Information Retrieval in Portalen

Gestaltungselemente, Praxisbeispiele und Methodenvorschlag

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

vorgelegt von

Stefan Kremer

aus
Deutschland

Genehmigt auf Antrag der Herren

Prof. Dr. Walter Brenner
und
Prof. Dr. Hubert Österle

Dissertation Nr. 2941

Difo-Druck GmbH, Bamberg 2004

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

St. Gallen, den 14. Juni 2004

Der Rektor:

Prof. Dr. Peter Gomez

Vorwort

Die vorliegende Arbeit ‚Information Retrieval in Portalen' ist das Ergebnis meiner mehr als dreijährigen Forschungstätigkeit in den Kompetenzzentren ‚Customer Knowledge Management' (CKM) und ‚Customer > Knowledge > Performance' (C>K>P) des Instituts für Wirtschaftsinformatik der Universität St. Gallen (IWI-HSG). Dabei profitierte ich vom kooperativen Ansatz des Forschungsprogramms ‚Business Engineering', der meine enge Zusammenarbeit mit anderen Wissenschaftlern förderte und mir die Möglichkeit bot, meine theoretische Überlegungen in der betrieblichen Praxis mit Forschungspartnern zu implementieren und zu reflektieren.

An dieser Stelle danke ich allen, die zum Gelingen dieser Arbeit beigetragen haben. Mein Dank gebührt zunächst Prof. Dr. Walter Brenner, geschäftsführender Direktor des Instituts, für die wissenschaftliche Betreuung und seine angenehme Art der Forderung und Förderung in einem praxisnahen Forschungsumfeld. Prof. Dr. Hubert Österle, Gründer des Instituts, danke ich herzlich für die Übernahme des Korreferats. Die Zusammenarbeit mit ihm hat die Ergebnisse meiner Arbeit stark geprägt.

Den Leitern der Kompetenzzentren CKM und C>K>P, Dr. Gerold Riempp und Dr. Lutz Kolbe, bin ich für die freundschaftliche Zusammenarbeit, ihre kritischen Fragen und hilfreichen Erfahrungen ebenso dankbar wie für das Vertrauen, mit dem sie mir ein sehr eigenständiges Arbeiten und die Verwirklichung meiner Ziele ermöglicht haben.

Stellvertretend für viele Forschungskollegen bedanke ich mich insbesondere bei meinen ‚Mitstreitern' Adrian Büren, Dr. Oliver Christ, Malte Dous, Malte Geib, Dr. Sandra Gronover, Oliver Kutsch, Enrico Senger, Harald Saloman, Ragnar Schierholz, Dr. Roland Schmid und Annette Reichold. Meinem Kollegen und Freund Dr. Henning Gebert danke ich ausserordentlich für die kurzweilige Zusammenarbeit in spannenden Projekten und erfolgreichen Workshops. Stefan Smolnik vom Groupware Competence Center der Universtität Paderborn verdanke ich interessante Erkenntnisse bei der Publikation unseres Hawaii-Artikels und die Einblicke, die wir durch die Evaluation seiner Topic-Map-Lösung an unserem Institut gewinnen konnten.

Den Geschäftsführern des Instituts, Dr. Ernst Ensslin und Dr. Dieter Zerndt, danke ich für ihre umfassende Hilfsbereitschaft in allen geschäftlichen aber auch privaten Angelegenheiten. Dem IWI Infrateam, Dani Seiler, Markus Handke und Michel Häusler, bin ich für ihre professionelle Unterstützung bei allen Hard- und Softwareproblemen zu jeder Tages- und Nachtzeit zu Dank verpflichtet.

Mein Dank gilt auch dem IWI Sekretariat, insbesondere Rita Bruderer und Barbara Rohner bei der Bewältigung des täglichen „Papier-, Formular- und Abrechnungskrieges" sowie der Bereitstellung von institutsnahen Parkmöglichkeiten, als familiäre Umstände häufige und schnelle Ortswechsel von mir gefordert haben.

Die studentischen Mitarbeiter Michael-Dominic Bauer, Oliver Metzner, Alexander Ostrowski, Christian Ruhse und Oliver Wüst unterstützen mich bei der Softwareevaluation, der Literaturrecherche, der Vorbereitung von Lehrveranstaltungen sowie der Organisation und Durchführung von Veranstaltungen des Instituts. Dafür danke ich ihnen ganz herzlich.

Für ihre tatkräftige Mitarbeit im institutsinternen Migrations- und Integrationsprojekt ‚KnowledgePort' danke ich Ralf Lutter, Martin Sander und Oliver Tunnat. Mit Sachverstand und Implementierungsexpertise haben sie in ihrer freiberuflichen Tätigkeit wesentlich zu der Qualität der heute im Einsatz befindlichen Wissensmanagementlösung beigetragen.

Praxisorientierte Forschung lebt von der Kooperation mit den Forschungspartnern aus der Wirtschaft. Mein besonderer Dank gilt zunächst Ingmar Blum und Martin Eberle, Zumtobel Staff AG, Walter Dyttrich und Tomasz Bugajski, Winterthur Versicherungen AG, sowie Martin Liebich, IMG AG und den jeweiligen Projektteams bei der konstruktiven Zusammenarbeit in den Information-Retrieval-Projekten und der Freigabe der daraus resultierenden Fallstudien. Gleichzeitig möchte ich mich bei Dr. Heinz-Gerd Kneip, Dr. Reinhard Aldag und Dr. Peter Kallas, BASF AG, und Klaus Zelmer, Bausparkasse Schwäbisch-Hall sowie den weiteren Mitarbeitern der jeweiligen Firmen bedanken, die ich im Rahmen meiner Kompetenzzentrumsarbeit als Key Account Manager begleiten und unterstützen konnte. Für interessante Diskussionen bei Workshops danke ich auch den Partnern der Kompetenzzentren CKM und C>K>P, insbesondere Pia Jaggi, Stefan Studer, Erik Krumdiek und Fredi Kuster, Helsana Versicherungen AG, Hans-Rudolf Häni, Dr. Marc Bider und Kuno Bürge, Credit Suisse AG, Martin Rothaut und Michael Heidt, Deutsche Telekom, Edwin Kölliker und Zsuzsanna Szalay, Swisscom IT Services sowie Thomas Pitz, Martin Werner und Sabine Vincze, Union Investment.

Während meiner freiberuflichen Tätigkeit für die IMG AG hat mir insbesondere die Zusammenarbeit mit Alexander Brochier, Dr. Alexander Lorani und Dr. Thomas Kaiser besondere Freude bereitet und hilfreiche Anregungen geliefert.

Schliesslich danke ich meinen Eltern, Rolf und Christa Kremer, ohne deren Unterstützung diese Arbeit nicht existieren würde. Meiner Frau Beatrix und unserer Tochter Katharina danke ich für ihre Begleitung und ihr Verständnis in allen Fällen, in denen wir der Erstellung dieser Arbeit Vorrang vor anderen Dingen geben mussten. Meiner Familie gebührt ein besonderer Dank, daher widme ich ihnen diese Arbeit.

St. Gallen, im Juni 2004 Stefan Kremer

Inhaltsübersicht

1. Einleitung 1

1.1. Ausgangslage und Handlungsbedarf 1

1.2. Entstehung und Einordnung der Arbeit 3

1.3. Ziele und Adressaten 4

1.4. Forschungsmethodik 5

1.5. Aufbau der Arbeit 7

2. Grundlagen 9

2.1. Business Engineering 10

2.2. Portale 15

2.3. Information Retrieval 19

2.4. Wissensmanagement 27

2.5. Konsequenzen für eine Information-Retrieval-Methode 33

3. Vergleich bestehender Methoden 34

3.1. Auswahlkriterien 35

3.2. Analyse 36

3.3. Übersicht der Bewertungen 51

3.4. Ableitung eines Information-Retrieval-Metamodells 52

4. Erfahrungen aus der Praxis 65

4.1. Zumtobel Staff – Customer Portal 66

4.2. The Information Management Group – Knowledge Pool 77

4.3. Winterthur Versicherungen – McB-Projektportal 87

4.4. Institut für Wirtschaftsinformatik – Knowledge Port 96

4.5. Erkenntnisse aus den Fallstudien 111

5. Methodenvorschlag für Information Retrieval in Portalen 120

5.1. Rollenmodell 121

5.2. Übersicht 122

5.3. Potenzialanalyse 124
5.4. Strategieplanung 132
5.5. Prozessplanung 145
5.6. Systementwurf 160
5.7. Konsolidierung und abschliessende Schritte 199
5.8. Dokumentationsmodell 205

6. Zusammenfassung und Ausblick 206
6.1. Ergebnisse der Arbeit 206
6.2. Ansätze zur Weiterentwicklung 211

Literaturverzeichnis 214

Inhaltsverzeichnis

1. Einleitung 1

1.1. Ausgangslage und Handlungsbedarf 1

1.2. Entstehung und Einordnung der Arbeit 3

1.3. Ziele und Adressaten 4

1.4. Forschungsmethodik 5

1.5. Aufbau der Arbeit 7

2. Grundlagen 9

2.1. Business Engineering 10

2.1.1. Einordnung und Abgrenzung 10

2.1.2. Drei-Ebenen-Modell 10

2.1.3. Geschäftsmodell des Informationszeitalters 11

2.1.4. Methodenentwicklung 13

2.1.5. Beitrag für die Dissertation 14

2.2. Portale 15

2.2.1. Einordnung und Abgrenzung 15

2.2.2. Portalfunktionen 16

2.2.3. Beitrag für die Dissertation 18

2.3. Information Retrieval 19

2.3.1. Einordnung und Abgrenzung 19

2.3.2. Indexierung und Suche 21

2.3.3. Portalsuchmaschinen 24

2.3.4. Beitrag für die Dissertation 26

2.4. Wissensmanagement 27

2.4.1. Einordnung und Abgrenzung 27

2.4.2. Wissensrepräsentation durch Informationsobjekte 28

2.4.3. Prozessunterstützung im CKM-Modell 29

2.4.4. Beitrag für die Dissertation 32
2.5. Konsequenzen für eine Information-Retrieval-Methode 33

3. Vergleich bestehender Methoden 34
3.1. Auswahlkriterien 35
3.2. Analyse 36
3.2.1. Extranet Development Lifecycle [Bayles 1998] 36
3.2.2. Engineering Enterprise Portals Framework [Finkelstein/Aiken 2000] 37
3.2.3. Methode zur Konzeption von Intranets [Kaiser 2000] 40
3.2.4. Implementierung von Unternehmensportalen [Bauer 2001] 42
3.2.5. Corporate Portals [Collins 2001] 44
3.2.6. Collaboration Portale [Puschmann 2003] 46
3.2.7. Integrierte Wissensmanagement-Systeme [Riempp 2003] 48
3.3. Übersicht der Bewertungen 51
3.4. Ableitung eines Information-Retrieval-Metamodells 52
3.4.1. Prozessunterstützung – Unterstützungspotenzial für Geschäftsprozesse 53
3.4.2. Retrieval-Leistungen – Definition des Funktionsumfangs 55
3.4.3. Prozessmanagement – Aufbau und Betrieb von Information Retrieval 57
3.4.4. Architekturelemente – Berücksichtigung relevanter Komponenten 58
3.4.5. Systemintegration – Anbindung bestehender Applikationen 61
3.4.6. Ganzheitlichkeit – Verbindung von Strategie, Prozessen und Systemen 63

4. Erfahrungen aus der Praxis 65
4.1. Zumtobel Staff – Customer Portal 66
4.1.1. Organisation und Ausgangssituation 66
4.1.2. Treiber und Herausforderungen für Information Retrieval 67
4.1.3. Unterstützungspotenzial 68
4.1.4. Benutzerschnittstelle und Retrieval-Leistung 68
4.1.5. Prozessmanagement 69

4.1.6. Systemintegration 71
4.1.7. Kosten- und Nutzenaspekte 75
4.1.8. Kritische Erfolgsfaktoren und geplante Weiterentwicklungen 75
4.2. The Information Management Group – Knowledge Pool 77
4.2.1. Organisation und Ausgangssituation 77
4.2.2. Treiber und Herausforderungen für Information Retrieval 79
4.2.3. Unterstützungspotenzial 79
4.2.4. Benutzerschnittstelle und Retrieval-Leistung 80
4.2.5. Prozessmanagement 81
4.2.6. Systemintegration 83
4.2.7. Kosten- und Nutzenaspekte 85
4.2.8. Kritische Erfolgsfaktoren und geplante Weiterentwicklungen 85
4.3. Winterthur Versicherungen – McB-Projektportal 87
4.3.1. Organisation und Ausgangssituation 87
4.3.2. Treiber und Herausforderungen für Information Retrieval 88
4.3.3. Unterstützungspotenzial 88
4.3.4. Benutzerschnittstelle und Retrieval-Leistungen 89
4.3.5. Prozessmanagement 90
4.3.6. Systemintegration 93
4.3.7. Kosten- und Nutzenaspekte 94
4.3.8. Kritische Erfolgsfaktoren und geplante Weiterentwicklungen 95
4.4. Institut für Wirtschaftsinformatik – Knowledge Port 96
4.4.1. Organisation und Ausgangssituation 96
4.4.2. Treiber und Herausforderungen für Information Retrieval 98
4.4.3. Unterstützungspotenzial 99
4.4.4. Benutzerschnittstelle und Retrieval-Leistung 99
4.4.5. Prozessmanagement 102
4.4.6. Systemintegration 105

4.4.7. Kosten- und Nutzenaspekte 109
4.4.8. Kritische Erfolgsfaktoren und geplante Weiterentwicklungen 110
4.5. Erkenntnisse aus den Fallstudien 111
4.5.1. Identifikation und Beurteilung des Unterstützungspotenzials 111
4.5.2. Analyse existierender Informationsstrukturen und Applikationen 113
4.5.3. Entwicklung einer Retrieval-Taxonomie 114
4.5.4. Spezifikation von Pflegeprozessen 117
4.5.5. Umsetzung von Anforderungen im Systementwurf 118

5. Methodenvorschlag für Information Retrieval in Portalen 120
5.1. Rollenmodell 121
5.2. Übersicht 122
5.3. Potenzialanalyse 124
5.3.1. Nutzenpotenziale identifizieren 125
5.3.2. Prozesse analysieren 126
5.3.3. Applikationen erfassen 128
5.3.4. Handlungsoptionen entwickeln 130
5.4. Strategieplanung 132
5.4.1. Leistungen festlegen 133
5.4.2. Architekturübersicht erstellen 136
5.4.3. Kosten und Nutzen bewerten 139
5.5. Prozessplanung 145
5.5.1. Retrieval-Prozess definieren 146
5.5.2. Retrieval-Taxonomie bereitstellen 147
5.5.3. Benutzer und Berechtigungen pflegen 152
5.5.4. Anwendungen integrieren 156
5.6. Systementwurf 160
5.6.1. Funktionen definieren 161
5.6.2. Suche 161

Suchanfrage 161
Suchergebnisanzeige 163
5.6.3. Navigation 167
5.6.4. Sicherheit 171
Authentifizierung und Single Sign On 171
Benutzerverwaltung 178
Autorisierung 180
5.6.5. Integration 184
5.6.6. Indizierung 192
5.6.7. Klassifikation 194
5.7. Konsolidierung und abschliessende Schritte 199
5.7.1. Software evaluieren 199
5.7.2. System implementieren 202
5.8. Dokumentationsmodell 205

6. Zusammenfassung und Ausblick 206
6.1. Ergebnisse der Arbeit 206
6.2. Ansätze zur Weiterentwicklung 211

Literaturverzeichnis 214

Abkürzungsverzeichnis

ASP	Active Server Pages (Microsoft)
ADB	Adressdatenbank (Zumtobel Staff)
API	Application Programming Interface
BASF	Badische Anilin- und Sodafabrik
BCI	Business Collaboration Infrastructure
BE	Business Engineering
BW	Business Warehouse (SAP)
BWL	Betriebswirtschaftslehre
bzw.	beziehungsweise
CA	Certificate Authority
ca.	circa
CC	Competence Center
CHF	Schweizer Franken
CIO	Chief Information Officer
CKM	Customer Knowledge Management (IWI-HSG)
C>K>P	Customer > Knowledge > Performance (IWI-HSG)
CM	Content Management
CMS	Content-Management-System
CORBA	Common Object Request Broker Architecture
CPS	Content Personalisation Server (Zumtobel Staff)
CRM	Customer Relationsship Management
CSS	Cascading Style Sheet
DB	Datenbank
d.h.	das heisst
DTD	Document Type Definition
EAI	Enterprise Application Integration
EBIT	Earnings Before Interest and Taxes
EP	Enterprise Portal (SAP)
ERP	Enterprise Resource Planning
EUR	Euro
etc.	et cetera
FAQ	Frequently Asked Questions
GB	Gigabyte

ggf.	gegebenenfalls
GTR	Global Text Retriever (IBM)
GUI	Graphical User Interface
HSG	Universität St. Gallen, Hochschule für Wirtschafts-, Rechts- und Sozialwissenschaften
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IBM	International Business Machines
ID	Identifikation
i.d.R.	in der Regel
IM	Instant Messaging
IMG	The Information Management Group
IO	Informationsobjekt
IR	Information Retrieval
IS	Informationssystem
ISO	International Organization for Standardization
IT	Informationstechnologie
IWI	Institut für Wirtschaftsinformatik
IWI-HSG	Institut für Wirtschaftsinformatik, Universität St. Gallen (HSG)
KM	Knowledge Management
K-Map	Knowledge Map (IBM)
KPort	Knowledge Port (IWI-HSG)
LDAP	Lightweight Directory Access Protocol
LDS	Lotus Discovery Server
LTPA	Lightweight Third Party Authentication
MB	Megabyte
McB	Management of Closed Blocks (Winterthur)
Mio.	Millionen
Mrd.	Milliarden
MS	Microsoft
NRPC	Notes Remote Procedure Call (IBM)
NTFS	New Technology File System (Microsoft)
ODBC	Open Database Connectivity
OSI	Open Systems Interconnection
PC	Personal Computer
PCIO	Profit Center International Organization (Zumtobel Staff)

PDA	Personal Digital Assistant
PDB	Produktdatenbank (Zumtobel Staff)
PDF	Portable Document Format
PKI	Public Key Infrastructure
PLM	Product Lifecycle Management
PROMET	Projektmethode
PwC	PricewaterhouseCoopers
RDF	Resource Description Framework
RMI	Remote Method Invocation
ROI	Return on Investment
s.	siehe
S.	Seite
SAP	Systeme, Anwendungen und Produkte in der Datenverarbeitung
SCM	Supply Chain Management
SQL	Structured Query Language
SSL	Secure Socket Layer
SSO	Single Sign On
TIBCO	The Information Bus Company
TOP	Technologie and Operations (Credit Suisse)
TREX	Text Retrieval and Extraction (SAP)
URL	Uniform Resource Locator
USA	United States of America
USB	Universal Serial Bus
u.U.	unter Umständen
vgl.	vergleiche
WAP	Wireless Access Protocol
WM	Wissensmanagement
WMS	Wissensmanagementsystem
WWW	World Wide Web
W3C	World Wide Web Consortium
XML	Extensible Markup Language
XSL	Extensible Stylesheet Language
XSLT	Extensible Stylesheet Language Transformations
z.B.	zum Beispiel
ZS	Zumtobel Staff

1. Einleitung

1.1. Ausgangslage und Handlungsbedarf

Unternehmen sehen sich beim Übergang in die Informationsgesellschaft einem *stetig wachsenden Informationsvolumen* gegenüber, das ohne den verstärkten Einsatz von Informationstechnologie nicht erschliessbar ist (vgl. [Fleisch/Österle 2004, 5], [Mertens 2003], [Brenner et al. 2002]). Sie setzen *Portale* als informationstechnische Werkzeuge ein, um Kunden, Mitarbeitern und Lieferanten relevante Inhalte und Funktionen aus unterschiedlichen unternehmensinternen und externen Applikationen in einem einheitlichen Frontend, z.B. einem Web-Browser, bereitzustellen (s. [Österle 2002a, 23], [Fleisch/Österle 2001b]).

Integrierte, personalisierte und in Echtzeit bereitgestellte Informationen unterstützen die Portalanwender bei der Durchführung ihrer *Kundenprozesse* (s. [Fleisch/Österle 2001a, 21], [Gurzki/Hinderer 2003]). Die gegen Null strebenden Grenzkosten elektronischer Dienstleistungen (s. [Fleisch 2000, 192]), eine höhere Kundenbindung, eine gesteigerte Prozesseffizienz und eine Kostenreduktion durch die verbesserte Verfügbarkeit von Informationen bilden das Nutzenpotenzial für den Portalbetreiber [s. Gillet 2001, 6].

Die Bereitstellung einer Vielzahl von Inhalten und Funktionen bestehender Anwendungen auf einer einheitlichen Portaloberfläche (Frontend) reduziert für die Anwender jedoch nicht a priori den *Aufwand der Informationssuche* (s. [Blessing 2001 , 106], [Kremer et al. 2003b, 2]. Suchmaschinen mit ihren Navigations- und Suchfunktionen kommt in Portalen eine besondere Bedeutung bei der effizienten und zeitnahen Erschliessung relevanter Informationen zu:

- 90% der in einer Studie befragten Unternehmen stufen Suchfunktionen in Portalen als ‚sehr wichtig' oder , wichtig' ein [s. Hagen 2000].
- In einer Studie der Yankee Group nennen 81% der befragten Unternehmen die Reduktion von Suchzeiten als den wichtigsten Grund zur Einführung eines Portals [s. Plumtree 2001, 12].
- CIO.de weist in einer Studie darauf hin, dass 90% der befragten Unternehmen bis zum Jahre 2005 Suchtechnologien in Portalen einführen werden [Surmacz 2003].

Die Anforderungen an die Einführung einer Portalsuchmaschine dürfen jedoch nicht unterschätzt werden. Projektverantwortliche sehen sich in den letzten Jahren mit veränderten technischen und organisatorischen Herausforderungen konfrontiert:

Einerseits erfordert die inhaltliche Integration einer zunehmenden Anzahl heterogener Anwendungen ein umfassendes *Verständnis der technologischen Möglichkeiten und*

Restriktionen dieser Backend-Systeme (s. [Ferber 2003, 31], [Raghavan 2002, 5]). Der heutige Einsatz verschiedener Datenbankanwendungen, Dateiserver, Content-Management-Systeme, E-Mail- und Groupware-Plattformen sowie einer Vielzahl weiterer dezentral organisierter Anwendungen mit ihren unterschiedlich stark strukturierten Inhalten, eigenen Zugriffsrechten und heterogenen Zielgruppen verhindert z.B. den direkten Einsatz der erfolgreichen Internet-Suchmaschine Google im betrieblichen Portalumfeld [s. Andrews 2003b, 5]. Die Bereitstellung relevanter Portalinhalte in Echtzeit verlangt beispielsweise eine einheitliche Semantik der integrierten Systeme [s. Fleisch/Österle 2004, 6] zu deren Umsetzung, z.B. in Form kontrollierter Suchvokabulare und anwendungsübergreifender Navigationsstrukturen, die Backend-Integration der inhaltsführenden Anwendungen (Content-Applikationen) notwendig ist.

Andererseits müssen Information-Retrieval-Lösungen heute gezielt an den *Anforderungen der Geschäftsprozesse* der jeweiligen Organisation ausgerichtet sein (s. [Riempp 2003, 4], [Abecker et al. 2002], [Delphigroup 2002]). Die Bereitstellung eines einzeiligen Eingabefeldes zur Eingabe eines beliebigen Suchbegriffs liefert den Anwendern selten die für ihre konkrete Aufgabenstellung benötigten Informationen (vgl. [Ferber 2003, 246], [Andrews 2003a], [Brewer 2002], [Gerick 2000]). 75% der in einer Wissensmanagementstudie Befragten (Portal- und Inhaltsverantwortliche sowie Endanwender) weisen der ‚Verbesserung von Suche' die höchste Priorität zu [s. Fank/Trojan 2003, 28]. 93,5% der befragten Unternehmen schätzen jedoch laut einer gemeinsamen Studie der Information Management Group AG und dem Institut für Wirtschaftsinformatik der Universität St. Gallen ihren „Kenntnisstand im Hinblick auf Portallösungen" nur als ‚sehr niedrig' bis ‚mittel' ein [IWI-HSG/IMG 2003, 39].

Trotz zahlreicher bestehender Problemlösungsansätze, u.a. aus den Bereichen der Betriebswirtschaftslehre (z.B. [Harris et al. 2003], [Goldratt 1990], [White 2002]), der Informatik (z.B. [Baeza-Yates/Ribeiro-Neto 1999], [Anderson/Pérez-Carballo 2001], [Gordon/Pathak 1999]), der Wirtschaftsinformatik (z.B. [Puschmann 2003], [Gurzki/Hinderer 2003]) und des Wissensmanagements (z.B. [Bach 2003], [Thiesse 2001]) steht den Verantwortlichen in Unternehmen bei der Einführung von Information Retrieval in Portalen kein umfassendes, den veränderten technischen und organisatorischen Rahmenbedingungen angepasstes methodisches Vorgehen zur Verfügung.

Die vorliegende Arbeit möchte mit einem Methodenvorschlag zur Analyse, Planung, Gestaltung und Implementierung von Information Retrieval in Portalen diese Lücke in Wissenschaft und Praxis schliessen.

1.2. Entstehung und Einordnung der Arbeit

Das vorliegende Dissertationsprojekt entstand im Rahmen des Kompetenzzentrums ‚Customer > Knowledge > Performance' (CC C>K>P, 2003 bis 2004) sowie des Vorgängers ‚Customer Knowledge Management' (CC CKM, 2001 bis 2002) des Instituts für Wirtschaftsinformatik der Universität St. Gallen (IWI-HSG). In dieser insgesamt vierjährigen Forschungskooperation erarbeiteten renommierte Unternehmen aus dem Europäischen Wirtschaftsraum in Zusammenarbeit mit dem IWI-HSG wissenschaftlich fundierte und gleichzeitig praxiserprobte Konzepte, Methoden und Lösungen in Bereichen des kundenorientierten Wissensmanagements. Ausgehend von der Wissensnutzung in betrieblichen Abläufen vertieften die Wissenschaftler und Praktiker gemeinsame Fragestellungen der Anwendung von Wissensmanagement-instrumenten wie beispielsweise Portalen und Suchmaschinen sowie der strategischen, organisatorischen und technischen Umsetzung von Wissensmanagementaktivitäten im Unternehmen. In über 14 Workshops und insgesamt 30 Praxisprojekten erarbeiteten die Forschungspartner in den beiden Kompetenzzentren entsprechende Grundlagen, Fallstudien sowie methodische Vorgehensweisen und setzten die gewonnenen Erkenntnisse in den Unternehmen um.

Der Kompetenzzentrumsansatz des Forschungsprogramms Business Engineering (BE) [s. Österle 1993] bildet die Grundlage dieser langjährigen Kooperation. Er verbindet konzeptionelle Forschung mit praktischer Beobachtung, die beispielsweise in Fallstudien beschrieben werden [s. Lee 1989] oder eine Implementierungsbegleitung in Form von partizipativer Aktionsforschung bei den Forschungspartnern ermöglichen [s. Whyte 1991]. Bei diesem kooperativen Forschungsansatz arbeiten immer mehrere Personen im Kompetenzzentrum an spezifischen Themenstellungen. Die vorliegende Arbeit weist daher Beziehungen zu folgenden abgeschlossenen Forschungsvorhaben am IWI-HSG auf:

- *Integrierte Wissensmanagementsysteme* (Dr. Gerold Riempp),
- *Method Engineering* (Prof. Dr. Thomas Gutzwiller),
- *Prozessorientiertes Wissensmanagement* (Dr. Frédéric Thiesse),
- *Collaboration Portale* (Dr. Thomas Puschmann),
- *Methode zur Konzeption von Intranets* (Dr. Thomas Kaiser) und
- *IT-gestütztes Kompetenzmanagement* (Dr. Henning Gebert).

1.3. Ziele und Adressaten

Die vorliegende Dissertation entwickelt eine Methode zur Analyse, Planung, Gestaltung und Implementierung von Information Retrieval in Portalen. Sie verfolgt damit folgende Ziele:

- Darstellung der *grundlegenden Beziehungen* zwischen Portalen, Information-Retrieval-Systemen und Wissensmanagement im Kontext des Business Engineering,
- Entwicklung eines *umfassenden Metamodells* durch Zusammenfassung relevanter Gestaltungselemente bestehender Portal- und Information-Retrieval-Methoden,
- Überprüfung des Metamodells durch eine *vergleichende Fallstudienanalyse*,
- Formulierung eines *Methodenvorschlags* in der Diktion des Method Engineering.

Diese Arbeit richtet sich an Adressaten aus Wissenschaft und Praxis, die sich mit den Potenzialen und Herausforderungen bei der Einführung von Information-Retrieval-Lösungen in Kunden-, Mitarbeiter- oder Lieferantenportalen konzeptionell oder in Umsetzungsprojekten auseinandersetzen:

- *Projektverantwortliche aus Fach- und IT-Bereichen* unterstützt die strukturierte Darstellung der Potenziale, Herausforderungen und kritischen Erfolgsfaktoren von Information-Retrieval-Lösungen bei der eigenen Planung. Die Methodenbestandteile auf Strategie-, Prozess- und Systemebene adressieren dabei sowohl betriebswirtschaftliche als auch technologische Fragestellungen. Darüber hinaus können Projektverantwortliche bei gleichartigen organisatorischen und technischen Rahmenbedingungen im eigenen Unternehmen Implementierungsdetails aus den umfassenden Beispielen der vier Praxisfälle direkt ableiten. Das Vorgehensmodell schliesst im Systementwurf mit einer Kriterienliste für eine Softwareauswahl ab, die sie beim individuellen Auswahlverfahren unterstützt bzw. relevante Beurteilungskriterien ableiten lässt.
- *Wissenschaftler*, die sich mit den Themenbereichen Portale, Information Retrieval und Wissensmanagement befassen, werden bei der Beurteilung theoretischer Integrationsansätze durch die Beschreibung realer Lösungen unterstützt. Die in der vorliegenden Arbeit enthaltenen Grundlagen, die Analyse und Bewertung der sieben bestehenden Portal- und Information-Retrieval-Methoden, die Darstellungen der vier Praxisfälle und der Methodenvorschlag selbst können in Lehrveranstaltungen Verwendung finden.
- *Implementierer von Information-Retrieval-Lösungen*, z.B. Mitarbeiter von prozess- und systemorientierten Beratungsgesellschaften unterstützt das Vorgehensmodell

bei der Durchführung von Praxisprojekten. Von der Potenzialanalyse über die Strategieplanung, Prozessplanung und den Systementwurf unterstützt die Methode durch die Angabe von Rollen und Techniken die sukzessive Erarbeitung von Ergebnissen.

- *Hersteller von Portalen und Suchmaschinen* kann das in dieser Arbeit enthaltene theoretische Rahmenwerk sowie die Umsetzungsbeispiele dahingehend unterstützen, ihre Lösungsansätze methodisch fundiert auf die Herausforderungen der betrieblichen Praxis abzustimmen.

1.4. Forschungsmethodik

Business Engineering bildet als angewandte und handlungsorientierte Wissenschaft die Schnittstelle zwischen betriebswirtschaftlicher und informationstechnischer Forschung ab [s. Ulrich 1984, 178]. Zielsetzung ist die Abbildung der betrieblichen Wirklichkeit des Forschungsgegenstandes in Form von Methoden und Handlungsoptionen. Die resultierenden Forschungsergebnisse sind in Form von umsetzbaren Handlungsempfehlungen in der Praxis validierbar [s. Österle et al. 1992, 35].

Die Mehrzahl der im Business Engineering erarbeiteten Ergebnisse basieren auf qualitativer Forschung (s. z.B. [Brenner 1993], [Brecht 1999], [Fleisch 2000], [Österle/Senger 2003]). Die Forschungsmethodik zur Erstellung dieser Arbeit orientiert sich an der Grundideen des ‚Participatory Action Research' (partizipative Aktionsforschung), welche die strikte Trennung zwischen Forscher und Forschungsobjekt aufhebt [s. Whyte 1991, 20]. Sie ermöglicht in Anlehnung an [Fleisch 2000, 21] folgenden arbeitsteiligen Prozess zwischen Wissenschaft und Praxis:

- Praxis und Wissenschaft definieren gemeinsam die Problemstellung.
- Die Wissenschaft strukturiert die Probleme und entwickelt Vorschläge für die Gestaltung der betrieblichen Wirklichkeit. Grundlagen hierfür sind theoretisches Wissen sowie Praxiserfahrungen.
- Wissenschaft und Praxis überprüfen anschliessend die Vorschläge und verfeinern sie weiter.
- Die Praxis wendet die Vorschläge an, d.h. sie gestaltet die betriebliche Wirklichkeit gemäss den Empfehlungen.
- Schliesslich werden die Ergebnisse gemeinsam überprüft und die Vorschläge entsprechend weiterentwickelt.

Die vorliegende Arbeit bezieht beim Entwurf der Methode zur Einführung von Information Retrieval in Portalen die Lösungen für die in Tabelle 1-1 dargestellten projektbezogenen Aufgabenstellungen ein.

Projektbezeichnung	Forschungspartner	Laufzeit	Aufgabenstellung
Teilprojekt Search	Zumtobel Staff	02/2003 – 12/2003	Einführung einer Suchmaschine im Kunden- und Mitarbeiterportal (Teilprojektleitung)
TOP Customer Portal	Credit Suisse	02/2003 – 03/2003	Klassifikationsgerüst für Inhalte eines Mitarbeiterportals
Rollenmodelle für Portale	Cambridge Technology Partners	01/2003 – 03/2003	Vorgehensmodell für Rollenmodelle in Portalen
Wissensportal SIRiUS	BASF	11/2001 – 12/2002	Softwareauswahlverfahren für eine Portalplattform
KnowledgePort	IWI-HSG	07/2001 – 12/2002	Aufbau von Mitarbeiter-, Extranet- und Internetportalen (Projektleitung)
IMG Knowledge Management	IMG	05/2001 – 08/2001	Klassifikationsgerüst für Inhalte eines Mitarbeiterportals
Projektportal McB	Winterthur Versicherungen	11/2000 – 06/2001	Implementierungsbegleitung für ein Projektportal
PwC Knowledge Management	Pricewaterhouse Coopers	07/1999 – 09/2000	Administration und Customizing einer Suchmaschine für ein Mitarbeiterportal

Tabelle 1-1: Projektumfeld der Arbeit

Die Kooperationen mit den Forschungspartnern Zumtobel Staff, Winterthur Versicherungen, The Information Management Group (IMG) und die Umsetzung am Institut für Wirtschaftsinformatik (IWI-HSG) umfassen eine Feldstudienzeit von mehr als 150 Personentagen. Aufgrund dieser Signifikanz und der Verfügbarkeit von validen Daten [s. Yin 1994, 33] stellt die vorliegende Arbeit diese vier Praxisfälle in Anlehnung an die von [Senger 2004] erarbeiteten Konzepte in Form von Fallstudienbeschreibungen dar.

1.5. Aufbau der Arbeit

Die vorliegende Dissertation umfasst sechs Kapitel (s. Abbildung 1-1). Der Einleitung (Kapitel 1) folgen die Kapitel 2 bis 4, deren Teilergebnisse zur Bearbeitung des jeweils nächsten Kapitels dienen. Kapitel 5 greift die in den vorherigen Kapiteln erarbeiteten Erkenntnisse erneut auf, um mit dem Methodenvorschlag den Kern des Dissertationsprojekts zu erarbeiten. Die Arbeit schliesst mit einer Zusammenfassung und einem Ausblick in Kapitel 6 ab.

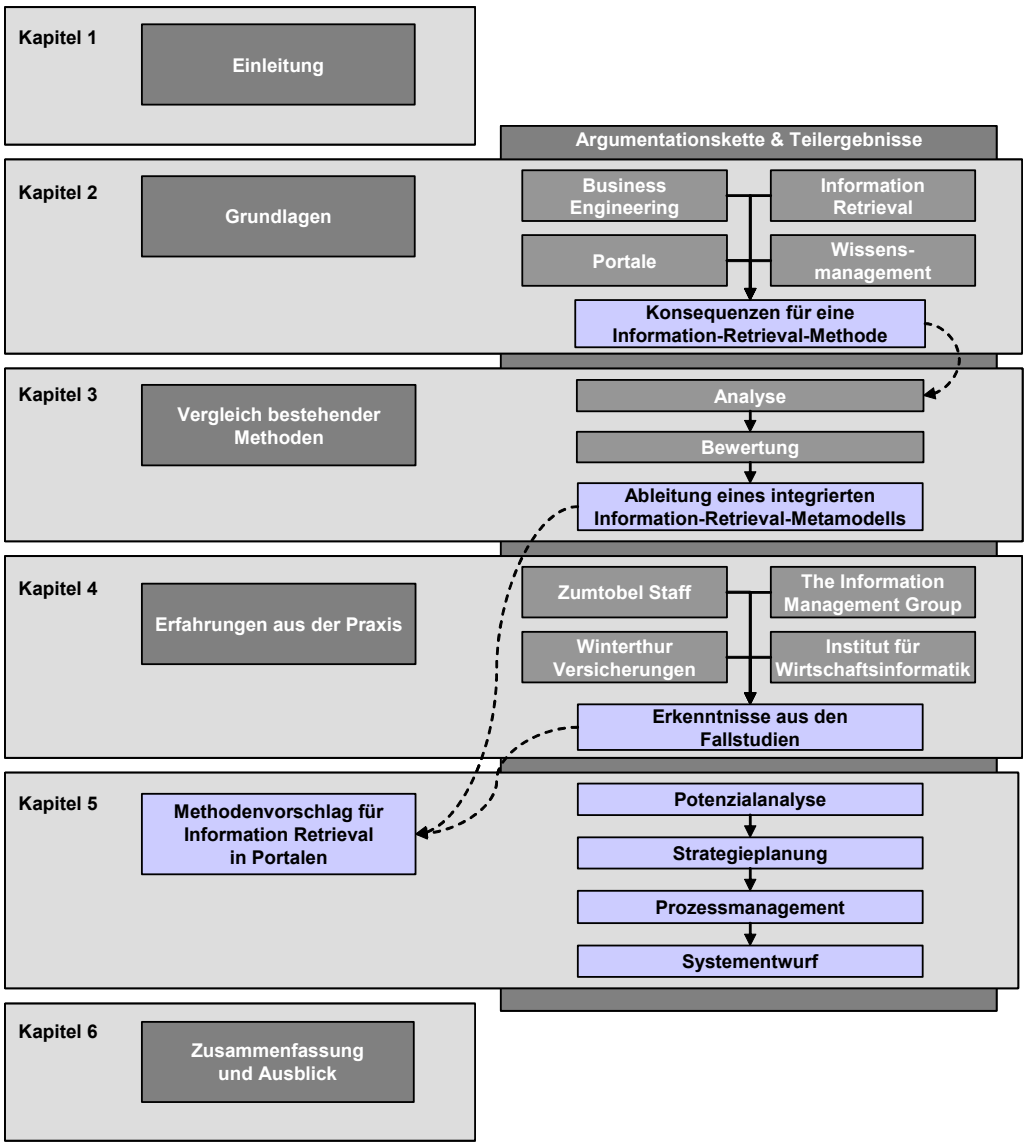


Abbildung 1-1: Übersicht über Aufbau, Argumentationskette und (Teil-)Ergebnisse des vorliegenden Dissertationsprojekts

Das *Kapitel 1* skizziert die *Ausgangslage* und leitet den *Handlungsbedarf* für eine Information-Retrieval-Methode ab. Die enthaltenen Abschnitte beschreiben die *Entstehung des Dissertationsprojekts* und ordnen es im *Kontext wissenschaftlicher Forschungsbereiche* ein. Die Formulierung von *Zielen und Adressaten* der vorliegenden Arbeit, der verwendeten *Forschungsmethodik* und ein Überblick über den *Aufbau der Dissertation* schliessen das vorliegende Kapitel 1 ab.

Das *Kapitel 2* vermittelt die für Information Retrieval in Portalen notwendigen Grundlagen und formuliert aus diesen *Konsequenzen für eine Information-Retrieval-Methode*. Ausgehend von den Anforderungen des Forschungsprogramms *Business Engineering* (BE) ordnen die enthaltenen Abschnitte die für die vorliegende Arbeit wesentlichen Begriffe ‚*Portale*' und ‚*Information Retrieval*' ein, grenzen sie ab und formulieren ihre jeweiligen Beiträge für die Dissertation. *Wissensmanagement* greift beide Begriffe auf, zeigt ihre Zusammenhänge und formuliert ihren Beitrag für die Unterstützung von Geschäftsprozessen im Sinne des Business Engineering.

Das *Kapitel 3* selektiert und analysiert *sieben bestehende Portal- und Information-Retrieval-Methoden* anhand der aus den Grundlagen in Kapitel 2 abgeleiteten Anforderungen der betrachteten Forschungsschwerpunkte Business Engineering, Portale, Information Retrieval und Wissensmanagement. Die enthaltenen Abschnitte beschreiben und bewerten die Methoden, skizzieren ihre jeweiligen Stärken und Lücken und leiten aus diesen die *Gestaltungselemente eines Information-Retrieval-Metamodells* für einen eigenen Methodenvorschlag ab.

Das *Kapitel 4* beschreibt *vier Praxisprojekte*, in denen der Autor der vorliegenden Arbeit die Umsetzung von Information-Retrieval-Lösungen aktiv begleitet hat. Die enthaltenen Abschnitte strukturieren die Darstellung jedes einzelnen Falls zunächst anhand der Gestaltungselemente des in Kapitel 3 erarbeiteten Metamodells. Die übergreifenden *Erkenntnisse aus den Fallstudien* dienen anschliessend gemeinsam mit dem aus den sieben wissenschaftlichen Methoden in Kapitel 3 abgeleiteten Metamodell als *Grundlage zur Entwicklung eines eigenen Methodenvorschlags*.

Das *Kapitel 5* schlägt eine *Methode für die Umsetzung* von Information Retrieval in Portalen vor. Die Grundlage dazu bildet das in Kapitel 3 entwickelte *Metamodell*. Die beschriebenen *Techniken* des *Vorgehensmodells* bieten Vorschläge für die Analyse, Planung, Gestaltung und Implementierung von Information Retrieval in Portalen auf den Ebenen Strategie, Prozesse und Systeme. *Vertiefende Beispiele* aus den Praxisprojekten verdeutlichen die Anwendung und die Ergebnisse der in den Techniken beschriebenen *Aktivitäten*. Sie ergänzen gleichzeitig die Fallstudien in Kapitel 4. Ein *Dokumentationsmodell* stellt anschliessend wesentliche Ergebnisdokumente des Methodenvorschlags dar und schliesst damit das Kapitel 5 ab.

Das *Kapitel 6* fasst wesentliche *Ergebnisse der Arbeit* zusammen und bietet mit einem Ausblick auf aktuelle Trends *Anregungen zur Weiterentwicklung*.

2. Grundlagen

Das vorliegende Kapitel vermittelt einen Überblick über die wissenschaftlichen Grundlagen, die als Basis für die Analyse, Planung, Gestaltung und Implementierung von Information-Retrieval-Lösungen in Portalen dienen.

Der Bezugsrahmen für die theoretische Auseinandersetzung mit dem Themenspektrum einer wissenschaftlichen Arbeit (s. [Yin 1994, 18], [Miles/Huberman 1994, 18], [Stake 1995, 15]) besteht im vorliegenden Dissertationsprojekt aus vier Schwerpunkten:

- Den Forschungsrahmen bilden die Grundsätze des *Business Engineering* (BE). Sie adressieren die ingenieurmässige Vorgehensweise bei der projektorientierten Umsetzung von Potenzialen der Informationstechnologie. Die resultierenden Lösungen auf Strategie-, Prozess- und Systemebene orientieren sich dabei konsequent an den Anforderungen der Kunden eines Unternehmens im Informationszeitalter.
- *Portale* realisieren als informationstechnologische Komponenten die integrierte Darstellung von Inhalten und Funktionen heterogener Systeme auf einer einheitlichen Benutzeroberfläche. Sie bilden die zentrale Schnittstelle zwischen Unternehmen im Informationszeitalter und ihren Kunden, Lieferanten und Mitarbeitern. Anwendungsübergreifende Navigations- und Suchmechanismen ermöglichen den verschiedenen Zielgruppen eine effiziente Erschliessung der im Portal vorhandenen Informationen.
- *Information Retrieval* adressiert die systematische Erschliessung relevanter Inhalte durch Indexierung und Suche. Der Einbezug intellektueller Leistungen bei der Indexierung verbessert die Erschliessbarkeit von Inhalten, z.B. durch Klassifikation mit beschreibenden Attributen (Metadaten). In ihrer Ausprägung als Suchmaschinen müssen Information-Retrieval-Systeme in Portalen eine Vielzahl heterogener, semi- oder unstrukturierter Informationsquellen integrieren.
- Mit Portalen und Suchmaschinen stellt das *Wissensmanagement* zwei informationstechnologische Komponenten zur bedarfsgerechten Informationsversorgung unterschiedlicher Zielgruppen bereit. Suchen und Finden von in Informationsobjekten abgebildetem Wissen ist die Grundlage für die Nutzung der Ressource ‚Wissen' in Geschäftsprozessen eines Unternehmens.

Die folgenden Abschnitte detaillieren die vier skizzierten Themenschwerpunkte und stellen ihren jeweiligen Beitrag für die Dissertation dar. Der Abschnitt 2.5 konsolidiert die gewonnen Erkenntnisse und formuliert Konsequenzen für eine Information-Retrieval-Methode.

2.1. Business Engineering

2.1.1. Einordnung und Abgrenzung

Business Engineering bezeichnet die „methoden- und modellbasierte Konstruktionslehre für Unternehmen des Informationszeitalters“ [Österle/Winter 2003, 7]. Die am Institut für Wirtschaftsinformatik der Universität St. Gallen entwickelte Forschungsdisziplin beschäftigt sich mit Herausforderungen, die bei der Transformation eines Unternehmens aus der Industrie- in die Informationsgesellschaft entstehen (s. [Österle/Winter 2003, 7], [Österle/Blessing 2003]). Business Engineering sieht Innovationen im Bereich der Informationstechnologie als so genannte ‚Enabler’, die bisherige unternehmerische Tätigkeiten verbessern (z.B. durch Qualitätssteigerungen, Kostensenkung und Beschleunigung) und neue Geschäftsfelder erschliessbar machen (z.B. durch die Integration von Informationen entlang der Wertschöpfungskette und den Datenaustausch in Echtzeit) [s. Riempp 2003, 61]. Business Engineering bringt betriebswirtschaftliches und informationstechnisches Wissen zusammen und verbindet es mit allen Aspekten der Transformation, von Darstellungsmitteln über Vorgehensmodelle bis hin zu kulturellen und politischen Gesichtspunkten (s. [Österle/Blessing 2000, 62], [Senger 2004, 29]). Neben dem fachlichen Entwurf von Geschäftslösungen adressiert das Business Engineering mit Projektmanagement und Change Management die Begleitung der Veränderungsprozesse [s. Österle/Blessing 2003].

2.1.2. Drei-Ebenen-Modell

Bei der Transformation von Unternehmen durch Potenziale der Informationstechnologie betrachtet das Business Engineering im fachlichen Entwurf drei Gestaltungsebenen:

- *Strategie*. „Die Ebene Strategie definiert die Position des Unternehmens im Markt und die daraus abgeleiteten Schlüsselentscheidungen für das Unternehmen und seine Geschäftsfelder“ [Österle 1995, 3].
- *Prozess*. „Die Ebene Prozess leitet aus der Strategie die Leistungen, den Ablauf, die Computerunterstützung und Führungsmittel ab und detailliert die Organisationsstruktur“ [Österle 1995, 5].
- *System*. „Die Ebene (Informations-)System konkretisiert den Prozessentwurf; sie liefert die Vorgabe für die organisatorische und informationstechnische Implementierung“ [Österle 1995, 6].

Durch die Betrachtung der Gestaltungselemente aller drei Ebenen integriert Business Engineering zentrale Bereiche der Wirtschaftsinformatik, der Organisationslehre und

des Technologiemanagements [s. Österle 1995, 14] und erlaubt dadurch einen ganzheitlichen Ansatz bei der Ableitung und Realisierung von Geschäftslösungen des Informationszeitalters (s. [Österle 1995], [Brenner 1995]).

2.1.3. Geschäftsmodell des Informationszeitalters

Das Geschäftsmodell des Informationszeitalters ist das Zielbild der durch die Business-Engineering-Forschung untersuchten Transformation von Unternehmen aus dem Industrie- in das Informationszeitalter. Die Unternehmen des Informationszeitalters sind danach nicht mehr produkt-, sondern kundenorientiert [s. Österle 1999, 18]. Sie richten ihr Handeln an den in so genannten Kundenprozessen zusammengefassten Anforderungen ihrer Kunden aus. Ein Kundenprozess bestimmt daher, welche Leistungen ein Unternehmen anbieten muss (s. [Porter 1984], [Ives/Learmonth 1984], [Österle 1995]). Da ein einzelnes Unternehmen i.d.R. nicht alle Leistungen eines Kundenprozesses erbringt, organisieren sich komplementäre Unternehmen in Wertschöpfungsnetzwerken, um mit ihren Leistungen einen Kundenprozess möglichst umfassend abzudecken [s. Fleisch 2000].

Die aus dem Geschäftsmodell abgeleitete Geschäftsarchitektur illustriert die wesentlichen Bestandteile unternehmerischen Handelns im Informationszeitalter. Sie sind in [Österle 2002a] ausführlich dargestellt. Die folgende Darstellung fokussiert ausschliesslich auf die im Kontext der vorliegenden Arbeit relevanten Elemente:

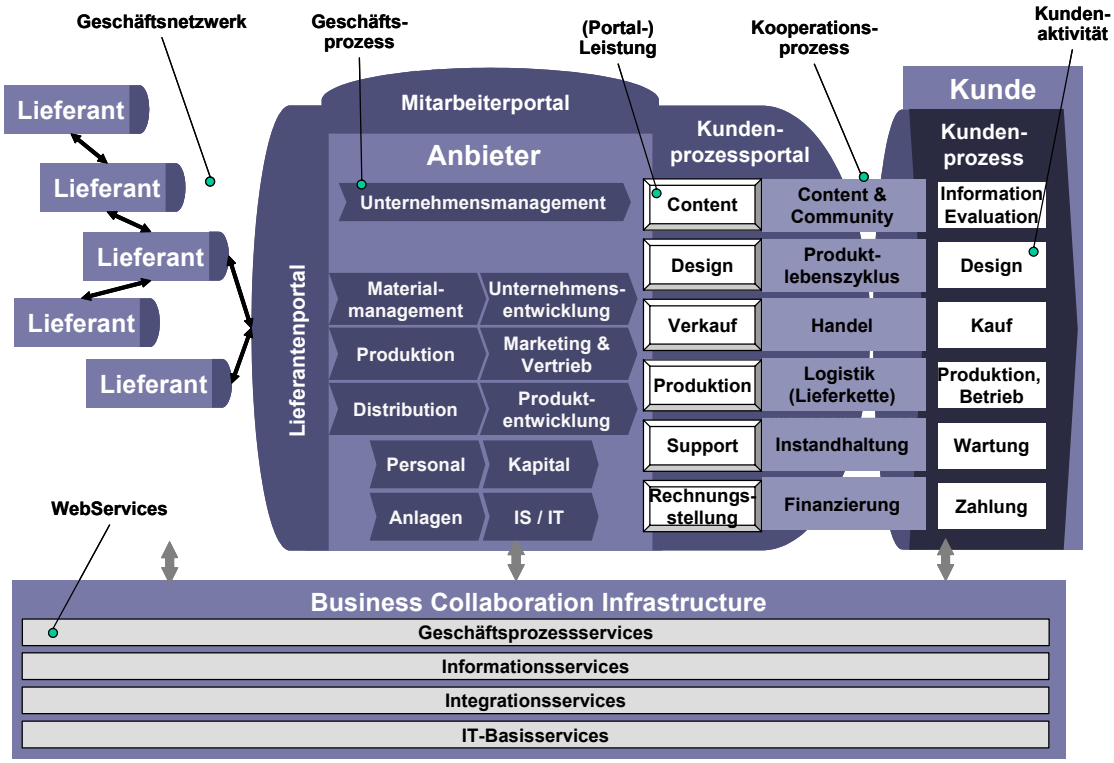


Abbildung 2-1: Geschäftsarchitektur des Informationszeitalters [Österle 2002b, 335]

Ein *Kundenprozess* fasst alle Aufgaben zusammen, die ein Kunde zur Befriedigung eines bestimmten Bedürfnisses durchläuft und für deren Unterstützung er Leistungen von Unternehmen bezieht. Dadurch bestimmt der Kundenprozess die Produkte und Dienstleistungen, die ein Anbieter seinen Kunden in *Kooperationsprozessen* offerieren kann. So genannte Lebenszyklusansätze, bei denen Kunden zur Bedürfnisbefriedigung einen Zyklus von *(Kunden-)Aktivitäten* durchlaufen, liefern Unternehmen Ansatzpunkte, Kundenprozesse zu identifizieren [Reichmayr 2002]. Kunde ist in der Geschäftsarchitektur des Informationszeitalters ein Leistungsbezieher, unabhängig davon, ob es sich um einen Endkunden oder eine in der Wertschöpfungskette nachgelagerte Organisationseinheit handelt [s. Senger 2004, 33]. Kunden bevorzugen in der Geschäftsarchitektur des Informationszeitalters dabei den Anbieter, der ihren Kundenprozess möglichst umfassend abdeckt [s. Österle 2002a, 22].

Die vollständige Abdeckung eines umfassenden Kundenprozesses durch ein einzelnes Unternehmen ist nicht möglich [Senger 2004, 34]. Ein Unternehmen im Informationszeitalter integriert daher seine eigenen Produkte und Dienstleistungen mit denen weiterer Lieferanten in einem *Geschäftsnetzwerk*. Diese Leistungsintegration nützt dem Kunden, da dieser im Idealfall pro Kundenprozess nur eine Geschäftsbeziehung unterhalten muss.

Das *Kundenprozessportal* bildet die Schnittstelle zwischen Kunden und Anbieter. Es fasst alle *(Portal-)Leistungen* zusammen, die ein Unternehmen für einen spezifischen Kundenprozess über elektronische Medien zur Verfügung stellen kann ([Davydov 2001, 182], [Schwarz 2000, 41]. Das Portal integriert dabei, z.B. als zentrale Anlaufstelle im Web-Browser, alle für den Kunden und seinen Kundenprozess relevanten Inhalte und Funktionen aus unterschiedlichen Applikationen (s. [Österle 2002a, 23], [Fleisch/Österle 2001b]). Die gegen Null strebenden Grenzkosten elektronischer Dienstleistungen (s. [Fleisch 2000, 192]), eine höhere Kundenbindung, eine gesteigerte Prozesseffizienz und eine Kostenreduktion durch die integrierte Verfügbarkeit von Informationen bilden das Nutzenpotenzial für den Anbieter (s. [Puschmann 2003, 1], [Gillet 2001, 6]).

Das *Mitarbeiterportal* stellt analog zum Kundenprozessportal den Beschäftigten des Anbieters als ‚internen Kunden' alle notwendigen Informationen und den Zugang zu betrieblichen Anwendungen bereit, die sie zur Leistungserbringung in *Geschäftsprozessen* benötigen.

WebServices sind standardisierte elektronische Dienstleistungen. Ein Unternehmen im Informationszeitalter kann WebServices an externe Dienstleister auslagern, wenn diese billiger als innerhalb des Unternehmens erbracht werden (z.B. Lohnabrechnung) oder zur Koordination zwischen Unternehmen des Geschäftsnetzwerks notwendig sind (z.B. Zahlungsverkehrsdienste) [s. Österle 2002a, 33]. Als klar abgrenzbare Prozessbestandteile übernehmen WebServices Aufgaben, die die beteiligten

Unternehmen in ähnlicher Form benötigen und daher günstig in elektronischer Form zukaufen können. WebServices müssen zeit- und/oder transaktionsbasiert verrechenbar und in Informationssystemen abbildbar sein [s. Reichmayr 2002, 103].

Die *Business Collaboration Infrastructure* dient zur Kommunikation und Leistungsintegration zwischen den Partnern des Geschäftsnetzwerks. Über organisatorische Vereinbarungen sowie technisch standardisierte Protokolle und Schnittstellen bietet sie allen Beteiligten die notwendige Infrastruktur, um miteinander über WebServices zu interagieren [s. Österle/Winter 2003, 9].

2.1.4. Methodenentwicklung

Methoden bezeichnen planmässige, begründete Vorgehensweisen zur Erreichung definierter Ziele im Rahmen festgelegter Prinzipien [s. Balzert 1996, 36]. Das Business Engineering hält Forschungsergebnisse vielfach in Methoden fest (s. z.B. [Gebert 2003], [Puschmann 2003], [Kaiser 2000], [Gronover 2003]). Dabei bietet das Method Engineering zur Darstellung von Business-Engineering-Methoden eine eigene Metamethode und ein eigenes Metamodell (s. [Gutzwiller 1994, 11], vgl. zur Metamodellierung [Ferstl/Sinz 1998, 117], [Scheer 1998]).

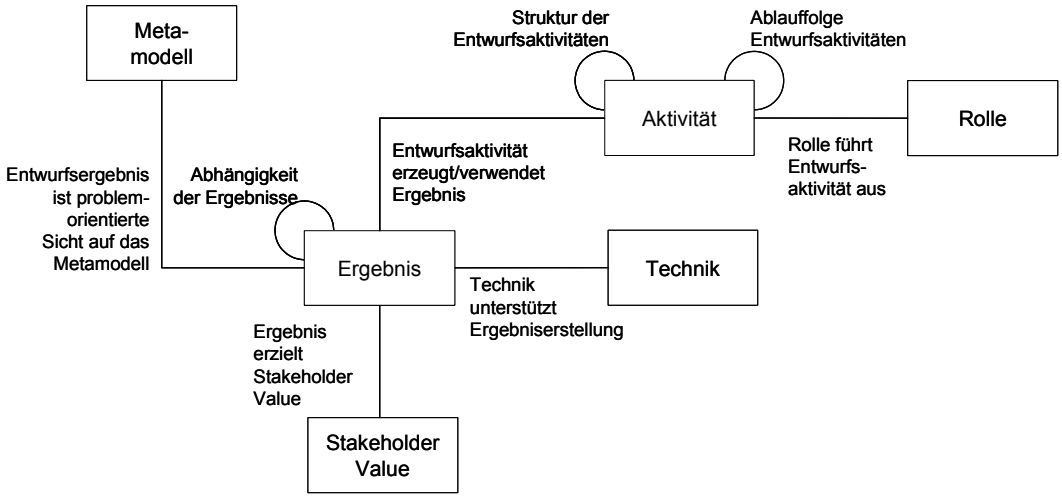


Abbildung 2-2: Metamodell des Method Engineering [Österle/Blessing 2000, 76]

Die Erstellung von *Ergebnissen*, z.B. in Form ganzer Geschäftslösungen aber auch einzelner Dokumente, spiegelt die Ergebnisorientierung des Business Engineering wider. Ergebnisse definieren sich durch ihren Wert gegenüber relevanten Interessengruppen (‚*Stakeholder Value*').

Ergebnisse werden durch *Aktivitäten* erzeugt und verwendet [s. Gebert 2003, 20]. Die Verantwortung für die Ausführung von Aktivitäten wird dabei von *Rollen* wahrgenommen, die den Mitwirkenden in einem Projekt zugeordnet sind. *Techniken* leiten und unterstützen dabei die Erstellung von Ergebnissen, z.B. in Form von Checklisten, die auf kritische Fragestellungen fokussieren.

Das *Metamodell* beschreibt die Gestaltungsobjekte des Business Engineering und deren Beziehungen. Es skizziert wesentliche Elemente der Wertschöpfung von Unternehmen auf Basis des Drei-Ebenen-Modells (s. Abschnitt 2.1.2).

