ) 2000
- [Longman 2008]
Longman, C., Why Master Data Management is Such a Challenge, in: DM Review, 18, 2008, Nr. 11, S. 18
- [Loos 1999]
Loos, P., Grunddatenverwaltung und Betriebsdatenerfassung als Basis der Produktionsplanung und -steuerung, in: Corsten, H., Friedl, B. (Hrsg.), Produktionscontrolling, Vahlen, München 1999, S. 227-252
- [Loser et al. 2004]
Loser, C., Legner, C., Gizanis, D., Master Data Management for Collaborative Service Processes, in: Jian, C. (Hrsg.), International Conference on Service Systems and Service Management, Research Center for Contemporary Management, Tsinghua University, Tsinghua 2004
- [Loshin 2008]
Loshin, D., Master Data Management, Elsevier Science & Technology Books, Burlington (MA) 2008
- [Low et al. 2001]
Low, W. L., Lee, M. L., Ling, T. W., A knowledge-based approach for duplicate elimination in data cleaning, in: Informations Systems, 26, 2001, Nr. 8, S. 585-606
- [Lutz 2000]
Lutz, J. C., EAI Architecture Patterns, in: EAI Journal, 2000, Nr. March, S. 64-73
- [Mädche et al. 2001]
Mädche, A., Staab, S., Studer, R., Ontologien, in: Wirtschaftsinformatik, 43, 2001, Nr. 4, S. 393-395
- [Madnick 1999]
Madnick, S. E., Metadata Jones and the tower of Babel: The challenge of large-scale semantic heterogeneity, Proceedings of the 3rd IEEE Meta-Data Conference, Bethesda (MD) 1999, S. 1-13
- [Maier 1998]
Maier, R., Nutzen und Qualität der Datenmodellierung - Ergebnisse einer empirischen Studie, in: Wirtschaftsinformatik, 40, 1998, Nr. 2, S. 130-140

- [March et al. 2000]
March, S. T., Hevner, A. R., Ram, S., Research Commentary: An Agenda for Information Technology Research in Heterogeneous Environments, in: Information Systems Research (ISR), 11, 2000, Nr. 4, S. 327-341
- [March/Smith 1995]
March, S. T., Smith, G. F., Design and natural science research on information technology, in: Decision Support Systems (DSS), 15, 1995, Nr. 4, S. 251-266
- [March/Storey 2008]
March, S. T., Storey, V. C., Design Science in the Information Systems Discipline: An Introduction to the Special Issue on Design Science Research, in: MIS Quarterly, 32, 2008, Nr. 4, S. 725-730
- [Marco 2000]
Marco, D., Building and Managing the Meta Data Repository. A Full Lifecycle Guide, John Wiley & Sons, Hoboken (NJ) 2000
- [Marco 2005]
Marco, D., CIOs and Meta Data Management - Part 1, in: DM Review, 15, 2005, Nr. 7, S. 10/34
- [Marco/Michael 2004]
Marco, D., Michael, J., Universal Meta Data Models, John Wiley & Sons, Indianapolis (IN) 2004
- [Marco/Smith 2006]
Marco, D., Smith, A. M., Metadata Management & Enterprise Architecture: Understanding Data Governance and Stewardship, Part II, in: DM Review, 16, 2006, Nr. 10, S. 17
- [Martin 1977]
Martin, J., Computer Data-Base Organization, 2. Aufl., Prentice-Hall, Englewood Cliff (NJ) 1977
- [Martin/Leben 1989]
Martin, J., Leben, J., Strategic Information Planning Methodologies, 2. Aufl., Prentice-Hall, Englewood Cliff (NJ) 1989
- [Mayr et al. 1987]
Mayr, H. C., Dittrich, K. R., Lockemann, P. C., Datenbankentwurf, in: Lockemann, P.C., Schmidt, J.W. (Hrsg.), Datenbankhandbuch, 1987, S. 481-557
- [Mayring 2008]
Mayring, P., Qualitative Inhaltsanalyse: Grundlagen und Techniken, Beltz, Weinheim 2008
- [McAfee 2005]
McAfee, A., Will Web Services Really Transform Collaboration?, in: MIT Sloan Management Review, 46, 2005, Nr. 2, S. 78-84
- [McComb 2003]
McComb, D., Semantics in Business Systems. The savvy manager's guide, Morgan Kaufmann, San Francisco (CA) 2003
- [Medjahed et al. 2003]
Medjahed, B., Benatallah, B., Bouguettaya, A., Ngu, A. H. H., Elmagarmid, A. K., Business-to-business interactions: issues and enabling technologies, in: The VLDB Journal — The International Journal on Very Large Data Bases, 12, 2003, Nr. 1, S. 59 - 85

