



Thema:

**Fachkonzeptuelle Entwicklung eines Informationssystems zur Unterstützung des
Präqualifikationsprozesses von kleinen und mittelständigen Unternehmen des
Maschinen- und Anlagenbaus**

Diplomarbeit

Arbeitsgruppe Wirtschaftsinformatik

Themensteller: Prof. Dr. Hans-Knud Arndt, Wirtschaftsinformatik III (FIN/ITI)

Betreuer: Dipl. -Wirtsch. -Inf. Stefan Breitenfeld, Wirtschaftsinformatik III
(FIN/ITI)

Vorgelegt von: Falk Ziolkowski

Abgabetermin: 07.11.07

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Verzeichnis der Abkürzungen und Akronyme	III
Abbildungsverzeichnis	V
Tabellenverzeichnis	VII
1 Einleitung	1
1.1 Problemstellung	1
1.2 Zielstellung	3
1.3 Vorgehensweise und Struktur der Arbeit	4
2 Grundlagen	8
2.1 Computerunterstützung für das Management	8
2.1.1 Information und Wissen	8
2.1.2 Informationssystem	9
2.1.3 Informationsbedarf, -angebot und -nachfrage	14
2.1.4 Informationsbedarfsanalyse	15
2.1.5 Management	17
2.1.6 Managementinformationssysteme	20
2.2 Enterprise Application Integration	24
2.2.1 Vorgehensweisen der Integration	27
2.2.2 Föderierte Informationssysteme	29
2.3 Computerlinguistik	31
2.3.1 Texttechnologie	32
2.3.2 Textkorpora	33
2.3.3 Information Retrieval und Information Extraction	34
3 Anforderungsanalyse	37
3.1 Ist-Analyse des Beantwortungsprozesses	37
3.2 Analyse der Präqualifikationsfragebögen	41
3.2.1 Vorbereitung der Analyse	41
3.2.2 Vorgehensweise der Analyse	44
3.2.3 Ergebnisse der Analyse	45
3.2.3.1 Fragebogenbereich: allgemeine Unternehmensdaten	47
3.2.3.2 Fragebogenbereich: Referenzen	48
3.2.3.3 Fragebogenbereich: Leistungsspektrum	49
3.2.3.4 Fragebogenbereich: Qualitätssicherung	50
3.2.3.5 Fragebogenbereich: Sicherheit, Gesundheit und Umwelt ..	51
3.2.3.6 Fragebogenbereich: Ansprechpartner	53

3.2.3.7	Fragebogenbereich: Mitarbeiter	54
3.2.3.8	Nachweis-Dokumente	54
3.3	Zusammenfassung und Schlussfolgerungen	55
4	Fachkonzept des Informationssystems	59
4.1	ARIS-Konzept	59
4.2	Soll-Prozess der Beantwortung eines Präqualifikationsfragebogens	62
4.2.1	Digitalisierung	63
4.2.1.1	Dokumentscan	64
4.2.1.2	OCR-Erkennung	65
4.2.1.3	Attribuierung	67
4.2.2	Vorverarbeitung	69
4.2.3	Semi-automatische Beantwortung	71
4.2.3.1	Antwortgenerierung	71
4.2.3.2	PQF-Antwortbearbeitung	73
4.2.3.3	PQF-Prüfung	76
4.2.4	Datenschema des Präqualifikationsfragebogens	77
4.3	Schemata der Datenbasis und Teilsysteme	78
4.3.1	Föderiertes Datenschema	78
4.3.2	Unternehmensverwaltung	81
4.3.3	Statistikverwaltung	84
4.3.4	Referenzverwaltung	86
4.3.5	Managementsystemverwaltung	88
5	Zusammenfassung und Ausblick	91
A	Tabellen der Informationsbedarfsanalyse	97
B	Abbildungen	108
	Literaturverzeichnis	111

Verzeichnis der Abkürzungen und Akronyme

ANSI	American National Standard Institut
ASME	American Society of Mechanical Engineers
BMVBS	Bundesministerium für Verkehr, Bau und Stadtentwicklung
CI	Coded Information
CIS	Chefinformationssystem
CL	Computerlinguistik
CO	Co-Reference Resolution
D&B	Dunn & Bradstreet
DBMS	Datenbankmanagementsysteme
DIN	Deutsches Institut für Normung e. V.
DMS	Dokumentenmanagementsystem
dpi	dots per inch
DSS	Decision Support System
DTD	Dokumenttyp-Definition
DV	Datenverarbeitung
EAI	Enterprise Application Integration
EDI	Electronic Data Interchange
EDV	elektronische Datenverarbeitung
eEPK	erweiterte ereignisgesteuerte Prozesskette
eERM	erweitertes Entity-Relationship-Modell
EIS	Executive Information System
EN	Europäische Norm
ES	Expertensystem
ESS	Executive Support System
EUS	Entscheidungsunterstützungssystem
FIS	Führungsinformationssystem
GATE	General Architecture for Text Engineering
GDSS	Group Decision Support System
HSE	Health Safety Environment
ICR	Intelligent Character Recognition
IE	Information Extraction
ISO	International Organization for Standardization
IT	Informationstechnik
kmU	kleine und mittelständige Unternehmen
MIS	Managementinformationssystem
MRS	Management Reporting System
MS	Managementsystem
MSS	Management Support System

MUS	Managementunterstützungssystem
NCI	None-coded Information
NE	Named Entity Recognition
OAS	Office Automation System
OCR	Optical Character Recognition
OHSAS	Occupational Health and Safety Assessment Series
OLAP	OnLine Analytical Processing
PQ	Präqualifikation
PQ-Verfahren	Präqualifikationsverfahren
PQF	Präqualifikationsfragebogen
QM	Qualitätsmanagement
SCC	Safety Certificate Contractors
ST	Scenario Templates
TE	Template Element Filling
TEI	Text Encoding Initiative
TIF	Tagged Image File
TPS	Transaction Processing System
TR	Template Relations
TRD	Technische Regeln für Dampfkessel
TVS	Transaktionsverarbeitungssystem
VDA	Verband der Automobilindustrie e. V.
VIS	Vorstandsinformationssystem
VOB/A	Vergabe- und Vertragsordnung für Bauleistungen Teil A
WYSIWYG	What You See Is What You Get
XML	eXtensible Markup Language
XPS	Expert System
XSS	Expert Support System

Abbildungsverzeichnis

Abb. 1.1:	Darstellung des Beantwortungsprozesses	2
Abb. 1.2:	Aufbau der Arbeit	6
Abb. 2.1:	Beziehung zwischen einem System und seiner Umwelt	10
Abb. 2.2:	Morphologischer Kasten: Systemeigenschaften	11
Abb. 2.3:	Modell computergestützter Informationssysteme	13
Abb. 2.4:	Beziehung zwischen Informationsangebot, Informationsnachfrage und Informationsbedarf	14
Abb. 2.5:	Unterscheidung von Ausführungs- und Managementsystem im Regelkreis und deren Verschachtelung	18
Abb. 2.6:	Begriffshierarchie computergestützter Managementinformationssysteme	21
Abb. 2.7:	Historische Entwicklung von Managementinformationssystemen . . .	22
Abb. 2.8:	Arten der Heterogenität	26
Abb. 2.9:	Allgemeine Architektur föderierter Informationssysteme	29
Abb. 2.10:	Fünf-Ebenen-Schema-Architektur	31
Abb. 3.1:	Ist-Analyse der Beantwortung eines Präqualifikationsfragebogens . . .	39
Abb. 3.2:	Struktur der Präqualifikationsfragebögen	45
Abb. 3.3:	Bereiche der Präqualifikationsfragebögen	46
Abb. 3.4:	Informationsbedarf des Bereichs allgemeinen Unternehmensdaten . . .	47
Abb. 3.5:	Informationsbedarf des Bereichs Referenzen	48
Abb. 3.6:	Informationsbedarf des Bereichs Leistungsspektrum	50
Abb. 3.7:	Informationsbedarf des Bereichs Qualitätssicherung	50
Abb. 3.8:	Informationsbedarf des Bereichs HSE	52
Abb. 3.9:	Informationsbedarf des Bereichs Ansprechpartner	53
Abb. 3.10:	Informationsbedarf des Bereichs Mitarbeiter	54
Abb. 3.11:	Informationsbedarf der Dokumente	55
Abb. 4.1:	ARIS-Haus mit Phasenkonzept	61
Abb. 4.2:	Übersicht des Beantwortungsprozesses der Präqualifikationsfragebögen	63
Abb. 4.3:	Teilprozesse der Digitalisierung	63
Abb. 4.4:	Teilprozess Dokumentscan	64
Abb. 4.5:	Teilprozess OCR-Erkennung	66
Abb. 4.6:	Teilprozess Attribuierung	68
Abb. 4.7:	Teilprozess Vorverarbeitung	70
Abb. 4.8:	Übersicht des Teilprozesses Beantwortung	71
Abb. 4.9:	Teilprozess Antwortgenerierung	72

Abb. 4.10: Teilprozess PQF-Antwortbearbeitung	74
Abb. 4.11: Teilprozess PQF-Prüfung	76
Abb. 4.12: Datenschema des Präqualifikationsfragebogens	77
Abb. 4.13: Föderiertes Datenschema	80
Abb. 4.14: Funktionsübersicht der Unternehmensverwaltung	81
Abb. 4.15: Datenschema der Unternehmensverwaltung	82
Abb. 4.16: Dokumentenmodell	83
Abb. 4.17: Funktionsübersicht der Statistikverwaltung	85
Abb. 4.18: Datenschema der Statistikverwaltung	85
Abb. 4.19: Funktionsübersicht der Referenzverwaltung	86
Abb. 4.20: Datenschema der Referenzverwaltung	88
Abb. 4.21: Funktionsübersicht der MS-Verwaltung	89
Abb. 4.22: Datenschema der MS-Verwaltung	90
Abb. 5.1: Übersicht des Fachkonzepts	92
Abb. B.1: Attributzuordnungsdiagramme der Präqualifikationsfragebögen	108
Abb. B.2: Attributzuordnungsdiagramme der Unternehmensverwaltung	108
Abb. B.3: Attributzuordnungsdiagramme der Dokumente	109
Abb. B.4: Attributzuordnungsdiagramme der Statistikverwaltung	109
Abb. B.5: Attributzuordnungsdiagramme der Referenzverwaltung	110
Abb. B.6: Attributzuordnungsdiagramme der MS-Verwaltung	110

Tabellenverzeichnis

Tab. 2.1: Übersicht zu Begriffen und Akronymen computerbasierter Informationssysteme	23
Tab. 3.1: Einteilung der Präqualifikationsfragebögen	42
Tab. A.1: Bereiche der Präqualifikationsfragebögen	97
Tab. A.2: Allgemeine Unternehmensdaten	98
Tab. A.4: Informationsbedarf des Bereichs Referenzen	98
Tab. A.5: Informationsbedarf des Bereichs Leistungsspektrum	99
Tab. A.6: Informationsbedarf des Bereichs Qualitätssicherung	99
Tab. A.7: Informationsbedarf des Bereichs HSE	100
Tab. A.8: Informationsbedarf des Bereichs Ansprechpartner	100
Tab. A.9: Informationsbedarf des Bereichs Mitarbeiter	101
Tab. A.10: Informationsbedarf der Dokumente	101
Tab. A.11: Informationsbedarf des Bereichs kommerzielle Aspekte	102
Tab. A.12: Informationsbedarf des Bereichs Versicherungen	102
Tab. A.13: Informationsbedarf des Bereichs Logistik	103
Tab. A.14: Informationsbedarf des Bereichs Garantie	103
Tab. A.15: Informationsbedarf des Bereichs Informationstechnologie	104
Tab. A.16: Informationsbedarf des Bereichs Gebäude	104
Tab. A.17: Informationsbedarf des Bereichs Subunternehmen	105
Tab. A.18: Informationsbedarf des Bereichs Handelsregister	105
Tab. A.19: Informationsbedarf des Bereichs Firmenstrategie	105
Tab. A.20: Informationsbedarf des Bereichs Marktpositionierung	106
Tab. A.21: Informationsbedarf des Bereichs Herstellung	106
Tab. A.22: Informationsbedarf des Bereichs Entwicklung	106
Tab. A.23: Informationsbedarf des Bereichs Management: Planung, Beschaffung und Controlling	107
Tab. A.24: Informationsbedarf des Bereichs Angebotsaktivität	107

1 Einleitung

1.1 Problemstellung

In verschiedenen Branchen bspw. der Automobilbranche, im Maschinenbau, der Chemiebranche oder dem Bauwesen werden Aufträge oder Projekte nur an geeignete Unternehmen vergeben. Diese Eignung eines Unternehmens wird durch die Auftraggeber über eine Lieferantenbewertung festgestellt. Dafür existieren verschiedene Verfahren, auf die an dieser Stelle nicht eingegangen wird. Neben der Kalkulation eines Angebots müssen die Unternehmen dem potenziellen Auftraggeber explizit gefragte unternehmensspezifische Daten (im Folgenden auch als Informationsbedarf bezeichnet) im Sinne einer Präqualifikation (PQ) zur Verfügung stellen. Dafür werden Präqualifikationsfragebögen verwendet. Eine Präqualifikation ist nach dem Bundesministerium für Verkehr, Bau und Stadtentwicklung (BMVBS) die „vorgelagerte auftragsunabhängige Prüfung der Eignungsnachweise [nach bestimmten Kriterien, die] auf Basis der in § 8 VOB/A definierten Anforderungen“ (BMVBS (2006a), S. 4) festgelegt wurden. Diese in der Vergabe- und Vertragsordnung für Bauleistungen Teil A (VOB/A) beschriebenen Anforderungen gelten für öffentliche Ausschreibungen und beziehen sich auf Eignungsnachweise in den Bereichen Fachkunde, Leistungsfähigkeit und Zuverlässigkeit des zu qualifizierenden Unternehmens. Dabei dürfen Nachweise und Informationen zum Umsatz, der Ausführung von Leistungen sowie die durchschnittliche Anzahl der Mitarbeiter über einen Zeitraum von drei Geschäftsjahren verlangt werden (BMVBS (2006b), S. 19 f.). Das Präqualifikationsverfahren (PQ-Verfahren) des Maschinen- und Anlagenbaus ist mit der Leitlinie des BMVBS vergleichbar.

Es gibt verschiedene Arten der Präqualifikation. Allgemein kann eine Präqualifikation zwischen unterschiedlichen Unternehmen, zwischen Unternehmen und Behörden sowie zwischen Unternehmen und Zertifizierungsstellen durchgeführt werden. Die Qualifikation für eine öffentliche Ausschreibung wird über eine separate Zertifizierungsstelle realisiert. Im Maschinen- und Anlagenbau sind PQ-Verfahren nicht standardisiert. Die meisten Bauvorhaben im Maschinen- und Anlagenbau werden in Form von Projekten durchgeführt. Dabei durchlaufen diese Projekte die folgenden vier Phasen: Anbahnungsphase, Definitionsphase, Abwicklungsphase und Abschlussphase (vgl. Thurow et al. (2006), S. 67). In der Anbahnungsphase versuchen Unternehmen Projekte bei Auftraggebern zu akquirieren bzw. Auftraggeber schreiben Projekte aus, um die sich Unternehmen bewerben. Ein Auftraggeber kann in diesem Zusammenhang eine Präqualifikation mit seinen potenziellen Lieferanten (Subunternehmen) durchführen und sich damit eine Lieferantenliste aufbauen (vgl. Thurow et al. (2006), S. 68). Dabei werden vom Auftraggeber Präqualifikationsfragebögen an die Subunternehmen geschickt, um Unternehmensdaten zu erfassen.

Für kleine und mittelständige Unternehmen (kmU) im Maschinen und Anlagenbau ist die Teilnahme an einem Präqualifikationsverfahren von existenzieller Bedeutung, weil nur durch diesen Eignungsnachweis eine Projektvergabe durch den potenziellen Auftraggeber erfolgen kann. Aus strategischer Sicht kann ein kmU durch die Bearbeitung von Projekten Gewinne erzielen und dadurch am Markt bestehen. Deshalb wird die Beantwortung der Präqualifikationsfragebögen vom Führungspersonal der kmU durchgeführt. Die Abteilungen, die für die Präqualifikation verantwortlich sind, variieren in den Unternehmen. Beispielsweise kann die Präqualifikation dem Einkauf, dem Qualitätsmanagement oder der Projektleitung untergeordnet werden. Mit den Präqualifikationsfragebögen können die Unternehmen vor und während der Angebotserstellung konfrontiert werden. Dabei erhält das Unternehmen neben der Anfrage zu einem Projekt zusätzlich ein Präqualifikationsfragebogen. Beide Dokumente werden nach der Bearbeitung wieder an den Auftraggeber gesendet. Der ausgefüllte Präqualifikationsfragebogen und das Angebot fließen beim Auftraggeber in den Prozess der Lieferantenbewertung ein.

Die Tätigkeiten, die ein kmU durchführen muss, um alle Unterlagen und Daten zusammenzustellen, insbesondere um den Präqualifikationsfragebogen auszufüllen, sind sehr zeitaufwendig und kostenintensiv. In Abb. 1.1 sind diese Tätigkeiten grob dargestellt.

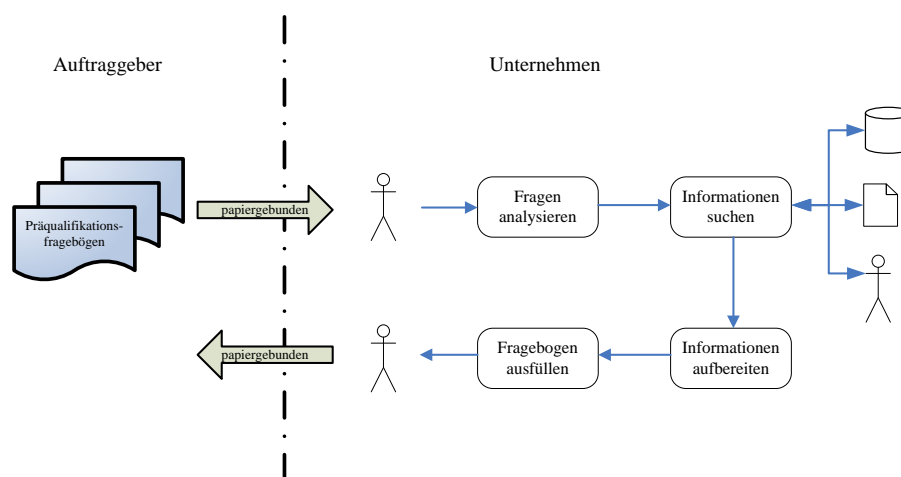


Abb. 1.1: Darstellung des Beantwortungsprozesses

Zunächst werden die Fragen analysiert, um zu erfahren welche Informationen für die Antworten benötigt werden. Das Problem dabei ist, dass die für den Beantwortungsprozess benötigten Informationen nicht zentral zur Verfügung stehen und zum Teil nicht digital vorliegen. Die benötigten Informationen betreffen unterschiedliche Bereiche und damit unterschiedliche Verantwortlichkeiten im Unternehmen. Auf Grund der verteilten Informationen im Unternehmen und der heterogenen Systemlandschaft,

z. B. Buchhaltungsprogramme, Kalkulationsprogramme, ERP-Systeme, verschiedene Datenbanken und Dokumentenablagen im Netzwerk oder auf Papier, bedarf es einer umfassenden und zeitaufwendigen Suche nach den benötigten aktuellen Informationen. Wenn der Bearbeiter die Informationen gefunden hat, müssen diese aufbereitet werden, bspw. wird eine Auflistung über die Anzahl der Mitarbeiter in Abhängigkeit ihrer Qualifikationen benötigt. Der Informationsbedarf des Auftraggebers kann unterschiedliche Granularitäten aufweisen. Es gibt Präqualifikationsfragebögen, in denen die Informationsangabe, z. B. Qualifikationen oder Umsätze etc., ohne Nachweis-Dokument erfolgt. In anderen Präqualifikationsfragebögen müssen die Angaben mit entsprechenden Dokumenten nachgewiesen werden. Der Nachweis durch Dokumente hängt vom Informationsbedürfnis des anfragenden Unternehmens ab und ist nicht vorhersagbar. Diese Dokumente müssen dann als Kopie dem Präqualifikationsfragebogen beigelegt werden. Nachdem der Bearbeiter alle Dokumente kopiert und alle Informationen beschafft hat, kann er den Präqualifikationsfragebogen ausfüllen und an den potenziellen Auftraggeber zurückschicken.

Den Aufwand zur Teilnahme am Präqualifikationsverfahren trägt das Unternehmen als potenzieller Auftragnehmer selbst. Dabei besteht die Gefahr, den Auftrag nicht zu erhalten. Auf Grund des Wettbewerbsdrucks muss das Unternehmen Möglichkeiten und Wege finden, wie die Angebotserstellungsprozesse und somit auch der Präqualifikationsprozess auf Seiten des Unternehmens verkürzt und die Kosten verringert werden können.

Im Rahmen eines Industrieprojektes des Fraunhofer IFF wurde von verschiedenen Unternehmen der Wunsch nach einer automatischen Beantwortung dieser zum Teil wiederkehrenden Präqualifikationsfragebögen geäußert. Der Beantwortungsprozess sollte durch ein Informationssystem unterstützt werden, um den Zeitaufwand und die Kosten dieses Prozesses zu senken, da es nicht sicher ist, ob sich durch diese Bemühungen ein kostendeckender Auftrag ergibt.

1.2 Zielstellung

Das Ziel dieser Diplomarbeit ist die fachkonzeptionelle Entwicklung eines Informationssystems zur Unterstützung des Managements bei der Präqualifikation ihres Unternehmens. Dabei liegt der Focus auf der semi-automatischen Beantwortung der Präqualifikationsfragebögen. Um eine Automatisierung zu erreichen ist es notwendig den Informationsbedarf eines Präqualifikationsfragebogens zu erkennen. Die Grundlage für die Beantwortung wird durch eine Analyse des Informationsbedarfs

verschiedener Präqualifikationsfragebögen erarbeitet. Darauf aufbauend wird ein Informationsmodell erstellt. Dieses Informationsmodell stellt den Informationsbedarf der Präqualifikationsfragebögen und damit das Informationsangebot des zu konzipierenden Informationssystems dar. Auf Grundlage dieses Informationsmodells kann die semi-automatische Beantwortung der Präqualifikationsfragebögen erfolgen. Dadurch wird der Bearbeitungsaufwand reduziert, das Kosteneinsparungspotenzial dieses Prozesses ausgeschöpft sowie die Interoperabilität zwischen verschiedenen Unternehmen erhöht und beschleunigt. Die dazu notwendigen Erkenntnisse zum Funktionsumfang und Informationsangebot des zu entwickelnden Informationssystems, werden durch die Analyse des Ist-Zustandes des Beantwortungsprozesses ermittelt.

Die konkreten Ausprägungen der Informationen in dem zu entwickelnden Modell sind branchen- bzw. unternehmensspezifisch, nicht aber das Informationsmodell selbst. Demzufolge wird darauf geachtet das Informationssystem so abstrakt zu gestalten, dass es in verschiedenen Branchen, die ein Verfahren zur Präqualifikation einsetzen, verwendet werden kann. Deshalb wird in dieser Arbeit am Beispiel der Präqualifikation von kmU im Maschinen- und Anlagenbau ein allgemeines Informationssystem zur Unterstützung der Beantwortung von Präqualifikationsfragebögen entwickelt. Dabei werden branchen- und unternehmensspezifische Eigenheiten abstrahiert. Die Betrachtung dieser spezifischen Eigenschaften, Fachbegriffe und Synonyme sind nicht primär Bestandteil dieser Arbeit.

1.3 Vorgehensweise und Struktur der Arbeit

Die Vorgehensweise orientiert sich an den Phasen des allgemeinen Softwarelebenszyklus. Dieses lineare Phasenmodell unterteilt sich in Problemstellung, Anforderungsanalyse, Entwurf, Implementation, Test, Einführung, Betrieb und Wartung (vgl. Ludewig und Lichter (2007), S. 146 f.; Stickel (2001), S. 113).

Zunächst wird auf Basis der Problembeschreibung eine Anforderungsanalyse durchgeführt. Ziel ist die Konzeption eines Informationssystems zur Unterstützung des Beantwortungsprozesses der Präqualifikationsfragebögen. Daraus leiten sich für die Anforderungsanalyse die Ist-Analyse des Beantwortungsprozesses und die Analyse des Informationsbedarfs ab.

Mit der Aufnahme des Ist-Prozesses und der Analyse der Probleme und Anforderungen wird das Soll-Konzept für die Prozessgestaltung des Informationssystems entwickelt. Durch eine Informationsbedarfsanalyse werden Erkenntnisse über das zu konzipierende Informationsangebot des Informationssystems ermittelt. Das Soll-Konzept und die Modellierung des Informationsangebots sind Teil der Entwurfsphase, in der die fachlichen

Anforderungen an das zu konzipierende Informationssystem beschrieben werden. Die im Anschluss daran folgenden Phasen des Softwarelebenszyklus sind nicht Bestandteil dieser Arbeit.

Die Anforderungsanalyse gliedert sich in zwei Teile. Der erste Teil, die Erarbeitung des Ist-Prozesses und die Ist-Analyse, erfolgte in einem Workshop mit vier kmU des Maschinen- und Anlagenbaus. Für die Informationsbedarfsanalyse wurden Präqualifikationsfragebögen beschafft und inhaltlich analysiert. Die Entwicklung des Fachkonzeptes erfolgt unter Verwendung der ARIS Vorgehensweise, auf welche an späterer Stelle genauer eingegangen wird.

Aus der beschriebenen Vorgehensweise leitet sich die Struktur dieser Arbeit ab. Zusammenfassend wird diese Struktur in der Abb. 1.2 dargestellt. Die vorliegende Arbeit gliedert sich in fünf Kapitel. Das erste Kapitel beschreibt die Problemstellung, das Ziel, die Vorgehensweise und die Struktur der Arbeit. Aufbauend auf der Problemstellung und der Vorgehensweise werden im Kapitel 2 die Grundlagen für das zu entwickelnde Fachkonzept durch eine Literaturrecherche erarbeitet.

Das Kapitel 2 gliedert sich inhaltlich in drei Abschnitte. Der Abschnitt 2.1 beschäftigt sich, auf Grund der Partizipation des Managements am Beantwortungsprozess, mit dem Forschungsfeld Computerunterstützung für das Management. Dabei werden die wesentlichen Grundbegriffe Information, Wissen, Informationssystem, Management und Managementinformationssysteme behandelt. Des Weiteren werden die Begriffe Informationsbedarf, -angebot und -nachfrage sowie die Methoden zur Informationsbedarfsanalyse beschrieben. Diese Erkenntnisse werden im Kapitel 3 wieder aufgegriffen. Auf Grund der zu erwartenden heterogenen Systemlandschaft, in der die benötigten Informationen verwaltet werden, widmet sich der Abschnitt 2.2 dem Forschungsfeld des Enterprise Application Integration. Der Focus dieses Abschnittes liegt auf den Vorgehensweisen zur Integration und dem Konzept der föderierten Informationssysteme. Um eine semi-automatische Beantwortung der Präqualifikationsfragebögen zu ermöglichen, sind semantische Analysen notwendig. Das Forschungsfeld, welches sich mit natürlichsprachlichen Phänomenen und deren Bearbeitung auseinandersetzt, ist die Computerlinguistik und wird im Abschnitt 2.3 dieser Arbeit behandelt. Es wird auf die Teilgebiete Texttechnologie, Textkorpora, Information Retrieval und Information Extraction eingegangen, um Methoden für das Fachkonzept des Informationssystems und die semi-automatische Erschließung der Präqualifikationsfragebögen zu adaptieren.

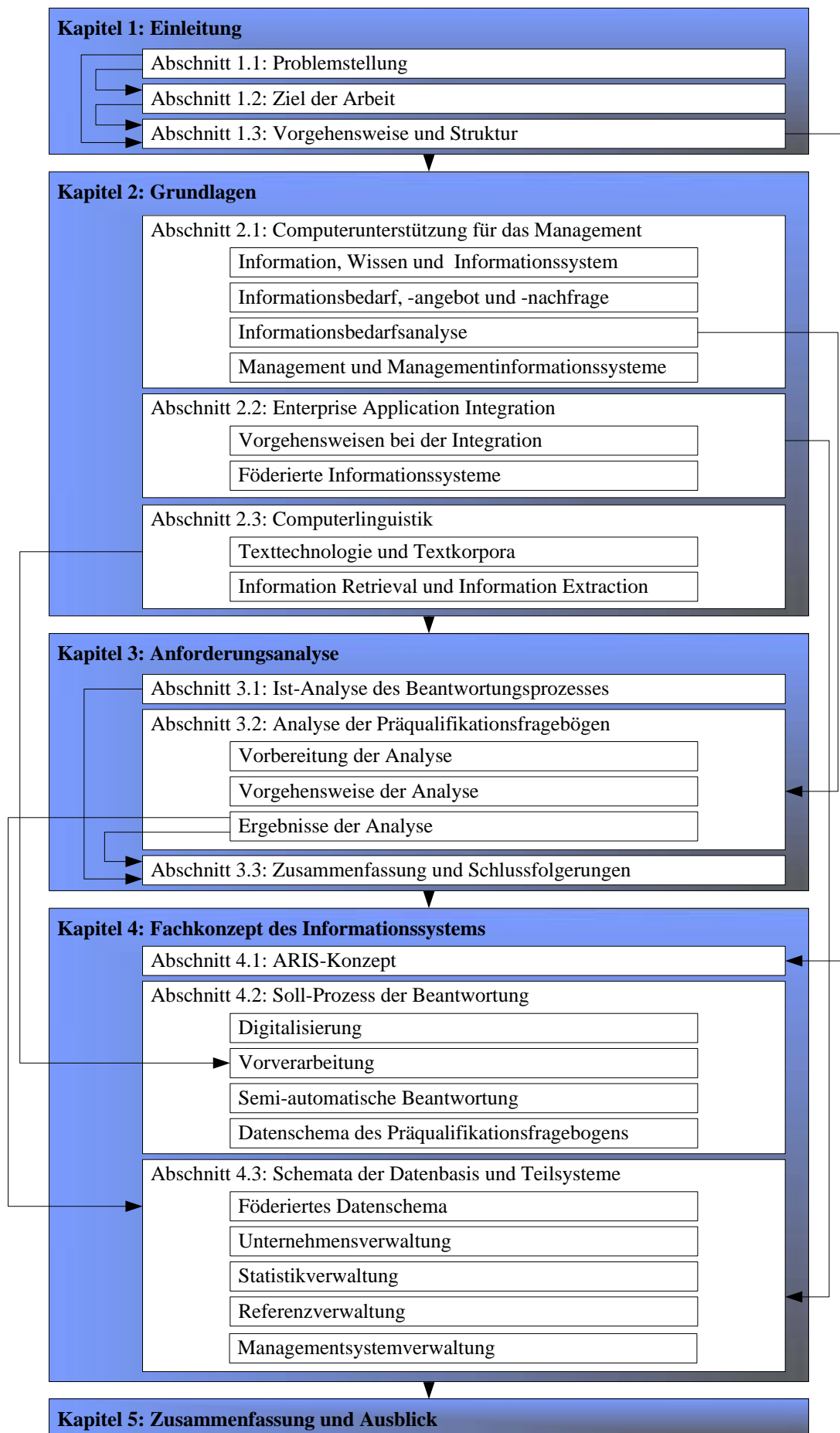


Abb. 1.2: Aufbau der Arbeit

Das Kapitel 3 gliedert sich in drei Abschnitte und beschäftigt sich mit den Vorgehensweisen und Ergebnissen der Anforderungsanalyse. Der Abschnitt 3.1 stellt den Ist-Prozess und die im Workshop genannten Anforderungen dar. Der zweite Abschnitt beschäftigt sich mit der Informationsbedarfsanalyse, die auf Basis der Präqualifikationsfragebögen durchgeführt wird. In diesem Zusammenhang wird im Abschnitt 3.2 die Vorgehensweise bei der Analyse und der Auswertung beschrieben. Die Ergebnisse dieser Analyse werden im Abschnitt 3.2.3 vorgestellt. Den Abschluss dieses Kapitels bildet der Abschnitt 3.3 mit einer Zusammenfassung der Analyseergebnisse und einigen Schlussfolgerungen für das Fachkonzept.

Das Kapitel 4 führt die bis dato gewonnenen Erkenntnisse zu einem Fachkonzept zusammen und ergänzt diese gegebenenfalls. Zunächst wird das verwendete Modellierungskonzept ARIS im Abschnitt 4.1 vorgestellt. Darauf aufbauend wird der Soll-Prozess der Beantwortung im Abschnitt 4.2 skizziert. Dieser gliedert sich in die Teilprozesse Digitalisierung, Vorverarbeitung, semi-automatische Beantwortung und das benötigte Datenschema der Präqualifikationsfragebögen. Das Kapitel endet mit dem Abschnitt 4.3. In diesem werden die Erkenntnissen des Enterprise Application Integration, die Ergebnisse der Informationsbedarfsanalyse und die Schlussfolgerungen des Abschnitts 3.3 aufgegriffen, um im Abschnitt 4.3.1 in der Top-Down-Vorgehensweise ein föderiertes Datenschema zu entwickeln. Aus diesem globalen Schema leiten sich die in den nachfolgenden Abschnitten beschriebenen Teilsysteme Unternehmens-, Statistik-, Referenz- und Managementsystemverwaltung sowie deren Teilschemata bzw. die benötigten Exportschemata ab.

Den Abschluss dieser Arbeit bildet das Kapitel 5 mit einer Zusammenfassung der Ergebnisse. Dabei werden die Probleme aufgegriffen, die während der fachkonzeptionellen Entwicklung des Informationssystems aufgetreten sind und Aufgaben für die weitere Forschung dargestellt.

2 Grundlagen

Die Beantwortung des Präqualifikationsfragebogens führt zu einer Positionierung des eigenen Unternehmens bei einem anfragenden Unternehmen. Mit dieser Positionierung ist eine mögliche Auftragserteilung verbunden. Dadurch ist die Bearbeitung eines Präqualifikationsfragebogens von strategischer Bedeutung und obliegt dem Management. Das zu entwickelnde Informationssystem soll das Führungspersonal bei dieser Aufgabe unterstützen. Aus diesem Grund werden in diesem Kapitel die Grundlagen der Managementunterstützung durch Informationssysteme diskutiert. Dabei werden, aufbauend auf den Definitionen von Information und Management, die Begriffe Informationssystem und schließlich Managementinformationssystem betrachtet. Der Begriff Managementinformationssystem lässt sich in die Begriffe Management und Informationssystem zerlegen. Letzterer setzt sich wiederum aus den Begriffen Information und System zusammen. Die Erläuterungen dieser Begriffe bilden den Anfang dieses Kapitels. Dabei wird auf die in dieser Arbeit verwendeten Methoden der Informationsbedarfsanalyse eingegangen.

In den darauf folgenden Abschnitten werden weitere Grundlagen für das zu entwickelnde Fachkonzept gelegt. Da sich die Umgebung des zu entwickelnden Informationssystems als heterogen darstellt und verschiedene Informationsquellen integriert werden müssen, werden das Forschungsfeld Enterprise Application Integration und das Konzept der föderierten Informationssysteme vorgestellt. Den Abschluss dieses Kapitels bildet ein Exkurs in die Computerlinguistik, um Grundlagen für die semi-automatische Textverarbeitung zu erläutern.

2.1 Computerunterstützung für das Management

2.1.1 Information und Wissen

Der Begriff der Information ist sehr alt und wird in den verschiedenen Wissenschaften mit unterschiedlichem Fokus betrachtet. In dieser Arbeit wird Information aus betriebswirtschaftlicher Sicht gesehen und als *zweckorientiertes Wissen* definiert. Diese Definition lässt sich auf Wittmann zurückführen (vgl. Alex (1998), S. 9; Alpar et al. (2005), S. 8) und beinhaltet den komplexen Begriff Wissen. Wissen kann als gesamter Erkenntnisstand eines Menschen, eines Unternehmens oder der Menschheit definiert werden (vgl. Rautenstrauch und Schulze (2003), S. 7). In der Definition von Wittmann wird Wissen durch seine Verwendung zur Erreichung eines bestimmten Ziels zur Information. In diesem Zusammenhang kann Wissen, wenn es im Unternehmen

„außerhalb des Menschen gespeichert wird, als 'Daten' bezeichnet“ (Alex (1998), S. 10) werden.

Information und Wissen werden als Daten codiert (vgl. Rautenstrauch und Schulze (2003), S. 8). Daten werden durch definierte Syntaxregeln aus Symbolen und Zeichen eines Zeichenvorrats gebildet (vgl. Bodendorf (2006), S. 1). Übertragene Daten werden zu Nachrichten und der Austausch von Nachrichten wird als Kommunikation bezeichnet (vgl. Alpar et al. (2005), S. 9). Kommunikation ist eine Handlung, bei der „Symbole bzw. Zeichen von einem Sender zu einem Empfänger weitergereicht werden“ (Knackstedt (2006), S. 35). Diese Zeichen werden vom Empfänger hinsichtlich ihrer Bedeutung und des Kontextes interpretiert und dadurch zu Wissen oder Information. Dabei ist nicht zwangsläufig sichergestellt, dass die verstandene Bedeutung der des Senders entspricht.

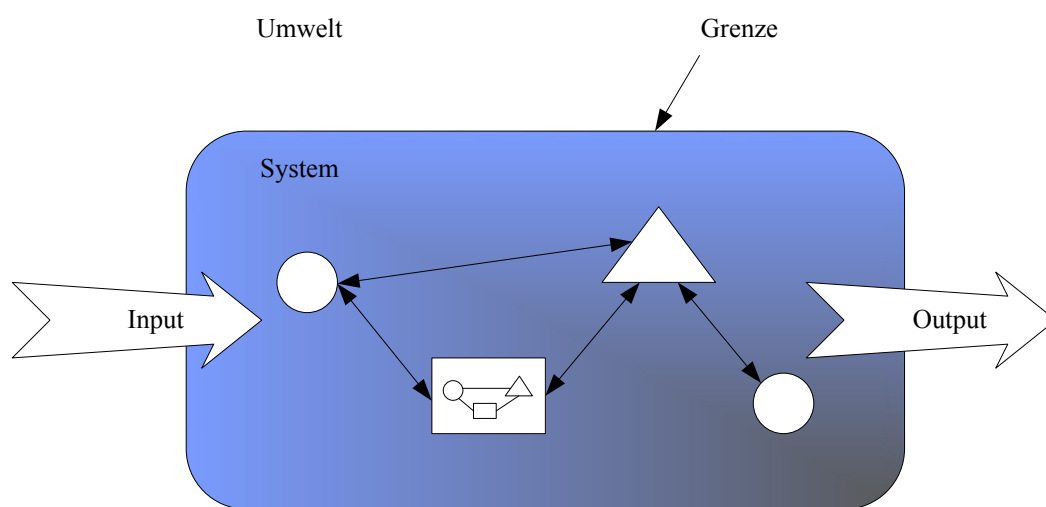
Die Definition der Information von Wittmann wird in Alpar et al. (2005) zu *Information als zusätzliches zweckorientiertes Wissen* erweitert und deutet damit ein Wachstum von Wissen an. Durch diese Betrachtungsweise wird Wissen „indirekt als Summe aller erhaltenen, verarbeiteten und als Wissen gespeicherte Informationen definiert“ (Alpar et al. (2005), S. 8). So entsteht aus Wissen Information und aus Information entsteht wiederum Wissen. Ob Wissen für eine Person zu Information wird, hängt per Definition vom Kontext, in dem das Wissen verwendet wird, ab. Wissen, welches bspw. zur Lösung eines Problems eingesetzt wird, stellt während der Verarbeitung Information dar. Die entstandene Lösung kann dann als Zuwachs von Wissen betrachtet werden. Somit ist neues Wissen entstanden, welches wiederum als Information in andere Probleme einfließen kann. Demzufolge wird zwischen Information und Wissen keine Hierarchie gebildet, wie es bei anderen Autoren der Fall ist¹ (vgl. Alpar et al. (2005), S. 9). Vielmehr handelt es sich bei Information und Wissen um zwei Zustände der gleichen Substanz (vgl. Alpar et al. (2005), S. 9).

2.1.2 Informationssystem

Der Begriff Informationssystem ist eine Zusammensetzung aus dem zuvor definierten Begriff Information und dem noch zu definierenden Begriff System. Ein System besteht nach Alpar et al. (2005) aus einer bestimmten Anzahl von Elementen, die untereinander

¹Unter der Annahme, dass Informationen Einfluss auf die Wahrnehmung einer Person und die Beurteilung von Sachverhalten haben, entsteht Wissen durch die zweckgerichtete Vernetzung von Information und ist ihr hierarchisch übergeordnet (vgl. Bodendorf (2006), S. 1 f.). Auch Bodendorf (2006) äußert sich kritisch zur vorgestellten Trennschärfe der Begriffe Daten, Information und Wissen. Er schlägt vor ein Kontinuum, mit den Polen Daten und Wissen, anzunehmen (vgl. Bodendorf (2006), S. 2). Diese werden mit Differenzierungsmerkmalen unterlegt. Allerdings befindet sich auch bei diesem Ansatz die Information nicht spezifiziert zwischen den beiden Polen und bildet damit implizit wiederum eine Hierarchie, bei der bloß die Übergänge nicht genau definiert sind.

verknüpft sind und sich von ihrer Umwelt abgrenzen. Ein System kann, muss aber nicht zwingend, mit seiner Umwelt kommunizieren. Aus seiner Umwelt kann ein System Inputs (Eingaben) erhalten. Die erhaltenen Inputs werden im System von den einzelnen Elementen verarbeitet. Das Ergebnis der Verarbeitung kann als Output vom System an seine Umwelt freigesetzt werden. Die Form des In- und Outputs, z. B. als materielle oder immaterielle Güter, ist vom jeweiligem System abhängig. Werden die einzelnen Elemente näher betrachtet, können sie ebenso Systeme darstellen und werden als Subsystem bezeichnet (vgl. Alpar et al. (2005), S. 18). Die Abb. 2.1 zeigt die schematische Darstellung eines Systems.