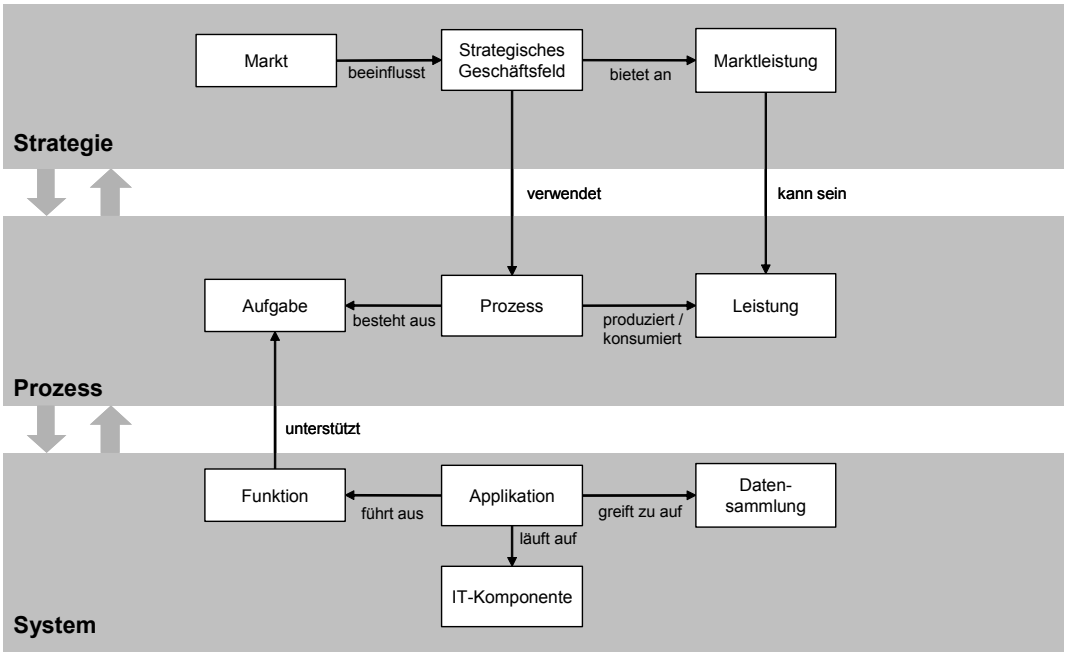


Abbildung 2-3: Metamodell des Business Engineering [Österle/Blessing 2000, 77]

2.1.5. Beitrag für die Dissertation

Das vorliegende Dissertationsprojekt gewinnt aus dem Forschungsrahmen des Business Engineering folgende Erkenntnisse:

- *Informationstechnologie*. Business Engineering leitet aus informations-technologischen Innovationen sowohl Verbesserungen bisheriger unter-nehmerischer Tätigkeiten als auch die Erschliessung neuer Geschäftsfelder ab. Die Umsetzung der Potenziale der Informationstechnologie erfolgt in Transformationsprojekten.

- *Ganzheitlichkeit*. Durch die gleichwertige Betrachtung von Trans-formationsprojekten auf Strategie-, Prozess- und Systemebene verbindet Business Engineering als ganzheitlicher Managementansatz betriebswirtschaftliches und informationstechnisches Wissen.

- *Kundenprozessorientierung*. Die in Kundenprozessen zusammengefassten Anforderungen der Kunden bilden den Ausgangspunkt eines Trans-formationsprojekts. Sie bestimmen das Handeln von Unternehmen im Informationszeitalter.

- *Portale*. Kundenprozessportale bilden die Schnittstelle zwischen Kunden und Anbieter. Sie integrieren die Inhalte und Funktionen unterschiedlicher Applikationen, die ein Unternehmen zur umfassenden Abdeckung eines spezifischen Kundenprozesses bündeln kann. Mitarbeiterportale adressieren ‚interne Kunden' bei ihrer Leistungserbringung.
- *Ingenieurmässiges Vorgehen*. Durch die Vorgabe von Methoden und die Entwicklung von Handlungsoptionen systematisiert Business Engineering den Transformationsprozess. Die Grundlage dafür bilden die Elemente des Method Engineering.

2.2. Portale

2.2.1. Einordnung und Abgrenzung

Die betriebliche Aufgabenerfüllung erfolgt i.d.R. nicht nur mit einer einzigen Anwendung [s. Österle et al. 1996, 3]. Vertikale Anwendungen für z.B. betriebswirtschaftliche Standardprozesse (Enterprise Resource Planning, ERP), Kundenbeziehungsmanagement (Customer Relationship Management, CRM), Management der Liefer- bzw. Wertschöpfungskette (Supply Chain Management, SCM) und Management von Produkten (Product Lifecycle Management, PLM) unterstützen gezielt definierte Abläufe der Anwender, während horizontale Anwendungen (z.B. E-Mail-, Groupware- und Content-Management-Systeme) ihre Funktionen prozessübergreifend zur Verfügung stellen (vgl. [Stahlknecht/Hasenkamp 2002, 330], [Mertens et al. 1998]). Eine durchgängige und umfassende Unterstützung von Kundenprozessen (s. Abschnitt 2.1.3) erfordert jedoch in vielen Fällen eine übergreifende Integration von Inhalten und Funktionen dieser in sich geschlossenen Anwendungen [s. Puschmann 2003, 32].

Applikationen zur integrierten Darstellung von Inhalten und Funktionen heterogener Systeme in einer einheitlichen Benutzerschnittstelle mit anwendungsübergreifenden Navigations- und Suchmechanismen werden seit Ende 1998 unter dem Begriff ‚Portal' diskutiert (s. [Shilakes/Tylman 1998], [White 2000]). [Puschmann 2003, 58] bietet einen Überblick zum Portalbegriff, der in der wissenschaftlichen Literatur und in den Darstellungen aus der Praxis mit einer Vielzahl fachlicher und technischer Definitionen belegt ist:

Autor	Definition
[Fleisch/Österle 2001a, 20]	„Ein Portal ist ein elektronisches Fenster auf Services (Dienstleistungen), die aus Sicht des Nutzers zusammengehören. Das elektronische Fenster ist die Benutzerschnittstelle eines Portals zu internen und externen Applikationen. Es betreibt die grafische bzw. audiovisuelle Integration (Frontend-Integration) von Services. [...] Portale verschaffen internen und externen Usern einen einfachen Zugang zu einem umfassenden Set an aufeinander abgestimmten Value Added Services. Der Nutzen für den Portal-User ist die Backend-Integration dieser Services.“
[Röhricht/Schlögel 2001, 163]	„Ein Portal ist eine innerhalb eines Browsers ablauffähige Applikation, die bisher unstrukturierte Informationen und Services klassifiziert und strukturiert, sowie einen personalisierten Zugriff auf Informationen und Services erlaubt.“
[Kalakota/Robinson 2001, 87]	„A portal [...] offers an aggregated set of services for a specific well-defined group of users – either value-added services to the market channel or decrease the transaction costs associated with the customer/supplier relationship.“
[Davydov 2001, 182]	„It [a portal] is not a product sold by a vendor, but a goal to be achieved through the integration of multiple products from multiple vendors. It is a concept of a unification platform that allows for a collection of application services to work together to facilitate access to a world of information. The ability to aggregate these services and to provide the necessary platform for them to work cooperatively are the real values of this concept, in general, and of corporate portals, in particular.“
[Höller et al. 1998, 16]	„Durch ihr Frontend – den Browser – eröffnen sie [Portale] den Zugriff auf unterschiedlichste Informationen und Ressourcen. Insbesondere der Zugriff auf Ressourcen unterschiedlichster Datenformate und -logik wie ERP-Systeme, Host-Anwendungen, Mail-Systeme und Unternehmensdatenbanken wird durch Portale auf ideale Weise zentral gebündelt und den Anwendern zur Verfügung gestellt.“
[Schwarz 2000, 41]	„Portale können als web-basierte, personalisierbare und integrierte Zugangssysteme zu Applikationen, Content und Services verstanden werden, die der ganzheitlichen Unterstützung von Kundenprozessen dienen.“

Tabelle 2-1: Eine Auswahl von Definitionen des Begriffs ‚Portal’ (in Anlehnung an [Puschmann 2003, 58])

Die vorliegende Arbeit versteht ein Portal als ein Informationssystem, das Benutzer entlang ihrer Prozesse durch die Integration und Bereitstellung von Inhalten und Funktionen bestehender, heterogener Anwendungen unter einer einheitlichen Benutzeroberfläche umfassend unterstützt. Portale stellen ihre Leistungen nicht nur gebündelt dar, sondern unterstützen auch ihre Nutzung durch eine Vielzahl von (Portal-)Funktionen [vgl. Österle 2002b].

2.2.2. Portalfunktionen

Applikationsfunktionen unterstützen die Durchführung betrieblicher Aufgaben im Rahmen von Prozessen (s. Metamodell des Business Engineering, Abbildung 2-3). Art und Umfang der durch ein Portal bereitgestellten Funktionen beeinflussen demnach die Prozessunterstützung von Kunden, Mitarbeitern und Lieferanten eines Unternehmens.

Die wissenschaftliche Literatur, Veröffentlichungen aus der Praxis sowie Darstellungen von Softwareherstellern weisen Portalen unterschiedliche Funktionen zu (vgl. [Dias 2001], [Davydov 2001], [Bauer 2001], [Bullinger et al. 2002], [IBM 2003],

[SAP 2003]). Obwohl teilweise unterschiedlich bezeichnet, lassen sich Portalfunktionen wie folgt zusammenfassen:

- Die *Präsentation* umfasst die Darstellung von Informationen gegenüber den Portalbenutzern. Sie ist abhängig von den Anzeige- und Eingabemöglichkeiten des Endgeräts, mit dem der Benutzer auf das Portal zugreift, z.B. Web-Browser, WAP-Handy, Smartphone oder Personal Digital Assistant (PDA) [s. Puschmann 2003, 76]. Greift der Portalbenutzer mit einem Web-Browser zu, so präsentiert das Portal die Benutzeroberfläche als grafische Benutzerschnittstelle (Graphical User Interface, GUI).

- Die *Navigation* umfasst die Bereitstellung von begrifflichen und grafischen Strukturen, die den Anwendern als Wegweiser eine schrittweise Bewegung innerhalb und zwischen Bereichen der Benutzeroberfläche ermöglicht [Riempp 2003, 219]. Die Portalnavigation fasst beispielsweise produkt- oder dienstleistungsbezogene Informationen und Funktionen eines Unternehmens zusammen.

- Die *Suche* erlaubt den Portalbenutzern die Formulierung eines Informationsbedürfnisses über eine Suchanfrage und liefert entsprechende Antworten in Form von Suchergebnissen zurück [Riempp 2003, 218]. Die Suchergebnisse verweisen durch elektronische Verknüpfungen (Hyperlinks) auf die Fundorte, die vom Benutzer in der Suchergebnisliste gesichtet, beurteilt und bei Bedarf aufgerufen werden.

- *Indizierung* (auch Indexierung) bezeichnet die Zuordnung von inhaltsbeschreibenden Begriffen, so genannten Deskriptoren, zu einem Dokument [s. Nohr 2000, 19]. Die Deskriptoren repräsentieren den Inhalt eines Dokuments und dienen zu dessen Erschliessung. Suchmaschinen (‚Search Engines') automatisieren diesen Vorgang, indem sie selbständig die Inhalte der im Portal angebundenen Anwendungen (Content-Applikationen) indizieren und den Benutzern zur Erschliessung umfassende Navigations- und Suchfunktionen bieten.

- Die *Klassifikation* fasst die im Portal enthaltenen Inhalte systematisch zusammen. Ein hierarchisches Klassifikationsschema, eine so genannte Taxonomie, ermöglicht den Portalbenutzern die strukturierte Erschliessung einer grossen Anzahl von Informationen aus heterogenen Quellen durch Navigation und Suche anhand homogener Klassen (s. [Gaus 2003], [Bailey 1994], [Logan 2001]).

- Die *Personalisierung* umfasst die Individualisierung der Kommunikation zwischen Portal und Benutzer, z.B. durch die Anpassung der Benutzeroberfläche durch den Anwender [s. Schackmann/Schü 2001].

- Die (Portal-)*Integration* umfasst die Einbindung bestehender Informationen und Funktionen anderer Anwendungssysteme. Sie schafft den wesentlichen Nutzen für den Anwender durch die Backend-Integration der im Frontend angebotenen Leistungen [s. Fleisch/Österle 2001a, 28].

- Die *Authentifizierung* stellt die elektronische Identität eines Portalanwenders eindeutig fest (s. [Raepple 2001, 5], [Zwerger/Schneider 2002, 153]). In diesem Zusammenhang kommt der einmaligen Authentifizierung (so genanntes Single Sign On) gegenüber allen im Portal integrierten Applikationen eine besondere Bedeutung zu. Beim Single Sign On übernehmen alle im Portal integrierten Anwendungen die vom Portal beim Login des Benutzers vorgenommene Authentifizierung.

- Die *Autorisierung* einer Anwendung prüft, ob ein Benutzer zur Durchführung einer bestimmten Aktion berechtigt ist (s. [Raepple 2001, 5], [Nusser 1998, 15]). Diese Prüfung, die so genannte Zugangskontrolle, basiert auf der Durchsetzung von Zugriffsrechten, die an eine Kombination aus Subjekt (z.B. Benutzer oder Applikation) und Objekt (z.B. Informationsobjekt, Applikation oder Gruppe von Objekten) gekoppelt sind (s. [Oppliger 1998, 55], [Ahuja 1996, 182]).

2.2.3. Beitrag für die Dissertation

Das vorliegende Dissertationsprojekt leitet aus den Definitionen des Portalbegriffs und den skizzierten Funktionen folgende Erkenntnisse ab:

- Portale führen als *integrierende Informationssysteme* bestehende Inhalte und Funktionen verschiedener anderer Anwendungen in einem *einheitlichen Frontend*, z.B. einem Web-Browser, zusammen. Sie bieten ihren Benutzern dadurch einen einfachen, integrierten Zugang zu einem umfassenden, *rollenspezifischen Angebot aufeinander abgestimmter elektronischer Dienste*.

- In ihrer Ausprägung als *Kundenprozessportale* stellen Portale ein *zentrales Gestaltungselement der Geschäftsarchitektur von Unternehmen im Informationszeitalter* dar (s. Abschnitt 2.1.3). Portalfunktionen unterstützen die Aufgabenerfüllung von Kunden, Mitarbeitern und Lieferanten.

- Mit *Suche und Navigation* stellen Portale ihren Anwendern im Portal-Frontend zwei Funktionen bereit, um die zur Durchführung von Geschäftsprozessen *benötigten Informationen selbständig zu finden*. Ihre Bereitstellung *erfordert eine Backend-Integration* vorhandener Applikationen durch die *Abbildung der Strukturen bestehender Inhalte* (Klassifikation), die *Berücksichtigung von Zugriffs-*

beschränkungen einzelner Applikationen (Authentifizierung, Autorisierung) und die *Erfassung heterogener Datenformate* (Indizierung).

2.3. Information Retrieval

2.3.1. Einordnung und Abgrenzung

Die zunehmende elektronische Informationsverarbeitung lässt die Menge der digital verfügbaren Inhalte immer schneller anwachsen (s. [Latham 2001], [Fleisch/Österle 2001a]). Obwohl die vernetzte betriebliche Arbeitswelt mit Content-Management-Systemen, E-Mail-Anwendungen, Datenbanken, Dateiservern etc. gleichzeitig den potentiellen Zugang von Kunden, Mitarbeitern und Lieferanten zu diesen Inhalten ermöglicht, schafft erst die Erschliessung und anschliessende Anwendung von Informationen einen wirtschaftlichen Nutzen (vgl. [Nohr 2003], [Ferber 2003]).

Information Retrieval (IR) adressiert die Erschliessung relevanter Informationen durch die Repräsentation, Speicherung, Organisation von und den Zugang zu elektronisch verfügbaren Daten (vgl. [Baeza-Yates/Ribeiro-Neto 1999, 1], [Salton/McGill 1987, 1], [Rijsbergen 1979, 3]). Zielsetzung und Handlungsoptionen ergeben sich aus der Abgrenzung des Daten- und Informationsbegriffs, den wissenschaftliche Autoren unterschiedlich definieren (vgl. [Nonaka/Takeuchi 1995, 20], [Rehäuser/Krcmar 1996, 4], [Probst et al. 1999, 36], [Schreyögg 2001]).

Die vorliegende Arbeit orientiert sich an der von [Riempp 2003] gewählten Abgrenzung: Daten sind demnach „syntaktisch geordnete Repräsentationen von etwas Wahrnehmbaren, beispielsweise einem Messwert, einer Telefonnummer oder einer Bilanzkennzahl. Im Umfeld von Computersystemen sind sie maschinenlesbar, d.h. in digitaler Form speicherbar“ [Riempp 2003, 76]. Informationen stellen „Daten in einem für Menschen bedeutungsvollen Kontext“ dar [Riempp 2003, 76]. Informationssysteme speichern demnach Daten in Form so genannter Informationsobjekte (z.B. Dokumente), die von Menschen in einen Kontext gestellt werden.

Die vorliegende Arbeit versteht ein Information-Retrieval-System als ein Informationssystem, das Benutzer bei der Erschliessung (Retrieval, Zurückgewinnung) benötigter Informationen durch die Repräsentation, Speicherung, Organisation von und den Zugang zu in text- bzw. dokumentenorientierten Anwendungen (so genannten Content-Applikationen) gespeicherten Informationsobjekten unterstützt.

Abbildung 2-4 zeigt das grundlegende Schema eines solchen Information-Retrieval-Systems [s. Ferber 2003, 21]:

- *Autor und Repräsentation von Objekten*. Die von einem Autor bereitgestellten Objekte sind im Information-Retrieval-System durch eine so genannte Repräsentation, ein Abbild charakteristischer Merkmale der Objekte, dargestellt.

Eine Suchmaschine repräsentiert die aus einer Content-Applikation, z.B. einem Content-Management-System indizierten Dokumente durch die im Suchindex enthaltenen Begriffe.

- *Nutzer und Repräsentation des Informationsbedarfs*. Der Nutzer als Suchender nach (Informations-)Objekten repräsentiert seinen Informationsbedarf im IR-System. In einer Suchmaschine formuliert er seinen Informationsbedarf beispielsweise durch die Eingabe entsprechender Suchbegriffe in einem Suchfeld. Die im Benutzerdialog eingegebenen Suchdetails repräsentieren im IR-System seinen Informationsbedarf.

- *Zuordnung von Objekten zum Informationsbedarf*. Das IR-System bestimmt auf Basis der Repräsentation des Informationsbedarfs und der Repräsentation aller Objekte eine Menge von Objekten, die dem Nutzer als Antwort auf seine Suchanfrage zurückgeliefert werden.

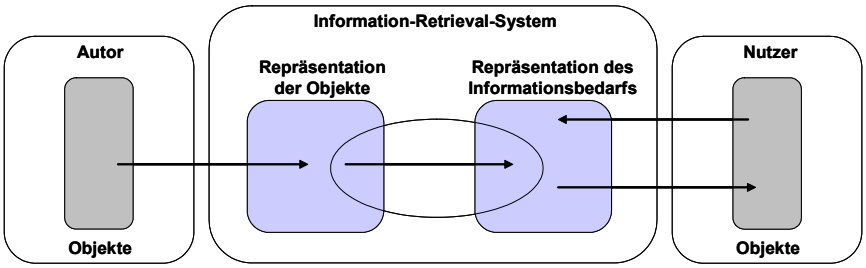


Abbildung 2-4: Schema eines Information-Retrieval-Systems [Ferber 2003, 25]

Seit Ende der 70er Jahre erweitern und verfeinern Forscher und Praktiker die ursprüngliche Zielsetzung von Information-Retrieval-Sytemen, Textbestandteile aus elektronischen Datenbeständen zu extrahieren und zur Suche bereitzustellen. Sie adressieren dabei unterschiedliche Bestandteile des in Abbildung 2-4 dargestellten Schemas eines Information-Retrieval-Systems, z.B. durch Aspekte wie Informations-modellierung, Dokumentenklassifikation, Suchalgorithmen, Systemarchitekturentwurf und Benutzerschnittstellengestaltung (s. [Fuhr 2002, 6], [GI 2004], [Baeza-Yates/Schäuble 2002]).

Die vorliegende Arbeit strebt im Kontext des Business Engineering eine ausgewogene Darstellung betriebswirtschaftlicher und informationstechnologischer Aspekte von Information Retrieval an. Suchalgorithmen als Forschungsgebiet der Informatik werden daher nicht betrachtet. Diese sind in entsprechender Fachliteratur detailliert dargestellt (vgl. [Nohr 2003, 33], [Baeza-Yates/Ribeiro-Neto 1999, 19], [Sparck Jones/Willett 1997, 305], [Ferber 2003, 33]).

2.3.2. Indexierung und Suche

Der schematische Ablauf beim Information Retrieval (IR) umfasst Indexierung und Suche als zwei komplementäre Aktivitäten [s. Sparck Jones/Willett 1997, 1]:

- *Indexierung*. Die Indexierung dient der inhaltlichen Erschliessung von in Content-Applikationen enthaltenen Informationsobjekten. Sie umfasst das begriffliche Erfassen des Inhalts (Inhaltsanalyse) und die Abbildung des Inhalts (Inhaltsrepräsentation) durch sprachliche Elemente in einem Index. Das Ziel der Indexierung ist es, die indizierten Inhalte suchbar und auffindbar zu machen.
- *Suche*. Die Suche umfasst Benutzeraktivitäten zum Lokalisieren relevanter Informationsobjekte, z.B. die Angabe von Suchbegriffen, die Navigation in Begriffsstrukturen (‚Browsing') aber auch die Prüfung von Suchergebnislisten und die Auswahl entsprechender Suchergebnisse.

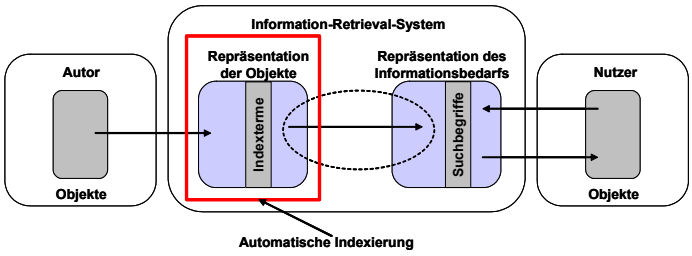
Während eine hohe Interaktion zwischen Benutzer und IR-System die Vorgänge der Suche bestimmt, beziehen verschiedene Indexierungsverfahren die Autoren von Informationen unterschiedlich stark ein (s. [Gaus 2003], [Nohr 2003, 19]).

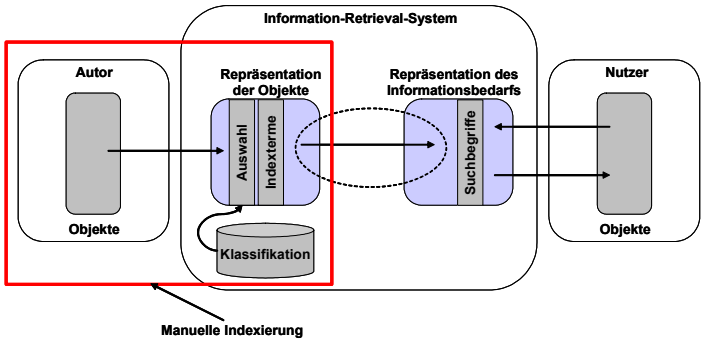
- Inhaltsanalyse und -repräsentation werden bei *intellektuellen (manuellen, gebundenen) Indexierungsverfahren* von Menschen vorgenommen. Sie streben eine korrekte und konsistente Repräsentation von Inhalten an, indem sie die durch Inhaltsanalyse erkannten Sachverhalte durch normierte Benennungen einer Indexierungssprache (kontrolliertes Vokabular), so genannte Deskriptoren, wiedergeben. Deskriptoren erlauben den Autoren beispielsweise die bewusste Beschreibung des Kontexts der geschäftlichen Situation, in dem ein Informationsobjekt entstanden ist, z.B. Ort, Zeit, Zielgruppe, Vorgang, Thema etc. (s. [Dey/Abowd 1999], [Klemke 2000]). Diese inhaltsbeschreibenden Daten (Metadaten) sind als Eigenschaften (Attribute) mit dem jeweiligen Informationsobjekt verbunden und in der Content-Applikation gespeichert (vgl. [Cathro 1997], [Berners-Lee 1997]).
- *Automatische (freie) Indexierungsverfahren* extrahieren dagegen selbständig ausgewählte Begriffe aus einem Inhalt. Sie speichern diese als so genannte Indexterme ab oder fügen Deskriptoren einer kontrollierten Indexierungssprache dem Inhalt als Inhaltsrepräsentation hinzu.
- Die *semiautomatische Indexierung* stellt eine Zwischenstufe dar, bei der maschinell erzeugte Indexterme durch Menschen kontrolliert und ggf. nachbearbeitet werden.

Unabhängig vom Indexierungsverfahren erzeugen von einem Autor manuell zugewiesene Deskriptoren oder von einem IR-System automatisch extrahierte Indexterme eine Beziehung (Relation) zwischen einem Informationsobjekt und einem sprachlichen Begriff. Da Begriffe gedankliche Abbildungen von Gegenständen, Sachen, Vorgängen, Ideen etc. sind, können sich Unterschiede bei der Interpretation durch Menschen ergeben [Gaus 2003, 56].

Die Linguistik (Sprachwissenschaft) adressiert diese Herausforderung mit der Lehre von Wortbedeutungen, der so genannten Semantik. Sie untersucht die Beziehung zwischen Objekten und den sie bezeichnenden Benennungen (Terme) [s. Rodi 1992, 298]. Im Gegensatz dazu untersucht die Syntax die formale Relation von Zeichen eines Begriffs, ohne seine Beziehung zu einem Objekt zu berücksichtigen [s. Bach 2003, 113]. Im betrieblichen Umfeld beschreiben z.B. Glossare die für einen Bereich relevanten Terme in ihrer Bedeutung (Semantik), während Taxonomien als hierarchische Ordnungsrahmen eine kontrollierte Syntax bieten und sich damit als Klassifikationsschema eignen [s. Kremer et al. 2003a]. Die Informatik bezeichnet begriffliche Beschreibungssysteme auch vielfach als Ontologien (s. [Guarino 1992, 196], [Staab 2002, 200], [Hesse 2002, 477]).

Tabelle 2-2 stellt die Wirkungsweise und Charakteristika verschiedener Indexierungs- und Suchverfahren im grundlegenden Schema für Information-Retrieval-Systeme (s. Abbildung 2-4) dar, die sich insbesondere durch die Kontrolle der Syntax sowie den Einbezug intellektueller Leistungen von Autor und Nutzer unterscheiden.

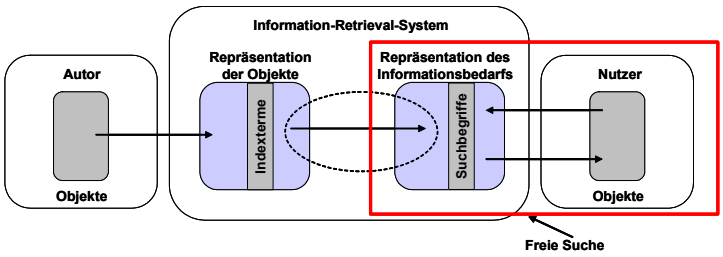
Indexierungsverfahren	Charakteristika
Automatische Indexierung	
 • Das IR-System analysiert die vom Autor bereitgestellten Objekte und repräsentiert diese durch selbständig gewählte Indexterme. • Das IR-System arbeitet sowohl für Autoren als auch für Nutzer als ‚Black Box', da die Indizierungsmechanismen für beide nicht ersichtlich sind. • Moderne IR-Systeme kontrollieren die Syntax von Suchbegriffen, z.B. durch den Vorschlag alternativer Suchterme bei Rechtschreibfehlern.	• Grosse Dokumentenmengen sind ohne automatische Indexierung nicht mehr zu bewältigen [Nohr 2003, 11]. • Bei der automatischen (freien) Indexierung fehlen intellektuelle Beiträge, somit entfällt auch jegliche terminologische Kontrolle durch Autoren oder Nutzer [Gaus 2003, 256]. • Intellektuelle Leistungen, z.B. das Finden von passenden Suchbegriffen zu vorhandenen Inhalten verlagern sich auf die Suchaktivitäten. • Zeitaufwand und Kosten der Recherche fallen überwiegend beim Nutzer an.

<table>
<tr><th colspan="2">Manuelle Indexierung</th></tr>
<tr><td>


Die Autoren ordnen bei der manuellen (gebundenen) Indexierung Objekte anhand eines systematischen Ordnungsschemas (Klassifizierung). Sie verwenden dabei relevante Deskriptoren aus einem kontrollierten Vokabular (Klassifikation) [Gaus 2003, 52].
Der Autor beeinflusst aktiv das IR-System, für den Nutzer ist es hingegen eine ‚Black Box'.

</td><td>

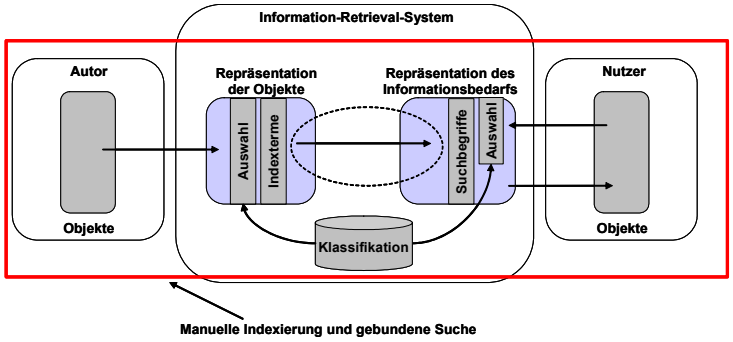
Der Kontext eines Objekts wird vom Autor durch die Klassifikation explizit beschrieben.
Resource Description Frameworks (RDFs) bieten eine Vielzahl vordefinierter Attribute zur Beschreibung von Inhalten, wie z.B. der Dublin Core Metadata Standard [Dublincore 2003].
Die Klassifikation bietet durch ein kontrolliertes Vokabular eine Syntaxkontrolle [Gaus 2003, 53].
Für Verschlagwortung sowie für Erstellung und Pflege der Verschlagwortungsliste (Taxonomie) entstehen Kosten [Altenhofen et al. 2002].

</td></tr>
<tr><th>Suchverfahren</th><th>Charakteristika</th></tr>
<tr><th colspan="2">Freie Suche</th></tr>
<tr><td>


Der Nutzer gibt bei der freien Suche (Freitextsuche, Volltextsuche, Stichwortsuche) beliebige Suchbegriffe ein, um seinen Informationsbedarf zu übermitteln.
Das IR-System liefert Hinweise auf Objekte zurück, deren Indexterme mit den Suchbegriffen übereinstimmen.

</td><td>

Die gezielte Informationssuche (Recherche) in grossen Datenbeständen erfordert vom Nutzer eine hohe fachliche und sprachliche Qualifikation [Gaus 2003, 261].
Eine terminologische Kontrolle der Syntax von Suchbegriffen findet nicht statt. Moderne IR-Systeme können jedoch beispielsweise Rechtschreibfehler des Nutzers automatisch korrigieren oder alternative Terme vorschlagen.

</td></tr>
<tr><th colspan="2">Gebundene Suche</th></tr>
<tr><td>


Autor und Nutzer verwenden für Indexierung und Suche dieselbe Klassifizierung.

</td><td>

Der Kontext eines Objekts wird vom Autor explizit beschrieben.
Der Kontext des Autors ist für den Nutzer durch die Auswahl von Klassifikationsmerkmalen ersichtlich bzw. aus den verwendeten Begriffen rekonstruierbar.
Autor und Nutzer verwenden dieselbe Syntax, die die Klassifikation mit den Bezeichnungen in den Auswahllisten zur Verfügung stellt.

</td></tr>
</table>

Tabelle 2-2: Wirkungsweise und Charakteristika unterschiedlicher Indexierungs- und Suchverfahren

Eine ausschliesslich manuelle Indexierung ist bei einem hohen Datenvolumen in den seltensten Fällen wirtschaftlich gerechtfertigt [Ferber 2003, 81]. Automatische Indexierung und Freitextsuche berücksichtigen hingegen nicht die in Organisationen

häufig vorliegenden Klassifizierungen von Inhalten bestehender Systeme. Diese Klassifizierungen bzw. die von Autoren in den Metadaten der Dokumente hinterlegten Klassifikationsmerkmale reflektieren den Kontext, indem ein Inhalt erstellt worden ist. Sie stellen für die Nutzer aussagekräftige Kriterien dar, um die Relevanz von gefundenen Inhalten schnell beurteilen zu können [s. Kremer et al. 2003b, 4]. In der betrieblichen Praxis hat daher eine Kombination aus intellektuellem Erschliessen durch Klassifikation und Freitextsuche auf Basis automatischer Indexierung eine hohe Bedeutung [s. Gaus 2003, 263].

2.3.3. Portalsuchmaschinen

Portale integrieren heterogene Informationsquellen und stellen den Benutzern Information-Retrieval-Funktionen wie Suche, Navigation, Indizierung, Klassifikation und Präsentation über Suchmaschinen (‚Search Engines') zur Verfügung (s. [Puschmann 2003, 75], [Riempp 2003, 220]). Suchmaschinen bilden dabei wesentliche Gestaltungselemente eines Information-Retrieval-Systems in zwei unterschiedlichen Ausprägungen ab (s. [Baeza-Yates/Ribeiro-Neto 1999, 191, 249], [Brin/Page 1998]):

- *Indexer*. Eine Suchmaschinenkomponente, der so genannte Crawler (auch Spider) durchsucht die in einer Organisation vorhandenen Quellen und Applikationen (z.B. Dateiserver, Intranetseiten und Datenbankinhalte) und stellt diese zur Weiterverarbeitung bereit [s. Bach 2003, 84]. Im nächsten Verarbeitungsschritt indiziert die Suchmaschine die gefunden Dokumente. Im Index sind die extrahierten Begriffe (Terme) mit Verweisen auf ihren jeweiligen Fundort abgelegt. Suchmaschinen nach dem Indexer-Prinzip implementieren eine automatische Indexierung, können jedoch auch Klassifikationsmerkmale indizierter Dokumente (Meta-Tags) auslesen und für eine gezielte Suche zur Verfügung stellen. Kommerzielle Anbieter von Indexer-Suchmaschinen sind beispielsweise Verity (K2, Ultraseek), Autonomy, IBM / Lotus (Lotus Discovery Server, Global Text Retriever, Lotus Domain Search) und Google (Google Search Appliance).

- *Broker*. Suchmaschinen nach dem Brokerprinzip legen keinen eigenen Suchindex an. Sie nutzen existierende Indizes angebundener Applikationen, indem sie die Suchanfrage des Benutzers an diese verteilen und die zurückgelieferten Ergebnisse vor der Anzeige aggregieren. Aufgrund ihrer Funktionsweise werden Suchmaschinen nach dem Brokerprinzip auch als föderierte Suche (auch ‚Federated Search' oder ‚Meta Search') bezeichnet. Broker-Suchmaschinen erfordern bestehende Indizes in den zu durchsuchenden Quellen, die automatisch oder manuell indiziert sind. Kommerzielle Broker-Suchmaschinen bieten beispielsweise IBM / Lotus (Lotus Extended Search) und Infopeople (intrence) an.

Die unterschiedlichen Funktionsweisen von Indexer- und Broker-Suchmaschinen führen zu in den in Tabelle 2-3 abgebildeten Vor- und Nachteilen:

Vorteile	Nachteile
Indexer-Suchmaschinen	
• Ermöglichen den Aufbau und die Bereitstellung einer dynamischen, begriffsbasierten Navigationsstruktur • Performance, z.B. Antwortzeitverhalten ist ausschliesslich von der Leistungsfähigkeit der Suchmaschine abhängig und mit dieser skalierbar • Erlaubt die Verknüpfung von Indexinformationen mit weiteren Details, z.B. der Bewertung von Suchergebnissen	• Erfordern spezifische Schnittstellen (Konnektoren) für die Indizierung unterschiedlicher Quellen • Index benötigt bis zu 80% des Speicherplatzes des indizierten Datenvolumens • Indizierung belastet die Infrastruktur, z.B. das Netzwerk und die zu indizierenden Applikationen
Broker-Suchmaschinen	
• Nutzen bestehende Indizes, benötigen somit keinen eigenen Speicherplatz • Ermöglichen die parallele Bearbeitung von Suchanfragen in den angefragten Applikationen • Sichern höchste Aktualität der Suchergebnisse, da direkt in den Inhalten der Quellen und Applikationen gesucht wird • Erlauben die automatische Umsetzung einer Suchanfrage in unterschiedlich komplexe Suchanfragen der angesprochenen Applikationen	• Erfordern eigene Suchmechanismen bzw. Indizes in den angebundenen Quellen • Nicht zum Aufbau einer dynamischen, begriffsbasierten Navigationsstruktur geeignet

Tabelle 2-3: Vor- und Nachteile unterschiedlicher Suchmaschinenprinzipien

Eine grosse Systemvielfalt, ein hohes Volumen unstrukturierter Daten und der Einsatz von Internettechnologien haben die Rahmenbedingungen für Information Retrieval im betrieblichen Umfeld in den letzten Jahren erheblich verändert (vgl. [Ferber 2003, 31, 285], [Baeza-Yates/Schäuble 2002], [Raghavan 2002], [Andrews 2003b]). Unternehmen sind beim Einsatz von Portalsuchmaschinen mit folgenden Herausforderungen bei der Integration in ein bestehendes Informationssystem konfrontiert:

- *Heterogene Objekte*. Informationsobjekte weisen unterschiedliche Formate auf, z.B. HTML, Microsoft Word und Adobe PDF. Unterschiedliche Länge, Granularität und Strukturinformationen erhöhen die Vielfalt von Dokumentenarten. Inhalte sind teilweise klassifiziert und in verschiedenen Sprachen verfasst. Heterogene Informationsobjekte erfordern daher eine Beschreibung der sie kennzeichnenden Merkmale, z.B. Format, Sprache, Schlüsselwörter, Autorenschaft und Erstelldatum [s. Ferber 2003, 246]. Portalsuchmaschinen müssen unterschiedliche Formate und Strukturinformationen indizieren, um den Anwendern in der Suchergebnisliste beispielsweise eine einheitliche Darstellung aller gefundenen Informationsobjekte zu geben.

- *Verteilte und redundante Datenspeicherung*. Informationen sind im betrieblichen Umfeld in einer Vielzahl von Applikationen verteilt und teilweise redundant gespeichert, z.B. in E-Mail-Systemen, Dateiservern und Content-Management-Systemen. Über entsprechende Schnittstellen, so genannte Konnektoren, müssen Portalsuchmaschinen die benötigten Anwendungen anbinden, um deren Inhalte zu indizieren.

- *Dynamik und Lebenszyklus von Portalinhalten*. Portalinhalte bestimmen ihren Wert durch ihre Aktualität. Sie unterliegen einem so genannten Lebenszyklus, der mit der Erstellung und Publikation beginnt und mit der Archivierung von Inhalten endet. Unternehmen müssen Portalsuchmaschinen diesen Anforderungen anpassen, indem die Indizes aktuell gehalten werden und neue Inhalte auch ab dem Zeitpunkt ihrer Publikation gesucht und gefunden werden können. Eine besondere Herausforderung stellen dabei beispielsweise Content-Management-Systeme dar, die ihre Inhalte erst zum Zeitpunkt der Abfrage durch einen Benutzer dynamisch generieren.

- *Zugriffsbeschränkungen*. Portale integrieren zielgruppenspezifisch Inhalte und Applikationen. Sind diese zugriffsbeschränkt, d.h. nur einer definierten Anwendergruppe zugänglich, so müssen auch Portalsuchmaschinen diese Restriktionen berücksichtigen. Anwender dürfen über Suche und Navigation ausschliesslich Inhalte finden, auf die sie in den indizierten Systemen auch direkt zugreifen können.

2.3.4. Beitrag für die Dissertation

Das Dissertationsprojekt gewinnt aus den Ansätzen des Information Retrieval folgende Erkenntnisse:

- Information Retrieval adressiert im betrieblichen Umfeld durch den Einsatz von Portalsuchmaschinen die *einfache und komfortable Erschliessung von text- bzw. dokumentenorientierten Informationsobjekten* aus einer Vielzahl bestehender, heterogener Anwendungen.

- Portalsuchmaschinen stellen den Benutzern dazu Information-Retrieval-Funktionen wie Suche, Navigation, Indizierung, Klassifikation und Präsentation zur Verfügung. Die *mögliche Vielfalt von Dokumentenformaten, Klassifikations-strukturen und Zugriffsberechtigungen* der zu integrierenden Applikationen erfordert eine genaue Analyse, Planung und Gestaltung bei der Implementierung einer Portalsuchmaschine.

- Die Autoren von Inhalten können *inhalts- und kontextbeschreibende Details (Metadaten) in den Meta-Tags von Informationsobjekten* hinterlegen. Diese erlauben dem Nutzer eine *gezielte, attributbasierte Suche* auch in grossen Informationsbeständen. Darüber hinaus stellen die Metadaten von Informationsobjekten für die Nutzer *aussagekräftige Kriterien* dar, um die Relevanz von gefundenen Inhalten im eigenen geschäftlichen Kontext schnell beurteilen zu können.
- Die *wirtschaftlichen Vorteile einer automatischen Indexierung* können mit den *verbesserten Suchmöglichkeiten einer manuellen Klassifikation* von Inhalten kombiniert werden. Die in den Meta-Tags eines Informationsobjekts von den Autoren gepflegten Klassifikationsmerkmale stehen den Nutzern bei einer Suche zur Verfügung und verfeinern beispielsweise eine Volltextsuche.

2.4. Wissensmanagement

2.4.1. Einordnung und Abgrenzung

Der Übergang vom Industrie- zum Informationszeitalter ist durch den Wechsel von körperlichen zu geistigen Tätigkeiten gekennzeichnet. In vielen Unternehmen ist Wissen ein wesentlicher oder ausschliesslicher Faktor bei der Erbringung von Dienstleistungen oder der Herstellung von Produkten. Seit Mitte der neunziger Jahre beschäftigt Forscher und Praktiker die Gestaltung und das Management der Ressource Wissen (‚Knowledge') im betrieblichen Umfeld.

Zielsetzung der verschiedenen Wissensmanagementmodelle ist die Erreichung dauerhafter Wettbewerbsvorteile (s. [Thiesse 2001], [von Krogh et al. 2000], [Probst et al. 1999], [Wiig 1999], [Demarest 1997], [Schüppel 1996], [Nonaka/Takeuchi 1995]). Zwecke und Vorgehensweisen für das Wissensmanagement resultieren aus der Definition des Begriffs ‚Wissen' (s. [Grover/Davenport 2001, 6], [McAdam/McCreedy 1999, 91], [Fahey/Prusak 1998, 265], [Demarest 1997, 2], [von Krogh/Roos 1995, 56]).

[Gebert 2003] liefert in seiner Arbeit eine Übersicht ausgewählter Definitionen:

Autor	Definition
[Senge 1990], [Kofmann/Senge 1993, 15]	"Knowledge is [...] capacity for effective action."
[Drucker 1994, 18]	"Knowledge is information that... [makes] an individual or an institution capable of different or more effective action."
[Nonaka 1994, 15]	"Knowledge is justified true belief that increases an entity's capacity for effective action."
[Wiig 1995, 19]	"Knowledge consists of truths and beliefs, perspectives and concepts, judgements and expectations, methodologies and know-how. Knowledge is accumulated, organized, [...] integrated and held over longer periods to be available(,) to be applied(,) to handle specific situations and problems."
[Sveiby 1997]	"I define knowledge as a capacity to act."
[Probst et al. 1999, 37]	„Die Nutzung von vernetzten Informationen in einem bestimmten Handlungsfeld."
[Franken/Gadatsch 2002, 4]	„Die Repräsentation der Welt – in Form von mentalen Mustern (Schemata) – die die Fähigkeit zum Denken und Handeln (das Handlungspotenzial) eines Menschen oder allgemeiner einer Handlungseinheit bestimmen."

Tabelle 2-4: Eine Auswahl von Definitionen des Begriffs ‚Wissen' [Gebert 2003, 46]

Sowohl die angelsächsischen als auch die deutschsprachigen Abgrenzungen beschreiben Wissen (‚Knowledge') durch eine handlungs- und umsetzungsorientierte Zielsetzung (‚Effective Action'), bei der Menschen vernetzte Informationen nutzen.

Die folgenden Abschnitte definieren den Zusammenhang von Wissen und Informationen für die vorliegende Arbeit und zeigen deren Anwendung in einem im Kontext des Business Engineering entstandenen Wissensmanagementmodells.

2.4.2. Wissensrepräsentation durch Informationsobjekte

Wissen zeichnet sich durch das Verständnis eines Sachverhalts (‚Know-why'), das Vermögen zu effektivem Handeln (‚Know-how') oder die Einsicht der Konsequenzen von Entscheidungen (‚Know-what') menschlicher Individuen aus (s. [Sanchez 1997], [Zahn et al. 2000]). Die Existenz von Wissen ist damit an Personen gebunden und bezieht seinen Wert aus der kontinuierlichen Reflektion, Kommunikation und Weiterentwicklung innerhalb und zwischen Individuen [Fahey/Prusak 1998]. Beim Austausch von Wissen versuchen Menschen, Ausschnitte ihrer so genannten mentalen Modelle an andere Menschen zu übertragen [Riempp 2003, 78]. Mentale Modelle sind dabei vereinfachte und auf wesentliche Komponenten reduzierte geistige Vorstellungen bezüglich der Realität.

[Polanyi 1966, 5] unterscheidet implizites und explizites Wissen von Individuen. Implizites Wissen sind unbewusste Teile mentaler Modelle eines Menschen, die sich z.B. in seinen Handlungsweisen oder intuitiven Entscheidungen ausdrücken [s. Riempp 2003, 332]. Explizites Wissen umfasst die Bestandteile eines mentalen

Modells, deren sich ein Individuum bewusst ist und die es gegenüber anderen Individuen beschreiben (explizieren) kann.

Wird (explizites) Wissen von Individuen getrennt, entstehen Informationsobjekte als stark verkürzende Abbildungen des menschlichen Wissens (s. [Riempp 2003, 80], [Gebert 2003, 28]). Sie repräsentieren Wissen im Kontext des explizierenden Individuums, z.B. in Form von elektronischen Dokumenten wie Berichten, Angeboten, Präsentationsfolien aber auch Webseiten, Portalinhalten, E-Mails oder Datenbankeinträgen. Sind das zugrundeliegende mentale Modell und der Kontext des Erstellers eines Informationsobjekts bekannt, so können andere Individuen Rückschlüsse auf das abgebildete Wissen ziehen, indem sie die vorliegenden Informationen interpretieren.

2.4.3. Prozessunterstützung im CKM-Modell

Das im Kompetenzzentrum ‚Customer Knowledge Management' (CKM) des Instituts für Wirtschaftsinformatik der Universität St. Gallen entwickelte CKM-Modell integriert Wissensmanagement in die Geschäftsprozesse von Unternehmen (s. [Kolbe et al. 2003], [Gebert et al. 2003]).

Im Unterschied zu anderen Wissensmanagementmodellen (vgl. [Probst et al. 1999], [Bullinger et al. 1998], [Zahn 1998]) besitzen die Wissensmanagementziele des CKM-Modells keinen eigenen Wertschöpfungsanspruch. Die so genannte Wissenspyramide adressiert in vier Handlungsfeldern (Säulen) ausschliesslich die Anforderungen und die Wertschöpfung von Geschäftsprozessen:

- Das Handlungsfeld *Inhalt* repräsentiert das von Individuen getrennte und in Geschäftsprozessen in Form von Informationsobjekten nutzbare explizierte Wissen (vgl. [Davenport/Marchand 2001], [Fahey/Prusak 1998]). Dabei können Informationsobjekte Wissen direkt repräsentieren (repräsentierende Informationsobjekte) oder auf Wissensträger verweisen (verweisende Informationsobjekte) [Gebert 2003, 48]. Im CKM-Modell bietet das Content Management entsprechende Prozesse und Systeme für die Erstellung, Pflege, Veröffentlichung, Überarbeitung und Archivierung von Inhalten (s. [Büren et al. 2003], [Christ 2002]). Integrierende Informationssysteme wie Portale stellen beispielsweise den Mitarbeitern eines Unternehmens prozessrelevante Inhalte rollenspezifisch zur Verfügung (s. Abschnitt 2.2). Suchmaschinen ermöglichen darüber hinaus eine effiziente Erschliessung von Inhalten aus einer Vielzahl vorhandener Wissensquellen (s. Abschnitt 2.3.3).

- Das Handlungsfeld *Kompetenz* umfasst das von Mitarbeitenden zur Erfüllung ihrer Aufgaben im Rahmen von Geschäftsprozessen benötigte, nichtexplizierbare (implizite) Wissen [s. Leonard 1998, 5]. Kompetenzmanagementprozesse mit den

Leistungen Kompetenzträgerangleichung, -transformation, -identifikation und -zuordnung sowie entsprechende Systeme (z.B. Skill-Management-Systeme und Expertenverzeichnisse) bieten im CKM-Modell entsprechende Handlungsalternativen [s. Gebert 2003].

- Das Handlungsfeld *Zusammenarbeit* adressiert das von Individuengruppen zur Zusammenarbeit benötigte (Gruppen-)Wissen (s. [Nonaka/Konno 1998], [Wenger 1997]). Das CKM-Modell gestaltet und fördert die Zusammenarbeit in Prozessen, Organisationseinheiten und überbetrieblichen Arbeitsgruppen beispielsweise durch Wissensnetzwerke (s. [Enkel et al. 2000], [Köhne et al. 2000]). Applikationen wie E-Mail- und Groupware-Systeme, Screen- und Application-Sharing-Systeme, Awareness / Instant Messaging, Video Conferencing und auch Workflow-Management-Systeme bieten entsprechende technische Unterstützung (s. zur detaillierten Darstellung der Begriffe [Riempp 2003, 210]).
- Das Handlungsfeld *Struktur* organisiert die in Unternehmen vorhandenen Wissensquellen, um die enthaltenen repräsentierenden und verweisenden Informationsobjekte organisationsweit verfügbar und leicht erschliessbar zu machen (s. [Budin 1990], [Kremer et al. 2003a]). Durch Informationsobjekte repräsentiertes Wissen finden und analysieren Mitarbeitende beispielsweise über Navigation und Suche [s. Rosenfeld/Morville 2002]. Sie beurteilen dabei den Kontext der bereitgestellten Informationsobjekte beispielsweise durch die Analyse von Metadaten, welche die Eigenschaften von Informationsobjekten beschreiben bzw. Informationsobjekte klassifizieren. Aufbau, Pflege und Bereitstellung von Klassifikationsmerkmalen adressiert das CKM-Modell beispielsweise mit den Aufgaben des Terminologiemanagements [s. Kremer et al. 2003b]. Portale und Suchmaschinen bilden im CKM-Modell Systeme zur Umsetzung des Handlungsfelds ‚Struktur'. Sie integrieren heterogene Wissensquellen und stellen den Benutzern Funktionen wie Suche, Navigation, Indizierung, Klassifikation und Präsentation zur Erschliessung prozessrelevanten Wissens zur Verfügung (s. Abschnitt 2.3.3).

Abbildung 2-5: Wissenspyramide des CKM-Modells

Die Wissenspyramide des CKM-Modells stellt zur Umsetzung der Handlungsfelder Inhalt, Kompetenz, Zusammenarbeit und Struktur vier aufeinander aufbauende Zielebenen bereit.

- Die erste Anforderung von Geschäftsprozessen ist häufig die Identifikation des für ihre Wertschöpfung wesentlichen Wissens [Gebert 2003, 49]. *Transparenzierung von benötigtem und bestehendem Wissen* in Form von Kompetenzen, Inhalten, Strukturen und Formen der Zusammenarbeit bilden die Grundvoraussetzung aller Wissensmanagementaktivitäten [s. Geib/Riempp 2002, 401].

- Mit der *aktiven Förderung des Wissensaustauschs* streben Prozessverantwortliche in einem zweiten Schritt Effizienzsteigerungen in den Geschäftsprozessen an. Ist relevantes Wissen in Form von Informationsobjekten im Unternehmen bereits vorhanden, so schliesst die aktive Steuerung der Verteilung die auf Basis der Wissenstransparenz ermittelten ‚Wissenslücken'.

- Die durch Wissenstransparenz und Wissensaustausch adressierten Wissensquellen erfordern eine *fortlaufende Pflege und Weiterentwicklung*, die Prozessverantwortliche durch die Entwicklung von aktuell oder künftig benötigtem Wissen steuern.

- Auf der obersten Zielebene der Wissenspyramide stellen die Prozessverantwortlichen die *Effizienz aller Wissensmanagementmassnahmen* sicher. Sie wägen dabei Kosten- und Nutzenaspekte der rechtzeitigen Bereitstellung spezifischer Informationsobjekte für das Unternehmen ab [s. Geib/Riempp 2002, 403].

2.4.4. Beitrag für die Dissertation

Das vorliegende Dissertationsprojekt leitet aus dem Wissensmanagement folgende Erkenntnisse ab:

- *Wissen* ist eine Eigenschaft menschlicher Individuen und an diese gebunden. Es drückt sich beispielsweise im Verständnis eines Sachverhalts, dem Vermögen zu effizientem Handeln oder der Einsicht in die Konsequenzen von Entscheidungen aus.
- *Informationsobjekte* sind stark verkürzte Abbildungen menschlichen Wissens, die von Individuen in Form elektronischer Dokumente erzeugt, gespeichert, verteilt, überarbeitet, gesucht, genutzt und gelöscht werden können. Andere Individuen können mit Hilfe von Informationsobjekten Rückschlüsse auf das ursprüngliche Wissen rekonstruieren, wenn ihnen der Kontext des Erstellers bekannt ist.
- Das *CKM-Modell* integriert Wissensmanagementaktivitäten in die Wertschöpfungskette eines Unternehmens über die vier Handlungsfelder Inhalt, Kompetenz, Zusammenarbeit und Struktur. Beim Management von Inhalten und Strukturen spielen Suchmaschinen mit ihren Funktionen wie Indizierung, Klassifikation, Suche und Navigation eine wesentliche Rolle.
- *Information Retrieval in Portalen* stiftet Nutzen durch eine verbesserte Erschliessbarkeit von Informationsobjekten (s. Abschnitt 2.3.3). Es trägt zur Transparenzierung bestehenden Wissens bei und bildet gemäss dem CKM-Modell die erste und damit wesentliche Grundlage zur Umsetzung von Wissens-management im betrieblichen Umfeld.

2.5. Konsequenzen für eine Information-Retrieval-Methode

Auf Basis der in den Abschnitten 2.1 bis 2.4 gewonnenen Erkenntnisse formuliert die vorliegende Arbeit Konsequenzen für die weitere Vorgehensweise. Eine Methode zur Analyse, Planung, Gestaltung und Implementierung von Information Retrieval in Portalen muss:

- die Gestaltung von Information-Retrieval-Funktionen wie Suche und Navigation an den Anforderungen der geschäftlichen Abläufe der Anwender ausrichten (*Prozessorientierung*),
- die Ausschöpfung von Möglichkeiten der modernen Informationstechnologie durch die Planung von Abläufen für Aufbau und Betrieb organisatorisch verankern (*Prozessmanagement*),
- das Potenzial von Suchmaschinen zur Verbesserung bisheriger unternehmerischer Tätigkeiten und/oder zur Erschliessung neuer Geschäftsfelder ausschöpfen (*Retrieval-Leistungen*),
- die beteiligten Systeme in Form von Portal, Suchmaschine und die inhaltsführenden Anwendungen im Backend aufeinander abstimmen (*Architekturelemente*),
- bei der Integration heterogener Anwendungen deren Struktur bzw. Klassifikationsmerkmale, bestehende Zugriffsbeschränkungen sowie unterschiedliche Inhaltsformate abbilden (*Systemintegration*),
- bei der Umsetzung eines Information-Retrieval-Projekts eine ganzheitliche Betrachtung auf Strategie-, Prozess- und Systemebene gewährleisten (*Ganzheitlichkeit*) sowie
- den systematischen Transformationsprozess durch eine umfassende Berücksichtigung aller Gestaltungselemente der Methodenentwicklung unterstützen (*Methodenelemente*).

Das folgende Kapitel 3 vergleicht bestehende Methoden anhand der oben genannten Anforderungen.

3. Vergleich bestehender Methoden

Methoden bezeichnen planmässige, begründete Vorgehensweisen zur Erreichung definierter Ziele im Rahmen festgelegter Prinzipien [s. Balzert 1996, 36]. Sie unterstützen z.B. die Realisierung eines Informationssystems durch die Zerlegung eines Problems in Teilbereiche (Hierarchisierung) oder die Entwicklung von Komponenten, die über definierte Schnittstellen mit ihrer Umwelt kommunizieren (Modularisierung).

Der Vorschlag einer eigenen Methode für Information Retrieval in Portalen ist nur dann sinnvoll, wenn keine Methode vorhanden ist oder existierende Methoden in Bezug auf die an sie gestellten Anforderungen unvollständig sind. Auf Basis der in Abschnitt 2.5 abgeleiteten Kriterien beurteilt dieses Kapitel existierende Portalmethoden in Anlehnung an [Legner 1999, 73] anhand des inhaltlichen Beitrags (inhaltliche Charakteristika) und der definierten Methodenelemente (strukturelle Charakteristika). Die Bewertungskriterien sind in Tabelle 3-1 zusammengefasst.

Kriterium	Beschreibung
Inhaltliche Charakteristika	
Prozessorientierung	Berücksichtigung der Anforderungen von wertschöpfenden Geschäftsprozessen
Prozessmanagement	Gestaltung von Pflegeprozessen für den Aufbau und Betrieb einer Information-Retrieval-Lösung
Retrieval-Leistungen	Darstellung von Such- und Navigationsfunktionen anhand ihrer Einsatzmöglichkeiten, Zielsetzungen und Restriktionen
Architekturelemente	Berücksichtigung der Abhängigkeiten von Portal- und Retrieval-Architekturelementen, um die mit diesen Wechselwirkungen verbundenen Potenziale und Risiken adressieren zu können
Systemintegration	Darstellung der Integration in bestehende Informationssysteme, um vorhandene Applikationen anzubinden
Ganzheitlichkeit	Abdeckung von Strategie-, Prozess- und Systemebene, um bei der Einführung von Retrieval in Portalen die erfolgreiche Transformation aller relevanten Bereiche herbeizuführen
Strukturelle Charakteristika	
Methodenelemente	Darstellung von Methodenbestandteilen im Sinne des Method Engineering, z.B. Vorgehen, Rollen, Techniken, Ergebnisse und Metamodell

Tabelle 3-1: Bewertungskriterien

Die folgenden Abschnitte skizzieren zunächst die Auswahlkriterien für die Untersuchung bestehender Methoden. Diese werden anschliessend analysiert und anhand der oben genannten Kriterien bewertet. Die Konstruktion eines integrierten Metamodells für Information Retrieval in Portalen formalisiert daraus abgeleitete Konsequenzen für die eigene Arbeit.

3.1. Auswahlkriterien

Für den Vergleich wurden bestehende Ansätze ausgewählt, die Information Retrieval in Portalen adressieren oder vergleichbare Informationssystemarchitekturen aufweisen. Sie müssen zusätzlich:

- Methoden bzw. Methodenbestandteile enthalten,
- die Integration bestehender Anwendungen berücksichtigen,
- sowohl Planungs- als auch Implementierungsaspekte abdecken und
- umsetzungsorientiert sein bzw. in der Praxis angewendet werden.

Der Vergleich basiert auf der Auswertung von Publikationen (Bücher, Zeitschriftenartikel, Konferenzbeiträge, Vortragspräsentationen) zu den einzelnen Methoden. Die Analyse wurde im November 2003 abgeschlossen. Sie verwendet ausschliesslich veröffentlichte bzw. zur Veröffentlichung freigegebene Materialien und umfasst die in Tabelle 3-2 aufgeführten Ansätze.