- [Melchert 2006]
Melchert, F., Methode zur Konzeption von Metadatenmanagementsystemen für das Data Warehousing, Dissertation, Universität St. Gallen, St. Gallen 2006
- [Mertens 2001]
Mertens, P., Lexikon der Wirtschaftsinformatik, 4. Aufl., Springer, Berlin 2001
- [Mertens 2007]
Mertens, P., Integrierte Informationsverarbeitung 1, 16. Aufl., Gabler, Wiesbaden 2007
- [Mertens/Holzner 1992]
Mertens, P., Holzner, J., WI - State of the Art. Eine Gegenüberstellung von Integrationsansätzen der Wirtschaftsinformatik, in: Wirtschaftsinformatik, 34, 1992, Nr. 1, S. 5-25
- [Moody/Shanks 1994]
Moody, D. L., Shanks, G., What Makes a Good Data Model? Evaluating the Quality of Entity Relationship Models, in: Loucopoulos, P. (Hrsg.), Proceedings of the 13th International Conference on the Entity-Relationship Approach, 881, Springer, Berlin/Heidelberg 1994, S. 94-111
- [Moraes Batista/Salgado 2007]
Moraes Batista, M., Salgado, A. C., Data Integration Schema Analysis: An Approach with Information Quality, in: Robert, M.A., O'Hare, R., Markus, M.L., Klein, B. (Hrsg.), Proceedings of 12th International Conference on Information Quality, Cambridge (MA) 2007,
- [Morgan 1997]
Morgan, D. L., Focus Groups as Qualitative Research, 2. Aufl., Sage Publications, Thousand Oaks (CA) 1997
- [Morris 2006]
Morris, J., Practical Data Migration, British Computer Society, Swindon 2006
- [Morschheuser et al. 1996]
Morschheuser, S., Raufer, H., Wargitsch, C., Challenges and Solutions of Document and Workflow Management in a Manufacturing Enterprise: A Case Study, Proceedings of the Hawaii Conference on System Sciences (HICSS '96), IEEE Computer Society Press 1996,
- [Moss 2007]
Moss, L. T., Enterprise Data Modeling: Lost Art or Essential Science?, in: Business Intelligence Journal, 12, 2007, Nr. 1, S. 7-12
- [Muljadi et al. 2006]
Muljadi, H., Takeda, H., Ando, K., Development of a Semantic Wiki-based Feature Library for the Extraction of Manufacturing Feature and Manufacturing Information, in: International Journal of Computer Science, 1, 2006, Nr. 4, S. 265-273
- [Müller/Gronau 2008]
Müller, C., Gronau, N., Wikis, in: Back, A., Gronau, N., Tochtermann, K. (Hrsg.), Web 2.0 in der Unternehmenspraxis - Grundlagen, Fallstudien und Trends zum Einsatz von Social Software, Oldenbourg, München 2008, S. 10-17
- [Müller/Freytag 2003]
Müller, H., Freytag, J.-C., Problems, Methods, and Challenges in Comprehensive Data Cleansing, Working Paper, 2003

- [Müller et al. 2006]
Müller, S., Kuhn, W., Meiler, C., Petrov, I., Jablonski, S., Integratives IT-Architekturmanagement, in: Reussner, R., Hasselbring, W. (Hrsg.), Handbuch der Software-Architektur, dpunkt.verlag, Heidelberg 2006, S. 187-209
- [Myers 2002]
Myers, M. D., Qualitative Research in Information Systems, <http://www.qual.auckland.ac.nz/index.htm>, 16.09.2002
- [Nohr et al. 2008]
Nohr, H., Roos, A. W., Vöhringer, A., Relationship Management von Verbundgruppen, in: Becker, J., Knackstedt, R., Pfeiffer, D. (Hrsg.), Wertschöpfungsnetzwerke - Konzepte für das Netzwerkmanagement und Potenziale aktueller Informationstechnologien, Physica-Verlag, Heidelberg 2008, S. 153-169
- [Nuseibeh et al. 1996]
Nuseibeh, B. A., Finkelstein, A., Kramer, J., Method Engineering for Multi-Perspective Software Development, in: Information and Software Technology, 38, 1996, Nr. 4, S. 267-274
- [O'Neil 2005]
O'Neil, B. K., Business Metadata: How To Write Definitions, <http://www.tdan.com/i032fe01.htm>, 27.08.2007
- [Oestereich 2004]
Oestereich, B., Objektorientierte Softwareentwicklung - Analyse und Design mit der UML 2.0, 6. Aufl., Oldenbourg, München/Wien 2004
- [OGC 2007]
OGC, ITIL - Service Transition, TSO, 2007
- [Olson 2003]
Olson, J., Data Quality - The Accuracy Dimension, Morgan Kaufmann, San Francisco (CA) 2003
- [OMG 2007]
OMG, Unified Modeling Language: Superstructure - Version 2.1.1, Object Management Group (OMG), Needham (MA) 2007
- [Ortner 1991]
Ortner, E., Informationsmanagement - Wie es entstand, was es ist und wohin es sich entwickelt, in: Informatik-Spektrum, 14, 1991, S. 315-327
- [Ortner 1999]
Ortner, E., Repository Systems. Teil 1: Mehrstufigkeit und Entwicklungsumgebung, in: Informatik-Spektrum, 22, 1999, S. 235-251
- [Ortner 2000]
Ortner, E., Terminologiebasierte, komponentenorientierte Entwicklung von Anwendungssystemen, in: Flatscher, R.G., Turowski, K. (Hrsg.), Proceedings des 2. Workshop komponentenorientierte betriebliche Anwendungssysteme (WKBA 2), Wien 2000, S. 1-20
- [Ortner et al. 1990]
Ortner, E., Rössner, J., Söllner, B., Entwicklung und Verwaltung standardisierter Datenelemente, in: Informatik Spektrum, 13, 1990, S. 17-30
- [Österle 1995]
Österle, H., Business Engineering: Prozess- und Systementwicklung, Band 1: Entwurfstechniken, 2. Aufl., Springer, Berlin 1995