Quelle: vgl. Alpar et al. (2005), S. 18

Abb. 2.1: Beziehung zwischen einem System und seiner Umwelt

Systeme lassen sich durch unterschiedliche Kriterien klassifizieren. Eine mögliche Systemklassifikation wird von Alpar et al. (2005) beschrieben. Diese Klassifikation beinhaltet die Kriterien Entstehung, Komponenten, Existenz, Umweltinteraktion, Verhalten, Anpassung und Steuerung. Die Entstehung eines Systems erfolgt entweder auf natürliche oder auf künstliche Weise. Wobei künstliche Systeme durch den Menschen entwickelt werden. Dabei können maschinelle oder natürliche Elemente, auch Komponenten genannt, verwendet werden. Aus diesem Grund können Systeme aus natürlichen, maschinellen oder gemischten Komponenten bestehen. Die Existenz eines Systems hängt von seiner Umsetzung, als abstraktes Gedankenkonstrukt oder konkret real existierendes System, ab. Anhand ihrer Umweltinteraktion werden Systeme als geschlossen oder offen eingeordnet. Geschlossene Systeme interagieren nicht mit ihrer Umwelt, während offene Systeme über bestimmte Schnittstellen mit ihrer Umwelt kommunizieren können. Das Verhalten von Systemen wird als deterministisch bezeichnet, wenn es vollständig vorhergesagt werden kann. Systeme, deren Verhalten teilweise

vorhersagbar ist, weisen ein stochastisches Verhalten auf. Das ist schon der Fall, wenn sich nur eine Komponente z. B. in Bezug auf ihren Ausfall einer gewissen Unsicherheit unterwirft. Von zufälligem Verhalten wird gesprochen, wenn keine Vorhersage möglich ist. Das Kriterium der Anpassung unterteilt sich in adaptiv und nicht-adaptiv. Adaptiv bedeutet, dass ein System auf Eingaben seiner Umgebung reagiert und sich verändert, also anpasst. Im Gegensatz dazu bedeutet nicht-adaptiv, dass ein System sich nicht an seine Umwelt anpasst oder sich nicht anpassen kann. Die Steuerung eines Systems kann mit oder ohne Rückkopplung erfolgen. Rückkopplung bedeutet in diesem Zusammenhang, dass auf Grund des Outputs, den ein System erzeugt, Änderungen der Eingaben vorgenommen werden, um die Abweichungen des Outputs von bestimmten Soll-Werten zu reduzieren. Zusammenfassend werden in der Abb. 2.2 diese in Alpar et al. (2005) beschriebenen Kriterien als morphologischer Kasten dargestellt.

Kriterium	Ausprägung		
Entstehung	natürlich		künstlich
Komponenten	maschinell	natürlich	maschinell und natürlich
Existenz	abstrakt		konkret
Umweltinteraktion	offen		geschlossen
Verhalten	deterministisch	stochastisch	zufällig
Anpassung	adaptiv		nicht-adaptiv
Steuerung	mit Rückkopplung		ohne Rückkopplung

Abb. 2.2: Morphologischer Kasten: Systemeigenschaften

Aus betriebswirtschaftlicher Sicht lässt sich ein Informationssystem „als ein aufeinander abgestimmtes Arrangement von personellen, organisatorischen und technischen Elementen verstehen, das dazu dient, Handlungsträger mit zweckorientiertem Wissen für die Aufgabenerfüllung zu versorgen“ (Picot und Maier (1992), S. 923). Danach kann ein Informationssystem als eine „formale Beschreibung des gesamten Kommunikationssystems eines Unternehmens“ (Arndt (1997), S. 137) gelten. Nach der obigen Systemdefinition lässt sich das gesamte Informationssystem des Unternehmens in Subsysteme gliedern, die einzeln oder als Ganzes betrachtet werden können. Die Subsysteme können formal oder informal sein. Als computergestützte Informationssysteme werden Systeme bezeichnet, die in Form eines Mensch-Maschine-Systems Teilprozesse durch Computer realisieren (vgl. Alex (1998), S. 13, Picot und Maier (1992), S. 923). Im weiteren Verlauf dieser Arbeit wird, auf Grund des zu entwickelnden Informationssystems und dessen Computerpartizipation, unter

einem Informationssystem ein computergestütztes, „künstliches, konkretes System, das aus maschinellen und natürlichen Elementen besteht und seine Nutzer mit Informationen versorgt“ (Alpar et al. (2005), S.28), verstanden.

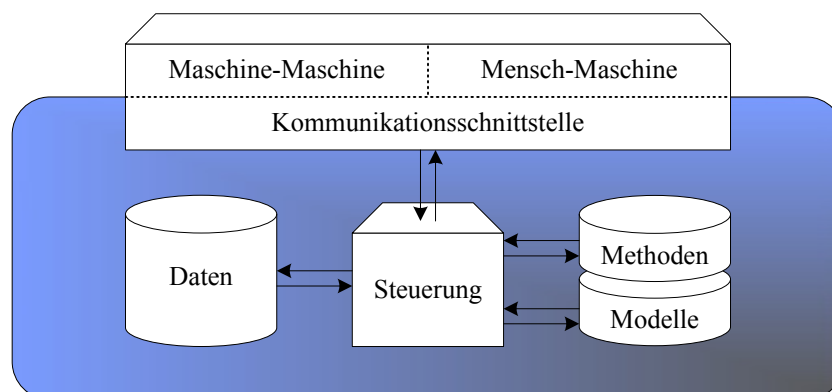
Die Aufgaben eines Informationssystems sind Erfassung, Klassifikation, Verarbeitung, Speicherung, Vermittlung und die Ausgabe von Informationen (vgl. Struckmeier (1997) S. 9). Diese Aufgaben werden von Arndt (1997) um die Wiedergewinnung und Auswertung von Informationen erweitert. Dadurch wird auf die benötigten Suchfunktionalitäten und Auswertungsmöglichkeiten computergestützter Informationssysteme hingewiesen. Für die Verteilung von Informationen werden das Push-Prinzip und das Pull-Prinzip unterschieden (vgl. Bodendorf (2006), S. 5). Beim Push-Prinzip werden Informationen proaktiv durch deren Urheber an die Empfänger verteilt. Im Gegensatz dazu geht die Initiative der Informationsbeschaffung beim Pull-Prinzip vom Empfänger aus. Dieser ruft die gewünschten Informationen an einer meist zentralen Stelle ab.

Informationssysteme werden nach wirtschaftlichen Kriterien mit einer „optimale[n] Bereitstellung von Information und Kommunikation“ (Knackstedt (2006), S. 36) gestaltet. Dabei ist der Begriff *optimal* zu relativieren, weil das Optimum meist nicht bekannt ist. Daher wird zu einem bestimmten Zeitpunkt ein Optimum angenommen, welches während der Entwicklung zu erreichen ist. Die Gestaltung eines Informationssystems ist „eine komplexe Aufgabe, [...] die umfangreiche Strukturierungsleistungen [...] erfordert“ (Knackstedt (2006), S. 36). Für diese Strukturierungsleistung und deren Kommunikation werden nach Knackstedt (2006) Modelle verwendet. Ein Modell ist nach klassischer Auffassung ein vereinfachtes Abbild der Realität. Nach konstruktivistischer Auffassung wird ein Modell aus der Vorstellungswelt des Modellierers geschaffen und stellt eine Repräsentation von relevant erscheinenden Sachverhalten dar (vgl. Knackstedt (2006), S. 36 f.). Nach einer neueren Diskussion in der Wirtschaftsinformatik tritt die folgende Auffassung in den Vordergrund: Ein Modell ist das „Ergebnis eines Konstruktionsprozesses, bei dem die Wahrnehmung von Inhalten eines ausgewählten Gegenstandes zweckorientiert repräsentiert wird“ (Alpar et al. (2005), S. 20). Diese Auffassung des Modellbegriffs erscheint für das zu entwickelnde Fachkonzept dieser Arbeit (siehe Kapitel 4) als zweckmäßig.

Der Aufbau eines Informationssystems lässt sich, als allgemeines Modell mit den fünf Komponenten Informationsbenutzer, Informationsobjekt, Informationsgenerator, Informationsvermittler und Informationssystemgestalter beschreiben. In diesem allgemeinen Modell geht die Informationsnachfrage vom Informationsbenutzer aus. Dieser fordert ein Informationsobjekt bei einem Informationsgenerator an. Der Generator

erzeugt dann das Informationsobjekt und stellt es dem Informationsbenutzer zur Verfügung. Wenn das Informationsobjekt nicht direkt dem Informationsbenutzer übertragen werden kann, wird ein Informationsvermittler dazwischen geschaltet, der die Informationsobjekte interpretiert. Nach Alex (1998) kommt es in der Praxis häufig zu Schwierigkeiten, die benötigten Informationsobjekte abzufragen oder vorzuhalten. Ursachen dafür sind bspw. ungenau formulierte Abfragen sowie nicht konsistente oder fehlenden Daten für die Generierung der Informationsobjekte. Um dem entgegenzuwirken, entwickelt der Informationssystemgestalter zielgerichtet, dem Informationsbedarf der Informationsbenutzer entsprechende Informationsgeneratoren. Im Sinne der Mensch-Maschine- und Maschine-Maschine-Kommunikation können der Informationsgenerator, -vermittler und -benutzer jeweils ein natürliches oder technisches Element des Informationssystems darstellen. Im weiteren Verlauf der Arbeit werden die Begriffe Informationsbenutzer und Benutzer synonym verwendet.

Computergestützte Informationssysteme bestehen aus den vier Komponenten Datenbank, Kommunikationsschnittstelle, Modell/Methodendatenbank und Ablaufsteuerung (siehe Abb. 2.3).



Quelle: vgl. Alex (1998), S. 16

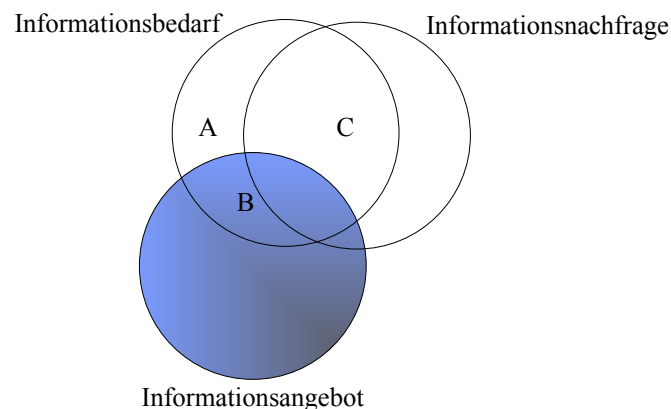
Abb. 2.3: Modell computergestützter Informationssysteme

Die Datenbasis des Informationssystems wird in Form von Daten in der Datenbank gespeichert. Sie enthält die logische Struktur der Daten, welche während der Entwicklung der Datenbank festgelegt wird. Die Modell/Methodendatenbank beinhaltet die für die Aufgaben des Informationssystems notwendigen Modelle und Methoden. Methoden sind Verfahren zur Problemlösung. Die gespeicherten Modelle stellen für die Lösung bestimmter Probleme die „Verknüpfung von Methoden, Daten und Steuerinformationen“ (Alex (1998), S. 17) her. Nach Alex (1998) stellt die Modell/Methodendatenbank die Funktionsdatenbank dar und enthält somit die Funktionen, Prozesse oder

Teilprozesse eines Informationssystems. Die Kommunikationsschnittstelle beinhaltet Elemente zur Maschine-Maschine- und Mensch-Maschine-Kommunikation. Über die Maschine-Maschine-Schnittstelle werden hauptsächlich Daten zwischen Anwendungen ausgetauscht. Die Darstellung der Information und die Realisierung der Eingaben durch den Benutzer werden über die Mensch-Maschine-Schnittstelle realisiert. Schließlich stellt die Ablaufsteuerung die zentrale Komponente im Informationssystem dar, die alle Komponenten miteinander verbindet.

2.1.3 Informationsbedarf, -angebot und -nachfrage

Die drei Begriffe Informationsbedarf, Informationsnachfrage und Informationsangebot stehen oft nicht deckungsgleich zueinander (siehe Abb. 2.4).



Quelle: vgl. Stickel (2001), S. 4

Abb. 2.4: Beziehung zwischen Informationsangebot, Informationsnachfrage und Informationsbedarf

Der Informationsbedarf wird in dieser Relation auch als objektiver Informationsbedarf bezeichnet (vgl. Stickel (2001), S. 5). Die Informationsnachfrage wird als subjektiv empfundener Informationsbedarf aufgefasst und „ergibt sich durch Aggregation der jeweiligen Nachfrage der Aufgabenträger im Unternehmen“ (Stickel (2001), S. 5). Der subjektiv empfundene Informationsbedarf wird auch als Informationsbedürfnis bezeichnet (vgl. Oppelt (1995), S. 203). Dem gegenüber steht das Informationsangebot, welches die Informationen repräsentiert, die das Informationssystem des Unternehmens zur Verfügung stellen kann. Allerdings gibt es einige Probleme, wie Abb. 2.4 verdeutlicht. Werden bspw. durch die Benutzer Informationen nachgefragt, die nicht der Aufgabenerfüllung dienen, entspricht der objektive Informationsbedarf nicht dem subjektiven Informationsbedürfnis. In diesem Fall ist die klare Beschreibung der Aufgabe und der zur Lösung benötigten Informationsobjekte unabdingbar. In der Abb. 2.4 wird

deutlich, dass das Informationsangebot zu umfangreich für eine bestimmte Aufgabe sein kann und es dadurch zu einer Informationsüberflutung kommt. Ein weiteres Problem ist die nicht vollständige Abdeckung des objektiven Informationsbedarfs. Die Gestaltung des Informationsangebotes ist eine wichtige Voraussetzung für die Effektivität eines Informationssystems. Das Optimum für ein Informationssystem stellt in diesem Zusammenhang die Gleichheit von Informationsnachfrage, Informationsbedarf und Informationsangebot dar. Die exakte Ermittlung des objektiven Informationsbedarfs ist in der Praxis oft mit Schwierigkeiten verbunden, weshalb ein kontinuierlicher Verbesserungsprozess mit einer „nahezu optimalen Lösung“ (Stickel (2001), S. 5) angestrebt wird.

2.1.4 Informationsbedarfsanalyse

Ein Informationssystem soll seinen Benutzern Informationen zur Verfügung stellen (vgl. Abschnitt 2.1.2). Aus diesem Grund ist es für diese Arbeit essenziell die Informationsnachfrage und den Informationsbedarf der potenziellen Benutzer zu ermitteln. Darauf aufbauend kann dann das Informationsangebot geplant werden. Die Ermittlung des Informationsbedarfs und die Bedeutung der Informationsbedarfsanalyse sind hinreichend bekannt, allerdings wird eine „unzureichende theoretische und methodische Absicherung des Vorgehens bemängelt“ (Knackstedt (2006), S. 342). Dies führt häufig zu nachträglichen Modifikationen und bei der Systemeinführung zu Zeitverzögerungen (vgl. Struckmeier (1997), S. 29). Aus diesem Grund werden an dieser Stelle die Informationsbedarfsanalyse und die in dieser Arbeit verwendeten Methoden kurz vorgestellt.

Die Informationsbedarfsanalyse ist die systematische Erhebung aufgabenbezogener, fachlicher und allgemeiner Informationen, die für die Informationsbenutzer notwendig sind, um die ihnen gestellte Aufgabe zu erfüllen und sich in der betrieblichen Umwelt zu orientieren (vgl. Struckmeier (1997), S. 28). Sie dient als Basis für die Gestaltung des Informationsangebots, welches von Informationssystemen zur Verfügung gestellt wird. Dabei beinhalten die Aufgaben der Informationsbedarfsanalyse die Analyse des Informationsflusses, der informationsbedarfsauslösenden Faktoren, des Zwecks der Information und des Nutzerverhaltens bezüglich Präferenzen, Gewohnheiten und Werte (vgl. Struckmeier (1997), S. 29). Für die Informationsbedarfsanalyse stehen in der Literatur verschiedene Methoden zur Verfügung, die im Folgenden genannt werden (vgl. Struckmeier (1997), S. 30 ff.; Knackstedt (2006), S. 342 f. und Beiersdorf (1995), S. 75):

- Fragebogenmethode,
- Interview,

- Brainstorming,
- Berichtsmethode,
- Konferenzmethode,
- Katalog-Methode,
- Entscheidungsfeldanalyse,
- Dokumentenanalyse,
- Aufgaben- und Problemanalyse,
- Analogieschlussmethode,
- Spiegelbildmethode,
- Blickaufzeichnung,
- Business-System-Planning,
- Ist-Ansatz,
- Nebenprodukttechnik,
- Null-Ansatz,
- Schlüsselindikatoren-Ansatz,
- Total-Ansatz,
- Kritische Erfolgsfaktoren sowie die
- Selbst- und Fremdbeobachtung.

Diese Liste erhebt keinen Anspruch auf Vollständigkeit, zeigt aber die Vielfalt der vorhandenen Methoden. Da der Informationsbedarf, den ein Informationssystem abdecken soll, sehr spezifisch ist, sind nicht alle Methoden gleichermaßen anwendbar. Im Folgenden werden deshalb nur die Methoden beschrieben, die sich auf das spezifische Problem dieser Arbeit anwenden und adaptieren lassen. Zu diesen Methoden gehören Ist-Ansatz, Konferenzmethode, Interview, Dokumentenanalyse sowie Aufgaben- und Problemanalyse.

Der Ist-Ansatz stellt eine Möglichkeit dar, einen Ist-Zustand zu erfassen, zu beschreiben, zu analysieren und zu bewerten (vgl. Struckmeier (1997), S. 39). Dabei werden zu Beginn der Durchführung die, für die Erhebung, benötigten Methoden festgelegt. Ein Nachteil dieser Methode ist die Zeitpunktbezogenheit. Es werden nur Zustände und Tatsachen erfasst, die zur Erhebung verfügbar sind. Damit stellt das Ergebnis des Ist-Ansatzes eine Momentaufnahme eines sich weiterentwickelnden Systems dar. Dennoch können aus dieser Momentaufnahme Erkenntnisse über das betrachtete System gewonnen werden. Die beim Ist-Ansatz ermittelten Aufgaben und Probleme können dann unter Verwendung der Aufgaben- und Problemanalyse genutzt werden, um Soll-Zustände festzulegen (vgl. Beiersdorf (1995), S. 76).

Die Konferenzmethode ist eine Methode, die in einer Gruppe mit Moderator durchgeführt wird, um eine einheitliche Sichtweise auf einen Ist-Zustand und die Anforderungen für

den Soll-Zustand zu erhalten (vgl. Beiersdorf (1995), S. 76). In der Praxis werden hierfür häufig Workshops durchgeführt. Der Vorteil gegenüber Einzelgesprächen liegt in der Beteiligung von mehreren Personen. Dadurch werden einseitige Sichtweisen vermieden. Als nachteilig ist der hohe Zeitaufwand bei der Vorbereitung und Durchführung zu erwähnen.

Die Interview-Methode ist eine flexible, mündliche Befragung, die offen oder strukturiert mit einer oder mehrerer Personen durchgeführt wird (vgl. Struckmeier (1997), S. 31). Strukturiert ist das Interview in dem Sinne, dass im Vorfeld Fragen fixiert werden, um eine bestimmte Problemstellung zu hinterfragen. Der Verlauf eines Gesprächs ist dann aber in diesem Rahmen flexibel. Die Nachteile dieser Methode liegen operativ im hohen Zeitaufwand, administrativ in der zeitlichen Koordination mit den Befragten und inhaltlich in der subjektiven Darstellung von allgemeinen Sachverhalten. Als Vorteile sind der persönliche Kontakt, die daraus resultierende Flexibilität und das geringe Risiko für Missverständnisse zu nennen (vgl. Knackstedt (2006), S. 343). Ein weiterer Vorteil ist die Darstellung des individuellen Informationsbedürfnisses des Befragten (vgl. Struckmeier (1997), S. 31).

Die Dokumentenanalyse ist die letzte der hier vorgestellten Methoden. Diese Methode dient grundlegend dazu, Dokumente daraufhin „zu analysieren, welcher Informationsbedarf sich aus ihnen für die Lösung konkret gestellte[r] Aufgabe[n] ergibt“ (Beiersdorf (1995), S. 84). Der Vorteil dieser Methode ist der Versuch den Informationsbedarf ohne subjektive Einflüsse von Personen zu ermitteln und damit eine Objektivierung zu erreichen. Dabei stellen Dokumente eine auf die Vergangenheit gerichtete Informationsquelle dar, was als Nachteil erwähnt werden muss. Deshalb ist je nach Problemstellung auf die Aktualität der Dokumente zu achten.

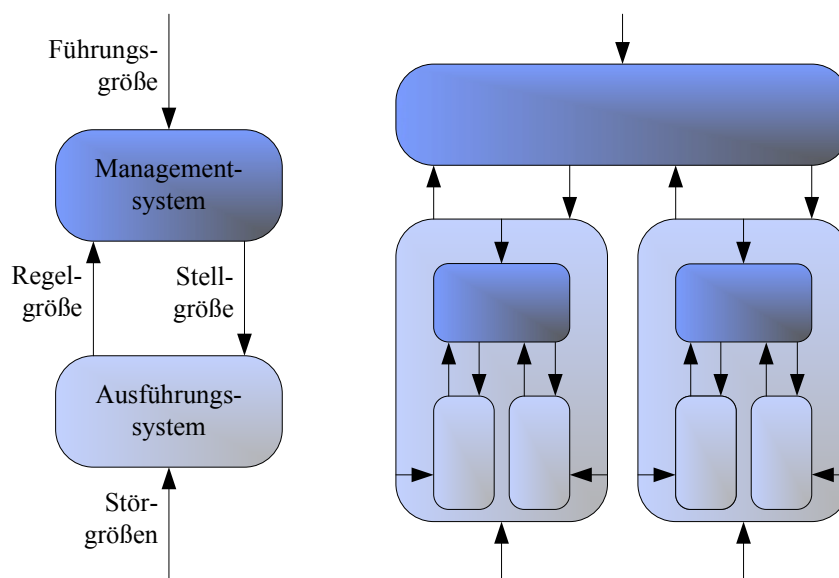
Jede der beschriebenen und nicht beschriebenen Methoden weisen verschiedene Mängel auf, daher wird in der Literatur die Kombination einzelner Methoden propagiert (vgl. Struckmeier (1997), S. 38; Knackstedt (2006), S. 344). Wie die Methoden angewendet und adaptiert werden, hängt von der jeweiligen Problemstellung und Situation ab (vgl. Kapitel 3).

2.1.5 Management

Der Begriff Management kann aus funktionalem, institutionellem und instrumentalem Blickwinkel betrachtet werden (vgl. Knackstedt (2006), S. 69). Die funktionale Sicht behandelt die Aufgaben des Managements, während die institutionelle Sicht sich mit den Personen, die die Managementaufgaben ausführen, beschäftigt. Schließlich widmet

sich die instrumentale Sicht dem Methodenspektrum des Managements und besonders den verwendeten Informationssystemen (vgl. Knackstedt (2006), S. 69). Im Folgenden werden die drei Sichtweisen zum Management näher betrachtet.

Unter funktionalen Gesichtspunkten und der Betrachtung des Managements als System unterteilt sich der Managementbereich in Managementsystem bzw. -prozess und Ausführungssystem bzw. -prozess (vgl. Knackstedt (2006), S. 69). Der Managementprozess beeinflusst aktiv seine Umwelt. Die Beziehungen zwischen Management- und Ausführungssystem werden im Regelkreismodell in Abb. 2.5 (linke Seite) dargestellt.



Quelle: vgl. Knackstedt (2006), S. 70

Abb. 2.5: Unterscheidung von Ausführens- und Managementsystem im Regelkreis und deren Verschachtelung

Das Managementsystem erhält Vorgaben in der Form von Führungsgrößen. Diese werden vom Managementsystem genutzt, um unter Verwendung der Stellgrößen das Ausführungssystem zu lenken. Die Verarbeitung der Stellgrößen erfolgt im Ausführungssystem. Dabei unterliegt das Ausführungssystem zusätzlichen, von außen kommenden, Störgrößen, auf die das Managementsystem keinen Einfluss hat. Die Störgrößen werden vom Ausführungssystem indirekt mit den erarbeiteten Ergebnissen an das Managementsystem in Form der Regelgrößen übermittelt. Unter Verwendung der Regelgrößen kann im Managementsystem ein Soll-Ist-Abgleich mit den Führungsgrößen erfolgen. Aus dieser Analyse ergeben sich wiederum neue Stellgrößen für das Ausführungssystem.

Die einstufige Darstellung der Unterscheidung von Management- und Ausführungssystem (Abb. 2.5 linke Seite) wird der Komplexität von Unternehmen nicht gerecht, weshalb die Unterscheidung der beiden Systeme rekursiv angewendet wird (vgl. Knackstedt (2006), S. 70). Die daraus resultierende Verschachtelung und Interaktion der Systeme wird ebenfalls in der Abb. 2.5 (rechte Seite) gezeigt.

Der interne Managementprozess kann in die Phasen Zielbildung, Problemerkennung, Planung, Implementierung und Kontrolle eingeteilt werden (vgl. Knackstedt (2006), S. 70 ff.). In der Phase der Zielbildung wird für den Managementprozess das Zielsystem entwickelt, indem von außen vorgegebene Ziele ermittelt, analysiert und geordnet werden. Auf Basis der festgelegten Ziele können Abweichungen und damit Handlungszwänge in der Phase der Problemerkennung festgestellt werden. In der Planung werden erkannte Probleme analysiert und mögliche Lösungen erarbeitet. Danach erfolgt die Bewertung der Problemlösungsvorschläge. Die Option mit dem bestmöglichen Zielerreichungsgrad wird am Ende der Planungsphase ausgewählt. Als Implementierung wird in diesem Zusammenhang die Umsetzung der ausgewählten Maßnahme durch deren Initialisierung bezeichnet. Dabei werden neue Führungs- bzw. Steuerungsgrößen für untergeordnete Management- bzw. Ausführungssysteme generiert. In den untergeordneten Managementsystemen kann sich die Vorgabe weiterer Zielgrößen fortsetzen. Dies ist aber nicht mehr Bestandteil der initiiierenden Management-Instanz (vgl. Knackstedt (2006), S. 71 f.), sondern obliegt der Autonomie des untergeordneten Systems. In der Kontrollphase werden die Teil- bzw. Endergebnisse, der in der Planungsphase festgelegten Maßnahme, geprüft. Dies erfolgt um festzustellen, ob durch die implementierte Maßnahme das geplante Ziel erreicht wurde. Eine weitere Gliederung des Managements aus funktionaler Sicht erfolgt nach der Fristigkeit der Planung in strategisches, taktisches und operatives Management (vgl. Knackstedt (2006), S. 72).

Die institutionelle Sicht betrachtet, im Gegensatz zur funktionalen Sicht, nur die Organisationseinheiten, die eine gewisse Entscheidungsgewalt besitzen. Dabei wirkt die funktionale Sicht bei der „Identifikation notwendiger aber nicht institutionell zugewiesener Aufgaben“ (Knackstedt (2006), S. 73) unterstützend. Die Gliederung des Managements erfolgt institutionell in oberes, mittleres und unteres Management (vgl. Staehle (1999), S. 89 ff.). Andere gebräuchliche Begriffe, die denselben Sachverhalt darstellen, sind Top-Management (vgl. Reckemmer (1999), S. 33), Fachbereichsmanagement und operatives Management (vgl. Knackstedt (2006), S. 73).

Die instrumentale Sicht beschäftigt sich mit den Managementtechniken und -methoden. Für das Management sind Informationen die wichtigste Entscheidungsgrundlage. Deshalb ist die Gestaltung von Informationssystemen eine der wichtigsten Aufgaben (vgl.

Knackstedt (2006), S. 74). Im folgenden Abschnitt wird auf Informationssysteme zur Managementunterstützung genauer eingegangen.

2.1.6 Managementinformationssysteme

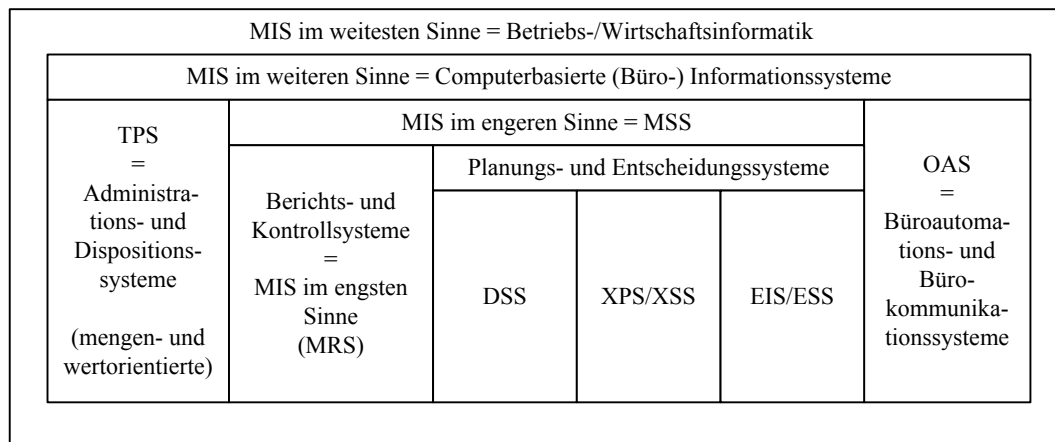
Der Begriff der Managementinformationssysteme (MIS) geht auf die Mitte der 60er Jahre zurück (vgl. Alex (1998), S. 29). Die ursprüngliche Idee der MIS war die Bereitstellung des gesamten, im Unternehmen von Transaktionssystemen erzeugten Datenvolumens. In diesem Ansatz, der auch als Total Systems Approach bezeichnet wird, sollte die Unternehmensführung, d. h. die Manager aller Ebenen, in Echtzeit über standardisierte Berichte auf alle Unternehmensdaten zugreifen können (vgl. Holten und Knackstedt (1997), S. 5). Diese anfängliche Idee scheiterte an den damaligen Rahmenbedingungen. In Holten und Knackstedt (1997) werden die folgenden Ursachen für das Scheitern genannt:

- Überzogene Erwartungshaltungen,
- inadäquate Realisierungsstrategien,
- keine einheitlichen Datenbanken, daraus folgten inkonsistente, fehlerhafte oder fehlende Daten für die Berichte,
- geringe Datenaktualität, durch störanfällige und langsame Datenverarbeitung (vgl. Vetschera (1995), S. 10 f.),
- mangelnde Akzeptanz durch die Benutzer,
- Fehleinschätzung des Benutzerverhaltens (vgl. Struckmeier (1997), S. 11),
- technische Probleme der Datenhaltung,
- Vernachlässigung der inhaltlichen Aspekte der Berichte,
- mangelnde Programmierkenntnisse der Benutzer, dadurch meist nur vordefinierte Berichte sowie
- Informationsüberflutung statt Filterung.

Diese Probleme konnten angesichts des technischen Entwicklungsstandes nicht gelöst werden.

So wandelte sich die inhaltliche Bedeutung des MIS-Begriffs in den 70er Jahren von einem totalen Informationssystem zu schrittweise integrierten Teillösungen (vgl. Holten und Knackstedt (1997), S. 6). Die Abschwächung der Anforderungen auf eine bestmögliche oder effektive Unterstützung der Managementebenen folgte aus der Erkenntnis, dass sich die Managementaufgabe als sehr komplex und wenig strukturierbar darstellt. Daher ist die vollständige Abbildung der Managementaufgabe kaum möglich.

In den folgenden Jahren entstanden verschiedene Begriffe und Techniken zur Unterstützung des Managements. Abb. 2.6² stellt einen Ausschnitt dieser Informationssysteme in einer Begriffshierarchie dar.



Quelle: in Anlehnung an Oppelt (1995), S. 9

Abb. 2.6: Begriffshierarchie computergestützter Managementinformationssysteme

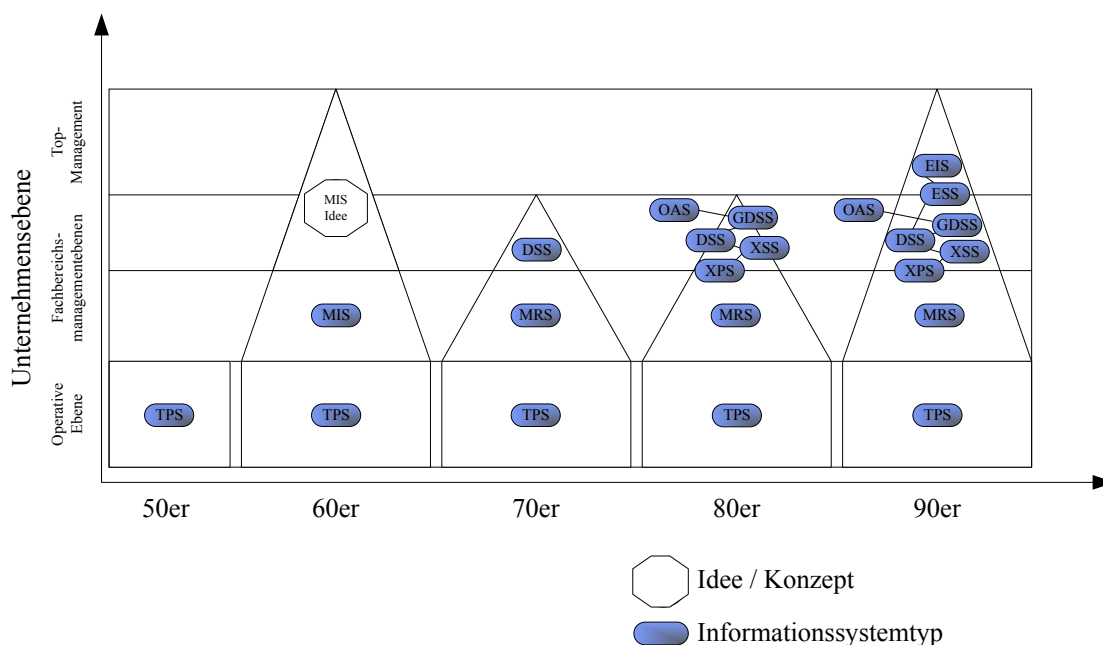
Nach Oppelt (1995) werden vier Ausdehnungen des MIS-Begriffs unterschieden. *Im weitesten Sinne* umfasst MIS alle Informationssysteme der Betriebs- und Wirtschaftsinformatik. *Im weiteren Sinne* werden alle computerbasierten (Büro-) Informationssysteme dem MIS-Begriff untergeordnet. Dazu gehören die Transaction Processing Systems (TPS), die MIS im engeren Sinne und die Office Automation Systems (OAS). *MIS im engeren Sinne* bezeichnet eine Menge von Informationssystemen zur Unterstützung der eigentlichen Managementaufgabe. Diese Systeme werden auch als Management Support Systems (MSS) bezeichnet. Zu den MSS gehören die Planungs- und Entscheidungssysteme sowie die Berichts- und Kontrollsysteme, welche auch als Management Reporting Systems (MRS) bezeichnet werden. Die Planungs- und Entscheidungssysteme unterteilen sich wiederum in Decision Support Systems (DSS), Executive Information Systems (EIS) bzw. Executive Support Systems (ESS) und die wissensbasierten Expert Systems (XPS) und Expert Support Systems (XSS).

Die MRS sind die „Nachfolger der in der Praxis realisierten MIS [..], welche der Idee der MIS nicht gerecht werden konnten“ (Holten und Knackstedt (1997), S. 6). Durch das Scheitern der ersten MIS-Generation ist der Begriff in der Literatur negativ besetzt (vgl. Alpar et al. (2005), S. 33). Aus diesem Grund streben einige Autoren die Ablösung

²Die Darstellung wurde, in Anlehnung an Holten und Knackstedt (1997), S. 10, um das Akronym XSS erweitert.

des MIS-Begriffs durch den EIS-Begriff an (vgl. Holten und Knackstedt (1997), S. 10; Struckmeier (1997), S. 12).

Die anfängliche Idee des totalen Informationssystems für alle Managementebenen scheiterte. Stattdessen wurden auf Grund der technischen Rahmenbedingungen Teilsysteme entwickelt, die die einzelnen Managementaufgaben unterstützen. Anfang der 80er Jahre wurden verschiedene Funktionalitäten einzelner System zu neuen Informationssystemen integriert, bspw. entstanden durch die Erweiterung der DSS um Gruppen-Funktionalitäten die Group Decision Support Systems (GDSS). Dies wird in der Abb. 2.7³, welche die historische Entwicklung der MIS bis in die 90er skizziert, durch die Verbindungslinien zwischen den Informationssystemtypen dargestellt.



Quelle: in Anlehnung an Holten und Knackstedt (1997), S. 12

Abb. 2.7: Historische Entwicklung von Managementinformationssystemen

Durch diese Weiterentwicklungen verschwimmen die Grenzen zwischen den Informationssystemen, wodurch die Einordnung bzw. die Abgrenzung erschwert wird. Die begriffliche Vielfalt wird zusätzlich durch die teilweise synonyme Verwendung von Begriffen, die Begriffsbildung im deutschen Sprachraum sowie durch Marketing-Aspekte von Informationssystemen negativ beeinflusst (vgl. Rechkemmer (1999), S. 21 f.;

³Es handelt sich bei dieser Darstellung um einen Ausschnitt des Originals. Weggelassen wurde, zu Gunsten einer besseren Übersicht, die EDV technische Entwicklung. Zusätzlich wurde das Akronym TVS durch TPS ersetzt, um die einheitliche Verwendung der angelsächsischen Akronyme zu verdeutlichen. Das im Original verwendete Akronym EXP wurde durch das in der Literatur gebräuchlichere XPS für Expert System ausgetauscht.

Stahlknecht und Hasenkamp (1998), S. 426 f.). Aus diesem Grund werden die hier verwendeten Begriffe und Akronyme des MIS-Bereichs im angelsächsischen und deutschen Sprachraum in der Tab. 2.1 gegenübergestellt.

Tab. 2.1: Übersicht zu Begriffen und Akronymen computerbasierter Informationssysteme

Angelsächsischer Sprachraum	Deutscher Sprachraum
OAS - Office Automation System	Büroautomationssystem Bürokommunikationssystem
TPS - Transaction Processing System	TVS - Transaktionsverarbeitungssystem Transaktionssystem Operatives System Administrationssystem
MSS - Management Support System	MUS - Managementunterstützungssystem
MRS - Management Reporting System	MIS - Managementinformationssystem Berichts- und Kontrollsystem
MIS - Management Information System(s)	Computergestütztes Informationssystem Betriebs-/Wirtschaftsinformatik
DSS - Decision Support System	EUS - Entscheidungsunterstützungssystem
XPS - Expert System	ES - Expertensystem
XSS - Expert Support System	Wissensbasiertes EUS
EIS - Executive Information System	FIS - Führungsinformationssystem CIS - Chefinformationssystem
ESS - Executive Support System	VIS - Vorstandsinformationssystem

Quelle: Holten und Knackstedt (1997) S. 11

Weitere Entwicklungen zur Unterstützung des Managements, wie z. B. Data Warehouse Konzepte, OnLine Analytical Processing (OLAP), Business Intelligence, Dokumenten-, Wissens- und Contentmanagementsysteme werden in der Abb. 2.7 nicht dargestellt. Diese Entwicklungen schließen sich am Ende der 90er Jahre an (vgl. Hansen und Neuman (2001), S. 509).

Der hier dargelegte historische Überblick zeigt die Vielfalt des MIS-Begriffs. Die negative Belastung führte zu verschiedenen Bezeichnungen. Seit den 90er Jahren wird von der dritten MIS-Generation gesprochen und das Akronym MIS erlebt eine Renaissance (vgl. Arndt (2006), S. 334 f.). In Hansen und Neuman (2001) werden MIS synonym mit MSS verwendet und als „Informationssysteme [bezeichnet], die für Führungskräfte eine adäquate Informationsversorgung und Entscheidungsunterstützung bieten“ (Hansen und Neuman (2001), S. 453). Unter *adäquat* wird in diesem Zusammenhang eine aufgabengerechte und benutzergerechte Informationsversorgung und „dem Stil des Managements entsprechende Präsentation und Bedienungsfreundlichkeit“ (Hansen und Neuman (2001), S. 453) verstanden. Ähnliche Definitionen werden von weiteren Autoren, wie Hanning (1998), S. 1, Elm (1972), S. 24 und Picot und Maier (1992), S. 931 verwendet. Des Weiteren kann das MIS-Konzept für die Unterstützung und die

Informationsversorgung von Managementsystemen eingesetzt werden (vgl. Arndt (2006), S. 336). Durch den technologischen Fortschritt und den zukünftigen Entwicklungen gelingt es sich der ursprünglichen Idee anzunähern. Dabei ist die vollständige Automatisierung auf Grund der Komplexität der Managementaufgabe schwer erreichbar. Deshalb wird in dieser Arbeit der MIS-Begriff auf Basis der Gemeinsamkeit aller vorgestellten Definitionen abgegrenzt. Ein Managementinformationssystem ist ein Informationssystem, welches das Management bei der Erfüllung seiner Aufgaben unterstützt.

2.2 Enterprise Application Integration

Integration ist die Zusammenführung eines Ganzen aus ehemals isolierten Einzelteilen (vgl. Mertens (2004) S. 1). Auf Informationssystem bezogen werden Menschen, Aufgaben und Technik integriert. Die Integration von Informationssystemen kann über Daten und Funktionen erfolgen (vgl. Scheer (1992), S. 1042 f.). Für die Datenintegration stehen unterschiedliche Möglichkeiten, wie Import und Export von Dateien und ein gemeinsamer Zugriff auf Datenquellen zur Auswahl. Der Austausch von Daten über Dateien ist eine asynchrone Kommunikation zwischen zwei Systemen. In diesem Zusammenhang stellt der Austausch über unterschiedliche Austauschformate und deren Nutzung ein Problem dar. Diese Problematik lässt sich unter der Verwendung von standardisierten Austauschformaten, wie bspw. mit der eXtensible Markup Language (XML), lösen (vgl. Winkeler et al. (2000), S. 30). Die XML stellt in diesem Zusammenhang „ein[en] de facto Standard für Datenbeschreibungssprachen“ (Reitz et al. (2003), S. 16) im Umfeld des Enterprise Application Integration dar und „ist eine Strukturbeschreibungssprache, die zum Datenaustausch in heterogenen Umgebungen verwendet wird“ (Winkeler et al. (2000), S. 30). Eine weitere Möglichkeit zur Datenintegration ist der gemeinsame Zugriff auf unterschiedliche Datenquellen. Ein integriertes Informationssystem greift gemeinsam mit den Altanwendungen auf deren Datenbasis zu. Diese Variante hat den Vorteil, dass ein integriertes System global über die benötigten Daten verfügen kann, ohne die Altanwendungen zu verändern. Ein Nachteil dieser Variante liegt im schreibenden Zugriff auf die Datenbasis, wenn die Altanwendung ebenfalls noch schreibenden Zugriff besitzt. Dies ist der Fall, wenn im integrierten System die Funktionalität des Altsystems nicht vollständig enthalten ist. Das bedeutet, dass im Unternehmen das Altsystem zur Bearbeitung der Daten weiter verwendet wird. Dies kann zu Konsistenzproblemen in den Datenbanken führen. Die vollständige Ablösung von Altsystemen ist häufig „aus Gründen des Investitionsschutzes nicht realistisch“ (Jung (2005), S. 83). Deshalb wird die Anwendungsarchitektur durch Integrationskomponenten um weitere Funktionen ergänzt.