Methode	Kurzbeschreibung
Extranet Development Lifecycle [Bayles 1998]	Ansatz zur Entwicklung eines Extranets, das interne mit externen Geschäftseinheiten auf Basis einer web-basierten Applikation verbindet
Engineering Enterprise Portals Framework [Finkelstein/Aiken 2000]	Ansatz zur Planung, Implementierung und Anwendung von Portalen auf Basis von XML
Methode zur Konzeption von Intranets [Kaiser 2000]	Konzepte und Lösungen zum Einsatz von Internet-Technologie zur Informationsunterstützung von Geschäftsprozessen innerhalb eines Unternehmens
Implementierung von Unternehmensportalen [Bauer 2001]	Ansatz für die Entwicklung eines Unternehmensportals, das interne mit externen Geschäftspartnern verbindet
Corporate Portals [Collins 2001]	Ansatz zur Ableitung von Portallösungen durch die Analyse von Portalfunktionen
Collaboration Portale [Puschmann 2003]	Ansatz zur Umsetzung von Portalen, die inner- und überbetriebliche Applikationen verbinden und Anwendern einen rollenbasierten Zugriff auf prozessrelevante Informationen gewähren
Integrierte Wissensmanagement-Systeme [Riempp 2003]	Ansatz zur Gestaltung und Integration von Informationssystemen für Wissensmanagement mit der Zielsetzung, eine effektive Gestaltung und Ausführung von Geschäftsprozessen in Organisationen zu unterstützen

Tabelle 3-2: Übersicht der analysierten Methoden

Der folgende Abschnitt analysiert diese Ansätze und bewertet sie hinsichtlich der in Tabelle 3-1 aufgeführten Kriterien.

3.2. Analyse

Die folgende Gliederung der analysierten Ansätze nach *Fokus*, *Architektur- und Gestaltungselemente*, *Strukturelle Elemente*, *Inhaltliche Elemente* und *Bewertung* orientiert sich an den Darstellungen von [Puschmann 2003].

3.2.1. Extranet Development Lifecycle [Bayles 1998]

Fokus: Im Mittelpunkt des Ansatzes von [Bayles 1998] steht die Entwicklung eines Portals, das interne mit externen Geschäftseinheiten auf Basis einer web-basierten Applikation verbindet. Bayles berücksichtigt dabei den gesamten Lebenszyklus dieses so genannten Extranets, z.B. neben der eigentlichen Erstellung auch die Wartung einer Applikation. Das Buch richtet sich an IT-Professionals wie Softwareentwickler oder -architekten sowie Fachbereiche, die sich mit dieser Themenstellung befassen.

Architektur und Gestaltungselemente: Das Buch enthält keine grafische Architekturdarstellung, führt jedoch die folgenden Gestaltungselemente (‚components') in Form einer Drei-Schicht-Architektur auf [vgl. Balzert 2001, 697]:

- Der *Data and System Services Tier* beinhaltet Infrastrukturelemente wie Netzwerkverbindung (‚connectivity'), Datenbanken (‚databases'), Sicherheit (‚security') und Zugriffskontrolle (‚access control').
- Der *Application Tier* setzt sich aus Groupware-Anwendungen und Applikationen zur verteilten Dokumentenbearbeitung zusammen. Dabei stehen Aspekte der Versionskontrolle von Inhalten im Vordergrund. Als weiterer Bestandteil stellt die ‚Dynamic Content Assembly' den Extranet-Benutzern Inhalte aus unterschiedlichen Quellen zur Verfügung.
- Der *Presentation Tier* enthält die Benutzerschnittstelle, aber auch Personalisierungsfunktionen, z.B. zur Veränderung des Layouts oder der angezeigten Inhalte.

Strukturelle Elemente: Das Vorgehensmodell besteht aus 7 Phasen und ist in 26 Schritte (Aktivitäten) detailliert (s. Abbildung 3-1). Das Ergebnismodell umfasst die im Laufe eines Projekts zu erstellenden Dokumente. Eine genaue Beschreibung, wie diese erstellt werden sollen, fehlt hingegen. Die organisatorischen Rollen sind für jeden Teilschritt angeführt, allerdings ist kein eigenes Rollenmodell ausgewiesen. Die Gestaltungselemente eines Extranets werden definiert, ohne deren Beziehungen untereinander aufzuzeigen. Ein Metamodell ist nicht dargestellt.

Inhaltliche Elemente: Die Methode definiert ihre Anforderungen auf Basis identifizierter Probleme in der internen und externen Kommunikation eines Unternehmens. Als Grundlage für die Gestaltung der Extranet-Architektur dient die

Analyse der IST-Situation eines Unternehmens in Form einer IS-Architektur. Darauf aufbauend definieren die Verantwortlichen eine SOLL-Architektur für das Extranet, indem sie die Komponenten, Funktionalitäten, Inhalte, die Benutzerschnittstelle und Gestaltungsrichtlinien bestimmen. Der Aufbau von Prototypen soll in einem nächsten Schritt dabei helfen, die Akzeptanz zu testen und Fehler bereits am Anfang zu beheben. Der Prototyp wird in der Implementierungsphase sukzessive um Funktionalitäten erweitert und mit technischen Anforderungen, wie Sicherheit, unterstützte Sprachen, Integration von Applikationen, benötigte Bandbreiten und Antwortzeitverhalten etc. abgeglichen. Die Bereitstellung des Produktivsystems ergänzt, im Sinne des Lifecycle-Ansatzes, eine Betriebs-, Mess- und Wartungsphase.

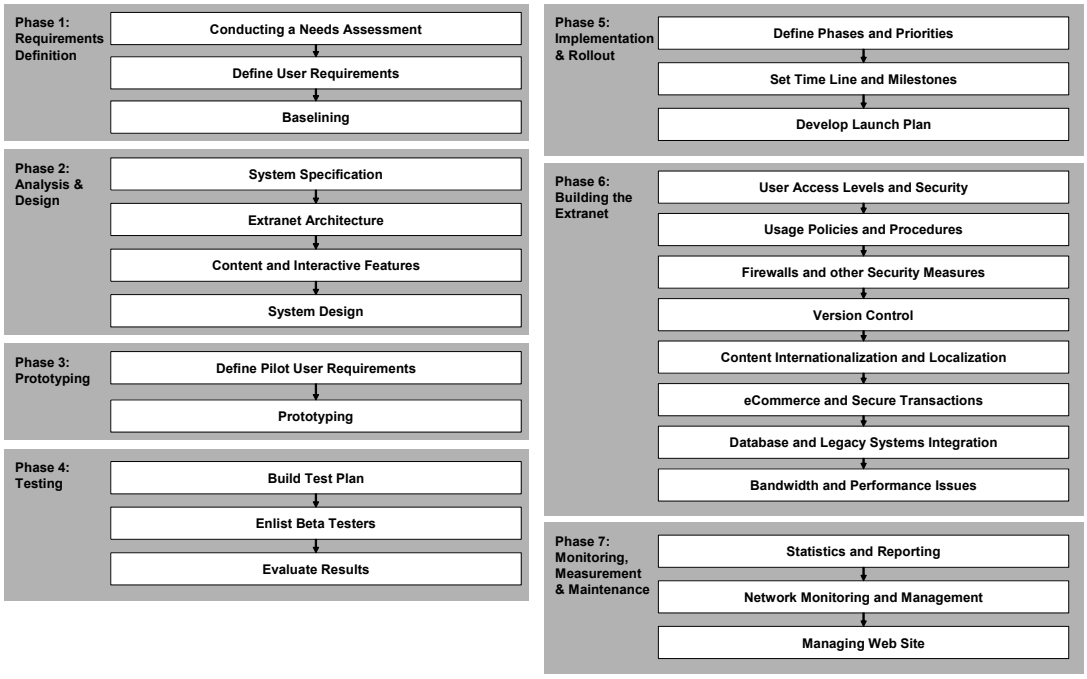


Abbildung 3-1: Vorgehensmodell Extranet Development Life Cycle [Bayles 1998, 62]

Bewertung: Die Stärken dieses Ansatzes liegen einerseits im frühen Einbezug von Mitarbeitern, Extranetnutzern und Management, dem Einsatz von Prototypen sowie der Beteiligung von Testanwendern. Dadurch adressiert die Methode eine enge Zusammenarbeit mit den unterschiedlichen Interessengruppen. Andererseits betont das Vorgehensmodell durch ‚Monitoring, Measurement & Maintenance' die Phase des Betriebs. Da die Architekturplanung auf Infrastrukturkomponenten begrenzt ist, weist die Methode inhaltliche Lücken auf. Für Information Retrieval relevante funktionale Aspekte werden daher nur am Rande behandelt. Die Methode macht weiterhin keine Aussagen zur Gestaltung von Navigation und Suche. Die strukturellen Schwächen sind das fehlende Meta- und Rollenmodell.

3.2.2. Engineering Enterprise Portals Framework [Finkelstein/Aiken 2000]

Fokus: Das Engineering Enterprise Portals Framework von [Finkelstein/Aiken 2000] umfasst die Planung, Implementierung und Anwendung von Portalen. Portalinhalte

werden auf Basis der Extensible Markup Language (XML) bereitgestellt. Der Schwerpunkt liegt dabei auf der Klassifikation von Inhalten mit Metadaten, die beispielsweise bei einer Portalsuche verwendet werden. Durch den technischen Fokus adressiert das Buch überwiegend Informationssystemarchitekten und Softwareentwickler.

Architektur und Gestaltungselemente: Abbildung 3-2 zeigt die Architektur des vorliegenden Rahmenwerks.

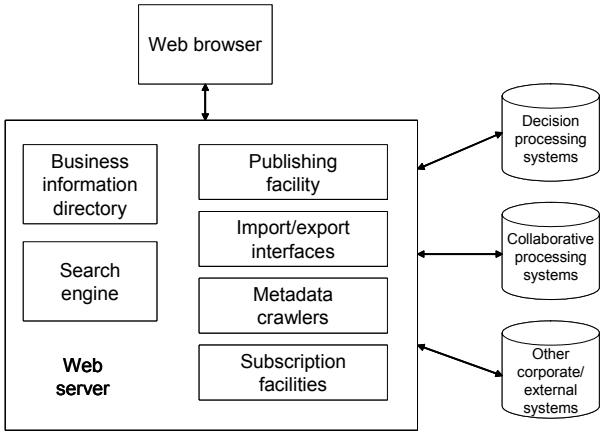


Abbildung 3-2: Enterprise Portal Architecture [Finkelstein/Aiken 2000, 491]

[Finkelstein/Aiken 2000, 491] nennen folgende Portalbestandteile (‚components'):

- Geschäftsrelevante Informationen sind in unterschiedlichen Anwendungen enthalten, z.B. zur Entscheidungsunterstützung (‚*Decision processing systems*'), für Kommunikation und Zusammenarbeit (‚*Collaborative processing systems*') oder weitere unternehmensinterne und -externe Systeme (‚*Other corporate / external systems*').

- Eine Suchmaschine (‚*Search engine*') bindet diese Applikationen an das Portal an, so dass die Komponente für die Indizierung (‚*Metadata crawler*') entsprechende Informationen aus den Anwendungen extrahieren und in einem Verzeichnis (‚*Business information directory*') speichern kann. Alternativ können Inhalte auch direkt aus anderen Applikationen in das Verzeichnis im- und exportiert werden (‚*Import/export interfaces*').

- Such- und Navigationsfunktionen für das ‚Business information directory' stellt den Portalbenutzern ein Webserver zur Verfügung. Die Anwender greifen dabei mit einem Webbrowser auf die Inhalte und Services des Portals zu, z.B. durch das Abonnement (‚*Subscription facility*') bestimmter Inhalte. Diese werden dann personalisiert bereitgestellt (‚*Publishing facility*').

Strukturelle Elemente: Das Vorgehensmodell besteht aus 3 Phasen mit insgesamt 21 Schritten (s. Abbildung 3-3). Die einzelnen Aktivitäten werden sehr ausführlich

beschrieben und weisen jeweils Ergebnisdokumente auf. Meta- und Rollenmodell sind nicht explizit dargestellt.

Inhaltliche Elemente: Die Methode beginnt mit der Erfassung der Benutzeranforderungen aus fachlicher Sicht, z.B. auf Basis bestehender Geschäftsabläufe. Aus diesen werden Datenmodelle und Metadaten abgeleitet, welche die Informationsbedürfnisse der Anwender repräsentieren. Eine Bewertung möglicher Lösungen erfolgt anhand des Unterstützungspotenzials für Geschäftsprozesse, für die Messgrössen definiert werden. In einem Kick-off Meeting werden die vom Projektteam erhobenen Anforderungen mit den Anwendern verifiziert. Aus dem logischen Architekturvorschlag entwirft das Projektteam schliesslich das physische Design des Portals. In der zweiten Phase beginnt die Entwicklung der Portallösung. Den Ausgangspunkt bildet eine Analyse bestehender Applikationen in Bezug auf ihre Inhalte, Datenstrukturen und Metadaten. Darauf aufbauend zeigt das Vorgehensmodell in der dritten Phase, wie Portale mit Hilfe von Metadaten aufgebaut werden können. Die Methode sieht eine Analyse des Nutzungsverhaltens der im Portal bereitgestellten Inhalte vor. Diese Ergebnisse dienen zur Definition von Verbesserungsmöglichkeiten und schliessen das Vorgehensmodell ab.

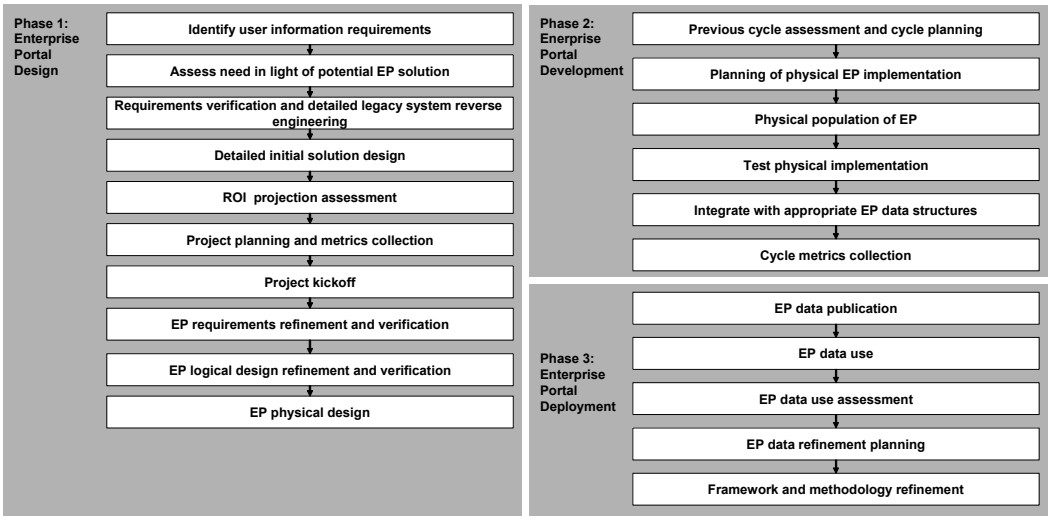


Abbildung 3-3: Vorgehensmodell für das Engineering Enterprise Portal Framework [Finkelstein/Aiken 2000, 172]

Bewertung: Eine Stärke der Methode liegt im umfassenden Vorgehensmodell. Der eigentliche Schwerpunkt liegt jedoch im technischen Bereich. Die Methode zur Rekonstruktion von Datenmodellen bestehender Systeme (so genanntes Reverse Engineering), bietet die Möglichkeit, Inhalte in neue, umfassende Applikationen zu migrieren. Dadurch scheiden Praxisfälle aus, bei denen Alt-Systeme trotz Integration in einem Portal erhalten bleiben sollen. Die Fokussierung auf XML als zentrales Dokumentenformat schränkt die Anwendbarkeit ein, solange die Konvertierungsprobleme mit in der Praxis verbreiteten Formaten, wie z.B. MS Office nicht gelöst sind. Die Methode lässt zudem die Explikation von Meta- und Rollenmodell vermissen.

3.2.3. Methode zur Konzeption von Intranets [Kaiser 2000]

Fokus: Die Methode von [Kaiser 2000] gibt Handlungsanweisungen zur Entwicklung von unternehmensinternen Portalen. Zielsetzung dieser Intranets ist die Unterstützung von Geschäftsprozessen mit Informationen. Sie beinhaltet sowohl konzeptionell-logische als auch technologiespezifische Aspekte so genannter Web-Applikationen. Die Arbeit richtet sich an Wissenschaftler und Praktiker, z.B. Business Engineers, Portalprojektleiter, IS-Manager und Berater.

Architektur und Gestaltungselemente: Kaiser stellt die Gestaltungselemente des Portals in Form eines Metamodells dar. Die gewählte Sicht zeigt die Applikationsarchitektur, welche die für die Bereitstellung und Nutzung von Informationen benötigten Bestandteile verdeutlicht.

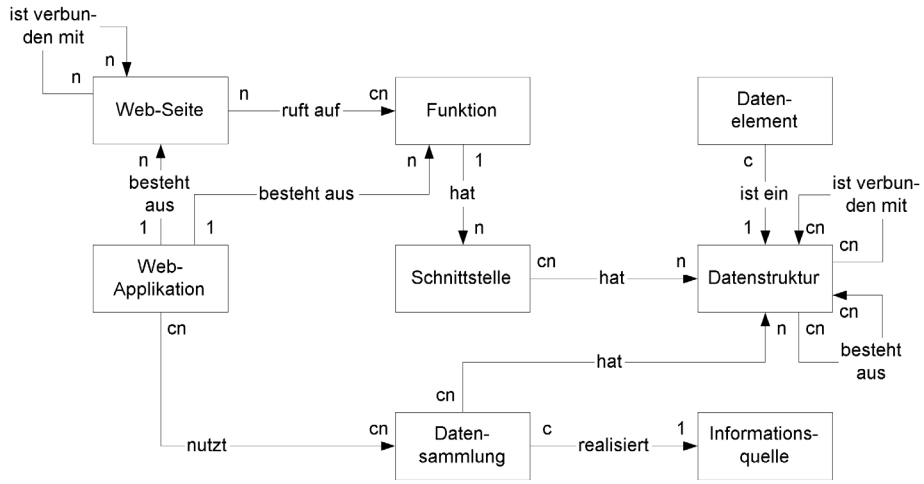


Abbildung 3-4: Metamodell der Portalapplikationsarchitektur nach [Kaiser 2000, 87]

- Das Portal, als *Webapplikation*, steht im Mittelpunkt der Applikationsarchitektur. *Webseiten* bilden die Benutzerschnittstelle der Webapplikation. Sie sind über Hyperlinks untereinander verbunden und erlauben dadurch beispielsweise die Navigation im Portal.

- *Funktionen* realisieren die Anwendungslogik einer Webapplikation. Sie werden aus der Benutzerschnittstelle heraus aufgerufen, z.B. als Suchanfrage in einen Webbrowser.

- *Datensammlungen* und *Schnittstellen* sind durch *Datenstrukturen* definiert und erlauben den Zugriff von Funktionen und Web-Applikationen. *Datenelemente* sind dabei Teile von Datenstrukturen, die die eigentlichen Portalinhalte repräsentieren.

Strukturelle Elemente: Das Vorgehensmodell besteht aus 6 Techniken (Phasen) mit insgesamt 26 Aktivitäten (s. Abbildung 3-5). Diese werden in 23 Ergebnisdokumenten festgehalten. Die Methode verfügt ausserdem über ein vollständiges expliziertes Rollenmodell und ein Metamodell mit vier verschiedenen Sichten (Organisations-, Informations- und Applikationsarchitektur sowie Infrastruktur).

Inhaltliche Elemente: Die *erste Phase* analysiert Webtechnologien hinsichtlich ihrer Potenziale zur Optimierung bestehender Prozesse. Hierfür werden konkrete Projektvorschläge für einzelne Prozesse anhand definierter Kriterien (Struktur, Informationsbedarf und -struktur, Zentralisierungsgrad, Aufgabenträger) erstellt und bewertet. In der *zweiten Phase* wird der Zentralisierungsgrad für das Intranet festgelegt, der die Gestaltung der Aufbau- und Ablauforganisation massgeblich beeinflusst. Die Definition der Prozessarchitektur, einer Rahmenorganisation sowie die Bestimmung von Richtlinien schliessen diese Phase ab. Die *dritte Phase* der Informationsstrukturierung erhebt die bestehenden Informationsquellen und deren Struktur, definiert Informationsobjekte sowie deren Zuordnung zu Kategorien (Klassifikation) und Intranetrollen. In der *vierten Phase* erhebt der Projektverantwortliche die bestehende Applikationsarchitektur. Auf Basis der bisherigen Anforderungen spezifiziert er die geplante Webapplikation in den Bereichen Datenstruktur, Navigationsstruktur, Oberflächengestaltung und Funktionsbeschreibung. Die Planung des Applikationsmanagements schliesst diese Phase ab. Die Infrastrukturplanung als *fünfte Phase* erfolgt analog zum Applikationsentwurf, konzentriert sich jedoch auf die rein informationstechnologischen Gestaltungsmerkmale. Die Sicherheitsplanung ordnet abschliessend den identifizierten Bedrohungen geeignete Sicherheitsmechanismen zu.

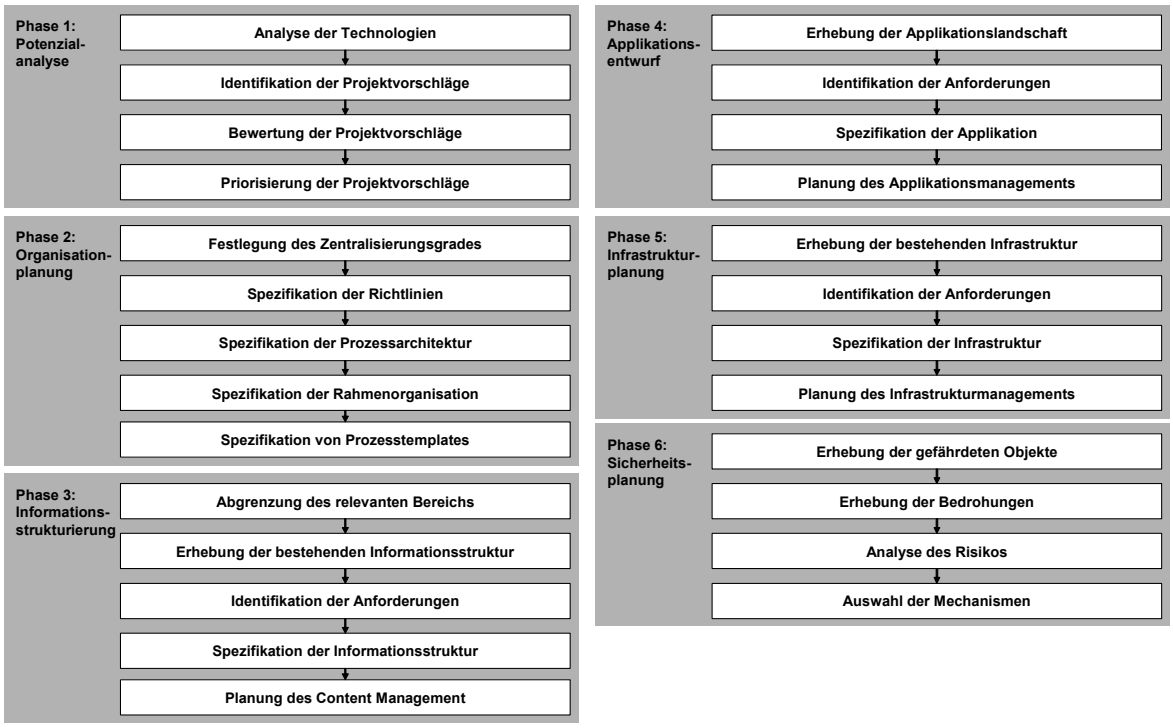


Abbildung 3-5: Vorgehensmodell der Methode zur Konzeption von Intranets [Kaiser 2000, 117]

Bewertung: Die Methode liefert wertvolle Hinweise für die Entwicklung von Information Retrieval in Portalen. Die Analyse bestehender Prozesse hinsichtlich ihres Informationsbedarfs, die Erfassung bestehender Informationsquellen und der darauf basierende Entwurf einer neu zu implementierenden Informations- und Applikationsarchitektur sind dabei hervorzuheben. Da die Methode umfassend auf die Konzeption von Intranets ausgelegt ist, bleiben jedoch spezifische Herausforderungen des Information Retrieval undokumentiert, z.B. die Berücksichtigung von Zugriffsbeschränkungen heterogener Datenbestände bei einer Portalsuche. Auch die Möglichkeiten und Restriktionen der heutigen Portal- und Suchmaschinentechnologien kann dieser Ansatz nicht berücksichtigen.

3.2.4. Implementierung von Unternehmensportalen [Bauer 2001]

Fokus: Im Mittelpunkt des umsetzungsorientierten Ansatzes von [Bauer 2001] steht die Entwicklung eines Unternehmensportals, das interne mit externen Geschäftspartnern verbindet. Das Buch zeigt Geschäftsmodelle, Ziele und Funktionen von Portalen auf. Es adressiert Praktiker und enthält Handlungsanweisungen, wie Portale aufgebaut und eingesetzt werden können, um bestehende Prozesse zu verbessern oder neue Geschäftsfelder zu erschliessen.

Architektur und Gestaltungsmerkmale: Das Werk enthält keine explizit dargestellte Portalarchitektur, weist jedoch Gestaltungselemente auf. Diese als wichtigste Portalfunktionen bezeichneten Bestandteile sind [Bauer 2001, 38]:

- *Personalisierung*. Anpassung von Inhalten und Layout durch den Benutzer,
- *Benutzerverwaltung und Sicherheitsservices*. Verwaltung von Rollen, Rechten, Gruppen und Benutzerinformationen,
- *Dynamische Inhalte und Webpublishing*. Datenbankbasierte Informationsbereitstellung,
- *Externe Webanwendungen und Integration von Unternehmensanwendungen*. Einbindung existierender unternehmensinter und externer Applikationen.

Strukturelle Elemente: Das Vorgehensmodell besteht aus einer Phase und ist in 4 Schritte (Aktivitäten) unterteilt (s. Abbildung 3-6). Die einzelnen Aktivitäten werden beschrieben, bleiben aber auf einer Übersichtsebene. Die Methode zeigt einige Ergebnisse in Form von Tabellen und Grafiken auf, beschreibt aber keinen Weg, wie diese zu erarbeiten sind. Die zugehörigen Rollen- und Metamodelle sind nicht explizit dargestellt.

Inhaltliche Elemente: Der Gestaltungsbereich der Methode beginnt mit der Potenzialanalyse, in deren Rahmen die Kernprozesse und -systeme ermittelt und auf Verbesserungen hin analysiert werden. Diesen beiden Phasen der Identifikation und Bewertung schliesst sich die Phase der Implementierung des Portals an. Dabei liegt der Fokus auf der nicht mehr erhältlichen Workplace-Technologie von SAP. Allerdings ist das Vorgehensmodell generisch gehalten, so dass es sich auch auf andere Portalapplikationen übertragen lässt. In der Phase der Implementierung ordnet der Portalinitiator den Mitarbeitern konkrete Rollen zu. Auf der Basis dieses Rollenmodells sowie der anfangs definierten Prozess- und Systemarchitektur wird anschliessend das Portal konfiguriert. Dabei wird jeder zukünftige Portalbenutzer hinsichtlich seiner Aufgaben und unterstützten Applikationen analysiert. Die daraus abgeleiteten Content-Applikationen werden schliesslich für das Portal angepasst und integriert.

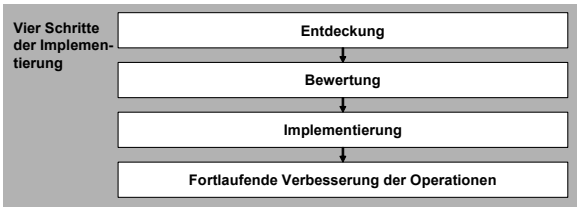


Abbildung 3-6: Vorgehensmodell Implementierung von Unternehmensportalen [Bauer 2001, 178]

Bewertung: Das Vorgehensmodell hebt die Analyse der Kundenprozesse durch den frühzeitigen Einbezug der Anwender hervor. Die dadurch entwickelte Portalarchitektur soll die Akzeptanz bei den Benutzern erhöhen. Der Ansatz verfolgt den Entwurf und die Implementierung in Form von modularen Komponenten. Dadurch soll eine hohe Wiederverwendungsquote in weiteren Portalen erreicht

werden. Den Erfolg der Lösung soll schliesslich die Unterstützung des Top-Managements garantieren. Trotz dieses Ansatzes bietet das Vorgehensmodell nur wenig Unterstützung bei der Planung und Implementierung von Retrieval in Portalen. Such- und Navigationsmöglichkeiten werden wie die anderen so genannten ‚Portaldienste' zwar aus Endbenutzersicht beschrieben, ohne jedoch die Abhängigkeiten und Wechselwirkungen der einzelnen Gestaltungselemente zu detaillieren. Das abschliessende Kapitel Technik gibt wenig Auskunft über die Umsetzungsmöglichkeiten und -herausforderungen der am Anfang beschriebenen Portallösungen. So bleibt beispielsweise die Anpassung und Integration von Content-Applikationen unklar. Darüber hinaus sind Methodenbestandteile wie Techniken, Aktivitäten oder Metamodell nicht explizit aufgeführt, was die Anwendung in anderen Fällen erschwert.

3.2.5. Corporate Portals [Collins 2001]

Fokus: Der Ansatz von [Collins 2001] definiert eine Portal-Lösung durch die Analyse von Portalfunktionen bzw. Gestaltungselementen. Corporate Portals adressiert mit der Erstellung einer Wirtschaftlichkeitsbetrachtung (‚Business Case') für den unternehmensinternen Einsatz von Portalen fachliche Entscheidungsträger.

Architektur und Gestaltungsmerkmale: Das Buch führt keine explizit dargestellte Portalarchitektur auf, stellt jedoch Gestaltungselemente dar. Diese als so genannte Portal Software Functions bezeichneten Bestandteile sind nach [Collins 2001, 37]:

- *Data Points and Integration* erlaubt einem Portalbenutzer den Zugriff auf eine Vielzahl unternehmensinterner und -externer Informationsquellen.
- *Taxonomy* bietet unternehmensspezifische Kategorien für Navigation und Suche, die sich an den betrieblichen Tätigkeiten ausrichten bzw. diese unterstützen.
- *Search capabilities* umfasst Komponenten, die Portalbenutzern das Finden von Informationen in Unternehmensanwendungen, im Internet und in externen Suchmaschinen ermöglicht.
- *Help features* unterstützen die Benutzer bei der Arbeit mit dem Portal, z.B. durch Hilfefenster.
- *Content Management* ermöglicht die Publikation und Verteilung von Inhalten.
- *Process and Action* sind Komponenten, die eine elektronische Vorgangs-bearbeitung (Workflows) unterstützen, z.B. durch die automatische Weiterleitung einer Aufgabe an den nächsten Bearbeiter.

- *Collaboration and Communication* erlauben die Zusammenarbeit und den Informationsaustausch mit weiteren Portalbenutzern, z.B. über E-Mail oder Diskussionsforen.
- *Personalization* ermöglicht jedem Benutzer die individuelle Organisation seines Arbeitsbereichs im Portal, z.B. durch das Hinzufügen, Verschieben oder Entfernen von Portlets.
- *Presentation* umfasst Portalkomponenten, die die Benutzerschnittstelle realisieren.
- *Administration* dient zur Konfiguration und Wartung des Portals, z.B. durch Datensicherung.
- *Security* beinhaltet die Verwaltung von Benutzern, Gruppen, Authentifizierungs- und Autorisierungsinformationen.

Strukturelle Elemente: Das Vorgehensmodell umfasst 6 Phasen (‚fundamental steps') mit insgesamt 15 Aktivitäten (s. Abbildung 3-7). Diese sind ausführlich beschrieben. Die Methode bietet zur Erarbeitung des Business Case ein Dokumentationsmodell, Checklisten und Rollenbeschreibungen.

Inhaltliche Elemente: Der Gestaltungsbereich der Methode beginnt mit der Portalstrategie (‚Corporate Portal Mission'), die ein Projektteam in der *ersten Phase* aus der Unternehmensstrategie ableitet. Der Projektleiter skizziert Portalfunktionen, deren Anwendung zu einer Verbesserung der Abläufe innerhalb der Organisation führen soll. Das Projektteam identifiziert anschliessend Experten aus den relevanten Geschäftsbereichen. Die *zweite Phase* beginnt mit der Analyse der vorhandenen Infrastruktur wie E-Mail-System, betriebswirtschaftliche Standardsoftware, vorhandene Portal-Anwendungen oder auch Suchmaschinen. Durch Umfragen oder Interviews werden dabei Mitarbeiter der IT-Abteilung involviert. Die Analyse der geschäftlichen Anforderungen und Informationsbedürfnisse startet die *dritte Phase*. Dabei werden von den Experten der einzelnen Geschäftsbereiche beispielsweise Anforderungen hinsichtlich der zu integrierenden Anwendungen, der Sicherheit aber auch der Navigations- und Suchfunktionen erhoben. Die *vierte Phase* umfasst das ‚Corporate Portal Storyboard', das Entwürfe für die Benutzerschnittstelle oder der Beschreibung von Arbeitsabläufen des zu realisierenden Systems mit den Anwendern simuliert. Kosten- und Nutzenaspekte adressiert die *fünfte Phase*. Das Unterstützungspotenzial liegt dabei in der Umsatzsteigerung, Kostenreduktion oder höherer Kundenzufriedenheit. Die *sechste und letzte Phase* schliesst die Methode mit einem Projektplan zur Umsetzung des Unternehmensportals ab.

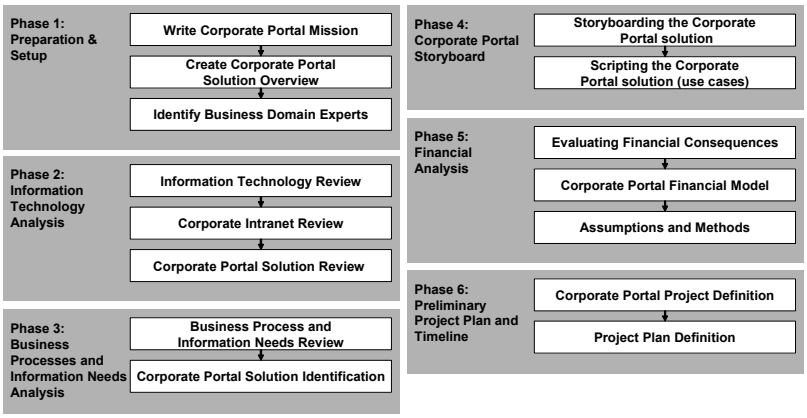


Abbildung 3-7: Vorgehensmodell Corporate Portals (nach [Collins 2001])

Bewertung: ‚Corporate Portals' bietet ein umfassendes Vorgehens- und Dokumentationsmodell zur Erstellung eines Business Case für ein Unternehmensportal. Checklisten zeigen die Orientierung der Methode an den geschäftlichen Anforderungen (‚Business process and information needs analysis'). Beispiele erleichtern die Verständlichkeit und Anwendbarkeit in der Praxis. Der gewählte Fokus adressiert primär fachliche Entscheidungsträger und deckt die Strategie- und Prozessebene ab. Das Buch enthält für diese Zielgruppe auch Beschreibungen von Portalfunktionen wie Suche, Navigation, Taxonomie oder Integration. Charakteristika der Implementierung oder Abhängigkeiten dieser Gestaltungselemente untereinander werden jedoch nicht aufgezeigt. Obwohl sich der Ansatz grundsätzlich an Geschäftsprozessen orientiert, bietet er keinen Vorschlag zur Nutzenmessung.

3.2.6. Collaboration Portale [Puschmann 2003]

Fokus: Der Ansatz von Puschmann sieht Portale als zentrale Benutzerschnittstellen im Unternehmen. Er fokussiert dabei auf so genannte Collaboration Portale, die inner- und überbetriebliche Applikationen verbinden und den Anwendern einen rollenbasierten Zugriff auf prozessrelevante Informationen gewähren. ‚Collaboration Portale' zeigt, welche Gestaltungselemente ein hierfür geeignetes Architektur- und Integrationsmodell beinhaltet und enthält Methodenbestandteile zur Umsetzung.

Architektur und Gestaltungsmerkmale: Die Architektur für ‚Collaboration Portale' fasst Gestaltungselemente für Information Retrieval im Bestandteil ‚Navigation' zusammen.

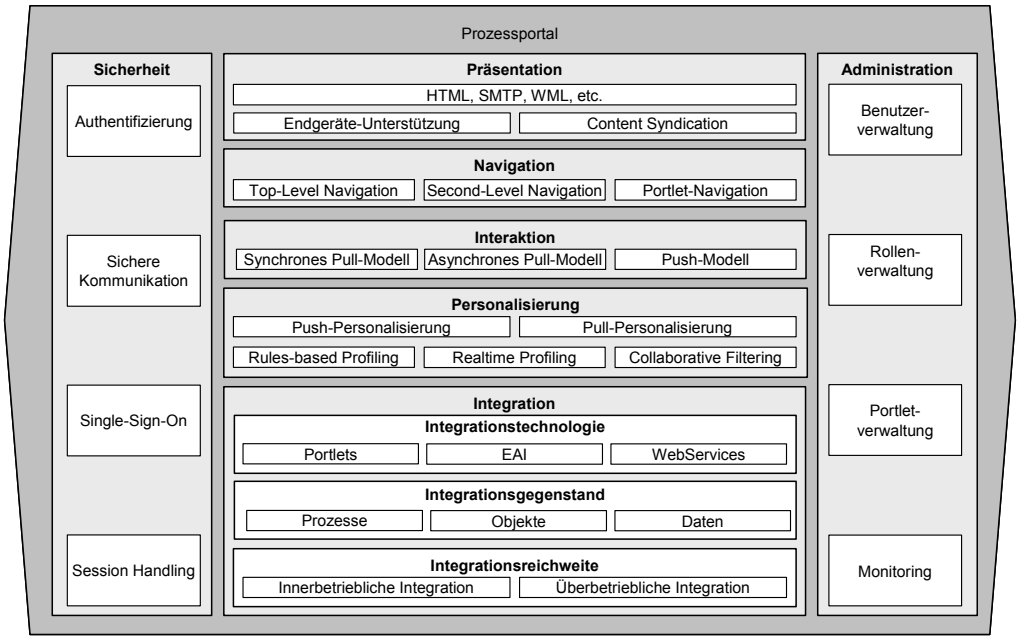


Abbildung 3-8: Portalarchitektur [s. Puschmann 2003, 75]

Strukturelle Elemente: Das Vorgehensmodell umfasst die beiden Techniken Portalarchitektur und Integrationsarchitektur (s. Abbildung 3-9) und ordnet sich dabei in die übergeordnete ‚Methode zur Entwicklung von Prozessportalen' [Alt et al. 2003] ein. Die einzelnen Techniken weisen zwischen vier und sechs Aktivitäten auf. Die Methode bietet ein Meta- sowie ein Dokumentationsmodell, Ergebnisse und Rollen. Sie stützt sich dabei im Wesentlichen auf das im Rahmen des Forschungsprogramms Business Engineering entstandene Methoden-Set PROMET [vgl. Blessing 2001, 160].

Inhaltliche Elemente: Der Gestaltungsbereich der Methode beginnt mit der Analyse der bestehenden Applikationsarchitektur. Ergebnis ist ein Aufgabenkettendiagramm, in dem die Anwendungen bzw. ihre Funktionen den gewünschten Abläufen (Ergebnis der Technik Kooperationsprozessarchitektur aus [Alt et al. 2003]) zugeordnet sind. Die nachfolgende Ermittlung der Integrationsbereiche dient der Technik ‚Portalintegration' dazu, Applikationsfunktionen und Portlets abzuleiten. Im Entwurf der SOLL-Applikationsarchitektur werden die bisher ermittelten Applikationsbausteine integriert dargestellt. In der nachfolgenden Entwicklung der Prozessportalarchitektur werden zentral und dezentral bereitzustellende Portalfunktionen festgelegt. Die abschliessende Umsetzung der Prozessportalarchitektur skizziert die Realisierungsmöglichkeiten. Die Technik ‚Integrationsarchitektur' analysiert zunächst die im Portal zu integrierenden Applikationen und bewertet deren Integrationsgrad. Um die Anzahl der Schnittstellen zwischen Portal und zu integrierenden Anwendungen gering zu halten, bietet die Methode die Einteilung in so genannte Integrationsmuster an. ‚Collaboration Portale' zeigt anhand des kommerziellen Produkts ‚TIBCO PortalBuilder', mit welchen Technologien die einzelnen Integrationsmuster umgesetzt werden können. Die nachfolgende Entwicklung der Integrationsarchitektur beschreibt jedes Integrationsmuster detailliert, z.B. durch die Spezifikation von Portlets zur

Standardsoftwareintegration oder Angaben zur Einbindung von Altanwendungen. Die abschliessende Toolauswahl bietet exemplarisch eine Bewertung von vier Portalprodukten anhand der geforderten Portalfunktionen.

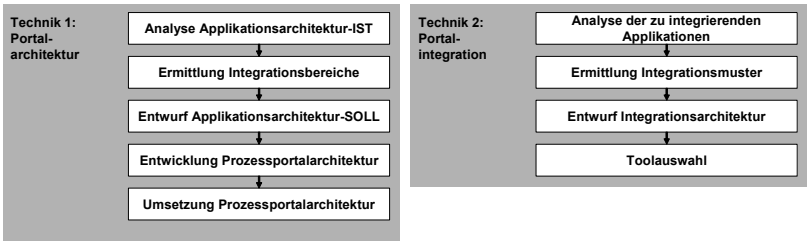


Abbildung 3-9: Vorgehensmodell Collaboration Portale

Bewertung: Obwohl die Methode detaillierte Aspekte von Information Retrieval fast vollständig ausspart, liefert sie mit ihren Aktivitäten und Ergebnissen zur Analyse von Prozessen und Applikationen dennoch wertvolle Grundlagen. Diese lassen sich auf die Planung und Implementierung von Portalsuche und -navigation übertragen.

3.2.7. Integrierte Wissensmanagement-Systeme [Riempp 2003]

Fokus: Der Ansatz von [Riempp 2003] untersucht die Gestaltung und Integration von Informationssystemen für Wissensmanagement (Wissensmanagementsysteme, WMS) mit der Zielsetzung, eine effektive Gestaltung und Ausführung von Geschäftsprozessen in Organisationen zu unterstützen. Für die vorliegende Arbeit relevante Aspekte von Portalen und Information Retrieval als wesentliche Bestandteile von WMS sind in dem Ansatz umfassend enthalten.

Architektur und Gestaltungsmerkmale: ‚Integrierte Wissensmanagement-Systeme' stellt eine Architektur für Wissensmanagementsysteme zur Verfügung. Diese beinhaltet die für die vorliegende Arbeit relevanten Komponenten von Information Retrieval in Portalen, z.B. Navigation und Suche. Riempp bezeichnet diese als Gestaltungsmerkmale des Orientierungsmanagements [Riempp 2003, 102].

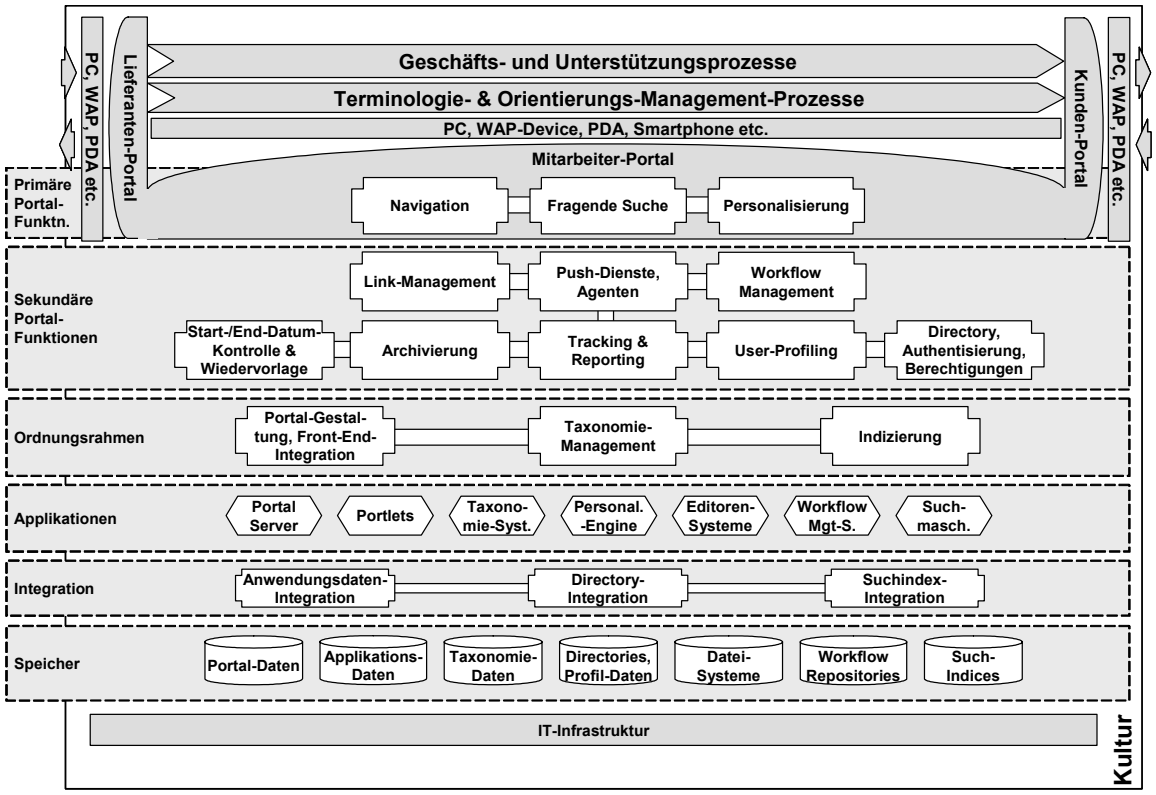


Abbildung 3-10: Architektur für Systeme des Orientierungs-Managements [s. Riempp 2003, 217]

- *Primäre Portalfunktionen* wie *Navigation*, *Suche* oder *Personalisierung* bieten den Portalanwendern zentrale Orientierungshilfen in den Portalinhalten.
- *Sekundäre Portalfunktionen* unterstützen Rollenträger mit Pflegeaufgaben, z.B. einen Applikationsverantwortlichen bei der Vergabe von Berechtigungen (Autorisierung).
- Ein *Ordnungsrahmen* bietet Rollenträgern Funktionen zur Portal- und Frontendgestaltung, zur Pflege von Begriffen für Navigation und Suche (Taxonomiemanagement) oder zur Indizierung von Inhalten.
- Die *Applikationen* stellen die benötigten Funktionen für die darüberliegenden Schichten bereit.
- Die *Integrationsschicht* bietet Funktionen für die Anbindung von Anwendungsdaten im Portal, die Kopplung verschiedener Verzeichnisdienste und das Indizieren und Suchen in relevanten Datenspeichern.
- Die *Speicherschicht* umfasst die Datensammlungen der verschiedenen integrierten Applikationen.

Strukturelle Elemente: In der Arbeit enthaltene Methodenelemente [s. Riempp 2003, 236-244] werden gemäss dem Vorgehensmodell des PROMET-Methodenkerns

zusammengefasst [vgl. Blessing 2001, 160]. Es umfasst die drei Phasen ‚Analyse', ‚Fachliche Konzeption' und ‚Systemkonzeption'. Die einzelnen Phasen enthalten Techniken sowie zugehörige Meta- und Rollenmodelle.

Inhaltliche Elemente: Die *erste Phase* des Vorgehensmodells beginnt mit der Analyse von Kundenprozessen. Dabei werden neben den Aufgaben des Prozesses selbst auch die damit verbundenen Wissensmanagementaufgaben erfasst, z.B. die Suche nach Inhalten. Diese Ergebnisse dienen dem Projektverantwortlichen zur Ableitung von Nutzenpotenzialen, z.B. der Verbesserung von Produkt- und Servicequalität, der Steigerung der Wettbewerbsfähigkeit und Marktstellung, der Erhöhung der Mitarbeiterzufriedenheit und besserer Wissenstransparenz. Die *zweite Phase* beginnt mit der Entwicklung einer auf Basis der Geschäftsstrategie abgeleiteten WM-Strategie. Dabei definiert der Projektverantwortliche Ziele, z.B. das Anbieten komfortabler Suchfunktionen und zugehörige Führungsgrössen, z.B. die Messung der Nutzung und der Anwenderzufriedenheit von Navigations- und Suchfunktionen. Der ‚Entwurf von Prozessen' legt Abläufe und Rollen zum Management von Navigation und Suche fest, z.B. den Aufbau von Suchindizes durch einen so genannten Search Manager. Anschliessend erfolgt in der *dritten Phase* die IS-Architekturplanung. Sie umfasst die Systeme, die Anwender beim Erschliessen von Informationsobjekten unterstützen (vgl. Abbildung 3-10).

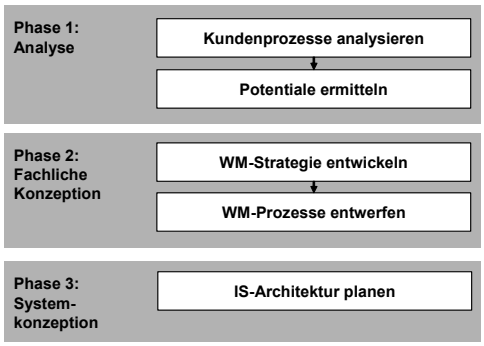


Abbildung 3-11: Vorgehensmodell für integrierte Wissensmanagement-Systeme

Bewertung: ‚Integrierte Wissensmanagement-Systeme' bietet einen umfassenden Lösungsansatz, Informationssysteme für Wissensmanagement zu gestalten. Zielsetzung ist es, betriebliche Prozesse durch die Integration von Inhalten, Zusammenarbeit, Kompetenz und Orientierung zu unterstützen. Orientierung umfasst dabei die im Rahmen des vorliegenden Dissertationsprojekts relevanten Komponenten für Suche, Navigation und Struktur in Portalen. Deren Architektur und Gestaltungselemente sind ausführlich hergeleitet. Auch wenn eine umfassende Methode für das Information Retrieval nicht enthalten ist, bieten die dargestellten Methodenelemente und zahlreichen Praxisbeispiele Handlungsempfehlungen zur Planung und Implementierung der Bereiche ‚Suche' und ‚Navigation' in Portalen.

3.3. Übersicht der Bewertungen

Tabelle 3-3 fasst die Charakteristika der untersuchten Ansätze zusammen.

Methode	Inhaltliche Charakteristika						Strukturelle Charakteristika				
	Prozessorientierung	Prozessmanagement	Retrieval-Leistungen	Architekturelemente	Systemintegration	Ganzheitlichkeit	Vorgehen	Rollen	Ergebnisse	Techniken	Metamodell
Extranet Development Lifecycle [Bayles 1998]	○	◑	○	○	◑	○	●	◑	◑	○	○
Engineering Enterprise Portals Framework [Finkelstein/Aiken 2000]	●	◑	◑	◑	◑	○	●	○	◑	◑	○
Methode zur Konzeption von Intranets [Kaiser 2000]	●	◑	●	◑	◑	●	◑	◑	◑	◑	◑
Implementierung von Unternehmensportalen [Bauer 2001]	◑	○		◑	◑	○	●	○	●	○	○
Corporate Portals [Collins 2002]	◑	○	●	◑	◑	○	●	◑	◑	●	○
Collaboration Portale [Puschmann 2003]	●	◑	○	◑	●	●	●	◑	●	●	◑
Integrierte Wissensmanagement-Systeme [Riempp 2003]	●	◑	●	◑	●	●	◑	◑	◑	◑	●

Legende: ● umfassend ◑ ansatzweise ○ nicht oder rudimentär beschrieben

Tabelle 3-3: Kurzübersicht Methodenvergleich

Bei den inhaltlichen und strukturellen Charakteristika decken die Erkenntnisse von [Riempp 2003], [Puschmann 2003] und [Kaiser 2000] weite Teile der in dieser Arbeit relevanten Aspekte umfassend bzw. ansatzweise ab. Da die genannten Arbeiten jedoch andere Themenschwerpunkte adressieren, weisen sie in Bezug auf Information Retrieval Lücken bei der Darstellung von Abhängigkeiten einzelner Architekturelemente, bei der Berücksichtigung von Pflegeprozessen und bei einzelnen Methodenbestandteilen auf. Aufgrund dieser Lücken wird in Kapitel 5 ein eigener Methodenvorschlag für Information Retrieval in Portalen erarbeitet, der die Stärken aller bewerteten Ansätze berücksichtigt und sie ergänzt bzw. verfeinert.

3.4. Ableitung eines Information-Retrieval-Metamodells

Der Vergleich der analysierten Methoden zeigt, dass kein Ansatz die beschriebenen Anforderungen an eine Methode für Information Retrieval in Portalen umfassend erfüllt. Diese Arbeit möchte daher durch die Entwicklung eines Methodenvorschlags auf Basis der untersuchten Ansätze diese Forschungslücke schliessen.

Die folgenden Abschnitte entwickeln dazu ein Modell, das die unterschiedlichen Information-Retrieval-Bestandteile der untersuchten Ansätze über die drei Ebenen des Business Engineering integriert. Die wesentlichen Gestaltungselemente und Beziehungen werden zunächst in Anlehnung an [Gebert 2003] in einem integrierten Metamodell in der Modellierungssprache des Business Engineering formalisiert. Alle verwendeten Elemente sind direkt aus den Elementen des BE-Metamodells abgeleitet und bilden wesentliche Bestandteile dieses Konzepts auf den Ebenen Strategie, Prozess und System ab (s. Abbildung 3-12).

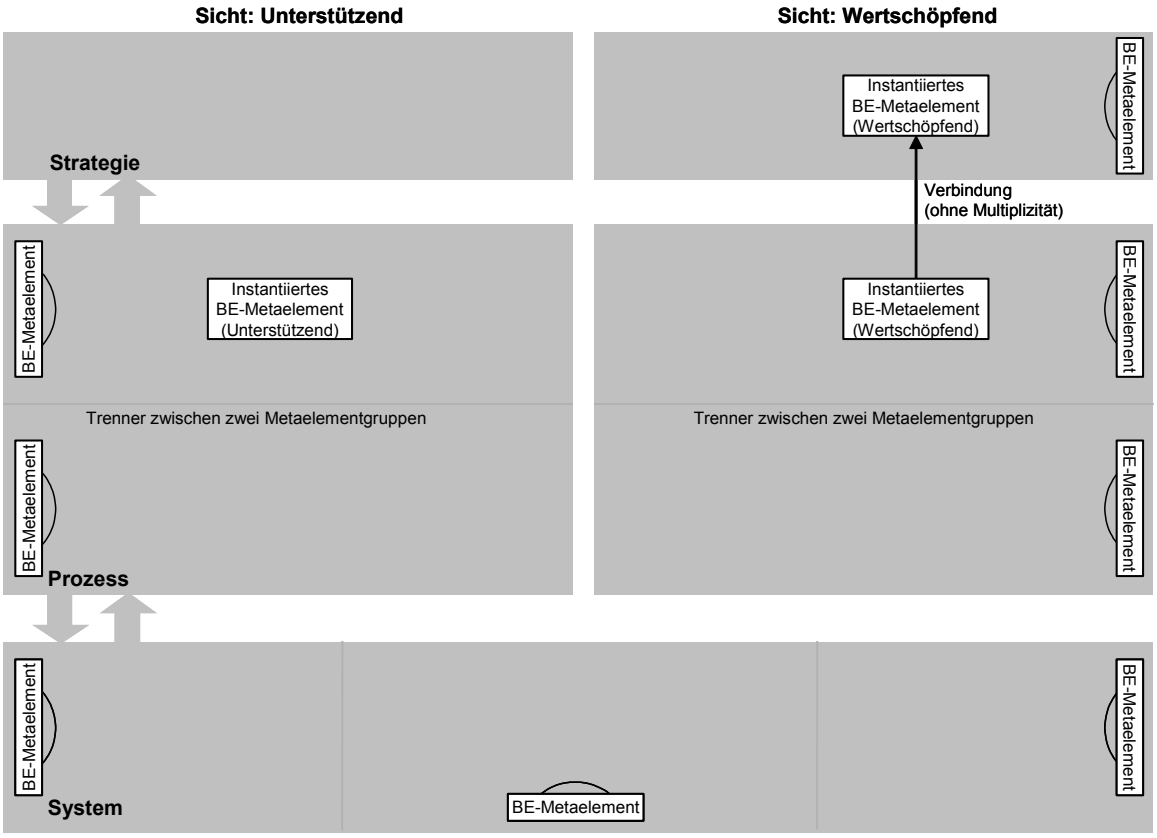


Abbildung 3-12: Darstellung und Gestaltungselemente des integrierten Metamodells (in Anlehnung an [Gebert 2003, 64])

Information Retrieval ist nicht selbst wertschöpfend (s. Abschnitt 2.4.3). Da es jedoch Unterstützungspotenzial für Prozesse der Wertschöpfungskette bietet, weist das Metamodell zwei Sichten auf. Die rechte Sicht beinhaltet relevante Elemente wertschöpfender Prozesse; die linke Sicht stellt die Elemente mit unterstützenden

Aufgaben dar. Die Informationstechnik verbindet die Systemebenen beider Sichten, da sie grundsätzlich wertschöpfende und unterstützende Funktionen integriert.

Alle Elemente des integrierten Metamodells sind direkt aus Elementen des BE-Metamodells abgeleitet. Sie dienen der Strukturierung und sind im Metamodell an den jeweiligen Seiten dargestellt. Zur besseren Übersicht wird auf die zugehörigen Instantiierungsverbindungen verzichtet. Die Zugehörigkeiten werden durch Meta-elementbereiche abgegrenzt. Wie bei der Darstellung des BE-Metamodells verzichtet auch das integrierte Metamodell für Information Retrieval in Portalen auf die Darstellung von Multiplizitäten der Verbindungen.

Für jede Anforderung an eine Methode des Information Retrieval in Portalen (s. Tabelle 3-1, S. 34) werden zunächst die von den analysierten Ansätzen gewählten Lösungen analysiert und ein gemeinsames Integrationsszenario konstruiert. Dabei werden ausschliesslich die Ansätze behandelt, die die jeweilige Anforderung berücksichtigen (s. Tabelle 3-3, S. 51). Das integrierte Metamodell verbindet anschliessend die resultierenden Elemente.

3.4.1. Prozessunterstützung – Unterstützungspotenzial für Geschäftsprozesse

Die untersuchten Ansätze orientieren sich unterschiedlich stark an den betrieblichen Abläufen, die sie durch Informationsidentifikation unterstützen sollen:

- [Bauer 2001] beginnt den Gestaltungsbereich seiner Methode mit der Darstellung von Geschäftsprozessen der Bereiche Ein- und Verkauf am Beispiel der fiktiven Firma ‚Schneiderlein AG'. Zielsetzung ist es, diese „internettauglich umzugestalten und durch ein Portal zugänglich zu machen" [Bauer 2001, 153]. Bei der anschliessenden Darstellung der Umsetzung in einem Portal bleibt jedoch undokumentiert, warum und wie beispielsweise die Funktion ‚Suchen und Finden' hilft, Kunden besser zu informieren oder Aufträge schneller abzuwickeln.
- Auch die Methode von [Collins 2001] orientiert sich an Geschäftsprozessen. Zur Erhebung der ‚Business process and information needs' schlägt sie beispielsweise eine Analyse relevanter Anwendungsfälle (Use Cases) vor. Durch Hinweise wie „make this [required] information easily available such that employees can find information they need" [Collins 2001, 115] werden jedoch weder Bezüge zu den Geschäftsprozessen hergestellt noch das Unterstützungspotenzial detailliert.
- Puschmann bezieht Prozessanforderungen aus der ‚Methode zur Entwicklung von Prozessportalen' [Alt et al. 2002], in deren Kontext sich seine Arbeit einordnet. Da der Fokus seines Ansatzes nicht auf Information Retrieval liegt, dokumentiert Puschmann ein damit verbundenes Unterstützungspotenzial jedoch nicht weiter.