- [Österle/Blessing 2003]
Österle, H., Blessing, D., Business Engineering Modell, in: Österle, H., Winter, R. (Hrsg.), Business Engineering, Springer, Berlin 2003, S. 65-85
- [Österle/Otto 2009]
Österle, H., Otto, B., A Method For Consortial Research, Working Paper BE HSG / CC CDQ / 6, Institute of Information Management, University of St. Gallen, St. Gallen 2009
- [Österle et al. 2007]
Österle, H., Winter, R., Höning, F., Kurpjuweit, S., Osl, P., Business Engineering: Core-Business-Metamodell, in: WISU - Das Wirtschaftsstudium, 36, 2007, Nr. 2, S. 191-194
- [Otto/Beckmann 2001]
Otto, B., Beckmann, H., Klassifizierung und Austausch von Produktdaten auf elektronischen Marktplätzen, in: Wirtschaftsinformatik, 43, 2001, Nr. 4, S. 351-362
- [Otto et al. 2002]
Otto, B., Beckmann, H., Kelkar, O., Müller, S., E-Business-Standards. Verbreitung und Akzeptanz, Fraunhofer IRB Verlag, Stuttgart 2002
- [Otto/Hüner 2009]
Otto, B., Hüner, K., Funktionsarchitektur für unternehmensweites Stammdatenmanagement, Arbeitsbericht BE HSG / CC CDQ / 14, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen 2009
- [Otto et al. 2009]
Otto, B., Hüner, K., Österle, H., Identification of Business Oriented Data Quality Metrics, in: Bowen, P., Elmagarmid, A.K., Österle, H., Sattler, K.-U. (Hrsg.), Proceedings of the 14th International Conference on Information Quality (ICIQ'09), Potsdam 2009, S. 122-134
- [Otto/Österle 2010]
Otto, B., Österle, H., Relevance through Consortium Research? Findings from an Expert Interview Study, in: Winter, R., Zhao, J.L., Aier, S. (Hrsg.), Global Perspectives on Design Science Research, Springer 2010, S. 16-30
- [Otto et al. 2008]
Otto, B., Wende, K., Schmidt, A., Hüner, K., Vogel, T., Unternehmensweites Datenqualitätsmanagement: Ordnungsrahmen und Anwendungsbeispiele, in: Dinter, B., Winter, R. (Hrsg.), Integrierte Informationslogistik, Springer, Berlin 2008, S. 211-230
- [Päivärinta et al. 2002]
Päivärinta, T., Tyrväinen, P., Ylimäki, T., Defining Organizational Document Metadata: A Case Beyond Standards, Proceedings of 10th European Conference on Information Systems (ECIS), Gdansk 2002,
- [Papazoglou 2007]
Papazoglou, M. P., Web Services: Principles and Technology, Longman, Harlow 2007
- [Papazoglou/van den Heuvel 2006]
Papazoglou, M. P., van den Heuvel, W.-J., Service-oriented design and development methodology, in: International Journal of Web Engineering and Technology, 2, 2006, Nr. 4, S. 412-442

[Parent/Spaccapietra 1998]

Parent, C., Spaccapietra, S., Issues and approaches of database integration, in: Communications of the ACM (CACM), 41, 1998, Nr. 5, S. 166-178

[Park/Ram 2004]

Park, J., Ram, S., Information Systems Interoperability: What belies Beneath?, in: ACM Transactions on Information Systems, 22, 2004, Nr. 4, S. 595-632

[Peppers et al. 2006]

Peppers, K., Tuunanen, T., Gengler, C. E., Rossi, M., Hui, W., Virtanen, V., Bragge, J., The Design Science Research Process: A Model for Producing and Presenting Information Systems Research, Proceedings of the DESRIST 2006, 2006, S. 83--106

[Pentcheva 2007]

Pentcheva, D., Context Driven Collaborative Service-oriented Business Process Modeling, Diploma, Berlin Institute of Technology, Berlin 2007

[Periasamy 1993]

Periasamy, K. P., The State and Status of Information Architecture: An Empirical Investigation, in: DeGross, J.I., Bostrom, R.P., Robey, D. (Hrsg.), Proceedings of the Fourteenth International Conference on Information Systems, ACM, Orlando (FL) 1993, S. 255-270

[Periasamy/Feeny 1997]

Periasamy, K. P., Feeny, D. F., Information architecture practice: research-based recommendations for the practitioner, in: Journal of Information Technology, 12, 1997, Nr. 3, S. 197-205

[Perry 1998]

Perry, C., Processes of a case study methodology for postgraduate research in marketing, in: European Journal of Marketing, 32, 1998, Nr. 9-10, S. 785-802

[Pezzini 2003]

Pezzini, M., Application Integration Scenario: Building an Enterprise Nervous System, Gartner Symposium ITxpo, Gartner, inc. 2003,

[Pienimäki 2005]

Pienimäki, T., A Business Application Architecture Framework in Manufacturing Industry, Tampere University of Technology, Tampere 2005

[Pohl 2008]

Pohl, K., Requirements Engineering - Grundlagen, Prinzipien, Techniken, 1. Aufl., dpunkt.verlag, Heidelberg 2008

[Pohland 2000]

Pohland, S., Globale Unternehmensarchitekturen - Methode zur Verteilung von Informationssystemen, Weissensee Verlag, Berlin 2000

[Price/Shanks 2005]

Price, R., Shanks, G. G., A semiotic information quality framework: development and comparative analysis, in: Journal of Information Technology (JIT), 2005, 2005, Nr. 20, S. 88-102

[Priglinger/Friedrich 2008]

Priglinger, S., Friedrich, D., Master Data Management Survey 08, BARC-Institut (Business Application Research Center), Würzburg 2008