Die Architektur der betrieblichen Informationssysteme bezieht sich nach Conrad et al. (2006) meist auf die drei Ebenen Geschäft, Anwendung und Technik. Die Geschäftsarchitektur beschreibt die Aufbauorganisation und Ablauforganisation, d.h. die Organisationsstruktur und Prozesse einer Organisationseinheit. Während die Anwendungsarchitektur die Implementierung der Geschäftsarchitektur durch Anwendungen im Unternehmen darstellt, definiert die technische Architektur die Informations- und Kommunikationsinfrastruktur.

Auf Basis dieses Drei-Ebenen-Modells werden verschiedene horizontale Integrationsansätze, die Geschäftsprozess Integration, die Enterprise Application Integration und die Middleware Integration, verfolgt. Bei der horizontalen Integration werden Prozesse, Anwendungen und Daten, die bisher nebeneinander existierten, verbunden. Dadurch werden Geschäftsprozesse optimiert und Anwendungen sowie Daten effizient genutzt. Die dabei entstehenden Synergieeffekte wirken sich positiv auf das Unternehmen aus. Bei der vertikalen Integration werden Geschäftsprozesse durch die Ebenen vertikal auf der technischen Ebene abgebildet. Dadurch wird eine effektive Unterstützung der Geschäftsprozesse durch die IT-Landschaft erzielt. Im Folgenden wird der Bereich des Enterprise Application Integration (EAI) näher erläutert. Dabei wird auf die Eigenschaften heterogener Systemlandschaften eingegangen, um anschließend das Konzept der föderierten Informationssysteme zur Datenintegration zu erläutern. Neben dem föderierten Informationssystem existieren verschiedene andere Konzepte, wie z. B. Data-Warehouse Systeme oder serviceorientierte Architekturen, die sich prinzipiell für die Datenintegration eignen (Jung (2005), S. 82). Auf diese Konzepte wird in dieser Arbeit nicht näher eingegangen.

Unter der EAI-Bezeichnung werden Methoden und Konzepte zusammengefasst, mit denen die Kommunikation und die Interoperabilität zwischen verschiedenen Anwendungen und Geschäftsprozessen unternehmensintern und -übergreifend automatisiert und beschleunigt werden können (vgl. Winkeler et al. (2000), S. 7 f.; Keller (2002), S. 7). Dabei beschäftigt sich EAI stets mit den zwei Problembereichen technische Realisierung und organisatorische bzw. soziale Integration (vgl. Conrad et al. (2006), S. 5 f.). Das Ziel des EAI ist es, einen einheitlichen Zugriff auf Informationssysteme und Daten zu schaffen. Dies wird durch die Integration heterogener Informationssysteme erreicht. Damit ist ein weiteres Ziel, die Konsistenz der Daten zu sichern, verbunden (vgl. Conrad et al. (2006), S. 6 und S. 11). Eine integrierte Ebene über einer gewachsenen IT-Systemlandschaft mit ihren einzelnen Altanwendungen und Datenbeständen ist durch die drei Eigenschaften Heterogenität, Autonomie und Verteilung der integrierten Altanwendungen gekennzeichnet. Dadurch ist jedes Integrationsproblem auf diese drei

Aspekte zurückzuführen. Altanwendungen werden im Folgenden als Bestandteil eines integrierten Informationssystems auch als Komponenten bezeichnet.

Autonomie

In einer integrierten IT-Systemlandschaft bedeutet Autonomie, dass unterschiedliche Komponenten unabhängig voneinander Aufgaben des Unternehmens erledigen und Daten verarbeiten (vgl. Conrad et al. (2006), S. 14). Dies resultiert aus der Entwicklung einer Komponente für einen bestimmten Zweck im Unternehmen. Nach Conrad et al. (2006) treten dabei drei verschiedene Arten der Autonomie auf. Die Entwurfs-Autonomie besagt, dass verschiedene Komponenten unabhängig von anderen Komponenten entworfen werden und dadurch bspw. unterschiedliche Datenmodelle aufweisen. Kommunikation-Autonomie ist gegeben, wenn eine Komponente mit verschiedenen Komponenten kommunizieren kann und darüber die Entscheidungsgewalt hat. Ausführungs-Autonomie hat eine Komponente, wenn sie selbst über die Ablaufplanung eingehender Anfragen entscheiden kann.

Heterogenität

Die Heterogenität ist in syntaktische, datenmodellbasierte und logische Bereiche gegliedert und resultiert aus der autonomen Entwicklung von Altanwendungen (vgl. Conrad et al. (2006), S. 14 f.). In der Abb. 2.8 wird diese Hierarchie schematisch dargestellt.

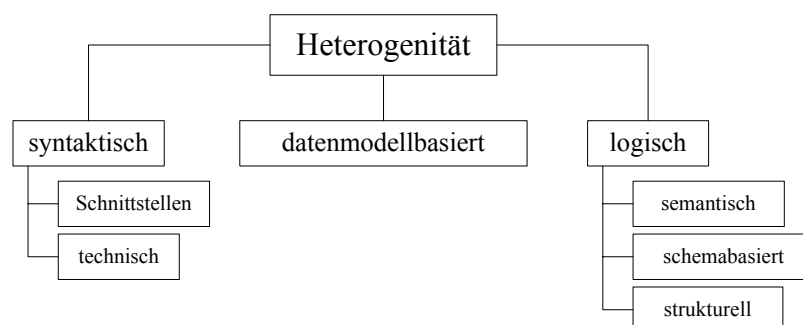


Abb. 2.8: Arten der Heterogenität

Die syntaktische Heterogenität bezieht sich auf technische Aspekte und Schnittstellen. Zur technischen Heterogenität gehören bspw. unterschiedliche Betriebssysteme oder Hardware-Plattformen. Die Schnittstellen-Heterogenität ist gegeben, wenn der Zugang zu einzelnen Komponenten über verschiedene Zugriffssprachen erfolgen muss. Die datenmodellbasierte Heterogenität beruht auf unterschiedlichen Paradigmen der Datenmodelle und damit auf der unterschiedlichen Semantik dieser Konzepte. So

kennen z. B. relationale Modelle das Prinzip der Vererbung von objektorientierten Modellen nicht. Die logische Heterogenität gliedert sich in semantische, schemabasierte und strukturelle Heterogenität. Daten und Schemata können, obwohl sie in derselben Notation formuliert sind, eine differenzierte Semantik aufweisen. Dieser Sachverhalt wird semantische Heterogenität genannt und basiert bspw. auf der unterschiedlichen Interpretation von Bezeichnungen. Diese semantischen Konflikte werden durch die Verwendung von Homonymen⁴ und Synonymen⁵ verursacht. Des Weiteren kann durch unterschiedliche Benutzer ein variierendes Verständnis der Begriffe entstehen. Die schemabasierte Heterogenität tritt auf, wenn Begriffe oder Sachverhalte durch unterschiedliche Modellierungskonstrukte des Datenmodells modelliert werden. Die strukturelle Heterogenität kann bei Elementen mit identischer Bedeutung auftreten, wenn sie semantisch homogen sind, im gleichen Datenmodell modelliert wurden, aber eine unterschiedliche Struktur aufweisen (vgl. Conrad et al. (2006), S. 15; S. 22 f. und S. 30 ff.).

Verteilung

Der Zugriff auf physisch verteilte Datenquellen kann aus technischer Sicht über Netzwerke unter Verwendung von Middleware erfolgen. Dadurch kann die physische Lokalisation der Datenquellen weitestgehend ignoriert werden und stellt somit heutzutage kein schwerwiegendes Problem dar (vgl. Conrad et al. (2006), S. 16).

2.2.1 Vorgehensweisen der Integration

Die Datenintegration ist die Basis für verschiedene Integrationsstrategien (vgl. Stickel (2001), S. 138). Da die Integrationsprobleme hauptsächlich auf der Heterogenität der zu integrierenden Altanwendungen und Datenquellen beruhen (Conrad et al. (2006), S. 22), müssen diese aufgelöst werden. Die Zusammenführung von unterschiedlichen Datenquellen erfolgt über die Schemaintegration. Dabei werden in der Literatur verschiedene Vorgehensweisen genannt, die im Folgenden beschrieben werden.

Die Top-Down Integration basiert auf der Vorgabe eines globalen Datenschemas, auch föderiertes Datenschema genannt. Das föderierte Datenschema wird wie ein neues Datenschema modelliert. In einigen Anwendungsbereichen sollte darauf geachtet werden, standardisierte Referenzschemata als föderierte Datenschemata zu verwenden (vgl. Conrad et al. (2006), S. 238). Diese haben verschiedene Vorteile. Durch die Nutzung von

⁴Eine Bezeichnung steht für verschiedene Begriffe, wie z. B. Bank steht für Geldinstitut und Sitzmöglichkeit.

⁵Verschiedene Bezeichnungen stehen für einen Begriff, wie z. B. Bücherei und Bibliothek.

Referenzschemata wird der Integrationsaufwand verringert und die Integration zwischen verschiedenen Unternehmen, die dasselbe Referenzschema verwenden, vereinfacht. Auf das föderierte Datenschema werden dann die lokalen Schemata der Komponentensysteme über Komponentenschemata und Exportschemata abgebildet. Durch diese Abbildung kann jedes lokale Schema als Sicht (Local-as-View) für das föderierte Schema verwendet werden (vgl. Conrad et al. (2006), S. 23 f. und S. 238 f.).

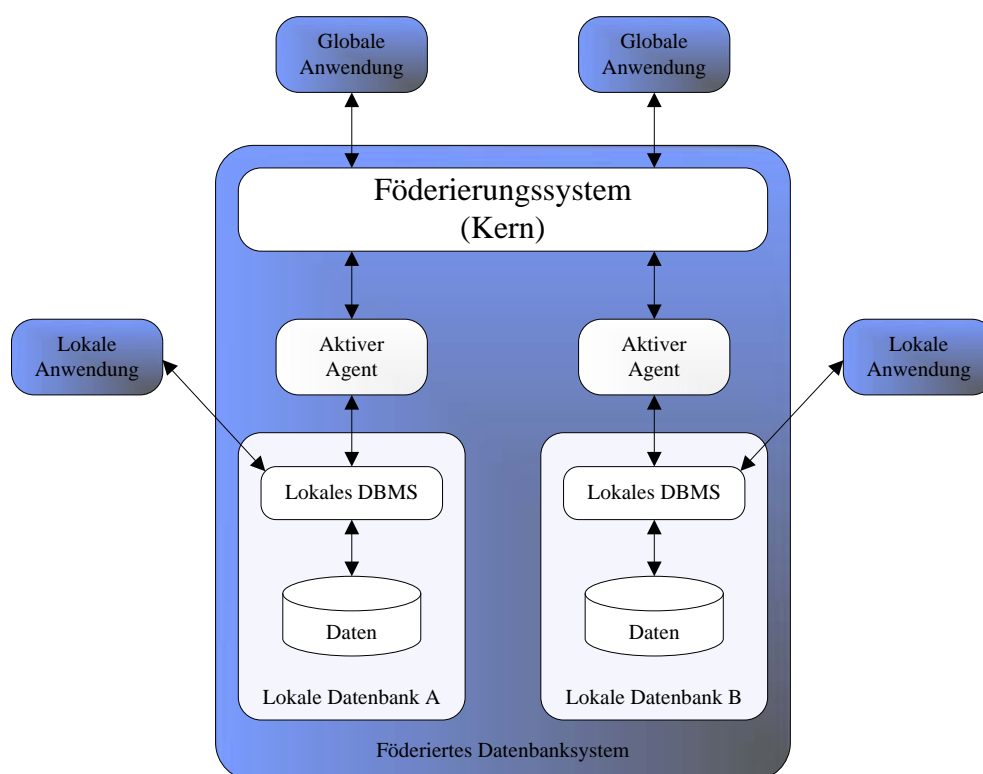
Bei der Bottom-Up Vorgehensweise werden zuerst die Komponentensysteme betrachtet und jedes lokale Schema in ein Komponentenschema transformiert. Dabei wird die Datenmodellheterogenität überwunden. Das bedeutet, jedem Komponentenschema liegt ein gemeinsames Datenmodell zugrunde. Für jedes Komponentenschema werden in einem Exportschema die Bereiche ausgewählt, die zu dem föderierten Datenschema integriert werden. Danach erfolgt die Analyse der Exportschemata. Gemeinsamkeiten werden festgestellt und die Schemaintegration zum föderierten Schema durchgeführt, bei der die strukturelle und semantische Heterogenität überwunden wird. Bei dieser Vorgehensweise werden die binäre und n-äre Integration unterschieden. Die binäre Integration erfolgt immer zwischen zwei Schemata. Entsprechend erfolgt die n-äre Integration zwischen mehr als zwei beteiligten Komponentensystemen gleichzeitig. Im n-ären Fall muss die Reihenfolge der zu integrierenden Schemata bestimmt werden. Hierbei gibt es die Möglichkeit des One Shot, d. h. alle Schemata werden gleichzeitig integriert. Eine andere Möglichkeit ist die iterative Vorgehensweise, bei der zwei oder mehrere Schemata Zwischenergebnisse bilden, welche wiederum schrittweise zum föderierten Schema integriert werden. Durch diese Vorgehensweise stellt das föderierte Schema eine Sicht über die integrierten lokalen Schemata (Global-as-View) dar (vgl. Conrad et al. (2006), S. 24 ff.).

Nach Conrad et al. (2006) werden die beiden Ansätze des Top-Down und Bottom-Up zum Jo-Jo-Ansatz kombiniert. Begonnen wird, wie beim Top-Down Ansatz, mit einem integrierten, globalen Datenschema. Dann wird zwischen den beiden Ansätzen gewechselt. Dieser Wechsel kann durch Anregungen aus der Bearbeitung der Altanwendungen zu Veränderungen des integrierten Start-Schemas führen. Die Kombination wird vorgeschlagen, um Synergieeffekte der beiden Ansätze für die Integration von alten und die Entwicklung von neuen Anwendungen zu nutzen (vgl. Conrad et al. (2006), S. 240 f.).

2.2.2 Föderierte Informationssysteme

Ein föderiertes Informationssystem ist ein Zusammenschluss von unterschiedlichen Komponentensystemen, die aus technischer Sicht, ihre Autonomie, Heterogenität und Verteilung bewahren. Es handelt sich beim föderierten Informationssystem um eine virtuelle Integration der beteiligten Datenquellen. Das bedeutet, die Daten werden nicht physisch integriert. Stattdessen wird „eine Schnittstelle auf ein unternehmensweites virtuelles Datenmodell geschaffen, unter dem alle Datenbanksysteme, unabhängig vom Hersteller, vom Modell und vom Schema, zusammengefasst werden“ (Reitz et al. (2003), S. 14). Somit wird für alle Anwendungen ein zentraler Zugriff auf die integrierten Daten ermöglicht.

Die allgemeine Architektur eines föderierten Informationssystems wird in Form eines föderierten Datenbanksystems in der Abb. 2.9 dargestellt.



Quelle: vgl. Conrad et al. (2006), S. 45

Abb. 2.9: Allgemeine Architektur föderierter Informationssysteme

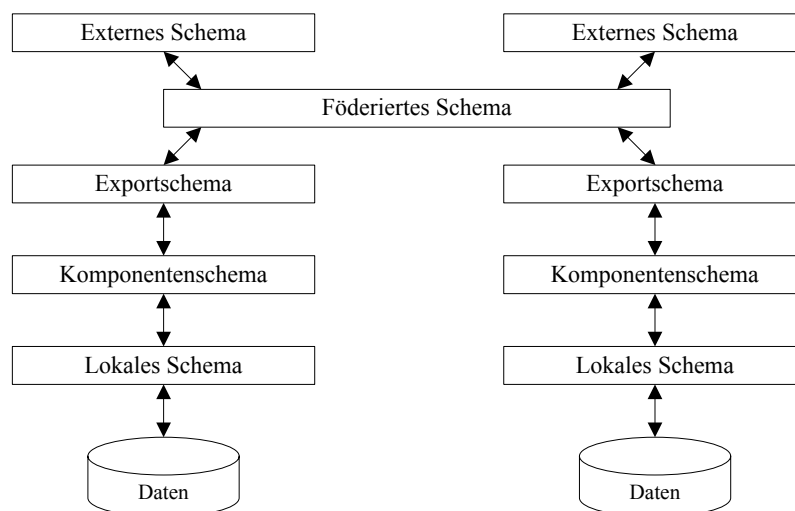
Der Zugriff globaler Anwendungen auf die lokalen Datenbankmanagementsysteme (DBMS) erfolgt über einen Föderierungskern. Für lokale Anwendungen ändert sich beim Zugriff auf die benötigten Datenbanken nichts. Der Föderierungskern ist ein übergeordnetes System und übernimmt verschiedene Integrationsaufgaben. Im Idealfall

stellt der Föderierungskerns ein föderiertes Datenbanksystem dar und implementiert die vollständige Funktionalität eines DBMS. Zu den Aufgaben eines DBMS gehören die Transaktionsverwaltung, Anfragebearbeitung und -optimierung, Integritätskontrolle und Recovery. Da der Aufwand zur vollständigen Umsetzung sehr hoch ist, wird in der Praxis „nur die Funktionalität im Föderierungskern realisiert, die im jeweiligen Anwendungskontext benötigt wird“ (Conrad et al. (2006), S. 46).

Die Kommunikation zwischen dem Föderierungskern und den lokalen DBMS erfolgt über spezielle Komponenten, die in der Abb. 2.9 als aktive Agenten bezeichnet werden. Sie übernehmen Übersetzungs- und Abbildungsaufgaben zwischen dem Föderierungskern und den Komponentensystemen und realisieren entfernte Systemaufrufe (vgl. Boles et al. (2004), S. 65). Je nach Aufgabenumfang können diese speziellen Komponenten auch andere Bezeichnungen, wie z. B. Adapter oder Wrapper haben. Der Vorteil dieser Komponenten liegt darin, dass „die Exportschnittstelle des zu integrierenden Systems nicht verändert werden muss, um die Kommunikation mit der 'Außenwelt' gewährleisten zu können“ (Winkeler et al. (2000), S. 24). Die aktiven Agenten wirken für die Komponentensysteme wie lokale Anwendungen und werden durch die Ausführungsautonomie der Komponentensysteme nicht bevorzugt (vgl. Conrad et al. (2006), S. 45).

Jedes Komponentensystem kann ein eigenes Datenmodell besitzen und Daten unterschiedlich strukturieren. Die Integration erfolgt über ein gemeinsames Schema auf Föderierungskernebene. Vereinfacht entspricht das Komponentensystem der Drei-Ebenen-Schema-Architektur, einem Standard für zentrale Datenbanksysteme des American National Standard Institut (ANSI). Der ANSI-Standard differenziert drei Ebenen. Die unterste Ebene bildet das interne Schema gefolgt vom konzeptuellen Schema und den externen Schemata als oberste Ebene. Das interne Schema beschreibt die physische Speicherung der Daten. Im konzeptuellen Schema werden die Daten in einem Datenmodell beschrieben, über das auf das interne Schema zugegriffen wird. Die externen Schemata sind Sichten (Views) auf das konzeptuelle Schema und erlauben Anwendungen einen Ausschnitt der Daten über das konzeptuelle Schema zu bearbeiten. Für föderierte Informationssysteme wird eine Fünf-Ebenen-Schema-Architektur, dargestellt in Abb. 2.10, verwendet.

Das lokale Schema entspricht dem konzeptuellen Schema des Komponentensystems und ist in dessen Datenmodell beschrieben. Darüber liegt das Komponentenschema, welches eine Übersetzung des lokalen Schemas in ein gemeinsames Datenmodell darstellt. Im Exportschema werden die Teile des Komponentenschemas bestimmt, die integriert werden sollen. Das föderierte Schema beschreibt den integrierten Datenbestand des



Quelle: vgl. Conrad et al. (2006), S. 49

Abb. 2.10: Fünf-Ebenen-Schema-Architektur

föderierten Informationssystems. Dabei werden alle Exportschemata berücksichtigt und redundante Daten für das föderierte Informationssystem virtuell aufgelöst. Die externen Schemata werden für globale Anwendungen definiert und enthalten die benötigte Sicht auf das föderierte Schema.

2.3 Computerlinguistik

Das Fachgebiet der Computerlinguistik (CL) erforscht die Methoden zur maschinellen Verarbeitung von natürlicher Sprache in gesprochener und geschriebener Form (vgl. Amtrup (2004), S. 2). Es setzt sich aus den Mutterdisziplinen Informatik und Linguistik zusammen. Die Methoden der Mutterdisziplinen dienen als Grundlage der computerlinguistischen Forschung. Allerdings hat die CL verschiedene eigenständige Methoden entwickelt und bietet dadurch ein sehr umfangreiches Forschungsfeld. Eine der wichtigsten Ressourcen der CL stellen Korpora dar. Korpora bestehen aus geschriebenen oder gesprochenen Texten, die mit linguistischen Informationen angereichert sind (vgl. Evert und Fitschen (2004), S. 406). Je nach Art des Inhalts werden die folgenden vier Korpustypen: Textkorpus, Sprachkorpus, multimedialer Korpus und Baubank unterschieden. Da in dieser Arbeit die Verarbeitung von geschriebener Sprache in Form der Präqualifikationsfragebögen notwendig ist, wird im Folgenden zuerst auf das Forschungsfeld Texttechnologie eingegangen. Darauf aufbauend wird die Ressource Textkorpora besprochen. Anschließend werden die Themenbereiche Information Retrieval und Information Extraction für die inhaltliche Erschließung und für die Suche nach Informationen in Textdokumenten vorgestellt.

2.3.1 Texttechnologie

Einen grundlegenden Teilbereich der CL stellt die Texttechnologie dar. Im Mittelpunkt steht die „linguistisch motivierte Informationsanreicherung und die Verarbeitung digital verfügbarer Texte durch standardisierte Auszeichnungssprachen“ (Rehm (2004), S. 138). Die Texttechnologie wird auch als „wissenschaftlich begründete Praxis“ (Lobin und Lemnitzer (2004), so zitiert in Rehm (2004), S. 138) bezeichnet. Dies begründet sich aus der Tatsache, dass die Texttechnologie, für konkrete Anwendungsfälle in unterschiedlichen Teilbereichen der CL, über keine ermittelbare Menge von aufeinander aufbauenden Theorien oder Methoden verfügt (vgl. Rehm (2004), S. 138). Die Schnittmenge und damit die einzige Gemeinsamkeit aller texttechnologischen Anwendungen stellt nach Rehm (2004) derzeit die XML dar. Die XML ist eine Metasprache, mit der verschiedene Markup- bzw. Auszeichnungssprachen spezifiziert werden können. Eine Auszeichnungssprache, die unter Verwendung der formalen Sprache XML entwickelt wurde, wird auch als XML-Anwendung bezeichnet.

Der Aufbau eines XML-Dokuments gliedert sich in Prolog und Dokumentelement (vgl. Young (2001), S. 46 f.). Der Prolog besteht aus der obligatorischen XML-Deklaration, der optionalen Dokumenttyp-Deklaration und optionalen Verarbeitungsanweisungen. Die Dokumenttyp-Deklaration wird auch als Dokumenttyp-Definition (DTD) bezeichnet. Die DTD beschreibt die Struktur eines XML-Dokuments und legt die als Markierungen verwendbaren Elemente fest (vgl. Young (2001), S. 111). Das Dokumentelement enthält den markierten Text.

Die Möglichkeiten der proprietären Syntax der DTD zur Spezifikation von Datenstrukturen und Restriktionen sind relativ eingeschränkt. Eine Alternative zur DTD stellt die Verwendung von Schemasprachen dar. Zu diesen Schemasprachen zählen bspw. XML-Schema, Relax NG oder Schematron. Mit deren Hilfe können komplexere Datentypen, Elemente und detailliertere Inhaltsmodelle modelliert werden. Ein weiterer Vorteil ist die Nutzung der XML-Notation innerhalb der Schemasprachen. Dadurch wird die durchgängige Verarbeitung mit unterschiedlichen XML-Werkzeugen ermöglicht (vgl. Rehm (2004), S. 142).

In der CL werden Texte mit linguistischen Informationen angereichert. Die Auszeichnung (oder Annotation) von arbiträren Informationen in Texten erfolgt durch Markupssprachen. Dabei stellt eine arbiträre Information eines Präqualifikationsfragebogens bspw. die Zeichenkette *Mitarbeiter* dar. Die linguistischen Informationen, die zu dieser Zeichenkette markiert werden, sind bspw. Wortart, Kasus oder Numerus. Die Anwendung von XML als Datenaustauschformat und die daraus resultierende einheitliche Darstellung von Ressourcen, ist ein weiteres Anliegen der CL. Die entstehenden Ressourcen sollen

als Datengrundlage für unterschiedliche computerlinguistische Anwendungen verwendet werden. Ein wesentliches Merkmal der Informationsmodellierung mit XML ist die strikte Abkopplung von Inhalt und Struktur eines Textes (vgl. Rehm (2004), S. 140). Dies wird durch eine Abstraktionsebene erreicht. An Stelle der, in WYSIWYG-Editoren (What You See Is What You Get) benutzten Formatierung der Textstellen mit Textsatz- oder Layout-Anweisungen, werden bei Auszeichnungssprachen die Textteile von Auszeichnungselementen⁶ umschlossen. Diese werden dann bei der Generierung einer Ausgabe-Version interpretiert und umgesetzt. Dadurch liegt der Focus innerhalb eines spezifischen Dokumententyps auf der Festlegung der logischen bzw. semantischen Funktion einer Textstelle und nicht auf deren typografischen Gestaltung (vgl. Rehm (2004), S. 138).

Das bekannteste Beispiel für die Anwendung von XML im Bereich der CL ist die automatische oder manuelle Annotation von Korpora. Dafür spezifiziert die Text Encoding Initiative (TEI) verschiedene DTD, die die Annotation auf vier Ebenen ermöglichen. Die erste Ebene spiegelt die Metadaten von Dokumenten wieder. Auf der zweiten Ebene werden textuelle Einheiten wie Kapitel oder Abschnitte erfasst. Innerhalb dieser textuellen Einheiten wird dann auf der dritten Ebene die Struktur von Sätzen und Wörtern markiert. Die vierte Ebene umfasst die Markierung syntaktischer oder morphologischer Einheiten (vgl. Rehm (2004), S. 146).

2.3.2 Textkorpora

Eine wichtige Eigenschaft von Textkorpora ist deren textuelle Zusammensetzung. Diese kann heterogen oder homogen in Bezug auf die Textsorte, die Domäne oder das Alter eines Textes sein (vgl. Evert und Fitschen (2004), S. 407). Der Umfang und die Textqualität sind von entscheidender Bedeutung für die praktische Verwendung und werden durch den Erstellungsprozess beeinflusst. Die Qualität des Textes wird durch die Tipp-, Grammatik- und Ausdrucksfehler oder die Fehler der linguistischen Vorverarbeitung, bestimmt. Der Erstellungsprozess eines Textkorpus kann, auf Grund der Tatsache das Texte aus Inhalt und Formatierungen bestehen, nicht vollständig automatisch durchgeführt werden (vgl. Evert und Fitschen (2004), S. 407). Es existieren keine einheitlichen Formatierungsregeln. Daraus resultieren unterschiedliche Arten der Kodierung. Dennoch lässt sich das Vorgehen abstrahieren und in die Phasen Aufbereitung, Vorverarbeitung und Anreicherung mit linguistischen Informationen einteilen (vgl. Evert und Fitschen (2004), S. 408 ff.).

⁶Auszeichnungselemente werden auch als Tags, Auszeichner, Markierungen, Elemente, Marken oder Etiketten bezeichnet.

In der Aufbereitung werden strukturelle Merkmale und Metainformationen des Quelltextes erkannt und in ein definiertes einheitliches Format (meist XML) umgewandelt. Dieser Schritt kann im Idealfall, wenn ein Text in einem standardisierten Format mit bekannter Auszeichnungssymbolik vorliegt, entfallen (vgl. Evert und Fitschen (2004), S. 408). Danach folgt die Vorverarbeitung, die sich in die Teilaufgaben Tokenisierung und die Satzgrenzenerkennung aufteilt. Durch die Tokenisierung (auch Segmentierung genannt) wird der Text in kleinere (Sätze) oder kleinste (Wörter) Einheiten zerlegt. Im Textkorpus ist die Grundeinheit das Token und stellt eine, von Leerzeichen oder Interpunktionen umgebene, Buchstaben- oder Ziffernfolge dar. Dabei können einige Probleme, wie Sonderzeichen oder zusätzliche Leerzeichen bei großen Zahlen, auftreten (vgl. Evert und Fitschen (2004), S. 409). Im Anschluss an die Tokenisierung erfolgt, unter Verwendung von Methoden zur Satzgrenzenerkennung, die Disambiguierung von Punkten als Satzende oder Tokenbestandteil.

Für die Anreicherung der Token mit linguistischen Informationen werden die Verfahren Part-of-Speech-Tagging, Lemmatisierung, Chunking und Parsing verwendet. Mit Hilfe des Part-of-Speech-Tagging lassen sich durch stochastische oder regelbasierte Tagger die Wortarten der Token bestimmen (vgl. Evert und Fitschen (2004), S. 411). Bei stochastischen Taggern werden Trainingskorpora benötigt, auf denen bestimmte Wortfolgen trainiert werden. Dabei ist zu berücksichtigen, dass je nach Textsorte und Domäne speziell angepasste Korpora unabdingbar für gute Ergebnisse sind (vgl. Evert und Fitschen (2004), S. 411). Zusätzlich wird beim Tagging ein Lexikon verwendet, um die Grundform der Worte zu bestimmen. Dieses Verfahren wird Lemmatisierung genannt und hat den Vorteil, dass bei späteren Anfragen nicht jede flektierte Form berücksichtigt werden muss. Auch in diesem Abschnitt können Probleme auftreten, weil es nicht immer eindeutig ist auf welcher Grundform ein Wort basiert. Beim Chunking und Parsing werden Tokengruppen mit strukturellen linguistischen Informationen versehen. Dazu gehören bspw. Nominalgruppen (die große, weite Welt), Verbalgruppen (ist gesehen worden) oder Phrasenstrukturen (vgl. Evert und Fitschen (2004), S. 412).

2.3.3 Information Retrieval und Information Extraction

Die grundlegenden Aufgaben des Information Retrieval beziehen sich auf die Indexierung und die Nutzung des Indexes, um Dokumente und die in ihnen gespeicherten Informationen wiederzufinden (vgl. Lorenz (2006), S. 22). Dabei kann die Anfrage selbst, durch die Suchbegriffe, als ein Dokument aufgefasst werden. Zu diesem Dokument wird ein Index erzeugt und bei der Suche mit den anderen Dokumentenindexen verglichen. Dadurch wird beim Information Retrieval die Ähnlichkeit von Dokumenten

zum Suchdokument bestimmt und die Dokumente mit den höchsten Ähnlichkeitsfaktoren in einer Liste dargestellt (vgl. Lorenz (2006), S. 22). Ob die Dokumente tatsächlich die benötigten Informationen enthalten, muss der Anfragende manuell ergründen.

Im Gegensatz dazu, wird beim Information Extraction direkt nach Informationen gesucht und diese ausgegeben. Information Extraction (IE) bezeichnet also Verfahren, die „gezielt domänenspezifische Informationen aus freien Texten aufspüren und strukturieren können, bei gleichzeitigem 'Überlesen' irrelevanter Informationen“ (Neumann (2004), S. 502). Dafür werden im Vorfeld genau die relevanten Informationen, die extrahiert werden sollen, in Lexikoneinträgen und Regeln spezifiziert. Die benötigten Informationen werden in Form von Templates modelliert und diese von IE-Systemen auf eine Menge von Textdokumenten angewendet. Das Ergebnis sind, mit den entsprechenden Informationen, instanziierte Templates, die dann für die Weiterverarbeitung genutzt werden können. Dabei gilt es generell, aber besonders im Deutschen, sprachliche Probleme zu berücksichtigen. Gleiche Sachverhalte können in unterschiedlichen Schreibweisen ausgedrückt werden. Aus diesem Grund sind Wortlisten oder Assoziationsregeln, welche die komplette Vielfalt erfassen, unrealistisch (vgl. Lorenz (2006), S. 80). Weitere prinzipiell lösbare Probleme sind lexikalische Analysen, Namenserkennung, partielle Syntaxanalysen und Referenzauflösungen. Diese haben durch die Komplexität der natürlichen Sprache zu einem „Kompromiss zwischen theoretischen Ansprüchen und pragmatischen Anforderungen“ (Neumann (2004), S. 505) geführt. Dieser Kompromiss führte zur Entwicklung und Verwendung von flachen Textverarbeitungsmethoden. Zu diesen gehören unter Anderen die bereits erwähnten Methoden Tokenisierung, Wortart-Tagging und Chunk-Parsing (vgl. Halama (2004), S. 218 ff.).

Die Aufgaben, die ein IE-System zu leisten hat, lassen sich in folgende, teilweise aufeinander aufbauende, Bereiche gliedern (vgl. Granitzer (2006), S. 28):

- Erkennung benannter Entitäten (Named Entity Recognition (NE)), z. B. Eigennamen, Orte oder Titel ,
- Auflösung von Koreferenzen (Co-Reference Resolution (CO)), bspw. persönliche Pronomen wie *er* oder *sie*,
- Füllung von Schablonen-Elementen (Template Element Filling (TE)), die beschreibende Informationen zu den in NE ermittelten Entitäten aufnehmen,
- Ermittlung von Schablonen-Relationen (Template Relations (TR)), bspw. Bezug einer Person zu einem Unternehmen sowie
- Identifizierung von Szenario-Schablonen (Scenario Templates (ST)), z. B. Szenarios und Ereignisse, an denen Ergebnisse aus TE und TR beteiligt sind.

Für diese Aufgaben werden maschinelle Lernverfahren und linguistische Analysen oder grammatikbasierte Ansätze verwendet, bei denen bspw. Gazetteers oder Lookup-Listen verwendet werden. Diese beinhalten die erkennbaren Entitäten (z. B. Ortsnamen). Die Vorteile liegen in der Sprachunabhängigkeit und hohen Verarbeitungsgeschwindigkeit. Als nachteilig sind das Problem der Vollständigkeit, die Nichterfassung morphologischer Variationen, der hohe manuelle Erstellungs- und Wartungsaufwand zu erwähnen (vgl. Granitzer (2006), S. 28 f.). Als ein Beispiel für eine vollständige Umgebung soll die Open Source Software GATE⁷ (General Architecture for Text Engineering) genannt werden. Dieses Framework bietet verschiedene Verfahren zur Analyse von Texten und Werkzeuge zur Erstellung von IE-Anwendungen (vgl. Granitzer (2006), S. 28) und wird von vielen kommerziellen Anwendungen als Basis genutzt (vgl. Lorenz (2006), S. 82).

Das IE ist sehr aufwendig und problembehaftet. Es stellt aber dennoch eine Möglichkeit für Umsetzung einer inhaltlichen Erschließung von Textdokumenten, wie die Präqualifikationsfragebögen, dar.

⁷[http : //gate.ac.uk/](http://gate.ac.uk/)

3 Anforderungsanalyse

In der Anforderungsanalyse werden die Anforderungen des zu entwickelnden Informationssystems analysiert und beschrieben. Diese Beschreibungen dienen der im Anschluss folgenden Entwicklung des Fachkonzeptes (siehe Kapitel 4). Die Anforderungsanalyse bildet ebenfalls die Grundlage, um die korrekte Erstellung des Fachkonzeptes zu prüfen. Dafür werden in der Anforderungsanalyse Kundenwünsche erfasst und Anforderungen an das Informationssystem abgeleitet. Um den Informationsbedarf des Prozesses der Präqualifikationsfragebogenbeantwortung zu ermitteln, ist es notwendig, eine Informationsbedarfsanalyse durchzuführen.

Dieses Kapitel gliedert sich inhaltlich in drei Teile und beschäftigt sich mit der Anforderungs- und Informationsbedarfsanalyse. Die gewählte Vorgehensweise entspricht der des Ist-Ansatzes (vgl. Abschnitt 2.1.4). Der Ist-Ansatz wird mit der Konferenzmethode, Interview-Technik, Dokumentenanalyse, Aufgaben- und Problemanalyse kombiniert. Dabei stellt die Aufgaben- und Problemanalyse den Übergang zum Fachkonzept dar. Im ersten Teil (Abschnitt 3.1) wird das betriebswirtschaftliche Problem hinter dem zu entwickelnden Fachkonzept beschrieben. Diese Erkenntnisse entstanden in einem Workshop unter Verwendung der Konferenzmethode. Dieser Workshop wurde mit vier kmU des Maschinen- und Anlagenbaus durchgeführt. Des Weiteren werden in diesem Abschnitt die Anforderungen an ein Informationssystem zur Unterstützung der Beantwortung von Präqualifikationsfragebögen (PQF) und der diesem Vorgang zugrunde liegende Ist-Prozess skizziert. Diese Ergebnisse wurden durch Interviews im Rahmen der PQF-Beschaffung verfeinert. Der zweite Teil dieses Kapitels beschäftigt sich mit der Dokumentenanalyse des Beantwortungsprozesses. Dabei wird im Abschnitt 3.2 auf die Vorgehensweise und die Durchführung der Dokumentenanalyse eingegangen. Danach erfolgt die Darstellung der Ergebnisse. Den Abschluss dieses Kapitels bildet eine Zusammenfassung der Ergebnisse und Schlussfolgerungen für das Fachkonzept.

3.1 Ist-Analyse des Beantwortungsprozesses

Die im Workshop definierten Anforderungen an das zu entwickelnde Informationssystem und der zugrunde liegende Prozess werden im Folgenden dargestellt. Um einen ersten Eindruck zu bekommen, wurde eine Ist-Analyse des Beantwortungsprozesses durchgeführt. Zu diesem Zweck wurden alle beteiligten Unternehmen in einem Workshop zum Ablauf des Prozesses befragt und die Ergebnisse anschließend diskutiert.

Ein Unternehmen wird mit den PQF in zwei Fällen konfrontiert. Der erste Fall geht vom Unternehmen selbst aus, indem das Unternehmen sich für ein ausgeschriebenes Projekt bewirbt und damit am Verfahren zur Präqualifikation des ausschreibenden Unternehmens teilnimmt. In diesem PQ-Verfahren wird dem anfragenden Unternehmen der PQF übermittelt. Im zweiten Fall wird ein PQ-Verfahren von einem Unternehmen mit oder ohne Projektbezug durchgeführt. Dieses PQ-Verfahren dient zum Aufbau einer neuen Bieterliste oder zu deren Aktualisierung. Der erste Fall ist nachfrageinitiiert ähnlich dem Pull-Prinzip, während der zweite Fall dem Push-Prinzip⁸ entspricht. In beiden Fällen wird der Informationsbedarf durch den PQF ausgelöst. Der ablaufende Prozess im Unternehmen, das den PQF beantwortet, ist in beiden Fällen identisch (siehe Abb. 3.1).

Der PQF wird per E-Mail, Fax oder postalisch an das Unternehmen gesendet. Ein Mitarbeiter nimmt den PQF entgegen und beginnt mit dessen Beantwortung. Er prüft, ob von diesem Unternehmen schon einmal ein PQF eingegangen ist und beantwortet wurde. Wenn ein früherer PQF vorhanden ist, können dessen Antworten eventuell wieder verwendet werden. Dies würde eine enorme Zeitersparnis bedeuten, da der Mitarbeiter nur noch die nicht mehr aktuellen Antworten des PQF überarbeiten müsste. Allerdings findet diese Prüfung, nach Angaben der Unternehmen, nicht vollständig bzw. gar nicht statt. Die Gründe hierfür liegen zum einen darin, dass bei der Suche meist nur die Papierablage des jeweiligen Büros einbezogen wird. Das bedeutet, es werden nur die Ordner mit aktuellen Projekten im Büro des Bearbeiters berücksichtigt, zum anderen sind ältere Dokumente entweder nicht mehr vorhanden oder sie lagern in einem Archiv. Diese meist nicht digitalen Archive werden vom Mitarbeiter nicht aufgesucht, weil ihm keine oder nur mangelhafte Informationen zu den eventuell vorhandenen PQF zur Verfügung stehen. Dadurch ist nach Angaben der Unternehmen der Suchaufwand für ein veraltetes eventuell nicht existierendes Dokument zu hoch.

Nachdem diese Prüfung abgeschlossen ist, wird der PQF gelesen und die einzelnen Fragen beantwortet. Dieser Prozess ist durch das Lesen und Verstehen der Inhalte der Fragen durch den Mitarbeiter gekennzeichnet. Im ersten Schritt bei der tatsächlichen Beantwortung wird die Frage gelesen. Danach prüft der Mitarbeiter, indem er nach den Informationen sucht, ob die benötigten Informationen⁹ und Nachweis-Dokumente zur Beantwortung verfügbar sind. Die möglichen Informationsquellen des Bearbeiters sind sein eigenes Wissen, der eventuell gefundene PQF, sein Arbeitsplatz, sein Rechner, das Netzwerk und Informationssysteme, auf die der Mitarbeiter zugreifen kann. Wenn es

⁸Vgl. in Abschnitt 2.1.2 Push- und Pull-Prinzip.

⁹Typischer Inhalt der Fragen, die von den Unternehmen im Workshop benannt wurden, sind bspw. Referenzen, Zertifikate oder Angaben zum Umsatz. Genauere Aussagen zum Informationsbedarf konnten im Workshop nicht ermittelt werden.

dem Bearbeiter nicht möglich ist die Frage selbst zu beantworten, muss er die Suche ausdehnen.