- Bayles führt in ihrem ‚Extranet Development Plan Template' zwar die Frage „What business problem will the extranet solve?" auf, verweist jedoch nicht auf die Analyse von Geschäftsprozessen als möglichen Ansatzpunkt. Nutzenaspekte eines Extranets nennt sie für die Bereiche ‚Infrastructure', ‚Sales and Marketing' sowie ‚Customer Service and Support Benefits' [s. Bayles 1998, 39-50]. Dies beinhaltet, dass Anwender sich die von ihnen benötigten Informationen selbst beschaffen müssen. Zugehörige Anforderungen an eine entsprechende Suchfunktion werden jedoch nicht vertieft.

- Kaiser erfasst gezielt die Informationsbedürfnisse rechercheintensiver Aufgaben bzw. unstrukturierter, wissensorientierter Prozesse [Kaiser 2000, 120]. Er sieht dabei Unterstützungspotenziale von Information Retrieval insbesondere im ‚schnellen Lokalisieren' von Informationen [Kaiser 2000, 123].

- Riempp sieht das „Lokalisieren relevanter Informationsobjekte" zur Unterstützung der Aufgaben eines Geschäftsprozesses als primäre Zielsetzung von Information Retrieval [Riempp 2003, 158]. Er stellt die Analyse von Prozessen und die Ermittlung von Potenzialen an den Anfang seiner Handlungsempfehlungen [Riempp 2003, 240]. Der Nutzen einer jeweiligen Lösung ergibt sich aus der Verbesserung von Geschäftsprozessen in Bezug auf Kosten-, Zeit- und Qualitätsaspekte, z.B. durch Verringerung von Suchzeiten.

- [Finkelstein/Aiken 2000] bewerten mögliche Lösungsalternativen anhand des Unterstützungspotenzials für Geschäftsprozesse. Sie definieren Messgrössen und schlagen die Analyse des Nutzungsverhaltens der im Portal bereitgestellten Inhalte vor.

Einige der analysierten Ansätze weisen bereits einen engen Bezug zu geschäftlichen Aufgaben auf. Sie stellen die Identifikation von Informationen als Leistung von Information Retrieval dar, die von Geschäftsprozessen konsumiert wird. Keiner der Ansätze sieht Information Retrieval als wertschöpfend an; das Unterstützungspotenzial definiert sich immer durch eine Qualitätssteigerung, Zeitersparnis und/oder Kostenreduktion der unterstützten Geschäftsprozesse.

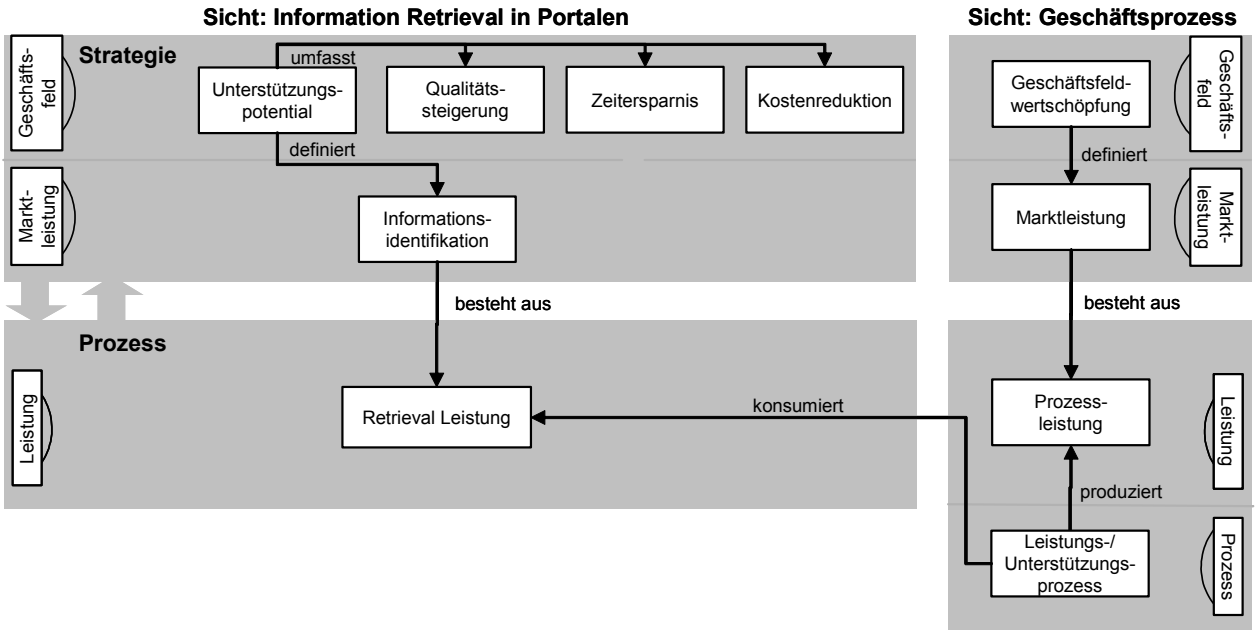


Abbildung 3-13: Prozessunterstützung durch Information Retrieval in Portalen

3.4.2. Retrieval-Leistungen – Definition des Funktionsumfangs

Die von den Geschäftsprozessen konsumierten Retrieval-Leistungen werden in den analysierten Ansätzen unterschiedlich detailliert beschrieben:

- [Bayles 1998] detailliert in seinem Ansatz weder Suche noch Navigation.
- Bauer skizziert mit Aussagen wie „Jedes Portal braucht eine gute ‚Suchen und Finden'-Funktion" [Bauer 2001, 82] den benötigten Funktionsumfang nur schemenhaft. Er streicht jedoch die Bedeutung der „Speicherung von Metadaten zu den Dokumenten und die Möglichkeit, nach diesen zu suchen" heraus [Bauer 2001, 112].
- Puschmann sieht Navigation als „wesentliches Element eines Portals" [Puschmann 2003, 82]. Suche ist für ihn ein Teil der Navigation. Er unterscheidet dabei Volltextsuche und Suchen „durch Browsen innerhalb einer vorgegebenen Begriffshierarchie (Taxonomie)" [Puschmann 2003, 82].
- Riempp differenziert unterschiedliche Vorgehensweisen bei der Indizierung von Inhalten, z.B. Volltextindizierung, attributive Indizierung oder taxonomiebildende bzw. taxonomieverwendende Indizierung. Abhängig von der Indizierungsart stehen den Anwendern unterschiedliche Leistungen für Navigation und Suche im Portal zur Verfügung [Riempp 2003, 224-227].
- Anwender lokalisieren relevante Informationen im Ansatz von Kaiser entweder durch „Navigation (Browsing) oder Suche (Searching)" [Kaiser 2000, 159]. Er

unterscheidet zwischen der Eingabe beliebiger Suchbegriffe und der „gezielten Suche auf einzelnen Strukturelementen [eines Informationsobjekts]“ [Kaiser 2000, 160].

- [Collins 2001] unterscheidet Navigation entlang einer Begriffshierarchie (‚Taxonomy’), Volltextsuche (‚Standard Search Features’) und Suche in Attributen eines Dokuments (‚Metadata Search’). Navigation und Metadatensuche erfordern einen durch das Unternehmen gepflegten Ordnungsrahmen.

Tabelle 3-4 ordnet die genannten Leistungen nach Navigation und Suche und differenziert in Anlehnung an [Collins 2001], ob dem Anwender ein Ordnungsrahmen zur Verfügung gestellt wird.

Ansatz	Navigation		Suche	
	Freie Begriffswahl	Kontrollierte Begriffe	Freie Begriffswahl	Kontrollierte Begriffe
Extranet Development Lifecycle [Bayles 1998]	◑	○	◑	○
Engineering Enterprise Portals Framework [Finkelstein/Aiken 2000]	◑	●	◑	●
Methode zur Konzeption von Intranets [Kaiser 2000]	●	●	●	●
Implementierung von Unternehmensportalen [Bauer 2001]	◑	○	◑	○
Corporate Portals [Collins 2002]	◑	●	◑	●
Collaboration Portale [Puschmann 2003]	●	●	◑	○
Integrierte Wissensmanagement-Systeme [Riempp 2003]	●	◑	●	●

Legende: ● umfassend ◑ ansatzweise ○ nicht oder rudimentär beschrieben

Tabelle 3-4: Retrieval-Leistungen der untersuchten Ansätze

Retrieval-Leistungen werden von Geschäftsprozessen einerseits in Form von Volltextsuche (Suche ohne Ordnungsrahmen), attributbasierter Suche (Suche mit Ordnungsrahmen), taxonomieorientierter Navigation (Navigation mit Ordnungsrahmen) oder als Kombination konsumiert. Die Bereitstellung von IR-Leistungen erfordert Aufbau- und Betriebsprozesse, z.B. die Pflege eines kontrollierten Vokabulars für eine Stichwortsuche, die Bereitstellung eines Suchformulars oder den Aufbau eines Navigationsgerüsts.

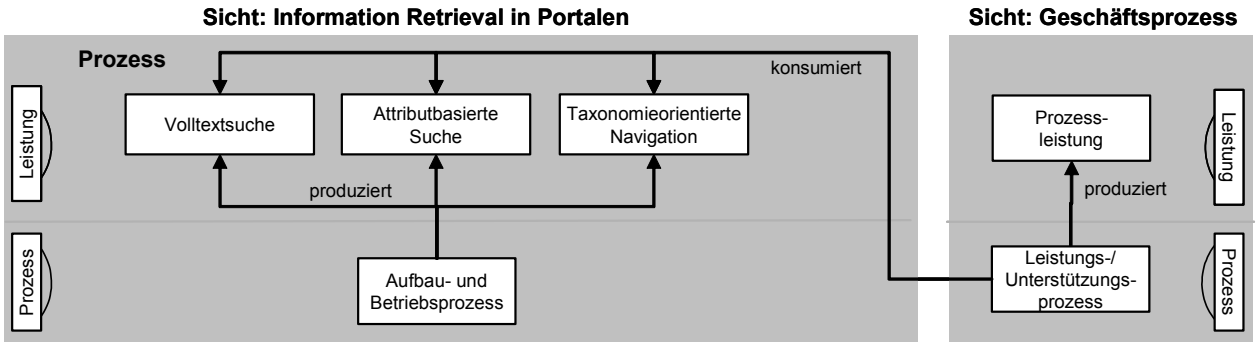


Abbildung 3-14: Retrieval-Leistungen

3.4.3. Prozessmanagement – Aufbau und Betrieb von Information Retrieval

Prozesse für den Aufbau und Betrieb von Information Retrieval-Lösungen, z.B. für die Abläufe zur Integration bestehender Anwendungen, die Pflege von Metadaten zur Klassifikation von Dokumenten oder die Aktualisierung von Suchmaschinenindizes gewichten die untersuchten Ansätze unterschiedlich stark:

- Sowohl [Puschmann 2003] als auch [Collins 2001] benennen organisatorische Rollen und beschreiben Schritte zur Planung und Umsetzung eines Portals. IR-spezifische Aspekte sowie Prozesse und Rollen zur fortlaufenden Pflege des neu geschaffenen Informationssystems sind nicht enthalten.

- [Finkelstein/Aiken 2000, 268] beschreiben einen Lebenszyklus von Portalinhalten (‚Data Life Cycle') mit verschiedenen Phasen, z.B. Inhaltsnutzung (‚data utilization'), -veränderung (‚data manipulation') oder Zuordnung von Metadaten (‚Metadata creation'). Dazugehörige organisatorische Rollen und Ablaufdetails werden nicht genannt.

- [Kaiser 2000] beschreibt die notwendigen Abläufe für Aufbau und Betrieb von webbasierten Informationssystemen und definiert dazugehörige Rollen wie ‚Autor', ‚Informationsverantwortlicher' oder ‚Leiter Inhaltsverantwortliche' für Intranetinhalte. Diese müssen für Portale angepasst und erweitert werden, da Information Retrieval i.d.R. eine Vielzahl verschiedener Content-Applikationen mit unterschiedlichen Informationsstrukturen integrieren muss. Auch Veränderungen der betrieblichen Begriffsverwendung (so genannter Term Lifecycle) und dadurch erforderliche Pflegeprozesse müssen ergänzt werden.

- [Riempp 2003] definiert Prozesse für so genannte Orientierungsfunktionen, die den Anwendern durch Suche oder Navigation die Erschliessung von Informations-objekten im Portal erleichtert. Die in der Arbeit aufgeführten Aufgaben-kettendiagramme zeigen exemplarisch Prozesse für die Klassifikation von Inhalten, Terminologiepflege und Bereitstellung von Navigations- und Suchfunktionen. Ein

Prozessmanagement für den Abgleich unterschiedlicher Zugriffsberechtigungen von im Portal zu integrierenden Content-Applikationen oder Abläufe zur Bereitstellung der Benutzerschnittstelle für Information Retrieval sind hingegen nicht detailliert.

Die bereitzustellenden Retrieval-Leistungen definieren die notwendigen Aufbau- und Betriebsprozesse. Für eine *Volltextsuche* sind beispielsweise der Aufbau und die Pflege eines entsprechenden Suchindex notwendig (‚Indexmanagement'). Retrieval-Leistungen wie attributbasierte Suche oder taxonomieorientierte Navigation stellen dem Anwender hingegen einen definierten Ordnungsrahmen zur Informationsidentifikation zur Verfügung. Die Benennung, Strukturierung und Bereitstellung dieses Ordnungsrahmens (Taxonomie) ist Aufgabe des ‚Terminologiemanagements' (vgl. [Hellmuth 1997]). Die Klassifikation von Inhalten auf Basis einer Taxonomie erweitert auch das Indexmanagement um ein entsprechendes Gestaltungselement.

Die Integration bestehender Anwendungen, existierender Zugriffsberechtigungen und bereitzustellender Benutzerschnittstellen ist in allen Fällen unverzichtbar. Sie sind im Unterstützungsprozess ‚Integrationsmanagement' zusammengefasst.

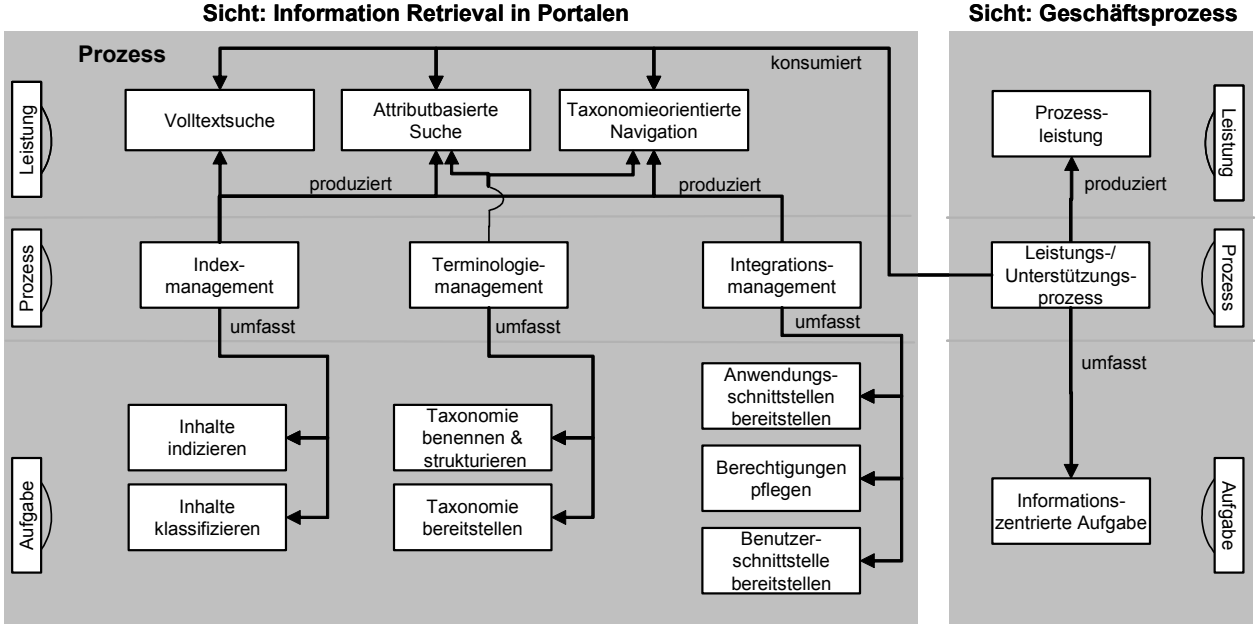


Abbildung 3-15: Aufbau- und Pflegeprozesse für Information Retrieval in Portalen

3.4.4. Architekturelemente – Berücksichtigung relevanter Komponenten

Die analysierten Ansätze führen z.B. Navigation, Suche, Authentifizierung, Autorisierung oder Integration als wesentliche Komponenten von Portal- und Retrieval-Architekturen auf. Resultierende Abhängigkeiten untersuchen sie in unterschiedlichem Masse.

- [Bauer 2001] nennt ‚Personalisierung', ‚Benutzerverwaltung', ‚Dynamische Inhalte und Webpublishing', ‚Externe Webanwendungen' und ‚Integration von Unternehmensanwendungen' als relevante Gestaltungselemente eines Portals. Er beschreibt diese, wie z.B. das ‚Suchen und Finden-Portlet', lediglich deskriptiv und zeigt keine Wirkungszusammenhänge auf, z.B. für die Anbindung dynamischer Inhalte an eine Suchfunktion.

- ‚Corporate Portals' von [Collins 2001] sieht ‚Search capabilities', ‚Taxonomy' und ‚Content Management' als wesentliche Bestandteile von Portalen. Abhängigkeiten der Portalfunktionen deutet die Autorin lediglich in einem Beispielprojektplan an. So ordnet sie die ‚set up security configuration' sowohl ‚Search' als auch ‚Content Management' zu [s. Collins 2001, 353].

- Funktionen wie ‚Suche', ‚Navigation', ‚Annotation von Informationen', ‚Personalisierung', ‚Erstellen und Versionieren' sowie ‚Kategorisierung von Informationen' sind bei Kaiser wesentliche Komponenten der Applikationsarchitektur [Kaiser 2000, 29]. Die „Verbesserung von Suchergebnissen durch Strukturierung schwach strukturierter bzw. unstrukturierter Informationsobjekte" [Kaiser 2000, 160] deutet dabei einen Zusammenhang zwischen Inhalten, ihrer Klassifikation und Suchfunktionen an.

- [Puschmann 2003] sieht ‚Präsentation', ‚Navigation', ‚Personalisierung', ‚Integration', ‚Administration' und ‚Sicherheit' als wesentliche Elemente einer Portalarchitektur [s. Puschmann 2003, 74-100]. Er beschreibt und erklärt diese ausführlich, zeigt jedoch keine Abhängigkeiten zwischen diesen Elementen auf. Zudem subsumiert er den für das vorliegende Dissertationsprojekt wichtigsten Bestandteil ‚Suche' im Portalelement ‚Navigation', eine Vereinfachung, die spezialisierte IR-Literatur nicht nachvollzieht [s. Baeza-Yates/Ribeiro-Neto 1999].

- [Finkelstein/Aiken 2000, 491] unterteilen ihre Architekturelemente in Suchmaschine (‚Search engine'), Indizes (‚Business information directory'), Funktionen für die Indizierung (‚Metadata crawlers') und bestehende Inhaltsquellen (‚Decision, collaborative, and other corporate and external systems').

- Die Architektur von [Riempp 2003, 217] stellt Gestaltungselemente (so genannte Funktionalitäten) für Wissensmanagementsysteme zur Verfügung, welche für die vorliegende Arbeit relevante Aspekte von Information Retrieval und Portalen als Benutzerschnittstellen, z.B. in Form von ‚Navigation', ‚Suche', ‚Indizierung' oder ‚Administration' beinhalten.

Die Architekturelemente sind in den untersuchten Ansätzen teilweise unterschiedlich bezeichnet, können jedoch unter den in Tabelle 3-5 dargestellten acht Begriffen zusammengefasst werden.

Ansatz	Architekturelemente								
	Präsentation	Navigation	Suche	Klassifikation	Personalisierung	Integration	Authentifizierung	Autorisierung	Sonstige
Extranet Development Lifecycle [Bayles 1998]	○	○	○	○	○	◑	◑	◑	Version Control, Translation, eCommerce, Secure Transactions
Engineering Enterprise Portals Framework [Finkelstein/Aiken 2000]	●	◑	●	●	●	◑	○	○	
Methode zur Konzeption von Intranets [Kaiser 2000]	◑	◑	◑	●	◑	◑	◑	◑	Annotationsfunktionen, Erstellen, Versionieren
Implementierung von Unternehmensportalen [Bauer 2001]	●	◑	●	●	●	●	◑	◑	
Corporate Portals [Collins 2002]	◑	◑	◑	◑	○	◑	◑	◑	Help Features, Content Management, Collaboration, Communication
Collaboration Portale [Puschmann 2003]	●	●	◑	○	●	●	●	●	Interaktion
Integrierte Wissensmanagement-Systeme [Riempp 2003]	●	◑	◑	●	●	◑	●	●	Link-Management, Archivierung, Tracking, Reporting, User-Profiling

Legende: ● umfassend ◑ ansatzweise ○ nicht oder rudimentär beschrieben

Tabelle 3-5: Architekturelemente der untersuchten Ansätze

Abbildung 3-16 ordnet die genannten Funktionen sowie Datensammlungen den sie umfassenden Applikationen zu. Aspekte der Personalisierung sind den anderen Bereichen zugeordnet, z.B. durch die Anpassung der Benutzeroberfläche im Architekturelement ‚Präsentation' oder die Filterung von Inhalten im Architekturelement ‚Suche'.

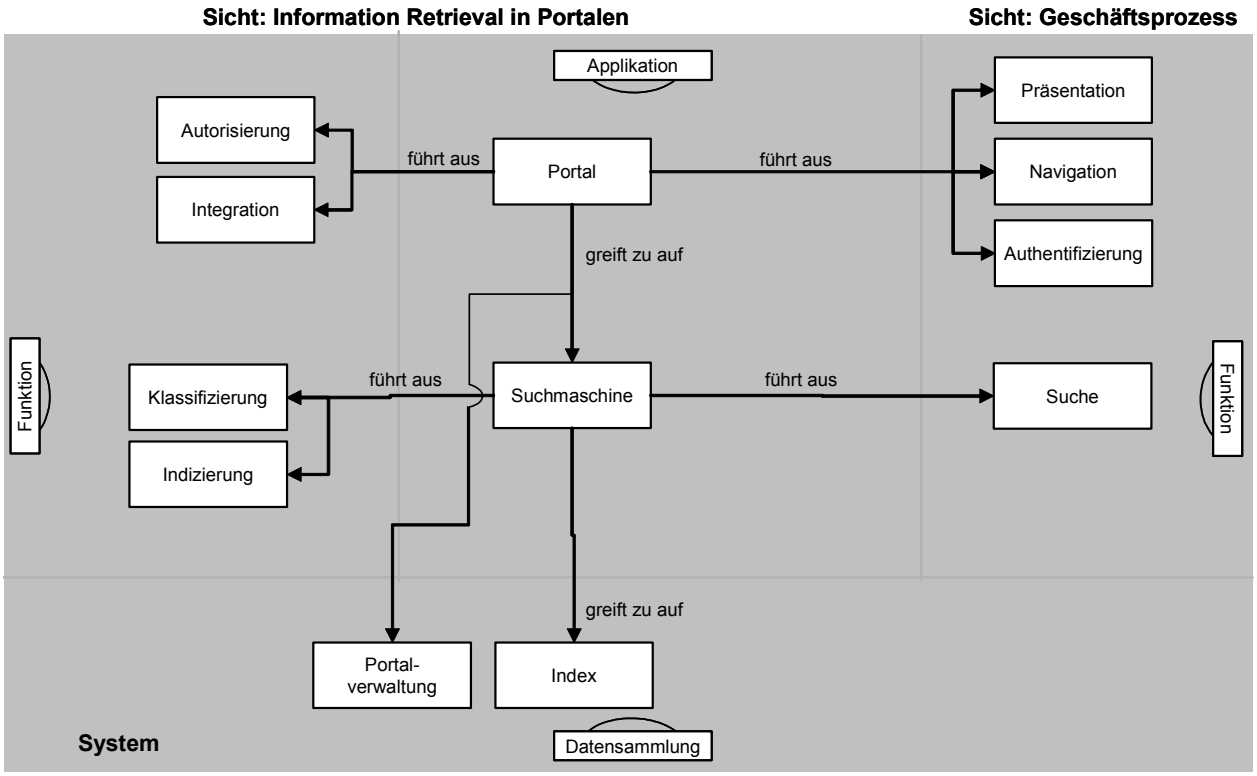


Abbildung 3-16: Architekturelement im integrierten Metamodell

3.4.5. Systemintegration – Anbindung bestehender Applikationen

Die untersuchten Ansätze bieten unterschiedliche Lösungen zur Integration von IR-Funktionen in ein bestehendes Informationssystem:

- [Bauer 2001] schlägt eine Analyse bestehender Applikationen vor. Details, Handlungsoptionen oder Lösungsalternativen sind jedoch nicht ausgeführt. So schreibt er z.B., dass „die Speicherung von Metadaten zu den Dokumenten und die Möglichkeit, nach diesen zu suchen“ wichtig ist [Bauer 2001, 112], zeigt jedoch dafür notwendige Integrationsschritte nicht auf.
- [Collins 2001] betont die Wichtigkeit, bestehende Applikationen zu erfassen und im Portal integriert bereitzustellen: „Developing an accurate profile of existing sources is an essential first step” [Collins 2001, 190]. Die Methode stellt dazu Checklisten zur Verfügung, z.B. für die Beschreibung vorhandener Informationsquellen.
- [Kaiser 2000] analysiert beim Vorgehen zur ‚Integration von Webapplikationen’ bestehende Systeme. Mit seinen umfassenden Analysedokumenten wie ‚Informationsquellenbeschreibungen’, ‚Informationsstrukturierung’, ‚Applikationsübersicht’ oder ‚Datenstrukturbeschreibung’ bietet er Ergebnisse an, die auch in anderen Integrationsmethoden verwendbar sind.

- [Bayles 1998] beschreibt Integration primär aus technologischer Sicht. Er konzentriert sich auf die Integration über so genannte Middleware-Technologien wie Common Object Request Broker Architecture (CORBA) oder Remote Method Invocation (RMI). Diese binden zu integrierende Systeme über Schnittstellen (so genannte Konnektoren) in eine übergreifende Integrationsarchitektur ein.

- Die ‚Analyse der zu integrierenden Applikationen' [s. Puschmann 2003, 168] fokussiert auf die Integration existierender Anwendungen bzw. Anwendungsbestandteile im Portal. Wesentliche Technologien dieser Integration sind für ihn Portlets, Enterprise Application Integration und WebServices, deren Zusammenhänge er exemplarisch in seinen ‚Bausteinen der Integrationsarchitektur' für den Fall einer ‚Lieferterminvorhersage' einordnet. Puschmann sieht weiterhin die so genannte ‚semantische Integration der beteiligten Systeme' als eine zentrale Herausforderung. Die semantische Integration ermöglicht verschiedenen Geschäftspartnern die Kommunikation über heterogene Informationssysteme. Er schlägt dazu vor, „Standards bezüglich Daten, Objekten und Prozessen zu definieren" [s. Puschmann 2003, 126], ohne jedoch die betroffenen Architekturelemente zu nennen oder entsprechende Lösungsschritte zu vertiefen.

- [Finkelstein/Aiken 2000] analysieren bestehende Applikationen in Bezug auf ihre Inhalte, Datenstrukturen und Metadaten. Gemeinsam mit den aus den fachlichen Benutzeranforderungen abgeleiteten Datenmodellen und Metadaten bilden diese Ergebnisse die Grundlage für die Integration von Systemen.

- Die Methodenbestandteile von [Riempp 2003] erfassen im Sinne einer Bestandsaufnahme im Unternehmen neben den vorhandenen Prozessen auch die dazugehörigen Informationssysteme. Dies umfasst die Aufstellung aller in den Geschäftsprozessen genutzten internen und externen Quellen von Informationsobjekten, die sie führenden Systeme und die jeweiligen Strukturen oder Klassifikationsgerüste. Die Integration dieser Applikationen ist ein eigener Bestandteil [s. Riempp 2003, 261]. Die Arbeit stellt für das Information Retrieval in Portalen relevante Integrationsansätze für Inhalte, Benutzerverzeichnisse und Suchindizes detailliert dar.

Alle Ansätze berücksichtigen vorhandene Systeme. Die Integration bestehender Content-Applikationen umfasst in den meisten Fällen die Analyse von Inhalten, deren Struktur sowie deren Zugriffsbeschränkungen. Die Anbindung von inhaltsführenden Anwendungen an eine Suchmaschine erfolgt dabei über eine Schnittstelle. Dieser so genannten Konnektor (auch ‚Connector', ‚Adapter' oder ‚Gateway') beeinflusst die Indizierung sowie die Informationssuche und -anzeige im Portal (s. Abbildung 3-17).

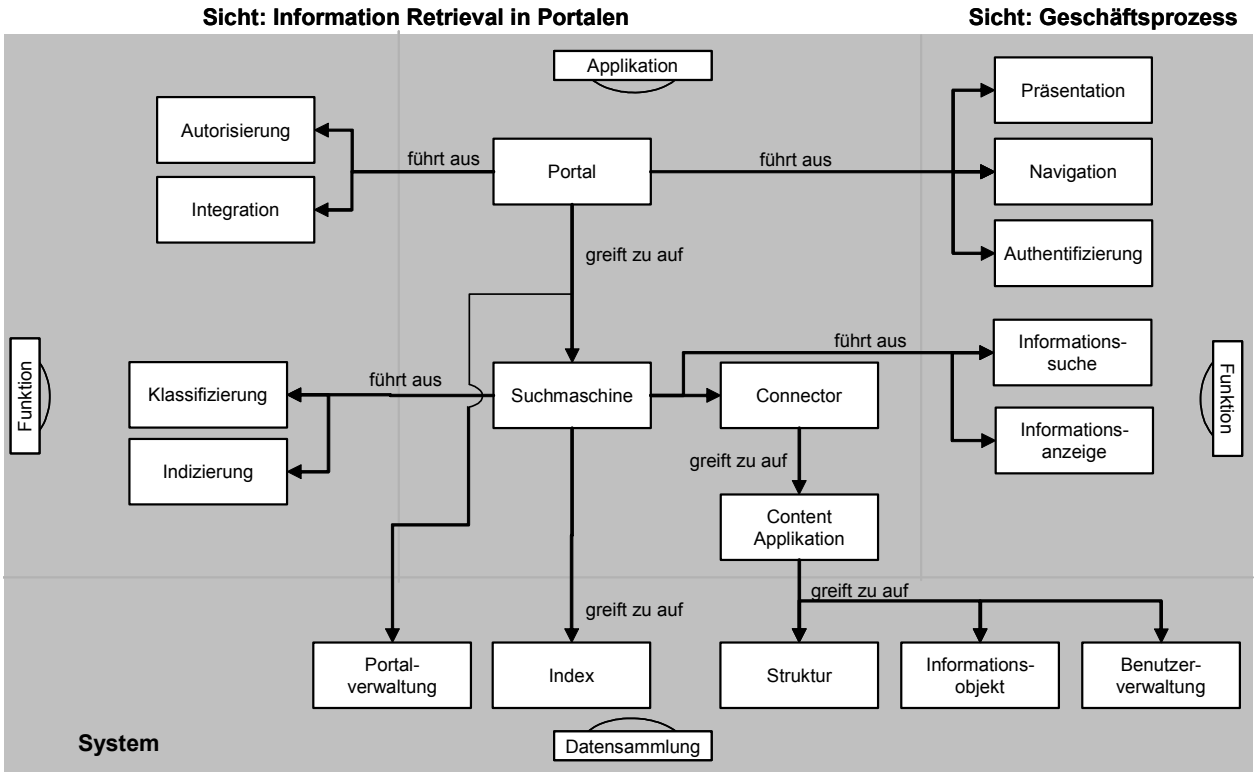


Abbildung 3-17: Systemebene des integrierten Metamodells

3.4.6. Ganzheitlichkeit – Verbindung von Strategie, Prozessen und Systemen

Die nachhaltige Nutzung technologischer Innovationen wie Portale oder Suchmaschinen erfordert einen ganzheitlichen Ansatz, der Strategie-, Prozess- und Systemkomponenten in einer Organisation berücksichtigt.

Tabelle 3-6 zeigt Schwerpunkte der untersuchten Ansätze.

Ansatz	Business-Engineering-Ebene		
	Strategie	Prozesse	Systeme
Extranet Development Lifecycle [Bayles 1998]	●	◑	◑
Engineering Enterprise Portals Framework [Finkelstein/Aiken 2000]	◑	○	●
Methode zur Konzeption von Intranets [Kaiser 2000]	●	●	●
Implementierung von Unternehmensportalen [Bauer 2001]	●	○	◑
Corporate Portals [Collins 2002]	●	◑	○
Collaboration Portale [Puschmann 2003]	◑	◑	●
Integrierte Wissensmanagement-Systeme [Riempp 2003]	●	●	●

Legende: ● umfassend ◑ ansatzweise ○ nicht oder rudimentär beschrieben

Tabelle 3-6: Schwerpunkte der untersuchten Ansätze

Die Verbindung der bisher hergeleiteten und beschriebenen Metamodellbestandteile führt zum integrierten Metamodell für Retrieval in Portalen (s. Abbildung 3-18).

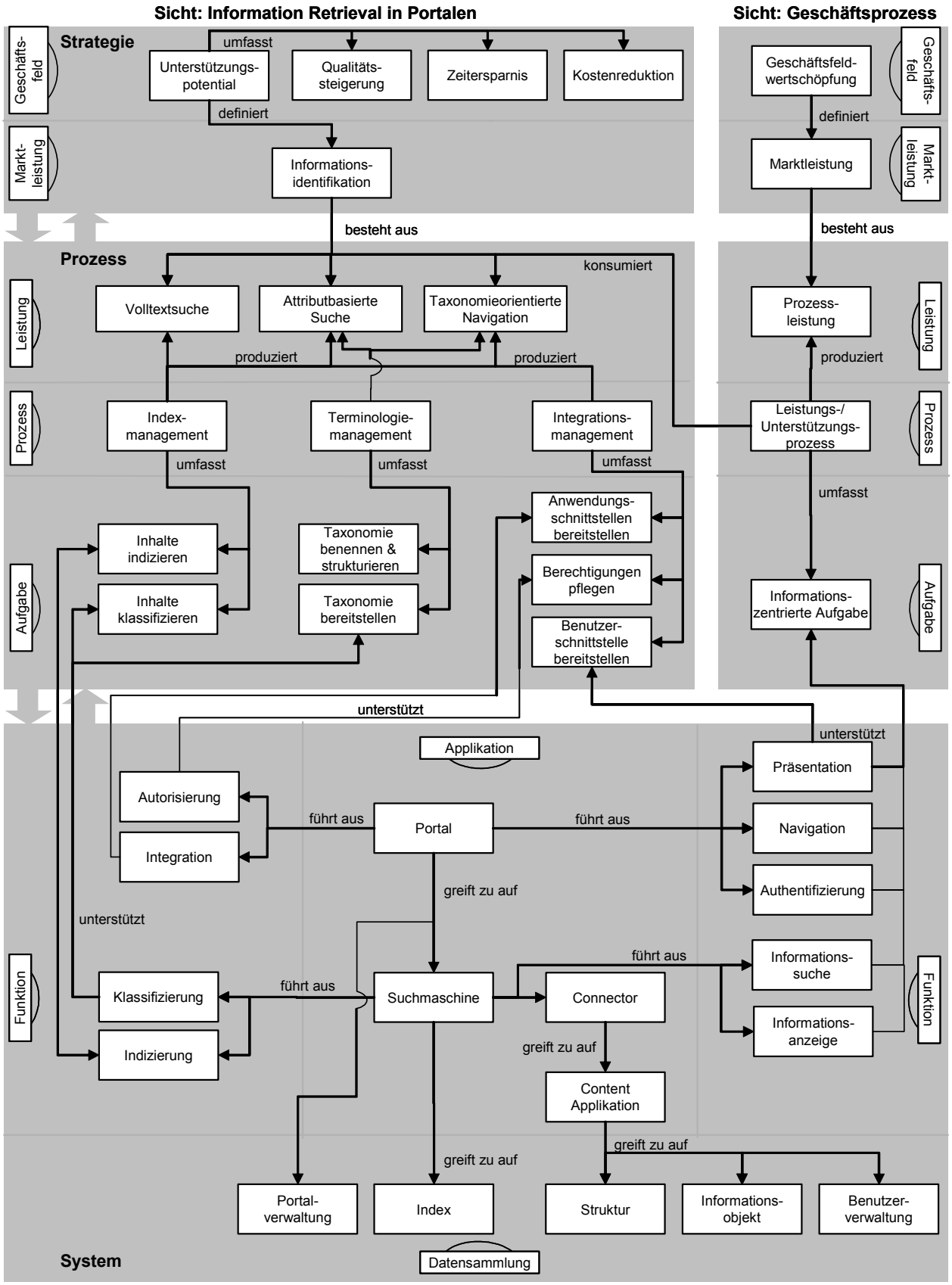


Abbildung 3-18: Integriertes Metamodell für Information Retrieval in Portalen

4. Erfahrungen aus der Praxis

Die folgenden Abschnitte stellen reale Lösungen für Information Retrieval in Portalen in vier unterschiedlichen Organisationen anhand von Fallstudienbeschreibungen vor. Sie wurden im Rahmen der Forschungstätigkeit der Kompetenzzentren ‚Customer Knowledge Management' (CKM) und ‚Customer>Knowledge>Performance' (C>K>P) des Instituts für Wirtschaftsinformatik der Universität St. Gallen aufgenommen. Um die Relevanz der aus der Literatur abgeleiteten Gestaltungselemente des Methodenvergleichs umfassend überprüfen zu können, sollten die untersuchten Lösungen:

- in der Unternehmenspraxis realisiert sein und im operativen Betrieb angewendet werden,
- eine grosse Varianz ihrer Zielsetzungen, organisatorischen und technischen Verankerungen aufweisen und
- Zugriff auf gesichertes Datenmaterial ermöglichen.

Die vier im vorliegenden Dissertationsprojekt enthaltenen Information-Retrieval-Lösungen erfüllen die genannten Kriterien. Dabei beschreiben sie organisatorische und technische Unterstützungspotenziale und Restriktionen von Information Retrieval in Portalen. Rechtliche und kulturelle Rahmenbedingungen werden nicht betrachtet.

Grundlagen der Darstellung bilden primär eigene Tätigkeiten des Autors im Rahmen der vier Projekte aber auch technische Implementierungen, Projektdokumentationen und Interviews mit den Projektverantwortlichen. Falls nicht anders vermerkt handelt es sich um originäre Abbildungen aus den untersuchten Praxisfällen. Aufgrund des begrenzten Umfangs konzentrieren sich die Beschreibungen auf die im Kontext der vorliegenden Arbeit relevanten Aspekte der jeweiligen Lösung.

Die Fallstudien führen jedes Unternehmen zunächst anhand der *Organisation und Ausgangssituation* sowie der *Treiber und Herausforderungen für Information Retrieval* ein. Anhand der Ausprägungen *Prozessorientierung bzw. Unterstützungspotential*, *Retrieval-Leistungen*, *Prozessmanagement* und *Systemintegration* erfolgt danach die weitere Darstellung anhand der Struktur zur Entwicklung des Metamodells für Information Retrieval in Portalen (s. Abschnitt 3.4). Abschliessend werden für jeden Praxisfall *kritische Erfolgsfaktoren und geplante Weiterentwicklungen* betrachtet.

Den Abschluss des Kapitels bilden die aus den untersuchten Fällen abgeleiteten Erkenntnisse sowie die daraus resultierenden Konsequenzen für die eigene Arbeit.

4.1. Zumtobel Staff – Customer Portal

„Die Einführung einer Suchmaschine kostet Geld. Das Nicht-Finden relevanter Informationen jedoch erheblich mehr.“

Ingmar Blum, Head of eBusiness Zumtobel Staff

4.1.1. Organisation und Ausgangssituation

Zumtobel Staff (ZS) ist ein weltweiter Anbieter von Lichtlösungen für alle Anwendungsbereiche professioneller Gebäudebeleuchtung, z.B. in Industrieanlagen, Büros, Shops, Museen, Parkhäusern, Hotels, Flughäfen und Krankenhäusern. Zu den Zielgruppen gehören neben dem Fachhandel, Installateuren, Lichtplanern und Architekten zunehmend auch Investoren, Projektentwickler, Betreiber und Verwalter grosser gewerblicher Immobilienprojekte. Im Geschäftsjahr 2000/2001 erwirtschaftete das Unternehmen mit Hauptsitz in Dornbirn, Österreich, einen Umsatz von EUR 486 Mio. Von den weltweit insgesamt 3`085 Mitarbeitern, die in 70 Ländern tätig sind, arbeiten rund 1`200 am Standort Dornbirn. Zumtobel Staff sieht sich als Lösungsanbieter, der mehr umfasst als die Produktion und den Verkauf von Produkten. Die an den Kundenprozessen ausgerichtete Unternehmensstrategie erfordert, dass relevante Informationen zielgruppengerecht im Kontext des jeweiligen Produkts angeboten werden, z.B. Berichte über Referenzinstallationen für Architekten, Datenblätter für Elektriker oder Veranstaltungshinweise für Investoren.

Die Notwendigkeit, mit allen Partnern enger zusammenzuarbeiten und ihnen rund um die Uhr auf allen Kontinenten diese hochwertigen Inhalte zur Verfügung zu stellen führte im Herbst 2000 zur Formulierung einer neuen eBusiness-Strategie. Die so genannte eStrategie sieht vor, Inhalte zielgruppenspezifisch zur Verfügung zu stellen und dadurch beispielsweise die Markteinführungszeit neuer innovativer Produkte zu verkürzen. Diese sollen als elektronische Dienstleistungen bereitgestellt werden, die sich durch gegen Null strebende Grenzkosten auszeichnen [vgl. Fleisch 2000, 192].

Ein zentrales Ergebnis des Initialprojekts im Oktober 2001 bildet die Portal-infrastruktur bei Zumtobel Staff (s. Abbildung 4-1). Die resultierende Applikation ist eine der weltweit ersten operativen Implementierungen des SAP Enterprise Portals.

Als erste Anwendung, die Zumtobel Staff als ‚eService’ bezeichnet, entwickelten die Projektverantwortlichen z.B. einen individuellen Produktkatalog (Produktdatenbank, PDB) mit personalisierten Preisen. Alle Zielgruppen können über diese Plattform die Verfügbarkeit von spezifischen Kundenlösungen prüfen und den Status ihrer Aufträge zu jeder Zeit abfragen.

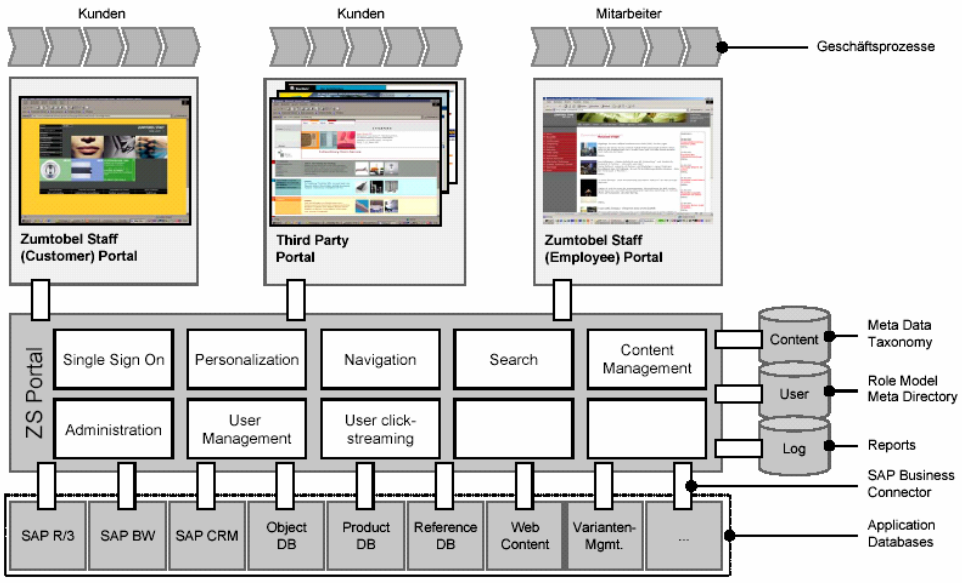


Abbildung 4-1: Portalinfrastruktur bei Zumtobel Staff

Auf Basis der im Initialprojekt entwickelten umfassenden Infrastruktur wurde das Portal sukzessive um zusätzliche Elemente und Dienstleistungen ergänzt. Ein Content-Management-System (CMS) stellt den Anwendern des Kundenportals beispielsweise ergänzende Informationen rund um die Lichtlösungen wie Montageanleitungen, Referenzprojekte, Produktneuigkeiten oder Veranstaltungshinweise zur Verfügung.

4.1.2. Treiber und Herausforderungen für Information Retrieval

- Die Kunden des Customer Portals konnten Inhalte und Dienste ausschliesslich über vorhandene Navigationseinträge und feste Menüstrukturen erreichen. Eine übergreifende Sicht, z.B. auf bestimmte Produkte, war in dem umfassenden Informationsangebot nicht möglich.
- Bestehende Anwendungen enthielten teilweise eigene Suchfunktionalitäten. Die Anwender mussten jedoch im Bedarfsfall in mehreren Applikationen nacheinander separate Suchanfragen absetzen, um z.B. ein vollständiges Bild einer bestimmten Lichtlösung zu bekommen.
- Andere Applikationen, wie z.B. das Content-Management-System, wiesen keine eigene Suchmöglichkeit auf. Die Projektleitung hatte beschlossen, auf die zusätzliche Lizensierung der mit dem CMS angebotenen Suchfunktionalität zu verzichten, da die Einführung einer übergreifenden Portalsuchmaschine bereits geplant war.

Das Anfang 2003 initiierte Projekt ‚Infrastructure NEXT - Teilprojekt Suche' sollte die Funktionen des Zumtobel Staff Portals um eine Suchmaschine zum anwendungsübergreifenden Finden von Inhalten erweitern. Die Projektleitung

entschied, die Suchmaschine in die vorhandene Portalarchitektur zu integrieren. Ihre grundsätzliche Funktionsweise sollte das Projektteam[1] dabei anhand eines Machbarkeitsnachweises (‚Proof of Concept') validieren, bei dem Benutzer mit einer einzigen Suchanfrage integriert Inhalte der Produktdatenbank (PDB) und die für das Portal in einem Content-Management-System (CMS) redaktionell gepflegten Inhalte finden können.

4.1.3. Unterstützungspotenzial

Das Unterstützungspotenzial durch Information Retrieval sahen die Verantwortlichen für ihr ‚ZS Customer Portal' in der Qualitätsverbesserung bestehender Dienstleistungen als professioneller ‚Anbieter von Lichtlösungen'. Eine übergreifende Suche über heterogene Applikationen und die einheitliche Darstellung der Suchergebnisse in einer Übersichtsliste reduziert zudem die Zeit, die Kunden für das Auffinden von Informationen benötigen. Existierende Anwendungen, die keine eigenen Suchmöglichkeiten bieten, können so den Kunden ihre Inhalte bereitstellen und diese themen- und zielgruppengerecht anbieten.

Aufgrund der strategischen Relevanz, den Kunden von Zumtobel Staff eine Portalsuche zur Verfügung zu stellen, verzichteten die Entscheidungsträger auf eine quantitative Bewertung des Unterstützungspotenzials.

Die folgenden Abschnitte skizzieren wesentliche Elemente der im Portal bereitgestellten Benutzerschnittstelle und die verfügbaren Retrieval-Leistungen. Anschliessend ist das im Rahmen von Anforderungsanalyse und Pflichtenheft erarbeitete Prozessmanagement dargestellt, gefolgt von der auf Basis der technischen Spezifikation durchgeführten Systemintegration.

4.1.4. Benutzerschnittstelle und Retrieval-Leistung

Die Portalnutzer können Suchanfragen auf jeder Portalseite eingeben. Abbildung 4-2 zeigt exemplarisch eine Suchanfrage nach dem Produkt ‚Xeno' und die dazugehörigen Suchergebnisse im Customer Portal von Zumtobel Staff. Die Suchmaschine führt eine Volltextsuche im Produktkatalog (PDB), in den Inhalten des Content-Management-Systems (CMS) sowie in einer angebundenen Adressdatenbank (ADB) durch. Die Suchergebnisse werden in einer konsolidierten Liste ausgegeben, deren Einträge mit den Inhalten der jeweiligen Applikation verlinkt sind. Die Anwender haben im dargestellten Beispiel eine einheitliche Sicht auf relevante Inhalte zum Produkt ‚Xeno' aus allen drei Anwendungen, die ohne Portalsuche nicht zur Verfügung steht.

[1] Ziel des vom Autor der vorliegenden Arbeit geleiteten ‚Teilprojekts Suche' war es, eine Suchmaschine auszuwählen und einzuführen.

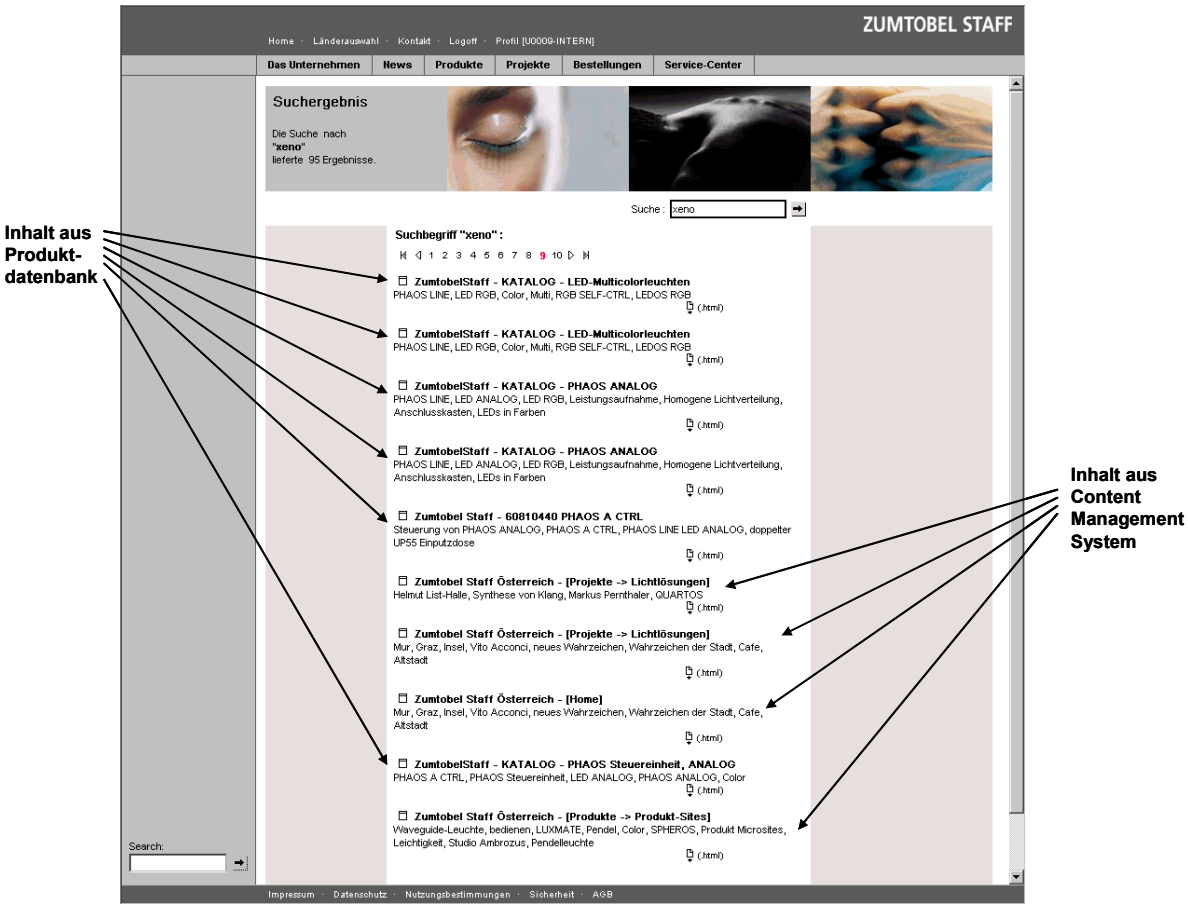


Abbildung 4-2: Suchergebnisse im Customer Portal von Zumtobel Staff

Die zuständigen Projektmitarbeiter erarbeiteten die Anforderungen an die dargestellte Benutzerschnittstelle und die Retrieval-Leistungen in mehreren Workshops mit Anwendervertretern. Der folgende Abschnitt stellt besondere Aspekte des Prozessmanagements dar, die Voraussetzungen zur Bereitstellung der gezeigten Retrieval-Leistungen sind.

4.1.5. Prozessmanagement

Das ‚ZS Customer Portal' besteht aus einzelnen sprachspezifischen Länderportalen. Das Portal für die Schweiz stellt seine Inhalte beispielsweise in Deutsch und Französisch zur Verfügung. Zusätzlich existieren verschiedene Benutzergruppen, die unterschiedliche Inhalte lesen dürfen (Personalisierung). Mehrsprachigkeit und personalisierte Inhalte erfordern, dass Anwender des Customer Portals auch über eine Suchanfrage ausschliesslich die Inhalte finden dürfen, die für sie relevant sind bzw. auf die sie Zugriff haben. Durch diese Anforderungen an eine Suchmaschine sind Zuordnungsregeln von Inhalten zu Anwendern bzw. Anwendergruppen notwendig.

Im Fall von Zumtobel Staff ist die Personalisierung durch eine Klassifizierung von Informationsobjekten realisiert. Jedes Objekt speichert dazu inhaltsbeschreibende Attribute in Form von HTML-Meta-Tags, die als Filter für eine personalisierte

Anzeige dienen. Die Mitarbeiter wählen diese Attribute bei der Erstellung oder Modifikation der Inhalte aus einer Liste mit vorgegebenen Begriffen aus.

Die Begriffe zur Klassifikation von und Suche nach Inhalten, die so genannte Retrieval-Taxonomie, wurden in mehreren Workshops mit den Verantwortlichen für das Content-Management-System und die Produktdatenbank erarbeitet. Tabelle 4-1 zeigt einen Ausschnitt aus der Retrieval-Taxonomie. In der ersten Umsetzung verwendet die Portalsuche ausschliesslich die Angaben zur Sprache, zum Länderportal, zur Quelle und zur Zielgruppe. Die anderen Attribute sollen erst in späteren Ausbaustufen genutzt werden.

Attribut	Wertebereich	Bemerkung
Sprache	DE, FR, IT, EN	Verfügbare Sprachen wie Deutsch, Französisch, Italienisch und Englisch
Länderportal	AT, CH, DE, FR, COM, PCIO	Verfügbare Länderportale für Österreich, Schweiz, Deutschland, Frankreich, internationale Auftritte (COM) sowie ein Extranet (PC International Organization)
Quelle	CMS, PDB, ADB	Integrierte Applikationen wie Content Management System, Produktdatenbank und Adress-datenbank
Informations-typ	News, Kalender, Produktkatalog, Download	
Zielgruppe	io.cust, io.empl	Kunden, Mitarbeiter
Anwendungs-bereich	Büro und Bildung, Industrie und Technik, Präsentation und Verkauf, Kunst und Kultur, Hotellerie und Gastronomie, Sport und Freizeit, Health & Care, Verkehrsbauten und Parkhäuser, Sicherheit und Vertrauen, Active Light	Anwendungsbereiche stimmen mit den Navigationseinträgen des Portals überein
Produktgruppe	Strahler und Stromschienen, Werfer Spiegel Systeme, Modulare Lichtsysteme, Down-/Uplights, Einbauleuchten, Anbau- und Pendelleuchten, Wallwasher, „Steh-, Tisch-, Wand- und LED-Leuchten“, „Lichtbänder, Einzellichtleisten“, Hallenleuchten, Leuchten höherer Schutzart, Gebäudesystemtechnik	Produktgruppen orientieren sich an den Einteilungen des Produkt-katalogs

Tabelle 4-1: Ausschnitt aus der Retrieval-Taxonomie bei Zumtobel Staff

Die Realisierung einer Volltextsuche erfordert keine Retrieval-Taxonomie (s. Abschnitt 3.4.3). Im ZS Customer Portal ist sie dennoch realisiert, um mit den Attributen ‚Sprache’, ‚Länderportal’ und ‚Zielgruppe’ die personalisierte Inhaltsanzeige für die unterschiedlichen Zielgruppen zu gewährleisten. Die Suchmaschine liest diese Attribute für jeden Benutzer beim Aufruf des Portals aus einem zentralen Benutzerverzeichnis aus und ermittelt anhand der Werte die dem jeweiligen Anwender zur Verfügung stehenden Suchindizes.

Definierte Prozesse für die Neuerstellung, Änderung und Löschung sind nicht vorgesehen, da diese Bestandteile der Retrieval-Taxonomie auch in absehbarer Zukunft unveränderlich sind. Die Einbindung der weiteren Attribute der Retrieval-

Taxonomie (‚Anwendungsbereich' und ‚Produktgruppe') in die bestehenden Anwendungen dient ausschliesslich einem Machbarkeitsnachweis. Sie sollen in späteren Ausbaustufen der Portalsuche für eine erweiterte Suchmaske genutzt werden.

CMS, PDB und ADB bestehen aus Datensammlungen, in der die Inhalte für alle Länder und alle Zielgruppen in allen verfügbaren Sprachen abgelegt sind. Sie implementieren folgendes Berechtigungskonzept:

- Leseberechtigungen in der PDB sind von der Länderzugehörigkeit eines Benutzers abhängig. Die Suchmaschine gibt beispielsweise bei einer Suchanfrage eines Anwenders des österreichischen Kundenportals ausschliesslich Inhalte zurück, die in einem speziell für diese Anwender gebildeten Teilindex (‚PDB_AT_Index') enthalten sind.
- Im CMS sind Leseberechtigungen sowohl länder- als auch zielgruppenspezifisch. Zusätzlich zur länderspezifischen Indizierung von Inhalten (analog zur PDB) werden hier durch einen Filtermechanismus bei der Anzeige von Suchergebnissen bei gleicher Suchanfrage für Kunden (Zielgruppe=io.cust) andere Inhalte ausgewiesen als für Mitarbeiter (Zielgruppe=io.empl).

4.1.6. Systemintegration

Die beschriebenen Anforderungen des Prozessmanagements sind in der Retrieval-Architektur des Customer Portals von Zumtobel Staff auf Systemebene mit folgenden Komponenten realisiert:

- Die *Suchmaschine* SAP TREX (Text REtrieval and eXtraction) ist ein integraler Bestandteil des Portals. Die Schnittstelle zwischen dem Benutzer und der Suchmaschine wird über Portlets im SAP Enterprise Portal (SAP EP) bereitgestellt, um in den vorhandenen Indizes zu suchen. Das SAP Content Management (CM) / Knowledge Management (KM) ist für die Kommunikation des SAP EP mit TREX zuständig. Im ZS Customer Portal realisiert das SAP CM/KM die Anbindung der Content-Applikationen wie Produktdatenbank (PDB) und Content-Management-System (CMS) für den Indizierungsprozess.
- In der *Produktdatenbank* (PDB) ruft der SAP Business Connector[2] die in einer SQL-Datenbank gespeicherten Inhalte für die Indizierung und Anzeige auf. Der auf Basis von ASP[3] programmierte ‚Renderer' wandelt die vom Business Connector

[2] Der Business Connector (BC) ist eine Middleware-Komponente von SAP. Er ermöglicht die Integration von SAP R/3-Systemen auf Basis offener Standards wie HTTP als Kommunikationsprotokoll und XML als Dokumentenformat.