- [Puschmann 2003]
Puschmann, T., Collaboration Portale - Architektur, Integration, Umsetzung und Beispiele, Dissertation, Universität St. Gallen, Difo-Druck, Bamberg 2003
- [Puschmann 2004]
Puschmann, T., Prozessportale - Architektur zur Vernetzung mit Kunden und Lieferanten, Springer, Berlin 2004
- [Quantz/Wichmann 2003]
Quantz, J., Wichmann, T., E-Business-Standards in Deutschland, Berlecon Research, Berlin 2003
- [Radcliffe et al. 2006]
Radcliffe, J., White, A., Newman, D., How to Choose the Right Architectural Style for Master Data Management, Gartner Group, Stamford (CT) 2006
- [Rahm/Bernstein 2001]
Rahm, E., Bernstein, P. A., A Survey of Approaches to Automatic Schema Matching, in: The International Journal on Very Large Data Bases (VLDB Journal), 10, 2001, Nr. 4, S. 334-350
- [Rahm/Do 2000]
Rahm, E., Do, H.-H., Data Cleaning: Problems and Current Approaches, in: IEEE Data Engineering Bulletin, 23, 2000, Nr. 4, S. 3-13
- [Rautenstrauch/Schulze 2003]
Rautenstrauch, C., Schulze, T., Informatik für Wirtschaftswissenschaftler und Wirtschaftsinformatiker, Springer, Berlin 2003
- [Redman 1996]
Redman, T. C., Data Quality for the Information Age, Artech House, Boston/London 1996
- [Redman 2001]
Redman, T. C., Data Quality. The Field Guide, Digital Press, Boston (MA) 2001
- [Reimers 2001]
Reimers, K., Standardizing the new e-business platform: Learning from the EDI experience, in: Electronic Markets, 11, 2001, Nr. 4, S. 231-237
- [Riebisch 2006]
Riebisch, M., Prozess der Architektur- und Komponentenentwicklung, in: Reussner, R., Hasselbring, W. (Hrsg.), Handbuch der Software-Architektur, dpunkt.verlag, Heidelberg 2006, S. 65-88
- [Rosemann 1996]
Rosemann, M., Komplexitätsmanagement in Prozessmodellen: Methodenspezifische Gestaltungsempfehlungen für die Informationsmodellierung, Gabler, Wiesbaden 1996
- [Rosemann 1999]
Rosemann, M., Gegenstand und Aufgaben des Integrationsmanagements, in: Scheer, A.W., Rosemann, M., Schütte, R. (Hrsg.), Integrationsmanagement, Institut für Wirtschaftsinformatik, Westfälische Wilhelms-Universität, Münster 1999, S. 5-18

[Rosemann/Schütte 1997]

Rosemann, M., Schütte, R., Grundsätze ordnungsmässiger Referenzmodellierung, in: Becker, J., Rosemann, Michael, Schütte, R. (Hrsg.), Entwicklungsstand und Entwicklungsperspektiven der Referenzmodellierung, Proceedings zur Veranstaltung vom 10. März 1997, Institut für Wirtschaftsinformatik, Westfälischen Wilhelms-Universität Münster, Münster 1997, S. 16-33

[Rosemann/Vessey 2008]

Rosemann, M., Vessey, I., Toward Improving the Relevance of Information Systems Research to Practice: The Role of Applicability Checks, in: MIS Quarterly, 32, 2008, Nr. 1, S. 1-22

[Rosenthal et al. 2004]

Rosenthal, A., Seligman, L., Renner, S., From Semantic Integration to Semantics Management: Case Studies and a Way Forward, in: ACM SIGMOD Record, 33, 2004, Nr. 4, S. 44-50

[Rowell 2008]

Rowell, M., Oracle, OAGIS and Open Standards, <http://www.oagi.org/oagi/downloads/meetings/2008-1202-RedwoodCity/Wednesday/2008%201202%20Oracle%20OAGIS%20and%20Open%20Standards.pdf>, 10.01.2010

[Rüegg-Stürm 2002]

Rüegg-Stürm, J., Das neue St. Galler Management-Modell: Grundkategorien einer integrierten Managementlehre - Der HSG-Ansatz, Haupt, Bern/Stuttgart/Wien 2002

[Ruh et al. 2001]

Ruh, W. A., Maginnis, F. X., Brown, W. J., Enterprise Application Integration, John Wiley & Sons, Inc., New York (NY) 2001

[Rupprecht 2007]

Rupprecht, T., Global Master Data & Product Lifecycle Management @ Beiersdorf, Tagungsunterlagen Stammdaten-Management Forum 2007, Frankfurt am Main 2007

[Russom 2006]

Russom, P., Master Data Management: Consensus-Driven Data Definitions for Cross-Application Consistency, White Paper, The Data Warehouse Institute (TDWI), Seattle (WA) 2006

[Ryu et al. 2006]

Ryu, K.-S., Park, J.-S., Park, J.-H., A Data Quality Management Maturity Model, in: ETRI Journal, 28, 2006, Nr. 2, S. 191-204

[SAP 2009]

SAP, Core-Datentypen und aggregierte Datentypen - SAP-Dokumentation, http://help.sap.com/saphelp_nwpi71/helpdata/de/45/614fc4c5293bdce1000000a1553f7/content.htm, 18.11.2009

[SBB Cargo 2009]

SBB Cargo, SBB Cargo im Jahr 2008 - Auszug aus dem Geschäftsbericht der SBB, http://www.sbbcargo.com/gb_cargo08_d_gzd.pdf, 05.10.2009

[Scannapieco et al. 2005]

Scannapieco, M., Missier, P., Batini, C., Data Quality at a Glance, in: Datenbank-Spektrum, 5, 2005, Nr. 14, S. 6--14