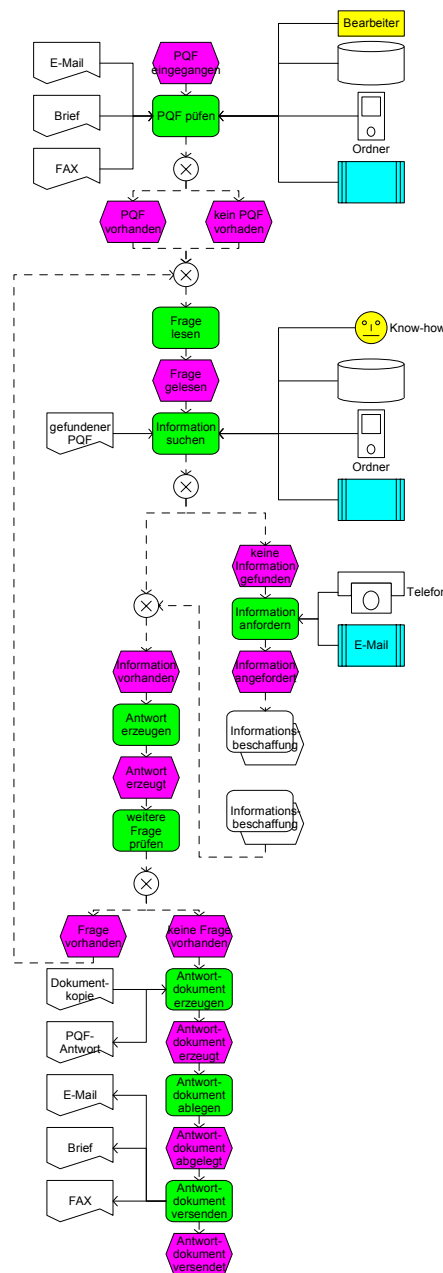


Abb. 3.1: Ist-Analyse der Beantwortung eines Präqualifikationsfragebogens

Der Bearbeiter versucht den Mitarbeiter im Unternehmen zu erreichen, der die Frage beantworten oder die benötigten Informationen beschaffen kann. Bei der Beantwortung hat der Bearbeiter also zwei Probleme hinsichtlich des Zugriffs auf die Informationen zu lösen. Wo ist die benötigte Information und wie kann auf die Information zugegriffen werden? Die Fragen nach dem Wo bzw. Wer, kann mit Hilfe von Erfahrung gelöst werden. In der Buchhaltung werden bspw. Angaben zum Umsatz verwaltet. Das zweite Problem wird unter Verwendung eines Adressbuches gelöst. Wenn der Mitarbeiter, der um die

Informationen gebeten werden soll krank bzw. im Urlaub ist, kann es passieren, dass der Bearbeiter auf die Antworten warten muss. Dieser Umstand verzögert die Beantwortung zusätzlich zur aufwendigen Suche. Der um die Informationen gebetene Mitarbeiter stellt neben seiner täglichen Arbeit die gewünschten Informationen zusammen, kopiert gegebenenfalls benötigte Nachweis-Dokumente und sendet diese an den Bearbeiter des PQFs zurück.

Die Beantwortung der PQF erfolgt digital, wenn der PQF in einem digital bearbeitbaren Format vorliegt. Ist dies nicht der Fall wird der PQF handschriftlich ausgefüllt. In beiden Fällen werden die Informationen vom Bearbeiter aufbereitet und in die gewünschte Form transformiert. Benötigte Nachweis-Dokumente werden, sofern sie noch nicht als Kopie vorliegen, kopiert oder eingescannt. Danach werden die Antworten in den PQF übertragen und die gewünschten Nachweis-Dokumente als Anlage hinzugefügt. Der beantwortete PQF wird abgelegt und an das anfragende Unternehmen per E-Mail, Fax oder Brief zurückgesendet. Die Ablage erfolgt auf unterschiedliche Arten. Elektronisch beantwortete Präqualifikationsfragebögen werden gespeichert und/oder ausgedruckt. Die nicht digital ausgefüllten PQF werden kopiert oder eingescannt.

Vor dem Hintergrund des beschriebenen Ist-Prozesses werden im Folgenden die Anforderungen an das zu entwickelnde Informationssystem erläutert. Der beschriebene Ablauf zeigt, dass der Beantwortungsprozess in einer heterogenen Systemlandschaft stattfindet und Unterstützung bzw. Informationen aus verschiedenen Bereichen des Unternehmens benötigt. Jeder Bereich kann unterschiedliche Ablagen und Systeme haben, in denen die gewünschten Informationen verwaltet werden. Daraus resultiert für die Bearbeitung der PQF ein hoher Zeitaufwand, welcher reduziert werden kann und sollte. Der verantwortliche Bearbeiter dieses Prozesses gehört dem Management an. Er kann Teilaufgaben an andere Mitarbeiter delegieren. Das zu entwickelnde Informationssystem soll den Bearbeiter eines Präqualifikationsfragebogens bei dessen Beantwortung unterstützen.

Der Prozess sollte effektiv und effizient durch ein Informationssystem unterstützt werden. Effektiv ist ein Informationssystem, wenn es die geforderten Funktionen aufweist und korrekt funktioniert. Effizient ist ein Informationssystem, wenn es mit gegebenen Betriebsmitteln sparsam umgeht (vgl. Ludewig und Lichter (2007), S. 42). Die geforderte Funktionalität ist eine semi-automatische Beantwortung der Präqualifikationsfragebögen, um dadurch den Bearbeitungsprozess zu verkürzen. Daraus soll sich ein Kosteneinsparungspotenzial erschließen und eine Entlastung des Führungspersonals und der beteiligten Mitarbeiter ergeben. Um dies zu erreichen, müssen die verschiedenen Informationsquellen der heterogenen Systemlandschaft integriert

werden. Die Bestandteile der heterogenen Systemlandschaft sind nicht bekannt, daher sollte das Konzept flexibel darauf reagieren können und plattformunabhängig sein. Der PQF kann in unterschiedlichen digitalen Formaten und in Papierform vorliegen. Daher ist die Verarbeitung unterschiedlicher Formate notwendig, damit das Informationssystem die PQF elektronisch verarbeiten kann. Zur Verarbeitung gehören die Digitalisierung, die Archivierung, das Retrieval, die Analyse der Fragen, die Suche von Informationen sowie der Vorschlag und die Generierung von Antworten. Der Bearbeiter sollte auf alle benötigten Informationen von seinem Arbeitsplatz aus zugreifen können.

3.2 Analyse der Präqualifikationsfragebögen

Das zentrale Dokument im Beantwortungsprozess stellt der PQF dar. Dieser enthält das Informationsbedürfnis des anfragenden Unternehmens. Dadurch wird die externe Informationsnachfrage für das Unternehmen, welches den PQF beantwortet, zum internen objektiven Informationsbedarf. Dieser Informationsbedarf soll durch das zu konzipierende Informationssystem gedeckt werden. Um eine genauere Vorstellung dieses Informationsbedarfes zu erlangen werden Präqualifikationsfragebögen, die aktuell von Unternehmen aus der Maschinen- und Anlagenbaubranche für die Präqualifikation genutzt werden, analysiert. Die Methodik und die Durchführung dieser Analyse werden im folgenden Abschnitt vorgestellt.

3.2.1 Vorbereitung der Analyse

Um den Informationsbedarf zu detaillieren, wurden Präqualifikationsfragebögen beschafft, analysiert und ausgewertet. Die Beschaffung der Fragebögen erfolgte in zwei Phasen. In der ersten Phase wurden, um die Kosten für diese Untersuchung so gering wie möglich zu halten und dennoch eine Arbeitsgrundlage für erste statistische Analysen zu erhalten, 30 Unternehmen des Maschinen- und Anlagenbaus ausgewählt. Unter diesen Unternehmen befanden sich Generalunternehmen, Subunternehmen und Kunden (Anlagenbetreiber). Die Anfrage erfolgte per E-Mail, in der die Unternehmen gebeten wurden, ihren eigenen Präqualifikationsfragebogen und die ihrer Kunden, sofern dieses gestattet war, dieser Untersuchung zur Verfügung zu stellen. Das Resultat der ersten Phase waren fünf Antworten. Drei Unternehmen waren nicht gewillt die Untersuchung zu unterstützen. Die anderen zwei Unternehmen übersendeten insgesamt fünf Präqualifikationsfragebögen. Auf Grund der geringen Rücklaufquote wurden in der zweiten Phase Telefonate mit den jeweiligen Unternehmen geführt. Bei diesen Telefonaten wurden zum einen das Thema dieser Arbeit und die Notwendigkeit der

PQF erläutert und zum anderen weitere Interviews zum Präqualifikationsprozess im Maschinen- und Anlagenbau geführt.

Insgesamt wurden daraufhin 26 unterschiedliche Präqualifikationsfragebögen von 15 Unternehmen zur Verfügung gestellt, wobei einige Unternehmen selbst keine Präqualifikation durchführen, aber mehrere Fragebögen ihrer Auftraggeber übermittelt haben. Für die PQF wurde eine Vorabanalyse durchgeführt, um die Beschaffenheit und Verwendbarkeit des Datenmaterials zu prüfen. Hierbei wurde eine stark abweichende Granularität des Informationsbedarfs einiger Fragebögen festgestellt. Um Erkenntnisse über die Ursache dieser Abweichungen zu gewinnen, wurden die Unternehmen und die dazugehörigen Präqualifikationsfragebögen in die Klassen Subunternehmen, Generalunternehmen, Kunde und Akkreditierungsstelle eingeteilt. Die Ergebnisse dieser Einteilung sind in der Tab. 3.1 dargestellt.

Tab. 3.1: Einteilung der Präqualifikationsfragebögen

PQF-Typ	Subunternehmen	Generalunternehmen	Kunde	Akkreditierungsstelle
PQF-Anzahl	7	16	2	1

Auf Grundlage dieser Einteilung wurden die Präqualifikationsfragebögen erneut untersucht und eine mögliche Erklärung für die unterschiedliche Granularität des Informationsbedarfs gefunden. Die PQF der Subunternehmen an deren Subunternehmen bspw. Lieferanten sind in der Granularität sehr inhomogen. Im Vergleich zu den PQF der Generalunternehmen und Kunden fällt der Informationsbedarf der Subunternehmen mehrheitlich geringer aus. Das Spektrum reicht von einfachen Fragen zu bestimmten Materialien und Herstellungsverfahren bis zur Granularität der Fragebögen von Generalunternehmen. Eine Ursache dafür ist, dass einige Subunternehmen, wenn sie Leistungen kaufen, die PQF der Generalunternehmen entweder komplett oder auszugsweise an ihre Lieferanten weiterreichen. Es gibt aber auch Subunternehmen, die ihren eigenen PQF entwickeln. Diese werden für den Aufbau von Lieferantenlisten, zur Prüfung der Lieferanten oder zur Sondierung des Lieferantenmarktes benutzt. Diese Listen sind für die kmU im Maschinen- und Anlagenbau besonders wichtig, um auf kurzfristig eingehende Auftragsanfragen, bei denen wenig Zeit für die Kalkulation und die Zusammensetzung der Kooperationspartner vorhanden ist, reagieren zu können.

Die PQF der Generalunternehmen spiegeln einen geringeren Informationsbedarf als der der Kunden, aber generell einen höheren Informationsbedarf als der der Subunternehmen wieder. Sie beziehen sich auf konkrete Projekte oder dienen dem Aufbau einer Bieterliste ohne Projektbezug. Zusätzlich zum Informationsbedarf der Subunternehmen erstreckt sich der Informationsbedarf mit Projektbezug auf bestimmte Arbeiten im Detail, z. B.

das Schweißen von Plastik, die Prüfung mit bestimmten Geräten, deren Zertifizierung oder die beauftragten Subunternehmen. Dienen die Präqualifikationsfragebögen dem Aufbau einer Bieterliste ohne Projektbezug, ist die Granularität des Informationsbedarfs geringer, z. B. die Angabe des Leistungsspektrums ohne speziell nach Einzelheiten des Plastikschiweißens zu fragen.

Die Präqualifikationsfragebögen der Kunden beziehen sich auf konkrete Projekte und sind in der Regel an Generalunternehmen gerichtet. Zusätzlich zum Informationsbedarf der Generalunternehmen fordern die Kunden genaue Informationen zu Planungsaktivitäten, Vorplanung und detaillierte Informationen über die Subunternehmen, die am Generalunternehmen gebunden sind. Die PQF der Kunden weisen dadurch den höchsten Informationsbedarf auf und sind dementsprechend am umfangreichsten.

Nach dieser Darstellung kann gesagt werden, dass der gesamte Informationsbedarf der Präqualifikation im Maschinen- und Anlagenbau von den Kunden bestimmt wird. Jede PQ-Stufe reicht ihren (Teil-)Informationsbedarf in Form der Präqualifikationsfragebögen an die nächst niedrigeren Stufe weiter. Der Aufbau der von den Generalunternehmen verwendeten PQF ähnelt einem Baukasten. Es gibt einen allgemeinen PQF, der für alle Gewerke¹⁰ verwendet wird. Dieser enthält den allgemeinen Informationsbedarf. Ergänzt wird der allgemeine PQF um gewerkespezifische und wahlweise projektspezifische PQF. In dieser Arbeit wird der allgemeine Informationsbedarf betrachtet, um ein breit einsetzbares Basis-Informationssystem zu schaffen, welches später um gewerke- und branchenspezifische Eigenheiten ergänzt werden kann.

Mit Hilfe dieser Erkenntnisse und dem Einsatzzweck der zu spezifizierenden Software, welche die kmU im Maschinen- und Anlagenbau bei der Bewältigung des Präqualifikationsprozesses unterstützen soll, wurden Entscheidungen über die Verwendung des Datenmaterials für die weitere Analyse getroffen. Die PQF der Kunden werden nicht verwendet, zum einen liegt dies in der zu hohen Granularität des Informationsbedarfs, zum anderen werden die kmU selten mit den PQF der Kunden, sondern vielmehr mit PQF von Generalunternehmen konfrontiert. Der Fragebogen der Akkreditierungsgesellschaft wird ebenfalls nicht verwendet, da es sich um eine direkte Zertifizierung mit gesetzlichen Regelungen handelt, die in keinem der anderen PQF vorkommen. Des Weiteren werden zwei PQF der Subunternehmen und ein PQF der Generalunternehmen ausgeschlossen. Die PQF der Subunternehmen haben eine zu geringe Granularität und würden im weiteren Verlauf der Arbeit das Ergebnis der Häufigkeitsanalyse verfälschen. Der ausgeschlossene PQF des Generalunternehmens ist nicht für die Aufnahme in eine Bieterliste geeignet, es handelt sich hierbei um einen

¹⁰Als Gewerke werden die einzelne Arbeitsbereiche im Handwerk bezeichnet (vgl. Kleinbauer et al. (2006), S. 73).

zusätzlichen Fragebogen zur Erlangung eines firmeninternen Zertifikates. Durch diese Selektion werden 17 Präqualifikationsfragebögen bei der Untersuchung berücksichtigt.

3.2.2 Vorgehensweise der Analyse

Die Dokumentenanalyse erfolgt durch eine inhaltliche Erschließung der PQF. Um den tatsächlichen Informationsbedarf der PQF zu ermitteln, werden alle strukturellen Bereiche, wie Mitarbeiter oder Referenzen, alle darin enthaltenen Dokumente und Informationsobjekte mit deren Attributen, in Tabellen erfasst. In diesen Tabellen werden die absoluten Häufigkeiten des Vorkommens aller Objekte pro untersuchten PQF notiert. Diese Untersuchung ergab 663 Informationsobjekte einschließlich der Dokumente und Attribute in verschiedenen Bereichen. Diese relativ hohe Anzahl der Informationsobjekte resultiert aus unterschiedlichen teilweise synonymen Bezeichnungen, Schreibweisen und unterschiedlicher Einordnung in die Bereiche der PQF. Um die Anzahl beherrschbar zu gestalten und die Qualität der Informationsobjekte zu erhöhen, werden diese Defekte durch abstrahieren, umsortieren und zusammenfassen aufgelöst. Das Ergebnis dieses Analyseschrittes sind 262 Informationsobjekte und 23 Dokumente in 21 Bereichen. Wobei die Dokumente zusätzlich in einer Tabelle separat aufgeführt werden, um die Betrachtung und Beschreibung zu erleichtern.

Im nächsten Schritt werden unter Verwendung der Häufigkeitsanalyse die relativen Häufigkeiten der Informationsobjekte, Dokumente und Bereiche der PQF berechnet, um festzustellen welche Informationsbedarfe von wie vielen Unternehmen nachgefragt werden. Es erscheint zweckmäßig, nicht jeden Informationsbedarf in das Fachkonzept aufzunehmen. Da einige Informationsbedarfe kaum nachgefragt werden und eine Informationsüberflutung zu vermeiden ist. Deshalb wird die Häufigkeitsanalyse durch eine ABC-Analyse ergänzt. Durch die ABC-Analyse werden den Informationsbedarfen Prioritäten zugeteilt. Die höchste Priorität hat die Klasse A, in ihr werden alle Informationsbedarfe mit einer relativen Häufigkeit von mehr als 70% erfasst. Die Informationsbedarfe, die eine relative Häufigkeit im geschlossenen Intervall zwischen 70% und 50% aufweisen, ordnen sich mit mittlerer Priorität in die Klasse B ein. Die niedrigste Priorität hat die Klasse C, in der die Informationsbedarfe mit weniger als 50% zusammengefasst werden.

Diese Einteilung stammt aus der Überlegung, dass es einen grundlegenden Informationsbedarf beim PQ-Verfahren geben muss, den es zu modellieren gilt. Da für die statistische Untersuchung die Grundgesamtheit der PQF relativ gering ist, werden die Klassen A und B verwendet, um dennoch ein brauchbares Ergebnis zu

erhalten. Dadurch wird ein Spektrum von 17 Informationsbedarfen der Klasse A und 26 Informationsbedarfen der Klasse B abgedeckt, um einen abstrahierten allgemein gültigen Informationsbedarf zu beschreiben. Im Folgenden werden die Ergebnisse dieser Analyse skizziert, um im Anschluss den dargestellten Informationsbedarf als Basis für das Informationsangebot des zu entwickelnden Informationssystems zu verwenden.

3.2.3 Ergebnisse der Analyse

In diesem Abschnitt werden die Ergebnisse der Analyse der PQF beschrieben. Diese Erkenntnisse dienen dazu einen allgemeinen Informationsbedarf, der durch die PQF beschrieben ist, sowie die allgemeine logische Struktur des PQF zu entwickeln. Zum Aufbau eines PQF gehören das Anschreiben, der Hauptteil und die Anlagen (siehe Abb. 3.2).

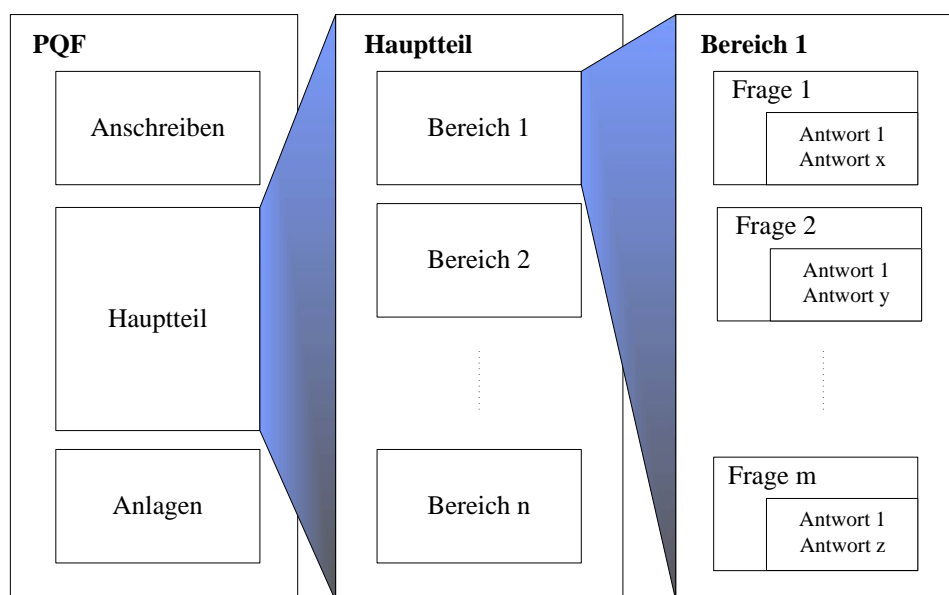


Abb. 3.2: Struktur der Präqualifikationsfragebögen

Das Anschreiben und die Anlagen enthalten für die Analyse des Informationsbedarfs keine relevanten Informationen und werden deshalb in dieser Arbeit nicht näher betrachtet. Der Hauptteil gliedert sich in verschiedene Fragebogenbereiche, die jeweils inhaltlich zusammengehörende Fragen beinhalten. Jede Frage kann eine oder mehrere Antworten enthalten. Die Fragen und zugehörigen Antworten charakterisieren den zu analysierenden Informationsbedarf.

Um eine bessere Darstellung der Ergebnisse zu erzielen, werden zuerst die logischen Bereiche, in welche die PQF gegliedert sind, vorgestellt und die Relevanz für die allgemeine Struktur unter Verwendung der Häufigkeitsanalyse bestimmt. Anschließend

werden die Informationsbedarfe in den Bereichen genauer betrachtet, die eine relative Häufigkeit von mehr als 50% besitzen. In einigen Fällen, auf die an geeigneter Stelle eingegangen wird, gibt es Informationsbedarfe, die eine relative Häufigkeit von mehr als 50% haben, sich aber in einem PQF-Bereich mit weniger als 50% befinden. Dies resultiert aus der nicht-standardisierten individuellen Gestaltung und Aufteilung der PQF durch die Unternehmen. Da diese Informationsobjekte notwendig sind, werden sie zweckorientiert in die logische Struktur des allgemeinen PQFs dieser Arbeit eingeordnet.

Das vollständige Ergebnis der Informationsbedarfserfassung mit den absoluten und relativen Häufigkeiten der Informationsbedarfe befindet sich im Anhang A (Tabellen A.1 bis A.24). Jede Tabelle stellt dabei einen Bereich und die darin enthaltenen Informationsbedarfe der PQF dar. Die Darstellung der Ergebnisse in diesem Abschnitt erfolgt unter Verwendung von Diagrammen, die eine Auswahl der Informationsbedarfe des jeweiligen Bereichs enthalten und mit den entsprechenden relativen Häufigkeiten dargestellt werden. Unter dem Diagrammpunkt *Sonstiges* sind die Informationsbedarfe der Bereiche zusammengefasst, die mit der dargestellten oder einer geringeren relativen Häufigkeit in den PQF vorkommen und somit der Klasse C zugeordnet werden. Da diese nicht relevant für den allgemeinen Informationsbedarf der PQF sind, werden sie auch nicht näher benannt. Im Diagramm der Abb. 3.3 sind die PQF-Bereiche mit ihren relativen Häufigkeiten in den PQF dargestellt.

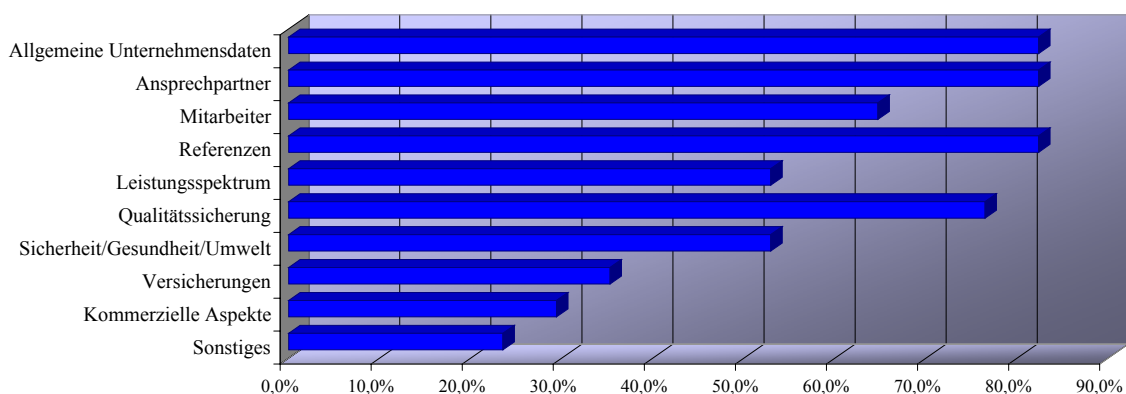


Abb. 3.3: Bereiche der Präqualifikationsfragebögen

Durch den gewählten Schwellenwert von 50% werden nur die PQF-Bereiche der Klasse A und B näher betrachtet. Diese sieben Bereiche allgemeine Unternehmensdaten, Referenzen, Leistungsspektrum, Qualitätssicherung, Sicherheit/Gesundheit/Umwelt, Ansprechpartner und Mitarbeiter stellen die kleinste, allgemeingültige Schnittmenge der PQF-Bereiche dar.

3.2.3.1 Fragebogenbereich: allgemeine Unternehmensdaten

Der Bereich allgemeine Unternehmensdaten erreicht eine relative Häufigkeit von 84,2% und wird daher der Klasse A zugeordnet. In diesem Bereich sind die folgenden Informationsbedarfe der Klassen A und B: Firmenname, Anschrift, Telefonnummer, Faxnummer, E-Mail-Adresse, Internetadresse, Rechtsform, Gründungsort, Besitzer und Organisation zu finden (siehe Abb. 3.4).

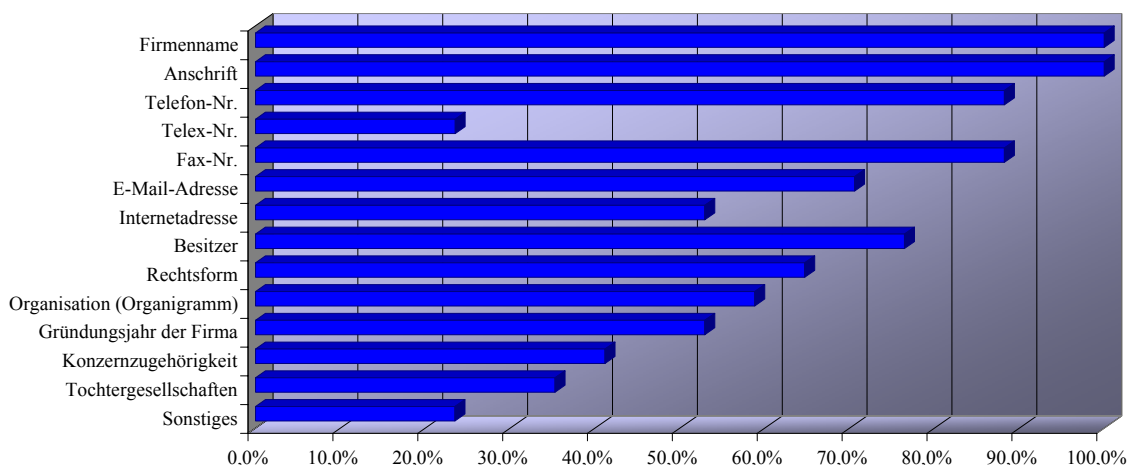


Abb. 3.4: Informationsbedarf des Bereichs allgemeinen Unternehmensdaten

Unter dem Informationsbedarf Organisation werden verschiedene Organigramme, die als Dokumente verlangt werden, zusammengefasst. Diese Organigramme stellen die Organisation des Unternehmens mit Ansprechpartnern und Führungspositionen, die Organisation auf Baustellen und/oder die Organisation im Unternehmensverband dar. Im Organigramm des Unternehmensverbands werden Standorte, Konzernzugehörigkeit, Tochtergesellschaften und Divisionen vermerkt. Damit werden die entsprechenden Informationsbedarfe trotz der zu geringen relativen Häufigkeit (siehe Tab. A.2 im Anhang) indirekt berücksichtigt, obwohl diese nicht explizit in das zu spezifizierende Informationsmodell aufgenommen werden.

Aus dem Bereich der kommerziellen Aspekte, der in die Klasse C eingeordnet ist, wird das Dokument Jahresabschlussbericht mit der relativen Häufigkeit von 52,9% dem Bereich allgemeine Unternehmensdaten zugeordnet, da es sich inhaltlich am besten in diesen Bereich integriert. Der Jahresabschlussbericht wird über die vergangenen drei bis fünf Geschäftsjahre gefordert, um die Entwicklung des Unternehmens beurteilen zu können. Zusätzliche Informationen, die im Zusammenhang mit dem Jahresabschlussbericht erfragt werden, sind jährliche Umsätze, durchschnittliche Umsätze über mehrere Jahre oder Gewinn- und Verlustrechnung. Diese Informationsbedarfe werden auf Grund ihrer Einordnung in die Klasse C nicht

explizit für das zu entwickelnde Informationsmodell berücksichtigt. Eine Ausnahme bildet der Umsatz als beschreibende Information zum Jahresabschlussbericht. Mit dieser Metainformation können die verschiedenen Informationsbedarfe zum Umsatz berechnet werden. Deshalb wird der Umsatz trotz einer relativen Häufigkeit von 41,2% in das Informationsmodell als Attribut zum Jahresabschlussbericht aufgenommen.

Ein weiteres Dokument, welches mit einer relativen Häufigkeit von 52,9% in den Analyseergebnissen vorkommt, ist das Versicherungszertifikat. In 35,3% der PQF wurde das Versicherungszertifikat in einem eigenen PQF-Bereich Versicherungen aufgeführt. Dieser Bereich wird auf Grund der Einordnung in die Klasse C nicht als eigenständiger Bereich für die allgemeine Struktur der PQF aufgenommen. Stattdessen wird das Versicherungszertifikat den allgemeinen Unternehmensdaten zugeordnet. Als typische Versicherungen im Anlagenbau werden hier bspw. die Produkthaftpflichtversicherung und die Arbeitgeberhaftpflichtversicherung genannt.

3.2.3.2 Fragebogenbereich: Referenzen

In fast jedem der untersuchten Präqualifikationsfragebögen wird nach den Referenzen oder der Erfahrung eines Unternehmens gefragt, dabei sind bei 82,4% die Referenzen als eigenständiger PQF-Bereich gekennzeichnet. In diesem Bereich gehören die abgeschlossenen Projekte mit 70,6% zur Klasse A und die Geschäftsbeziehungen mit 58,8% zur Klasse B der Informationsbedarfe (siehe Abb. 3.5).

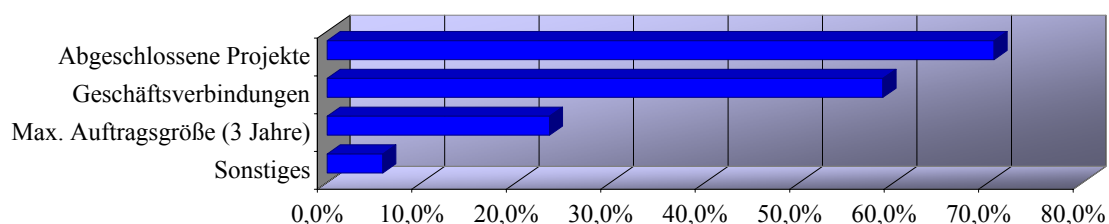


Abb. 3.5: Informationsbedarf des Bereichs Referenzen

Ein Projekt wird für einen bestimmten Standort und mit einer (eindeutigen) Bezeichnung durch einen Kunden beauftragt und meist von einem Generalunternehmen federführend geleitet. Das Generalunternehmen bindet für bestimmte Aufgaben im Projekt Unternehmen durch einen Vertrag. Das gebundene Unternehmen selbst kann Subunternehmen für die beauftragte Art der Leistung verpflichten. Daraus leitet sich der erfragte Prozentsatz der Eigenleistung ab. Weitere gefragte Merkmale eines Projektes sind das eingesetzte Aufsichtspersonal, die Auftragsgröße und die insgesamt geleisteten Arbeitsstunden. Zu den Arbeitsstunden können noch Angaben zur Spitze und zum

Durchschnitt während des Projektes notwendig sein. Die Unfälle, die bei der Arbeit zur Erfüllung der Projektauftrages passieren, werden in einem projektspezifischen Dokument, der Vorfall-Statistik, erfasst. In der Vorfall-Statistik werden neben den normalen und meldepflichtigen Unfällen, mit oder ohne Todesfolge, auch Vorfälle wie z. B. die Nichteinhaltung von Terminen, Nichterbringung von Leistungen und Mängelanzeigen¹¹ festgehalten. Genauere Untersuchungen der Vorfall-Statistik in Bezug auf die Unterschiede einzelner Gewerke des Maschinen- und Anlagenbaus sind zu spezifisch und werden in dieser Arbeit nicht durchgeführt.

Eine weitere beschreibende Information zu einem Projekt ist dessen zeitliche Planung, z. B. indem geplante und tatsächliche Anfangs- und Endtermine erfragt werden. Auf Grund der zu niedrigen relativen Häufigkeit werden laufende Projekte, nicht explizit als Informationsbedarf erfasst. Unter den Geschäftsbeziehungen werden alle Kontakte zu Unternehmen, d. h. Kunden, Subunternehmen und Generalunternehmen einschließlich der vertraglichen Verhältnisse verstanden. Ein Teil dieser Informationsbedarfe lässt sich durch die abgeschlossenen Projekte bestimmen, indem aus abgeschlossenen Projekten Kunden, Generalunternehmen, Subunternehmen oder Vertragsarten ermittelt werden.

3.2.3.3 Fragebogenbereich: Leistungsspektrum

Der Bereich Leistungsspektrum ist mit einer relativen Häufigkeit von 52,9% in den PQF vertreten. Für diesen Bereich werden in den PQF die folgenden Synonyme verwendet: Liefer- und Leistungsprogramm, Produktionsprogramm, Service und Leistungskatalog, Fähigkeiten, Leistung, Leistungsspektrum, Produktpalette, Dienstleistungen und Performance. Auf Grund der begrifflichen Vielfalt, die durch die unterschiedlichen Unternehmen verwendet werden, wird der zusammenfassende Begriff Leistungsspektrum als Bezeichnung für diesen Bereich ausgewählt. Dieser Begriff kann sowohl für Dienstleistungen als auch für Produkte verwendet werden.

In diesem Bereich werden die Informationsbedarfe der Leistungsart berücksichtigt. Zur Leistungsart des Unternehmens gehören die Arbeitsgebiete, die mit eigenem Personal und mit Hilfe von Subunternehmen bearbeitet werden können (siehe Abb. 3.6). Das Arbeitsgebiet, das mit eigenem Personal durchgeführt wird, ist das Kernarbeitsgebiet und wird zu 100% in den untersuchten PQF erfragt. Das Arbeitsgebiet der Subunternehmen wird mit einer relativen Häufigkeit von 52,9% gefordert. Beispiele für die in den PQF genannten Leistungen sind Oberflächen- bzw. Korrosionsschutz, Glühen, Prüfung, Lieferung von Apparaten und Werkstoffen. Eine genauere Betrachtung der einzelnen

¹¹Mängelanzeigen sind Dokumente, in denen qualitative Mängel und fehlende Leistung dokumentiert werden.

Leistungen bzw. Leistungsarten erfolgt nicht in dieser Arbeit, da diese zu den spezifischen Eigenschaften von Gewerken im Maschinen- und Anlagenbau zählen.

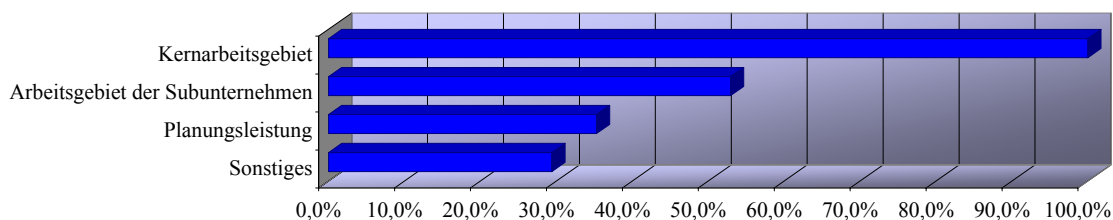


Abb. 3.6: Informationsbedarf des Bereichs Leistungsspektrum

3.2.3.4 Fragebogenbereich: Qualitätssicherung

Der Bereich Qualitätssicherung oder Qualitätsmanagement (QM) ist mit einer relativen Häufigkeit von 76,5% in den PQF vorhanden und enthält die folgenden Informationsbedarfe: QM-System, QM-Zertifizierung, QM-Handbuch, QM-Beauftragter und Vorgehensweisen zum Qualitätsmanagement (siehe Abb. 3.7).

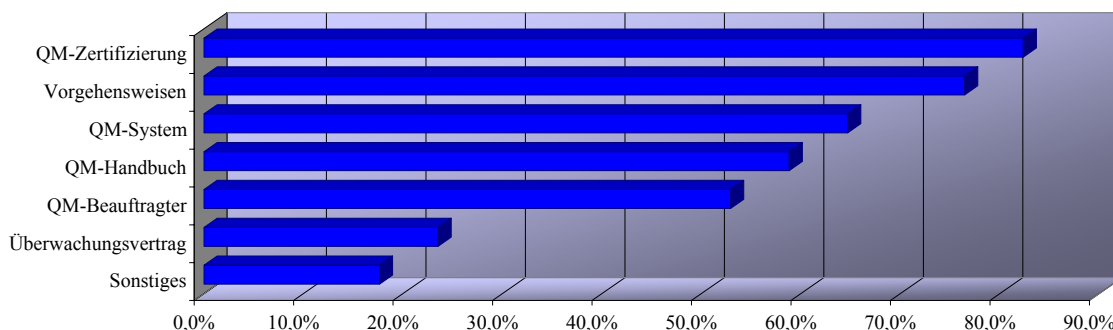


Abb. 3.7: Informationsbedarf des Bereichs Qualitätssicherung

Das QM-System, auch Qualitätssicherungssystem oder Qualitätskontrollsystem genannt, gehört mit einer relativen Häufigkeit von 64,7% in die Klasse B. Hierbei wird nach der Existenz eines QM-Systems gefragt und ob dieses formalisiert ist. Formalisiert bedeutet in diesem Fall, dass das QM-System, ob zertifiziert oder nicht, in irgendeiner Form dokumentiert (schriftlich festgehalten) und dadurch nachvollziehbar ist. Ferner bedeutet es, dass die Prozesse immer auf die gleiche Art und Weise ablaufen und geprüft werden. In 82,4% aller untersuchten PQF wird nach einer QM-Zertifizierung gefragt. Bei dieser QM-Zertifizierung, z. B. nach DIN EN ISO 9001:2000 oder VDA 6.1, handelt es sich um ein Nachweis-Dokument, welches durch das präqualifizierende Unternehmen mit Hilfe eines PQF angefordert wird. Ein weiterer Informationsbedarf der Klasse A,

der mit einer relativen Häufigkeit von 76,5% betrachtet wird, sind Vorgehensweisen im Qualitätsmanagement. Diese umfassen eine Menge von Maßnahmen zur Kontrolle, Überwachung und Steuerung des QM. Einige dieser erfragten Maßnahmen werden im Folgenden beispielhaft genannt:

- Durchführung von Qualitätsprüfungen, wie z. B. Produktprüfungen während und nach dem Herstellungsprozess
- interne/externe Prüfvorrichtungen und Prüfverfahren, wie z. B. Drucktest, Belastungstest, zerstörungsfreie Prüfung mit Röntgenstrahlen oder Zylinderbruchstellentest
- Prüfung und Kalibrierung der Inspektionsausrüstung
- Durchführung von internen/externen Audits, Kundensystemaudits, Verfahrens- und Prozessaudits
- Maßnahmen für Qualitätskontrolle der Subunternehmen
- Auswahlverfahren für Subunternehmen
- QM-Schulungen für Mitarbeiter
- Arbeitsanweisung für die interne/externe Dokumentendurchsicht mit den Subunternehmen
- Arbeitsanweisung für Handhabung und Ablage von Informationen von Subunternehmen
- Arbeitsanweisung für die Vorbereitung der Herstellungsunterlagen
- Arbeitsanweisung für mechanische Fertigung, Prüfung und Übergabe

In 58,8% der untersuchten PQF wird erfragt, ob ein QM-Handbuch vorhanden ist. Als Nachweis kann der Index des QM-Handbuches als Dokument gewünscht werden, dies ist aber nur in 5,9% der PQF der Fall. Der QM-Beauftragte gehört mit 52,9% in die Klasse B und wird mit Kontaktdaten, Vor- und Nachname erfragt. Da der QM-Beauftragte auch als Ansprechpartner im Unternehmen dient, befinden sich detaillierte Angaben zu dessen Attributen im Abschnitt 3.2.3.6.

3.2.3.5 Fragebogenbereich: Sicherheit, Gesundheit und Umwelt

Der Bereich Sicherheit, Gesundheit und Umwelt (HSE¹²) wird mit einer relativen Häufigkeit von 52,9% als eigenständiger Bereich betrachtet. In diesem Bereich werden die Informationsbedarfe HSE-Zertifizierung mit 64,7%, Vorgehensweisen mit 64,7% und HSE-Handbuch mit 52,9% in den PQF abgefragt (siehe Abb. 3.8).

¹²Das Akronym HSE stammt von den englischen Begriffen Health für Gesundheit, Safety für Sicherheit und Environment für Umwelt.