[3] Active Server Pages (ASP) ist eine von Microsoft entwickelte Technologie zur Erzeugung serverseitig dynamisch generierter Webseiten.

gelieferten XML-Inhalte durch eine XSL-Transformation[4] in das HTML-Format um. Der Renderer stellt diese HTML-Seiten gegenüber der Suchmaschine durch die Angabe von Parametern in der URL der ASP-Seite für die Indizierung und Anzeige bereit.

- Die Inhalte des *Content-Management-Systems* (CMS) bzw. des Content Personalisation Server (CPS) der Firma Reddot stehen TREX über ein spezielles Portlet zur Verfügung. Ein als Dispatcher bezeichnetes Servlet[5] generiert dazu aus den Inhalten des CPS entsprechende HTML-Seiten.
- Die *Adressdatenbank* (ADB) verwendet dieselben Systemkomponenten wie die Produktdatenbank. Aus Übersichtsgründen ist sie in Abbildung 4-3 nicht separat aufgeführt.
- Alle Applikationen beziehen ihre Benuterzdaten aus einer zentralen Benutzerverwaltung (*Corporate LDAP*).

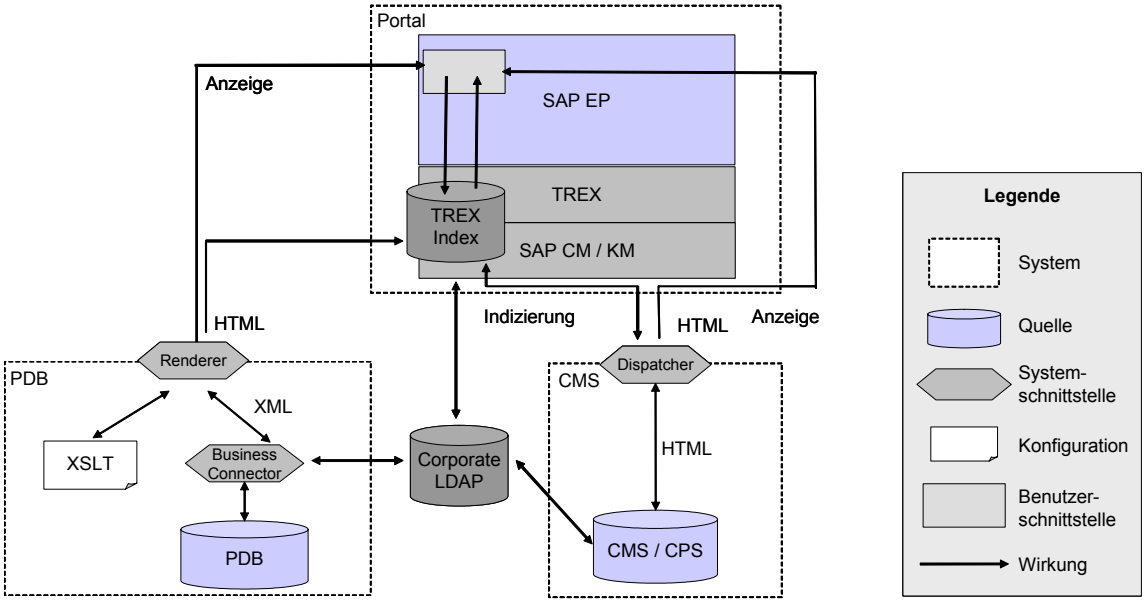


Abbildung 4-3: Retrievalarchitektur bei Zumtobel Staff

Die Suchmaschine TREX benötigt die Angabe einer URL der zu indizierenden HTML-Seite(n). Die Inhalte, die unter dieser Adresse zu finden sind, werden indiziert. TREX ist dabei in der Lage, weiterführenden Links zu folgen, die sich in der anfänglichen HTML-Seite befinden. Über eine entsprechende Verlinkung können so

[4] Extensible Stylesheet Language for Transformations (XSLT) ist eine offizielle Empfehlung des World Wide Web Consortiums (W3C). Sie bietet eine flexible, leistungsfähige Sprache, mit der XML-Dokumente umgewandelt werden können, z.B. in HTML-Dokumente, ein weiteres XML-Dokument oder eine PDF-Datei [Tidwell 2002, 17].

[5] Servlets sind Java-Komponenten, die Web-Clients wie z.B. Web-Browsern serverseitigen Code zur Verfügung stellen [vgl. Balzert 1996, 946].

alle relevanten Inhalte der PDB oder des CMS in den Suchmaschinenindex aufgenommen werden.

Abbildung 4-4 zeigt schematisch den vollständigen Ablauf von der Indizierung bis zur Anzeige von Suchergebnissen am Beispiel der deutschsprachigen Seiten des Customer Portal Schweiz (ohne Produktkatalog).

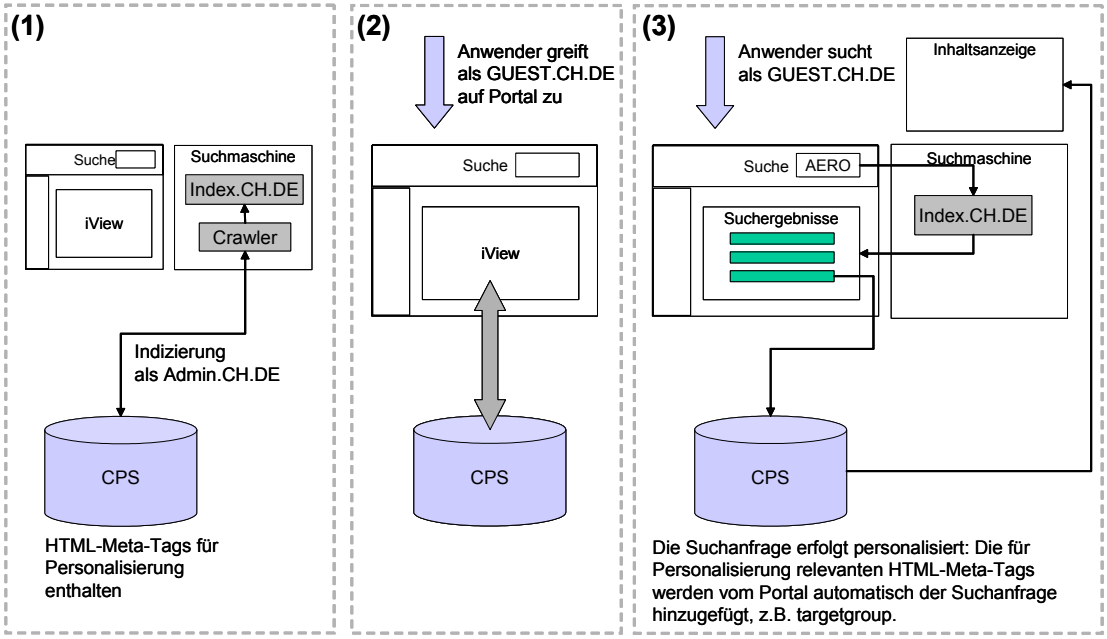


Abbildung 4-4: Indizierung und Suche im CMS von Zumtobel Staff

TREX greift mit einer Administrator-ID (‚Admin.CH.DE') zu, die leseberechtigt auf alle Inhalte des CPS ist, die das deutschsprachige Customer Portal Schweiz beinhaltet und generiert dabei einen Index (‚Index.CH.DE'). Beim Zugriff eines anonymen Benutzers wird vom Portal automatisch eine Authentifizierung vorgenommen, z.B. als ‚GUEST.CH.DE' für die deutschsprachigen Seiten des Schweizer Portals. Nach dem erstmaligen Aufruf der Portalseiten sind alle anonymen Benutzer als ‚GUEST.CH.DE' im Portal eingeloggt (2). Aufgrund einer Zuordnung im Benutzerprofil im zentralen Verzeichnisdienst des Portals (Coporate LDAP) erfolgen Suchanfragen dieser Benutzer in Folge ausschliesslich in den öffentlich verfügbaren Inhalten im ‚Index.CH.DE' (3). Diese Zuordnung zwischen der Sprachwahl des Portalbenutzers und dem verwendeten Index musste in allen Suchanfragefenstern (einfach, erweitert) programmiert werden.

Abbildung 4-5 zeigt die technische Umsetzung der Klassifizierung von Inhalten des CMS mit Attributen der Retrieval-Taxonomie. Über den Eintrag ‚Schlagworte' in der oberen Menüzeile von Reddot wählt ein Redakteur in der Rubrik ‚Crawler' ein Länderkürzel aus. Reddot generiert automatisch eine Startseite für die Suchmaschine TREX, die auf alle Portalseiten mit gleichem Crawler-Attribut verweist. Auf diese Weise bestimmt ein Redakteur, welche Seiten indiziert werden. Für jedes Informationsobjekt innerhalb einer Portalseite gibt der Redakteur über ‚Edit

Personalisation' an, welche Zielgruppe den jeweiligen Inhalt lesen darf. Die beiden Rollen ‚io.cust' und ‚io.empl' entstammen der Retrieval-Taxonomie.

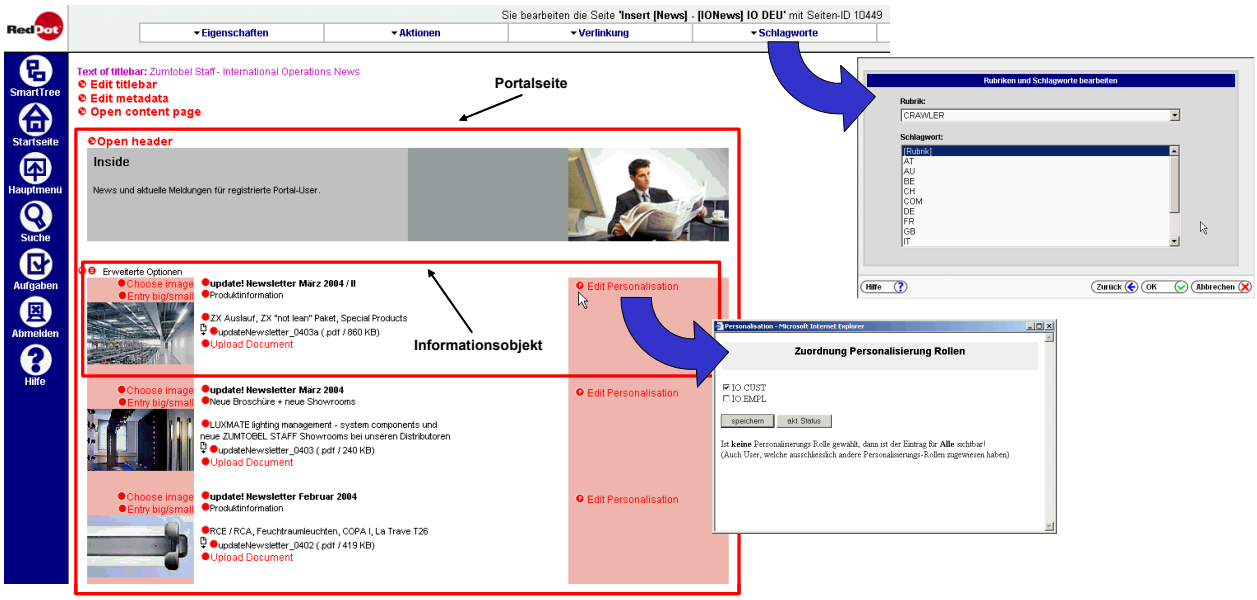


Abbildung 4-5: Klassifikation im Content-Management-System von Zumtobel Staff

Benutzerverwaltung und Authentifizierung übernehmen die beteiligten Applikationen als Basisdienste vom SAP Portal bzw. aus dem zentralen Benutzerverzeichnis (Corporate LDAP). Abbildung 4-6 zeigt exemplarisch die Abfrage dieser Benutzerdaten, die im XML-Format ausgegeben werden. Sie enthalten pro Benutzer beispielsweise die Angabe der Benutzerrolle, z.B. ‚io.empl' für Mitarbeiter, welche die personalisierte Suchergebnisanzeige steuert.

```
<?xml version="1.0" ?>
- <personalization user="skremer" default_language="EN">
  - <index name="PDB-AT-EN">
      <source>PDB</source>
      <filter type="land">AT</filter>
      <filter type="language">EN</filter>
    </index>
  - <index name="PRIVATEWEB_P42-EN">
      <source>PRIVATEWEB_P42</source>
      <filter type="language">EN</filter>
    </index>
  - <index name="CMS-COM-EN">
      <source>CMS</source>
      <filter type="site">COM</filter>
      <filter type="language">EN</filter>
    </index>
    <role>IO.EMPL</role>
    <browserlang>de-ch</browserlang>
  </personalization>
```

Abbildung 4-6: Ausgabe von Benutzerdaten im XML-Format aus dem zentralen Verzeichnis bei Zumtobel Staff

Tabelle 4-2 zeigt das Mengengerüst der Retrieval-Lösung im Februar 2004.

Content-Applikation	Anzahl Informationsobjekte	Änderungsfrequenz	Indexgrösse
Produktkatalog (PDB)	Ca. 20`000 HTML-Seiten, 300 Teaser Files mit einem Gesamtvolumen von 350 MB, 900 Manual Files mit einem Gesamtvolumen von 250 MB	Ca. 2`000 neue Dokumente pro Jahr (können jedoch auch 500 an einem Tag sein) Ca. 700 Löschungen pro Jahr	Ca. 1 GB
Content-Management-System (CMS)	Ca. 2`000 bis 5`000 Content-Objekte mit einem Gesamtvolumen von 500 MB, davon 250 MB in 15 ZIP-Dateien	Ca. 250 neue Dokumente pro Jahr Ca. 250 bestehende Dokumente, die pro Jahr entfernt werden	Ca. 60 MB für die Länderportale IT, AT und COM (PCIO)
Adressdatenbank (ADB)	Ca. 100 Content-Objekte	- vernachlässigbar -	- vernachlässigbar -

Tabelle 4-2: Mengengerüst für indizierte Inhalte bei Zumtobel Staff

4.1.7. Kosten- und Nutzenaspekte

Bei der Infrastruktur konnte bestehende Hard- und Software genutzt werden. Die Suchmaschine SAP TREX war bereits vor Projektbeginn im Lizenzpreis des SAP Enterprise Portals enthalten. Die Umsetzung und Einführung verursachte Personalkosten in Höhe von 138 Personentagen. Diese verteilen sich zu 35% auf die Analyse-, zu 45 % auf die Konzeptions- und zu 20 % auf die Implementierungsphase.

Führungsgrössen zur Nutzungskontrolle waren nicht definiert. Die Rückmeldungen von Anwendern des Systems sind jedoch sehr positiv. Besondere Beachtung findet die hohe Geschwindigkeit, mit der Suchanfragen ausgeführt werden. Viele Anwender äusserten daraufhin, auf die applikationsspezifischen Suchmechanismen in Zukunft evtl. sogar verzichten zu können.

Eine Protokollierungsfunktion zeichnet alle Suchanfragen auf, so dass auch zu einem späteren Zeitpunkt eine Auswertung, z.B. nach den am häufigsten verwendeten Suchbegriffen möglich ist.

4.1.8. Kritische Erfolgsfaktoren und geplante Weiterentwicklungen

Für die Einführung einer Retrieval-Lösung im Customer Portal von Zumtobel Staff sahen die Beteiligten folgende Faktoren als erfolgskritisch an:

- Den Kunden des Customer Portals sollte eine *intuitiv bedienbare Suchmaschine* bereitgestellt werden. Die Umsetzung einer im Internet bereits weit verbreiteten und bekannten Volltextsuche berücksichtigt dies.

- Die Bereitstellung der Portalsuche sollte in *evolutionären Schritten* erfolgen. Dazu würde die übergreifende Suche zunächst auf die Integration der Produktdatenbank und des Content-Management-Systems beschränkt.
- Die Verwendung von SAP TREX innerhalb der vom SAP Enterprise Portal geprägten Infrastruktur stellt sicher, dass die ausgewählte Suchmaschine zu einem späteren Zeitpunkt grundsätzlich auch *innerhalb anderer Teilkonzerne der Zumtobel AG eingesetzt* werden können.
- SAP TREX erleichtert die Integration der Suchmaschine in die bestehende Portal-Infrastruktur, da beispielsweise die bestehende Portalbenutzerverwaltung oder Funktionsschnittstellen des Business Connectors genutzt werden können.

Derzeit prüfen die Projektverantwortlichen den umfassenden Ausbau der Information-Retrieval-Lösung in den Punkten:

- Bereitstellung einer ‚Erweiterten Suche', die ein Finden von Inhalten auf Basis ausgewählter Attribute der Retrieval-Taxonomie erlaubt, z.B. nach Produktgruppe,
- Anbindung weiterer Applikationen, z.B. Referenzprojekte und die Orderstatusverwaltung des SAP R/3-Systems,
- Anpassung der Suchmaschine zur Nutzung in einem Mitarbeiterportal,
- Verwendung der Suche auf Webseiten, die derzeit ohne die Portalplattform betrieben werden und
- Einbindung der Suche in weiteren Anwendungen, um z.B. den Benutzern bei der Navigation im Produktkatalog relevante Dokumente aus anderen Quellen automatisch anzubieten.

4.2. The Information Management Group – Knowledge Pool

„Über die Taxonomieorientierte Suche gelangen die Berater sehr zielgerichtet zu den für die effiziente Ausübung der Projektarbeit erforderlichen Dokumenten. Das Web-Interface der IMG-Wissensmanagementlösung ermöglicht ihnen auch beim Kunden jederzeit Zugang zu qualitätsgesicherten Informationen.“

Martin Liebich, Vice President – Chief Information Officer

4.2.1. Organisation und Ausgangssituation

Die ‚The Information Management Group’ (IMG) ist ein international tätiges Beratungsunternehmen mit weltweit mehr als 600 Beratern. 1989 als Spin-off des Instituts für Wirtschaftsinformatik St. Gallen (IWI-HSG) gegründet, konzentriert sich IMG auf die Implementierung informationstechnisch gestützter, unternehmensweiter Prozesse und zwischenbetrieblicher Geschäftsnetzwerke zur Steigerung der Leistungsfähigkeit ihrer Kunden, z.B. in den Bereichen ‚Business Networking’, ‚Enterprise Resource Planning’, ‚Customer Relationship Management’ und ‚Portal Technologie’. Sie versteht sich als umfassender Dienstleister, der ein Kundenprojekt von der Strategiephase bis zur Implementierung begleiten kann. Die IMG konzentriert sich auf Kunden aus den Branchen Finanzdienstleister, Konsumentenprodukte/Handel, Chemie, Pharma, Automobil sowie Hochtechnologie, Ingenieurswesen und Konstruktion. Das Unternehmen hat seinen Hauptsitz in St. Gallen, Schweiz, und verfügt über Niederlassungen in Europa, USA und Asien.

Die Mitarbeiter der IMG erbringen ausschliesslich immaterielle, sehr wissensintensive Dienstleistungen. Die Geschäftsleitung erkannte, dass diese mit der im Jahr Frühjahr 2001 erreichten Personalstärke von fast 600 Mitarbeitern nicht mehr ausreichend unterstützt wurde:

- Die in Projekten erstellten Dokumente (Angebote, Business Cases, Checklisten, Erfahrungen, Präsentationen, Arbeitsberichte etc.) wurden bislang nicht ausreichend als Basis für die Arbeit in anderen Projekten unternehmensweit nutzbar gemacht.
- Das starke Personalwachstum erschwerte es, den Austausch von Informationen und die Vermeidung von Doppelspurigkeiten über die bisher intensiv gepflegten informellen Netzwerke weiter zu betreiben.
- Eine unternehmensweite Ablage, Verteilung und Findung relevanten Wissens in der Vielzahl heterogener Systeme mit individuellen Strukturen und Abläufen war nicht möglich.

Das im Rahmen einer umfassenden Wissensmanagementinitiative initiierte Projekt ‚Communities & Knowledge Pool' hatte das Ziel, den Mitarbeitern das Ablegen und Wiederfinden von elektronischen Inhalten durch die Strukturierung des Informationsangebotes und Bereitstellung eines effizienten, unternehmensweiten Informationssystems für Wissensmanagement zu erleichtern.

Ein zentrales Ergebnis war die Knowledge-Management-Architektur in Abbildung 4-7. Die Komponenten dieser Architektur sollten die Mitarbeiter bei ihren Prozessen wie folgt unterstützen: Communities, Datenbankapplikationen auf Basis von Lotus Notes, bilden die Grundlage des Dokumentenaustauschs bei der Zusammenarbeit in Projekten. Sie können von jedem Mitarbeiter beantragt werden (‚Request Form') und dienen einem klar definierten Zweck und Personenkreis. Neben der Unterstützung von Projektmitarbeitern werden Communities auch zentralen Organisationseinheiten der IMG für die strukturierte Ablage von Dokumenten zu einem bestimmten Thema oder Service bereitgestellt. Die Anwender einer Community können Dokumente einstellen, bearbeiten, publizieren und lesen. Bei der Kategorisierung der Inhalte verwenden sie eine einheitliche und IMG-weit gültige Taxonomie. Wichtige und für einen grösseren Benutzerkreis relevante Dokumente sendet jeder Mitarbeiter aus einer Community in einen zentralen Pool, das Portal ‚Knowledge Pool'. Das Portal ermöglicht eine kombinierte Volltextsuche und attributbasierte Suche in den veröffentlichten Dokumenten. Bei der attributbasierten Suche stehen die in der Taxonomiedatenbank hinterlegten Begriffe zur Verfügung.

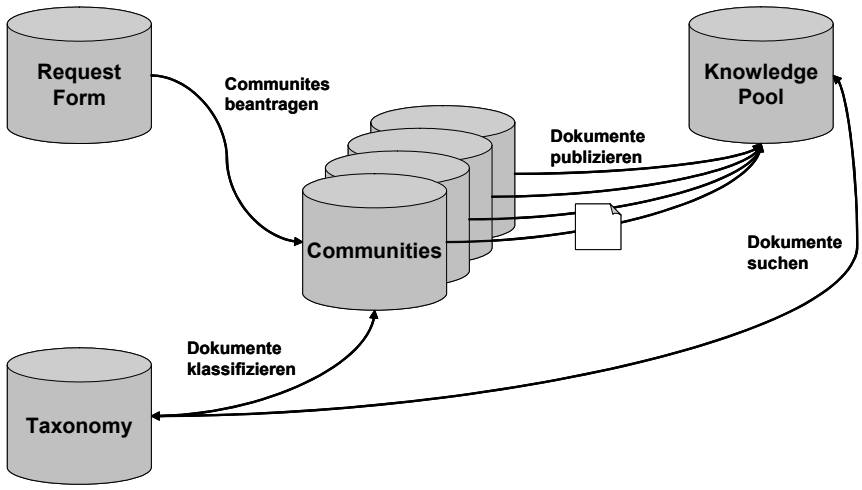


Abbildung 4-7: Knowledge-Management-Architektur der IMG

Zur Unterstützung der Zusammenarbeit auf Projekten liessen die Wissensmanagementverantwortlichen zunächst die Community-Datenbanken entwickeln. Im Januar 2002 stellten sie diese im Unternehmen zur Nutzung bereit.

4.2.2. Treiber und Herausforderungen für Information Retrieval

- Die Communities boten den beteiligten Projektmitgliedern bereits eigene Ablage-, Navigations- und Suchfunktionen. Eine übergreifende Recherche über alle Communities, z.B. nach gewissen Dienstleistungen oder Technologien war jedoch nicht möglich.
- Die Strukturierung von Inhalten im Kontext einzelner Projekte war aufgrund der begrenzten Themenstellung und dem überschaubaren Nutzerkreis vergleichsweise einfach. Für die umfassende Ablage und Suche im Portal ‚Knowledge Pool' musste hingegen eine unternehmenseinheitliche Taxonomie implementiert werden.
- Die Publikation von Inhalten im Knowledge Pool musste bestehende Zugriffsbeschränkungen der Communities abbilden.

Anfang 2002 entschieden die Verantwortlichen, die geplante KM-Architektur in einer zweiten Ausbaustufe um das Portal ‚Knowledge Pool' und eine Suchmaschine zum übergreifenden Finden von Inhalten zu erweitern[6]. Die Projektleitung beschloss, auch Portal, Suchmaschine und Taxonomiedatenbank auf Basis der bereits für die Umsetzung der Community-Datenbanken verwendete Groupware-Plattform Lotus Notes zu entwickeln.

4.2.3. Unterstützungspotenzial

Die Bereitstellung von ‚Knowledge Pool' und Suchmaschine vervollständigt die geplante KM-Architektur der IMG. Der Knowledge Pool enthält damit alle unternehmensweit aus den Communities publizierten Informationsobjekte. Als zentrale Anlaufstelle dient er den Beratern zur Navigation und datenbankübergreifenden Suche in qualifizierten Inhalten.

Das Unterstützungspotenzial durch Information Retrieval besteht somit in:

- der *Zeitersparnis* durch kürzere Recherchezeiten, da die Informationsobjekte für alle Mitarbeiter schnell verfügbar sind,
- der *Qualitätssteigerung* der Arbeitsergebnisse durch den besseren Zugriff auf bestehende Dokumente mit gleicher Themenstellung und
- der *Kostenreduktion* durch Vermeidung von Doppelarbeiten und Fehlern aus Unkenntnis sowie der zeitnahen Verfügbarkeit von Informationsobjekten.

[6] Der Autor der vorliegenden Arbeit nahm die Rolle eines Projektmitarbeiters bei der Erarbeitung und Abstimmung der Retrieval-Taxonomie wahr, die zur Klassifizierung von Dokumenten in den Communities und zur attributbasierten Suche im Portal dienen sollte.

Wegen des grossen Handlungsdrucks, in kurzer Zeit eine hochwertige Lösung zur Unterstützung der Mitarbeiter bei ihren informationsintensiven Aufgaben bereitzustellen, verzichteten die Entscheider bewusst auf eine quantitative Potenzialanalyse.

4.2.4. Benutzerschnittstelle und Retrieval-Leistung

Die Mehrzahl der Berater arbeitet vor Ort bei Kunden, wo sie keine Verbindung zum Firmennetzwerk haben. Die Komponenten der KM-Architektur stehen daher sowohl online im Web-Browser als auch offline im Lotus Notes Client zur Verfügung. Die Retrieval-Leistungen in Form von taxonomieorientierter Navigation, Volltextsuche und attributbasierter Suche sind dabei identisch.

Im Web-Browser dient das Portal ausschliesslich der Informationsfindung. Abbildung 4-8 zeigt exemplarisch eine Suchanfrage und die dazugehörigen Suchergebnisse im Portal der IMG. Die Suchmaschine führt in diesem Fall eine attributbasierte Suche durch. In Listenform werden die Suchergebnisse angezeigt, die auf im ‚Knowledge Pool' vorhandene Informationsobjekte verweisen.

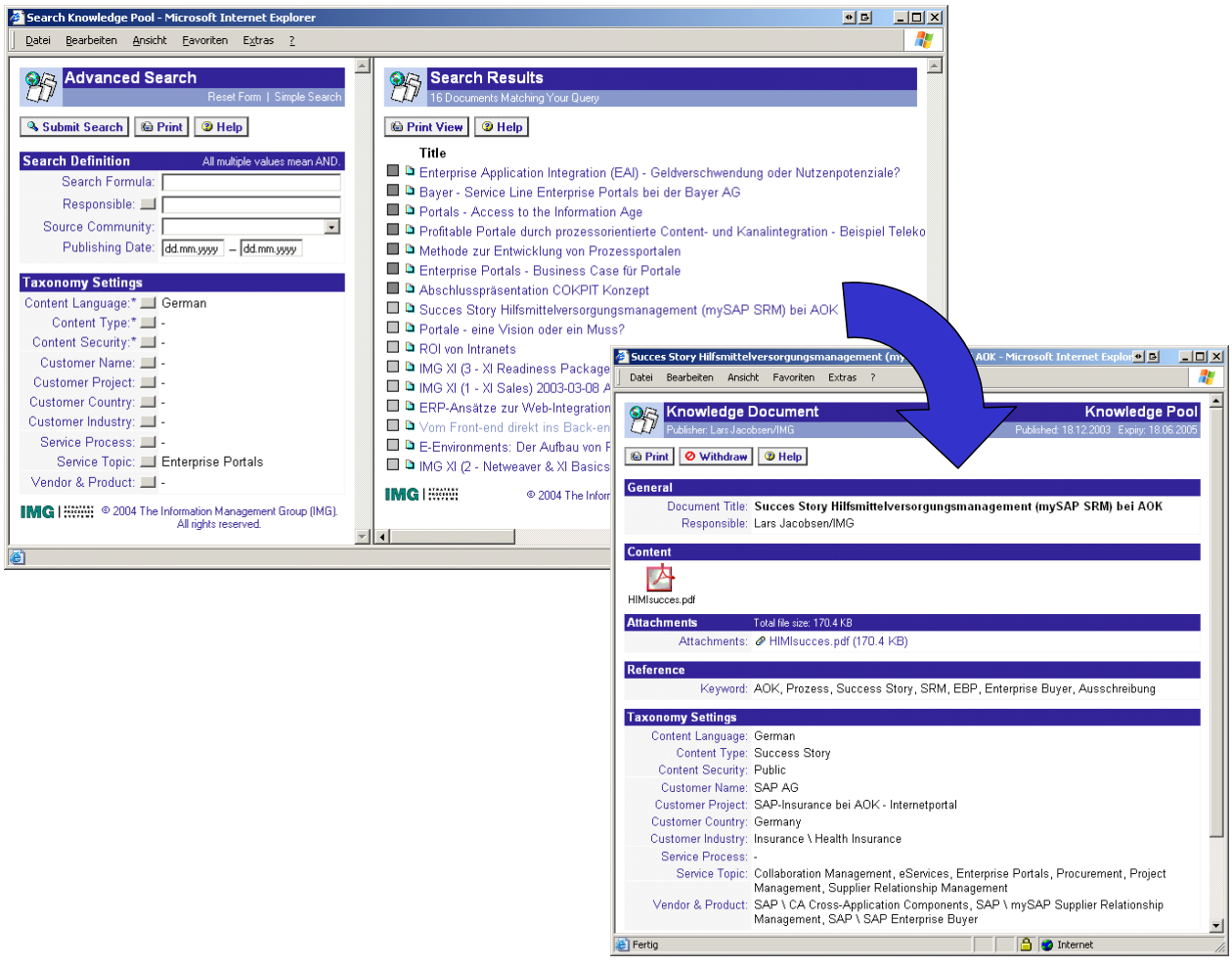


Abbildung 4-8: Information Retrieval im Knowledge Pool der IMG

Während Volltextsuchanfragen durch direkte Eingabe eines gesuchten Begriffs ausgeführt werden können, wurde zur benutzergerechten Formulierung komplexer Suchanfragen eine Eingabemaske für die ‚Erweiterte Suche' bereitgestellt. Im Benutzerdialog stehen Schaltflächen zur Auswahl entsprechender Werte aus der Retrieval-Taxonomie zur Verfügung (s. Abbildung 4-8, obere Grafik, Bereich ‚Taxonomy Settings').

4.2.5. Prozessmanagement

Basis dieser umfassenden Retrieval-Leistungen bildet eine unternehmensweit gültige Taxonomie. Sie dient einerseits dazu, Dokumente in den Community-Datenbanken zu klassifizieren. Zum anderen stehen die Attribute und Werte in den ‚Taxonomy Settings' den Portalbenutzern bei der attributbasierten Suche bereit (s. Abbildung 4-8).

Tabelle 4-3 zeigt einen Ausschnitt aus der Retrieval-Taxonomie, die in mehreren Arbeitssitzungen vom Projektteam erarbeitet und mit einzelnen Fachbereichen innerhalb des Unternehmens abgestimmt wurde.

Attribut	Wertebereich
Responsible	- alle verfügbaren Mitarbeiter -
Source Community	- alle verfügbaren Communities -
Publishing Date	Start- und Enddatum
Content Language	Chinese, Dutch, English, French, German, Italian, Japanese, Polish, Portuguese, Russian, Spanish
Content Type	Article, Best Practice, Brochure, Business Blueprint, Business Case, Checklist, Code, Concept, Contract, Direction, FAQ, Guideline, Lessons Learned, Link, Manual, Market Report, Media, Method, Presentation, Price List, Procedure Model, Process Description, Project Plan, Project Review, Proposal, Report, Success Story, Technology Report, Template, Training Material, White Paper, Working Paper
Content Security	Confidential, Internal Use, Public
Customer Name	- alle verfügbaren Kundennamen -
Customer Project	- alle verfügbaren Kundenprojekte -
Customer Country	Austria, Belgium, China, Denmark, Finland, France, Germany, Great Britain, Italy, Japan, Luxemburg, Norway, Poland, Portugal, Spain, Sweden, Switzerland, The Netherlands, USA
Customer Industry	Aerospace & Defense, Automotive, Banking, Chemicals, Communications, Consumer Products, Electronics & High Tech, Engineering & Construction, Fashion & Sports, Financial Services, Gardening & Landscaping, Health Services, Higher Education & Research, Insurance, Media & Entertainment, Metals & Mining, Mill Products, Other Manufacturer, Other Retailer, Other Service Provider, Partner, Pharmaceuticals, Professional Services, Public Services, Transportation & Logistics, Travel & Hospitality, Unknown Segments, Utilities & Energy Services
Service Process	Asset Management, Distribution, Enterprise Development, Enterprise Management, Finance & Controlling, Human Resources, Information Management, Manufacturing / Production, Marketing & Sales, Materials Management, Product Innovation & Design, Commerce, Content & Community, Finance Chain, Maintenance & Repair, Product Life Cycle, Supply Chain

Service Topic	Business Enhancement, Business Model of the Information Age, Business Networking, Business Process Outsourcing, Business Process Redesign, Change Management, Collaboration Management, Content Management, Customer Relationship Management, Customer Support, Data Warehouse Services, eBusiness in General, eServices, Enterprise Application Integration, Enterprise Management, Enterprise Portals, Enterprise Resource Planning, Financial Consolidation, Information Process Management, IS/IT Management, Knowledge Management, Master Data Management, Mobile Commerce, Packaged Software Implementation, Performance Measurement & Benchmarking, Procurement, Product Lifecycle Management, Project Management, Public Key Infrastructure, Realtime Management, Software Development, Strategy Development, Supplier Relationship Management, Supply Chain Management
Vendor & Product	Arcplan, Ascential, Baan, Bea, BridgePoint, Business Objects and Acta, Chordiant Software, Cognos, Crystal Reports, Descartes, Documentum, Gauss, HAHT Commerce, Hyperion, IBM, Impress, Informatica, Intershop, Microsoft, MIS, Oracle, PeopleSoft, Peregrine, SAP, SAS, Siebel Systems, SUN, Tibco, webMethods

Tabelle 4-3: Retrieval-Taxonomie der IMG

Die Abstimmung der Branchenwerte (‚Customer Industry') und der Dienstleistungsschwerpunkte (‚Service Topic') nahm einen Grossteil der insgesamt fast 20 Personentage zur Erstellung der Retrieval-Taxonomie in Anspruch, da diese verschiedensten Anforderungen gerecht werden muss und durch die Veränderung der geschäftlichen Tätigkeit einem Wandel unterliegen.

Prozesse für die Neuerstellungen, Änderungen und Löschungen von Werten der Retrieval-Taxonomie berücksichtigt die Information-Retrieval-Lösung der IMG wie folgt:

- Jeder Mitarbeiter kann die Erstellung eines neuen Taxonomiewertes beantragen. Die Community-Datenbank unterstützt diesen Ablauf durch die Bereitstellung einer Schaltfläche, die einen entsprechenden Antrag per E-Mail an eine Kommission weiterleitet. Diese prüfen, ob der Wert bereits durch einen anderen Begriff abgedeckt ist oder im Konflikt mit anderen Werten steht. Bei einem positiven Entscheid nimmt die Kommission diesen Begriff in der Taxonomiedatenbank auf. Der Antragsteller wird sowohl bei der Aufnahme als auch bei der Ablehnung per E-Mail mit der jeweiligen Begründung benachrichtigt. In 2003 wurden in diesem Prüfprozess ca. 10 bis 15 neue Werte aufgenommen.
- Eine Änderung von Begriffen wird im System nicht nachgepflegt, d.h. bestehende Werte werden nicht durch andere Werte aktualisiert. Somit findet auch keine Reklassifizierung von existierenden Dokumenten statt.
- Begriffe werden nicht gelöscht. Sie stehen weiterhin in Auswahllisten zur Klassifikation von Dokumenten innerhalb der Communities und in der erweiterten Suche zur Verfügung.

Die Klassifikation von Dokumenten mit Werten der Retrieval-Taxonomie erfolgt in allen Communities durch den Zugriff auf die zentrale Taxonomiedatenbank. Im

Bereich ‚Taxonomie Settings' wählt der Autor eines Dokuments in umfassenden Benutzerdialogen die gewünschten Werte aus.

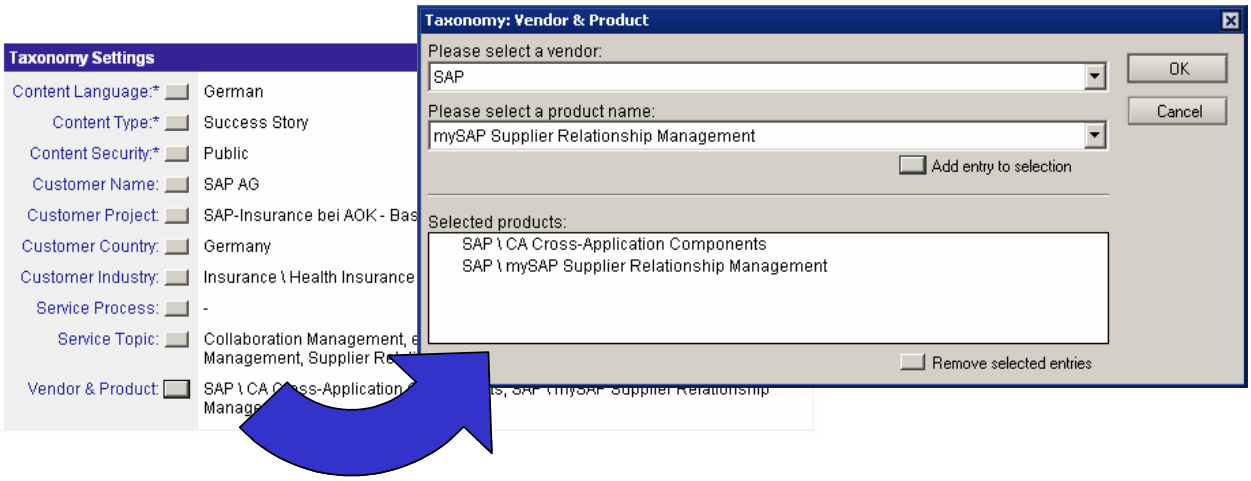


Abbildung 4-9: Klassifikation eines Informationsobjekts bei IMG

Die Attribute ‚Customer Name' und ‚Customer Project' stellen eine Besonderheit dar, da ihre Werte nicht von der Taxonomiedatenbank geliefert werden. Die Retrieval-Lösung der IMG greift hier auf eine andere bestehende Anwendung zu. Das so genannte ‚Account Management' beinhaltet alle verkaufsrelevanten Informationen und stellt auch die in den Communities und im Knowledge Pool benötigten Werte für Kundennamen und Projektbezeichnungen bereit.

Für die Integration von Anwendungen und die Berücksichtigung von Zugriffsberechtigungen im Portal waren keine zusätzlichen Pflegeprozesse notwendig, da Portal, Suchmaschine und Content-Applikation in einer einzigen Applikation, dem ‚Knowledge Pool' umgesetzt sind. Diese Vorgehensweise ermöglicht die Nutzung eines gemeinsamen Benutzerverzeichnisses oder eines einheitlichen Authentifizierungsmechanismus in allen Anwendungen.

4.2.6. Systemintegration

Die Umsetzung der KM-Architektur der IMG basiert auf der konsequenten Nutzung der vorhandenen Lotus-Notes-Infrastruktur. Communities, Knowledge Pool und Taxonomie-Datenbank sind allesamt Domino-Applikationen. Die Communities basieren auf einer gemeinsamen Datenbankschablone (‚Template') für Lotus Notes Release 5.0.x. Sie befinden sich auf einem zentralen Lotus Domino 5.0.x Server, der den Beratern den Zugriff per Lotus Notes Client oder Web-Browser erlaubt.

Portal, Suchmaschine und Content-Applikation sind in einer einzigen Applikation realisiert. Dadurch entfällt z.B. der Aufwand für die Integration von Authentifizierung, Autorisierung oder Benutzerverwaltung.

Durch die Verwendung eines einheitlichen Templates für die Communities existiert nur ein Typ von Anwendungsschnittstelle gegenüber dem Portal bzw. der Suchmaschine. Grundlage des Austausches von Informationsobjekten zwischen den beteiligten Applikationen ist das in Abbildung 4-10 dargestellte gemeinsame Datenmodell. Die attributbasierte Suche ist z.B. dadurch erleichtert, dass identische Felder in allen Anwendungen gleich benannt sind.

Dimension	Attribute	Taxonomy Item	Multi value	Mandatory
Content	Language	x		x
	Type	x	x	x
	Subject			
	Body			
	Category		x	x
	Free Keyword		x	
Customer	Name	x		
	Project	x		
	Country	x	x	
	Industry	x	x	
Service	Process	x	x	
	Topic	x	x	
Vendor	Vendor Name	x	x	
	Product Name	x	x	
Modification Info	Creator			x
	Created			x
	Last Modifier			x
	Last Modified			x
	History		x	x
Security	Additional Authors		x	
	Readers		x	
Misc. Info	Community Name			x
	Project Name			
	Responsible		x	x
	Submitted Date			x
	Expiry Date			x
Internal Fields	Form			x
	Document UID			x
	Build No.			x
	Comm. Replica ID			x
	Document Status			x
	Security	x		x

Legende:
Taxonomy Item: Attribut für Suche und Navigation
Multi value: Mehrfachfeld
Mandatory: Pflichtfeld

Abbildung 4-10: Datenmodell der Community-Datenbanken und des Knowledge Pool der IMG

Zur Indizierung und Suche kommt im Knowledge Pool der IBM Global Text Retriever (GTR) zur Verwendung. Er ist seit dem Release 5.0.3 ein integraler Bestandteil des Lotus Domino Servers und stellt Volltextsuche und attributbasierte Suche innerhalb jeder Domino-Applikation bereit. Abbildung 4-11 zeigt den Administrationsdialog zum Erstellen, Aktualisieren oder Löschen des Index. Bei der erstmaligen Erstellung gibt der Administrator an, ob in den Dokumenten vorhandene Dateianhänge oder auch verschlüsselte Felder mit indiziert werden.

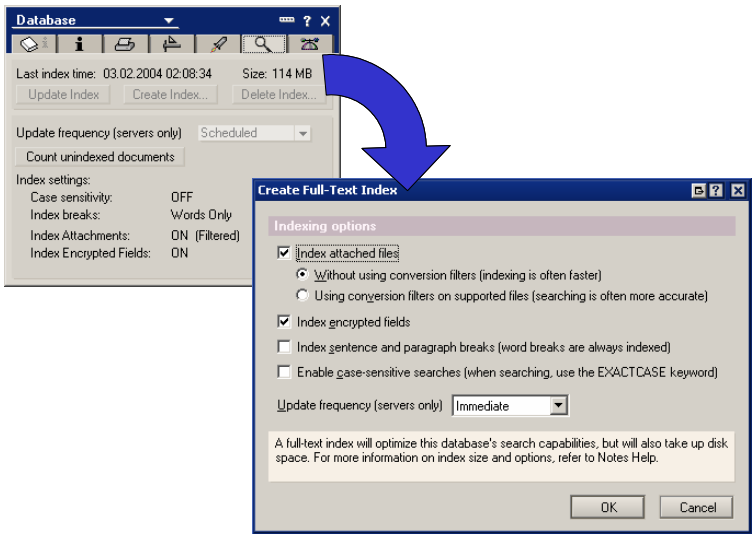


Abbildung 4-11: Indexmanagement im ‚Knowledge Pool' von IMG

Die Aktualisierung des Index erfolgt i.d.R. zeitgesteuert durch den Domino Server, so dass hier keine weiteren Pflegeaufwände entstehen.

Tabelle 4-4 zeigt das Mengengerüst der Retrieval-Lösung im Februar 2004.

Content-Applikation	Anzahl Informationsobjekte	Änderungsfrequenz	Indexgrösse
Knowledge Pool (enthält die aus den Communities übermittelten Dokumente)	Ca. 600 Dokumente mit einem Gesamtvolumen von ca. 550 MB	Ca. 20 neue oder geänderte Dokumente pro Monat	Ca. 115 MB

Tabelle 4-4: Mengengerüst für indizierte Inhalte bei IMG

4.2.7. Kosten- und Nutzenaspekte

Die Planung, Umsetzung und Einführung des Portals ‚Knowledge Pool' mit seinen Retrieval-Funktionen und der dazugehörigen Taxonomiedatenbank verursachte im Rahmen des gesamten KM-Projekts anteilig interne Personalkosten in Höhe von 50 Personentagen. Der externe Programmierer brauchte für die technische Umsetzung ca. 30 weitere Personentage. Führungsgrössen zur Nutzungskontrolle waren nicht definiert. Die integrierte Nutzungsmessung des Knowledge Pools weist jedoch im Durchschnitt 10'000 Lesezugriffe pro Woche aus.

4.2.8. Kritische Erfolgsfaktoren und geplante Weiterentwicklungen

Für Aufbau und Betrieb der Information Retrieval-Lösung sahen die Verantwortlichen der IMG folgende Faktoren als erfolgskritisch:

- Alle KM-Anwendungen sollten *intuitiv und einfach bedienbar* sein, z.B. auch die Suchfunktionen im ‚Knowledge Pool'. Die Nutzung und Weiterentwicklung der bei allen Mitarbeitern bekannten Infrastruktur auf Basis von Lotus Notes

berücksichtigt die Integration dieser neuen Anwendungen in die bestehende Arbeitsumgebung.

- Die Erstellung, Ablage und Auffindung von Inhalten sollten mit *einfach gehaltenen Prozessen* erfolgen. Durch die Verwendung von Datenbankschablonen sind die Abläufe in allen Communities gleich. Durch die Verwendung einer gemeinsamen Taxonomie wird das Ablegen und Auffinden von Informationsobjekten vereinheitlicht.
- Um *keine neuen Schnittstellen* zu schaffen, sollte die bestehende Infrastruktur weitestgehend genutzt werden können. Die Umsetzung auf Basis von Lotus Notes / Domino berücksichtigt diese Forderung.

Mit Sachstand vom Frühjahr 2004 ist keine umfassende Weiterentwicklung der Information-Retrieval-Lösung bei der IMG geplant.

4.3. Winterthur Versicherungen – McB-Projektportal

„Die Kombination aus Navigation und Suche ist für die Projektmitglieder ein unabdingbares Hilfsmittel, in der Vielzahl der Inhalte und in grossen Projekten punktgenau die Dokumente zu finden, die sie für ihren aktuellen Arbeitsschritt benötigen."

Walter Dyttrich, McB Portal Manager, Winterthur Versicherungen

4.3.1. Organisation und Ausgangssituation

Die Winterthur Versicherungen (Life & Pensions) bietet weltweit Versicherungen zum Schutz von Angehörigen im Todesfall und zur Sicherung der Altersversorgung von Privatpersonen an. Sie agiert dabei als eigenständiger Geschäftsbereich im Verbund der Credit Suisse Group, in welche die Winterthur Versicherungen seit 1997 integriert ist. Ende 2001 erwirtschafteten 7'400 Mitarbeiter einen Umsatz von CHF 8'713 Mrd. Mit einem Prämienvolumen von CHF 17,4 Mrd. steht die Winterthur Versicherungen in der Schweiz an der zweiten Stelle, in Europa auf Platz Zehn in der Versicherungsbranche.

Lebens- und Altersvorsorgeversicherungen sind Produkte mit einer Laufzeit von mehreren Jahrzehnten. Die Vertragsbedingungen jedes einzelnen Versicherungsvertrages werden dabei zum Zeitpunkt des Abschlusses mit dem Versicherungsnehmer festgelegt und gelten i.d.R. für die gesamte Laufzeit. Neuverträge werden mit den jeweils aktuellen Standardkonditionen abgeschlossen, im Versicherungsportfolio bleiben aber weiterhin die bis dahin geschlossenen alten Versicherungsverträge enthalten. Letztere, in der Versicherungsbranche als ‚Closed Blocks' bezeichneten Altverträge, sind bei der Winterthur Life & Pensions aufgrund der dynamischen Marktentwicklung und einer zunehmenden Globalisierung in den letzten Jahren stark angewachsen. Zur Reduktion der damit verbundenen erheblichen Verwaltungskosten initiierte Winterthur Life & Pensions Ende 2000 die Initiative ‚Management of Closed Blocks' (McB) mit dem Ziel, Verwaltungsprozesse, versicherungsmathematische Modelle und informationstechnische Systeme zur Bewirtschaftung der Altverträge sukzessive zu vereinfachen und zu vereinheitlichen. Diese Standardisierung ist jedoch im Regelfall nur länderspezifisch möglich, da unterschiedliche Rechtsvorschriften zu beachten sind. Gleichwohl können viele länderspezifische Erfahrungen der McB-Aktivitäten auf andere Länder übertragen werden.

Die Konsolidierung der bisherigen Strukturen wurde durch ein Beraterteam von ca. 25 Personen getragen. Diese McB-Consultants waren Spezialisten, die besondere Kompetenzen in den Bereichen Verwaltungsprozesse, versicherungsmathematische Modelle und informationstechnische Systeme aufwiesen (so genannte ‚Centers of Competence') und als interne Berater die McB-Projekte in den einzelnen Ländern betreuten. Ein McB-Consultant betreute dabei i.d.R. parallel mehrere McB-Projekte in

unterschiedlichen Ländern und wurde durch lokale Spezialisten der jeweiligen Landesgesellschaft unterstützt. Die McB-Consultants, welche sich an häufig wechselnden Standorten in ganz Europa aufhielten, mussten zu jeder Zeit auf das Wissen der anderen Kollegen zurückgreifen können. Dieser Austausch erfolgte hauptsächlich dokumentenbasiert. Die Effektivität der McB-Initiative hing dabei auch von der schnellen Verfügbarkeit der Arbeitsergebnisse (z.B. Prozessbeschreibungen) für die unmittelbar Beteiligten, wie auch für andere interessierte Projektmitglieder, ab.

4.3.2. Treiber und Herausforderungen für Information Retrieval

- In der frühen Projektphase konnte der Wissensaustausch innerhalb des McB-Projekts systemtechnisch durch den Versand von E-Mails unterstützt werden. Jeder Empfänger bekam die Unterlagen wie Vorlagen, Statusberichte oder Projektreports einzeln zugeschickt. Bei der Weiterbearbeitung entstanden dadurch häufig unterschiedliche Versionen eines Dokuments oder einzelne Mitarbeiter (z.B. nach Neueintritt) verfügten nicht über alle relevanten Unterlagen.
- Dokumente wurden teilweise auch auf einem zentralen Fileserver abgelegt. Dieser ermöglichte jedoch nur den Zugriff innerhalb der Schweiz und war somit im internationalen Rahmen der McB-Initiative nicht verwendbar. Daneben wirkt sich die hohe Netzbelastung negativ auf die Geschwindigkeit des Datentransfers aus.
- Neben dem Dokumentenaustausch waren zeitnahe Informationen über den Status der einzelnen McB-Projekte für das zentrale Projektmanagement unverzichtbar. Voraussetzung für Wiederverwendbarkeit und Vergleichbarkeit war die Definition eines einheitlichen Vorgehensmodells für alle McB-Projekte, dessen Struktur im verwendeten System abgebildet werden musste. Diese Anforderungen konnten weder durch den Versand von E-Mails noch durch die Nutzung einer Fileserverstruktur vollständig abgedeckt werden.

Aufgrund der zunehmenden und wiederholten Nachfragen aus den beteiligten Ländern nach projektspezifischen Informationen und der erwarteten Entwicklung des Informationsaufkommens mit der McB-Initiative entschied sich die Projektleitung im November 2000 für den Aufbau eines Projektportals im Winterthur Intranet. Die Plattform ‚McB-Portal' sollte allen Beteiligten (z.B. Projektzentrale, Projektteams und projektexternen Anspruchsgruppen) das von ihnen benötigte Wissen prozessorientiert über Navigation und Suche zur Verfügung stellen.

4.3.3. Unterstützungspotenzial

Für die Verantwortlichen standen qualitative Nutzenaspekte im Vordergrund:

- Das Projektportal erleichtert den international agierenden McB-Beratern den Zugriff auf bestehende Ergebnisse, Checklisten und Arbeitsanleitungen.
- Doppelspurigkeiten und Fehler aus Unkenntnis können vermieden werden, da die Inhalte des Portals grundsätzlich allen Beteiligten zur Verfügung stehen.
- Die Kommunikation der beteiligten Berater wird untereinander verbessert, da mit dem McB-Portal ein gemeinsamer Ordnungsrahmen für das Ablegen und Auffinden von Inhalten geschaffen wird.

Das aufgabenbezogene Ablegen und Finden von Dokumenten reduziert zudem die Zeit für Informationsrecherche durch eine an den Prozessen der Beteiligten ausgerichtete Navigation und Suche. Die Nutzung des Portals ist von allen internationalen Standorten aus möglich und verringert dadurch beispielsweise die Kosten für den Versand von E-Mails.

4.3.4. Benutzerschnittstelle und Retrieval-Leistungen

Im McB-Portal stehen den Mitarbeitern Volltextsuche, attributbasierte Suche und eine taxonomieorientierte Navigation zur Verfügung. Die Navigation ist hierarchisch aufgebaut, wobei die oberste Navigationsebene aus Einträgen für sechs Leistungs- und Unterstützungsprozesse besteht.

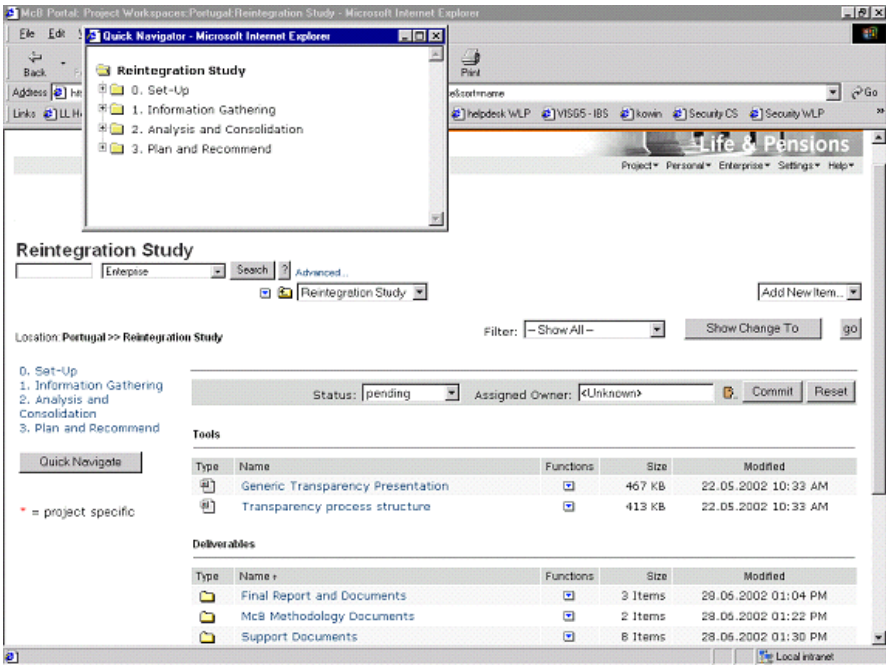


Abbildung 4-12: Prozessorientierte Navigation im McB-Portal

Jeder Prozess ist dabei in weitere Bestandteile untergliedert (Phase, Aktivität, Aufgabe). Abbildung 4-12 zeigt dies exemplarisch für den Prozess ‚Reintegration Study' mit den Phasen ‚0. Set-Up', ‚1. Information Gathering', ‚2. Analysis and Consolidation' und ‚3. Plan and Recommend'. Auf allen Navigationsebenen werden

vorhandene Beschreibungen, Hilfsmittel und Arbeitsdokumente zur Unterstützung der jeweiligen Aufgabe angezeigt. Über eine einblendbare Navigationsübersicht (‚Quick Navigator') können gezielt einzelne Prozesse, Phasen, Aktivitäten und Aufgaben innerhalb eines Projekts angesteuert werden.

Suchanfragen sind im gesamten Portal unter Berücksichtigung der Benutzerberechtigungen sowohl im Volltext als auch attributbasiert möglich. Der Suchraum kann zusätzlich auf bestimmte Projekte eingeschränkt werden. Abbildung 4-13 zeigt die realisierte Kombination aus Volltextsuche nach dem Begriff ‚Benutzertest' und attributbasierter Suche nach den Dokumenteneigenschaften ‚Competence = IT'.

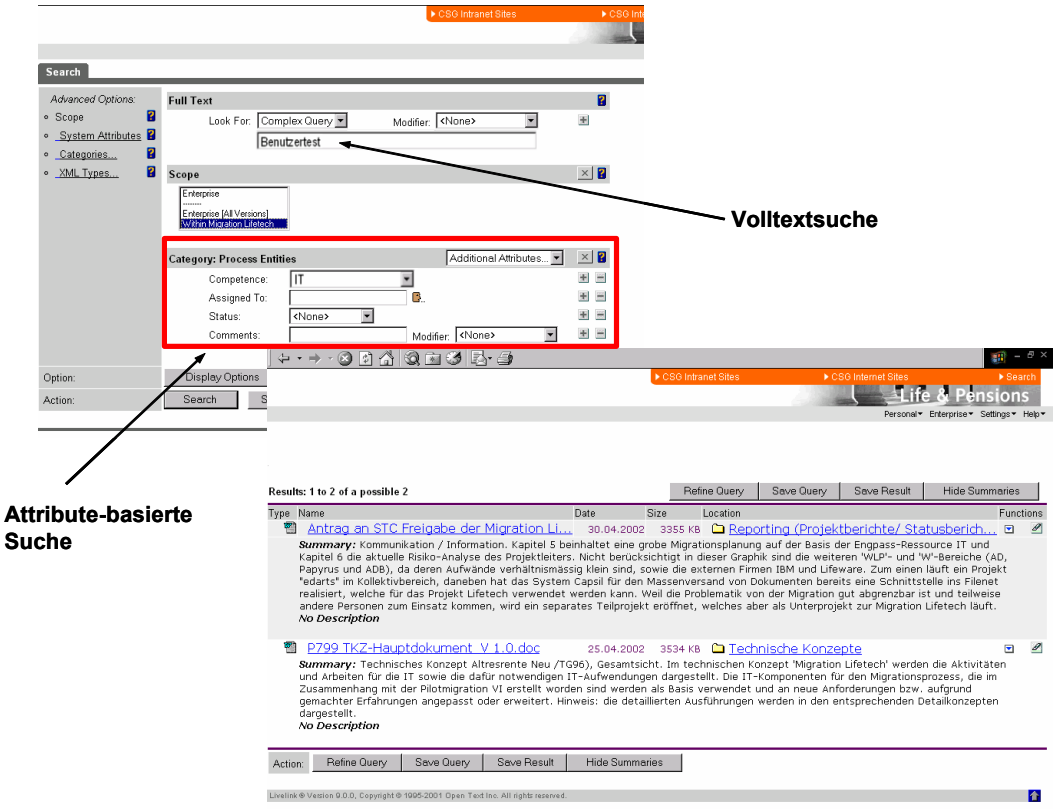


Abbildung 4-13: Suche im Projektportal der Winterthur Versicherung

Die zuständigen Projektmitarbeiter erarbeiteten die Anforderungen an die dargestellte Benutzerschnittstelle und die Retrieval-Leistungen in mehreren Workshops mit Anwendervertretern. Der folgende Abschnitt stellt besondere Aspekte des Prozessmanagements dar, die Voraussetzungen zur Umsetzung der gezeigten Retrieval-Leistungen sind.