- [Schaffert et al. 2007]
Schaffert, S., Bry, F., Baumeister, J., Kiesel, M., Semantic Wiki, in: Informatik-Spektrum, 30, 2007, Nr. 6, S. 434--439
- [Scheer 1998]
Scheer, A.-W., ARIS - Vom Geschäftsprozess zum Anwendungssystem, 3. Aufl., Springer, Berlin et al. 1998
- [Scheer 2001]
Scheer, A.-W., ARIS - Modellierungsmethoden, Metamodelle, Anwendungen, Springer, Berlin et al. 2001
- [Scheer/Loos 1996]
Scheer, A.-W., Loos, P., 15 Jahre Erfahrung mit unternehmensweiter Datenmodellierung und Datenbankeinsatz, in: Datenbank-Rundbrief (Mitteilungen der GI-Fachgruppe 2.5.1), 17, 1996, S. 1-12
- [Schelp/Winter 2006]
Schelp, J., Winter, R., Method Engineering - Lessons Learned from Reference Modeling, in: Chatterjee, S., Hevner, A. (Hrsg.), Proceedings of the 1st International Conference on Design Science in Information Systems and Technology (DESRIST 2006), Claremont (CA) 2006, S. 555-575
- [Schemm 2007]
Schemm, J., Fallstudie Mars Inc.: Global Data Synchronization in der Konsumgüterindustrie, Arbeitsbericht BE HSG / CC CDQ / 5, Insitut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen 2007
- [Schemm 2008]
Schemm, J., Stammdatenmanagement zwischen Handel und Konsumgüterindustrie - Referenzarchitektur für die überbetriebliche Datensynchronisation, Dissertation, Universität St. Gallen, Difo-Druck, Bamberg 2008
- [Schissler et al. 2004]
Schissler, M., Zeller, T., Mantel, S., Überbetriebliche Integration von Anwendungssystemen: Klassifikation von Integrationsproblemen und -lösungen, in: Bartmann, D., Mertens, P., Sinz, E.J. (Hrsg.), Überbetriebliche Integration von Anwendungssystemen - FORWIN-Tagung 2004, Shaker, Aachen 2004, S. 1-20
- [Schmidt 2009a]
Schmidt, A., Fallstudie inet-logistics – Integration von Jet Aviation und Fiege, Arbeitsbericht BE HSG / CC CDQ / 16, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen 2009a
- [Schmidt 2009b]
Schmidt, A., Fallstudie SBB Cargo – Einführung eines unternehmensweiten Stammdatenmanagements, Arbeitsbericht BE HSG / CC CDQ2 / 22, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen 2009b
- [Schmidt et al. 2010a]
Schmidt, A., Hüner, K., Grewe, A., Fallstudie Deutsche Telekom AG – Einheitliche Datenarchitektur als Grundlage für unternehmensweites Datenqualitätsmanagement, Arbeitsbericht BE HSG / CC CDQ2 / 23, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen 2010a

- [Schmidt/Osl 2008]
Schmidt, A., Osl, P., A Method for Establishing Transparency on Information Objects, Working Paper BE HSG / CC CDQ / 2, Institute of Information Management, University of St. Gallen, St. Gallen 2008
- [Schmidt/Otto 2008]
Schmidt, A., Otto, B., A Method for the Identification and Definition of Information Objects, in: Robert, M.A., O'Hare, R., Markus, M.L., Klein, B. (Hrsg.), Proceedings of the 13th International Conference on Information Quality, Cambridge (MA) 2008, S. 214-228
- [Schmidt et al. 2010b]
Schmidt, A., Otto, B., Österle, H., Integrating Information Systems: Case Studies on Current Challenges, in: Electronic Markets, 20, 2010b, Nr. 2, S. 161-174
- [Schmitt/Saake 2005]
Schmitt, I., Saake, G., A comprehensive database schema integration method based on the theory of formal concepts, in: Acta Informatica, 41, 2005, Nr. 7/8, S. 475-524
- [Schnabel 2007]
Schnabel, F., Metamodellbasierte Transformation von Geschäftsprozessen in service-orientierte Architekturen, Diplomarbeit, Universität Karlsruhe (TH), 2007
- [Scholz/Tietje 2002]
Scholz, R. W., Tietje, O., Embedded Case Study Methods. Integrating quantitative and qualitative Knowledge, Sage Publications, Thousand Oaks (CA) 2002
- [Schüngel 1995]
Schüngel, M., Stand der Unternehmensdatenmodellierung in der Praxis - Eine empirische Untersuchung, Diplomarbeit, Westfälischen Wilhelms-Universität Münster, Münster 1995
- [Schwinn 2005]
Schwinn, A., Entwicklung einer Methode zur Gestaltung von Integrationsarchitekturen für Informationssysteme, Dissertation, Universität St. Gallen, Difo-Druck, Bamberg 2005
- [Sciore et al. 1994]
Sciore, E., Siegel, M., Rosenthal, A., Using Semantic Values to Facilitate Interoperability Among Heterogeneous Information Systems, in: ACM Transactions on Database Systems, 19, 1994, Nr. 2, S. 254-290
- [Sen 2004]
Sen, A., Metadata management: past, present and future, in: Decision Support Systems, 37, 2004, Nr. 1, S. 151-173
- [Senger/Österle 2004]
Senger, E., Österle, H., PROMET Business Engineering Case Studies (BECS) Version 2.0, Working Paper BE HSG / BECS / 1, Institute of Information Management, University of St. Gallen, St. Gallen 2004
- [Sfintesco 2009]
Sfintesco, A., Elektronische B2B-Integration in der Logistik - Herausforderungen, Lösungsansätze, Verbesserungspotenziale, Masterarbeit, Universität St. Gallen, St. Gallen 2009