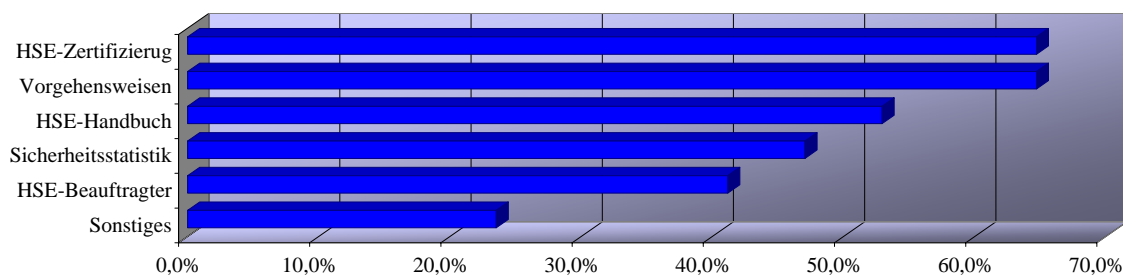


Abb. 3.8: Informationsbedarf des Bereichs HSE

Unter die HSE-Zertifizierung fallen u. a. die Nachweis-Dokumente für bspw. Umwelt, die ISO 14001 und für Arbeitsschutz OHSAS 18001 oder SCC. Im Folgenden werden einige Beispiele für Vorgehensweisen im Bereich HSE aufgezählt:

- Weitergabe von Informationen an Kunden über Gefahrstoffe (in Bezug auf Transport, Behandlung und Lagerung)
- Durchführung von Schwachstellenanalyse und Standard-Risikoanalyse
- Feedback an Mitarbeiter über Schwachstellen
- HSE-Schulungen für Mitarbeiter
- HSE-Schulungen zu Umweltrisiken und Arbeitsschutz
- Ausgabe von Sicherheitsbroschüre mit Anweisungen
- Schulungsintervalle
- Substanzmissbrauchspolitik
- Prüfungen/Überwachung der Geschäftsführer, Baudirektoren, Standort- und Projektmanager
- Einhaltung von HSE-Gesetzen
- Ermittlung der Umweltrisiken
- Berichterstattung an das Generalunternehmen über Unfälle
- Durchführung von Sicherheitsinspektionen der Projekte (Inspektor und Intervalle)
- Verwendung von Anreizsystemen für Sicherheit
- Sicherung des Verhaltens der Subunternehmen

Weitere Informationsbedarfe, die nicht explizit durch ihre Einordnung in die Klasse C in das Informationsmodell eingehen, sind Kennzahlen für Sicherheit oder synonym die Sicherheitsstatistik. Dennoch werden im Folgenden mögliche Kennzahlen für die Sicherheit genannt, da diese in den Präqualifikationsfragebögen an anderer Stelle erfragt und durch die abgeschlossenen Projekte im Bereich Referenzen als begleitendes Dokument Vorfall-Statistik in das Informationsmodell aufgenommen werden. Beispiele für geforderte Kennzahlen, die als Metadaten für die Vorfall-Statistik genutzt werden können, sind:

- 1000-Mann-Quote¹³,
- Anzahl von Verletzungen und Krankheiten,
- Anzahl Vorfälle,
- Anzahl ärztlicher Behandlungen,
- Anzahl dauerhafter Teilinvalidität,
- Todesrate,
- Anzahl der geleisteten Arbeitsstunden der letzten drei Jahre sowie
- totale und eingeschränkte Ausfalltage.

3.2.3.6 Fragebogenbereich: Ansprechpartner

Angaben zum Bereich Ansprechpartner oder Kontaktpersonen sind mit 82,4% in den PQF vertreten. Das Diagramm in Abb. 3.9 zeigt die einzelnen Häufigkeiten der fachspezifischen Ansprechpartner für die abstrahierten Unternehmensbereiche. Die benötigten Ansprechpartner lassen sich grob in technische, kaufmännische und juristische Bereiche, Qualitätsmanagement und HSE einteilen. Der technische Bereich, der kaufmännische Bereich und der Ansprechpartner für das Qualitätsmanagement weisen eine relative Häufigkeit von mehr als 50% auf.

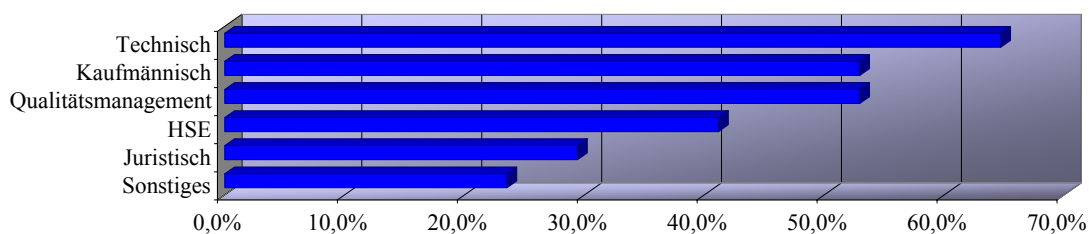


Abb. 3.9: Informationsbedarf des Bereichs Ansprechpartner

Im Detail werden die Ansprechpartner für die Unternehmensbereiche Technik, Produktion, Fertigung, Verkauf, Einkauf, Planung, Konstruktion und Logistik im Unternehmen in den PQF aufgeführt. Diese wurden zum technischen und kaufmännischen Bereich zusammengefasst. Die Ansprechpartner für die Bereiche Qualitätsmanagement und HSE werden in den PQF ohne entsprechenden Qualitäts- bzw. HSE-Bereich unter dem Bereich Ansprechpartner aufgeführt. Wenn diese eigenständigen Bereiche vorhanden sind, werden die Ansprechpartner in diesen PQF-Bereichen (siehe Abschnitt 3.2.3.4 und Abschnitt 3.2.3.5) aufgeführt. Allerdings wird der Informationsbedarf des Ansprechpartners HSE auf Grund der

¹³Die 1000-Mann-Quote ist die Unfallrate meldepflichtiger Unfälle pro 1000 Mitarbeiter. Als meldepflichtig werden Unfälle betrachtet, wenn sie zu einer Arbeitsunfähigkeit von mindestens drei Tagen oder zum Tode führen.

zu geringen Häufigkeit nicht direkt in das Informationsmodell aufgenommen. Weitere beschreibende Informationen der Ansprechpartner sind Vor- und Nachname, Position, Firmenzugehörigkeit und die dazu gehörenden Kontaktdaten. Zu den Kontaktdaten eines Ansprechpartners gehören Telefonnummer, Faxnummer und die E-Mail-Adresse.

3.2.3.7 Fragebogenbereich: Mitarbeiter

Der Bereich Mitarbeiter oder Personal ist Bestandteil von 64,7% der untersuchten PQF und wird durch die vier Informationsbedarfe Mitarbeiteranzahl gesamt, Mitarbeiteranzahl gewerblich, Kapazität und Qualifikationen bestimmt (siehe Abb. 3.10).

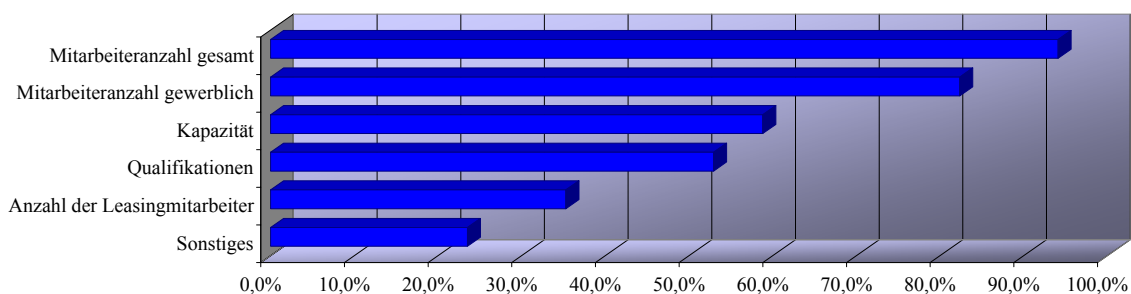


Abb. 3.10: Informationsbedarf des Bereichs Mitarbeiter

Die Gesamtanzahl der Mitarbeiter setzt sich aus den Leasingmitarbeitern und den eigenen gewerblichen Mitarbeitern zusammen. Die gewerblichen Mitarbeiter werden auch als Stammpersonal bezeichnet. Weitere Angaben zur Gesamtanzahl sind Aufschlüsselungen der Mitarbeiter nach Berufsbezeichnungen, wie z. B. Schweißer, Monteure, Schweißaufsicht und Prüfaufsicht bzw. nach Bereichen, wie z. B. kaufmännischer oder technischer Bereich. Diese Informationsbedarfe sind, ebenso wie die Fluktuationsstatistik und das Durchschnittsalter, von dynamischer Natur und können aus gepflegten Stammdaten der Personalabteilung ermittelt bzw. berechnet werden. Die Kapazität der Mitarbeiter bezieht sich auf die verfügbaren und die an andere Projekte bereits gebundenen Arbeitsstunden des Unternehmens. Die Dokumente, die im Bereich der Mitarbeiter benötigt werden, sind deren Qualifikationsnachweise.

3.2.3.8 Nachweis-Dokumente

In den PQF werden verschiedene Dokumente als Nachweise verlangt (siehe Tab. A.10 im Anhang). Die Dokumente, die als Anlage zu den PQF gefordert werden, sind Zertifikate mit 88,2%, fachspezifische Zulassungen bzw. Fachbetriebsnachweise mit

64,7%, Organigramme und Reports mit jeweils 58,8% und Versicherungszertifikate mit 52,9% (siehe Abb. 3.11). Die Prospekte und Produktkataloge fallen mit 29,4% ebenso wie die unter Sonstiges zusammengefassten Dokumente in die Klasse C und werden deshalb nicht näher beschrieben.

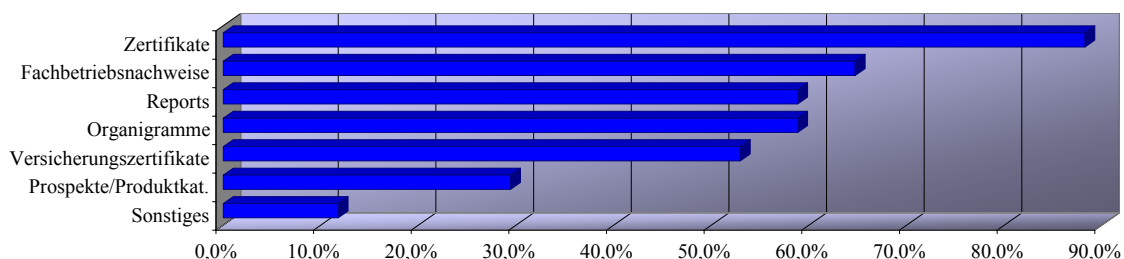


Abb. 3.11: Informationsbedarf der Dokumente

In die Gruppe der Zertifikate gehören behördlich bescheinigte Kompetenzen und Unternehmensakkreditierungen, wie z. B. QM-Zertifikat oder ASME Certificates of Authorization. Zu den fachspezifischen Zulassungen bzw. Fachbetriebsnachweisen gehören Berufsqualifikationen und Zertifikate der Mitarbeiter, z. B. TRD 100¹⁴, oder DIN 8563¹⁵. Die Zertifikate und Qualifikationen werden jeweils mit den Informationen Art der Zertifizierungen, Zertifizierungsstelle, Zertifikatnummer und Gültigkeit verlangt. Die zusätzlich gefragten Informationen zum Versicherungszertifikat sind Versicherungsgesellschaft, Versicherungsart, Versicherungsnummer sowie Deckungshöhe und Gültigkeit. Des Weiteren werden als Anlage zu den PQF folgende Dokumente benötigt: der Jahresabschlussbericht, der Dunn & Bradstreet Report¹⁶, die HSE-Jahresabschlussberichte und die Vorfall-Statistiken. Die genannten Dokumente werden als Reports (Berichte) zusammengefasst.

3.3 Zusammenfassung und Schlussfolgerungen

Zusammenfassend lässt sich die Struktur eines PQF in Anschreiben, Hauptteil und Anlagen gliedern (siehe Abb. 3.2). Das Anschreiben enthält u. a. beschreibende Informationen zum Projekt, Kunden und zeitliche Angaben, welche als Metadaten extrahiert werden können. Der Hauptteil unterteilt sich in unterschiedliche Bereiche (vgl. Abschnitt 3.2.3). Jeder Bereich kann mehrere Fragen beinhalten, die verschiedene Antwortmöglichkeiten vorgeben oder Platzhalter für die Antworten enthalten.

¹⁴Die technischen Regeln für Dampfkessel 100 beschäftigen sich mit den allgemeinen Grundsätzen für Werkstoffe zur Dampfkesselherstellung.

¹⁵Die Sicherung der Güte von Schweißarbeiten wird durch die DIN 8563 gewährleistet.

¹⁶Beim D&B-Report handelt es sich um Unternehmensberichte und Ratings, die von der Firma Dunn & Bradstreet Inc. erstellt werden.

Die Informationsbedarfsanalyse zeigte, dass die folgenden Bereiche: allgemeine Unternehmensdaten, Ansprechpartner, Mitarbeiter, Referenzen, Leistungsspektrum, Qualitätssicherung und Sicherheit/Gesundheit/Umwelt der Klasse A und B als relevant eingestuft werden. Somit bilden diese mit ihren Informationsbedarfen die Datenbasis für den Beantwortungsprozess.

Anhand der Ist-Analyse des Beantwortungsprozesses wurden folgende Probleme festgestellt:

1. Die PQF weisen unterschiedliche Formate auf, die teilweise nicht digital verarbeitet werden können.
2. Die Archivierung der PQF und deren Wiederverwendung ist unzureichend.
3. Die Weitergabe von Fragen an andere Mitarbeiter ist für die Beantwortung notwendig.
4. Der Beantwortungsprozess benötigt Informationen, die in unterschiedlichen heterogenen Anwendungen im Unternehmen verwaltet werden.
5. Einige der benötigten Informationen werden nicht direkt in den Anwendungen verwaltet, wie z. B. ob ein QM-System formalisiert ist.

Das erste, zweite und dritte Problem kann durch die Umgestaltung und Modellierung des Soll-Prozesses der PQF-Beantwortung gelöst werden. Das vierte Problem, die fehlenden Informationen und heterogene Systemlandschaft kann durch die Konzipierung von Teilsystemen zur Informationspflege und durch die Integration bestehender Anwendungen beseitigt werden. Da genaue Kenntnisse über die zukünftige Systemlandschaft fehlen, wird der Top-Down Ansatz für die Konzeption der integrierten Datenbasis gewählt und im Folgenden drei mögliche Szenarien beschrieben. Aus diesen drei Szenarien wird ein allgemein gültiges Szenario festgelegt. Anhand dieses Szenarios werden die benötigten Prozesse und Schemata konzipiert, die als Datengrundlage für den Hauptprozess des Informationssystems dienen.

Im ersten Szenario werden alle benötigten Informationsbedarfe durch unterschiedliche Anwendungen im Unternehmen verwaltet und gepflegt. Das zu konzipierende Informationssystem benötigt den Zugriff auf die Datenbasis der Fremdanwendung. In diesem Fall müssen Schnittstellen konzipiert werden. Durch einen Pull-Mechanismus kann das Informationssystem zur Laufzeit die benötigten Informationen aus der Altanwendung anfordern. Da dies zu Performanceproblemen der Altanwendung führen kann, ist auch ein Push-Mechanismus für die Speicherung der benötigten Informationen in die Datenbank des Informationssystems durch die Fremdanwendung vorstellbar. Dabei entsteht einerseits das Problem der redundanten Datenhaltung. Andererseits hat diese Möglichkeit den Vorteil, durch eine geringere Anzahl von Transaktionen

in der Fremdanwendung dessen Performance nicht zu beeinträchtigen. Um keine Konsistenzprobleme durch redundante Daten zu erhalten, erfolgt der Zugriff virtuell über ein globales Datenschema. Die Datenpflege, also schreibende Zugriffe, verbleibt bei den Altanwendungen.

Das zweite Szenario ist dadurch gekennzeichnet, dass im Unternehmen nur teilweise die benötigten Anwendungen mit den entsprechenden Funktionalitäten vorhanden sind. Daraus folgt, dass nicht alle benötigten Informationsbedarfe zur Verfügung stehen. Deshalb müssen die fehlenden Anwendungen konzipiert oder bestehende Anwendungen um die benötigte Funktionalität erweitert werden. Darauf aufbauend kann die Integration der vorhandenen Informationen über ein globales Datenschema, wie im ersten Anwendungsfall, erfolgen.

Im dritten Szenario werden im Unternehmen keine der benötigten Informationsbedarfe verwaltet. In diesem Fall kann eine Anwendung konzipiert werden, in der alle benötigten Informationsbedarfe verwaltet und als Datenbasis zur Verfügung stehen.

Aus diesen drei Anwendungsfällen wird im Folgenden ein allgemeiner Anwendungsfall festgelegt, der die Probleme der heterogenen Systemlandschaft und die nicht existierenden Anwendungssysteme zur Informationspflege berücksichtigt. Es wird davon ausgegangen, dass in einem Unternehmen der zweite Anwendungsfall am häufigsten vorliegt. Allerdings ist es nicht möglich, die fehlenden Anwendungen in der Systemlandschaft zu charakterisieren. Aus diesem Grund wird die Annahme aus dem dritten Szenario, dass keine Anwendung vorhanden ist der Konzeption zu Grunde gelegt. Demnach werden alle Anwendungen und Prozesse beschrieben, die für das Informationsangebot benötigt werden. Wahlweise können dann bei der Einführung die benötigten Komponenten zusammengestellt werden. Die Integration erfolgt über ein globales Datenschema. Diese Vorgehensweise erlaubt, das zu entwickelnde Informationssystem in jede beliebige Systemlandschaft zu integrieren und schließt die drei oben erwähnten Szenarien ein. Das bedeutet, wenn in einem Unternehmen alle Anwendungen vorhanden sind, werden diese über das globale Schema integriert. Fehlende Anwendungen werden der Systemlandschaft hinzugefügt und schließlich über das globale Schema verbunden. Bei der Konzeption dieser integrierten Datenbasis und ihrer Anwendungen wird die Top-Down Vorgehensweise angewendet. Dabei wird aus den Informationsbedarfen der Analyse ein globales Datenschema entworfen. Auf Basis dieses Schemas werden Teilschemata entwickelt die ebenfalls als Exportschemata für bestehende Anwendungen verwendet werden können. Zu den Teilschemata werden die benötigten Datenverwaltungsfunktionen beschrieben.

An das Fachkonzept zur Unterstützung des Beantwortungsprozesses von Präqualifikationsfragebögen werden zusammenfassend die folgenden drei Anforderungen gestellt. Das Fachkonzept soll:

- die Digitalisierung der Präqualifikationsfragebögen und Verarbeitung unterschiedlicher Textformate ermöglichen, um die automatischen Analysen durchführen zu können.
- die semi-automatische Beantwortung der Präqualifikationsfragebögen unterstützen, um die Dauer dieses Prozesses zu verkürzen und das Führungspersonal zu entlasten.
- die Integration der verschiedenen Teilsysteme gewährleisten, um eine einheitliche Datengrundlage für die Beantwortung zu schaffen und die Flexibilität des Konzeptes zu erhöhen.

Diese Anforderungen werden im folgenden Kapitel aufgegriffen und umgesetzt.

4 Fachkonzept des Informationssystems

Nachdem im Kapitel 3 der Ist-Prozess und das benötigte Informationsangebot ermittelt wurden, werden im folgenden Kapitel diese Erkenntnisse genutzt, um auf fachkonzeptueller Ebene ein Informationssystem zur Unterstützung des Beantwortungsprozesses von Präqualifikationsfragebögen zu entwickeln. Dieses Kapitel gliedert sich in drei Abschnitte. Im Abschnitt 4.1 wird auf die verwendete Methodik und Modelle des Fachkonzepts eingegangen. Darauf aufbauend werden im zweiten und dritten Abschnitt das Informationssystem und seine Teilsysteme beschrieben. Der Abschnitt 4.2 beschäftigt sich mit der Soll-Konzeption des Beantwortungsprozesses, während Abschnitt 4.3 die Datenmodellierung des Informationsangebots und das benötigte Konzept zur Datenpflege beschreibt. Da es für die Gestaltung eines föderierten Datenschemas in Bezug auf die PQF kein Referenzschema gibt, wird aus den aufgenommenen Informationsbedarfen ein eigenständiges Schema entwickelt und die benötigten Funktionalitäten beschrieben.

4.1 ARIS-Konzept

Bei der Gestaltung von Informationssystemen werden je nach Vorgehensmodell und in Abhängigkeit der Nähe zur informationstechnischen Lösung unterschiedliche¹⁷ Modelle erstellt, die im nachfolgenden Schritt in andere der Problemlösung dienende Modelle überführt werden (vgl. Knackstedt (2006), S. 40). Vorgehens- und Phasenmodelle verknüpfen die Aufgaben, die zur Entwicklung notwendig sind. Phasenmodelle sind gegenüber Vorgehensmodellen weniger detailliert (vgl. Knackstedt (2006), S. 96).

Im Fachkonzept werden die betriebswirtschaftlich-inhaltlichen Anforderungen an das zu entwickelnde System beschrieben (vgl. Knackstedt (2006), S. 116). Damit stellt das Fachkonzept die Schnittstelle zwischen Fachanwendern und IT-Experten der DV-Abteilungen dar. Für Managementinformationssysteme ist „festzulegen, welche Führungskräfte für welche Managementaufgaben welche Daten benötigen“ (Knackstedt (2006), S. 97). Dafür sind entsprechende Informationsbedarfsanalysen durchzuführen, wie es in Kapitel 3 dieser Arbeit geschehen ist. Die im Fachkonzept erarbeiteten betriebswirtschaftlichen Anforderungen werden im DV-Konzept zu informationstechnischen Lösungskonzepten verarbeitet (vgl. Knackstedt (2006), S. 98). Dabei wird eine Informationssystemarchitektur aus identifizierten Hard- und Softwaresystemen zusammengefügt. Für diese Zusammenstellung kann eine Orientierung an verschiedenen Referenzmodellen erfolgen. Dies ist aber nicht Bestandteil dieser Arbeit. Die Umsetzung des im Fachkonzept beschriebenen Informationsbedarfs in

¹⁷Die Unterschiede beruhen auf den verwendeten Modellierungstechniken bzw. -methoden.

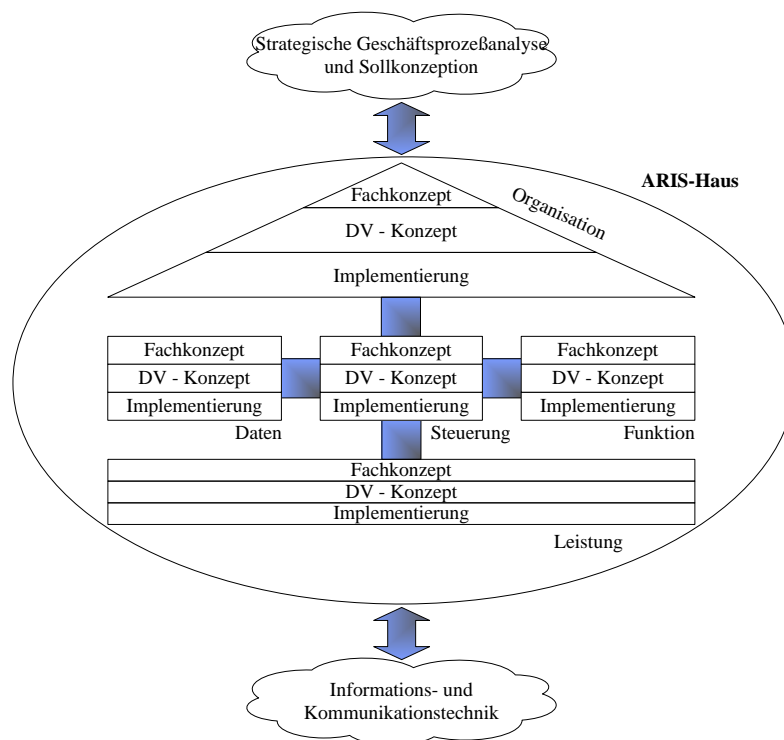
das Datenbankmodell des gewählten DBMS bildet einen weiteren Schwerpunkt des DV-Konzeptes (vgl. Knackstedt (2006), S. 98). In der Phase der Implementierung wird die reale Umsetzung der im DV-Konzept entwickelten Vorgaben durchgeführt. Dabei werden Hardware-Komponenten aufgebaut, Programmcode eigener Anwendungen erstellt, Anwendungssysteme installiert und angepasst (vgl. Knackstedt (2006), S. 99 f.). Die Betriebsphase beschäftigt sich mit der Instandhaltung der Leistungsfähigkeit des Systems. Dies schließt streckenweise bei der Lösung von Problemen die vorangegangenen Phasen ein (vgl. Knackstedt (2006), S. 100).

Zur Unterstützung der Ablauforganisation der Entwicklung von Anwendungssystemen werden die Methoden und Vorgehensmodelle des Software Engineering verwendet. Sequentielles Vorgehen bedeutet, dass die Phasen aufeinander folgen und die vorangegangenen Ergebnisse einer Phase zur Transformation, als Input der aktuellen Phase genutzt werden (vgl. Stickel (2001), S. 112). Ein Rückschritt zu einer vorherigen Phase ist unzulässig (vgl. Knackstedt (2006), S. 100). Auf Grund ihrer Komplexität sind bei den meisten Entwicklungen Korrekturen notwendig, deshalb ist das sequentielle Vorgehen nicht geeignet. Häufig werden erst im Ausführungssystem Schwächen und Änderungswünsche erkannt (vgl. Knackstedt (2006), S. 101). Ein iteratives Vorgehen schließt die Wiederholung von Phasen ein und eignet sich daher besser zur Entwicklung von Anwendungen und Informationssystemen.

Das ARIS-Konzept ist ein sichtenorientiertes Modellierungskonzept für die Entwicklung integrierter Informationssysteme, welches von der IDS Prof. Scheer GmbH entwickelt wurde. Die ARIS-Sichten bilden sich durch den ähnlichen semantischen Zusammenhang ihrer Modellierungsgegenstände (vgl. Scheer (2002), S. 33). Es gibt die fünf folgenden Sichten: Organisations-, Funktions-, Daten-, Leistungs- und Steuerungssicht. Die Organisations-sicht umfasst alle Organisationseinheiten der Aufbauorganisation, sowie menschliche und maschinelle Aufgabenträger. In der Funktionssicht werden Funktionen, ihre Ziele und die Anwendungssoftware zusammengefasst. Die Datensicht enthält Umfelddaten, Nachrichten und Ereignisse. Die Leistungssicht beinhaltet „alle materiellen und immateriellen Input- und Output-Leistungen einschließlich der Geldflüsse“ (Scheer (2002), S. 36). Durch die Steuerungssicht werden die Sichten miteinander verbunden und die Beziehungen zwischen ihnen, also das gesamte Prozessmodell dargestellt (vgl. Scheer (2002), S. 36). Diese fünf Sichten bilden das ARIS-Haus (siehe Abb. 4.1).

Das ARIS-Phasenmodell orientiert sich am Softwarelebenszyklus. Aus der betriebswirtschaftlichen Problemstellung wird das Fachkonzept unabhängig von der technischen Umsetzung entwickelt. Das Fachkonzept bildet die Grundlage für das

DV-Konzept gefolgt von der technischen Implementierung. Das ARIS-Phasenmodell wird mit dem ARIS-Konzept kombiniert (siehe Abb. 4.1).



Quelle: Scheer (2002), S. 41

Abb. 4.1: ARIS-Haus mit Phasenkonzept

Das bedeutet, jede Sicht wird stufenweise vom Fachkonzept bis zur Implementierung verfeinert. Der Vorteil der Sichten besteht darin, dass für jede Sicht und jede Phase unterschiedliche Modelle verwendet werden können. In dieser Arbeit werden auf fachkonzeptioneller Ebene in der Funktionssicht Funktionsbäume, in der Datensicht erweiterte Entity Relationship Modelle (eERM) und in der Steuerungssicht erweiterte ereignisgesteuerte Prozessketten (eEPK) verwendet. Bei den Objekten der Leistungssicht handelt es sich um Informationsobjekte, die in der Datensicht modelliert werden, um Redundanzen zu vermeiden. Auf Grund der unterschiedlichen Unternehmen wird auf ein Organigramm verzichtet da nur ein Bearbeiter für die Prozesse modelliert wird. Die Organisationssicht sollte individuell vor der Einführung mit dem Rollenkonzept im konkreten Unternehmen modelliert werden.

Für die Angabe der Kardinalitäten im eERM können verschiedene Notationen verwendet werden. Die ursprüngliche Notation des Entity-Relationship-Ansatzes ist die 1:1-, 1:n- sowie n:m-Beziehung (vgl. Chen (1991), S. 35 ff.). Diese Darstellung der Maximalwerte schließt kein Vorkommen einer Entität in der Beziehung mit ein (vgl. Heuer und Saake

(2000), S. 70) und ermöglicht dadurch nicht die Modellierung von Beziehungen an denen Entitäten beteiligt sein müssen. Nach Heuer und Saake (2000) wird dies durch die (min., max.)-Notation realisiert. Auf Grund des Abstraktionsniveaus des Fachkonzepts wird im Folgenden die Chen-Notation verwendet.

Funktionsbäume stellen eine Funktionshierarchie dar und werden in prozess- oder objektorientiert unterteilt (vgl. Scheer (1998), S. 25 f.). Dabei bedeutet prozessorientiert, dass die beteiligten Funktionen zum selben Prozess gehören aber an unterschiedlichen Informationsobjekten ausgeführt werden. Objektorientierte Funktionsbäume weisen auf Funktionen hin die am selben Informationsobjekt verrichtet werden.

4.2 Soll-Prozess der Beantwortung eines Präqualifikationsfragebogens

Die technische Unterstützung der PQF-Beantwortung auf Grundlage der Analyse in Kapitel 3 steht im Mittelpunkt des zu entwickelnden Informationssystems. Um den PQ-Prozess zu unterstützen werden verschiedene Prozesse sowie funktionale und fachliche Anforderungen an den Prozess der PQF-Beantwortung dargestellt und beschrieben. Die folgenden Abschnitte beschreiben den Soll-Prozess der PQF-Beantwortung.

Die Ist-Analyse des Prozesses zeigte u. a. folgende Probleme:

- Die PQF weisen unterschiedlichen Formate auf, die teilweise nicht digital verarbeitet werden können.
- Die Archivierung der PQF und deren Wiederverwendung ist unzureichend.
- Die Weitergabe von Fragen an andere Mitarbeiter ist zu deren Beantwortung notwendig.
- Der Beantwortungsprozess benötigt Informationen, die in unterschiedlichen heterogenen Anwendungen im Unternehmen verwaltet werden.

Diese Probleme können im Soll-Prozess der Beantwortung gelöst werden. Die PQF lassen sich grob in die zwei Klassen CI-Dokumente und NCI-Dokumente einordnen (vgl. Klingelhöller (2001), S. 61). Das Akronym CI steht für das englische *Coded Information* und bedeutet, dass ein Dokument für die Anwendungen elektronischer Datenverarbeitung (EDV) in lesbarer also kodierter Form vorliegt. Im Gegensatz dazu steht das Akronym NCI für das englische *None-coded Information* und bezieht sich auf Dokumente, die für EDV-Anwendungen in nicht lesbarer also nicht kodierter Form vorliegen. Dabei bezieht sich *lesbar* auf die automatische Verarbeitung. Um die automatische oder zumindest durchgängig elektronische Verarbeitung zu gewährleisten, müssen die PQF digitalisiert und in ein bearbeitbares Format umgewandelt werden. Nachdem die

PQF in einem einheitlichen Format vorliegen, kann die semi-automatische inhaltliche Erschließung beginnen. Auf der Erschließung aufbauend kann die Beantwortung der PQF erfolgen. Nachdem alle Fragen beantwortet sind, wird der PQF geprüft, archiviert und versendet. Zusammenfassend gliedert sich der Hauptprozess in die Teilprozesse Digitalisierung, Vorverarbeitung und (semi-automatische) Beantwortung. Diese werden als Wertschöpfungsdiagramm in Abb. 4.2 dargestellt und im Folgenden näher beschrieben.

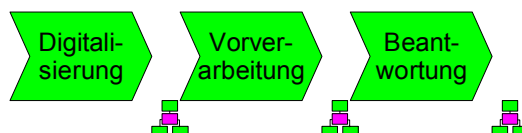


Abb. 4.2: Übersicht des Beantwortungsprozesses der Präqualifikationsfragebögen

4.2.1 Digitalisierung

Die Ist-Analyse zeigte, dass die Übertragung der PQF von den anfragenden zu den befragten Unternehmen per E-Mail, Fax und/oder Brief erfolgt. Dabei stellen papiergebundene Briefe und Faxe NCI-Dokumenten dar. Um die semi-automatische Verarbeitung zu ermöglichen, müssen diese Dokumente digitalisiert werden. Dazu werden diese Dokumente eingescannt und der Text mit Hilfe einer Optical Character Recognition (OCR) Software in kodierter Form abgelegt. Da in diesem Fachkonzept das Einscannen von Dokumenten in verschiedenen Prozessen notwendig ist, wird an dieser Stelle ein abstrakter Prozess zum Digitalisieren von Dokumenten beschrieben. Dieser Teilprozess ist an den Archivierungsprozess der Dokumentenmanagementsysteme (DMS) angelehnt (vgl. Limper (2001), S. 70 ff.).

Der Digitalisierungsprozess gliedert sich in die in Abb. 4.3 dargestellten Teilprozesse Dokumentenscan, OCR-Erkennung und Attribuierung.



Abb. 4.3: Teilprozesse der Digitalisierung

4.2.1.1 Dokumentscan

Der Teilprozess Dokumentscan kann durch zwei Ereignisse beginnen und wird in der Abb. 4.4 in Form einer eEPK dargestellt. Das erste Ereignis betrifft den Fall, dass ein *Dokument nicht digital vorhanden* ist und digitalisiert werden muss. Das zweite Ereignis kann nach einer Sichtprüfung des gescannten Images eintreten und wird zur Darstellung der Rekursion des Dokumentscanprozesses modelliert, zusätzlich muss das Originaldokument verfügbar sein. Durch diese Ereignisse wird die Funktion *Dokument scannen* ausgelöst.

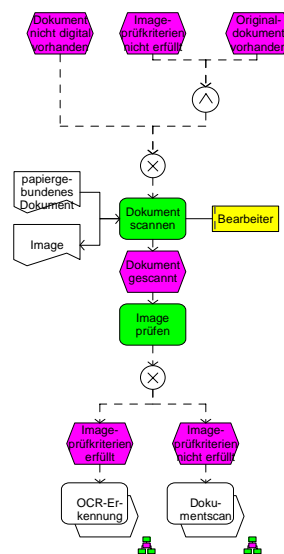


Abb. 4.4: Teilprozess Dokumentscan

In dieser Funktion wird ein papiergebundenes Dokument seitenweise unter Verwendung eines Scanners von einem Bearbeiter eingelesen und ein Image des Dokuments erstellt. Die verwendete Auflösung wird als dots per inch (dpi) angegeben. Für die im Anschluss durchgeführte OCR-Erkennung sind bei Schwarzweißdokumenten 400 dpi und bei Farbdokumenten bis zu 1200 dpi notwendig, um die Ergebnisqualität zu verbessern (vgl. Götzer et al. (2004), S. 202).

Der Bearbeiter wird an dieser Stelle als Personentyp modelliert, da dieser Prozess von unterschiedlichen Mitarbeitern im Unternehmen ausgeführt werden kann. Nach der Erstellung des Images wird dieses durch den Bearbeiter einer Sichtprüfung unterzogen, dabei werden die folgenden Imageprüfkriterien angewendet (vgl. Limper (2001), S.70):

- Vollständigkeit,
- Lesbarkeit,
- Ausrichtung,
- Kontrast sowie

- Schärfe.

Wenn das Dokument diese *Imageprüfkriterien nicht erfüllt*, wird das Image verworfen und das Dokument erneut gescannt. Die Speicherung des Image erfolgt üblicherweise im Tagged Image File (TIF) Format (vgl. Götzer et al. (2004), S. 218). Nachdem das Image alle *Imageprüfkriterien erfüllt*, wird der zweite Teilprozess die OCR-Erkennung durchgeführt.

4.2.1.2 OCR-Erkennung

Die OCR-Erkennung dient der Erstellung eines CI-Dokuments. Als Input für diesen Prozess dienen NCI-Dokumente, die in einem Image-Format vorliegen und nicht direkt maschinell verarbeitet werden können (vgl. Götzer et al. (2004), S. 72). Im Folgenden wird das NCI-Dokument auch als Image bezeichnet. Das Ergebnis dieses Prozesses „ist von der Qualität der Vorlage und der Leistungsfähigkeit der Software“ (Götzer et al. (2004), S. 72) abhängig. Als OCR oder Intelligent Character Recognition (ICR) Verfahren werden Algorithmen bezeichnet, die Textdaten aus Image-Dokumenten extrahieren und in maschinenlesbare Form umwandeln (vgl. Götzer et al. (2004), S. 177 f.). Dabei stellt ICR erweiterte Funktionen zur Verfügung. Beide Verfahren werden auf Grund ihres fließenden Übergangs als OCR-Verfahren zusammengeführt. Der Teilprozess OCR-Erkennung ist in Abb. 4.5 als eEPK dargestellt und wird im Folgenden beschrieben.

Die erste Funktion dieses Teilprozesses wird durch das Ereignis *Imageprüfkriterien erfüllt* ausgelöst. Dieses Ereignis kann in den zwei Teilprozessen Dokumentscan und OCR-Erkennung selbst eintreten. Die durch diese Ereignisse ausgelöste Funktion *OCR-Erkennung durchführen*, wird von einem Bearbeiter überwacht. Dieser ist für alle Funktionen dieses Prozesses verantwortlich. Nachdem das Image die OCR-Software durchlaufen hat, wird das OCR-Ergebnis vom Bearbeiter geprüft. Diese manuelle Qualitätskontrolle ist notwendig, da die „Fehleranfälligkeit der OCR-Verfahren bei mind. 1% liegt“ (Götzer et al. (2004), S. 179). Die Prüfung erfolgt auf dem Bildschirm am Arbeitsplatz des Bearbeiters. Der Bearbeiter entscheidet darüber, ob die Prüfkriterien erfüllt sind oder nicht. Die Prüfkriterien beziehen sich auf den korrekt erkannten Text. Dafür stehen dem Bearbeiter die Rechtschreib- und Grammatikprüfung zur Fehlerquotenermittlung zur Verfügung. Wenn die *OCR-Prüfkriterien erfüllt* sind, handelt es sich um ein maschinell bearbeitbares CI-Dokument, welches im nachfolgenden Teilprozess attribuiert und gespeichert wird

Sind die *OCR-Prüfkriterien nicht erfüllt*, wird vom Bearbeiter eine von drei *Vorgehensweise bestimmt*, welche durch die drei Ereignisse *OCR-Ergebnis manuell*,

automatisch und *nicht korrigierbar* ausgelöst werden. Die gewählte Vorgehensweise hängt vom Zustand des OCR-Ergebnisses und dem Verwendungszweck des Dokuments ab. Bei der Digitalisierung eines PQF ist eine OCR-Erkennung zwingend notwendig, um eine semi-automatische Verarbeitung zu ermöglichen.

Die erste Möglichkeit ist die Funktion *OCR-Ergebnis manuell korrigieren*. Diese wird vom Bearbeiter gewählt, wenn nur wenige Fehler im OCR-Ergebnis sind, die manuell mit geringem Aufwand beseitigt werden können. Für diese Funktion wird ein Texteditor benötigt, indem das OCR-Ergebnis bearbeitet werden kann. Der Teilprozess ist nach der Korrektur beendet und das Zwischenergebnis wird an den Teilprozess Attribuierung weitergereicht.

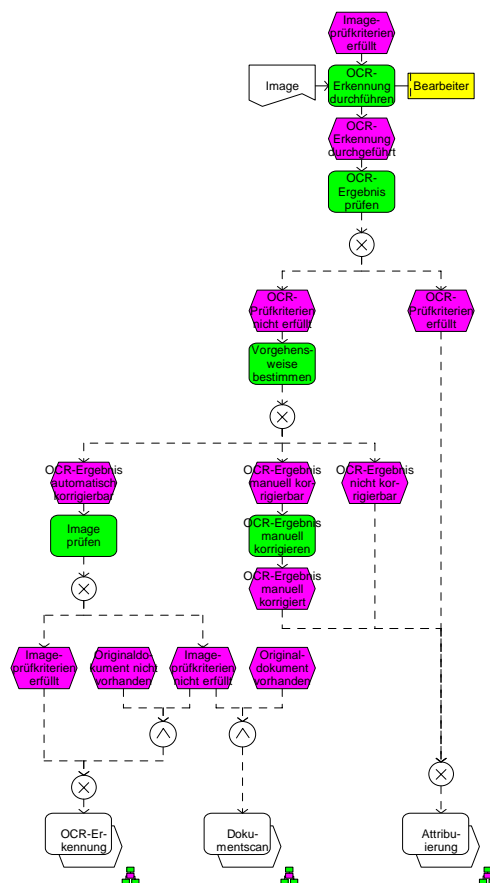


Abb. 4.5: Teilprozess OCR-Erkennung

Die zweite Möglichkeit ist eine erneute automatische OCR-Erkennung. Diese wird vom Bearbeiter ausgewählt, wenn der manuelle Aufwand zu groß wäre und der Bearbeiter eine automatische Verbesserung der Ergebnisse für möglich hält. In dem Fall wird das zur Bearbeitung verwendete Image durch die Funktion *Image prüfen* anhand der Image-Prüfkriterien aus dem Teilprozess Dokumentscan untersucht. Diese Prüfung ist notwendig, da durch die menschliche Sichtprüfung nach dem Scannen Fehler übersehen

werden können, die dann wiederum zu einem schlechten OCR-Ergebnis führen. Wenn die *Imageprüfkriterien erfüllt* sind, bedeutet das, dass ein Fehler bei der OCR-Erkennung aufgetreten ist und der Teilprozess wiederholt werden muss. Sind die *Image-Prüfkriterien nicht erfüllt* ist es notwendig, das Dokument noch einmal zu scannen. Hierbei können zwei zusätzliche Ereignisse eintreten. Wenn das *Originaldokument vorhanden* ist, kann das Dokument erneut gescannt werden, d. h. der Teilprozess Dokumentscan wird gestartet. Gegebenenfalls werden die Einstellungen beim Scannen verändert. Das erstellte Image geht wiederum in den OCR-Erkennungsprozess ein. Ist jedoch das *Originaldokument nicht vorhanden*, kann das Image nicht durch erneutes Scannen verbessert werden. Durch die erneute Ausführung des OCR-Erkennungsprozesses und der Anpassung von Einstellungen in der Funktion *OCR-Erkennung durchführen*, ist es dennoch unter Umständen möglich, das OCR-Ergebnis zu verbessern.

Die dritte Möglichkeit ist, dass das OCR-Ergebnis nicht manuell oder automatisch verbessert werden kann, da bspw. schon mehrere Iterationsschritte dieses Teilprozesses durchlaufen wurden. In diesem Fall tritt das Ereignis *OCR-Ergebnis nicht korrigierbar* ein. Dieses Ereignis beendet den Teilprozess der OCR-Erkennung und löst den Teilprozess Attribuierung aus. Dieser Teilprozess erlaubt die bis dato erarbeiteten Zwischenergebnisse zur weiteren Verarbeitung im Informationssystem zu speichern.

Das Ergebnis im Teilprozess OCR-Erkennung kann zwischen einem perfekt und nicht erkannten Text variieren. Im letzten Fall ist die Attribuierung des Images von besonderem Stellenwert, weil über das Image als NCI-Dokument keine Volltextsuche stattfinden kann, um es im Informationssystem wiederzufinden.