4.3.5. Prozessmanagement

Die Bereitstellung von Dokumenten in der gezeigten Navigationshierarchie oder das gezielte Auffinden spezifischer Inhalte erfordert die Klassifizierung der Informationsobjekte. Zu Projektbeginn analysierte das McB-Team in mehreren

Workshops die Anforderungen von Betreibern und Anwendern des zukünftigen Portals. Dabei wurden Prozesse, Leistungen und Informationsbedürfnisse identifiziert und in einem Grobkonzept dargestellt. Das Projektteam spezifizierte in Workshops die Leistungsprozesse über Anwendungsfallanalysen (‚Use Case Analysis') aus und bildete sie in einer Projektmethode mit den hierarchischen Stufen Prozess, Phase, Aktivität und Aufgabe ab. Diese Attribute und die zugehörigen Werte bilden den Kern der Retrieval-Taxonomie (s. Tabelle 4-5).

Attribut	Wertebereich	Bemerkung
Competence	Actuary, Business Organisation, IT, Operations, Legal, Finance / Accounting, Testing, Data quality, Conversion	Gruppen von Spezialisten
Assigned to	-	Zuständiger Bearbeiter
Status	pending, in process, completed, not applicable	Bezeichnung unterschiedlicher Fertigstellungszustände von Projekten, Phasen, Aktivitäten oder Aufgaben
Created by	-	Ersteller eines Dokuments
Creation Date	-	Erstelldatum eines Dokuments
Country	-	Länderbezeichnung
Project	z.B. Migration-UK-Lifesystem	Projektbezeichnung
Process	Transparency, System Migration, Re-Engineering, Knowledge Management, Country relationship management, European Operational Council, McB strategy development, People management, Process and quality management, McB Project management, Performance measurement	Bezeichnung der Kernprozesse der McB-Initiative
Phase	Ausschnitt für Migration: 1. Feasability, 2. Preparation & Development, 3. Conversion, 4. Cleanup	Bezeichnung der Phasen der McB-Initiative
Activity	Ausschnitt für 4. Cleanup: 4.1 Monitor receiving environment, 4.2 Reorganizse workforces, and 4.3. Dispose of the sending system	Bezeichnung der Aktivitäten der McB-Initiative
Task	Ausschnitt für 4.3 Dispose of the sending system: 4.3.1 Switch off the system, 4.3.2 Dispose of hardware, 4.3.3 Dispose of software	Bezeichnung der Aufgaben der McB-Initiative

Tabelle 4-5: Ausschnitt aus der Retrieval-Taxonomie der Winterthur Versicherungen

Die Benutzerschnittstelle für ein Dokument zeigt Abbildung 4-14, in der ein Projektmitarbeiter z.B. den Status des Informationsobjektes durch Zugriff auf die Retrieval-Taxonomie ändert.

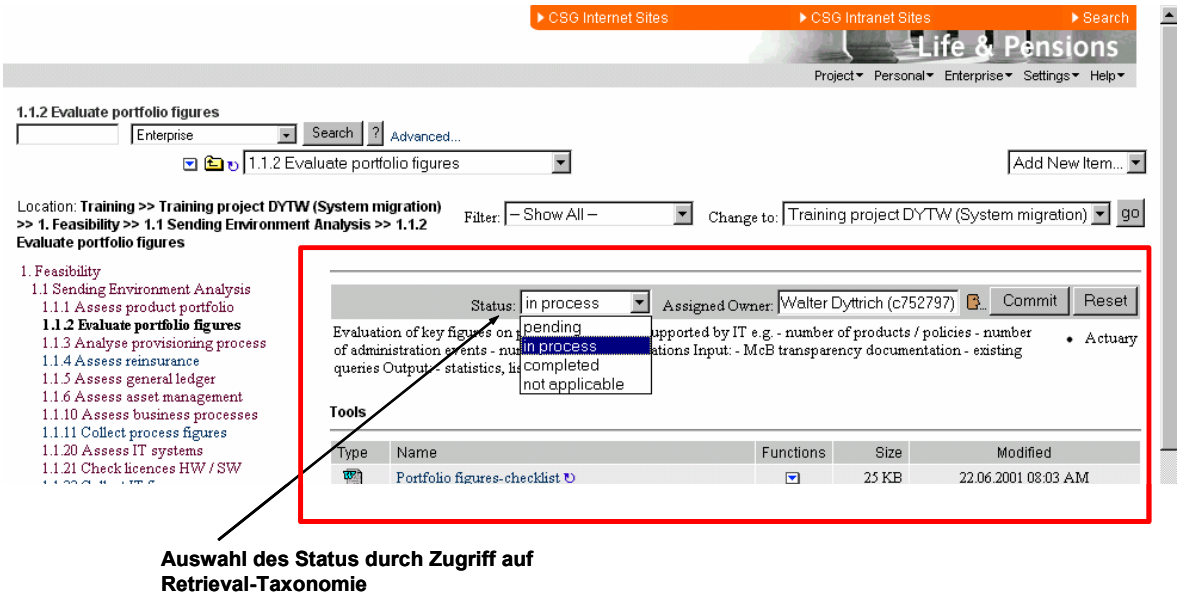


Abbildung 4-14: Benutzerschnittstelle für Informationsobjekt im McB-Portal (Entwurf)

Neuerstellungen, Änderungen und Löschungen von Attributen oder Werten der Retrieval-Taxonomie werden dabei wie folgt berücksichtigt:

- Wesentliche Werte der Retrieval-Taxonomie, z.B. für die hierarchischen Attribute ‚Phase', ‚Activity' und ‚Task' sind in so genannten ‚Blueprints' zusammengefasst. Sie erlauben, bewährte Strukturierungen einzelner Projekte wiederzuverwenden um neue Projekte mit einer festen Sollstruktur anzulegen. Auch für bereits angelegte Projekte kann eine Aktualisierung auf Basis eines ‚Blueprints' erfolgen, sofern in den einzelnen Phasen innerhalb von Projekten noch keine Anpassungen vorgenommen worden sind.

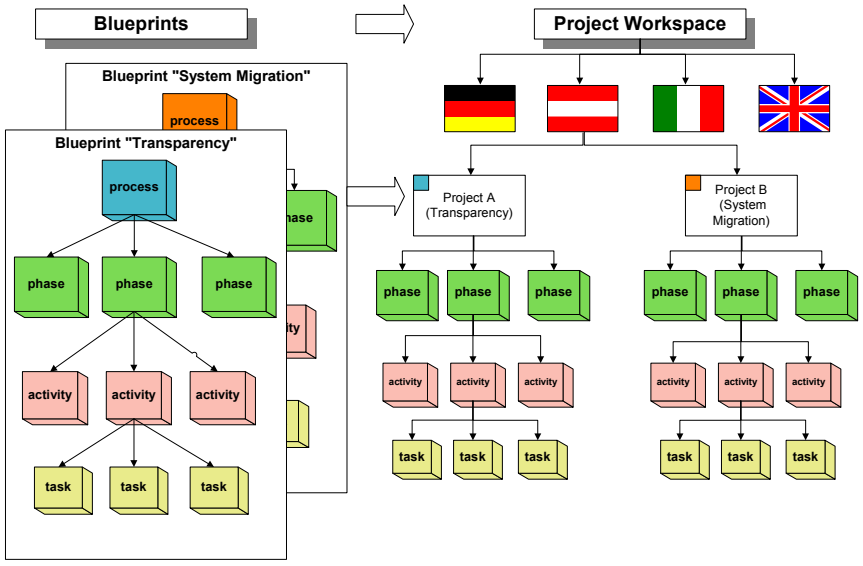


Abbildung 4-15: Pflege von Strukturinformationen im McB-Portal

- Berechtigungen zum Hinzufügen, Modifizieren oder Löschen einzelner Attribute und Werte der Retrieval-Taxonomie sind an Rollen gebunden. So kann beispielsweise ausschliesslich der Rolleninhaber ‚Project Manager' eine Projektbezeichnung ändern.
- Die Aktualisierung von Dokumenteneigenschaften muss im Bedarfsfall manuell erfolgen. Dabei greifen die Anwender zwar auf die veränderten Werte der Retrieval-Taxonomie zu, die Reklassifizierung wird jedoch nicht weiter vom System unterstützt.

Für die Indizierung der im McB-Portal verfügbaren Inhalte sind keine Prozesse definiert. Dieser Vorgang erfolgt automatisch und war bereits für die gesamte Plattform ‚LifeLink' aktiviert. Auch der Abgleich verschiedener Benutzergruppen war nicht erforderlich, da das McB-Portal auf dem bestehenden Intranetverzeichnis aufsetzt. Über Navigation und Suche stehen dem Kernteam der McB-Consultants alle Inhalte zur Verfügung. Die Mitglieder einzelner Projekte haben hingegen ausschliesslich Zugriff auf die Dokumente ihres jeweiligen Projekts.

4.3.6. Systemintegration

Da die Winterthur Life & Pensions mit Opentext Livelink (Release 9) bereits eine leistungsfähige Intranetplattform besass, entschied der Leiter der McB-Initiative von Projektbeginn an, die Anpassung und Erweiterung des vorhandenen Intranets ‚LifeLink' an die spezifischen Anforderungen von McB vorzunehmen. Um die Umsetzbarkeit des Konzepts frühzeitig zu prüfen, führten die Projektbeteiligten der Winterthur und des IWI-HSG zwei Arbeitssitzungen mit technischen Ansprechpartnern der Firma Opentext durch. Da die grafischen Darstellungsmöglichkeiten von ‚LifeLink' für die Gestaltung der gewünschten geographischen Landkarte zur Auswahl einzelner Länder nicht ausreichten, wurde die Einstiegsseite in das Portal separat mit dem Content-Management-System Obtree erstellt.

Mit der Umsetzung von Portal, Suchmaschine und Content-Applikation durch Opentext Livelink entfallen weitere Aufwände für die Integration von Authentifizierung, Autorisierung oder Benutzerverwaltung. Die System-verantwortlichen fokussierten daher auf die Datenmodellierung, welche die im Portal abzubildenden Informationsobjekte (z.B. Vorlagen, Beispieldokumente, Arbeitsunterlagen und Ergebnisse) hinsichtlich ihrer Eigenschaften und Methoden festlegt. Resultat war die Spezifikation eines objektorientierten Modells mit einer Superklasse ‚McB InfoObject' sowie verschiedener Subklassen, u.a. eines universellen Containers (‚McB_Document') für die Aufnahme von Dokumenten. Dieser enthielt alle prozessrelevanten Attribute, z.B. für Prozess, Phase, Aktivität und Aufgabe (Attribute McB_IO_Process, _Phase, _Activity und _Task) (s. Abbildung 4-16).

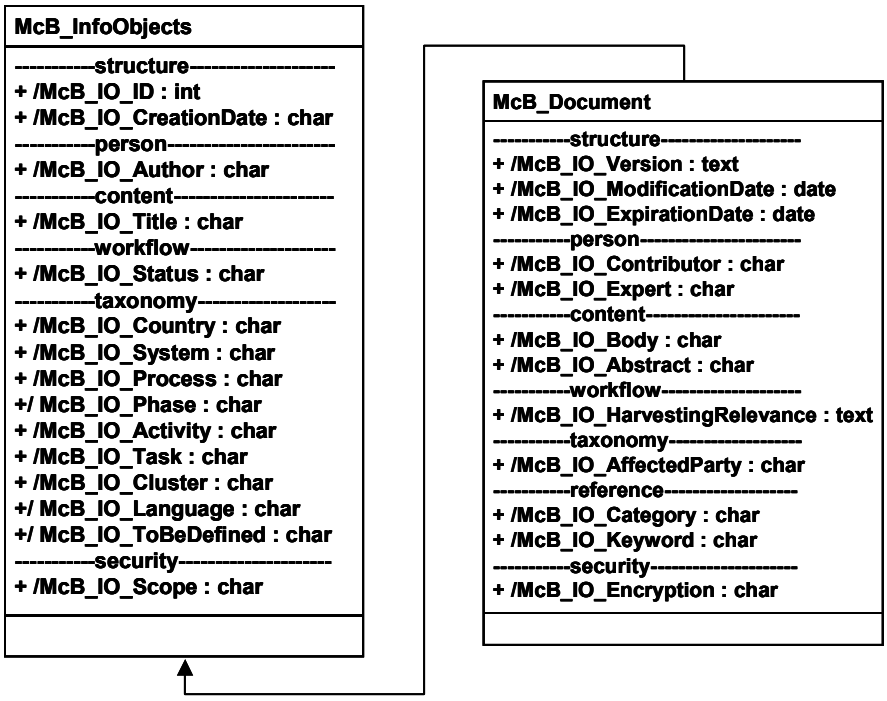


Abbildung 4-16: Ausschnitt aus dem Objekt- und Datenmodell des McB-Portals

Erst durch diese Mehrfachklassifizierung von Informationsobjekten, z.B. nach Prozess, nach Phase, nach Land und nach Autor sind im Portal unterschiedliche Sichten und Strukturierungen möglich. Darüber hinaus können diese Informationsobjekte auch über attributbasierte Suchanfragen gefunden werden.

Tabelle 4-6 zeigt das Mengengerüst der Retrieval-Lösung im Oktober 2002.

Content-Applikation	Anzahl Informationsobjekte	Änderungsfrequenz	Indexgrösse
McB-Portal	Ca. 5'000 Dokumente	- keine Angabe -	- keine Angabe -

Tabelle 4-6: Mengengerüst für indizierte Inhalte bei Winterthur Versicherungen

4.3.7. Kosten- und Nutzenaspekte

Im Oktober 2002, nach mehr als eineinhalb Jahren Betrieb, beinhaltet das McB-Portal mehr als 5'000 Dokumente und wird in 21 Projekten in 9 europäischen Ländern eingesetzt. Dabei dienen die aufeinander abgestimmten Gestaltungselemente für Struktur, Navigation und Suche im Portal ca. 25 Beratern sowie ca. 130 Mitarbeitern der Landesgesellschaften als zentrale Merkmale des Retrievals im Portal. Durch die konsequente Nutzung und kontinuierliche Weiterentwicklung wurde beispielsweise die SOLL-Struktur des Prozesses ‚Portfolio Migration' bereits achtmal erfolgreich wiederverwendet bzw. verbessert.

Über eine jährliche Kundenzufriedenheitsumfrage bei 25 bis 30 Kunden der McB-Initiative wird zur Performance-Messung ein Index ermittelt. In diesem sind Messgrössen für alle Dienstleistungen von McB, auch das Portal, enthalten. Die

Umfrageergebnisse werden ausgewertet, Massnahmen vorgeschlagen und mit den internen Kunden diskutiert.

Der Realisierungsaufwand für das McB-Portal liegt unter 100 Personentagen. Aufgrund der bereits vorhandenen Plattform ‚LifeLink' lag der eigentliche Entwicklungsaufwand für das McB-Portal bei nur 25 Personentagen.

4.3.8. Kritische Erfolgsfaktoren und geplante Weiterentwicklungen

- Für die Anwender war die *Verwendung einer vorgegebenen Projektmethode* und dazugehöriger Vorlagen für die Projektdurchführung wichtig. Die an den Arbeitsabläufen ausgerichteten Retrieval-Funktionen wie prozessorientierte Navigation und Suche stellen dies sicher.
- Durch die Bereitstellung aller relevanten Inhalte in einer zentral verfügbaren Applikation konnten die Verantwortlichen einen *weltweiten Zugriff auf Dokumente* ermöglichen.
- Aus Sicht der Betreiber stand eine *hohe Benutzerakzeptanz und Anwendung der Plattform* im Vordergrund. Die Nutzungskennzahlen zeigen, dass die Anwender das McB-Portal intensiv in ihren Arbeitsalltag einbeziehen.
- Eine *zeitnahe Realisierung mit geringem Implementierungsaufwand* sollte die Kosten der Lösung begrenzen. Durch die Verwendung und Anpassung des vorhandenen Systems Opentext Livelink wurde dies gewährleistet.

4.4. Institut für Wirtschaftsinformatik – Knowledge Port

„Das KnowledgePort-Projekt hat verschiedene heterogene Informationsbestände auf einer neuen, leistungsfähigeren Plattform zusammengeführt. Auf dieser Basis hat das deutlich verbesserte Information Retrieval die Adressaten-spezifische Erschliessung der Inhalte entscheidend vorangebracht."

Dr. Gerold Riempp, Projektleiter CC CKM und Knowledge Port (Phase I)

4.4.1. Organisation und Ausgangssituation

Das 1989 gegründete Institut für Wirtschaftsinformatik der Universität St. Gallen (IWI-HSG) forscht an Themen des Informationsmanagements und verwandter Gebiete. Es besteht aus 4 Professoren mit ihren Lehrstühlen, die insgesamt 12 Projektleiter, 38 wissenschaftliche Mitarbeiter und rund 50 studentische Mitarbeiter umfassen. Die Forschung erfolgt im gemeinsam getragenen Forschungsprogramm ‚Business Engineering HSG' in enger und langjähriger Kooperation mit europäischen Unternehmen (so genannte Forschungspartner). Ein Kompetenzzentrum (Competence Center, CC) fasst für die Dauer von zwei Jahren mehrere Forschungspartner zur gemeinsamen Erforschung eines bestimmten Themenschwerpunktes zusammen. Ein Projektleiter und ein Team von wissenschaftlichen Mitarbeitern erarbeiten in einem Kompetenzzentrum mit den Praxisvertretern Konzepte, Methoden und Lösungen im Rahmen regelmässiger Workshops sowie in gemeinsamen Projekten. Die Arbeitsergebnisse in den Kompetenzzentren sind, ebenso wie die Publikationen und Inhalte der Lehrveranstaltungen, immateriell. Die Mitarbeiter erstellen sie mit Textverarbeitungs-, Tabellenkalkulations- und Präsentationsgrafikprogrammen aber auch mit der universitätsweit verfügbaren Groupware-Plattform Lotus Notes.

Im Spätherbst 2000 erhielt der Projektleiter des Kompetenzzentrums ‚Customer Knowledge Management' (CKM) die Aufgabe, das Wissensmanagement (WM) des Instituts zu reorganisieren.

- Die bis dahin gewachsenen organisatorischen und technischen Strukturen wurden immer weniger der Grösse des Instituts gerecht. Der Umgang mit Informationen wurde dezentral in den unterschiedlichen Projektteams geregelt.
- Der systematische Wissensaustausch zwischen den einzelnen Kompetenzzentren des Instituts war gering. Synergien bei der Bearbeitung gleicher Forschungsthemen wurden unzureichend genutzt.
- Die Anforderungen der Forschungspartner hinsichtlich Qualität, kurzer Antwortzeiten und detaillierten Arbeitsergebnissen wuchsen. Andererseits waren

die Möglichkeiten aktueller Informationssysteme zur Unterstützung der Abläufe nicht ausreichend genutzt.

Mit dem daraufhin gebildeten Team[7] führte der Projektleiter zunächst eine Erhebung der IST-Situation sowie der Anforderungen durch. Bei strukturierten Befragungen von Mitarbeitern, in Workshops sowie Interviews mit Führungskräften wurden die erleichterte Orientierung und der verbesserte Zugang zu Informationsobjekten wiederholt als wichtigste Anforderungen genannt.

Ein zentrales Ergebnis der nach der Erhebung durchgeführten Konzeptionsphase war im Frühjahr 2001 die in Abbildung 4-17 dargestellte Architektur für das so genannte Portal ‚Knowledge Port'. ‚Knowledge Port' sollte die bisherigen heterogenen Applikationen in ein neues, auf die Bedürfnisse der Mitarbeiter abgestimmtes Informationssystem konsolidieren und u.a. umfassende Navigations- und Suchfunktionen bereitstellen.

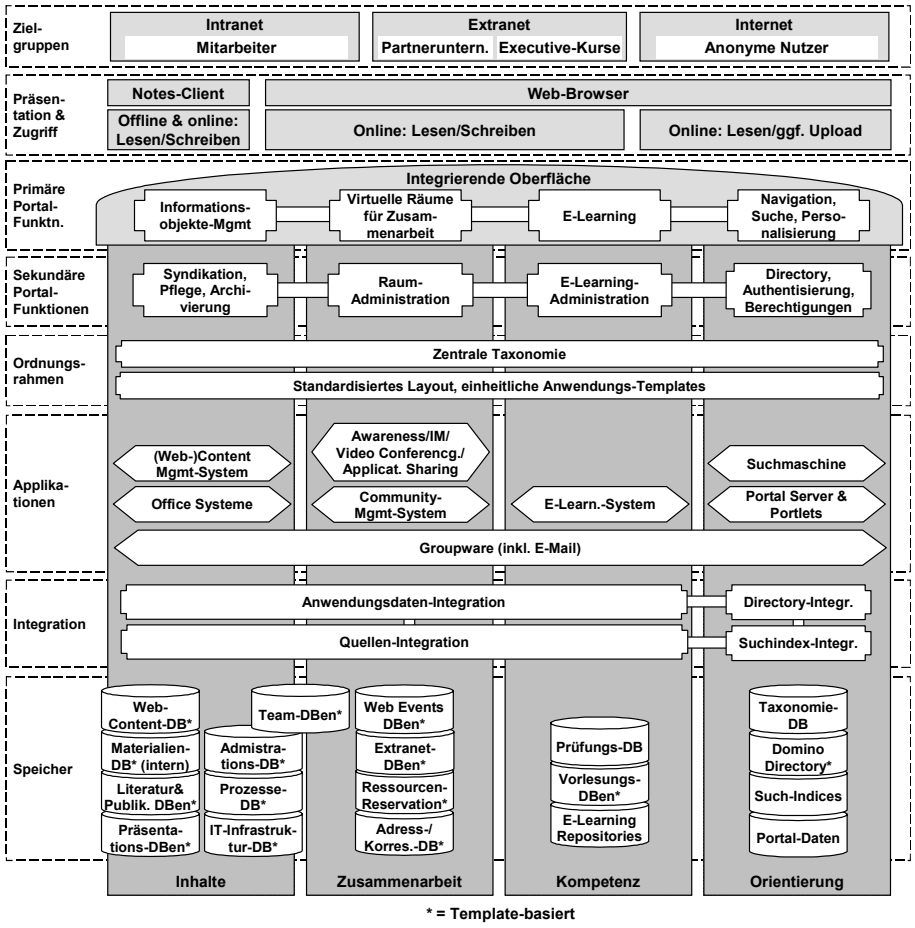


Abbildung 4-17: Architektur des Wissensmanagement-Systems bei IWI-HSG [s. Riempp 2003, 53]

[7] Der Autor der vorliegenden Arbeit gestaltete zunächst von Juni bis Dezember 2002 als Projektmitglied, später als verantwortlicher Projektleiter, die Implementierung.

4.4.2. Treiber und Herausforderungen für Information Retrieval

Folgende Aspekte stellten Treiber für die Umsetzung von ‚Knowledge Port' dar:

- Die Vielzahl der vorhandenen Systeme wie Fileserver, Datenbanken, Referenzverwaltungen etc. erlaubte es bisher keiner Anwendergruppe, eine übergreifende, z.B. thematische Sicht über das vorhandene Informationsangebot zu gewinnen. Bei der vorbereitenden Recherche eines Workshops zum Thema ‚Suchmaschinen' stellte sich beispielsweise heraus, dass die Informationsfindung im Internet mit der Suchmaschine Google einfacher war als die Erschliessung der institutseigenen Datenbestände, obwohl bereits Präsentationen und Artikel zu diesem Thema vorhanden waren.

- Die teilweise in Content-Applikationen vorhandenen Suchmechanismen waren nicht aufeinander abgestimmt. Die Anwender mussten z.B. Suchsyntax oder Suchoperatoren jeder einzelnen Anwendung kennen, um ihre Suchanfrage zu formulieren. Nach einigen Versuchen brachen die Anwender die Suche oft vollständig ab.

- Die Verschlagwortung von Dokumenten nahm jeder Mitarbeiter nach seinem eigenen Verständnis vor. Dies führte dazu, dass Inhalte entweder gar nicht oder mit individuell gewählten Begriffen klassifiziert wurden. Im ‚worst case' blieben die Schlüsselwörter bei der Verwendung einer alten Präsentation als Vorlage für ein neues Arbeitsergebnis bestehen. Angaben zum Autor oder zum Thema waren dadurch falsch und für eine gezielte Suche nutzlos.

- Verschiedene Zielgruppen (Intranet, Extranet, Internet) wurden mit einer Vielzahl unterschiedlicher Systeme konfrontiert. Eine integrierte Benutzeroberfläche mit standardisiertem Layout für den Zugriff auf die sie betreffenden Informationen stand ihnen nicht zur Verfügung.

Mitte 2001 setzte das Projektteam die neue Architektur schrittweise um. Über Navigation und Suche in ‚Knowledge Port' sollten allen Zielgruppen die von ihnen gesuchten Informationsobjekte komfortabel, schnell und entsprechend ihrer Zugriffsrechte zur Verfügung gestellt werden. Besondere Herausforderungen bestanden dabei insbesondere darin:

- eine logisch integrierte Sicht auf die physisch verteilten Content-Applikationen zu ermöglichen,

- die für Klassifikation, Volltextsuche sowie attributbasierte Suche und Navigation institutsweit gültige Taxonomie zu erarbeiten,

- weitestgehend vorhandene oder bereits über die Universität verfügbare Applikationen bzw. lizensierte Softwarepakete zu verwenden und
- zur Vermeidung von redundanten Benutzerdaten und für die zentrale Authentifizierung aller Anwender das vorhandene Domino Directory als führenden Verzeichnisdienst zu nutzen.

4.4.3. Unterstützungspotenzial

Das Unterstützungspotenzial durch Information Retrieval sahen die Verantwortlichen bei IWI-HSG überwiegend in einer möglichen Qualitätssteigerung. Die Retrieval-Leistungen ermöglichen eine bessere Nutzung von inhaltlichen Synergien, z.B. zwischen den einzelnen Kompetenzzentren. Vorhandene Inhalte zu einem bestimmten Thema sind für alle Anwendergruppen leicht auffindbar und erleichtern damit den Informationsaustausch innerhalb des Instituts, mit den Forschungspartnern sowie der Öffentlichkeit. Über Navigation und Suche werden allen Beteiligten verbesserte bzw. erweiterte elektronische Dienstleistungen angeboten. Ein weiterer Nutzenaspekt besteht darin, durch ein eigenes Portal- und Retrieval-Projekt die Erfahrungen der Mitarbeiter in diesem Bereich aufzubauen.

Die folgenden Abschnitte skizzieren wesentliche Elemente der Umsetzung von ‚Knowledge Port'. Ausgehend von den bereitgestellten Benutzerschnittstellen und verfügbaren Retrieval-Leistung für die Zielgruppen ‚Intranet', ‚Extranet' und ‚Internet' werden die dafür notwendigen Prozesse von Terminologie-, Integrations- und Indexmanagement dargestellt. Die Darstellung ausgewählter technischer Aspekte der Systemintegration schliesst die Fallstudie ab.

4.4.4. Benutzerschnittstelle und Retrieval-Leistung

Um allen drei Zielgruppen (Intranet, Extranet, Internet) die notwendigen Inhalte und Services über eine einheitliche Portalplattform zur Verfügung zu stellen, führte das Projektteam im Herbst 2002 einen Machbarkeitsnachweis (‚Proof of Concept') mit zwei führenden Software-Herstellern durch. Beide Anbieter implementierten ihre Portal- und Retrieval-Produkte in einer Testumgebung des Instituts. Während ein Hersteller trotz erfolgreichem Funktionsnachweis den Produktiveinsatz durch hohe Lizenzkosten unmöglich machte, erfüllten im anderen Fall die Leistungsmerkmale der Software nicht die gestellten Anforderungen. Als Zwischenlösung implementierte das Projektteam daher für die Zielgruppen jeweils unterschiedliche Content-Applikationen, Portale und Suchmaschinen in einem einheitlichen Layout. Intranet-, Extranet- und Internet-Anwender greifen jedoch bereits heute auf eine gemeinsame Taxonomie und ein zentrales Benutzerverzeichnis zu, um eine spätere Integration in einer gemeinsamen Portalplattform zu erleichtern.

- *Internet*. Die Öffentlichkeit greift auf redaktionell erstellte Inhalte sowie Publikationen der Institutsmitarbeiter über das so genannte IWI-Web zu (http://www.iwi.unisg.ch). Die Benutzerschnittstelle bietet dabei eine kombinierte Volltextsuche und attributbasierte Suche (s. Abbildung 4-18).

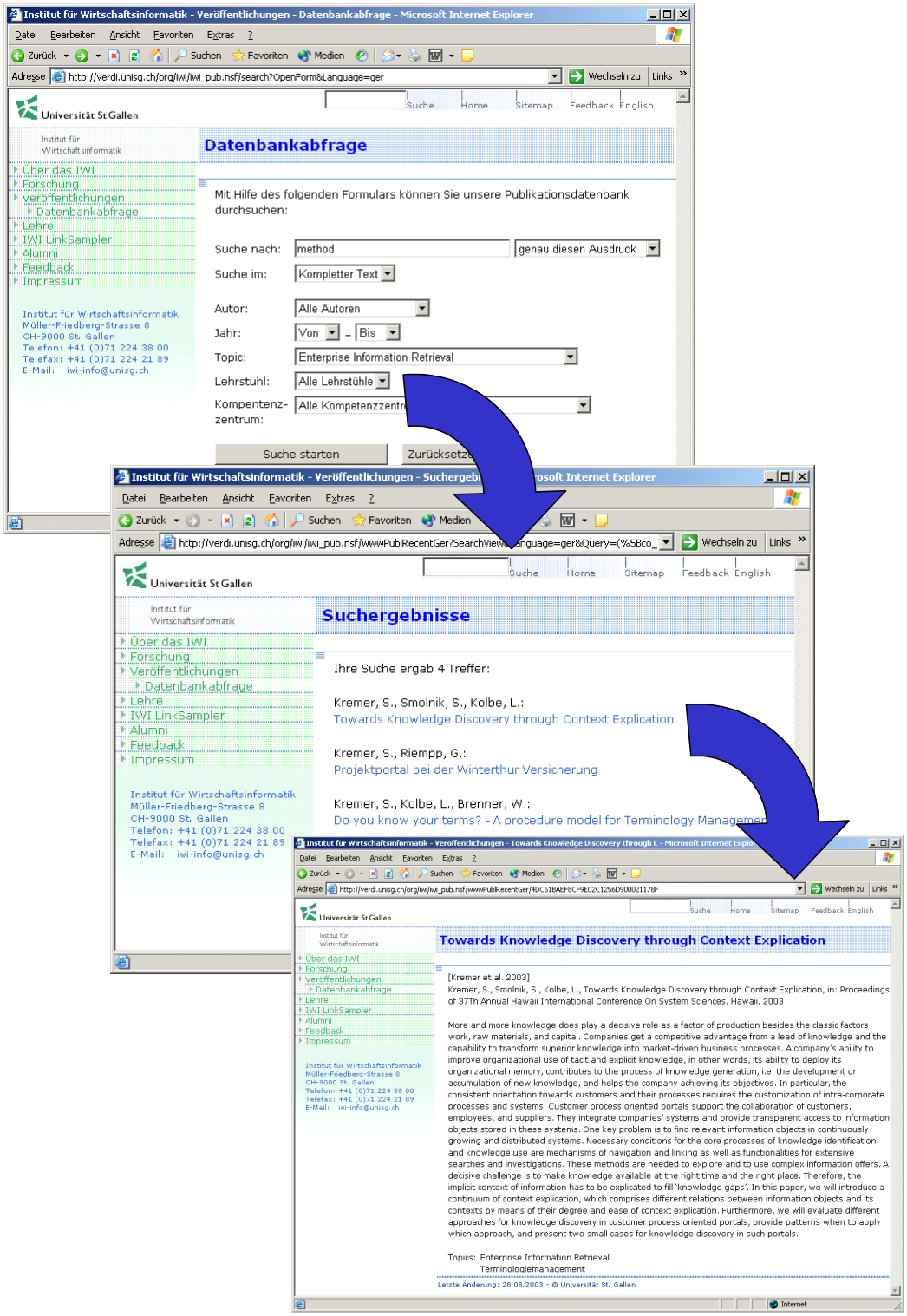


Abbildung 4-18: Retrieval im Internet-Auftritt von IWI-HSG

- *Extranet*. Den Forschungspartnern steht in ihrer Extranet-Plattform eine Volltextsuche zur Verfügung. Diese erfolgt sowohl in den Dokumenten als auch in den enthaltenen Präsentationen oder mit Textverarbeitungsprogrammen erstellten Inhalten.

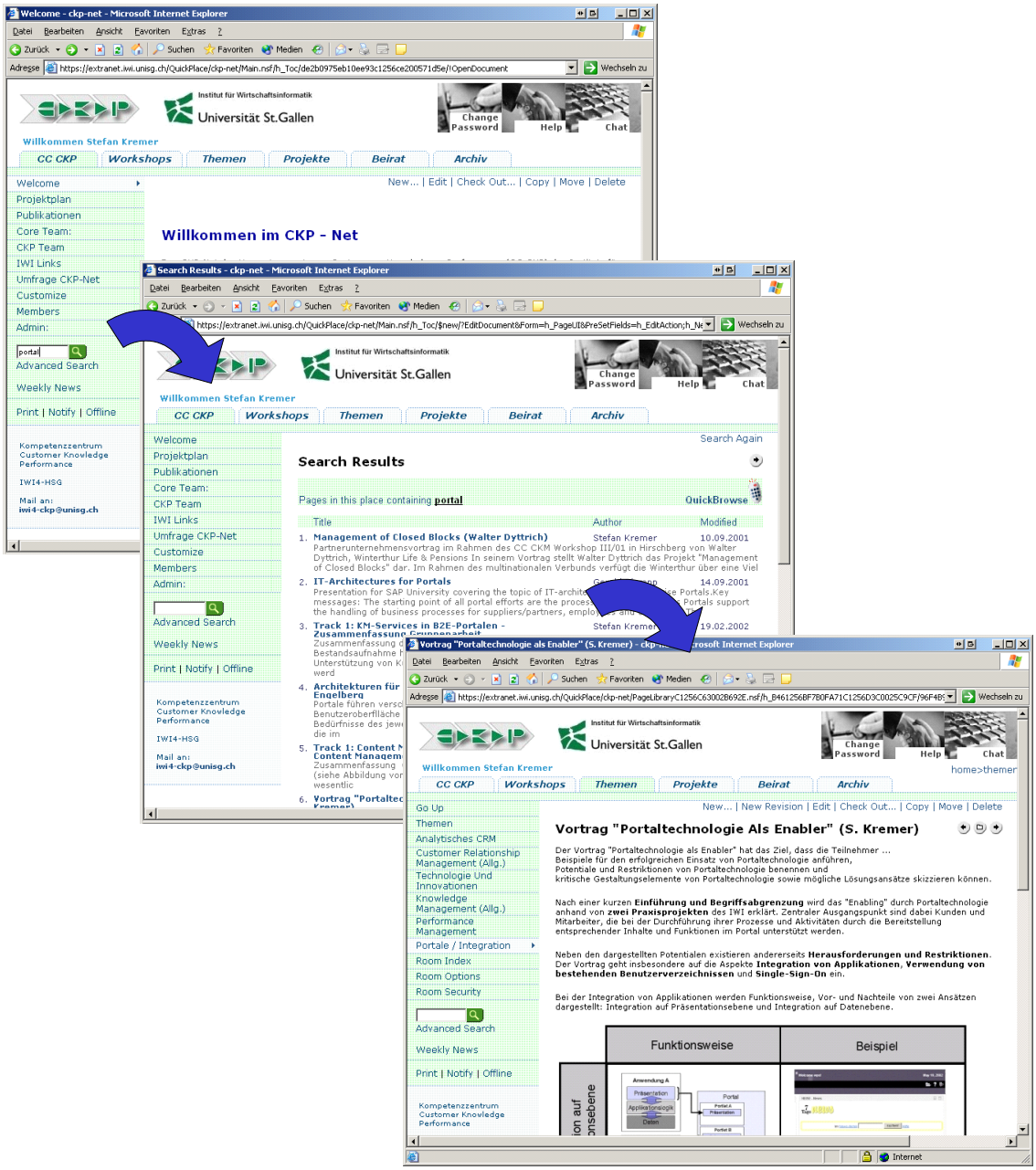


Abbildung 4-19: Retrieval im Extranet-Portal von IWI-HSG

Die Navigation im Extranet des Kompetenzzentrums ‚Customer > Knowledge > Performance' (C>K>P) bietet beispielsweise zwei unterschiedliche Sichten auf die enthaltenen Dokumente: die Navigation über ‚Workshops' führt chronologisch die Veranstaltungen des CCs auf. Die Einträge in den jeweiligen Workshop-Agenden verlinken dabei auf die Inhalte im Navigationsbereich ‚Themen'.

- *Intranet*. Die Institutsmitarbeiter arbeiten mit dem ‚KPort LinkSampler'. Dieser stellt Volltextsuche und attributbasierte Suche über alle internen Content-Applikationen wie Team-Datenbanken, Literatur- und Publikationsdatenbanken sowie Präsentationsdatenbanken zur Verfügung.

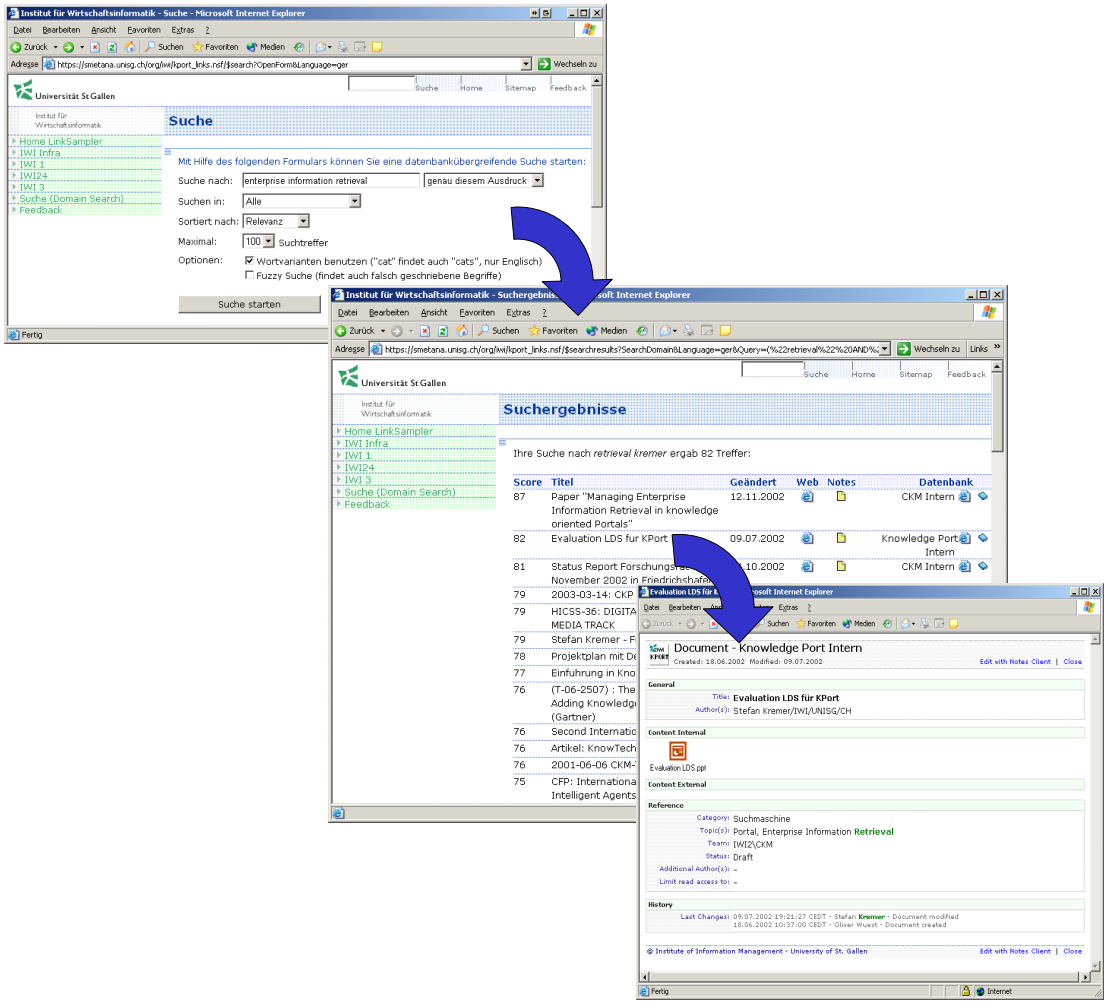


Abbildung 4-20: Retrieval im Internet-Auftritt von IWI-HSG

Eine Besonderheit stellen die Verknüpfungen in der Suchergebnisliste dar. Die Einträge können wahlweise im Web-Browser oder, falls installiert, mit dem Notes-Client geöffnet werden.

4.4.5. Prozessmanagement

Um eine attributbasierte Suche in den Anwendungen von ‚Knowledge Port' zur ermöglichen, ist die in Tabelle 4-7 dargestellte Retrieval-Taxonomie erforderlich. Nur durch diese ist die gezielte Erschliessung von Dokumenten auf Basis ihrer Eigenschaften wie Autor, Themenschwerpunkt oder Organisationseinheit möglich. Zentrales Merkmal stellt dabei das Attribut ‚Topic' dar, das den thematischen Schwerpunkt eines Dokuments kennzeichnet. Es adressiert dabei das Bedürfnis der

Anwender, bestehende Informationen zu ihren überwiegend themenbezogenen Tätigkeiten zu finden.

Attribut	Wertebereich	Bemerkung
Author	- offen -	alle Mitarbeiter des Instituts
Year	- offen -	Jahreszahlen
Chair	IWI 1, IWI 2, IWI 3, IWI 4	Lehrstuhl
Competence Center	• IWI 1: BPM, DWS, BAI, EMBE, DW2, AIM, RMA • IWI 2: BN, I-NET, TCC, PRO, ECC, PRO-BM, BKM, eBN, PSI, CRM, CRIM, iBN, BN2, M-Lab, TI, CKM • IWI 3: E-Learning, Knowledge Networks, Learning Center, Knowledge Source • IWI 4: CKM, C>K>P, IIM	Bezeichnung aller 32 abgeschlossenen oder laufenden Kompetenzzentren; aufgeführt nach Lehrstühlen
Topic	Ausschnitt für IWI 4 / C>K>P: Administration, Enterprise Application Integration, Enterprise Information Retrieval, IT-Marketing, M-Commerce, M-Payment, Methode, Portal, Privacy, Real-time Management, Sicherheit, System, Terminologiemanagement, Web Services, Wirtschaftsinformatik (allgemein)	Thematische Schwerpunkte aller 32 abgeschlossenen oder laufenden Kompetenzzentren; insgesamt 450 Begriffe

Tabelle 4-7: Ausschnitt aus der Retrieval-Taxonomie bei IWI-HSG

Die ursprünglich stark heterogene Systemumgebung aus Fileservern, Datenbanken, Literaturverwaltungssystemen und individuellen Ablagen reduzierten die Projektverantwortlichen durch die Entwicklung einer Standard-Datenbankschablone (Template) auf Basis von Lotus Notes. Für jedes Kompetenzzentrum wurde aus diesem Template eine so genannte Team-Community generiert, die als zentrale Umgebung für die interne Arbeit in den CCs dient.

Aus diesem Standard wurden weitere Content-Applikationen abgeleitet. So verfügen Literatur- und Publikationsdatenbank oder Präsentationsverwaltungsdatenbank über dieselbe Kernfunktionalität. Durch diese Standardisierung der verwendeten Applikationen ist die Anwendungsschnittstelle für die Suchmaschinenintegration immer identisch. Felder mit gleicher Bedeutung, z.B. Topic sind in allen Anwendungen gleich bezeichnet.

Alle Applikationen innerhalb von ‚Knowledge Port' greifen zum Zweck der inhaltlichen Klassifikation auf eine zentrale Taxonomie-Datenbank zu, in der die Topics gepflegt sind. Abbildung 4-21 zeigt die Schaltfläche sowie den Dialog zur Auswahl des Topics im Bereich ‚Reference', der in jeder Content-Applikation von ‚Knowledge Port' enthalten ist.

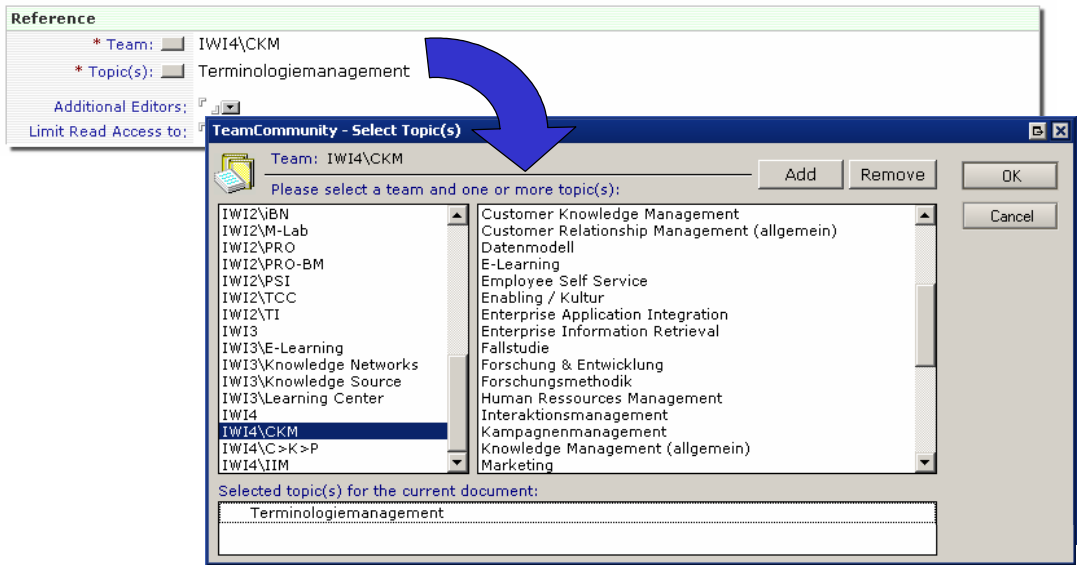


Abbildung 4-21: Klassifikation eines Informationsobjekts bei IWI-HSG

Topics werden dezentral von verantwortlichen Mitarbeitern der einzelnen Kompetenzzentren gepflegt. Diese so genannten Taxonomy-Editors greifen schreibend auf die zentrale Taxonomie-Datenbank zu und Erstellen, Modifizieren oder Deaktivieren Topics (s. Abbildung 4-22).

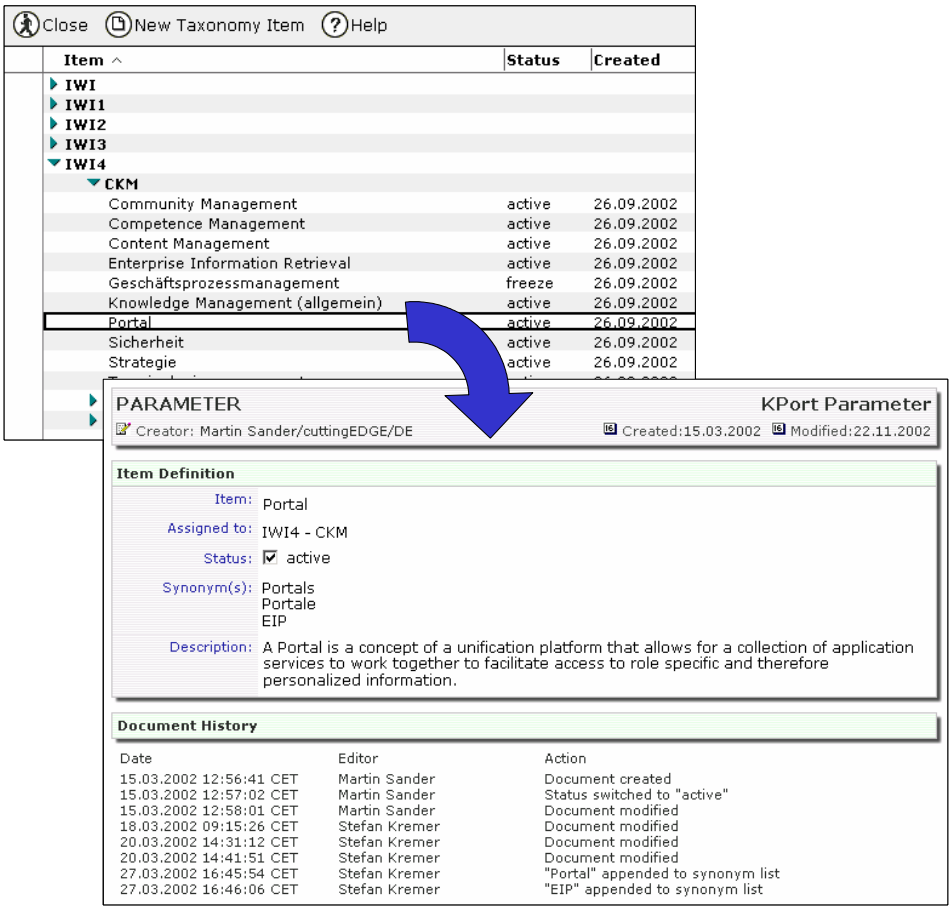


Abbildung 4-22: Taxonomie-Datenbank von IWI-HSG

Jeder Taxonomieeintrag besteht dabei aus einem Hauptwort (‚Item'), Synonymen, der organisatorischen Zuordnung bezüglich der Pflege sowie einem Status. Letzterer kann die Werte ‚draft' für Entwürfe, ‚active' für nutzbare Einträge sowie ‚freeze' für früher genutzte Einträge annehmen. Die Anwender in den Content-Applikationen können im Bereich ‚Reference' (s. Abbildung 4-21) ausschliesslich Einträge mit Status ‚active' verwenden.

Durch die weitgehende Nutzung der bereits vorhandenen Groupware-Plattform Lotus Notes/Domino konnte das bestehende Benutzerverzeichnis für die zentrale Authentisierung aller Anwender genutzt werden. Extranets und der KPort LinkSampler greifen auch für die Pflege der Autorisierung auf das Domino Directory zu, um existierende Benutzer oder Gruppen in den Zugriffskontrolllisten der Applikationen zu verwalten.

4.4.6. Systemintegration

Die beschriebenen Retrieval-Leistungen und das Prozessmanagement sind bei IWI-HSG mit folgenden Komponenten auf Systemebene umgesetzt:

- Die *Suchmaschine* IBM Global Text Retriever (GTR) ist ein integraler Bestandteil aller drei Portale (IWI-Web, KPort LinkSampler, Extranets). Sie stellt als Komponente des Lotus/Domino-Servers (ab Release 5) über die so genannte Domain Search eine datenbankübergreifende Volltextsuche (VT) zur Verfügung. Ihre Suchanfrage- und Ergebnismasken können durch HTML-Programmierung angepasst werden. Darüber hinaus berücksichtigt die Domain Search bei der Anzeige von Suchergebnissen die benutzerbezogenen Zugriffsrechte. Für die attributbasierte Suche (AB) in den im IWI-Web enthaltenen Publikationen führt der GTR zusätzlich eine feldbasierte Indizierung innerhalb der Datenbank (IWI-Publications) durch.
- Das *IWI-Web* ist ein auf Basis des Internet-Auftritts der Universität St. Gallen weiterentwickeltes System (Lotus Notes/Domino R5). Die enthaltenen redaktionell erstellten Inhalte (Web Content), z.B. Veranstaltungshinweise, Lehrveranstaltungen oder Prüfungstermine werden gemeinsam mit den Publikationen des Instituts (IWI-Publications) von der Suchmaschine über eine proprietäre Lotus-Notes-Schnittstelle indiziert und als Suchergebnisse im HTML-Format angezeigt.

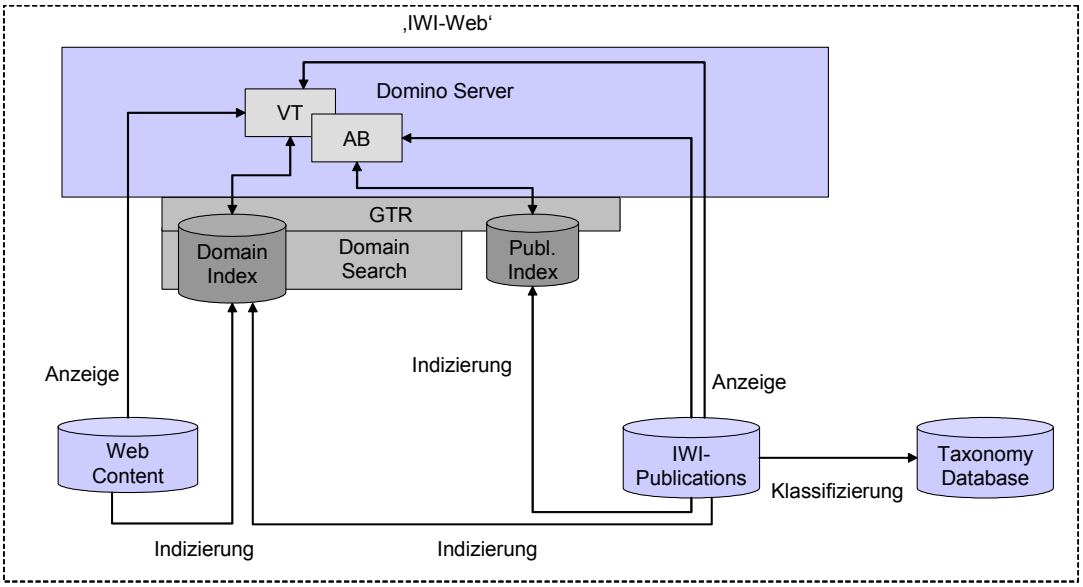


Abbildung 4-23: Retrieval-Architektur der Portalkomponente ‚IWI-Web'

- Das *Team-Community*-Template ist eine Eigenentwicklung auf Basis von Lotus Notes/Domino R5. Auf dem Template basierende Datenbanken verfügen als führende Content-Applikation des Instituts über eine komfortable Web-Schnittstelle, so dass die Anwender mit einem Web-Browser lesend auf die Inhalte zugreifen können. Die Suchmaschine indiziert die Inhalte, wie beim IWI-Web, über eine Notes-spezifische Schnittstelle unter Berücksichtigung von Zugriffsbeschränkungen.

- *Literatur- und Publikationsdatenbanken*, *Präsentationsdatenbank* und *IWI-Publications* stellen Weiterentwicklungen auf Basis des Team-Community-Templates dar. Ihr Verhalten bei Indizierung und Suche ist analog zu dem der Team-Datenbanken.

- Der *KPort LinkSampler* basiert auf dem Template des IWI-Webs. Auch hier stellt die Domain Search auf Basis des IBM Global Text Retriever die datenbankübergreifende Volltextsuche (VT) zur Verfügung. Dabei indiziert der GTR alle Team-Communities sowie die Literatur- und Publikationsdatenbanken der Lehrstühle.

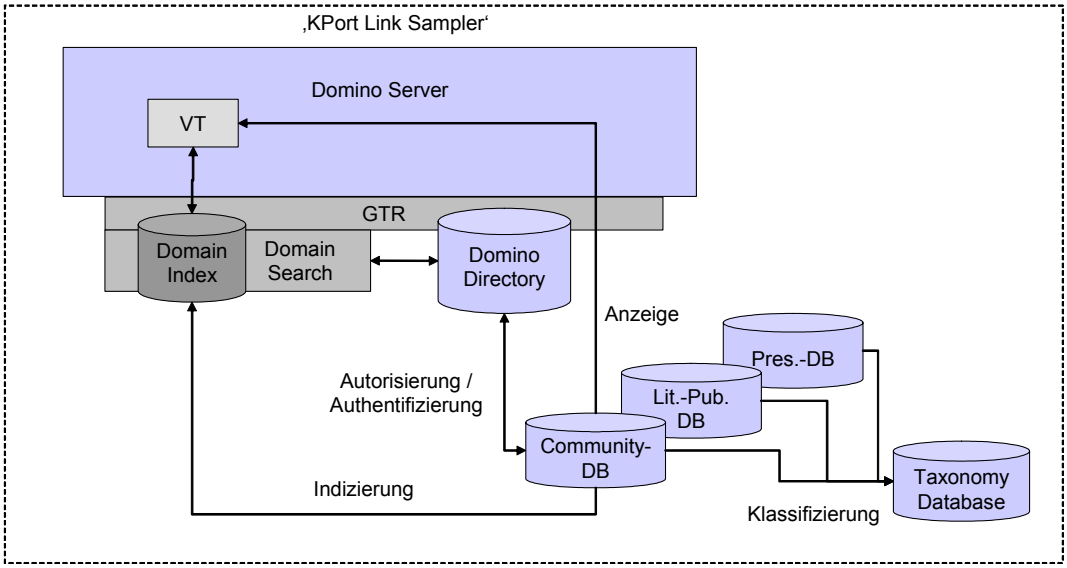


Abbildung 4-24: Retrieval-Architektur für die Portalkomponente ‚KPort Intranet'

- Die *Taxonomiedatenbank* enthält das kontrollierte Vokabular des Instituts. Zur Klassifizierung von Dokumenten greifen die Mitarbeiter aus Team-Communities, Literatur- und Publikationsdatenbanken und den Präsentationsdatenbanken auf diese zu. Im IWI-Web stehen Teile der Taxonomie, insbesondere für den thematischen Schwerpunkt (‚Topic') und die organisatorische Einordnung (‚Chair' bzw. ‚Competence Center') eines Dokuments dem Anwender zur attributbasierten Suche zur Verfügung.

- Das *Domino Directory* dient als zentrales Benutzerverzeichnis der Authentifizierung. Alle Content-Applikationen greifen auch für die applikationsbezogene Pflege der Zugriffskontrolle von Benutzern und Gruppen auf das Domino Directory zu. Der Global Text Retriever prüft bei Indizierung und Anzeige von Suchergebnissen auf Basis des Domino Directory, ob ein zugriffsgeschütztes Dokument in der Suchergebnisliste angezeigt wird.

- Die *Extranets* sind auf Basis von Lotus Team Workspace 3.0 (ehemals Quickplace) implementiert. Die Indizierung und Suche in den Inhalten des auf Lotus/Domino-Applikationen basierenden Team Workspace erfolgt für Benutzer und Administratoren vollständig transparent. Die in den Extranets enthaltenen Inhalte werden z.B. automatisch vom Global Text Retriever indiziert und stehen für eine Volltextsuche zur Verfügung. Die Retrieval-Architektur ist vergleichbar mit der in Abbildung 4-24.

Die vergleichsweise einfache Integration der Komponenten basiert zum einen auf der herstellerzentrierten Produktauswahl, getrieben u.a. durch bereits verfügbare Lizenzen oder durch das vorhandene Domino Directory. Ein weiterer entscheidender Faktor ist die Implementierung des Kern-Datenmodells in allen Content-Applikationen. So heisst

z.B. das Textfeld für den Titel eines Dokuments in den Team-Communities, Literatur- und Publikationsdatenbanken und der Präsentationsverwaltung immer ‚Subject'. Abbildung 4-25 zeigt einen Ausschnitt, der auch das Attribut ‚Topic' der Retrieval-Taxonomie abbildet.