- [Shankaranarayan/Even 2004]
Shankaranarayan, G., Even, A., Managing Metadata in Data Warehouses: Pitfalls and Possibilities, in: Communications of AIS (CAIS), 14, 2004, S. 247-274
- [Shankaranarayanan/Even 2006]
Shankaranarayanan, G., Even, A., The Metadata Enigma, in: Communications of the ACM, 49, 2006, Nr. 2, S. 88-94
- [Shankaranarayanan et al. 2003]
Shankaranarayanan, G., Ziad, M., Wang, R. Y., Managing Data Quality in Dynamic Decision Environments: An Information Product Approach, in: Journal of Database Management, 14, 2003, Nr. 4, S. 14-32
- [Shanks et al. 2003]
Shanks, G., Tansley, E., Weber, R., Using Ontology To Validate Conceptual Models, in: Communications of the ACM (CACM), 46, 2003, Nr. 10, S. 85-89
- [Shanks/Darke 1999]
Shanks, G. G., Darke, P., Understanding corporate data models, in: Information & Management, 35, 1999, S. 19-30
- [Sheth/Larson 1990]
Sheth, A. P., Larson, J. A., Federated Database Systems for Managing Distributed, Heterogeneous, and Autonomous Databases, in: ACM Computing Surveys, 22, 1990, Nr. 3, S. 183-236
- [Siebertz 2009]
Siebertz, C., Zentrales Stammdatenmanagement bei Oerlikon Textile mit SAP NetWeaver MDM, Frankfurt am Main 2009
- [Siegel/Madnick 1991]
Siegel, M., Madnick, S. E., A Metadata Approach to Resolving Semantic Conflicts, Proceedings of the 17th International Conference on Very Large Data Bases (VLDB), Barcelona 1991, S. 133-145
- [Simsion 2007]
Simsion, G., Data Modeling: Theory and Practice, Technics Publications LLC, Bradley Beach (NJ) 2007
- [Sinz 1999]
Sinz, E. J., Architektur von Informationssystemen, in: Rechenberg, P., Pomberger, G. (Hrsg.), Informatik-Handbuch, Carl Hanser, München/Wien 1999, S. 1035-1046
- [Smith/McKeen 2008]
Smith, H. A., McKeen, J. D., Developments in Practice XXX: Master Data Management: Salvation Or Snake Oil?, in: Communications of the AIS (CAIS), 23, 2008, Nr. 4, S. 63-72
- [Smith/Potter 2009]
Smith, M., Potter, K., IT Spending and Staffing Report, 2009, ID Number: G00164940, Gartner, Stamford, USA 2009
- [Sommer 2005]
Sommer, M., Modellierung betrieblicher Informationssysteme, in: Budäus, D. (Hrsg.), Governance von Profit- und Nonprofit-Organisationen in gesellschaftlicher Verantwortung, Deutscher Universitätsverlag/GWV Fachverlage GmbH, Wiesbaden 2005, S. 217-239

- [Souzis 2005]
Souzis, A., Building a Semantic Wiki, in: IEEE Intelligent Systems, 20, 2005, Nr. 5, S. 87--91
- [Spaccapietra et al. 1992]
Spaccapietra, S., Parent, C., Dupont, Y., Model Independent Assertions for Integration of Heterogeneous Schemas, in: The International Journal on Very Large Data Bases (VLDB Journal), 1, 1992, Nr. 1, S. 81-126
- [Spath et al. 2009]
Spath, D., Weisbecker, A., Kokemüller, J., Stammdaten-Managementsysteme 2009: Eine Marktübersicht zu aktuellen Systemen, Fraunhofer-Institut für Arbeitswirtschaft und Organisation (IAO), Fraunhofer Verlag, Stuttgart 2009
- [Sperling 2007]
Sperling, S., Konzeption einer Methode zum Integrationsmanagement bei Unternehmenszusammenschlüssen auf der Basis von multiperspektivischen Unternehmensmodellen, Logos, Berlin 2007
- [Spiegler 2000]
Spiegler, I., Knowledge management: a new idea or a recycled concept?, in: Communications of the AIS, 3, 2000, S. Article 14
- [Stadlbauer 2007]
Stadlbauer, F., Zwischenbetriebliche Anwendungsintegration. IT-Management in Unternehmensnetzwerken, Deutscher Universitätsverlag, Wiesbaden 2007
- [Stahlknecht/Hasenkamp 2002]
Stahlknecht, P., Hasenkamp, U., Einführung in die Wirtschaftsinformatik, 10. Auflage. Aufl., Springer, Berlin 2002
- [Strahringer 1996]
Strahringer, S., Metamodellierung als Instrument des Methodenvergleichs: Eine Evaluierung am Beispiel objektorientierter Analysemethoden, Shaker, Aachen 1996
- [Strauch 2002]
Strauch, B., Entwicklung einer Methode für die Informationsbedarfsanalyse im Data Warehousing, Dissertation, Universität St. Gallen, Difo-Druck, Bamberg 2002
- [Stuhec 2005]
Stuhec, G., SAP Global Data Types Based on CCTS, <https://www.sdn.sap.com/irj/servlet/prt/portal/prtroot/docs/library/uuid/b602d790-0201-0010-e3a8-9e4ddfc45d17>, 10.01.2010
- [Stuhec 2007]
Stuhec, G., Using CCTS Modeler Warp 10 to Customize Business Information Interfaces, SAP AG, 2007
- [Susman/Evered 1978]
Susman, G. I., Evered, R. D., An Assessment of the Scientific Merits of Action Research, in: Administrative Science Quarterly, 23, 1978, Nr. 4, S. 582-603
- [Tannenbaum 2001]
Tannenbaum, A., Metadata Solutions, Addison-Wesley, Boston et al. 2001
- [ter Hofstede/Verhoef 1997]
ter Hofstede, A. H. M., Verhoef, T. F., On the Feasibility of Situational Method Engineering, in: Information Systems, 22, 1997, Nr. 6/7, S. 401-422