4.2.1.3 Attribuierung

Die digitalen Dokumente werden vom Informationssystem erfasst und indiziert. Indizierung, Indexierung, Verschlagwortung und Attribuierung haben in diesem Kontext dieselbe Bedeutung. Über die bei der Indizierung vergebenen Attribute können die Dokumente im Archivsystem wiedergefunden werden. Die Indizierung kann manuell oder automatisch erfolgen. Das automatische Indizieren kann durch Extraktion der benötigten Informationen aus CI-Dokumenten durchgeführt werden. Eine weitere Möglichkeit ist, dass ein im Unternehmen erstelltes Dokument durch die Anwendung selbst mit seinen Attributen im Archivsystem abgelegt wird (vgl. Klingelhöller (2001), S. 11). Die Indizierung kann und sollte auf Grund des benötigten Fachwissens auch in den jeweiligen Fachabteilungen erfolgen. Die Funktionen die von Archivsystemen unterstützt werden sind das Erfassen, die Indizierung, die Archivierung und das Wiederfinden von

Dokumenten (vgl. Klingelhöller (2001) S. 13). Archivsysteme stellen eine wesentliche Komponente in einem DMS dar (vgl. Klingelhöller (2001) S. 16 ff.).

In der ersten Funktion des Teilprozesses Attribuierung (siehe eEPK der Abb. 4.6) wird die Klasse eines Dokuments bestimmt. Für diesen Teilprozess ist wiederum ein Bearbeiter aus dem jeweiligen Fachbereich verantwortlich. Diese Funktion kann semi-automatisch ablaufen. Die automatische Einordnung in die Dokumentklassen erfolgt auf Basis der Struktur und des Inhalts der CI-Dokumente. Das Ergebnis wird vom Bearbeiter geprüft und gegebenenfalls korrigiert. Dazu ist es notwendig, dass das Dokument auf dem Bildschirm des Bearbeiters angezeigt wird und er das Dokument lesen darf.

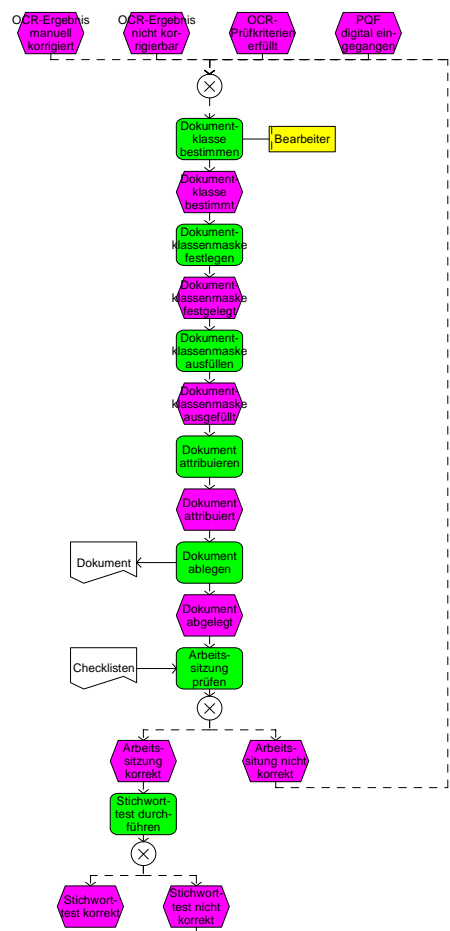


Abb. 4.6: Teilprozess Attribuierung

Die festgestellte Dokumentklasse bestimmt die Metadaten, die zu einem Dokument erfasst werden. Die Metadaten beschreiben das Dokument näher, erleichtern die Suche und ermöglichen das Wiederfinden des Dokuments im Ablagesystem. Beispielsweise können Angaben zum Autor oder Erscheinungszeitpunkt als Metadaten verwendet werden (vgl. Lorenz (2006), S. 66). Um die Metadaten einzugeben, wird die entsprechende Maske in der Funktion *Dokumentklassenmaske ausfüllen* aufgerufen. Jede

Dokumentenklasse besitzt allgemeine und spezielle Metadaten. Diese werden nicht im Fachkonzept beschrieben, da es sich bei der Auswahl eines Standards, z. B. dem Dublin Core, um eine technische Festlegung handelt, die nach dem Vorgehensmodell dem DV-Konzept zugeschrieben wird.

Die Metadaten können semi-automatisch vom Informationssystem vorgeschlagen werden. Der Bearbeiter prüft die Angaben, wenn es notwendig ist, korrigiert er sie. Nachdem die Metadaten festgelegt und die Dokumentklassenmaske ausgefüllt ist, wird semi-automatisch das *Dokument indiziert* und anschließend abgelegt. Die Ablage des Dokuments und seiner Metadaten kann getrennt erfolgen (vgl. Götzer et al. (2004), S. 186). Die Metadaten werden in einem für häufige Transaktionen optimiertes Datenbanksystem mit Verweis auf das jeweilige Dokument gespeichert.

Die Dokumente können in einer Archiv-Datenbank abgelegt werden. Die Ablage in einer Dokumentendatenbank ist für diese Zwecke optimal, aber nicht zwingend notwendig. Nach der Ablage wird der Bearbeiter seine *Arbeitssitzung prüfen*, dabei stellt er fest, ob alle Schritte durchgeführt und das Dokument und die Metadaten vollständig bearbeitet wurden. Wenn dies nicht der Fall ist, wird die Arbeitssitzung wiederholt. Die Prüfung kann mit Hilfe des Systems durch Checklisten unterstützt werden. Ist die Arbeitssitzung korrekt durchgeführt, erfolgt ein Stichworttest. Dabei stellt der Bearbeiter fest, ob er das Dokument im Ablagesystem wiederfinden kann und ob es angezeigt wird. Wenn der Stichworttest fehlschlägt, wird der Teilprozess Attribuierung erneut durchgeführt. Andernfalls endet dieser Teilprozess. Das Ergebnis der digitalen Aufbereitung ist ein kodiertes Dokument, wie z. B. ein PQF, das in einem standardisierten Format abgelegt wird.

4.2.2 Vorverarbeitung

Um die Inhalte der Texte maschinell zu erschließen, ist es notwendig, die Texte für nachfolgende Analysen vorzubereiten (vgl. Granitzer (2006), S. 25). Der Prozess der Vorverarbeitung dient der strukturellen Analyse und der Normalisierung der Präqualifikationsfragebögen und ist an den Prozess zur Erstellung von Textkorpora aus Abschnitt 2.3.2 angelehnt. Die Normalisierung schließt in diesem Prozess die Überführung verschiedener Formate in ein einheitliches zur Bearbeitung notwendiges Format ein. Dabei sind verschiedene Formate zu berücksichtigen.

Die Vorverarbeitung beginnt mit der strukturellen Analyse und Markierung der Merkmale und Layoutanweisungen in einem PQF (siehe Abb. 4.7). Der PQF wird mit seinen Metadaten in ein einheitliches Textformat transferiert. Nachdem die Struktur im

PQF markiert wurde, werden im nächsten Schritt die einzelnen Bereiche und die enthaltenen Fragen untersucht. Die Fragen sind in natürlicher Sprache formuliert. Dabei ist zu beachten, dass einige Fragen aus einem Wort, einer Wortgruppe, einem Satz oder mehreren Sätzen bestehen können. Ebenso sind sprachliche Eigenschaften u. a. Mehrdeutigkeit und Vagheit (vgl. Granitzer (2006), S. 25) problematisch. In der folgenden lexikalischen Analyse werden unter Verwendung der Tokenisierung und Satzgrenzenerkennung die Texte der Bereiche zerlegt und markiert. Danach werden durch Part-of-Speech-Tagger, Lemmatisierung und Chunk-Prasing die markierten Segmente mit linguistischen Informationen angereichert.

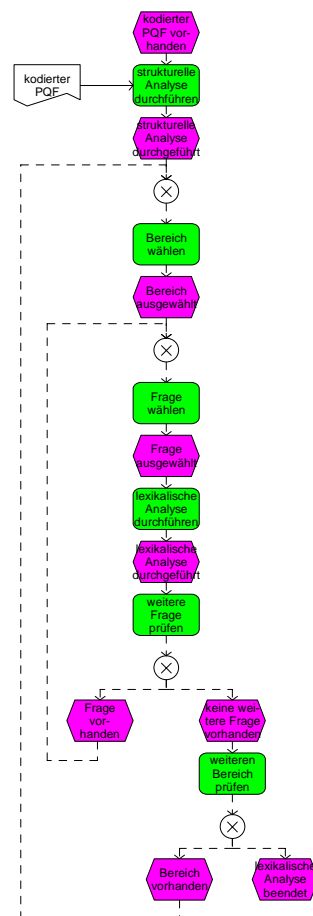


Abb. 4.7: Teilprozess Vorverarbeitung

Der vorverarbeitete PQF wird mit den Markierungen zur Weiterverarbeitung im Informationssystem abgelegt. Die Ablage der PQF erfolgt in einem standardisierten und textbasierten Format.

4.2.3 Semi-automatische Beantwortung

Der Teilprozess Beantwortung gliedert sich, wie in der Abb. 4.8 dargestellt, in die drei Teilprozesse Antwortgenerierung, PQF-Antwortbearbeitung und PQF-Prüfung. Im Folgenden werden diese Teilprozesse unter Verwendung von eEPK Modellen beschrieben.



Abb. 4.8: Übersicht des Teilprozesses Beantwortung

4.2.3.1 Antwortgenerierung

Der Teilprozess Antwortgenerierung beginnt nach der Vorverarbeitung mit einer semantischen Analyse eines PQFs (siehe Abb. 4.9). Dabei werden die Informationsbedarfe der Bereiche und Fragen identifiziert. Im Rahmen der semantischen Analyse wird mit Hilfe von Schlüsselwörtern, Synonymlisten und Phrasen der semantische Inhalt der Bereiche und Fragen bestimmt. Ziel ist es, den Informationsbedarf einer Frage zu erkennen, um die Suche nach diesen Informationen zu ermöglichen. Die erkannte Semantik wird zu den jeweiligen Fragen zu den Strukturelementen gespeichert.

Die automatische Informationssuche erstreckt sich über die PQF-Datenbank und über den integrierten Datenbestand des Unternehmens, der in Abschnitt 4.3 beschrieben wird. Durch die Funktion *PQF prüfen* wird geklärt, ob ein älterer PQF des anfragenden Unternehmens in der PQF-Datenbasis vorhanden ist. Die Prüfung erfolgt anhand der PQF-Metadaten, unter Berücksichtigung des anfragenden Unternehmens. Wenn bereits ein PQF vorhanden ist, wird die Funktion *PQF vergleichen* ausgeführt. Durch diese Funktion werden die Fragebögen auf Bereichs- und Fragenebene untersucht, um Gemeinsamkeiten und Unterschiede festzustellen. Die gespeicherten Antworten von identischen und ähnlichen Fragen werden in den PQF nach der Prüfung der Gültigkeit übertragen. Die Frage gilt dann als beantwortet. Dabei weisen identische Fragen den gleichen Wortlaut und Aufbau auf, während ähnliche Fragen einen unterschiedlichen Wortlaut und Aufbau haben, sich aber semantisch auf denselben Informationsbedarf beziehen. Die Wiederverwendung von Antworten sichert die Konsistenz des Auftretts gegenüber den anfragenden Unternehmen und reduziert die Bearbeitungszeit.

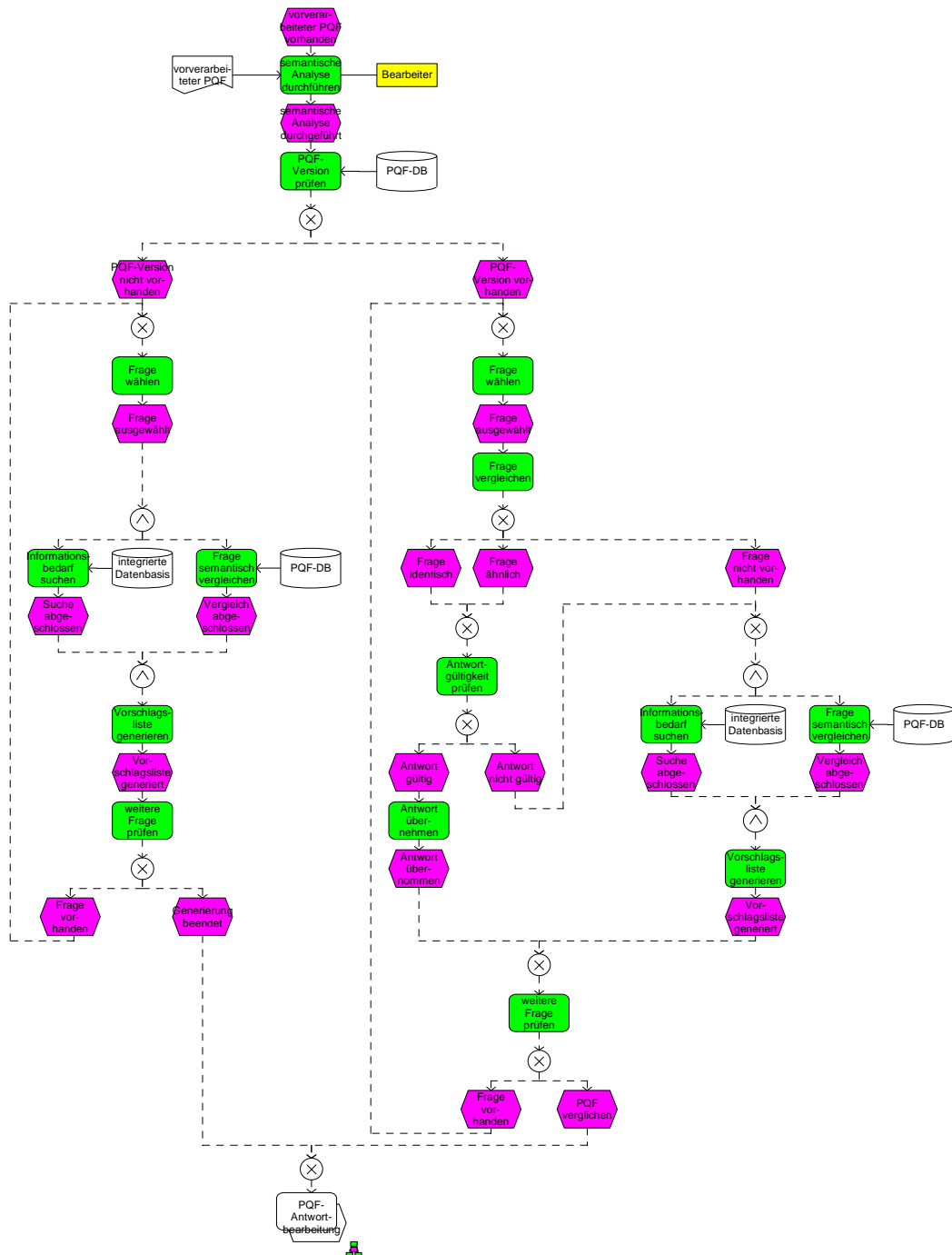


Abb. 4.9: Teilprozess Antwortgenerierung

Sind die Antworten nicht mehr gültig oder wurde kein abgelegter PQF des anfragenden Unternehmens gefunden, wird die Suche des benötigten Informationsbedarfs auf semantischer Ebene, über alle gespeicherten PQF-Antworten und auf den integrierten Datenbestand, erweitert. Aus beiden Funktionen werden die Antworten und Informationen verwendet, um eine Vorschlagsliste für die jeweilige Beantwortung der Frage zu erzeugen. Die im Folgenden durch den Bearbeiter geprüft wird. Eine Frage, für die eine Vorschlagsliste existiert, gilt als nicht beantwortet. Weitere Informationsbedarfe, wie Nachweis-Dokumente, werden als Referenz auf das jeweilige Dokument oder mit einem Vermerk „nicht vorhanden“ der Frage hinzugefügt. Wenn keine Antworten oder Informationen gefunden wurden, ist die Vorschlagsliste leer. Im nachfolgenden Teilprozess die Antwortbearbeitung der unbeantworteten Fragen.

4.2.3.2 PQF-Antwortbearbeitung

Der Teilprozess der PQF-Antwortbearbeitung erfolgt für jede nicht beantwortete Frage im aktuellen Präqualifikationsfragebogen (siehe Abb. 4.10). Zu Beginn dieses Teilprozesses wird eine nicht beantwortete Frage ausgewählt. Da jede Frage beantwortet werden muss, kann die Auswahl automatisch erfolgen. Die PQF-Antwortbearbeitung wird beendet, wenn keine unbeantworteten Fragen, also Fragen mit einer Vorschlagsliste, vorhanden sind. Mit der ausgewählten, unbeantworteten Frage wird dem Bearbeiter die Vorschlagsliste vom Informationssystem angezeigt. An dieser Stelle teilt sich der Prozess durch die folgenden zwei Ereignisse: *Vorschlag vorhanden* und *Vorschlagsliste ist leer*. Im ersten Fall, wenn mindestens ein *Vorschlag vorhanden* ist, wird durch den Bearbeiter in der Funktion *Vorschlag auswählen* die bestmögliche Antwort für den PQF ermittelt und im nächsten Schritt geprüft. Dabei kann eine *Antwort inhaltlich und formal richtig*, *inhaltlich richtig aber formal falsch* oder *inhaltlich falsch sein*. Ist die *Antwort inhaltlich und formal richtig* kann sie in den PQF übernommen werden. Wenn die *Antwort inhaltlich richtig aber formal falsch* ist, wird sie in die gewünschte Form, u. a. Freitext, Listen oder Tabellen, überführt und kann anschließend in den PQF aufgenommen werden. Das dritte Ereignis *Antwort ist inhaltlich falsch* löst, wie das Ereignis *Vorschlagsliste ist leer*, die *manuelle Suche* nach dem Informationsbedarf aus.

Für diese Funktion stehen dem Bearbeiter verschiedene Mechanismen zur Verfügung. Der Bearbeiter kann eigene Anfragen an das Informationssystem und seine integrierte Datenbasis formulieren. Für diese Suche kann ein Expertenmodus die Eingabe von Anfrage-Statements ermöglichen. Da aber nicht jeder Bearbeiter mit Anfragesprachen an Datenbanken vertraut ist, wird eine benutzerfreundliche Art der Informationssuche angeboten. Für die Suche nach gespeicherten PQF wird ein Abfrageformular

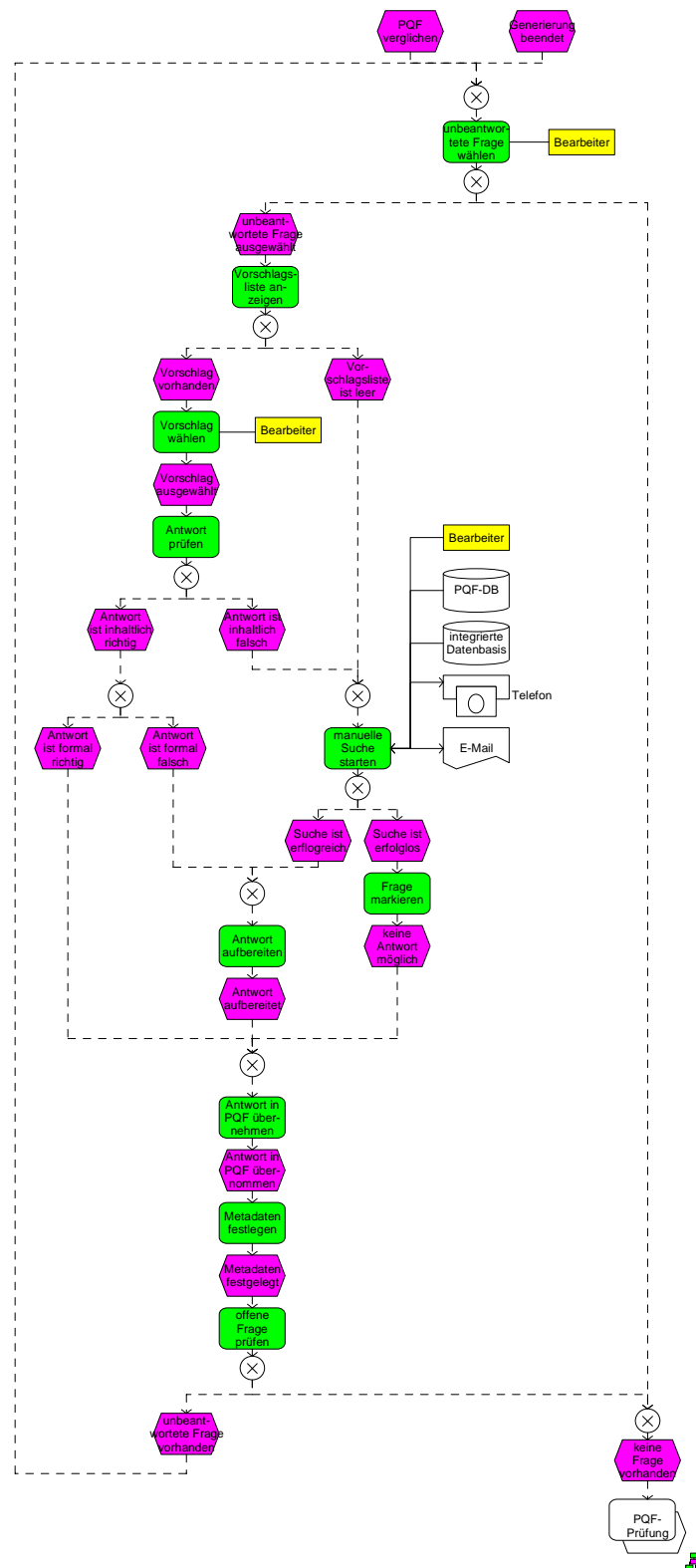


Abb. 4.10: Teilprozess PQF-Antwortbearbeitung

bereitgestellt, mit dessen Hilfe der Bearbeiter nach ein oder mehreren Metadaten der PQF suchen kann. Die Ergebnisse werden dem Bearbeiter auf dem Bildschirm angezeigt. Für die Suche nach dem Informationsbedarf in der integrierten Datenbasis, wird die Methode des Information-Browsings in einem graphischen Benutzerinterface verwendet. Dabei gelangt der Bearbeiter über die Auswahl des Unternehmens in die einzelnen Informationsbereiche und schließlich zu den Informationsobjekten und deren Attribute. Für die Auswahlmöglichkeiten der obersten Ebenen sollten große, leicht zu interpretierende Symbole verwendet werden. Auf unterster Ebene werden dann die einzelnen Informationen dargestellt. Ist die Suche erfolgreich, werden die gefundenen Informationen ausgewählt und in einen Editor zur Antwortaufbereitung übernommen. Wie im Ist-Prozess steht dem Bearbeiter frei, ebenfalls über Telefonate die Suche nach den Informationen durchzuführen. Eine weitere Möglichkeit ist die Weiterleitung der Frage an eine zuständige Abteilung im Unternehmen. An dieser Stelle entsteht, wie im Ist-Prozess, ein Bruch, dessen zeitliche Komplexität nicht überschaubar ist. Aus diesem Grund sollte diese Funktion, obwohl sie die gängige Praxis darstellt und deshalb hier modelliert wurde, nicht genutzt werden. Es kann auch vorkommen, dass es zu einigen Fragen, die bspw. zu einem anderen Gewerk gehören, keine Informationen im Unternehmen gibt. In diesem Fall ist die *Suche erfolglos* und die Frage wird im PQF mit „keine Informationen vorhanden“ markiert und dementsprechend im PQF beantwortet.

Nachdem die Antwort in die gewünschte Form gebracht und in den PQF aufgenommen wurde, werden semi-automatisch Metadaten für die Frage und die Antwort erzeugt. Die automatisch vergebenen Metadaten beziehen sich auf die Semantik des Informationsbereichs und des Informationsbedarfs, die Art der Darstellung und auf die zeitliche Gültigkeit der Antwort. Diese Metadaten werden für die automatische Suche nach den Informationsbedarfen in der PQF-Datenbank und der Antwortübernahme benötigt. Diese Angaben werden nochmals geprüft, gegebenenfalls korrigiert und gespeichert. Als Richtwert für die Gültigkeit könnte an dieser Stelle ein Jahr dienen. Die Gültigkeit ist vom Informationsobjekt der Antwort und dessen Halbwertszeit abhängig. Genauere Untersuchungen zum Aktualitätsverfall wurden in dieser Arbeit nicht durchgeführt.

Schließlich wird geprüft, ob noch weitere Fragen nicht beantwortet sind. Ist das der Fall, wird die Funktion *unbeantwortete Frage wählen* erneut aufgerufen. Wurden hingegen alle Fragen beantwortet, ist keine weitere Frage vorhanden und der Teilprozess PQF-Antwortbearbeitung ist beendet.

4.2.3.3 PQF-Prüfung

Im Teilprozess PQF-Prüfung, dargestellt in Abb. 4.11, wird der beantwortete PQF nochmals geprüft. Er beginnt durch das im vorherigen Teilprozess erzeugte Ereignis *keine Frage vorhanden*. Diese Prüfung schließt alle Antworten des PQFs ein und erlaubt dem Bearbeiter im Bedarfsfall deren Korrektur. Die Antworten können mit den im Teilprozess PQF-Antwortbearbeitung beschriebenen Vorgehen korrigiert werden. Wenn die vollständige Bearbeitung des PQFs festgestellt wird, kann der PQF mit den Antworten im Informationssystem abgelegt werden. Aus den abgelegten Antworten wird das Antwortdokument, in der Funktion *PQF verschicken*, erzeugt, welches an das anfragende Unternehmen zurückgesendet wird. Hier sollten vier Fälle berücksichtigt werden, die sich aus den unterschiedlichen Formaten und Präferenzen der anfragenden Unternehmen ergeben.

Der erste Fall ist die Nutzung des Originaldokuments, wenn dieses eine Bearbeitung zulässt. Dabei werden die gespeicherten Antworten in das Originaldokument übertragen. Die Nachweise werden bei dieser Variante aus der Datenbasis in ein vom anfragenden Unternehmen bestimmtes Format exportiert und dann dem PQF angehängt.

Im zweiten Fall wird das Originaldokument nicht bearbeitet. Die Antworten werden als Anlage hinzugefügt. Dabei wird ein Antwortdokument mit Bezug auf die Fragen des PQF erzeugt und an das Originaldokument angehängt. Ebenso verhält es sich mit den benötigten Nachweisen.

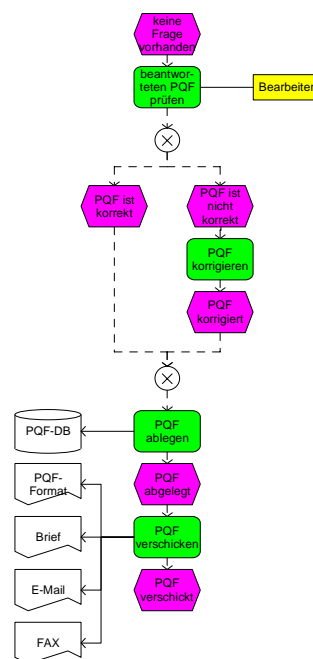


Abb. 4.11: Teilprozess PQF-Prüfung

Der dritte Fall stellt eine Kombination der beiden zuvor genannten Fälle dar. Dabei muss entschieden werden, welche Fragen im Originaldokument und welche in der Anlage beantwortet werden. Dieses Vorgehen kann wiederum semi-automatisch erfolgen. Nachdem eine Antwort in das Originaldokument eingefügt wurde, wird bspw. das Layout auf unerwünschte Verschiebungen untersucht. Dies bedeutet, dass für die Antwort nicht genügend Platz im Dokument vorhanden ist und die Antwort ausgelagert werden muss. In den dargestellten drei Fällen ist die Übertragung des PQF mit den benötigten Anlagen an das anfragende Unternehmen per E-Mail, Brief oder Fax, je nach Wunsch, möglich. Es können durchaus kombinierte Übertragungen stattfinden.

Der vierte Fall ist die digitale Übertragung der PQF in einem einheitlichen standardisierten Format ohne Berücksichtigung des Originaldokuments. Diese Variante hat einige Vorteile. Durch diese Übertragung findet zwischen den Unternehmen kein Medienbruch statt. Die gesammelten Informationen können zur Auswertung direkt, ohne zeitaufwendige und fehleranfällige Digitalisierungs- und Extraktionsprozesse, in das eigene Informationssystem übernommen werden. Dadurch wäre eine Effizienzsteigerung des gesamten PQ-Prozesses möglich. Dem anfragenden Unternehmen ist es somit möglich, sich schneller für ein Subunternehmen zu entscheiden. Daraufhin mit der Kalkulation früher zu beginnen und demnach schneller ein Angebot beim Kunden einzureichen, welches dann schließlich zu einem Auftrag führen kann.

4.2.4 Datenschema des Präqualifikationsfragebogens

Das Datenschema der PQF beschreibt, wie die PQF in der PQF-Datenbank abgelegt werden (siehe Abb. 4.12).

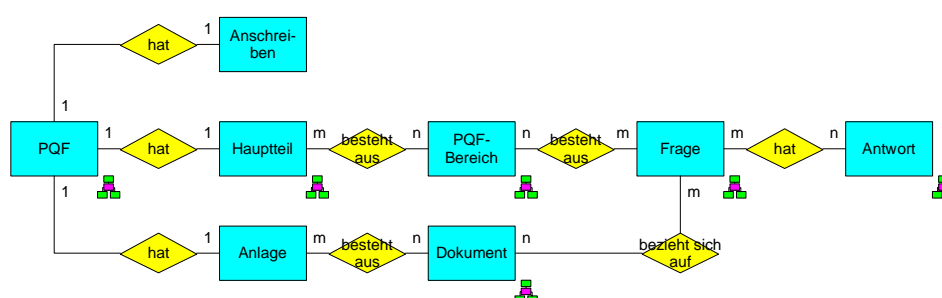


Abb. 4.12: Datenschema des Präqualifikationsfragebogens

Der PQF gliedert sich in die bereits erwähnten Teile Anschreiben, Hauptteil und Anlagen. Der Hauptteil beinhaltet ein oder mehrere P-QF-Bereiche und die wiederum eine oder mehrere Fragen. Zu einer Frage können eine oder mehrere Antworten gehören. Der PQF, die Fragen und die Antworten enthalten bestimmte Metadaten,

die im Attributzuordnungsdiagramm der Abb. B.1 im Anhang skizziert werden. Dabei ist zu berücksichtigen, dass der Wortlaut von Fragen und Antworten die syntaktische Struktur von Phrasen und die Semantik enthält. Diese Art der Speicherung soll den Grad der Wiederverwendung von Antworten fördern. Das bedeutet, dass eine Antwort mehreren Fragen zugeordnet werden kann. Die Fragen unterscheiden sich voneinander durch unterschiedliche Formulierungen sowie eventuell unterschiedlichen Bezeichnungen, die durch die semantische Analyse in den Informationsbedarf aufgelöst wurde. Im Anlagenteil werden die benötigten Antwortdokumente einschließlich der Nachweis-Dokumente aufgeführt, welche sich auf die im PQF enthaltenen Fragen beziehen.

4.3 Schemata der Datenbasis und Teilsysteme

Die Analyse in Kapitel 3 zeigte die Relevanz der Bereiche allgemeine Unternehmensdaten, Ansprechpartner, Mitarbeiter, Referenzen, Leistungsspektrum, Qualitätssicherung und Sicherheit/Gesundheit/Umwelt. Die Informationen zu den entsprechenden Bereichen sind in einer heterogenen Systemlandschaft verteilt. Da diese Systemlandschaft als unbekannt einzustufen ist, werden ein föderiertes Datenschema und die benötigten Teilsysteme modelliert. Die dabei verwendete Vorgehensweise entspricht dem Top-Down-Ansatz. In den folgenden Abschnitten werden die Daten und Teilsysteme beschrieben, welche die integrierte Datenbasis des Informationssystems bilden soll.

4.3.1 Föderiertes Datenschema

Der zentrale Entitätstyp ist das Unternehmen (siehe Abb. 4.13). Der Besitzer eines Unternehmens kann als Entitätstyp zu Unternehmen und Person spezialisiert werden. Dabei kann jeder Besitzer mehrere Unternehmen und jedes Unternehmen einen oder mehrere Besitzer haben (n:m-Beziehung). Einem Unternehmen können verschiedene Kontaktdaten und Ansprechpartner zugeordnet werden (1:n-Beziehung). Ein Ansprechpartner ist für einen oder mehrere Unternehmensbereiche zuständig. Zwischen diesen beiden Entitätstypen besteht eine m:n-Beziehung, weil ebenfalls jedem Unternehmensbereich mehrere Ansprechpartner zugeordnet werden können. Zum Ansprechpartner gehören, wie zum Unternehmen, Kontaktdaten. Der Entitätstyp Ansprechpartner ist eine Person und erbt dessen Attribute, auf die an späterer Stelle noch eingegangen wird. Eine Person gehört im Unternehmen einer Berufsgruppe an. Es müssen aber nicht alle Personen im Unternehmen in diesem Datenmodell erfasst werden, obwohl dadurch die aggregierten Informationen des PQF-Bereich Mitarbeiter berechnet

werden können (vgl. Abschnitt 3.2.3.7). Um die benötigten Informationsbedarfe zu befriedigen, werden die Entitätstypen Mitarbeiterstatistik und Berufsgruppe eingeführt und im Abschnitt 4.3.3 näher erläutert. Den Berufsgruppen werden unterschiedliche Qualifikationen zugeordnet. Dabei kann durch die m:n-Beziehung eine Qualifikation in verschiedenen Berufsgruppen verwendet werden. Zusätzlich bezieht sich eine oder mehrere Qualifikationen auf ein oder mehrere Leistungsbereiche, die wiederum dem Unternehmen zugeordnet werden.

Neben dem Qualifikationsdokument werden in diesem Modell noch weitere Dokumente berücksichtigt. Der Entitätstyp Dokument vererbt seine Attribute an die Entitätstypen Versicherung, Report, Managementsystem-Zertifizierung, Organigramm, Qualifikation und den PQF. Der PQF wird im integrierten Datenschema (Abb. 4.13) nicht aufgeführt, weil er vom Hauptprozess verwaltet wird. Der Entitätstyp Report teilt sich in die Entitätstypen Vorfall-Statistik, DUNS-Report, Jahresabschluss-, Umwelt- und Sicherheitsbericht. Dabei werden die Versicherungen, Reports und Organigramme dem Unternehmen direkt zugeordnet.

Weiterhin gehören zu einem Unternehmen Referenzen und Managementsysteme (MS). Die Referenzen spalten sich in die zwei Entitätstypen Projekte und Geschäftsbeziehungen. Projekte werden für einen Kunden mit oder ohne Generalunternehmen aber mit einem oder mehreren Subunternehmen durchgeführt. General-, Subunternehmen und Kunden werden zu Unternehmen generalisiert und stellen die Rollen von Unternehmen dar. Ein Projekt betrifft verschiedene Leistungsbereiche eines Unternehmens. Die Leistungsbereiche bilden das Leistungsspektrum und beinhalten dessen Informationsbedarfe (vgl. Abschnitt 3.2.3.3). Zu einem Projekt wird das Dokument Vorfall-Statistik hinzugefügt, dabei ist dieses Dokument an ein spezifisches Projekt gebunden. Dies bedeutet, dass einem Projekt eine Vorfall-Statistik zugeordnet werden kann. Dies kann während des Projektes oder nach dessen Beendigung hinzugefügt werden. Geschäftsbeziehungen, die nicht durch ein Projekt dargestellt werden können, bestehen zwischen mindestens zwei Unternehmen.

Ein Unternehmen kann ein oder mehrere Managementsysteme besitzen. Um auf die Informationsnachfrage der PQF reagieren zu können, wird dem Managementsystem ein Unternehmensbereich und dessen Ansprechpartner, sowie mehrere Vorgehensweisen zugeordnet. Zu jedem Managementsystem können verschiedene Dokumente, wie Reports, MS-Handbuch und die MS-Zertifizierung, verwaltet werden. Wenn ein Managementsystem durch eine Software im Unternehmen unterstützt wird, kann diese durch den Entitätstyp Sicherungssystem beschrieben werden.

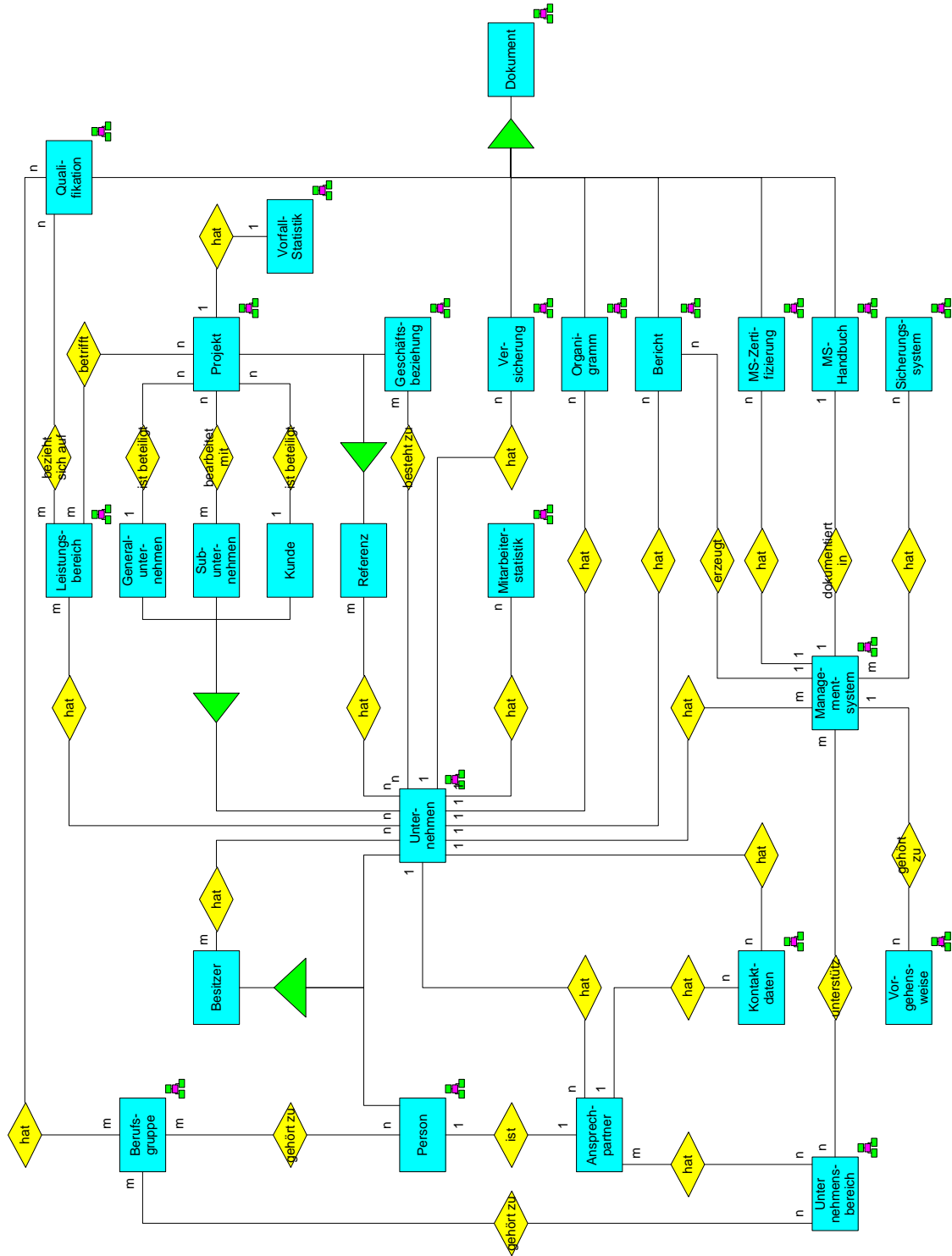


Abb. 4.13: Föderiertes Datenschema

Aus dem beschriebenen föderierten Datenschema lassen sich vier Teilsysteme ableiten. In den folgenden Abschnitten werden die Funktionalitäten dieser Teilsysteme zur Datenverwaltung und das jeweilige Datenschemata beschrieben.

4.3.2 Unternehmensverwaltung

Im Teilsystem Unternehmensverwaltung werden die allgemeinen Unternehmensdaten gepflegt. Bei diesem und den anderen Teilsystemen, für die Referenzverwaltung, Mitarbeiterstatistik und Managementsysteme, handelt es sich um reine Informationsverwaltungssysteme. Diesen Systemen liegen sequenzielle Prozesse zu Grunde, deshalb wird auf die Darstellung durch ePKs verzichtet. Stattdessen werden die benötigten Funktionen als Funktionsbaum (siehe Abb. 4.14) und die dazugehörigen Datenschemata als eERM (siehe Abb.4.15) dargestellt.

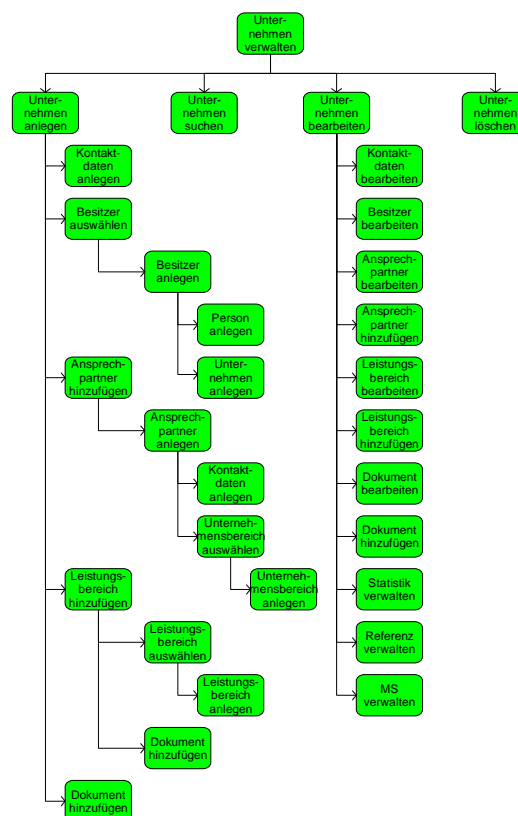


Abb. 4.14: Funktionsübersicht der Unternehmensverwaltung

Der Funktionsbaum der Unternehmensverwaltung gliedert sich objektorientiert in die Funktionen *Unternehmen anlegen*, *suchen*, *bearbeiten* sowie *löschen*. Unter Verwendung der Suche kann ein Unternehmen in der Datenbasis gefunden, angezeigt und dann gelöscht oder bearbeitet werden. Die Bearbeitung erfolgt prozessorientiert an den Objekten, die zu einem Unternehmen angelegt werden. Dies erfolgt durch die Funktion

Unternehmen anlegen. Dabei werden die Attribute eines Unternehmens erfasst (siehe Attributzuordnungsdiagramme in Abb. B.2). Zu den in dieser Funktion erfassten Attributen gehören der Name, die Anschrift, die Rechtsform des Unternehmens, dessen Gründungsort und die Internetadresse. Die Anschrift setzt sich aus den Informationen Straße, Hausnummer, Postleitzahl, Ort, Bundesland und Postfach zusammen.