Attribut-Name	Datentyp	Nutzer-Inter-aktion	Multi-plizität	Zweck
Sektion „Content"				
Subject	Text	P	E	Titel des Informationsobjektes (InOb)
Body	Richtext	O	E	Interner Inhalt des InOb
BodyExternal	Richtext	O	E	Inhalt des InOb für Externe
Sektion „Reference"				
Topic	Keyword	P	M	Inhaltliche Verschlagwortung
Categories	Keyword	P	M	Ordnungsstruktur
Team	Text	V	E	Team-Kontext der Autoren
Sektion „Workflow"				
DateCreation	Date	A	E	Datum der Erstellung
DateLastModification	Date	A	E	Datum der letzten Änderung
DatePublication	Date	V	E	Datum der Veröffentlichung
DateExpiry	Date	V	E	Verfallsdatum / Wiedervorlage
Destination	Keyword	V	M	Zielgruppe(n)
Creator	Username	A	E	Urheber
Authors	Username	V	M	Weitere Autoren
Status	Keyword	P	E	Content-Lebenszyklus (s.o.)
History	Text	A	M	Bearbeitungshistorie
Sektion „Security"				
Reader	Username	O	M	Leseberechtigung
AdditionalEditors	Username	O	M	Weitere Bearbeitungsrechte
Sektion „Administration"				
Build	Text	A	E	Version des Datenmodells
UID	Text	A	E	Eindeutige Kennung
Version	Text	O	E	Version des InOb
Legende: P = Pflichteingabe, O = optionale Eingabe, V = Vorschlag von System, A = automatische Belegung durch System, E = Einzelwert, M = multiple Werte				

Abbildung 4-25: Kern-Datenmodell der Content-Applikationen bei IWI-HSG

Bei neu hinzukommenden Content-Applikationen, z.B. einer Team-Community für ein weiteres Kompetenzzentrum, ist lediglich eine Konfiguration durch den Datenbankadministrator notwendig, um einen Suchindex anzulegen. In den so genannten Datenbankeigenschaften (s. Abbildung 4-26, linke Grafik) muss eine Schaltfläche aktiviert werden, um die Datenbank in den Domain-Search-Index aufzunehmen. Der Domino-Server, auf dem die Domain Search läuft, untersucht alle Datenbanken periodisch auf dieses Merkmal. Ist die entsprechende Eigenschaft aktiviert, legt der Server selbständig ein entsprechendes Indexdokument an (s. Abbildung 4-26, rechte Grafik). Der Index wird in Folge automatisch aktualisiert, sobald die Institutsmitarbeiter neue Dokumente einstellen oder bestehende Inhalte löschen.

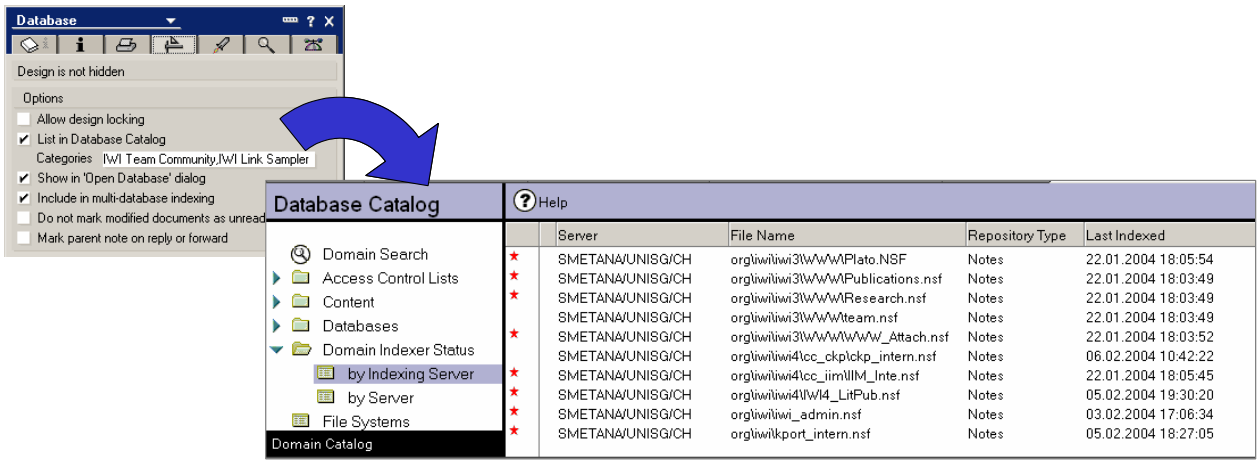


Abbildung 4-26: Aktivierung der Indexierung für Content-Applikationen bei IWI-HSG

Nach der erfolgreichen Einführung und dem Betrieb weist die skizzierte Lösung im Februar 2004 das in Tabelle 4-8 abgebildete Mengengerüst auf:

Content-Applikation	Anzahl Informationsobjekte	Volumen in MB	Änderungsfrequenz	Indexgrösse
CKP-Intern	630 Dokumente	1’945	2'640 Änderungen in 2003	120 MB
BN2-Intern	1’025 Dokumente	2’020	3’960 Änderungen in 2003	80 MB
IIM-Intern	1 Dokument	2	Keine Änderung in 2002 / 2003	Kein Index
AIM-Intern	150 Dokumente	42	440 Änderungen in 2003	Kein Index
CKM-Intern	850 Dokumente	1’210	660 Änderungen in 2002 / 2003	93 MB
IWI-Admin	260 Dokumente	420	Ca. 550 Änderungen in 2003	70 MB
IWI Literature & Publications	10'250 Dokumente	2'740	Ca. 16’000 in 2003	200 MB
IWI-Web	460 Dokumente	35	Ca. 1’760 Änderungen in 2003	Kein Index
IWI-Publications	1’320 Dokumente	410	Ca. 10’780 Änderungen in 2003	140 MB
KPort LinkSampler	70 Dokumente	7	Ca. 1’320 Änderungen in 2003	Kein Index
Extranet CKP-Net	4’290 Dokumente	2’160	Ca. 440 Änderungen in 2003	1’220 MB
Extranet BN2-Net	2’870 Dokumente	1’180	Ca. 880 Änderungen in 2003	635 MB

Tabelle 4-8: Mengengerüst für indizierte Inhalte bei IWI-HSG

4.4.7. Kosten- und Nutzenaspekte

Da das Direktorium des Instituts als Projektauftraggeber zu keinem Zeitpunkt eine Quantifizierung forderte, basieren die Nutzenaspekte auf den überwiegend positiven bis sehr positiven Rückmeldungen der Anwender. Hervorgehoben wird dabei die Erleichterung der Informationsfindung durch die Verfügbarkeit und Präzision der übergreifenden Suche. Folgende Kennzahlen von 2003 zeigen die intensive Nutzung:

- In den 9 im aktiven Einsatz befindlichen Team-Communities und 3 Literatur- und Publikationsdatenbanken sind insgesamt über 14’000 auf Basis der zentralen Taxonomie klassifizierte Dokumente enthalten.
- In den im IWI-Web verfügbaren 1'066 Publikationen der Institutsmitarbeiter wurde in 2003 von externen Nutzern 267-mal über Volltextsuche und attributbasierte Suche recherchiert.
- Der KPort LinkSampler wurde von Juli bis Dezember 2003 durchschnittlich 115-mal am Tag aufgerufen. Dabei haben die Mitarbeiter insgesamt 173 datenbankübergreifende Suchanfragen über die 9 Team-Communites und 3 Literatur- und Publikationsdatenbanken gestartet.
- Die Analyse der Protokolldatei weist im IWI-Web im Jahresdurchschnitt 18 Anwendersitzungen pro Tag aus. Im gesamten Jahr erfolgten 118 integrierte Suchanfragen über die Inhalte aus IWI-Web und Publikationen.

Für das Projekt ‚Knowledge Port’ war ein Gesamtbudget im unteren sechsstelligen Bereich (Schweizer Franken) vorgesehen. In den fast 1,5 Jahren Planung und Implementierung wurden dabei insgesamt 210 interne und 50 externe Personentage aufgewendet. Da ‚Knowledge Port’ ein vollständig neues, umfassendes Informationssystem implementiert hat, ist der reine Information Retrieval-Anteil nur schwer zu benennen. Zählt man neben Planung und Implementierung der eigentlichen Suchmaschine auch die Taxonomiedatenbank und die Entwicklung des Kern-Datenmodells dazu, liegt der Anteil schätzungsweise bei 20 bis 25 %.

4.4.8. Kritische Erfolgsfaktoren und geplante Weiterentwicklungen

- Die Definition von Anwendungsstandards für Funktionalität, Layout und Zugriffskonzept hat die Nutzung für die Anwender übersichtlicher und komfortabler gemacht.
- Die nahtlose Integration der Applikationen wurde durch die Schaffung eines einheitlichen Datenmodells und die Nutzung der Software-Systeme eines Herstellers erheblich erleichtert.
- Die Anbindung aller Anwendungen an das zentrale Benutzerverzeichnis vereinfacht die Administration und Pflege der beteiligten Personen, Gruppen und Rollen.
- Die gemeinsame Nutzung einer zentralen Taxonomiedatenbank für die Verschlagwortung von Inhalten und die Auswahl von Werten für eine attributbasierte Suche erhöhen die Treffergenauigkeit von Suchanfragen.

Die Projektverantwortlichen stehen weiterhin mit den erwähnten Portalsoftwareanbietern im Gespräch, um die bestehenden Lösungskomponenten auf Basis einer einheitlichen Portalplattform für alle Zielgruppen zu integrieren. Zusätzlich steht organisatorisch eine Konsolidierung der mittlerweile über 2'700 Topics an, mit denen die thematischen Schwerpunkte von Dokumenten klassifiziert werden. Dabei müssen Gruppen gebildet werden, in die die bestehenden Topics eingeordnet werden können. So lassen sich detaillierte Bezeichnungen wie ‚Knowledge-Management-Modelle', ‚Knowledge-Management-Prozesse', ‚Knowledge-Management-Strategien', ‚Knowledge Management', ‚Knowledge Management System', ‚Knowledge-Management-Architektur' beispielsweise in den Oberbegriff ‚Knowledge Management' einordnen. Diese Topic-Gruppen müssen über einen Suchmaschinenthesaurus umgesetzt werden, so dass die Suche eines Internet-Benutzers nach ‚Knowledge Management' alle mit den in dieser Gruppe enthaltenen Begriffen klassifizierten Dokumente zurückliefert.

4.5. Erkenntnisse aus den Fallstudien

Die Lösungen in den untersuchten Organisationen zeigen unterschiedliche Zielsetzungen und Ergebnisse bei der Planung und Implementierung von Information Retrieval in Portalen. Die folgende vergleichende Fallstudienanalyse überprüft anhand dieser Praxislösungen die Relevanz der im integrierten Metamodell beschriebenen Gestaltungselemente.

4.5.1. Identifikation und Beurteilung des Unterstützungspotenzials

Während die untersuchten Gestaltungsmodelle aus der Forschung zu Beginn eines IR-Projekts teilweise stark auf die Beschreibung von Systemfunktionen wie Indizierung oder Suche fokussieren, identifizieren und beurteilen die in der vorliegenden Arbeit untersuchten Organisationen das Unterstützungspotenzial aus unterschiedlichen Blickwinkeln. In allen Fällen waren Ineffizienzen in der täglichen Arbeit der Auslöser für ein Projekt:

- Mit der Einführung einer Suchmaschine im Customer Portal von Zumtobel Staff stand der *verbesserte Zugriff der Kunden auf relevante Informationen und damit eine umfassendere Unterstützung ihrer Prozesse* im Vordergrund. Durch die Bereitstellung einer übergreifenden Suchmöglichkeit über die redaktionell erstellten Inhalte des Content-Management-Systems (CMS) und die strukturierten Daten der Produktdatenbank (PDB) ist es allen Kunden möglich, mit einer einzigen Suchanfrage einen umfassenden Überblick über sämtliche angebotenen Inhalte, z.B. zu einem bestimmten Produkt, zu gewinnen.

- Ausgangspunkt der McB-Portalinitiative bildet die Unterstützung von Prozessen über Navigation und Suche, die im Rahmen einer Anforderungsanalyse mit Anwendervertretern modelliert worden sind. Die mit den Merkmalen der sechs erhobenen Leistungs- und Unterstützungsprozesse klassifizierten Dokumente wie Checklisten, Vordrucke oder Aufgabenbeschreibungen können von den McB-Consultants über Navigation und Suche gezielt gefunden werden. Sie unterstützen damit punktuell jeden einzelnen Arbeitsschritt im Rahmen des McB-Projekts und führen sowohl zu einer *Steigerung der Ergebnisqualität als auch zu einer Reduktion der Bearbeitungszeiten*.

- Als zentrale Anlaufstelle dient der ‚Knowledge Pool' der IMG den Beratern zur Navigation und datenbankübergreifenden Suche in qualifizierten Inhalten. Das Unterstützungspotenzial durch Information Retrieval besteht in der *Zeitersparnis durch kürzere Recherchezeiten*, der *Qualitätssteigerung der Arbeitsergebnisse durch den besseren Zugriff auf bestehende Dokumente* mit gleicher Themenstellung und der *Kostenreduktion durch Vermeidung von Doppelarbeiten*, *Vermeidung von Fehlern aus Unkenntnis* und der *zeitnahen Verfügbarkeit von Informationsobjekten*.

- Die IR-Lösung bei IWI-HSG adressiert eine *Qualitätssteigerung durch den verbesserten Zugriff auf Informationen*. Das System ‚Knowledge Port' konsolidiert die bestehende heterogene Systemlandschaft für die drei im Rahmen einer Anforderungsanalyse definierten Zielgruppen (Intranet, Extranet, Internet). Das kontrollierte Vokabular der attributbasierten Suche im IWI-Web unterstützt dabei z.B. die Orientierung der Öffentlichkeit in der Begriffswelt und den Inhalten des Instituts. Die gezielte Identifikation von Inhalten über das Attribut ‚Topic' erleichtert den Institutsmitarbeitern, z.B. in Forschungsprozessen ihre zumeist themenbezogene Recherche in den über 14'000 internen Dokumenten.

Die untersuchten Praxisfälle sehen das Unterstützungspotenzial von Information Retrieval in Portalen primär in der Qualitätssteigerung und/oder Zeitersparnis. Bemerkenswert ist dabei, dass in keinem Praxisbeispiel von den Verantwortlichen eine quantitative Beurteilung dieser Aspekte gefordert war. Somit hing die Realisierung der untersuchten Retrieval-Projekte nicht primär von finanziellen Vorteilen ab. Ausgangspunkt war vielmehr ein zunehmender Leidensdruck innerhalb der Organisationen [s. Senger 2004], z.B. den Kunden oder Anwendern die Informationsidentifikation zu erleichtern.

Eine gründliche Erhebung der Anwenderanforderungen entlang ihrer Strategie, Prozesse und Systeme ist zu Beginn eines jeden IR-Projektes unabdingbar. Nur auf diese Weise können die Bedürfnisse umfassend berücksichtigt und in den zu schaffenden Abläufen und Applikationen abgebildet werden. In allen Praxisbeispielen

erfolgt die Analyse der Anwenderbedürfnisse über die strukturierte Befragung von repräsentativen Anwendervertretern. Bei Winterthur, IMG und IWI führten die Projektmitarbeiter zusätzlich Workshops durch und interviewten einzelne Führungskräfte.

Umfassende Nutzungskennzahlen für Information Retrieval sind nur bei IWI-HSG verfügbar und wurden dort auch erst nach der Umsetzung der Lösung definiert. Alle Praxisbeispiele verzichteten also auf die Formulierung von Kennzahlen vor dem eigentlichen Projektbeginn. Die nachträgliche Definition und Extraktion relevanter Kennzahlen, z.B. aus Zugriffsprotokollen von Web-Servern, stellte sich bei IWI-HSG als aufwändig heraus.

Das integrierte Metamodell für Information Retrieval berücksichtigt die umfassende Identifikation und Beurteilung von Unterstützungspotenzialen durch seine ganzheitliche Betrachtung von Strategie-, Prozess-, und Systemebene. Sie sind für die Geschäftsfeldstrategie in den Begriffen ‚Qualitätssteigerung', ‚Zeitersparnis' und ‚Kostenreduktion' zusammengefasst (vgl. Abbildung 3-13, S. 55). Im Bereich der Quantifizierung von Unterstützungspotenzialen stösst das Metamodell an seine Grenzen und verweist auf eine detaillierte Analyse der jeweiligen informationszentrierten Aufgaben. Entsprechende Kennzahlen, z.B. Durchlaufzeit, müssen demnach von den Leistungsprozessen vorgegeben werden, die Information Retrieval unterstützt.

4.5.2. Analyse existierender Informationsstrukturen und Applikationen

In allen dargestellten Praxisfällen wurden die vorhandenen Informationsstrukturen und Anwendungen bereits zu Projektbeginn analysiert. Die Bewertung der Integrationsfähigkeit vorhandener Systeme anhand einiger weniger Merkmale, z.B. verfügbare Ausgabeschnittstellen und -formate, vorhandene Informationsstrukturen oder existierende Benutzerverzeichnisse kann als spezifisch für den Aufbau einer Information Retrieval-Lösung angesehen werden. Sie beeinflussen die Auswahl der zur Verwendung kommenden Systeme wie Portal, Suchmaschine oder Content-Applikationen erheblich:

- Bereits vor Projektbeginn stand bei der Winterthur Versicherungen aufgrund von Budgetrestriktionen und internen Richtlinien fest, dass Portal- und Suchfunktionen ausschliesslich durch die Anpassung und den Ausbau der vorhandenen Plattform Opentext Livelink bereitgestellt werden können. Da die Konzeption des McB-Portals mit der Ausweitung der McB-Projektinitiative weitestgehend zeitgleich erfolgte, konnte eine umfassende Analyse bestehender Systeme auf ein Minimum reduziert werden. Bestehende Informationsstrukturen wurden mit geringem

Aufwand aus den Erfahrungen der ersten McB-Projekte und der rudimentären Ablagestruktur auf dem gemeinsamen Fileserver abgeleitet.

- Auch die Verantwortlichen beim IMG wollten die bestehende Lotus-Notes-Infrastruktur soweit wie möglich nutzen. Eine detaillierte Analyse der ‚gewachsenen' Ablagestrukturen in Mail-Datenbanken, Dateiablagen oder Projektlaufwerken erfolgte jedoch nicht, da eine umfassende Migration von Altdatenbeständen für den Aufbau des neu zu erstellenden Informationssystems nicht notwendig war.
- Durch die grosse Systemvielfalt bei IWI-HSG zeichnete sich bereits zu Projektbeginn ab, dass bei der Integration der Inhalte in einem Portal und der Anbindung an eine Suchmaschine eine Vielzahl heterogener Schnittstellen entstehen würde. Unterschiedliche Zugriffsbeschränkungen, Datenstrukturen und Ausgabeformate führten daher frühzeitig zu der Entscheidung, ein neues und umfassendes Informationssystem auf Basis der am Institut vorhandenen Groupware-Plattform Lotus Notes zu implementieren. Dies war jedoch, wie auch bei IMG, nur möglich, da eine Vielzahl von Altdaten aufgrund der teilweise begrenzten Gültigkeit nicht migriert bzw. die Migration aufgrund des geringen Datenvolumens nicht systemtechnisch unterstützt werden musste.
- Das Portal bei Zumtobel Staff war durch die zwei führenden Applikationen Content-Management-System (CMS) und Produktdatenbank (PDB) geprägt. Im Vergleich zu den Praxisfällen von Winterthur, IMG und IWI-HSG konnten deren Inhalte jedoch aufgrund des stark erweiterten Funktionsumfangs (z.B. Bestellprozesse in der PDB) und des damit verbundenen Investitionsschutzes nicht einfach in ein neues System überführt werden. Bereits der Projektauftrag enthielt darüber hinaus die Anweisung, nur bei unzureichendem Funktionsumfang auf die Nutzung der bereits mit dem Portal lizensierten Suchmaschine SAP TREX zu verzichten.

Das integrierte Metamodell für Information Retrieval bildet die Analyse existierender Informationsstrukturen und Anwendungen durch die Komponenten der Systemebene ab (vgl. Abbildung 3-17, S. 63). Die Wirkungszusammenhänge der beteiligten Applikationen, Funktionen und Datensammlungen zeigen einzelne Interventionspunkte für die Gestaltung der Integration auf, ohne den Gesamtzusammenhang zu vernachlässigen.

4.5.3. Entwicklung einer Retrieval-Taxonomie

In allen Praxisfällen werden Inhalte bei der Erstellung manuell klassifiziert, um eine gezielte Informationsidentifikation, z.B. zu einem Themenschwerpunkt, einem

Produkt oder einer Dienstleistung zu ermöglichen. Die Retrieval-Taxonomie als Klassifikationsgerüst bzw. Ordnungsrahmen weist dabei in Art und Umfang teilweise grosse Unterschiede auf:

- Zur Klassifikation von Inhalten im McB-Portal verwenden die Mitarbeiter der Winterthur Versicherungen maximal 11 Attribute, z.B. ‚Project', ‚Process' oder ‚Status'. Die Wertebereiche schwanken zwischen 6 und 18 Begriffen.
- Die Retrieval-Taxonomie bei IWI-HSG weist 5 Klassifikationsmerkmale auf. Eine gezielte Informationssuche ist dadurch im Portal nach dem Autor, dem publizierenden Lehrstuhl, dem Kompetenzzentrum, dem Veröffentlichungszeitraum und dem inhaltlichen Schwerpunkt (‚Topic') eines Dokuments möglich. Pro Attribut existieren zwischen 4 und 25 Einträge.
- Inhalte werden bei der IMG mit bis zu 13 Attributen klassifiziert. Die unterschiedlichen Dimensionen berücksichtigen dabei z.B. Sprach- und Datumsangaben, Informationen zu Produkten und Dienstleistungen oder Bezeichnungen für Kunden, Projekte und Branchen. Die einzelnen Attribute weisen Wertebereiche zwischen 11 und 37 Einträgen auf.
- Obwohl bei Zumtobel Staff ausschliesslich eine Volltextsuche zur Verfügung gestellt wird, erfolgt auch hier die Klassifikation von Inhalten. Um den unterschiedlichen Zielgruppen Inhalte gemäss ihrer Zugriffsberechtigungen bereitzustellen, werden Dokumente mit Attributen für ‚Land', ‚Sprache' und ‚Zielgruppe' klassifiziert.

Die beschriebenen Lösungen stellen ihre Retrieval-Taxonomie für eine attributbasierte Suche (AB) oder eine taxonomieorientierte Navigation (TO) zur Verfügung. Eine Volltextsuche (VT) wird in allen Fällen abgebildet.

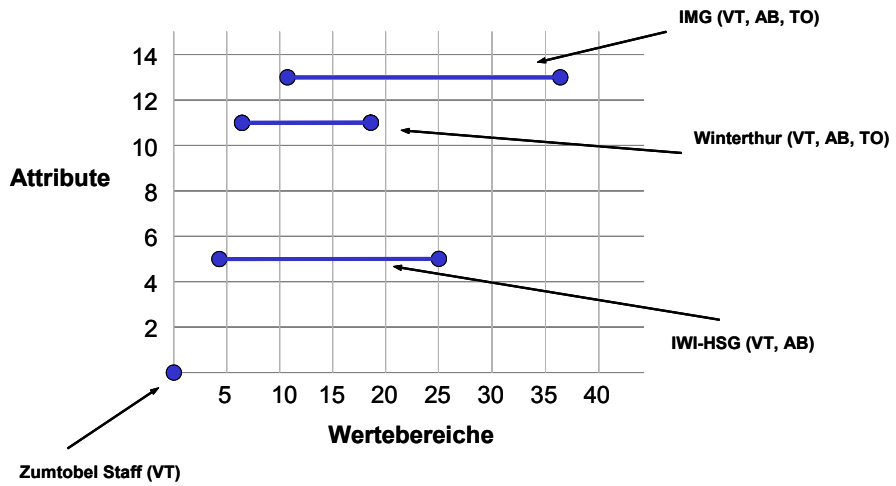


Abbildung 4-27: Vergleich des Umfangs der Retrieval-Taxonomie bei den Praxisfällen

Alle Ansätze weisen als gemeinsame Herausforderung auf, dass sie ‚unterschiedliche Begriffswelten' zwischen dem Ersteller eines Informationsobjekts und einem Suchenden abstimmen müssen. Je umfangreicher die Retrieval-Taxonomie ist, desto exaktere Suchanfragen können die Anwender formulieren. Die Erstellung und Pflege einer Vielzahl von Attributen und Wertebereiche steigert den Aufwand für die Betreiber. Die Abstimmung der mit 5 Attributen eher kleineren Retrieval-Taxonomie bei IWI-HSG nahm mit den Anwendervertretern bereits mehrere Personentage in Anspruch.

- Je weniger Transparenz beim Portalbetreiber über die Zielgruppe und deren ‚Begriffswelt' existiert, umso schlanker fällt die Retrieval-Taxonomie aus. Zumtobel Staff adressiert beispielsweise eine Vielzahl von Benutzern mit unterschiedlichen Informationsbedürfnissen. Die Bereitstellung einer Volltextsuche oder geplante Weiterentwicklungen hinsichtlich einiger weniger produktbeschreibender Inhaltsattribute berücksichtigen diese Ausgangssituation.

- Themen, Produkte, Dienstleistungen und Kundennamen unterliegen bei den Mitarbeitern der IMG einem gemeinsamen Begriffsverständnis. Sie entstammen z.B. dem Dienstleistungsportfolio, der Aufbauorganisation, den Unternehmenspräsentationen oder der betriebswirtschaftlichen Standardsoftware. Aufbau (Syntax) und Bedeutung (Semantik) sind den beteiligten Personen weitestgehend bekannt. Die hohe Transparenz dieser ‚Begriffswelt' erlaubt eine umfangreiche Retrieval-Taxonomie für die attributbasierte Suche und die taxonomieorientierte Navigation. Darüber hinaus können viele Attribute und Werte zur Klassifikation von Inhalten aus bereits bestehenden Systemen übernommen werden, z.B. Kunden- und Projektbezeichnungen aus der Projektverwaltung.

- IWI-HSG bedient sowohl bekannte Anwender (Intranet und Extranet) als auch unbekannte Zielgruppen (Internet). Diese duale Zielsetzung führt zu einer Retrieval-Taxonomie, die lediglich 3 inhaltsbeschreibende Attribute vorsieht. Die Mitarbeiter von IWI-HSG kennen die Bedeutung von ‚Lehrstuhl', ‚Kompetenzzentrum' und ‚Topic' in den Auswahllisten. Von anonymen Anwendern kann sie bei einer attributbasierten Suche im Internet-Auftritt von IWI-HSG im Bedarfsfall mit geringem Aufwand erschlossen werden.

Bemerkenswert ist, dass in allen Projekten zu Beginn eine umfassende Retrieval-Taxonomie mit einer möglichst grossen Anzahl von Attributen angedacht war. Die damit verbundenen Erhebungs- und Abstimmungsaufwände führten jedoch immer dazu, dass in der finalen Lösung ausschliesslich die für das Unternehmen wichtigsten Dimensionen in der Retrieval-Taxonomie abgebildet waren.

Das integrierte Metamodell für Information Retrieval adressiert die Entwicklung einer Retrieval-Taxonomie sowohl auf System- als auch auf Prozessebene (s. Abbildung

3-15, S. 58 und Abbildung 3-17, S. 63). Es zeigt z.B., dass Attribute und Werte aus bestehenden Datensammlungen verwendet werden können (‚Struktur'). Darüber hinaus bildet es den Pflegeprozess ‚Terminologiemanagement' ab, der die Benennung, Strukturierung und Bereitstellung einer Retrieval-Taxonomie umfasst.

4.5.4. Spezifikation von Pflegeprozessen

Die untersuchten Modelle aus der Forschung bieten Gestaltungshinweise für die Festlegung von Pflegeprozessen, z.B. für den Betrieb einer Retrieval-Taxonomie, für Abläufe zur systematischen Integration von Anwendungen oder Prozesse für den fortlaufenden Abgleich sich ändernder Benutzerberechtigungen (s. Abschnitt 3.2). Die Notwendigkeit zur Spezifikation dieser Pflegeprozesse zeigt sich auch in den skizzierten Fällen der betrieblichen Praxis:

- Nach dem ersten Produktivbetrieb im Juni 2001 stellte sich bei der Winterthur Versicherungen heraus, dass die vorgeschlagene Klassifikationsstruktur im McB-Portal zu detailliert war. Die Mitarbeiter konnten nicht mehr exakt festlegen, welche ihrer Dokumente sie sinnvollerweise mit welchen Werten der Retrieval-Taxonomie klassifizieren sollten. Ein Strukturierungsvorschlag für das Ablegen und Auffinden von Vorlagen und Dokumenten im Portal war zwar gewünscht, sollte jedoch einen geringeren Umfang aufweisen. Die Klassifikationsstruktur wurde daraufhin vereinfacht.

- Die Liste der verfügbaren Themenschwerpunkte (‚Topics') für eine attributbasierte Suche in den Publikationen von IWI-HSG wuchs in 1,5 Jahren operativem Betrieb auf über 2'700 Begriffe an. Begriffsänderungen können zwar durch Synonymlisten abgebildet werden, diese sind jedoch in der Suchmaschine noch nicht abgebildet. Internet-Benutzer finden daher bei der themenspezifischen Suche nach Dokumenten in einer Auswahlliste mehr als 580 Topics.

- Da der Hersteller des Content-Management-Systems (CMS) häufig neue Software-Updates bereitstellte, musste die Anpassung der Anwendungsschnittstellen in der heterogenen Systemumgebung von Zumtobel Staff häufig ad-hoc erfolgen. Die Nutzung der jeweils neuesten CMS-Version war für einen stabilen Systembetrieb notwendig, erschwerte aber andererseits die Spezifikation der Suchmaschine.

- Bei Winterthur Versicherungen, IMG und IWI-HSG war keine umfassende Planung von Abläufen zur Bereitstellung der Anwendungsintegration oder für die Benutzerschnittstellen notwendig, da in diesen Fällen die Information Retrieval-Lösungen auf Basis einheitlicher Plattformen bereitgestellt wurden. Das McB-Portal der Winterthur Versicherungen auf Basis von Opentext Livelink und die

Lösungen von IMG und IWI-HSG auf Basis von Lotus Notes boten bereits vor Projektbeginn entsprechende interne Schnittstellen.

- IMG und IWI-HSG konnten bestehende Berechtigungskonzepte erweitern. Die benötigten Benutzer, Rollen und Benutzergruppen lagen bereits in zentralen Benutzerverzeichnissen vor. Die im so genannten Domino Directory der Groupware-Plattform Lotus Notes von den zentralen Administrationen gepflegten Angaben wurden in beiden Organisationen z.B. bereits für die Adressierung von E-Mails verwendet. Auch Winterthur Versicherungen und Zumtobel Staff konnten bestehende Benutzerkonten für ihre Portale verwenden. Während im McB-Portal der Winterthur Versicherungen auch existierende Benutzergruppen des Systems ‚LifeLink' Verwendung fanden, mussten die für die Lesebeschränkungen notwendigen Zielgruppen ‚io.cust' und ‚io.empl' im Zumtobel Staff Benutzerverzeichnis neu eingepflegt werden. Dazu wurde jedem registrierten Benutzer manuell eine Benutzerrolle zugeordnet.

Das integrierte Metamodell adressiert die für Aufbau und Betrieb einer IR-Lösung notwendigen Pflegeprozesse. Sie sind auf Prozessebene über das Index-, Terminologie- und Integrationsmanagement abgebildet bzw. durch ihre zugehörigen Aufgaben wie z.B. ‚Berechtigungen pflegen', ‚Taxonomie benennen & strukturieren' oder ‚Inhalte klassifizieren' verfeinert (s. Abbildung 3-15, S. 58).

4.5.5. Umsetzung von Anforderungen im Systementwurf

Die technische Umsetzung von Navigation und Suche in den untersuchten Praxisfällen konfrontierte die Verantwortlichen von vorhandenen und neu zu schaffenden Systemen mit unterschiedlichen Herausforderungen:

- Die Integration der bestehenden Content-Applikationen wie Fileserver, Datenbanken oder individuelle Ablagesysteme wäre aufgrund der Vielzahl der dabei entstehenden heterogenen Schnittstellen bei IMG und IWI-HSG nicht wirtschaftlich gewesen. Beide Systementwürfe sahen daher vor, ein neues umfassendes System bereitzustellen, in dem Portal, Suchmaschine und Inhalte bereits aufeinander abgestimmt sind, z.B. durch eine einheitliche Inhaltsklassifikation oder ein gemeinsames Benutzerverzeichnis. Bei Zumtobel Staff dagegen mussten das Content-Management-System und die Produktdatenbank aufgrund des Investitionsschutzes in das IR-Konzept integriert werden.

- Zur Bereitstellung eines Single Sign Ons (SSO), bei dem ein am Portal authentifizierter Benutzer ohne weitere Login-Aufforderungen auch über die Suchmaschine auf geschützte Inhalte zugreifen kann, ist im Systementwurf bei der Winterthur Versicherungen keinerlei Anpassung notwendig. Das McB-Portal bietet

in einer einzigen Applikation Portal-, Suchmaschinen- und Inhaltsverwaltungsfunktionen. Eine leistungsfähige Retrieval-Lösung konnte dadurch ohne grosse Integrationsschwierigkeiten ‚aus einem Guss' bereitgestellt werden. Während bei IMG und IWI-HSG Portal und Suchmaschine bereits durch einfache Konfigurationseinstellungen ein gemeinsames Benutzerverzeichnis auf Basis von Lotus Notes nutzten, erforderte das SSO bei Zumtobel Staff die umfassende Anpassung aller beteiligten Applikationen durch Programmierung.

- Für die Implementierung der Information-Retrieval-Systeme aller Praxislösungen bildet die Definition eines zentralen und durchgängig genutzten Datenmodells sowie einer darauf abgestimmten Retrieval-Taxonomie die Basis für eine konsistente Navigation sowie für eine präzise Volltextsuche und attributbasierte Suchfunktion.

Die Gestaltungselemente ‚Portal', ‚Suchmaschine' und ‚Content-Applikation' sowie die ihnen zugeordneten Funktionen und Datensammlungen berücksichtigen im integrierten Metamodell für Information Retrieval den Systementwurf (s. Abbildung 3-17, S. 63). In den Praxisfällen auftretende Herausforderungen sind stark abhängig von der jeweiligen IT-Architektur. Hier stösst das Metamodell aufgrund der hohen Varianz technologischer Komponenten an seine Grenzen.

5. Methodenvorschlag für Information Retrieval in Portalen

Die im Rahmen der vorliegenden Arbeit untersuchten Forschungsansätze (s. Kapitel 3) sowie Praxisbeispiele (s. Kapitel 4) zeigen die Herausforderungen bei der Einführung und dem Betrieb von Information Retrieval in Portalen:

- Die Mitarbeiter von Zumtobel Staff standen vor der Herausforderung, eine vom Unternehmen lizensierte Suchmaschine zu verwenden, ohne jedoch die umfassende Umsetzung der Anforderungen zu vernachlässigen. Zusätzlich musste die neue Information-Retrieval-Lösung so modular und flexibel aufgebaut sein, dass die Integration weiterer Anwendungen zu einem späteren Zeitpunkt ohne grösseren Implementierungsaufwand möglich ist.
- Die Verantwortlichen bei IWI-HSG mussten aufgrund der Ergebnisse des Machbarkeitsnachweises mit zwei Softwareherstellern die Einführung einer kommerziellen Portalplattform verschieben. Die im Rahmen des Projekts erstellte Information-Retrieval-‚Zwischenlösung' sollte jedoch so umfassend geplant und implementiert werden, dass sie zu einem späteren Zeitpunkt in eine kommerzielle Portalsoftware integriert werden kann.
- Die Klassifikationsstruktur des Winterthur-Portals sollte an den Standardprozessen der Mitarbeiter ausgerichtet werden, aufgrund von Erfahrungen aus abgeschlossenen Projekten und neuen Anforderungen aber auch kontinuierlich anpassbar sein. Die Umsetzung der erforderlichen Navigations- und Suchfunktionen musste aufgrund von Budgetrestriktionen und internen Richtlinien mit einem im Unternehmen bereits vorhandenen System erfolgen.
- Die Information-Retrieval-Lösung der IMG musste den Mitarbeitern ein unternehmensweit einheitliches System zur Verfügung stellen, ohne jedoch deren bisherige individuelle Ablagestrukturen und dezentrale Arbeitsweise zu vernachlässigen.

Das folgende Kapitel detailliert wesentliche Elemente zur Planung und Implementierung von Information Retrieval in Portalen anhand einer Methode im Sinne des Methodenengineerings (s. Abschnitt 2.1.4). Die enthaltenen Vorgehens- und Rollenmodelle beschreiben die Integration und Bereitstellung von Inhalten aus Content-Applikationen durch eine Portalsuchmaschine im Rahmen eines Projekts.

Obwohl sich wesentliche Bestandteile der Methode in der Praxis bewährt haben (s. Praxisfälle in Kapitel 4), muss von einem Methodenvorschlag gesprochen werden, da eine durchgängige und umfassende Anwendung noch aussteht.

5.1. Rollenmodell

Eine Rolle wird von einer Geschäftseinheit, einem Mitarbeiter oder einem Kunden wahrgenommen. Sie ist mit Aufgaben, Kompetenzen und Verantwortung verbunden und ordnet im Rahmen der Methode den Aufgabenträgern konkrete Aktivitäten zu.

Das Rollenmodell weist zwei Dimensionen auf:

Dimension ‚Stufe'
• *Entscheider*. Personen, die Entscheidungen treffen können oder müssen
• *Verantwortlicher*. Zuständige Person für das termingerechte Erreichen und Weiterleiten von Ergebnissen
• *Unterstützer*. Personen, die inhaltliche Aussagen treffen können und als Informationslieferanten in das Projekt eingebunden sind
Dimension ‚Gruppe'
• *Initiatoren*. Personen, die ein Information-Retrieval-Projekt initiieren und die Umsetzung massgeblich prägen
• *Anwender*. Personen, die als externe Kunden oder interne Mitarbeiter Abnehmer für die über das Portal angebotenen Information-Retrieval-Leistungen sind.

Tabelle 5-1: Rollenmodell

Die Gruppe der Anwender ist für die gesamte Methode homogen. Sie setzt sich aus Fachbereichsverantwortlichen oder -mitarbeitern zusammen, die stellvertretend für eine Kunden- oder Mitarbeitergruppe stehen. Die Zusammensetzung der Initiatoren mit Unterstützungsfunktion ist spezifisch und den jeweiligen Techniken vorangestellt.

Das Beispiel in Tabelle 5-2 veranschaulicht die Anwendung des Rollenmodells innerhalb der Techniken.

Stufe	Initiatoren	Anwender (Mitarbeiter, Kunden)
Entscheider	Rolle, z.B.Führungsverantwortlicher	Rolle, z.B. Fachbereichsverantwortlicher
Verantwortlicher	Rolle, z.B. Projektleiter	Rolle, z.B. Fachbereichsvertreter
Unterstützer	Rolle, z.B. Anwendungsentwickler	Rolle, z.B. Fachbereichsmitarbeiter

Tabelle 5-2: Beispiel für die Anwendung des Rollenmodells

5.2. Übersicht

Die vorliegende Methode unterstützt Verantwortliche bei der Analyse, Planung und Implementierung von Information Retrieval in Portalen im betrieblichen Umfeld. Sie orientiert sich an den Zielsetzungen der von [Thiesse 2001, 106-194] entwickelten Methode zur Einführung von Wissensmanagement, die im Rahmen weiterer Arbeiten bereits erfolgreich adaptiert worden ist (vgl. [Gebert 2003, 135-224] [Bach et al. 2000, 117], [Blessing 2001 , 32]). Die Methode umfasst vier Techniken (s. Abbildung 5-1):

- In der *Potenzialanalyse* identifizieren die Initiatoren Einsatzmöglichkeiten und Restriktionen von Information Retrieval in Portalen. Sie konsolidieren die Ergebnisse in Form von Handlungsempfehlungen für ein zukünftiges Information-Retrieval-Projekt.
- In der *Strategieplanung* legen die Entscheider die Retrieval-Leistungen fest. Sie skizzieren die zur Umsetzung erforderliche Informationssystemarchitektur, schätzen Kosten- und Nutzenaspekte ab und schlagen eine Umsetzungsplanung vor. Die Ergebnisse der Strategieplanung bilden die Entscheidungsgrundlage für die Umsetzung einer Information-Retrieval-Lösung.
- In der *Prozessplanung* leiten die Verantwortlichen Prozesse zur Integration von Content-Applikationen, zur Entwicklung eines Benutzer- und Berechtigungs-modells für die Portalsuche und zur Entwicklung und Pflege einer Retrieval-Taxonomie ab. Die Prozessplanung schliesst damit die Anforderungsanalyse ab.
- Im *Systementwurf* spezifizieren die Projektmitarbeiter die zur Umsetzung der fachlichen Anforderungen notwendigen Systemfunktionen. Die Ergebnisse des technischen Systementwurfs bilden das Pflichtenheft, auf dessen Basis die Entscheider eine Softwareauswahl treffen.

Die in Abbildung 5-1 dargestellten Techniken bauen aufeinander auf (von oben nach unten). Auch die Bearbeitung der Aktivitäten bei Potenzialanalyse, Strategie- und Prozessplanung erfolgt sequentiell (von links nach rechts). Innerhalb des Systementwurfs ist keine feste Reihenfolge vorgeschrieben. Ein erfolgreiches Projekt muss jedoch alle Techniken und Aktivitäten durchlaufen. Dabei ist es Aufgabe der Projektverantwortlichen, die notwendige Detaillierung der einzelnen Schritte zu bestimmen.

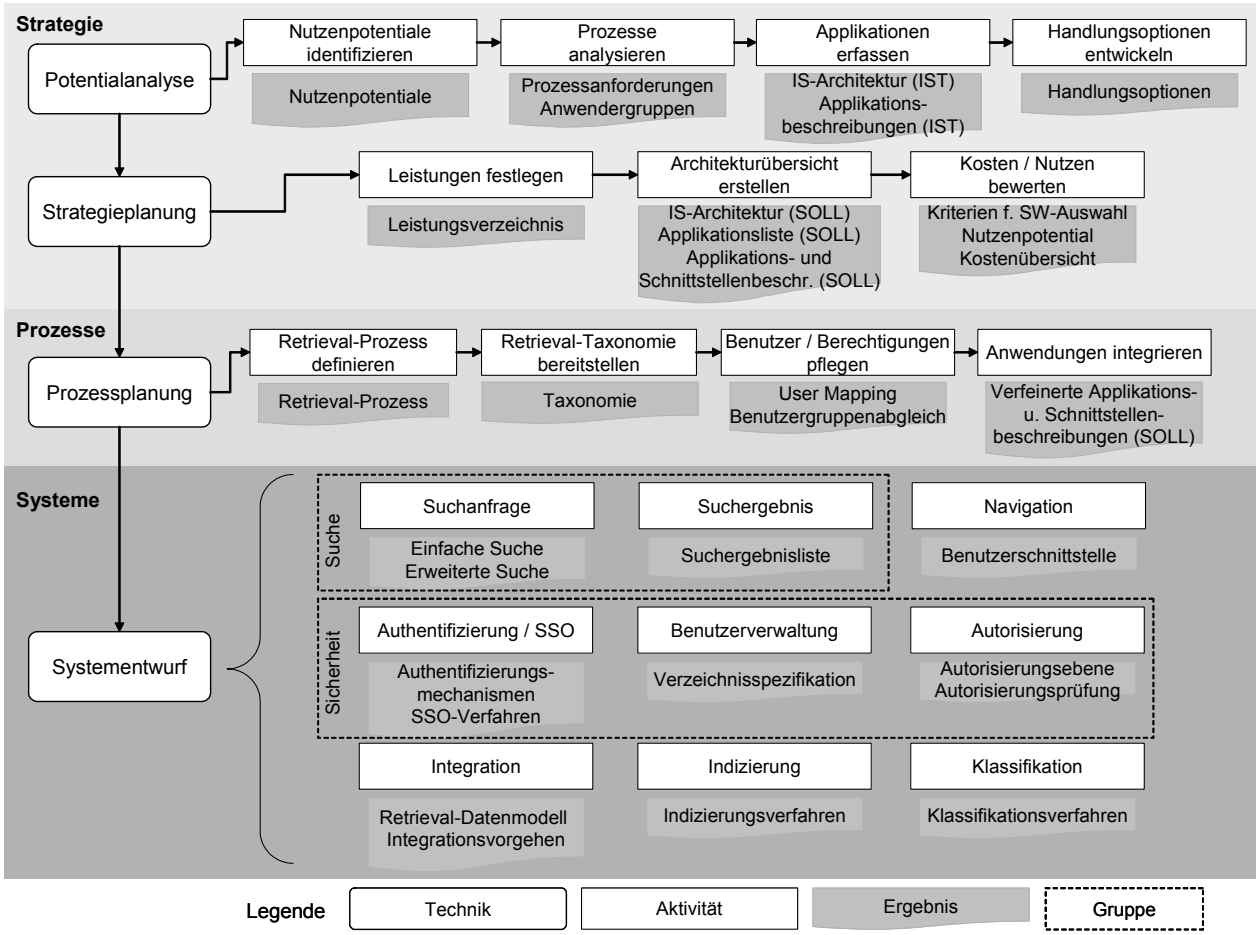


Abbildung 5-1: Vorgehensmodell für Information Retrieval in Portalen

Bei der folgenden Beschreibung der Techniken des Vorgehensmodells orientiert sich die Ausgestaltung an den Anforderungen der inhaltlichen, personellen und wirtschaftlichen Angemessenheit [s. Grochla 1982, 299-300]. Der Aufbau stellt beispielsweise sicher, dass die zur Bearbeitung einer Technik erforderlichen Informationen verfügbar sind und der Durchführungsaufwand in einem sinnvollen Verhältnis zum Nutzen der Technik steht.

Die enthaltenen Beispiele illustrieren den Methodenvorschlag. Sie stammen ausschliesslich aus den Fallstudien (s. Kapitel 4) und detaillieren bzw. vertiefen diese.

Die Methode setzt den Schwerpunkt auf Information-Retrieval-Aspekte. In Praxisprojekten relevante Aktivitäten des Projektmanagements und Change Management (z.B. Projektphasenplan, Verantwortlichkeitsmatrix, Risikoanalyse, Kommunikationsmassnahmen und Schulungen) behandelt der Methodenvorschlag nicht explizit. Projektmanagement und Change Management beeinflussen den Projekterfolg, sind jedoch nicht in ausreichendem Masse Information-Retrieval-spezifisch. Detaillierte Angaben zur Begleitung von Veränderungsprozessen finden sich in [IMG 1997a, 117], [Blessing 2001, 166] und [Ulrich/Fluri 1995, 205].

5.3. Potenzialanalyse

Kurzbeschreibung der Technik: Das Fehlen von Suchfunktionen in einem Portal kann bereits einen ausreichenden Anlass darstellen, der einen Verantwortlichen zur Auslösung eines Information-Retrieval-Projekts bewegt [vgl. Senger 2004, 353]. Die Potenzialanalyse hilft Unternehmen, diesen ‚Leidensdruck' anhand von IR-spezifischen Schwachstellen bei der betrieblichen Aufgabenerfüllung zu konkretisieren. Kritische Erfolgsfaktoren und Führungsgrössen ermöglichen die Operationalisierung der Zielsetzung und eine Messung der Veränderung nach der Einführung einer Suchmaschine. Die Verantwortlichen dokumentieren dazu die Anforderungen kundenorientierter Prozesse an eine Portalsuche. Neben den Einsatzmöglichkeiten identifizieren sie auch mögliche Restriktionen durch die frühzeitige Analyse der Integrationsfähigkeit vorhandener Systeme. Die Potenzialanalyse zeigt Handlungsoptionen auf, die eine Strategieplanung und damit ein Information-Retrieval-Projekt auslösen können.

Ziele: Die Technik zur Potenzialanalyse hat zum Ziel:

- Nutzenpotenziale für Information Retrieval in Portalen identifizieren,
- retrievalintensive Prozesse und die Informationsbedürfnisse der Anwender zu erfassen,
- Informationsobjekte und die sie führenden Anwendungen zu erfassen, die mit dem Informationsbedarf retrievalintensiver Aufgaben korrespondieren,
- Integrationsmöglichkeiten für diese Content-Applikationen zu bewerten und daraus
- Handlungsoptionen für die Umsetzung aufzuzeigen.

Voraussetzungen (Input): - keine -

Ergebnisse (Output): Die Ergebnisse der Technik Potenzialanalyse bestehen aus Prozessanforderungen, der IS-Architektur (IST), Applikationsübersichten und -beschreibungen (IST) sowie den daraus abgeleiteten Handlungsoptionen.

Rollen: Die Technik Potenzialanalyse wird unter Berücksichtigung der in Tabelle 5-3 beschriebenen Rollen angewendet.

Stufe	Initiatoren	Anwender (Mitarbeiter, Kunden)
Entscheider	Führungsverantwortlicher, Prozessverantwortlicher	Fachbereichsverantwortlicher
Verantwortlicher	Projektleiter	Fachbereichsvertreter, Applikationsverantwortlicher
Unterstützer	Projektteam, Anwendungsentwickler, IT	Fachbereichsmitarbeiter

Tabelle 5-3: Rollen der Technik Potenzialanalyse

5.3.1. Nutzenpotenziale identifizieren

Die analysierten Praxisfälle weisen Ziele und Nutzenpotenziale auf, die zu unterschiedlichen Handlungen in den Unternehmen geführt haben (s. Abschnitt 4.5.1):

- Mit der Einführung einer Suchmaschine im Customer Portal von Zumtobel Staff stand der *verbesserte Zugriff der Kunden auf relevante Informationen*, *umfassendere Unterstützung ihrer Prozesse* und somit eine *Erhöhung der Kundenzufriedenheit* im Vordergrund.
- Die Ziele des McB-Portals bestanden in *Steigerung der Ergebnisqualität* und der *Reduktion der Such- und Recherchezeiten*. Ausgangspunkt der McB-Portalinitiative der Winterthur Versicherungen war die Unterstützung jedes einzelnen Arbeitsschrittes der Mitarbeiter durch die Erschliessung relevanter Inhalte über Navigation und Suche.
- Die IR-Lösung bei IWI-HSG adressiert eine *Qualitätssteigerung der Arbeitsergebnisse* durch den *schnelleren Zugriff auf Informationen*. Komfortable Suchfunktionen im IWI-Web sollen die *Aussenwirkung des Instituts verbessern*.

Diese Ziele lassen sich in Anlehnung an ([Bullinger et al. 1997, 16], [KPMG 2001, 15], [Maier 2002, 317], [Heisig/Vorbeck 2001, 104]) in die Faktoren Qualität, Zeit, Kosten einteilen.

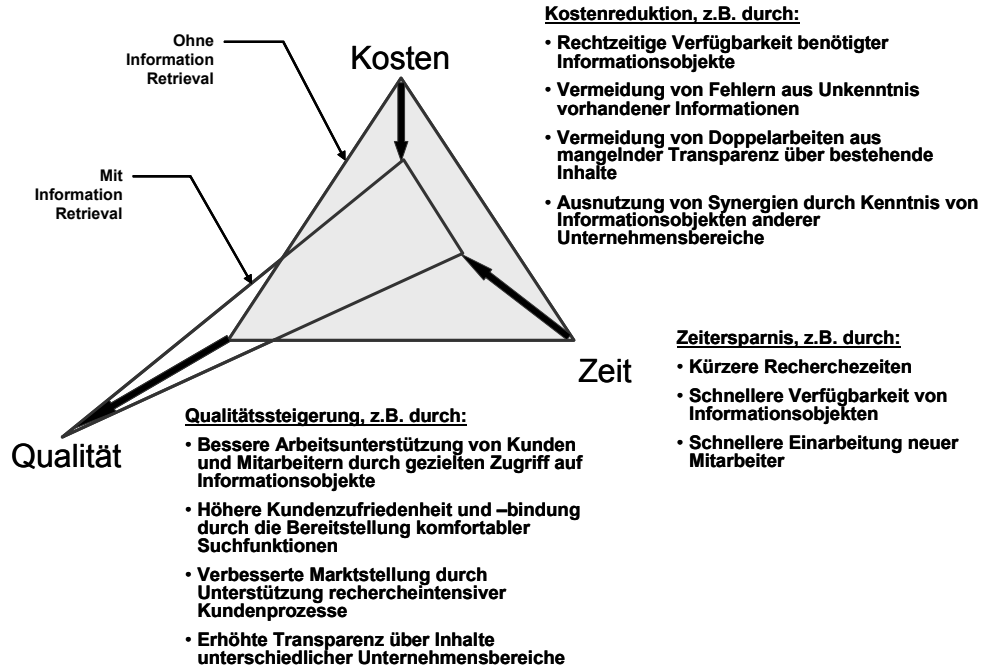


Abbildung 5-2: Mögliche Ziele und Erfolgsfaktoren von Information Retrieval in Portalen (in Anlehnung an [Riempp 2003, 246])

Zur Umsetzung der Ziele wie Qualitätssteigerung, Zeitersparnis oder Kostenreduktion identifizieren die Projektinitiatoren gemeinsam mit den Anwendern kritische

Erfolgsfaktoren. Diese müssen erfüllt sein, damit das jeweilige Ziel überhaupt erreicht werden kann. Am Beispiel einer Portalsuchmaschine wäre dies die bessere Arbeitsunterstützung von Mitarbeitern durch den gezielten Zugriff auf Informationsobjekte, die höhere Kundenzufriedenheit durch die Bereitstellung von Suchfunktionen, kürzere Recherchezeiten etc.

Den Grad der Zielerreichung können die Projektinitiatoren anhand von Messgrössen erheben, die aus den kritischen Erfolgsfaktoren abgeleitet sind, z.B. ‚Zeitdauer für Informationsrecherche zur Erstellung eines Angebots' oder ‚Anzahl Suchanfragen durch Kunden'. Durch die Vorgabe von Sollwerten ergeben sich daraus konkrete Führungsgrössen. Diese beschreiben gemäss [Hess 1996, 121] einen messbaren Sachverhalt und operationalisieren einen kritischen Erfolgsfaktor [s. Kaiser 2000, 91], z.B. ‚Anteil Informationsrecherche < 20% des Gesamtaufwands für die Erstellung eines Angebots' oder ‚Anzahl Suchanfragen > 80 pro Tag'.

Der Projektleiter hält Ziele, kritische Erfolgsfaktoren und Führungsgrössen in einer Tabelle fest. Er stimmt diese mit den Entscheidern ab. Sie bilden die Basis für die Entwicklung weiterer Handlungsoptionen.

5.3.2. Prozesse analysieren

Kernaufgabe von Unternehmen ist die Erstellung von Leistungen, die im Rahmen von Prozessen geschaffen werden (s. [Thiesse 2001, 141], Grundlagen in Abschnitt 2.1). Ein Prozess setzt sich dabei aus einzelnen Aufgaben zusammen, zu deren Durchführung Informationen benötigt werden.

Die Erfolge von Information Retrieval-Lösungen können nur mittelbar erfasst werden, z.B. anhand von Messgrössen wie der Prozessdurchlaufzeit [vgl. North 1999, 149]. Daher müssen die Projektverantwortlichen aus der Vielzahl der geschäftlichen Prozesse diejenigen identifizieren, die durch IR-Ziele (Qualitätssteigerung, Zeitersparnis, Kostenreduktion) verbessert werden können.

Unternehmen bilden ihre Prozesse und Aufgaben häufig in Prozesslandkarten [s. IMG 1997a, 37-38] oder Aufgabenkettendiagrammen [s. Schulze 2000, 142-150] ab. Abbildung 5-3 zeigt eine vereinfachte Prozesslandkarte des Instituts für Wirtschaftsinformatik.

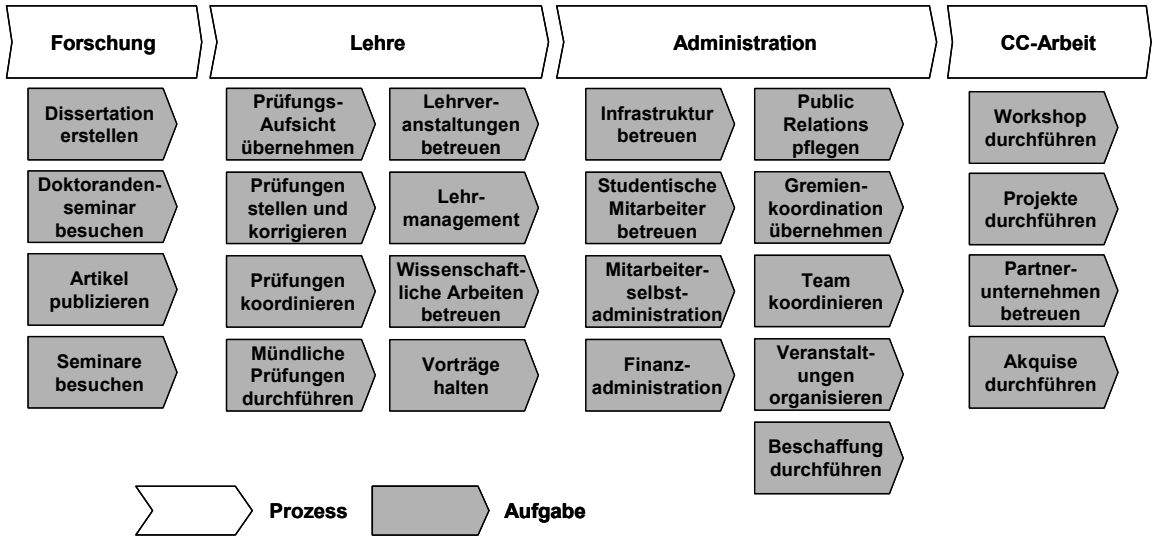


Abbildung 5-3: Prozesslandkarte von IWI-HSG (vereinfacht)

Nicht alle der dargestellten Tätigkeiten benötigen in gleichem Umfang Informationen. Während bei hoch strukturierten Aufgaben (z.B. Erfassung eines Zahlungsauftrags in der Finanzadministration) meist nur wenige, bekannte Informationen benötigt werden, nimmt bei schwach strukturierten Aufgaben (z.B. Publikation eines Artikels) die Lokalisierung, Beschaffung, Interpretation und Aufbereitung einen wesentlich grösseren Raum ein [vgl. Krüger 1994, 94]. Die Unterstützungsmöglichkeit durch Information Retrieval ist damit umso höher, je grösser der Anteil der informations- bzw. recherchebezogenen Tätigkeiten bei der Ausführung einer Aufgabe ist (s. [Kaiser 2000, 121], [Thiesse 2001, 114], [Eppler et al. 1999]).

Recherchebezogene Tätigkeiten identifizieren das Projektteam im Rahmen einer Prozessanalyse [s. Blessing 2001, 32 und 160], z.B. durch Untersuchung von Prozessbeschreibungen, Workshops mit Prozessverantwortlichen oder durch die Befragung von Mitarbeitern. Ein Beispiel[8] illustriert das konkrete Vorgehen:

> *In der Vorstudie führten die Projektmitarbeiter bei IWI-HSG eine Befragung aller Mitarbeiter durch. Ein elektronischer Fragebogen erfasste die Bewertung der Häufigkeit (0 = selten, 6 = häufig) und individuellen Wichtigkeit (0 = unwichtig, 6 = wichtig) der einzelnen Aufgaben innerhalb der Prozesse des Instituts. Der Aufwand (0 = kein Aufwand, 6 = sehr hoch) für die zur Erfüllung der Aufgaben notwendige Informationsrecherche wurden von den ca. 100 Befragten bei den Aufgaben 1 und 3 als ‚sehr hoch', bei der Aufgabe 2 als ‚hoch' und bei der Aufgabe 4 als ‚gering' eingestuft.*

[8] Beispiele sind durch kursive Schreibweise und einen vertikalen Balken auf der linken Seite gekennzeichnet.