- [Thangarathinam et al. 2004]
Thangarathinam, T., Wyant, G., Gibson, J., Simpson, J., Metadata Management: the Foundation for Enterprise Information Integration, in: Intel Technology Journal, 8, 2004, Nr. 4, S. 337-344
- [Thome/Sollbach 2007]
Thome, G., Sollbach, W., Grundlagen und Modelle des Information Lifecycle Management, Springer, Berlin 2007
- [Topi/Ramesh 2002]
Topi, H., Ramesh, V., Human Factors on Research on Data Modeling: A Review of Prior Research, An Extended Framework and Future Research Directions, in: Journal of Database Management, 13, 2002, Nr. 2, S. 3-19
- [Tozer 1999]
Tozer, G., Metadata Management for Information Control and Business Success, Artech House, Norwood (MA) 1999
- [Truong/Dustdar 2009]
Truong, H.-L., Dustdar, S., On Analyzing and Specifying Concerns for Data as a Service, Proceedings of the 2009 IEEE Asia-Pacific Services Computing Conference (APSCC), Singapore 2009, S. 87-94
- [Tuomi 1999]
Tuomi, I., Data Is More Than Knowledge: Implications of the Reversed Knowledge Hierarchy for Knowledge Management and Organizational Memory, in: Journal of Management Information Systems, 16, 1999, Nr. 3, S. 103-117
- [Ulrich 1984]
Ulrich, H., Die Betriebswirtschaftslehre als anwendungsorientierte Sozialwissenschaft, in: Ulrich, H., Dyllick, T., Probst, G. (Hrsg.), Management, Haupt, Bern 1984, S. 170-195
- [UN/CEFACT 2009a]
UN/CEFACT, Core Component Library (UN/CCL) - Version 09A, <http://www.unece.org/cefact/codesfortrade/unccl/CCL09A.xls>, 01.11.2009
- [UN/CEFACT 2009b]
UN/CEFACT, Core Components Technical Specification, Version 3.0, <http://www.unece.org/cefact/codesfortrade/CCTS/CCTS-Version3.pdf>, 05.10.2009
- [Uschold/Gruninger 2004]
Uschold, M., Gruninger, M., Ontologies and Semantics for Seamless Connectivity, in: SIGMOD Record, 33, 2004, Nr. 4, S. 58-64
- [Vetere/Lenzerini 2005]
Vetere, G., Lenzerini, M., Models for semantic interoperability in service-oriented architectures, in: IBM Systems Journal, 44, 2005, Nr. 4, S. 887-903
- [Vogel 2009]
Vogel, T., Serviceorientiertes Business Networking - Referenzarchitektur und Gestaltungsprinzipien, Dissertation, Universität St. Gallen, Difo-Druck, Bamberg 2009

- [Vogel et al. 2007]
Vogel, T., Legner, C., Au, C., Augenstein, C., Löhe, J., Schnabel, F., Wittmer, J., SOA for Automotive: Konzept m:n-fähiger Web Services für das überbetriebliche Änderungsmanagement, BE HSG / CC BN3 / Automotive 1, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen 2007
- [Vogel et al. 2008]
Vogel, T., Schmidt, A., Lemm, A., Österle, H., Service and Document Based Interoperability for European eCustoms Solutions, in: Journal of Theoretical and Applied Electronic Commerce Research (JTAER), 3, 2008, Nr. 3, S. 17-37
- [Vogler 2004]
Vogler, P., Prozess- und Systemintegration: Evolutionäre Weiterentwicklung bestehender Informationssysteme mit Hilfe von Enterprise Application Integration, Habilitation, Universität St. Gallen, St. Gallen 2004
- [Voß/Gutenschwager 2001]
Voß, S., Gutenschwager, K., Informationsmanagement, Springer, Berlin 2001
- [Vossen 1999]
Vossen, G., Datenbankmodelle, Datenbanksprachen und Datenbankmanagement-Systeme, 3. Aufl., Oldenbourg, München 1999
- [Wagter et al. 2005]
Wagter, R., van den Berg, M., Luijpers, J., van Steenbergen, M., Dynamic Enterprise Architecture: How to Make It Work, John Wiley & Sons, Hoboken (NJ) 2005
- [Wand/Wang 1996]
Wand, Y., Wang, R. Y., Anchoring data quality dimensions in ontological foundations, in: Communications of the ACM (CACM), 39, 1996, Nr. 11, S. 86-95
- [Wand/Weber 2002]
Wand, Y., Weber, R., Research Commentary: Information Systems and Conceptual Modeling - A Research Agenda, in: Information Systems Research (ISR), 13, 2002, Nr. 4, S. 363-376
- [Wang 1998]
Wang, R. Y., A Product Perspective on Total Data Quality Management, in: Communications of the ACM (CACM), 41, 1998, Nr. 2, S. 58-65
- [Wang/Madnick 1989]
Wang, R. Y., Madnick, S. E., The Inter-Database Instance Identification Problem in Integrating Autonomous Systems, Proceedings of the Fifth International Conference on Data Engineering, IEEE Computer Society, Washington, D.C. 1989, S. 46-55
- [Wang/Strong 1996]
Wang, R. Y., Strong, D. M., Beyond Accuracy: What Data Quality Means to Data Consumers, in: Journal of Management Information Systems, 12, 1996, Nr. 4, S. 5-34
- [Wang et al. 2000]
Wang, R. Y., Ziad, M., Lee, Y. W., Data Quality, Springer, Boston (MA) 2000

[Weber 2009]

Weber, K., Data Governance-Referenzmodell - Organisatorische Gestaltung des unternehmensweiten Datenqualitätsmanagements, Dissertation, Universität St. Gallen, Difo-Druck, Bamberg 2009

[Weber et al. 2009]

Weber, K., Otto, B., Österle, H., Data Governance: Organisationskonzept für das konzernweite Datenqualitätsmanagement, in: Hansen, H.R., Karagiannis, D., Fill, H.-G. (Hrsg.), Business Services: Konzepte, Technologien, Anwendungen; Proceedings der 9. Internationalen Tagung Wirtschaftsinformatik, Band 1, Österreichische Computer Gesellschaft, Wien 2009, S. 589-598