Nachdem die allgemeinen Informationen angelegt sind, werden die Kontaktdaten angelegt. Für ein Unternehmen müssen mehrere Kontaktdaten, personenabhängige und personenunabhängige Kontaktdaten, verwaltet werden. An dieser Stelle ist der personenunabhängige Kontakt des Unternehmens gemeint. Zu den Kontaktdaten gehören E-Mail, Telefon- und Faxnummer (abgebildet im Attributzuordnungsdiagramm in Abb. B.2).

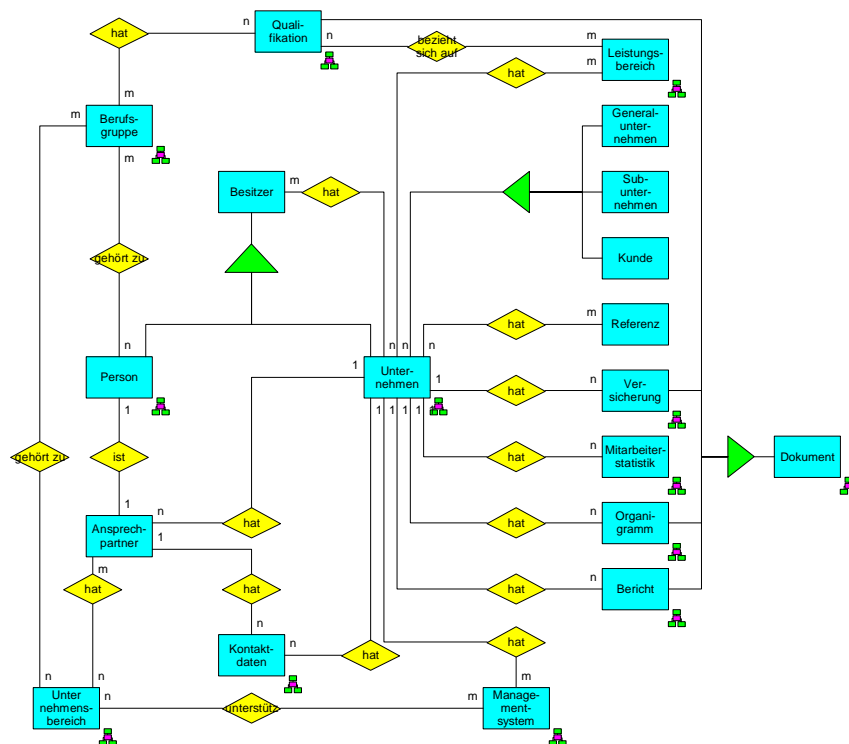


Abb. 4.15: Datenschema der Unternehmensverwaltung

Eine weitere Information, die zum Unternehmen festgehalten wird, betrifft dessen Besitzer. Die Besitzer eines Unternehmens können Unternehmen (z. B. Holdings) oder Personen sein. In der Funktion *Besitzer hinzufügen* werden Personen oder Unternehmen ausgewählt. Wenn die benötigten Instanzen noch nicht in der Datenbasis vorhanden sind, werden sie durch die Funktionen *Person anlegen* oder die oben beschriebene Funktion *Unternehmen anlegen* erzeugt. Personen werden mit den Attributen Anrede, Nachname, Vorname und Geschlecht angegeben.

Nachdem die Besitzerdaten gespeichert sind, werden die Ansprechpartner des Unternehmens mit den zugehörigen Kontaktdaten und den Unternehmensbereichen angelegt (siehe Abb. 4.15). Ein Ansprechpartner ist eine Person und erbt somit deren Attribute. Zusätzlich hat ein Ansprechpartner personenabhängige Kontaktdaten, wie oben beschrieben, und einen Zuständigkeitsbereich. Wenn der benötigte Unternehmensbereich nicht in der Datenbasis vorhanden ist, wird dieser angelegt.

Zum Unternehmen werden Leistungsbereiche erstellt, welche mit ihren Attributen registriert werden. Ein Leistungsbereich hat eine Bezeichnung und eine kurze Beschreibung. Dann ist die Art der Leistung, d. h. ob es sich bei der Leistung um Eigen- oder Fremdleistung handelt, auszuwählen. Zu einem Leistungsbereich können Dokumente zugeordnet werden. Diese Qualifikationsdokumente werden gescannt, attribuiert und mit Bezug zum Leistungsbereich abgespeichert, dies erfolgt ebenfalls durch den Teilprozess Digitalisierung (siehe Abschnitt 4.2.1).

Zu einem Unternehmen gehören verschiedene Dokumente (siehe Abb. 4.16). Die Dokumente, die nicht im Unternehmen erstellt werden, müssen, wenn sie nicht digital verfügbar sind, digitalisiert werden (vgl. Abschnitt 4.2.1). Zu diesen Dokumenten gehören die Versicherungszertifikate, DUNS-Bericht, Qualifikations- und MS-Zertifizierungsdokumente.

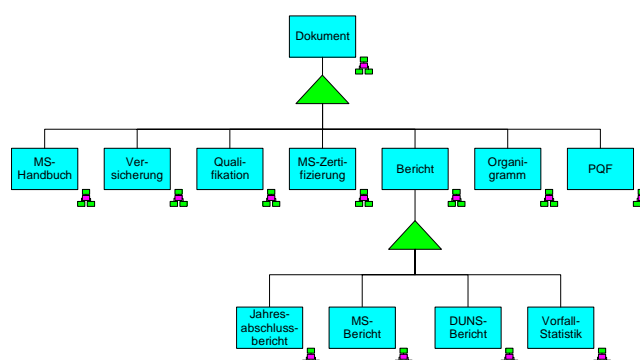


Abb. 4.16: Dokumentenmodell

Die Attribute eines Dokuments sind die Art des Dokuments, seine Gültigkeit, eine Beschreibung und ein Standardmetadatensatz. Auf diesen Metadatensatz wird an dieser Stelle nicht eingegangen, weil verschiedene Standards, wie z. B. Dublin Core, existieren und deren Festlegung in das DV-Konzept einzuordnen ist. Die Attribute des Entitätstyps Dokument werden an die untergeordneten Entitäten vererbt und durch diese um die spezifischen Metadaten erweitert (siehe Abb. B.3 im Anhang).

Das Versicherungszertifikat besitzt die zusätzlichen Attribute Versicherungsart, Deckungshöhe, Versicherungsgesellschaft, Versicherungsnummer und Abschlussdatum.

Das Dokument MS-Zertifizierung hat die Attribute Zertifizierungsstelle und Bezeichnung, während das Qualifikationsdokument zusätzlich noch einen Inhaber aufweist. Der Entitätstyp Bericht wird um Angaben zum Datum ergänzt. Dieses Datum wird wieder an alle untergeordneten Entitätstypen vererbt. Der DUNS-Bericht wird um die DUNS-Nummer ergänzt.

Die Dokumente, die vom Unternehmen selbst angefertigt werden, können direkt als kodierte Dokumente mit Attributen versehen und in die Datenbasis des Informationssystems aufgenommen werden. Dazu sollten die bestehenden Prozesse zur Dokumentenablage im Unternehmen angepasst werden. Im Grunde wird im Anschluss an den Erstellungsprozess des jeweiligen Dokuments der Attribuierungsprozess ausgeführt (siehe Abb. 4.6 auf der Seite 68) und das Dokument gespeichert. Die Attribuierung erfolgt idealerweise durch das jeweilige Anwendungssystem. Ist die Anpassung des Fremdsystems nicht möglich, ist es notwendig, die Metadaten semi-automatisch aus dem Dokument zu extrahieren. Diese Erweiterung betrifft den Erstellungsprozess der folgenden Dokumente: Organigramme, MS-Handbuch, Vorfall-Statistik, Jahresabschluss-, Umwelt- und Sicherheitsbericht. Die spezifischen Metadaten des Jahresabschlussberichts sind die Informationen Umsatz, Jahr sowie Währung (vgl. Abschnitt 3.2.3.1). Der Vorfall-Statistik werden projektspezifische Unfallkennzahlen zugeordnet (vgl. Abschnitt 3.2.3.5). Des Weiteren werden von den Managementsystemen Berichte erzeugt, die um die spezifischen MS-Kennzahlen ergänzt werden. Die vollständige Darstellung dieser spezifischen Metadaten ist nicht Bestandteil dieser Arbeit.

Zum Unternehmen werden auch Referenzen, Statistik und Managementsysteme zugeordnet. Diese Bereiche werden in den separaten Teilsystemen der nächsten Abschnitte beschrieben.

4.3.3 Statistikverwaltung

Die benötigten Informationen zum Bereich Mitarbeiter, wie in Abschnitt 3.2.3.7 beschrieben, sind aggregierte Informationen. Dabei handelt es sich um die Anzahl der Mitarbeiter gesamt und gewerblich aufgeschlüsselt nach Berufen und Bereichen des Unternehmens, das Durchschnittsalter, verfügbare Kapazität, gebundene Kapazität, Fluktuationsstatistik und Qualifikationen.

Um die benötigten Informationsbedarfe zu befriedigen, wird ein Teilsystem entwickelt, das diese statistischen Informationen über die Entitätstypen Mitarbeiterstatistik und Berufsgruppe verwaltet. Zu diesen Informationsbedarfen werden verschiedene

Funktionen benötigt (siehe Abb. 4.17). Die Mitarbeiterstatistik wird mit den Attributen Jahr, Quartal, Gesamtmitarbeiteranzahl, Durchschnittsalter, Gesamtarbeitsstunden und der Anzahl von Festangestellten angelegt und dem Unternehmen zugeordnet.

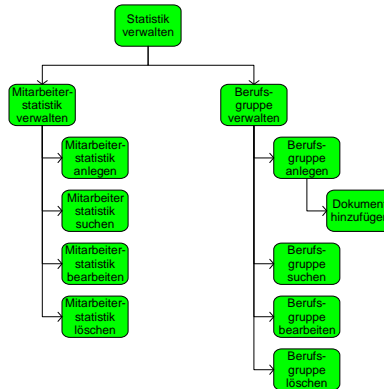


Abb. 4.17: Funktionsübersicht der Statistikverwaltung

Für jede Berufsgruppe werden die Attribute Bezeichnung, Mitarbeiteranzahl, Durchschnittsalter, Anzahl der Festangestellten, Jahr und Quartal angelegt (siehe Abb. B.4 im Anhang). Eine Berufsgruppe kann mehrere Qualifikationen haben und wird mehreren Bereichen im Unternehmen zugeordnet. Des Weiteren werden Such-, Bearbeitungs- und Löschfunktionen benötigt. Das beschriebene Datenschema wird in der Abb. 4.18 dargestellt. Bei den hier dargestellten Informationsbedarfen handelt es sich um zusammengefasste Informationen, die in der Personalabteilung verwaltet werden. Diese Abteilung müsste die Statistiken quartalsweise erstellen und speichern. Um den Aufwand gering zu halten, könnte dies durch einen Agenten automatisch erfolgen. Somit sind auch die persönlichen Daten eines Mitarbeiters vor unbefugtem Zugriff geschützt, da der Hauptprozess nur auf die Statistiken zugreift.

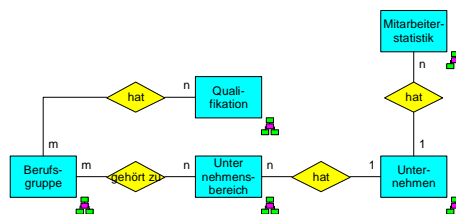


Abb. 4.18: Datenschema der Statistikverwaltung

Aus diesen Informationen lassen sich die Informationsbedarfe der PQF berechnen. Die gesamte Mitarbeiteranzahl und die gewerbliche Anzahl der Mitarbeiter in Form der Festangestellten sind gegeben. Die Differenz aus der Gesamtanzahl und den Festangestellten ergibt die Anzahl der Leasingmitarbeiter. Die Kapazität ist gegeben und bezieht sich auf die gesamten Arbeitsstunden. Die gebundene Kapazität ergibt sich aus der

Summe, der in den laufenden Projekten veranschlagten Arbeitsstunden. Aus der Differenz zwischen gesamtter und gebundener Kapazität ergibt sich die verfügbare Kapazität. Die Fluktuationsstatistik über die Anzahl der Mitarbeiter lässt sich über die Jahre und Quartale ermitteln. Durchschnittsalter und Qualifikationen sind ebenfalls gegeben.

4.3.4 Referenzverwaltung

Zum Bereich der Referenzen gehören in den PQF die Projekte und Geschäftsbeziehungen. Für die Referenzverwaltung werden die folgenden funktionalen Anforderungen aufgestellt. Die benötigten Funktionen der Referenzverwaltung gliedert sich prozessorientiert in die zwei Hauptfunktionen *Projekt verwalten* und *Geschäftsbeziehung verwalten* (siehe Abb. 4.19).

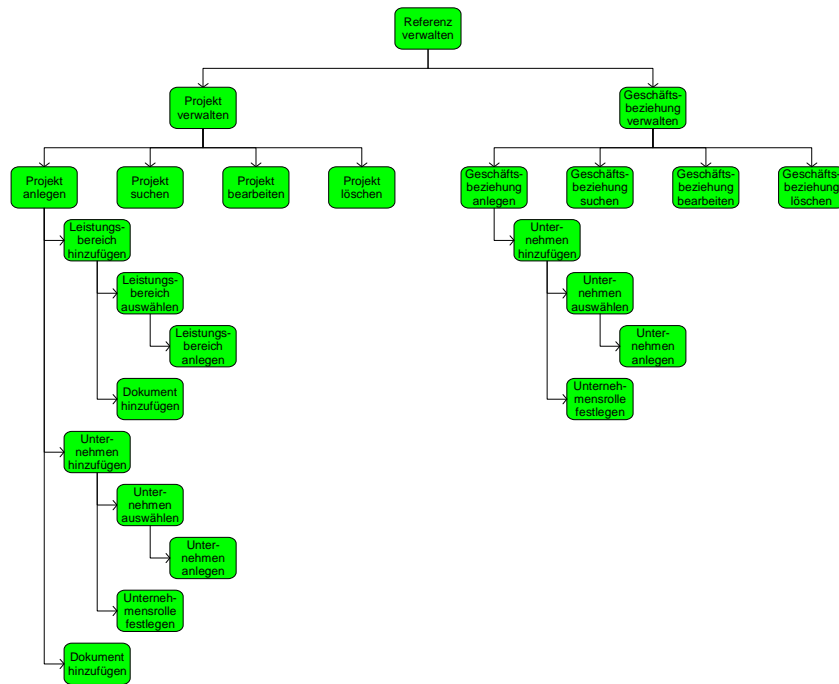


Abb. 4.19: Funktionsübersicht der Referenzverwaltung

Objektorientiert werden die Funktionen *Projekt anlegen*, *bearbeiten*, *suchen* und *löschen* der Funktion *Projekt verwalten* untergeordnet. Wenn ein Projekt in das Informationssystem aufgenommen werden soll, wird die Funktion *Projekt anlegen* ausgeführt. In dieser Funktion wird eine Eingabemaske aufgerufen, in der die Attribute eines Projekts Bezeichnung, Standort, Auftragsgröße, Vertragsart, geplantes Anfangsdatum und das geplante Enddatum eingegeben werden. Zu den einzugebenden Plandaten gehören ebenfalls die Angaben zur Anzahl des Aufsichtspersonals und geplante Arbeitsstunden. Bei abgeschlossenen Projekten werden zusätzlich die tatsächlich

geleisteten Arbeitsstunden, der Prozentsatz der Eigenleistung, das Anfangsdatum und das Enddatum hinzugefügt. Dies kann durch die Funktion *Projekt bearbeiten* erfolgen. Die angegebenen Attribute werden in der Abb. B.5 auf der Seite 110 im Anhang dargestellt.

Der Funktion *Projekt anlegen* sind die Funktionen *Leistung hinzufügen*, *Unternehmen hinzufügen* und *Dokument hinzufügen* prozessorientiert untergeordnet. Aus dem Leistungsspektrum des Unternehmens werden die Leistungsbereiche, die das Projekt betreffen, ausgewählt. Wenn der entsprechende Leistungsbereich nicht im System vorhanden ist, wird die Leistung, wie in Abschnitt 4.3.2 beschrieben, angelegt und ausgewählt. Ebenso verhält es sich mit den Unternehmen. Ein Projekt wird zwischen verschiedenen Unternehmen durchgeführt. In der Funktion *Unternehmen hinzufügen* wird das beteiligte Unternehmen ausgewählt und falls nicht vorhanden angelegt. Nach dem das Unternehmen ausgewählt ist, wird ihm eine der folgenden Rollen: General-, Subunternehmen oder Kunde zugeordnet. Für die Auswahlfunktionen sind entsprechende Suchmasken notwendig. Einem Projekt können verschiedene Dokumente hinzugefügt werden, an dieser Stelle wird die Vorfall-Statistik dem Projekt zugeordnet. Nachdem ein Projekt mit den zur Verfügung stehenden Informationen angelegt ist, kann es durch die Funktion *Projekt suchen* wiedergefunden werden. Die Suche kann über die Projektbezeichnung, Auftragsgröße, beteiligte Unternehmen, Leistungsbereiche, Zeitspanne der Abschlüsse und über den Status der Projekte durchgeführt werden. Der Status der Projekte leitet sich aus den Attributen Anfangs- und Enddatum ab. Projekte können „geplant“, „aktiv“ und „beendet“ sein. Die beendeten Projekte haben ein Anfangs- und ein Enddatum, aktive Projekte haben ein Anfangs-, aber kein Enddatum und geplante Projekte haben kein Anfangs- und kein Enddatum. Die Suchergebnisse werden in einer Liste angezeigt und ausgewählt.

Geschäftsbeziehungen, die nicht in einem Projekt erfasst sind, werden in der Hauptfunktion *Geschäftsbeziehungen verwalten* gepflegt (siehe Abb. 4.19). Die objektorientiert, untergeordneten Funktionen sind *Geschäftsbeziehung anlegen*, *suchen*, *bearbeiten* und *löschen*. Geschäftsbeziehungen werden mit den Attributen Anfangsdatum, Bezeichnung und Vertragsart angelegt (siehe Abb. B.5 auf der Seite 110 im Anhang). Zu einer Geschäftsbeziehung werden Unternehmen hinzugefügt. Falls das Unternehmen nicht in der Datenbasis erfasst ist, wird das Unternehmen neu angelegt. Dies erfolgt nach demselben Muster, wie in Abschnitt 4.3.2 beschrieben. In der Funktion *Geschäftsbeziehung suchen* wird eine Maske zur Verfügung gestellt, durch die nach Vertragsart, Bezeichnung und Unternehmen gesucht werden kann. Die Ergebnisse werden in Form einer Liste angezeigt und können durch die Funktion *Geschäftsbeziehung bearbeiten* verändert werden.

Das Datenschema der Referenzverwaltung wird im eERM der Abb. 4.20 dargestellt. Eine Geschäftsbeziehung besteht immer zwischen zwei verschiedenen Unternehmen. Da nicht jede Geschäftsbeziehung durch ein Projekt erfasst werden kann, wurde das Konstrukt Unternehmen hat Referenz, spezialisiert zu Geschäftsbeziehung, welche zu Unternehmen besteht modelliert. Dadurch ist es möglich nicht nur die Geschäftsbeziehungen des eigenen Unternehmens, sondern auch Geschäftsbeziehungen von Geschäftspartnern darzustellen und zu verwalten.

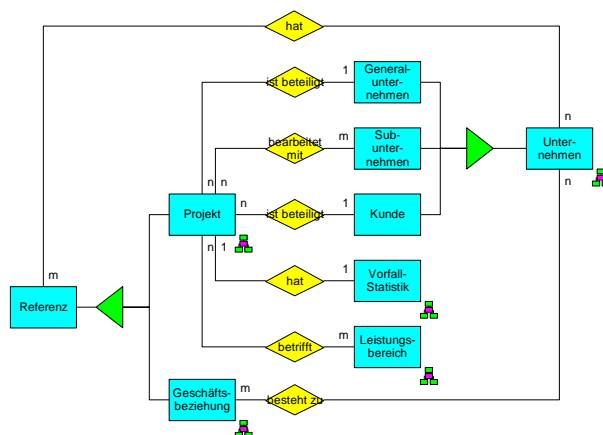


Abb. 4.20: Datenschema der Referenzverwaltung

4.3.5 Managementsystemverwaltung

Der Prozess der PQF-Beantwortung, wie im Abschnitt 3.2.3 beschriebenen, benötigt Informationen zum Qualitäts-, Gesundheits-, Sicherheits- und Umweltmanagement. Bei diesen vier Bereichen handelt es sich um Managementsysteme. Während der Analyse der Präqualifikationsfragebögen zeigte sich, dass in Zukunft ebenfalls Angaben zu anderen Managementsystemen, wie z. B. Vorgehensweisen zum Risikomanagement, von den Kunden verlangt werden (vgl. Abschnitt 3.2.3.5). Diese Annahme resultiert daraus, dass teilweise Angaben zum Risikomanagement im Bereich Sicherheits- und Umweltmanagement erfragt wurden. Allerdings wird das Risikomanagement noch nicht als eigenständiges Managementsystem in den PQF erfasst. Um die Erweiterbarkeit in diesem Bereich zu gewährleisten, wird hier eine abstrakte Verwaltung der Managementsysteminformationen auf Basis der Analyse aus Abschnitt 3.2 konzipiert. Der Vorteil einer abstrakten MS-Verwaltung liegt in der Erfassung verschiedener Managementsysteme. Es ist dadurch nicht notwendig, für jedes Managementsystem eine eigene Komponente zu konzipieren, sodass zukünftige MS ebenfalls in die Datenbasis aufgenommen werden können. Ein Nachteil dieser Vorgehensweise ist der Verlust von spezifischen Detailinformationen zu den einzelnen Managementsystemen.

Im Folgenden Abschnitt werden die Funktionen der Managementsystemverwaltung dargestellt und beschrieben. Die Funktion *MS verwalten* gliedert sich objektorientiert, wie die Abb. 4.21 zeigt, in die Funktionen *MS anlegen*, *bearbeiten*, *suchen* und *löschen*. Prozessorientiert ordnen sich die Funktionen *Verantwortlichen festlegen*, *Dokument hinzufügen*, *Sicherungssystem anlegen* und *Vorgehensweise hinzufügen* der Funktion *MS anlegen* unter.

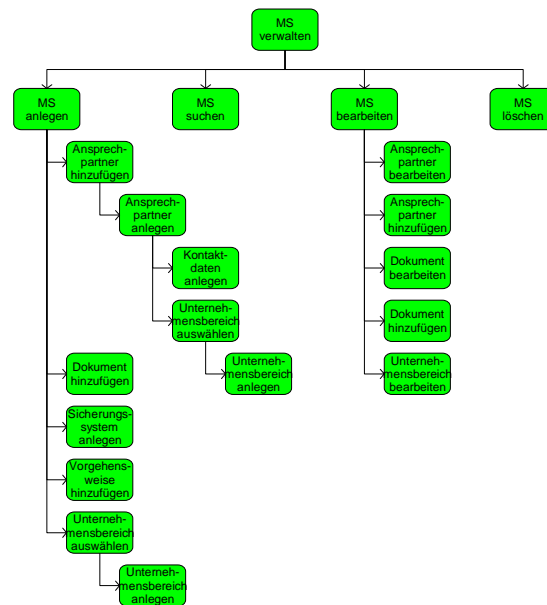


Abb. 4.21: Funktionsübersicht der MS-Verwaltung

Ein Managementsystem wird mit den Attributen Bezeichnung und Beschreibung angelegt (siehe Abb. B.6 auf der Seite 110 im Anhang). Zusätzlich wird ein Unternehmensbereich und ein verantwortlicher Ansprechpartner für das Managementsystem festgelegt (siehe Abb. 4.22). Sollte dieser Mitarbeiter noch nicht in der Datenbank vorhanden sein, wird er neu angelegt. Jedem Managementsystem können verschiedene Dokumente hinzugefügt werden. Diese Dokumente sind die MS-Zertifizierung und das Handbuch des Managementsystems. Jedem Managementsystem werden ein oder mehrere Zertifizierungsdokumente zugeordnet. Einem Managementsystem muss aber keine Zertifizierung zugeordnet werden, dies ist z. B. der Fall, wenn sich das Managementsystem im Aufbau befindet und noch nicht zertifiziert ist. In diesem Fall kann ein Managementsystem ohne Zertifizierung vorhanden sein. Wahlweise kann das Handbuch als Dokument vollständig oder teilweise aufgenommen werden, z. B. wird nur der Index des Handbuchs im Informationssystem als Dokument hinterlegt. Zu einem Managementsystem sollten noch weitere Dokumente, wie z. B. Reports, Programme und Protokolle, hinzugefügt werden.

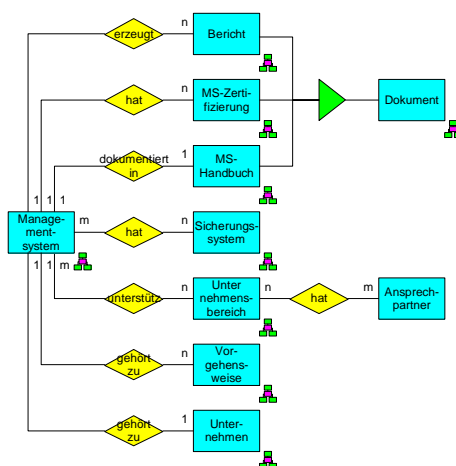


Abb. 4.22: Datenschema der MS-Verwaltung

Zu einem Managementsystem gehören Sicherungssysteme. Diese Anwendungssysteme werden mit Name, Hersteller und Art hinterlegt. Des Weiteren werden zu einem Managementsystem Vorgehensweisen gepflegt, diese werden mit den Attributen Bezeichnung, Unterbereich, Beschreibung und dokumentiert angelegt (siehe Abb. B.6 auf der Seite 110 im Anhang). Das dokumentiert-Attribut stellt den folgenden Sachverhalt dar. Wenn eine Vorgehensweise im Handbuch des jeweiligen Managementsystems aufgeschrieben wurde, ist diese dokumentiert, andernfalls ist die Vorgehensweise im Unternehmen existent aber nicht dokumentiert. Dadurch kann im PQF-Beantwortungsprozess auf eine Frage, ob eine Vorgehensweise dokumentiert ist, mit „ja“ oder „nein“ geantwortet werden.

Mit der Funktion *MS suchen* können die in der Datenbasis angelegten Managementsysteme über eine Suchmaske gefunden und ausgewählt werden. In den Informationen der Managementsysteme kann wiederum nach Vorgehensweisen gesucht werden, um die Informationen für die Beantwortung der PQF zu erhalten. Ist ein Managementsystem ausgewählt, können durch die Funktion *MS bearbeiten* Aktualisierungen an den Informationen vorgenommen werden.

Die vorgeschlagene Komponente zur Verwaltung von Managementsystemen ist rudimentär und auf die Bedürfnisse des PQ-Prozesses abgestimmt. Es wird kein Anspruch auf die vollständige Funktionalität eines implementierten MIS zur Unterstützung für bspw. Qualitätsmanagement erhoben. Allerdings sind hier auch unterstützende Funktionen, wie z. B. Erinnerungsfunktionen zur Erneuerung einer MS-Zertifizierung oder an Prüfungstermine, denkbar.

5 Zusammenfassung und Ausblick

Allgemein erfolgt eine Präqualifikation vor der eigentlichen Vergabe eines Projektes, um die Eignung eines Unternehmens festzustellen und das Risiko einer Fehlentscheidung zu verringern. Die Datenerfassung wird von den Auftraggebern über Präqualifikationsfragebögen realisiert. Diese werden den potenziellen Auftragnehmern zur Beantwortung zugesendet. Die Beantwortung der Präqualifikationsfragebögen ist zeitaufwendig und kostenintensiv. Außerdem können die Kosten der Beantwortung nicht direkt einem Projekt zugeordnet werden, weil die Auftragsvergabe erst nach der Präqualifikation erfolgt. Dabei kann es vorkommen, dass kein Auftrag an das Unternehmen vergeben wird. In diesem Fall können die Kosten nicht gedeckt werden. Deshalb ist es wichtig, die Beantwortung von Präqualifikationsfragebögen zu automatisieren, um so wenigstens die Kosten zu verringern und den Prozess zu beschleunigen.

Am Beispiel der Präqualifikation im Maschinen- und Anlagenbau wurde in dieser Arbeit eine Analyse des Beantwortungsprozesses und der, in der Praxis verwendeten, Präqualifikationsfragebögen durchgeführt (siehe Kapitel 3). Bei den Präqualifikationsfragebögen handelt es sich um Dokumente, mit denen Informationen über Unternehmen gesammelt werden. Außerdem gibt das anfragende Unternehmen Einblick in seine Entscheidungskriterien. Dabei stellen die gewünschten Informationen und deren logische Gruppierung die Struktur der Dokumente dar. Die Präqualifikationsfragebögen unterteilen sich in die Bereiche allgemeine Unternehmensdaten, Ansprechpartner, Mitarbeiter, Referenzen, Leistungsspektrum, Qualitätssicherung und Sicherheit/Gesundheit/Umwelt. Die darin enthaltenen Fragen stellen den Informationsbedarf des Beantwortungsprozesses dar.

Anhand dieser Analysen erfolgte im Kapitel 4 die fachkonzeptuelle Entwicklung eines Informationssystems zur Unterstützung der Beantwortung von Präqualifikationsfragebögen (siehe Abb. 5.1). Die Präqualifikationsfragebögen werden von den potenziellen Auftraggebern in papiergebundener oder digitaler Form an die Unternehmen übermittelt. Die papiergebundenen Dokumente stellen ein Problem bei der computerisierten Verarbeitung dar. Dieses Problem wurde im Fachkonzept durch eine vorgelagerte Digitalisierung unter Verwendung von Scanner-Hardware und OCR-Software gelöst. Dadurch werden NCI-Dokumente zu CI-Dokumenten transformiert und können somit einer automatischen Analyse unterzogen werden. Dabei wird in der Vorverarbeitung die Struktur der Dokumente analysiert, um Fragebogenbereiche und Fragen zu identifizieren. Danach werden die Fragen analysiert und mit linguistischen Informationen angereichert. Die linguistischen Informationen werden verwendet, um Rückschlüsse auf den Informationsbedarf der Frage zu ziehen.

Anhand der identifizierten Informationsbedarfe können die Fragen semi-automatisch beantwortet werden.

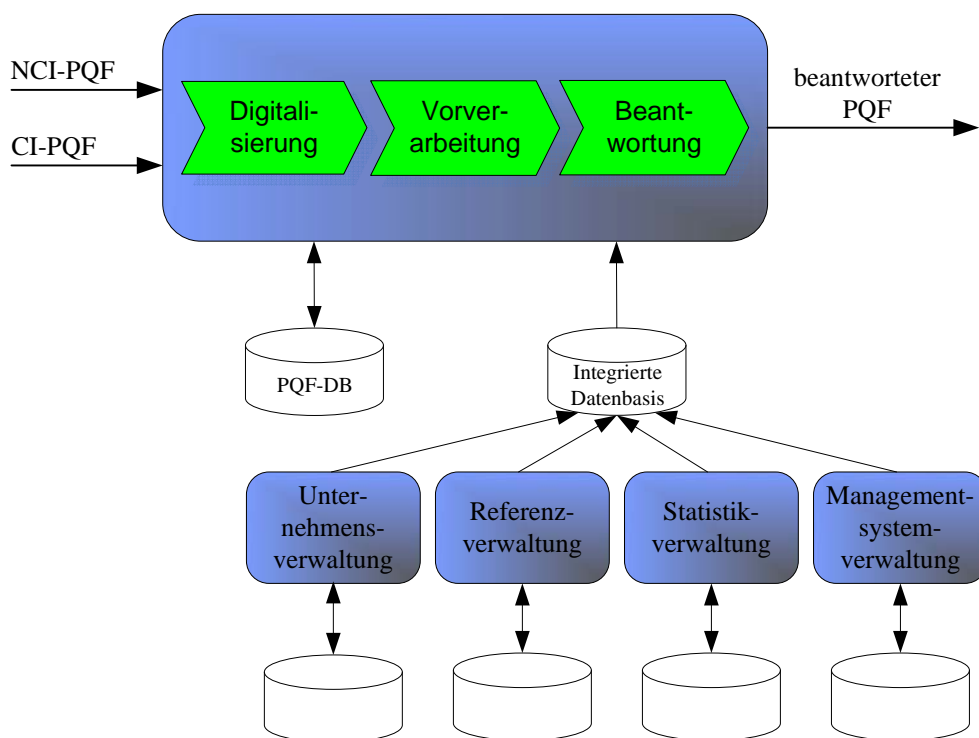


Abb. 5.1: Übersicht des Fachkonzepts

Für die Informationssuche stehen dem konzipierten Informationssystem die Datenbank der Präqualifikationsfragebögen und eine integrierte, globale Datensicht als Informationsangebot zur Verfügung. Durch Vergleiche mit vorhandenen Präqualifikationsfragebögen, deren Fragen und Antworten sowie der automatischen Informationsbedarfssuche, werden Antwortvorschläge generiert. Wenn keine Antworten gefunden werden, wurde ebenfalls eine manuelle Such- und Bearbeitungsfunktion sowie die Weiterleitung der Frage konzipiert. In den nächsten Schritten werden die Antworten übernommen und die benötigten digital vorliegenden Nachweis-Dokumente zusammengestellt. Danach kann der beantwortete Präqualifikationsfragebogen digital oder papiergebunden an den Auftraggeber zurückgesendet werden. Die integrierte Datenbasis entsteht virtuell durch ein föderiertes Datenschema, welches die lokalen Anwendungen einer heterogenen Systemlandschaft zusammenführt. Zu diesem Zweck wurden einerseits die benötigten Exportschemata und andererseits die Funktionalitäten der lokale Anwendungen zur Datenverwaltung konzipiert.

Während der Konzeption sind verschiedene Probleme hauptsächlich im Bereich der Computerlinguistik aufgetreten, die im Folgenden beschrieben werden. Dabei werden Aufgaben für die weitere Forschung aufgezeigt.

Ein Problem betrifft die syntaktische Analyse der Präqualifikationsfragebögen und deren Ablage-Format im Teilprozess Vorverarbeitung. Als Beispiel wird hier XML gewählt, da durch XML beliebige Strukturen in Texten, z. B. die Phrasenstruktur eines Satzes, markiert werden können. Das folgende Beispiel einer Dokumenttyp-Definition für die Phrasenstruktur ist aus Rehm (2004) entnommen und verdeutlicht die hier entstehende Problematik und Komplexität der syntaktischen Analyse.

```
<!ELEMENT s (np, vp)>          <!ELEMENT v (#PCDATA)>
<!ELEMENT np (det, n)>        <!ELEMENT n (#PCDATA)>
<!ELEMENT vp (v, np*)>        <!ELEMENT det (#PCDATA)>

<!ATTLIST np kasus (nom|akk|dat|gen) #REQUIRED>
<!ATTLIST v arg (1s|2s|3s|1p|2p|3p) #REQUIRED>
```

Diese DTD kann nur für einfache Sätze, die Artikel (det), Nomen (n) und Verben (v) enthalten, verwendet werden. Durch das Schlüsselwort ELEMENT werden in diesem Beispiel sechs Auszeichnungselemente und deren Inhaltsmodelle deklariert. Die deklarierten Elemente stellen im graphentheoretischen Sinne einen Baum mit dem Wurzelement s dar. Um die Verbindung zwischen Elementen herzustellen, werden in einer DTD verschiedene Konnektoren verwendet. Der Konnektor (,) erzwingt im Inhaltsmodell von s die Sequenz der Elemente np gefolgt von vp. Dies bedeutet, dass in einer Dokumenten-Instanz das Wurzelement s die beiden aufeinanderfolgenden Elemente np und vp enthält. Im Gegensatz dazu steht der Konnektor (|) für eine entweder-oder-Beziehung zwischen den beteiligten Elementen. Das Vorkommen der Elemente wird im Inhaltsmodell mit Hilfe von Okkurrenzindikatoren (*), (?) und (+) modelliert. Wobei kein Okkurrenzindikator hinter einem Element als einmaliges Vorkommen interpretiert wird. Dies ist der Fall im Inhaltsmodell von s. Das Symbol (*) deutet auf eine beliebige Häufigkeit (einschließlich Null) hin, wie im Inhaltsmodell vom Element vp das Element np*. Weitere Okkurrenzindikatoren sind das (?) für Häufigkeiten von (0 bis 1) und das (+) für Elemente mit der Häufigkeit (1 bis n), welche im Beispiel nicht verwendet werden. In der im Beispiel dargestellten DTD beinhalten die Elemente s, np und vp weitere Elemente und werden dadurch als Containerelemente bezeichnet. Im Gegensatz dazu handelt es sich bei den Elementen v, n und det um Datenelemente, welche mit dem Schlüsselwort #PCDATA modelliert werden. Diese werden zur Markierung arbiträrer Informationen in Texten genutzt. Zu jedem Element können Attribute bestimmt werden, die einen beschreibenden Charakter haben. In der dargestellten DTD werden Attribute durch das Schlüsselwort ATTLIST eingeleitet, gefolgt von dem Element, welches die Attribute erhalten soll. Im Beispiel wird zum Element np das Attribut kasus

deklariert. Hierbei handelt es sich um einen Aufzählungstyp, bei dem die erlaubten Werte (nom|akk|dat|gen) definiert werden. Das Schlüsselwort #REQUIRED bedeutet, dass diese Attribute obligatorisch sind. Ein Attribut ist optional, wenn das Schlüsselwort #IMPLIED benutzt wird.

Für den deutschsprachigen Raum gibt es keine Parser, die einen vollständigen Syntaxbaum für jeden Satz erstellen können (vgl. Evert und Fitschen (2004), S. 412). Die Anwendung der DTD auf den Satz „Ist ein Qualitätssicherungssystem vorhanden“, der in einem der Präqualifikationsfragebögen vorkommt, zeigt, dass die oben beschriebene DTD nicht ausreicht. Um den Satz darstellen zu können, muss die DTD erweitert werden. Bei diesem Satz handelt es sich um eine Frage und weist eine andere Struktur als eine Aussage auf. Fragen können bspw. mit einem Fragewort oder, wie in diesem Fall, mit einem Verb anfangen, gefolgt von einem Nomen und einem Adjektiv. Daraus ergeben sich die folgenden Änderungen an der oben beschriebenen DTD. Das Wurzelement s erfährt eine Modifikation seines Inhaltsmodells zu ((qw|v), np). Das Inhaltsmodell von np wird um ein Adjektiv zu (det, n, a) erweitert, um den Bezug des Adjektivs zum Substantiv nicht zu verlieren. Zusätzlich muss noch das Adjektiv durch <!ELEMENT a (#PCDATA)> als Datenelement der DTD hinzugefügt werden. Mit dieser modifizierten DTD lässt sich dann die folgende XML-Dokument-Instanz, mit der Markierung der Phrasenstruktur des Satzes „Ist ein Qualitätssicherungssystem vorhanden“, modellieren.

```
<?xml version="1.0" encoding="UTF-8">
<s>
  <v arg="3s">Ist</v>
  <np kasus="nom">
    <det>ein</det>
    <n>Qualitätssicherungssystem</n>
    <a>vorhanden</a>
  </np>
</s>
```

Die vorgenommenen Änderungen sind auf den gewählten Beispielsatz anwendbar. Eine Allgemeingültigkeit für die Annotation von Fragen wird in diesem Zusammenhang nicht erhoben. Vielmehr handelt es sich um ein Beispiel für die Komplexität der rein textuellen Aufbereitung der Präqualifikationsfragebögen. Darauf aufbauend kann die Analyse der allgemeinen Struktur, bspw. durch die Nomen der PQF, erfolgen, um die einzelnen Fragebogenbereiche und Fragenkomplexe zu identifizieren. Innerhalb der Bereiche und

Fragen können dann semantische und schließlich pragmatische Erkenntnisse für die automatische Beantwortung der Präqualifikationsfragebögen ermittelt werden.

Ein weiteres Problem für die semantische Analyse stellt die Verwendung unterschiedlicher Fachbegriffe und Formulierungen in den Präqualifikationsfragebögen dar. Syntaktisch verschiedene Fragen können semantisch denselben Informationsbedarf aufweisen. Beispiele für dieses Problem sind synonym verwendete Begriffe, wie Umsatz, Ertrag oder Erlös sowie für die Mitarbeiteranzahl das synonym gebrauchte Wort Personalanzahl oder semantisch äquivalente Schreibweisen, wie z. B. Anzahl der Mitarbeiter. Ein Lösungsansatz für die Auflösung semantischer Defekte natürlicher Sprache stellen die Konzepte: semantisches Wortnetz, Ontologie oder TopicMap dar. Zur Umsetzung eines dieser Konzepte ist es notwendig, genaue Kenntnisse über eine domänenspezifische Fachsprache zu ermitteln. Diese wurden auf Grund der Branchen- und Unternehmensspezifik in dieser Arbeit nicht berücksichtigt. Des Weiteren müssen die möglichen Darstellungsformen und Verfahren bewertet und in das Konzept zur Beantwortung von Präqualifikationsfragebögen integriert werden.

Im Rahmen der semantischen Analyse wird mit Hilfe von Schlüsselwörtern, Synonymlisten und Phrasen der semantische Inhalt der Bereiche und Fragen bestimmt. Hierbei sollte auch die Mehrsprachigkeit berücksichtigt werden. Durch die zunehmende Globalisierung werden häufiger Präqualifikationsfragebögen in englischer Sprache verfasst. Für das Erkennen einer anderen Sprache müssen die verwendeten Phrasen, Schlüsselwörter und die Grammatik für die jeweilige Sprache erfasst und implementiert werden. Danach kann ein Mapping mit dem Begriffskatalog erfolgen. Dadurch könnte das Informationssystem auch auf anderssprachige Präqualifikationsfragebögen reagieren. Die Antwortgenerierung muss dann ebenfalls an die andere Sprache angepasst werden.