[Wedekind 2001]

Wedekind, H., Bestandsdaten, Bewegungsdaten, Stammdaten, in: Mertens, P., Back, A., Becker, J., König, W., Krallmann, H., Rieger, B., Scheer, A.-W., Seibt, D., Stahlknecht, P., Strunz, H., Thome, R., Wedekind, H. (Hrsg.), Lexikon der Wirtschaftsinformatik, Springer, Berlin 2001, S. 72

[Weigand/van den Heuvel 2002]

Weigand, H., van den Heuvel, W.-J., Cross-organizational workflow integration using contracts, in: Decision Support Systems (DSS), 33, 2002, Nr. 3, S. 247-265

[Welge/Al-Laham 2008]

Welge, M. K., Al-Laham, A., Strategisches Management: Grundlagen - Prozess - Implementierung, 3. Aufl., Gabler, Wiesbaden 2008

[Wende et al. 2009]

Wende, K., Schmidt, A., Hüner, K., Ofner, M., Data Quality Management - Definitions and State of the Art, Working Paper BE HSG/ CC CDQ/ 1, Institute of Information Management, University of St. Gallen, St. Gallen 2009

[White 2008]

White, A., Governance of Master Data Starts With the Master Data Life Cycle, Gartner Group, Stamford (CT) 2008

[White et al. 2006]

White, A., Newman, D., Logan, D., Radcliffe, J., Mastering Master Data Management, Gartner Group, Stamford (CT) 2006

[White et al. 2008]

White, A., Radcliffe, J., Steenstrup, K., Bitterer, A., Beyer, M. A., Wilson, D., Rayner, N., Chandler, N., Newman, D., Hype Cycle for Master Data Management 2008, Gartner Group, Stamford (CT) 2008

[White et al. 2009]

White, A., Radcliffe, J., Wilson, D. R., Vendor Guide: Master Data Management, Gartner Group, Stamford (CT) 2009

[Wilde/Hess 2007]

Wilde, T., Hess, T., Forschungsmethoden der Wirtschaftsinformatik. Eine empirische Untersuchung, in: Wirtschaftsinformatik, 49, 2007, Nr. 4, S. 280-287

[Winter 2003]

Winter, R., Modelle, Techniken und Werkzeuge im Business Engineering, in: Österle, H., Winter, R. (Hrsg.), Business Engineering, Springer, Berlin etc. 2003, S. 87-117

[Winter 2004]

Winter, R., Architektur braucht Management, in: Wirtschaftsinformatik, 46, 2004, Nr. 4, S. 317-319

[Winter 2009]

Winter, R. (Hrsg.), Management von Integrationsprojekten - Konzeptionelle Grundlagen und Fallstudien aus fachlicher und IT-Sicht, Berlin 2009

[Winter/Fischer 2007]

Winter, R., Fischer, R., Essential Layers, Artifacts, and Dependencies of Enterprise Architecture, in: Journal of Enterprise Architecture, 3, 2007, Nr. 2, S. 7-18

[Wise 2008]

Wise, L., The Intrinsic Value of Master Data Management, in: DM Review, 18, 2008, Nr. 7, S. 8-10

[Wu et al. 1997]

Wu, B., Lawless, D., Bisbal, J., Richardson, R., Grimson, J., Wade, V., O'Sullivan, D., The Butterfly Methodology: A Gateway-free Approach for Migrating Legacy Information Systems, Proceedings of the 3rd IEEE Conference on Engineering of Complex Computer Systems (ICECCS97), IEEE Computer Society, Como 1997, S. 200-205

[Yin 2002]

Yin, R. K., Case Study Research. Design and Methods, 3. Aufl., Sage Publications, London 2002

[Yu 2007]

Yu, H., Context Driver Principle, Diploma, University Karlsruhe, Karlsruhe 2007

[Ziegler/Dittrich 2004]

Ziegler, P., Dittrich, K. R., Three Decades of Data Integration - All Problems Solved?, Proceedings of the IFIP 18th World Computer Congress (WCC 2004), Toulouse 2004, S. 3-12

[ZVEI 2006]

ZVEI, Klassifizierung und Produktbeschreibung in der Elektrotechnik- und Elektronikindustrie, Zentralverband Elektrotechnik- und Elektronikindustrie e.V. (ZVEI), Frankfurt am Main 2006

Lebenslauf

Persönliche Daten

Geburtsort Dresden (Deutschland)
Nationalität Deutsch

Ausbildung

2000 – 2006 Technische Universität Dresden (Deutschland)
Studium der Wirtschaftsinformatik
2002 – 2003 Ecole des Sciences Commerciales d’Angers, Angers (Frankreich)
Auslandsstudium, Strategisches Management
2006 – 2010 Universität St. Gallen (Schweiz)
Doktorandenstudium der Wirtschaftswissenschaften

Berufstätigkeit

2002 Robert Bosch GmbH, Stuttgart (Deutschland)
Praktikum
2004 Lehrstuhl für Wirtschaftsinformatik, insbesondere Systementwicklung, Prof. Esswein, Dresden (Deutschland)
Studentischer Mitarbeiter
2005 – 2006 Lehrstuhl für Marktorientierte Unternehmensführung, Prof. Töpfer, Dresden (Deutschland)
Studentischer Mitarbeiter
2006 – 2010 SAP Research, St. Gallen (Schweiz)
Research Assistant
2006 – 2010 Institut für Wirtschaftsinformatik, Universität St. Gallen (Schweiz),
Lehrstuhl Prof. Dr. Hubert Österle
Wissenschaftlicher Mitarbeiter