Da die Präqualifikationsfragebögen in unterschiedlichen Formaten übertragen werden können, ist die Normalisierung sehr komplex, wenn es sich dabei um Dokumente handelt, die keine Trennung zwischen Struktur, Layout und Inhalt vornehmen. Dieses Problem könnte mit einem standardisierten Übertragungsformat, bspw. auf Basis der XML, gelöst werden. Mit dem Electronic Data Interchange (EDI) existiert ein elektronisches, strukturiertes und branchenübergreifendes Austauschformat für Geschäftsdaten (vgl. Götzer et al. (2004), S. 220). Allerdings muss dieses Austauschformat auf dessen Anwendbarkeit im Bereich der Präqualifikation und für die Übertragung von Präqualifikationsfragebögen geprüft werden.

Auf Grund dieser Probleme wird im Folgenden eine zweistufige Vorgehensweise zur Umsetzung vorgeschlagen. Die erste Stufe beinhaltet eine praktikable Lösung zur semi-automatischen Beantwortung der Präqualifikationsfragebögen. Dabei werden

die oben beschriebenen Probleme nicht vollständig gelöst. Ein Schwerpunkt bei der Entwicklung liegt bei der möglichen Integration von Analyse- und Beantwortungskomponenten. Das in dieser Arbeit entwickelte Fachkonzept wird nach der ARIS-Vorgehensweise im DV-Konzept verfeinert. In dieser Phase erfolgt die Festlegung auf konkrete Techniken und eine genauere Spezifikation des Informationssystems. Danach erfolgen die Phasen der Implementation und die Einführung bei den entsprechenden Unternehmen. Es werden die Digitalisierung, die Datengrundlage und die manuelle Suche sowie Teile der Vorverarbeitung und automatischen Suchfunktion in der PQF-Datenbank implementiert und das Rechte-Konzept sowie das Organigramm für das jeweilige Unternehmen erstellt. In Abhängigkeit der Systemlandschaft des jeweiligen Unternehmens werden für das föderierte Datenschema die benötigten Exportschemata bzw. die benötigten Teilsysteme implementiert. Mit dieser Lösung können die Präqualifikationsfragebögen schneller beantwortet werden. Zusätzlich wird dadurch während der Nutzung ein PQF-Korpus aufgebaut. Dieser Korpus bildet eine wichtige Grundlage für linguistische, statistische Untersuchungen. Diese Ergebnisse werden in der zweiten Stufe, der vollständigen Automatisierung, in Bezug auf die Trainierbarkeit von Taggern und die Anpassung der linguistischen Analysen auf die Präqualifikationsfragebögen umgesetzt. Die konzipierten Komponenten können dann mit dem Informationssystem der ersten Stufe integriert werden.

Das konzipierte Informationssystem kann ebenfalls zur Verwaltung, der durch ein eigenes PQ-Verfahren gesammelten Daten über potenzielle Subunternehmen dienen. Dafür wird erstens die PQF-Datenbank eingesetzt, um die Präqualifikationsfragebögen der Subunternehmen zu speichern. Zweitens kann das globale Schema verwendet werden, um eine Datenbank zu implementieren, welche die gesammelten Daten aus den gespeicherten Präqualifikationsfragebögen für die Partnerauswahl vorhält.

Das entwickelte Konzept ist prinzipiell für alle Branchen, bei denen eine Präqualifikation auf Basis von Präqualifikationsfragebögen erfolgt, einsetzbar. Damit ist das Ziel dieser Arbeit, die fachkonzeptuelle Entwicklung eines Informationssystems zur Unterstützung des Präqualifikationsprozesses von kmU am Beispiel des Maschinen- und Anlagenbaus, erreicht. Das zweistufige Ausbaukonzept ermöglicht die Umsetzung einer praktikablen Lösung und stellt weitere Aufgaben an die Forschung bis zur vollständigen Automatisierung des Beantwortungsprozesses.

Anhang

A Tabellen der Informationsbedarfsanalyse

Tab. A.1: Bereiche der Präqualifikationsfragebögen

Bereich	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Allgemeine Unternehmensdaten	x	x	x	x	x	x	x	x	x	x	x	x	x			x		82,4%	A
Ansprechpartner	x		x	x	x	x	x	x	x		x	x	x		x	x	x	82,4%	A
Referenzen	x	x	x	x	x	x	x	x	x	x	x			x		x	x	82,4%	A
Qualitätssicherung	x				x	x	x	x	x	x	x	x	x	x		x	x	76,5%	A
Mitarbeiter	x	x			x	x	x	x	x	x		x		x		x		64,7%	B
Leistungsspektrum	x	x				x	x	x	x		x	x					x	52,9%	B
Sicherheit/Gesundheit/Umwelt	x						x		x	x	x	x	x			x		52,9%	B
Versicherungen	x					x				x		x	x	x				35,3%	C
Kommerzielle Aspekte	x									x				x		x	x	29,4%	C
Logistik und Auftragsabwicklung									x	x	x					x		23,5%	C
IT - elektronische Bearbeitung und Kommunikation						x					x	x				x		23,5%	C
Gebäude und Einrichtungen							x		x	x								17,6%	C
Zulieferer/Subunternehmer						x						x	x					17,6%	C
Handelsregistereintragungen	x					x												11,8%	C
Garantie													x			x		11,8%	C
Firmenstrategie																x		5,9%	C
Marktpositionierung																x		5,9%	C
Herstellung/Montage																x		5,9%	C
Entwicklung																x		5,9%	C
Management, Planung, Beschaffung und Controlling													x					5,9%	C
Angebotsaktivitäten													x					5,9%	C

Tab. A.2: Allgemeine Unternehmensdaten

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Firmenname	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	100,0%	A
Anschrift	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	100,0%	A
Telefon-Nr.	x	x	x	x	x		x	x	x	x	x			x	x	x	x	88,2%	A
Telex-Nr.		x						x								x	x	23,5%	C
Fax-Nr.	x	x	x	x	x		x	x	x	x	x	x	x	x	x		x	88,2%	A
E-Mail-Adresse	x		x				x		x	x	x	x	x	x	x	x	x	70,6%	A
Internetadresse	x		x						x	x	x	x	x		x	x		52,9%	B
Besitzer		x	x	x	x		x	x	x	x		x	x	x	x	x		76,5%	A
Rechtsform	x	x	x			x		x		x	x	x	x	x		x		64,7%	B
Organisation (Organigramm)		x			x	x			x	x	x	x	x	x		x		58,8%	B
Gründungsjahr der Firma		x	x		x	x	x	x			x			x		x		52,9%	B
Konzernzugehörigkeit										x	x	x	x	x		x	x	41,2%	C
Tochtergesellschaften										x	x	x	x	x		x	x	35,3%	C
Dun&Bradstreet (DUNS) Nr.										x	x			x		x		23,5%	C
Firmenstandorte											x	x		x		x		23,5%	C
Strukturänderungen												x	x	x				17,6%	C
Firmenerweiterungen		x										x						11,8%	C
Gründungsort						x										x		11,8%	C
Jahre unter jetzigem Namen														x		x		11,8%	C
Aktionäre													x				x	11,8%	C
Firmenbeschreibung												x		x				11,8%	C
SAP Kreditoren-Nr.											x							5,9%	C
Gründungsgrund																x		5,9%	C
Division, Sparte, Abteilung																x		5,9%	C
Generalunternehmen												x						5,9%	C
Subunternehmen												x						5,9%	C
Vertragsarten als Subunternehmen														x				5,9%	C

Tab. A.4: Informationsbedarf des Bereichs Referenzen

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Abgeschlossene Projekte	x	x	x		x	x	x	x		x		x	x	x			x	70,6%	A
Geschäftsverbindungen	x		x	x		x				x	x	x		x		x	x	58,8%	B
Max. Auftragsgröße (3 Jahre)			x			x	x							x				23,5%	C
Laufende Projekte														x				5,9%	C
Min. Auftragsgröße (3 Jahre)														x				5,9%	C
Gesamtanzahl abgeschlossener Verträge (3 Jahre)														x				5,9%	C

Tab. A.5: Informationsbedarf des Bereichs Leistungsspektrum

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Kernarbeitsgebiet	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	100,0%	A
Arbeitsgebiet der Subunternehmen	x	x	x							x	x	x	x	x			x	52,9%	B
Planungsleistung	x		x			x	x	x		x								35,3%	C
Arbeitsvorbereitung	x					x		x				x	x					29,4%	C
Materialleistung			x			x	x	x										23,5%	C
Planungsart/-system	x										x	x	x					23,5%	C
fachspezifische Zulassungen bzw. Fachbetriebsnachweise	x						x					x						17,6%	C
Anteil Eigenfertigung/ Fremdbezüge																x	x	11,8%	C
Lohnleistung			x			x												11,8%	C
Prozentsatz der Wertschöpfung										x								5,9%	C
F&E Aktivitäten der Produkte										x								5,9%	C
Schulungen für Stammkunden																x		5,9%	C

Tab. A.6: Informationsbedarf des Bereichs Qualitätssicherung

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
QM-Zertifizierung	x	x			x	x	x	x	x	x	x	x	x	x		x	x	82,4%	A
Vorgehensweisen	x	x		x	x	x		x		x		x	x	x	x	x	x	76,5%	A
QM-System	x	x		x	x	x	x					x	x	x		x	x	64,7%	B
QM-Handbuch		x			x		x	x	x			x	x	x		x	x	58,8%	B
QM-Beauftragter/- Ansprechpartner	x	x			x				x	x		x		x	x		x	52,9%	B
Überwachungsvertrag	x							x				x					x	23,5%	C
Politik/Programm										x		x		x				17,6%	C
Eigenständige Organisationseinheit für Qualitätssicherung												x		x		x		17,6%	C
Notfallpläne									x			x				x		17,6%	C
Philosophie												x		x				11,8%	C
Geplante Einführung										x						x		11,8%	C
Qualitätstechniken									x			x						11,8%	C
qualitätsbezogene Dokumentation		x														x		11,8%	C
CE-Konformität																x		5,9%	C
Qualitätssicherungsvereinbarung																x		5,9%	C
Spezielle Programme/Ziele zur Steigerung der Qualität dokumentiert																x		5,9%	C
Prüfungsintervall des Systems extern										x								5,9%	C
Prüfungsintervall des Systems interne Revision der Prozesse										x								5,9%	C
Laboreinrichtungen									x									5,9%	C
Auszeichnungen und Zertifikate von Kunden																x		5,9%	C
Auszeichnungen von Zertifizierungsstellen																x		5,9%	C
Bestätigung: Qualitätsstandard, so dass der Kunde keine weitere Prüfung durchführen muss.																x		5,9%	C
Abweichungsrate bei anderen Projekten														x				5,9%	C
Beschwerdemanagement																x		5,9%	C
Schwierigste Qualitätsangelegenheit in der Vergangenheit														x				5,9%	C
Erklärung: Bereitschaft für Lieferantenaudits											x							5,9%	C

Tab. A.7: Informationsbedarf des Bereichs HSE

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
HSE-Zertifizierung	x			x		x	x		x	x	x	x	x	x		x		64,7%	B
Vorgehensweisen	x			x		x	x		x	x	x	x	x	x		x		64,7%	B
HSE-Handbuch	x			x		x	x		x			x	x	x		x		52,9%	B
Sicherheitsstatistik	x			x			x			x	x	x	x	x				47,1%	C
HSE-Beauftragter	x			x			x		x			x	x	x				41,2%	C
HSE-Richtlinie									x			x	x			x		23,5%	C
HSE-Jahresabschlussbericht										x		x	x			x		23,5%	C
Report-System HSE intern/extern									x	x		x				x		23,5%	C
Notfallpläne									x	x				x		x		23,5%	C
Politik/Programm										x		x		x				17,6%	C
Philosophie												x		x				11,8%	C
Bestätigung: Subunternehmen erfüllen Standards										x						x		11,8%	C
HSE-Erklärung												x	x					11,8%	C
Ziele/Programme für die Verbesserung der Leistung HSE									x							x		11,8%	C
UM-System									x							x		11,8%	C
Erfahrungsänderungsrate														x				5,9%	C
Bestätigung: Ausschluss bestimmter Materialien										x								5,9%	C
Bestätigung: bestimmtes Verpackungsmaterial										x								5,9%	C
HSE-Auszeichnung													x					5,9%	C
Vergabe persönlicher Schutzausrüstung												x						5,9%	C
Stellungnahmen der Geschäftsleitung zum Sicherheits- und Null-Unfallmanagement in Bezug auf das Projekt												x						5,9%	C

Tab. A.8: Informationsbedarf des Bereichs Ansprechpartner

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Technisch	x	x			x				x	x	x		x	x	x	x	x	64,7%	B
Kaufmännisch	x	x			x				x	x				x	x	x	x	52,9%	B
Qualitätsmanagement	x	x			x				x	x		x		x	x		x	52,9%	B
HSE	x			x			x		x			x	x	x				41,2%	C
Juristisch	x	x				x					x			x				29,4%	C
Führungskräfte			x					x			x		x					23,5%	C
Fertigung										x					x		x	17,6%	C
Produktion		x														x		11,8%	C
Terminplanung						x											x	11,8%	C
Schweißaufsicht															x		x	11,8%	C
Logistik																x		5,9%	C
Konstruktion										x								5,9%	C
Prüfaufsicht															x			5,9%	C
Montage															x			5,9%	C

Tab. A.9: Informationsbedarf des Bereichs Mitarbeiter

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Mitarbeiteranzahl gesamt	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	94,1%	A
Mitarbeiteranzahl gewerblich	x	x	x	x	x		x	x	x		x	x	x	x	x		x	82,4%	A
Kapazität		x			x	x		x	x	x	x		x	x		x		58,8%	B
Qualifikationen		x			x	x	x			x			x	x		x	x	52,9%	B
Anzahl der Leasingmitarbeiter	x	x				x					x		x	x				35,3%	C
Mitarbeiter von Subunternehmen	x	x									x		x					23,5%	C
Führungskräfte				x				x			x		x					23,5%	C
Aufsichtsdichte	x	x					x											17,6%	C
Schulungsprogramme und Ausbildung										x				x		x		17,6%	C
Gewerkschaftsverband										x		x						11,8%	C
Streikgeschichte										x		x						11,8%	C
Fluktuationsstatistik												x	x					11,8%	C
Durchschnittsalter										x								5,9%	C
Prozentualer Anteil der Mitarbeiter in Bereichen																x		5,9%	C
Mitarbeiteranzahl beim größten Auftrag der letzten 3 Jahre														x				5,9%	C
Höchste Mitarbeiteranzahl in einem Auftrag des letzten Jahres														x				5,9%	C
Anreizsysteme																x		5,9%	C

Tab. A.10: Informationsbedarf der Dokumente

Dokumente	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Zertifikate	x			x	x	x	x	x	x	x	x	x	x	x	x	x	x	88,2%	A
Fachbetriebsnachweise	x			x	x	x	x	x	x	x		x	x	x			x	64,7%	B
Reports		x			x	x	x		x	x		x	x	x		x		58,8%	B
Organigramme		x			x	x			x	x	x	x	x	x		x		58,8%	B
Versicherungszertifikate	x		x			x			x	x		x	x	x		x		52,9%	B
Prospekte und Produktkataloge								x	x	x		x				x		29,4%	C
Angestelltensicherheitsregelbroschüre												x		x				11,8%	C
List of Codes														x				5,9%	C
Sicherheitsplan												x						5,9%	C
Index QM-Handbuch												x						5,9%	C
QM-Politik und QM-Statement												x						5,9%	C
Qualitätsplan												x						5,9%	C
Beispiel Projektplan												x						5,9%	C
Unternehmenskonten mit der Unterschrift eines Wirtschaftsprüfers												x						5,9%	C
Sicherheitshandbuch und Verfahren												x						5,9%	C
Lebenslauf des Führungspersonals												x						5,9%	C
Firmenbuchauszug											x							5,9%	C
Referenzliste abgeschlossener Projekte														x				5,9%	C
Referenzliste laufender Projekte														x				5,9%	C
Stückliste mit Baugeräten und Ausrüstung														x				5,9%	C
Bonusvereinbarungen																	x	5,9%	C
Rabattvereinbarung																	x	5,9%	C
Typische Reparaturdokumente											x							5,9%	C

Tab. A.11: Informationsbedarf des Bereichs kommerzielle Aspekte

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Jahresabschlussbericht		x			x	x	x			x		x	x	x		x		52,9%	B
Bankverbindung						x	x	x				x	x	x		x	x	47,1%	C
Jahresumsatz									x	x	x	x	x	x		x		41,2%	C
Durchschnitt Jahresumsatz über 3 Jahre	x	x	x		x			x	x		x							41,2%	C
Auftragsreichweite über 1 Jahr		x							x		x	x				x		29,4%	C
Grundkapital		x				x						x						17,6%	C
Rechtsstreitigkeiten (über 5 Jahre)												x	x	x				17,6%	C
Unbeglichene Änderungsaufträge (Nachtragsforderungen)												x	x	x				17,6%	C
Geschäftsfelder														x		x		11,8%	C
Marktanteil										x						x		11,8%	C
Versäumnis Arbeit zu beenden												x		x				11,8%	C
Finanzamt	x																	5,9%	C
Steuernummer	x																	5,9%	C
Umsatzsteuer-Ident-Nr.																	x	5,9%	C
Freistellungsbescheinigung (§48 b EStG)	x																	5,9%	C
Gewinn und Verlustrechnung																x		5,9%	C
Durchschnitt Verkauf über 3 Jahre										x								5,9%	C
Durchschnitt Verlust über 3 Jahre										x								5,9%	C
Durchschnitt Gewinn über 3 Jahre										x								5,9%	C
Return on Investment																x		5,9%	C
Durchschnitt Jahresumsatz ohne Subunternehmen	x																	5,9%	C
Investitionsvolumen																x		5,9%	C
Entwicklungskosten																x		5,9%	C
Eigenkapital und Vermögenswerte																x		5,9%	C
Anlagevermögen													x					5,9%	C
Barvermögen													x					5,9%	C
Eigenkapitalquote													x					5,9%	C
Gewinn vor Steuer													x					5,9%	C
Nettobetriebsergebnis													x					5,9%	C

Tab. A.12: Informationsbedarf des Bereichs Versicherungen

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Versicherungszertifikat	x		x			x			x	x		x	x	x		x		52,9%	B
Berufsgenossenschaft	x		x	x		x												23,5%	C

Tab. A.13: Informationsbedarf des Bereichs Logistik

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
System für Liefertreue an Kunden									x	x						x		17,6%	C
Lieferbedingungen									x							x		11,8%	C
Preislisten/ Embargos eigener Produkte									x							x		11,8%	C
Produkte (gilt für Vertreiber und Hersteller)									x			x						11,8%	C
Managementinstrumente zur Auftragsabwicklung (Auftragsverfolgung, Risikobewertung, Abwicklung)										x	x							11,8%	C
Logistikkonzept																x		5,9%	C
Logistikvereinbarungen Ship to Stok mit Lieferanten																x		5,9%	C
Lagerhaltung FIFO																x		5,9%	C
Lieferzeiten																x		5,9%	C
Lieferantenerklärung (Langfristig/einmalig) Stellungnahme über Einhaltung von Zollbestimmungen und Richtlinien der EU wenn der Kunde es wünscht (Bereitschaft)																x		5,9%	C
Organisationseinheit Logistik									x									5,9%	C
Feinplanung									x									5,9%	C
System zur Sicherung des Versands									x									5,9%	C
Prozentsatz der pünktlichen Lieferungen										x								5,9%	C
Häufigkeit des Überrollen der Produktionsplanung									x									5,9%	C
Entscheidungskompetenzen der Ansprechpartner									x									5,9%	C

Tab. A.14: Informationsbedarf des Bereichs Garantie

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Bankgarantie													x			x		11,8%	C
Konzernbürgschaft													x					5,9%	C
Garantiekondition																x		5,9%	C
Garantiperiode																x		5,9%	C

Tab. A.15: Informationsbedarf des Bereichs Informationstechnologie

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Software für Materialsteuerung und Beschaffung						x			x	x		x	x		x			35,3%	C
Elek. Planungssysteme, Termin und Projektsteuerung						x			x	x	x		x					29,4%	C
Aufmaß- und Abrechnungssysteme	x					x		x			x							23,5%	C
Beschreibung: Hardware und Netzwerkumgebung innerhalb der Herstellungsanlage						x			x	x			x					23,5%	C
Dokumentationssysteme (für gestellte Arbeit)						x							x			x		17,6%	C
Büro-Standardsoftware						x							x					11,8%	C
Formate für Datenbankaustausch und Downloadfähigkeiten											x		x					11,8%	C
Systeme und Schnittstellen für Datenübertragung in ein projektbezogenes System zur Fertigstellung													x			x		11,8%	C
Dokumenten-Managementsysteme								x					x					11,8%	C
E-Business/Internetkommunikation											x					x		11,8%	C
System zur Datenarchivierung und Schnittstellen zu anderen bekannten Systemen													x					5,9%	C
Erfahrung mit dem Austausch von Dokumenten in elektronischer Form													x					5,9%	C
Datentransfer durch EDI																x		5,9%	C
Erfahrung mit speziellen Systemen													x					5,9%	C
Erklärung bestimmte Systeme einzuführen													x					5,9%	C

Tab. A.16: Informationsbedarf des Bereichs Gebäude

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Maschinen, Baugeräte, Werkzeuge		x							x		x	x	x	x				35,3%	C
Produktionskapazität									x	x	x	x	x			x		35,3%	C
Fertigungsfläche gesamt									x	x		x	x	x				29,4%	C
Eigentumsverhältnis									x		x	x		x		x		23,5%	C
Vorgehensweise bei Kapazitätsengpässen										x		x	x			x		23,5%	C
Materialien						x		x				x				x		23,5%	C
Fabrikgröße												x		x				11,8%	C
Verwendete Kapazität										x		x						11,8%	C
geplante Anschaffungen und Ausbau										x		x						11,8%	C
Bürogröße														x				5,9%	C
Einrichtungen außer Haus												x						5,9%	C
Einkaufspolitik																x		5,9%	C
Einkaufsvolumen im Verhältnis zum Verkauf																x		5,9%	C
Logistikinfrastruktur														x				5,9%	C

Tab. A.17: Informationsbedarf des Bereichs Subunternehmen

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Firmenname						x						x	x					17,6%	C
Vertragsart						x				x		x						17,6%	C
Zertifizierung						x							x					11,8%	C
Arbeitsgebiet													x					5,9%	C
Beschaffungszeit und Liefervolumen													x					5,9%	C

Tab. A.18: Informationsbedarf des Bereichs Handelsregister

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Eintragsnummer	x		x			x												17,6%	C
Registergericht	x					x												11,8%	C
Geschäftslizenzen														x				5,9%	C
Handwerksrolle			x															5,9%	C

Tab. A.19: Informationsbedarf des Bereichs Firmenstrategie

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Vision							x									x		11,8%	C
Strategische Mitgliedschaften (Verbände)						x										x		11,8%	C
Strategische Verträge																x		5,9%	C
Strategische Partner																x		5,9%	C

Tab. A.20: Informationsbedarf des Bereichs Marktpositionierung

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Kerngeschäftsfelder																x		5,9%	C
Geplante Kernsegmente																x		5,9%	C
Globale Strategie																x		5,9%	C
Marktanteile und Wachstumsrate (Weltmarkt und Kerngeschäftsfelder)																x		5,9%	C
Verteilung des Verkaufs der letzten Jahre in bestimmten Regionen																x		5,9%	C
Hauptkonkurrenten für Kerngeschäftsfelder																x		5,9%	C
Marktführerschaft in Marken/Produkten																x		5,9%	C
Key Account Management																x		5,9%	C
Warum wollen Sie uns als Kunde?																x		5,9%	C
Warum glauben Sie, sollten wir mit Ihnen arbeiten?																x		5,9%	C

Tab. A.21: Informationsbedarf des Bereichs Herstellung

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Shutdown der Produktion zur Wartung, an Feiertagen oder Ferien																x		5,9%	C
Welcher Fertigungsstandort ist der Hauptinteressent?																x		5,9%	C
Prozentsatz der Eigenfertigung von Produkten																x		5,9%	C
Prozentsatz der (eigenen) Montage von Produkten																x		5,9%	C
Verantwortliche Organisationseinheit für Herstellung und Montage																x		5,9%	C
Hauptkriterien der Produktkosten																x		5,9%	C

Tab. A.22: Informationsbedarf des Bereichs Entwicklung

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Prozentsatz der Eigenentwicklung von Produkten																x		5,9%	C
Prozentsatz der Eigenentwicklung von Produktionsprozessen																x		5,9%	C
Hauptforschungsgebiete																x		5,9%	C
Entwicklungszyklus neuer Produkte																x		5,9%	C

Tab. A.23: Informationsbedarf des Bereichs Management: Planung, Beschaffung und Controlling

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Vorgehensweisen (Management- und Überwachungsprozeduren)												x	x					11,8%	C

Tab. A.24: Informationsbedarf des Bereichs Angebotsaktivität

Informationsbedarf	PQF 01	PQF 02	PQF 03	PQF 04	PQF 06	PQF 07	PQF 08	PQF 09	PQF 10	PQF 12	PQF 13	PQF 14	PQF 15	PQF 16	PQF 17	PQF 18	PQF 19	rel. Häufigkeit	Kategorie
Projekte, für die geboten wird													x					5,9%	C

B Abbildungen

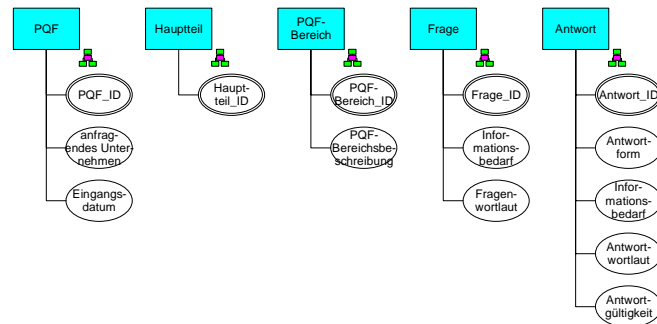


Abb. B.1: Attributzuordnungsdiagramme der Präqualifikationsfragebögen

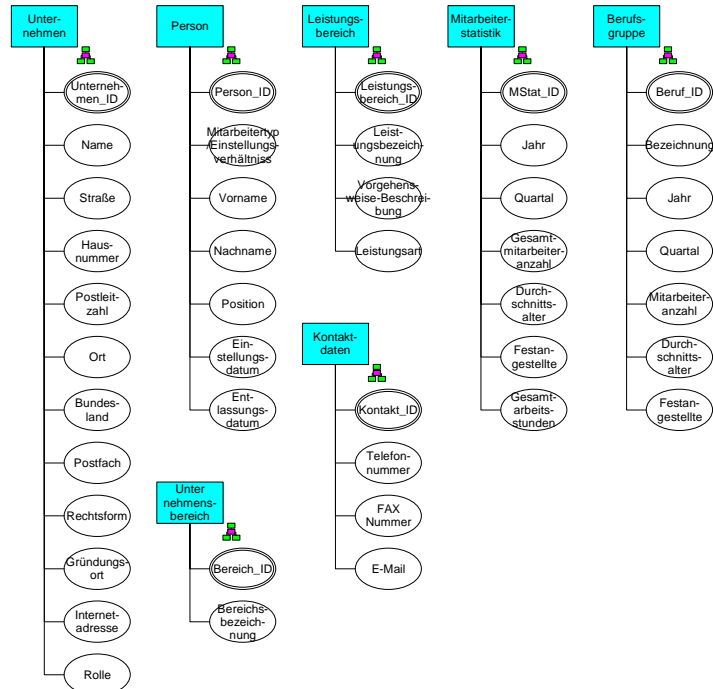


Abb. B.2: Attributzuordnungsdiagramme der Unternehmensverwaltung

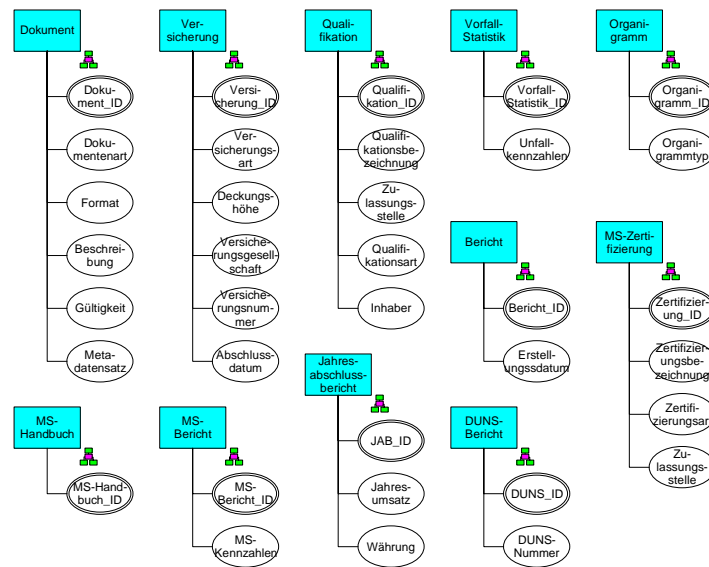


Abb. B.3: Attributzuordnungsdiagramme der Dokumente

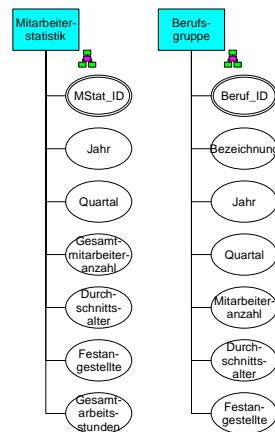


Abb. B.4: Attributzuordnungsdiagramme der Statistikverwaltung

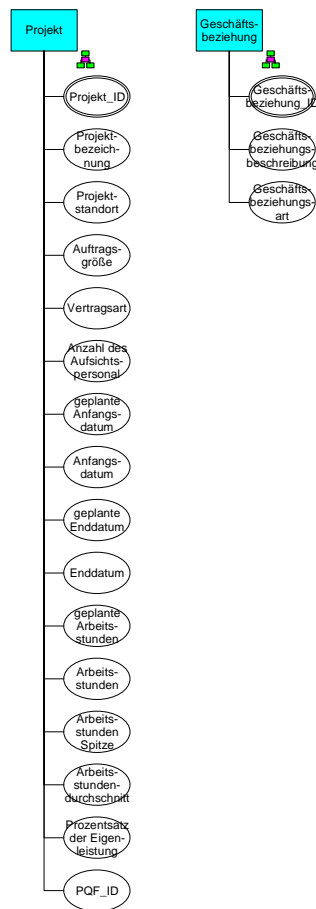


Abb. B.5: Attributzuordnungsdiagramme der Referenzverwaltung

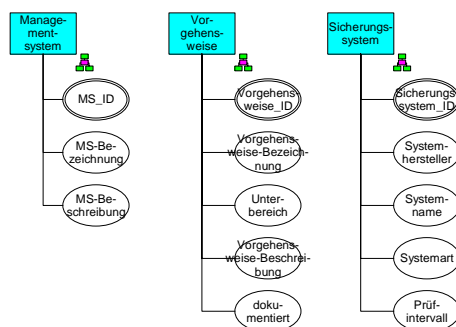


Abb. B.6: Attributzuordnungsdiagramme der MS-Verwaltung

Literaturverzeichnis

- Alex, B. (1998): Künstliche neuronale Netze in Management-Informationssystemen: Grundlagen und Einsatzmöglichkeiten. Gabler, Wiesbaden.
- Alpar, P., Grob, H. L., Weimann, P. und Winter, R. (2005): Anwendungsorientierte Wirtschaftsinformatik: Strategische Planung, Entwicklung und Nutzung von Informations- und Kommunikationssystemen. Vieweg, Wiesbaden, 4. Auflage.
- Amtrup, J. W. (2004): *Aspekte der Computerlinguistik*. In K.-U. Carstensen, C. Ebert, C. Endriss, S. J. Jekat, R. Klabunde und H. Langer (Hrsg.), Computerlinguistik und Sprachtechnologie: Eine Einführung, Kapitel 1.1, S. 01–15. Spektrum, München, 2. Auflage.
- Arndt, H.-K. (1997): Betriebliche Umweltinformationssysteme: Gestaltung und Implementierung eines BUIS-Kernsystems. Gabler, Wiesbaden.
- Arndt, H.-K. (2006): *Managementinformationssysteme: Informations- und kommunikationstechnische Unterstützung von Managementsystemen*. Zeitschrift für wirtschaftlichen Fabrikbetrieb (ZWF), Band 101 (2006), Nr. 6, S. 332–337.
- Beiersdorf, H. (1995): Informationsbedarf und Informationsbedarfsermittlung im Problemlösungsprozeß „Strategische Unternehmensplanung“. Rainer Hampp, München [u. a.].
- BMVBS (Hrsg.) (2006a): Leitlinie des Bundesministeriums für Verkehr, Bau und Stadtentwicklung für die Durchführung eines Präqualifizierungsverfahrens vom 25. April 2005 - in der Fassung vom 30.11.2006. BMVBS - Bundesministeriums für Verkehr, Bau und Stadtentwicklung, Berlin.
- BMVBS (Hrsg.) (2006b): Vergabe- und Vertragsordnung für Bauleistungen (VBO) Teil A: Allgemeine Bestimmungen für die Vergabe von Bauleistungen - in der Fassung vom 20.03.2006. BMVBS - Bundesministeriums für Verkehr, Bau und Stadtentwicklung, Berlin.
- Bodendorf, F. (2006): Daten- und Wissensmanagement. Springer, Berlin [u. a.], 2. Auflage.
- Boles, C., Friebe, J. und Luhmann, T. (2004): *Typische Integrationsszenarien und deren Unterstützung durch Web Services und andere Technologien*. In W. Hasselbring und M. Reichert (Hrsg.), EAI-Workshop 2004 - Enterprise Application Integration, S. 57–67. Gesellschaft für Informatik, Gito, Berlin.

- Chen, P. P. S. (1991): *Der Entity-Relationship-Ansatz zum logischen Datenbankentwurf*. In P. P. S. Chen und H.-D. Knöll (Hrsg.), *Der Entity-Relationship-Ansatz zum logischen Systementwurf*, S. 15–109. BI-Wiss.-Verl., Mannheim [u. a.].
- Conrad, S., Hasselbring, W., Koschel, A. und Tritsch, R. (2006): *Enterprise Application Integration: Grundlagen, Konzepte, Entwurfsmuster, Praxisbeispiele*. Spektrum, München, 1. Auflage.
- Elm, W. A. (1972): *Das Management-Informationssystem als Mittel der Unternehmensführung*. de Gruyter, Berlin [u. a.].
- Evert, S. und Fitschen, A. (2004): *Textkorpora*. In K.-U. Carstensen, C. Ebert, C. Endriss, S. J. Jekat, R. Klabunde und H. Langer (Hrsg.), *Computerlinguistik und Sprachtechnologie: Eine Einführung*, Kapitel 4.1, S. 406–413. Spektrum, München, 2. Auflage.
- Granitzer, M. (2006): *Knowminer: Konzeption und Entwicklung eines generischen Wissenserschliessungsframeworks*. Dissertation, Universität Graz.
- Götzer, K., Schneiderath, U., Maier, B. und Komke, T. (2004): *Dokumenten-Management: Informationen im Unternehmen effizient nutzen*. dpunkt, Heidelberg, 3. Auflage.
- Halama, A. (2004): *Flache Satzverarbeitung*. In K.-U. Carstensen, C. Ebert, C. Endriss, S. J. Jekat, R. Klabunde und H. Langer (Hrsg.), *Computerlinguistik und Sprachtechnologie: Eine Einführung*, Kapitel 3.3, S. 218–231. Spektrum, München, 2. Auflage.
- Hanning, U. (1998): *Einsatz von Managementinformationssystemen in Marketing und Vertrieb*. In U. Hanning (Hrsg.), *Managementinformationssysteme in Marketing und Vertrieb*. Schäffer-Poeschel, Stuttgart.
- Hansen, H. R. und Neuman, G. (2001): *Wirtschaftsinformatik I: Grundlagen betrieblicher Informationsverarbeitung*. Lucius und Lucius, Stuttgart, 8. Auflage.
- Heuer, A. und Saake, G. (2000): *Datenbanken: Konzepte und Sprachen*. mitp, Bonn, 2. Auflage.
- Holten, R. und Knackstedt, R. (1997): *Führungsinformationssysteme - Historische Entwicklung und Konzeption*. Technischer Bericht 55, Institut für Wirtschaftsinformatik der Westfälischen Wilhelms-Universität Münster.
- Jung, R. (2005): *Anforderungen an Informationsobjekttypen als Basis von Architekturentscheidungen bei der Datenintegration*. In R. Lenz, U. Hasenkamp,

- W. Hasselbring und M. Reichert (Hrsg.), EAI-Workshop 2005 - Enterprise Application Integration, S. 82–89. Gesellschaft für Informatik, Gito, Berlin.
- Keller, W. (2002): Enterprise Application Integration: Erfahrungen aus der Praxis. dpunkt, Heidelberg.
- Kleinbauer, M., Thurow, M. und Urbansky, A. (2006): *Klassifikation von Projekten für den verfahrenstechnischen Anlagenbau*. In M. Schenk (Hrsg.), Industriearbeitskreis "Kooperation im Anlagenbau": 3. Arbeitsbericht, S. 73–119. Fraunhofer IRB Verlag, Stuttgart.
- Klingelhöller, H. (2001): Dokumentenmanagementsysteme: Handbuch zur Einführung. Springer, Berlin [u. a.].
- Knackstedt, R. (2006): Fachkonzeptionelle Referenzmodellierung einer Managementunterstützung mit quantitativen und qualitativen Daten: Methodische Konzepte zur Konstruktion und Anwendung. Logos, Berlin.
- Limper, W. (2001): Dokumenten-Management: Wissen, Informationen und Medien digital verwalten. Dt. Taschenbuch-Verl. [u.a.], Orig.-Ausg. - München.
- Lobin, H. und Lemnitzer, L. (2004): *Text(e) technologisch*. In H. Lobin und L. Lemnitzer (Hrsg.), Texttechnologie: Anwendungen und Perspektiven, S. 1–9. Stauffenburg, Tübingen.
- Lorenz, S. (2006): Konzeption und prototypische Realisierung einer begriffsbasierten Texterschließung. Dissertation, Universität Trier.
- Ludewig, J. und Lichter, H. (2007): Software Engineering: Grundlagen, Menschen, Prozesse, Techniken. dpunkt, Heidelberg, 1. Auflage.
- Mertens, P. (2004): Integrierte Informationsverarbeitung 1: Operative Systeme in der Industrie. Gabler, Wiesbaden, 14. Auflage.
- Neumann, G. (2004): *Informationsextraktion*. In K.-U. Carstensen, C. Ebert, C. Endriss, S. J. Jekat, R. Klabunde und H. Langer (Hrsg.), Computerlinguistik und Sprachtechnologie: Eine Einführung, Kapitel 5.5, S. 502–510. Spektrum, München, 2. Auflage.
- Oppelt, R. U. G. (1995): Computerunterstützung für das Management: Neue Möglichkeiten der computerbasierten Informationsunterstützung oberster Führungskräfte auf dem Weg von MIS zu EIS? Oldenbourg, München [u.a.].

- Picot, A. und Maier, M. (1992): *Computergestützte Informationssysteme*. In E. Frese (Hrsg.), Handwörterbuch der Organisation, Band 2., S. 923–936. Poeschel, Stuttgart, 3. Auflage.
- Rautenstrauch, C. und Schulze, T. (2003): *Informatik für Wirtschaftswissenschaftler und Wirtschaftsinformatiker*. Springer, Heidelberg.
- Rechkemmer, K. (1999): *Topmanagement-Informationssysteme: Betriebswirtschaftliche Grundlagen*. Lucius und Lucius, Stuttgart.
- Rehm, G. (2004): *Texttechnologische Grundlagen*. In K.-U. Carstensen, C. Ebert, C. Endriss, S. J. Jekat, R. Klabunde und H. Langer (Hrsg.), *Computerlinguistik und Sprachtechnologie: Eine Einführung*, Kapitel 2.5, S. 138–147. Spektrum, München, 2. Auflage.
- Reitz, D., Schmietendorf, A., Dumke, R., Dimitrov, E., Lezius, J. und Schlosser, T. (2003): *Aspekte des empirischen Software Engineering im Umfeld von Enterprise Application Integration Lösungen*. Technischer Bericht Preprint Nr.5, Fakultät für Informatik, Universität Magdeburg.
- Scheer, A.-W. (1992): *Computergestützte Integration*. In E. Frese (Hrsg.), Handwörterbuch der Organisation, Band 2., S. 1042–1051. Poeschel, Stuttgart, 3. Auflage.
- Scheer, A.-W. (1998): *ARIS - Modellierungsmethoden, Metamodelle, Anwendungen*. Springer, Berlin [u. a.], 3. Auflage.
- Scheer, A.-W. (2002): *ARIS - Vom Geschäftsprozess zum Anwendungssystem*. Springer, Berlin [u. a.], 4. Auflage.
- Staehele, W. H. (1999): *Management: Eine verhaltenswissenschaftliche Perspektive*. Vahlen, München, 8. Auflage.
- Stahlknecht, P. und Hasenkamp, U. (1998): *Einführung in die Wirtschaftsinformatik*. Springer, Berlin [u. a.], 8. Auflage.
- Stickel, E. (2001): *Informationsmanagement*. Lehrbücher Wirtschaftsinformatik, Oldenbourg, München [u. a.].
- Struckmeier, H. (1997): *Gestaltung von Führungsinformationssystemen: Betriebswirtschaftliche Konzeption und Softwareanforderungen*. Gabler, Wiesbaden.
- Thurrow, M., Kleinbauer, M. und Urbansky, A. (2006): *Der Kooperationslebenszyklus im Anlagenbau*. In Industriearbeitskreis "Kooperation im Anlagenbau": 3. Arbeitsbericht, S. 63–72. Fraunhofer IRB Verlag, Stuttgart.

- Vetschera, R. (1995): Informationssysteme der Unternehmensführung. Springer, Berlin [u. a.].
- Winkeler, T., Raupach, E. und Westphal, L. (2000): *EAI - Enterprise Application Integration: Die Pflicht vor der E-Business-Kür.* In Leitfaden E-Business. Nummer 20, PriceWaterhouseCoopers GmbH.
- Young, M. J. (2001): XML - Schritt für Schritt. Microsoft Press, Unterschleißheim, 2. Auflage.

Abschließende Erklärung

Ich versichere hiermit, dass ich die vorliegende Diplomarbeit selbständig, ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Zernitz, den 31. Oktober 2007