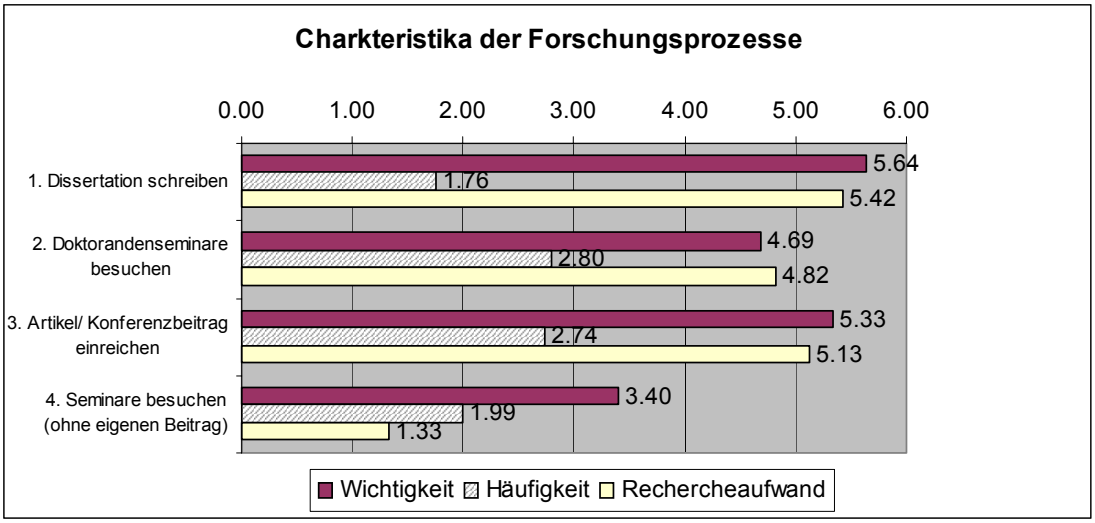


Abbildung 5-4: Retrieval-Aspekte in den Forschungsprozessen von IWI-HSG

Nach der ersten Identifikation möglicher Prozess-Kandidaten für ein Retrieval-Projekt erfolgt eine grobe Beurteilung des Handlungsbedarfs. Die Projektinitiatoren konsolidieren die Ergebnisse der Anforderungsanalyse in einer Liste der Prozessanforderungen (s. [Gebert 2003, 144], [Schulze 2000, 146]). Sie umfasst:

- die Bezeichnung der Aufgabe sowie der betroffenen Anwendergruppe,
- den mit der Erfüllung der Aufgabe verbundenen Informationsbedarf und
- die derzeitigen Schwachstellen bei der Informationsbeschaffung.

Aufgabe	Anwendergruppe	Informationsbedarf	Derzeitige Schwachstellen
Artikel / Konferenzbeitrag einreichen	Professoren, Projektleiter, Assistenten	Prozessbeschreibung Thematische Übersicht über vorhandene Publikationen Personenbezogene Übersicht über vorhandene Publikationen Relevante Konferenzen Ansprechpartner, die diesen Prozess durchlaufen haben	Informationen über mehrere Applikationen verteilt Informationen teilweise unstrukturiert Datensammlungen unbekannt

Tabelle 5-4: Auszug aus den Prozessanforderungen bei IWI-HSG

5.3.3. Applikationen erfassen

Suchmaschinen integrieren die Inhalte vorhandener Applikationen und stellen den Anwendern diese zur Bearbeitung informationsintensiver Aufgaben im Portal zur Verfügung. Eine Methode zur Einführung von Information Retrieval muss daher das im Unternehmen existierende Informationssystem einbeziehen. Die vollständige

Neuentwicklung (‚Greenfield Approach') ist in der Praxis nicht durchsetzbar [Riehm 1997, 192].

Zur Beurteilung der Integrationsfähigkeit existierender Anwendungen erfordert die Potenzialanalyse:

- eine Übersicht der für die Durchführung der zuvor erfassten Prozesse genutzten Applikationen und Datensammlungen,
- ihre Zusammenhänge und Wechselwirkungen sowie
- eine erste Beschreibung der Anwendungscharakteristika, z.B. Ausgabeformate, Strukturinformationen oder Zugriffsbeschränkungen.

Die Projektinitiatoren greifen dazu auf im Unternehmen vorhandene Unterlagen zu, z.B. IT-Landkarten, Applikationslisten und Applikationsbeschreibungen oder erheben diese Informationen in Zusammenhang mit den Fachabteilungen (vgl. [IMG 1997b], [Kaiser 2000, 178], [Alt et al. 2002, 80]). In der Applikationsbeschreibung (IST) halten sie die Charakteristika jeder zu integrierenden Anwendung fest (s. [Kaiser 2000, 175 und 186], [Thiesse 2001, 104]):

- die Bezeichnung und kurze Beschreibung der Anwendung,
- die Benennung der enthaltenen Informationsobjekte und
- Angaben zu Ausgabeformaten, Strukturinformationen (vorhanden, teilweise vorhanden, nicht vorhanden, nicht vorgesehen) sowie ihre Qualität (kontrollierte Begriffe, freie Begriffswahl) und Zugriffsbeschränkungen.

Anwendung	Informationsobjekt	Ausgabeformat	Strukturinformation	Zugriffs-beschränkung
Abstract Datenbank	Literatureintrag	Lotus Notes Dokument	Vorhanden; freie Begriffswahl	Professoren, Projektleiter, Assistenten
Endnote (Referenzverwaltung)	Referenz	Proprietär	Teilweise vorhanden; freie Begriffswahl	Professoren, Projektleiter, Assistenten
Verzeichnis ‚Publikationen' auf Dateiserver	Publikation	Microsoft Word, Adobe PDF	Teilweise vorhanden; freie Begriffswahl	Professoren, Projektleiter, Assistenten
Recherchedatenbanken der HSG	Dokument	HTML	Vorgesehen; kontrollierte Begriffe	Professoren, Projektleiter, Assistenten
Organisationshandbuch	Prozess-beschreibung	Lotus Notes Dokument	Vorhanden; kontrollierte Begriffe	Professoren, Projektleiter, Assistenten

Abbildung 5-5: Applikationsbeschreibung (IST) bei IWI-HSG

5.3.4. Handlungsoptionen entwickeln

Der letzte Schritt der Potenzialanalyse beantwortet die Frage, ob und wie die in den Prozessen vorhanden Potenziale mit einer IR-Lösung adressiert werden können. Die Projektinitiatoren ordnen dazu die in den Applikationen vorhandenen Informationsobjekte den in den Aufgaben entstehenden Informationsbedarfen zu.

Tabelle 5-5 zeigt die Zuordnung von Informationsobjekten zum Informationsbedarf für die Aufgabe ‚Artikel / Konferenzbeitrag einreichen':

	Relevante Applikation und Informationsobjekt				
Informationsbedarf	Abstract Datenbank	Endnote	Verzeichnis ‚Publikationen' auf Dateiserver	Recherche-DBs der HSG	Organisations-handbuch
1) Thematische Übersicht über vorhandene Publikationen	Literatur-eintrag	Referenz	Publikation	Dokument	
2) Personenbezogene Übersicht über vorhandene Publikationen	Literatur-eintrag	Referenz	Publikation	Dokument	
3) Prozessbeschreibung					Prozess-beschreibung
4) Relevante Konferenzen	Literatur-eintrag				
5) Ansprechpartner, die diesen Prozess durchlaufen haben			Publikation		

Legende: enthalten nicht enthalten

Tabelle 5-5: Zuordnung von Informationsobjekten zum Informationsbedarf

Die zur umfassenden Bearbeitung der Aufgabe benötigten Informationen sind in diesem Fall über fünf verschiedene Anwendungen verteilt. Aus Sicht einer übergreifenden Portalsuche entstehen folgende Herausforderungen und Handlungsoptionen:

- *A) Verteilte Inhalte erschliessen*. Die Inhaltsfindung bei 1) und 2) erfordert von den Mitarbeitern separate Recherchen in vorhandenen Anwendungen. Relevante Informationsobjekte zur Deckung desselben Informationsbedarfs sind in mehreren Applikationen verteilt vorhanden. Hier schafft bereits die Bereitstellung einer *Volltextsuche* Abhilfe.
- *B) Multiple Sichten abbilden*. Die Darstellung von 1) und 2) erfordert unterschiedliche Sichten (Thema, Person) auf dieselben Inhalte einer Applikation. Diese können im Vergleich zu A) nur mit einer *attributbasierten Suche oder taxonomieorientierten Navigation* umgesetzt werden.

- *C) Heterogene Informationsobjekte integrieren*. In den Anwendungen existiert eine Vielzahl unterschiedlicher Informationsobjekte (Literatureintrag, Referenz, Publikation, Dokument, Prozessbeschreibung). Sollen diese bei der Informationsrecherche unterschieden werden, so ist wie bei B) eine *attributbasierte Suche* oder *taxonomieorientierte Navigation* als Handlungsoption empfehlenswert.
- *D) Strukturinformationen erschliessen*. Die Informationsbedarfe 4) und 5) lassen sich nur über Details der Informationsobjekte (Konferenzen, Ansprechpartner) beantworten. Wie bei B) und C) ist auch hier eine strukturierte Erschliessung der Informationsobjekte über *attributbasierte Suche* oder *taxonomieorientierte Navigation* erforderlich.

In Anlehnung an [Thiesse 2001, 120] umfassen die Handlungsoption für ‚Volltextsuche', ‚attributbasierte Suche' und ‚taxonomieorientierte Navigation' Aussagen über Ziele, Motivation und Einflussfaktoren eines möglichen Retrieval-Projekts. Die Projektinitiatoren dokumentieren die ‚Handlungsoptionen' in Form einer Liste. Sie halten die Potenziale zur Qualitätssteigerung, Zeitersparnis sowie Kostenreduktion für jede Handlungsoption fest.

5.4. Strategieplanung

Kurzbeschreibung der Technik: In der Strategieplanung legen die Projektverantwortlichen die zukünftige Prozessunterstützung durch die Information-Retrieval-Lösung fest. Auf Basis der erfassten Applikationen skizzieren sie die zu implementierende Systemarchitektur. Die Projektverantwortlichen bewerten die Nutzenaspekte und schätzen die Kosten der einzuführenden Lösung ab.

Ziele: Die Technik Strategieplanung verfolgt die Ziele:

- ein Leistungsverzeichnis für Retrieval zu definieren,
- die zukünftige Systemarchitektur zu skizzieren,
- den Anpassungsbedarf bei vorhandenen Anwendungen zu formulieren und
- die durch Information Retrieval erreichbaren Nutzeneffekte sowie resultierenden Kosten zu beurteilen.

Voraussetzungen (Input): Die Strategieplanung setzt Prozessanforderungen, IS-Architektur (IST), Applikationsbeschreibungen (IST) und Handlungsoptionen der Technik Potenzialanalyse (s. Abschnitt 5.3) voraus.

Ergebnisse (Output): Die Ergebnisse der Technik Strategieplanung sind ein Retrieval-Leistungsverzeichnis, eine Architekturübersicht (SOLL), Schnittstellen- und Applikationsbeschreibungen (SOLL) sowie eine Kosten-/Nutzenbetrachtung.

Rollen: Die Technik Strategieplanung wird unter Berücksichtigung der in Tabelle 5-6 beschriebenen Rollen angewendet.

Stufe	Initiatoren	Anwender (Mitarbeiter, Kunden)
Entscheider	Führungsverantwortlicher, Prozessverantwortlicher	Fachbereichsverantwortlicher
Verantwortlicher	Projektleiter	Fachbereichsvertreter, Applikationsverantwortlicher
Unterstützer	Projektteam, Anwendungsentwickler, Controller, Systemarchitekt, Softwarehersteller	Fachbereichsmitarbeiter

Tabelle 5-6: Rollen der Technik Strategieplanung

5.4.1. Leistungen festlegen

Die Leistung von Information Retrieval ist die Identifikation von Informationen. Sie wird von Prozessen in Form von Volltextsuche, attributbasierter Suche, taxonomieorientierter Navigation oder einer Kombination konsumiert, um informationszentrierte Aufgaben auszuführen (vgl. Grundlagen in Abschnitt 2.3.3, Metamodell in Abschnitt 3.4 und Praxisbeispiele in Kapitel 4).

Mit den Retrieval-Leistungen legen die Projektinitiatoren jedoch nicht nur Art und Umfang der Prozessunterstützung fest. Je umfassender die von den Anwendern gewünschten Leistungen sind, desto höher sind auch die organisatorischen und technischen Anforderungen an die später zu implementierende Lösung.

- A) Beschränkt sich die Retrieval-Leistung auf eine *Volltextsuche*, so müssen die Verantwortlichen Prozesse für Index- und Integrationsmanagement spezifizieren (s. Abbildung 3-15, S. 58). Diese umfassen den Aufbau und die Pflege von Suchindizes, die Anbindung bestehender Anwendungen, den Abgleich existierender Zugriffsbeschränkungen oder die Bereitstellung von Benutzerschnittstellen. Bei einer Volltextsuche wird kein kontrolliertes Vokabular zur Verfügung gestellt. Aufgaben des Terminologiemanagements wie Benennung, Strukturierung und Bereitstellung von Begriffen sowie die manuelle Klassifikation von Inhalten entfallen daher (s. Abbildung 3-17, S. 63). Die Anforderungen an die anzubindenden Applikationen sind vergleichsweise gering.

  Eine anwendungsübergreifende Suche über alle im Portal bereitgestellten Inhalte, z.B. zu einem bestimmten Produkt, war bei Zumtobel Staff nicht möglich. Die Projektauftraggeber entschieden daher, die Qualität der den Kunden über das Portal angebotenen Dienstleistungen durch eine Volltextsuche über den Produktkatalog und die redaktionellen Inhalte des Content-Management-Systems (CMS) zu steigern.

  Das Integrationsmanagement umfasst bei Zumtobel Staff die einmalige Bereitstellung von Anwendungs- und Benutzerschnittstelle durch einen externen Implementierer. Sowohl das CMS der Firma Reddot als auch der bei Zumtobel Staff auf Basis einer SQL-Datenbank entwickelte Produktkatalog sind über eine HTTP-Schnittstelle in das SAP Enterprise Portal integriert. Das Portal stellt für die Volltextsuche ein Feld für die Eingabe einer oder mehrerer Suchbegriffe zur Verfügung.

  Das Indexmanagement ist auf die einmalige Erstellung rollenspezifischer Indizes beschränkt. Diese sind so angelegt, dass sie existierende Zugriffsbeschränkungen des Produktkatalogs und des CMS abbilden. Die tägliche Aktualisierung der

Indizes erfolgt automatisch. Die Suchmaschine SAP TREX erfasst dabei alle in Produktkatalog und CMS vorhandenen Informationsobjekte, z.B. Produktbeschreibungen, Installationsanweisungen oder Veranstaltungshinweise.

- B) Fordern die Anwender eine *attributbasierte Suche*, so müssen die Retrieval-Verantwortlichen zusätzlich zu A) für den Aufbau und die Pflege der Retrieval-Taxonomie auch Prozesse des Terminologiemanagements implementieren (s. Abbildung 3-15, S. 58). Die im Portal bereitgestellten Applikationen müssen strukturierbare Informationsobjekte beinhalten, die von den Mitarbeitern entsprechend klassifiziert und von der Suchmaschine ausgelesen werden können (s. Abbildung 3-17, S. 63).

Die IR-Lösung bei IWI-HSG adressiert eine Qualitätssteigerung der Arbeitsergebnisse durch den schnelleren Zugriff auf Informationen. Die gezielte Identifikation von Dokumenten war mit einer reinen Volltextsuche im Bestand von über 14'000 Informationsobjekten jedoch nicht gewährleistet. Die Projektverantwortlichen entschieden daher, Attribute für die Klassifikation von Dokumenten und die anschliessende Suche nach diesen Inhalten zu implementieren.

Die Begriffe für die Verschlagwortung von Dokumenten wurden durch Vertreter der Kompetenzzentren erarbeitet (zur Organisation des Instituts siehe Abschnitt 4.4). Die neu geschaffene Rolle des ‚Taxonomy-Managers' ist für die Erstellung, Überarbeitung oder Archivierung relevanter Begriffe des jeweiligen Kompetenzzentrums zuständig.

Zentrale, institutsweite Vorgaben existieren lediglich in Form von Handlungsempfehlungen. Diese wurden vom Projektteam ausgearbeitet und an die Kompetenzzentren kommuniziert. Sie geben anhand einiger weniger Regeln einen Rahmen zur Selektion von Begriffen vor, z.B. ‚Bevorzugt Begriffe im Singular erfassen' oder ‚Bevorzugt deutschsprachige Begriffe verwenden'.

- C) Die *taxonomieorientierte Navigation* erfordert, analog zur B), sowohl ein Integrations-, Index- als auch Terminologiemanagement. Hierbei liegt der Fokus der Begriffstrukturierung auf der Festlegung hierarchischer Beziehungen zwischen den Begriffen, die im Portal die Navigationshierarchie abbilden.

Abbildung 5-6 zeigt die Designstudie einer taxonomieorientierten Navigation bei IWI-HSG. Die Begriffe im linken Bereich bilden die Navigationshierarchie ab. Die oberste Navigationsebene ist im Beispiel durch die Kurzbezeichnungen der vier

Lehrstühle des Instituts abgebildet. Diese organisatorischen Merkmale werden auf der zweiten Ebene durch die Auswahl eines Kompetenzzentrums, z.B. CC CKM, verfeinert. Unterhalb dieser Navigationsebene verzweigt der Benutzer in die vier Tätigkeitsschwerpunkte eines Kompetenzzentrums. Für die Themenschwerpunkte ‚CC-Arbeit', ‚Organisation', ‚Forschung' und ‚Lehre' werden im rechten Bereich die zugehörigen Dokumente angezeigt.

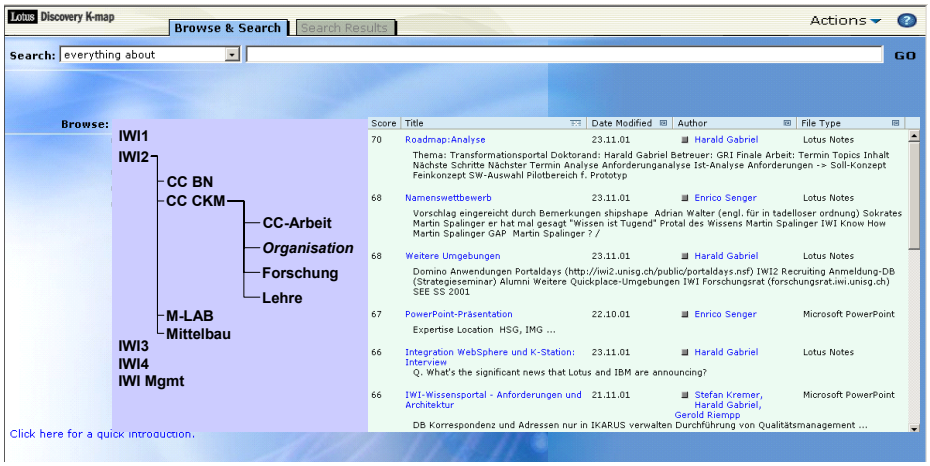


Abbildung 5-6: Taxonomieorientierte Navigation mit dem Lotus Discovery Server (Designstudie; Evaluation bei IWI-HSG)

Die Dokumente sind den Navigationseinträgen nicht fest zugeordnet sondern werden aus dem Index der Suchmaschine anhand von Regeln generiert. Diese Zuordnungsregeln sind bei einer taxonomieorientierten Navigation im Rahmen des Terminologiemanagements organisatorisch zu erarbeiten. Die Abbildung einer bestehenden Organisationsstruktur ist vergleichsweise einfach. Hier konnte das Projektteam auf bestehende Unterlagen wie Organigramme zurückgreifen. Von einer institutsweiten, themenbasierten Navigation sahen die Verantwortlichen bei IWI-HSG aufgrund der Vielzahl der Begriffe, ihrer Beziehungen und dem damit verbundenen Koordinationsaufwand ab.

Die Projektmitarbeiter haben die Anwendergruppen, deren Aufgaben und aufgabenbezogenen Informationsbedürfnisse sowie die benötigten Content-Applikationen in der Technik Potenzialanalyse (s. Abschnitt 5.3) bereits erfasst, Potenziale identifiziert und Handlungsoptionen entwickelt. Sie konsolidieren die Ergebnisse und ergänzen sie um die geforderte Retrieval-Leistung.

1.1 Ziele

1.1.1 Projektziele und -nutzen

Gemäss dem „Infrastructure Projektauftrag" (Version 1 vom 04.03.03) verfolgt das Teilprojekt Suche das Ziel, durch Bereitstellung einer übergreifenden Suchfunktion (auch Suchservice, oder Suche) im Customer Portal zum systematischen Ausbau der Portal Infrastruktur beizutragen.

Durch die Einbindung eines Suchservices werden Anwender (Kunden) des Customer Portals in allen Prozessen beim Auffinden von gewünschten Informationen unterstützt. Durch eine übergreifende Suche werden Suchzeiten verringert, die Service-Qualität erhöht und somit die Kundenprozesse besser unterstützt.

1.1.2 Systemziele

Abgeleitet aus den Projektzielen ergeben sich die im Folgenden skizzierten Systemziele:

- Das Teilprojekt Search umfasst in der Ausbaustufe P2.3 die Implementierung einer Volltextsuche im Customer Portal (auf Basis des SAP Enterprise Portal 5.0 SP5, SAP EP) über die Inhalte des noch zu implementierenden Content Management Sytems (CMS) und der bereits im Produktivbetrieb befindlichen Produktdatenbank (PDB, auf Basis von MS SQL).
- Die Suchanfragen werden zunächst nur über ein Einzeleingabefeld im Portal gestartet (sogenannte Simple Search). Die Implementierung einer erweiterten Suche (Advanced Search) wird im Rahmen dieser Projektphase nicht angestrebt (out of scope), ist jedoch in den Anforderungen, dem Detailkonzept (Pflichtenheft) und dem Softwareauswahlverfahren bereits zu berücksichtigen.

Abbildung 5-7: Festlegung von Retrieval-Leistungen im Projektauftrag von Zumtobel Staff (Markierungen einzelner Textteile vom Autor vorgenommen)

5.4.2. Architekturübersicht erstellen

Information-Retrieval-Lösungen können umfangreich und komplex sein. Sie setzen sich aus vielen Komponenten zusammen und erfordern ein planerisches Vorgehen. Ein Mittel dieser Planung ist die Informationssystemarchitektur. Die IS-Architektur beschreibt die logische Struktur eines Informationssystems. Sie zeigt die notwendigen Softwarekomponenten sowie die zwischen ihnen bestehenden Integrationsbeziehungen auf (vgl. [Thiesse 2001, 173], [Alt et al. 2002, 80]). Die Projektverantwortlichen erstellen die IS-Architektur mit folgenden Zielsetzungen:

- *Überblick.* Die Architekturübersicht gibt allen Projektbeteiligten einen Überblick über das zukünftige System. Sie bildet bereits bestehende Komponenten, zu erstellende neue Architekturbestandteile sowie existierende und zukünftige Schnittstellen ab.

- *Gestaltungselemente*. Der Schwerpunkt eines Retrieval-Projekts kann auf einzelne Komponenten der IS-Architektur fokussieren, ohne Abhängigkeiten zu anderen Architekturelementen zu vernachlässigen. Dadurch wird beispielsweise der Vergleich von Anforderungen an einzelne Gestaltungselemente mit den Leistungsmerkmalen von unterschiedlichen Softwareanbietern erleichtert.

- *Abhängigkeiten*. An den Verbindungspunkten einzelner Komponenten der Architektur entstehen Schnittstellen. Die Projektmitarbeiter ziehen diese heran, um das Realisierungsvorhaben im Unternehmen abzustimmen, technische Integrationsbeziehung zu detaillieren oder die Anpassungs- und Betriebskosten zu bewerten.

Eine applikatorische Architekturübersicht entsteht unter Berücksichtigung von Applikationen, Datensammlungen, Funktionen und Integrationsbeziehungen (s. [Österle 1995, 289], [Kaiser 2000, 201]). Die vorliegende Arbeit verwendet dazu die Bestandteile des Metamodells für Information Retrieval (s. Abschnitt 3.4):

- Applikationen: Portal, Suchmaschine, Content-Applikation(en)
- Datensammlungen: Informationsobjekte, Struktur, Benutzerverwaltung
- Funktionen: Indizierung, Suche, Navigation, Anzeige, Klassifikation, Integration, Authentifizierung, Autorisierung, Präsentation
- Integrationszusammenhänge: z.B. Suchmaschine ‚führt' Indizierung aus

Das Projektteam hat die bestehende Applikationslandschaft in Form der Architektur (IST) und den Applikationsbeschreibungen (IST) der Technik Potenzialanalyse (s. Abschnitt 5.3.3) erhoben. Für die Erstellung der Architekturübersicht greifen die Projektinitiatoren diese Ergebnisse auf. Sie ergänzen die Architektur (IST) um fehlende Applikationen und Datensammlungen, ordnen den Komponenten Funktionen zu und skizzieren die Integrationszusammenhänge (vgl. [Puschmann 2003, 109], [Blessing 2001, 164], [Kaiser 2000, 185-195]). Bereits im Unternehmen vorhandene Komponenten benennen die Projektmitarbeiter in der grafischen Architekturübersicht namentlich, z.B. mit Produktbezeichnungen.

> *Im Retrieval-Projekt bei IWI-HSG sollten Inhalte der Applikationen ‚Datenbank Server', ‚Extranet Server' und ‚File Server' von der Suchmaschine indiziert und den Benutzern über ein Portal bereitgestellt werden. Die Architekturübersicht sah vor, dass Suchmaschine, Portal und ‚File Server' dabei auf eine gemeinsame Benutzerverwaltung zugreifen (‚Domain Controller'). Durch die Synchronisation von Benutzer- und Gruppeninformationen sollte diese mit der Benutzerverwaltung von ‚Datenbank Server' und ‚Extranet Server', dem ‚Domino Directory', abgeglichen werden.*

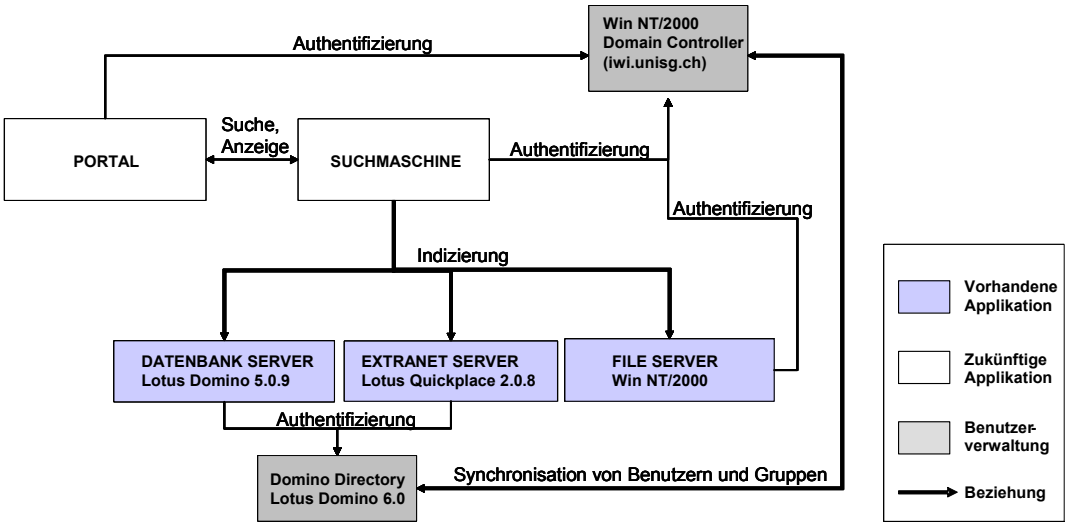


Abbildung 5-8: Architekturübersicht in der Planungsphase bei IWI-HSG

Die vorliegende Architekturübersicht diente beispielsweise zur Koordination der Entwickler einzelner Komponenten, zur Absprache mit der zentralen Informatik der Universität oder zur Definition von Schnittstellen. Sie war darüber hinaus eine wesentliche Grundlage für die Gespräche mit Softwareherstellern.

Die Verantwortlichen leiten zusammen mit Applikationsverantwortlichen, Anwendungsentwicklern und Systemarchitekten aus der IS-Architektur (SOLL) folgende Ergebnisse ab:

- *Applikationsliste (SOLL)*. Die Applikationsliste (SOLL) führt alle Anwendungen auf, aus denen das neue Informationssystem besteht.
- *Applikationsbeschreibungen (SOLL)*. Die Applikationsbeschreibung (SOLL) dokumentiert die zu realisierenden Anwendungen durch Angabe ihrer Eigenschaften, z.B. Bezeichnung der Anwendung, Funktionsumfang, enthaltene Informationsobjekte, Berechtigungskonzept etc.
- *Schnittstellenbeschreibungen (SOLL)*. Die Schnittstellenbeschreibung (SOLL) dokumentiert in Form von Regeln oder Eigenschaften die zukünftige Komponente einer Applikation, die dem Austausch von Informationen mit anderen Applikationen dient. Sie enthält beispielsweise Angaben zu Ein- und Ausgabeformaten oder der eingesetzten Technologie.

Aus dem SOLL-IST-Vergleich von Applikationslisten, Applikationsbeschreibungen und Schnittstellenbeschreibungen resultiert der Anpassungs- und Erweiterungsbedarf des bestehenden Informationssystems.

5.4.3. Kosten und Nutzen bewerten

Die Projektverantwortlichen haben gemeinsam mit Anwendervertretern die Anforderungen an eine Lösung für Information Retrieval in Portalen definiert, indem sie Nutzenpotenziale identifiziert, Handlungsoptionen entwickelt, Leistungen festgelegt, die zukünftige IS-Architektur abgeleitet und Anpassungsbedarfe an bestehende Anwendungen formuliert haben. Ein Pflichtenheft konsolidiert diese Informationen, indem es die fachlichen Anforderungen an die zu implementierende Lösung umfasst [vgl. Balzert 2001, 113].

Je unterschiedlicher die zu integrierenden Content-Applikationen in Bezug auf Dokumentenformate, Ausgabeschnittstellen, Struktur- und Klassifikationsmerkmale sowie Zugriffsbeschränkungen sind, desto höher sind die zu erwartenden Kosten einer Information-Retrieval-Lösung. Darüber hinaus steigen die voraussichtlichen Aufwände, wenn zusätzlich zur Volltextsuche auch Retrieval-Leistungen wie eine attributbasierte Suche oder eine taxonomieorientierte Navigation von den Anwendern gefordert werden.

Zur Bewertung der bei Aufbau und Betrieb einer Information-Retrieval-Lösung entstehenden Kosten ist der Vergleich von Realisierungsvarianten notwendig. Da kommerzielle Suchmaschinensoftware in unterschiedlichen Preissegmenten eine Vielzahl von Funktionen bietet (s. [Andrews 2003b], [Beall/Hodges 2002], [Ramos 2002]), ist in den meisten Unternehmen eine vollständige Neuprogrammierung nicht sinnvoll. Das Projektteam fokussiert daher im Softwareauswahlverfahren (Toolevaluation) auf die Anschaffung einer am Markt verfügbaren Portalsuchmaschine.

Die Toolevaluation verfolgt das Ziel, aus mehreren konkurrierenden Softwareprodukten das den Anforderungen bestmöglich entsprechende auszuwählen [Alt et al. 2002, 106]. Dabei orientiert sich das Projektteam an einem Auswahlprozess für Standardsoftware [vgl. Stahlknecht/Hasenkamp 2002, 304]):

- *Angebotseinholung*. Die Angebotseinholung basiert auf dem Pflichtenheft. Sie umfasst neben den zu realisierenden Retrieval-Leistungen und der bestehenden Infrastruktur auch die Angabe augenblicklicher bzw. zu erwartender Mengengerüste (Datenvolumen, Zugriffe etc.). Die Ergebnisse ‚Retrieval-Leistungen', ‚IS-Architektur (SOLL)', ‚Applikationsbeschreibungen (SOLL)' und ‚Schnittstellenbeschreibungen (SOLL)' liefern den Projektverantwortlichen die in Tabelle 5-7 aufgeführten Kriteriengruppen zur Strukturierung der Angebotseinholung:

Kriteriengruppe	Beschreibung
Indizierung & Crawler	Indizierung unterschiedlicher Formate (z.B. HTML, MS Office Dokumente, Adobe PDF oder ZIP-Dateien) und Quellen (z.B. HTTP, Lotus Notes oder Documentum) unter Angabe des jeweiligen Datenvolumens
Architektur & Integration	Funktionen zur Integration (z.B. zur Einbindung in ein bestehendes Portal, zur Berücksichtigung eines führenden Benutzerverzeichnisses oder Mechanismen für Single Sign On)
Content Organization	Funktionen für Aufbau und Pflege einer Taxonomie (z.B. automatische Generierung einer Navigationshierarchie, Extraktion von Schlüsselwörtern oder Navigation anhand von Inhaltsmetadaten)
Suchanfrage	Möglichkeiten für Informationsidentifikation (z.B. Volltextsuche, attributbasierte Suche oder taxonomieorientierte Navigation sowie Anpassung der Benutzerschnittstelle)
Suchergebnisse	Art und Umfang der Anzeige von gesuchten Inhalten (z.B. Hervorheben des Suchbegriffs im gefundenen Dokument, Gruppierung von Suchergebnissen oder Anpassung der Suchergebnisliste)
Performance & Skalierbarkeit	Leistungsmerkmale (z.B. Antwortzeit für Suchanfragen, maximale Anzahl gleichzeitiger Anwender oder Lastverteilung durch verteilte Systemkomponenten)
Client Types	Verfügbare Anwendungen für Endbenutzer und Administratoren (z.B. Web-Browser, Windows-Programme oder Java Applikationen)
Verfügbarkeit & Support	z.B. Produkthistorie, geplante Releases, Referenzinstallationen oder Telefonhotline
Preis & Lizenzierung	Lizenzmodell (z.B. pro Server, pro Benutzer etc.) und damit verbunden Anschaffungs- und Wartungskosten (z.B. jährliche Updates)

Tabelle 5-7: Kriteriengruppen für die Softwareauswahl

- *Grobbewertung der Angebote*. Nach dem Eingang der Angebote bewertet das Projektteam diese grob. Zur Verkürzung des Auswahlprozesses empfiehlt es sich, K.O.-Kriterien zu definieren, deren Eintreten zum Aussortieren des untersuchten Softwarepakets führt [vgl. Büren et al. 2003]. Kommerzielle Suchsoftware bietet heute vielfältige Schnittstellen und Erweiterungsmöglichkeiten, mit denen sich unterschiedlichste Anforderungen realisieren lassen. Die Bewertung sollte daher weiterhin unterscheiden, ob Produkte die Anforderungen ohne weitere Programmierung (‚Out-of-the-box'), über individuelle Projektlösungen oder gar nicht erfüllen. Sowohl der Systementwurf der Praxisbeispiele (s. Abschnitt 4.5.5) als auch empirische Erhebungen bei [Senger 2004, 185] zeigen, dass die Auswahl von Komponenten eines Herstellers dem Zusammenführen heterogener Systeme verschiedener Anbieter (‚Best-of-Breed') in Integrationsprojekten vorzuziehen ist.

- *Feinbewertung und Endauswahl*. Die Grobbewertung hat die in Betracht kommenden Angebote bereits reduziert. Sie erlaubt eine frühe Abschätzung der produktbezogenen Kosten in dieser Projektphase. Eine Feinbewertung der unterschiedlichen Suchmaschinen erfolgt erst nach Durchführung der Prozessplanung und Abschluss des Systementwurfs. Der vorliegende Methodenvorschlag greift dies in Abschnitt 5.7.1 auf. Die dort enthaltene Tabelle 5-28 listet ausführlich die in den Praxisfällen verwendeten Kriterien auf.

Das Projektteam fasst die Grobbewertung der Softwareprodukte im Produktvergleich zusammen.

Anbieter	TREX (SAP EP 5.0 SP5)	Verity Ultrasearch (ehemals Inktomi Enterprise Search)	ht://Dig	Lucene 1.2
Charakter	Enterprise Portalkomponente für Indizierung und Suche	Kommerzielle Suchmaschine	OpenSource	OpenSource
Hersteller / Organisation	SAP AG	Verity (Deutschland) GmbH	www.htdig.org	http://jakarta.apache.org/lucene/docs/index.html
Service AT	SAP AT	Verity (Deutschland) GmbH	keiner	keiner
Lizensierungsmodell	In SAP Enterprise Portal enthalten	abhängig von der Anzahl der indizierten Dokumente und der Portalzielgruppe, z.B. 25.000 EURO für insgesamt 25.000 Dokumente aus PDB und CMS, Nutzung ausschliesslich im Customer Portal	GNU Open License	OpenSource
Lizenzpreis ZS (Schätzung)	Keine zusätzlichen Lizenzkosten	ca. 25.000 EURO	entfällt	entfällt
Wartungskosten an Hersteller p.a. (Schätzung)	Keine zusätzlichen Wartungskosten	20% vom Lizenzpreis	entfällt	entfällt
Technische Plattform	Windows 2000 Advanced Server (SP 3)	Windows NT 4.0 und höher (intelbasiert), Solaris 2.6, Solaris 7 und Solaris 8 auf einer Sun Sparc, Microsoft Windows 2000, Red Hat Linux 6.0 oder höher	Release 3.1.6 (Februar 2002), Unix, C bzw. C++ Compiler notwendig, CGI	Unabhängig, Servlet Engine notwendig
Implementierungsausfwand Suche und Portalintegration	anonymisiert			
Marktposition DACH			entfällt	entfällt
Anforderungserfüllung Summary				
Indizierung & Crawler	Web-Crawler für Indizierung von HTML-Content vorhanden (auch aus dynamischen Quellen); weitere Formate: XML, text, RTF, MS Office, PDF etc.	Web-Crawler für Indizierung von HTML-Content vorhanden (auch aus dynamischen Quellen); HTML, XML, text, RTF, MS Office, Adobe Acrobat PDF, PostScript, FrameMaker, Lotus SmartSuite, WordPerfect und über 100 andere Datenformate	Web-Crawler für Indizierung von HTML-Content vorhanden (auch dynamiche Quellen), auch zugriffgeschützte Seiten (Username + Password), kein PDF	Web-Crawler für Indizierung von HTML-Content vorhanden ("spindle" oder "LARM"); PDF etc. separat zu implementieren, Basiskomponenten jedoch bereits als OpenSource vorhanden (PDF Parser, z.B. PDFBox)
Architektur & Integration	Integraler Bestandteil von SAP EP; TREX nutzt somit auch SSO- und LDAP-Lösungen des Portals	Kann laut Anbieter in SAP EP eingebunden werden (Aufwand von 2-3 Wochen), SSO bisher ausschliesslich mit Netegrity Siteminder	Keine Informationen zu Integrationsmöglichkeiten in SAP EP verfügbar	Integration in SAP EP kann programmiert werden, da Java APIs vorhanden
Suchfähigkeiten	Einfache oder mehrfache Suchbegriffe, logische Operatoren, Phrasensuche, Attributbasierte Suche, Suche nach Wortstämmen, Wildcards	Einfache oder mehrfache Suchbegriffe, logische Operatoren, Phrasensuche, Attributbasierte Suche, Suche nach Wortstämmen, Suche nach Meta-Daten, Wildcards	Einfache oder mehrfache Suchbegriffe, logische Operatoren, Word Stemming, keine Phrasensuche, keine Attributsuche, keine Wildcards	Einfache oder mehrfache Suchbegriffe, logische Operatoren, Word Stemming (limitierte auf Deutsch und Englisch), Attributsuche, Wildcards, Fuzzy- und Proximitysuche, keine Phrasensuche
Anzeigefähigkeiten	Suchanfragemaske und -ergebnismaske können mit EP 5 SP5 kaum modifiziert werden (ggfls. Java iView Programmierung); Verbesserte Anpassungsmöglichkeiten mit "Flexible UI" im EP 6	Suchanfragemaske und -ergebnismasken können modifiziert werden	Suchanfragemaske und -ergebnismasken können modifiziert werden (HTML-Templates)	Suchanfragemaske kann modifiziert werden (Servlet, JSP)
Performance & Skalierbarkeit	gut skalierbar		gering skalierbar, nur für kleinere Seiten geeignet	Sehr hohe Performance; Loadbalancing ggfls. auf Application Server realisierbar)
Client Types	Web-Browser	Web-Browser	Web-Browser	Web-Browser, Java Client
Support	SAP	Verity	OpenSource-Community	OpenSource-Community

Abbildung 5-9: Grobbewertung Produktvergleich bei Zumtobel Staff (Implementierungsaufwand anonymisiert)

Für die Aufstellung der in einem Information-Retrieval-Projekt entstehenden Kosten verwenden die Projektverantwortlichen bevorzugt ein im jeweiligen Unternehmen bereits bewährtes Kalkulationsschema. Ist ein solches bisher nicht vorhanden, empfiehlt sich die Einteilung nach Kostenarten, da diese als Teilmengen der Gesamtkosten eine ressourcenorientierte Sicht nach der Art der verbrauchten Güter und Dienstleistungen ermöglichen [s. Britzelmaier 1999, 104]. Tabelle 5-8 zeigt exemplarisch die Einteilung in Personal- und Sachkosten, die neben den einmalig auftretenden Kosten auch laufende Kosten aufschlüsselt.

	Einmalig	Laufend
Personalkosten	Entwicklungskosten EDV-Bereich • Analyse • Konzeption • Implementierung • Test und Integration • Dokumentation Entwicklungskosten Fachabteilungen • Analyse • Konzeption Entwicklungskosten Fremdfirmen Einführungskosten • Schulung • Organisatorische Umstellung …	Bedienungskosten Wartungskosten • Programmpflege • Datensicherung • Hardware Externe Dienstleistungen …
Sachkosten	Fremdsoftware • Kauf / kumulierte Nutzungskosten • Anpassung Hardware Material Räume …	Betriebskosten • Hardware • Netzwerk Software-Lizenzen Externer Service …

Tabelle 5-8: Kostenartenschema (in Anlehnung an [Kargl 2000])

Die Kosten- und Nutzenbetrachtungen der vier Praxisbeispiele (s. Kapitel 4) geben einen Überblick über Umfang und Aufteilung der Kosten. Während die laufenden Kosten der in den Unternehmen untersuchten Information-Retrieval-Lösungen in allgemeinen Infrastrukturkosten enthalten sind und nicht weiter aufgeschlüsselt werden, erlaubt die Betrachtung der einmaligen Kosten folgende Aussagen:

- Die *Entwicklungskosten* (EDV-Bereich, Fachabteilung und Fremdfirmen) entsprachen in den Praxisfällen einem Äquivalent von 50 bis 150 Personentagen.
- *Analyse- und Konzeptionsphasen* verusachten mit ca. 60% bis 75% den grössten Anteil der Entwicklungskosten.
- Die Projektverantwortlichen griffen zur Implementierung auf bereits vom Unternehmen *lizensierte Software* zu. (z.B. SAP Enterprise Portal, Opentext Livelink oder Lotus Notes). Eine anteilige Verrechnung der Sachkosten für die Fremdsoftware erfolgte nicht.
- In zwei Praxisfällen musste zum Betrieb der Suchmaschine *zusätzliche Hardware* in Form von Servern in der Grössenordnung zwischen 15'000 und 30'000 EUR angeschafft werden.

Können Projektverantwortliche zur Umsetzung einer Information-Retrieval-Lösung nicht auf bereits im Unternehmen vorhandene Software zugreifen, so müssen sie zusätzliche Sachkosten berücksichtigen. Bei der Nutzung von OpenSource-Suchmaschinen (z.B. ht://Dig und Lucene) sind diese zu vernachlässigen. Erfordert die Umsetzung der Anforderungen jedoch den Kauf einer kommerziellen Suchmaschine (z.B. Verity und Autonomy), so sind Anschaffungskosten von 250'000 EUR und mehr sowie laufende Kosten zwischen 15% und 20% der Anschaffungskosten zu berücksichtigen.

Während die Kostenarten im oben skizzierten Schema vergleichsweise leicht erfassbar sind, ist der Nutzen von Retrieval-Lösungen nur schwer zu beziffern. Traditionelle Investitionsrechnungsverfahren bewerten den Nutzen einer Investition ausschliesslich auf Basis des finanziellen Erfolges, z.B. über die Gesamtkapitalrentabilität (‚Return-on-Investment', ROI). Diese Finanzkennzahl setzt den erwirtschafteten Gewinn einer Investition ins Verhältnis zum betriebsnotwendigen Vermögen [s. Thommen 2002, 634]. Der eigentliche ‚Gewinn' der im Rahmen von Information Retrieval in Portalen identifizierbaren Informationsobjekte wird jedoch erst durch ihren Nutzen im Geschäftsprozess definiert (s. [Bach 2003, 130], [Applehans et al. 1998, 45], [Blessing 2001, 122]). Die rein quantitative Bewertung einer IR-Lösung ist für einen Investitionsentscheid daher weder angebracht noch ausreichend.

Die Technik Potenzialanalyse mit ihren Zielen, kritischen Erfolgsfaktoren sowie Mess- und Führungsgrössen in Bezug auf Qualität, Zeit und Kosten (s. Abschnitt 5.3.1) bietet den Projektbeteiligten ein qualitatives und quantitatives Bewertungsraster zur Abschätzung von Nutzenpotenzialen. [Riempp 2003, 247] weist in diesem Zusammenhang auf folgende Besonderheiten hin:

- *Qualitätssteigerung*. Die systematische Erfassung und Bewertung von Qualitätssteigerungen ist nur möglich, wenn ein Referenzwert vorliegt, der den Vergleich von Qualitätsmerkmalen mit und ohne IR-Lösung erlaubt. Anhaltspunkte liefern hier Berichte über Qualitätssteigerungen von Prozessleistungen durch den Einsatz von Portalsuchmaschinen, Anwender- und Kundenzufriedenheitsumfragen und das Benchmarking gegen Mitbewerber.

- *Kostenreduktion*. Die finanziellen Einsparungen durch IR-Lösungen sind schwer ermittelbar, da die Kosten für den ineffizienten Umgang mit Informationen nicht in der betrieblichen Kostenrechnung ausgewiesen sind. Sie sind in Kosten für Arbeitszeit, Gemeinkosten, Gewährleistungen etc. ‚versteckt'. [Kraus 2003] und [Davenport/Prusak 1998]) verweisen auf eine Bewertung von falschen Entscheidungen, Doppelspurigkeiten oder entgangenen Möglichkeiten aufgrund von unzureichenden Informationsidentifikationsmechanismen.

- *Zeitersparnis*: Verbrauchte Zeiträume für Informationssuche vor und nach der Einführung von Information Retrieval können durch Befragung von Kunden oder Zeiterfassung von Mitarbeitern erhoben werden. Die Bewertung der Zeitersparnis setzt jedoch voraus, dass Mitarbeiter die verfügbare Zeit für andere Tätigkeiten ergebniswirksam einsetzen.

Wichtiger als eine umfassende Investitionsrechnung sind für ein Information-Retrieval-Projekt somit eine ausreichende Transparenz der mit der Lösung verbundenen Kosten sowie die Festlegung von Mess- und Führungsgrössen der adressierten Ziele. Nur diese Grössen können die Projektverantwortlichen beeinflussen und beurteilen. Tabelle 5-9 formuliert Beispiele für die Ableitung von Nutzenpotenzialen und zugehörigen Mess- und Führungsgrössen.

Ziel	Nutzenpotenzial	Messgrösse	Führungsgrösse
Qualitätssteigerung	Höhere Kundenzufriedenheit durch Bereitstellung von Suchfunktionen	Bewertung der Suchfunktionen bei der jährlichen Kundenzufriedenheitsumfrage	Bewertung mit ‚gut' oder ‚sehr gut' bei mehr als 70% der Rückläufer
Kostenreduktion	Keine Doppelarbeiten durch Erhöhung der Transparenz über bestehende Inhalte	Anzahl der Fälle, bei denen die Identifikation von Informationen mit Hilfe einer Suchmaschine Doppelarbeiten im Unternehmen vermieden hat	Mehr als 10 dokumentierte Fälle innerhalb von 6 Monaten
Zeitersparnis	Kürzere Recherchezeiten der Mitarbeiter bei der Angebotserstellung	Zeitdauer für Informationsrecherche zur Erstellung eines Angebots	Informationsrecherche weniger als 20% des Gesamtaufwands für die Erstellung eines Angebots

Tabelle 5-9: Beispiel für Ziele, Nutzenpotenziale, Mess- und Führungsgrössen eines Information-Retrieval-Projekts

Die Projektverantwortlichen stellen zum Abschluss der Strategieplanung folgende Unterlagen zusammen:

- Ergebnisübersicht des groben Produktvergleichs,
- Anforderungen und Kriterienkatalog für die Feinbewertung und Endauswahl einer Suchmaschinensoftware,
- Kalkulationsschema für Anschaffungs- und Betriebskosten und
- Ziele und Nutzenpotenziale, kritische Erfolgsfaktoren sowie Mess- und Führungsgrössen.

Sie legen diese den Entscheidern vor und erwirken einen Beschluss zur weiteren Durchführung des Information-Retrieval-Projekts.

5.5. Prozessplanung

Kurzbeschreibung der Technik: In der Prozessplanung definieren die Verantwortlichen Abläufe, mit denen die Retrieval-Leistungen der Strategieplanung organisatorisch umgesetzt werden sollen. Sie skizzieren den zukünftigen Retrieval-Prozess und leiten Aktivitäten für den Aufbau und die Pflege einer Retrieval-Taxonomie, zur Integration von Applikationen und zur Entwicklung eines Benutzer- und Berechtigungsmodells für die Portalsuche ab. Die Ergebnisse schliessen die fachliche Konzeption ab.

Ziele: Die Technik Prozessplanung hat die Zielsetzung:

- den Ablauf des Retrieval-Prozesses zu skizzieren,
- Attribute und Werte der Retrieval-Taxonomie zu erarbeiten,
- unterschiedliche Benutzer- und Berechtigungsverwaltungen der beteiligten Applikationen abzugleichen,
- Anwendungsintegration existierender Anwendungen zu detaillieren und
- Prozesse zu definieren, die Aufbau und Betrieb von Retrieval-Taxonomie, Benutzer- und Berechtigungspflege und Anwendungsintegration regeln.

Voraussetzungen (Input): Die Technik Prozessplanung benötigt das Retrieval-Leistungsverzeichnis, die Architekturübersicht (SOLL) und die Applikations- und Schnittstellenbeschreibungen (SOLL) aus der Technik Strategieplanung (s. Abschnitt 5.4).

Ergebnisse (Output): Die Retrieval-Taxonomie, eine Zuordnungstabelle (‚User Mapping') für die Benutzerverwaltung, der Abgleich von Benutzergruppen, verfeinerte Applikations- und Schnittstellenbeschreibungen (SOLL) sowie zugehörige Aufgabenkettendiagramme mit Erstell- und Pflegeprozessen bilden die Ergebnisse der Technik Prozessplanung.

Rollen: Die Technik Prozessplanung wird unter Berücksichtigung der in Tabelle 5-10 beschriebenen Rollen angewendet.

Stufe	Initiatoren	Anwender (Mitarbeiter, Kunden)
Entscheider	Führungsverantwortlicher, Prozessverantwortlicher	Fachbereichsverantwortlicher
Verantwortlicher	Projektleiter	Fachbereichsvertreter, Applikationsverantwortlicher
Unterstützer	Projektteam, Review-Team, Terminologe, Search Manager, Portalarchitekt, Retrieval-Architekt, Anwendungsentwickler, Administrator, Sicherheits-planer, Directory Manager, Suchmaschinenentwickler	Fachbereichsmitarbeiter, Autor

Tabelle 5-10: Rollen der Technik Prozessplanung

5.5.1. Retrieval-Prozess definieren

Zielsetzung von Information Retrieval in Portalen ist die verbesserte Informationsfindung prozessrelevanter Inhalte. IR-Lösungen stellen den Anwendern dazu Retrieval-Leistungen wie Volltextsuche, attributbasierte Suche oder taxonomieorientierte Navigation über Portalsuchmaschinen bereit.

Durch die Definition des Retrieval-Prozesses legen die Projektverantwortlichen die zukünftige Interaktion der beteiligten Rollen und Anwendungen fest. Die Technik Prozessplanung definiert dazu:

- *Autor* oder *Content Manager* als Ersteller von Informationsobjekten,
- *Search Manager* für Einrichtung, Betrieb und Optimierung von Suchmaschinen,
- *Nutzer* bzw. *Anwendervertreter* als Endanwender von Retrieval-Funktionen,
- *Content-Applikation* als inhaltsführende Anwendung,
- *Suchmaschine* als Anwendung für die Indizierung von Inhalten sowie die Bereitstellung von Such- und Navigationsfunktionen.

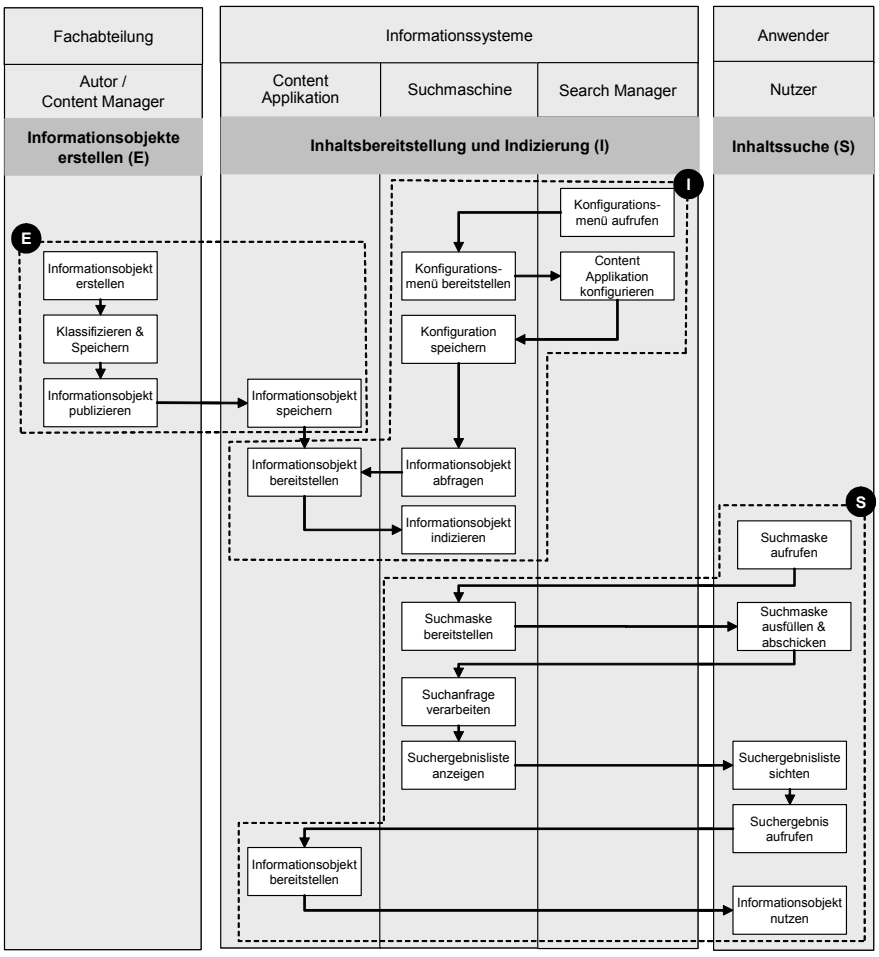


Abbildung 5-10: Aufgabenkettendiagramm für den Retrieval-Prozess

Die Suchmaschine fungiert als Bindeglied zwischen den Anwendern als Informationsnachfragern und den indizierten Applikationen als Informationsanbietern:

- *Suchmaske*. Die Suchmaske dient als Eingabeschnittstelle, in der die Anwender ihren Informationsbedarf formulieren. Sie bildet den Dialog zwischen Nutzer und IR-System ab. Die Qualität der Suchergebnisse hängt davon ab, wie gut die Repräsentation des Informationsbedarfs vom Anwender in der Suchmaske formuliert werden kann.
- *Suchergebnisliste*. Die Suchergebnisanzeige bildet die Ausgabeschnittstelle, anhand derer die Anwender die Relevanz der für sie gefundenen Inhalte in einem ersten Schritt beurteilen. Die Suchergebnisanzeige muss für eine schnelle Beurteilung den Kontext der gefundenen Informationsobjekte möglichst umfassend wiedergeben, z.B. durch die Angabe von Autoren und Stichworten.
- *Klassifikation*. Bereits die Klassifikation bei der Erstellung von Inhalten kann sich an den Informationsbedürfnissen der Nutzer orientieren. Dabei erhöht eine gemeinsame Klassifikation durch eine begrenzte Anzahl von Begriffen die Genauigkeit bei der Formulierung von Suchanfragen.

Die Projektverantwortlichen halten den Retrieval-Prozess als ein Teilergebnis der Prozessplanung fest.

5.5.2. Retrieval-Taxonomie bereitstellen

Eine Taxonomie dient zur Klassifikation von Begriffen und ihren Beziehungen in einem (meist hierarchischen) Ordnungssystem (s. [Bailey 1994], [Felber/Budin 1989]. Sie setzt sich aus den relevanten Begriffen eines Bereichs, die durch eine oder mehrere baum- oder auch netzwerkförmige Strukturen geordnet ist, zusammen. Eine Taxonomie zeigt die Beziehungen zwischen den Begriffen, bietet eine kontrollierte Benennung und eignet sich als Klassifikationsschema [vgl. Riempp 2003, 188].

Durch die Festlegung von Begriffen ermöglicht eine Retrieval-Taxonomie, elektronische Dokumente in einem Portal einheitlich zu klassifizieren und durch Navigation und Suche leichter auffindbar zu machen (s. [Wayne 1991], [Moriarty 1990], [Rich 1992]). Die Abbildungen der Benutzerschnittstellen der Praxisfälle bei IMG (s. S. 80), Winterthur Versicherungen (s. S. 90) und IWI-HSG (s. S. 100) illustrieren die Umsetzung der Retrieval-Taxonomie im Portal.

[Hellmuth 1997, 22] bezeichnet die Tätigkeiten zur Bildung und Pflege von Taxonomien als Terminologiemanagement. Terminologiemanagement erstellt für einen abgegrenzten Bereich (z.B. Organisation, Branche, Fachgebiet) ein begriffliches Ordnungssystem, pflegt es kontinuierlich und bietet es möglichst komfortabel zur Nutzung bei Suche oder Navigation an [s. Kremer/Riempp 2001, 59].

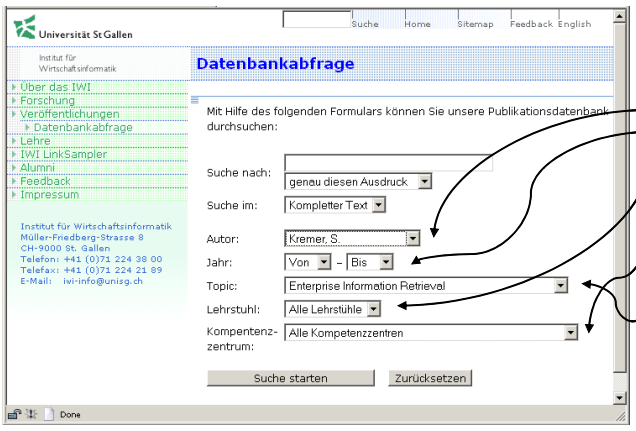
Tabelle 5-11 zeigt exemplarisch die Anzahl von Attributen und Werten der untersuchten Praxisfälle in Abhängigkeit von der benötigten Retrieval-Leistung. Der Aufwand für die erstmalige Erstellung einer Retrieval-Taxonomie ist nicht zu unterschätzen. In den Praxisfällen bei IMG, Winterthur Versicherungen und IWI-HSG waren dazu jeweils mehrere Workshops mit verschiedenen Fachvertretern und anschliessende Abstimmungsrunden notwendig. Nicht selten vergeht dabei eine ganze Stunde für die Einigung auf einen Wert [s. Kremer/Riempp 2001, 59].

Fallstudie	Anzahl Attribute	Anzahl Werte je Attribut	Retrieval-Leistung
Zumtobel Staff	0	0	Volltextsuche
The Information Management Group	13	11 – 37	Volltextsuche, attributbasierte Suche, taxonomieorientierte Navigation
Winterthur Versicherungen	11	6 – 18	Volltextsuche, attributbasierte Suche, taxonomieorientierte Navigation
Institut für Wirtschaftsinformatik	5	4 – 25	Volltextsuche, attributbasierte Suche

Tabelle 5-11: Mengengerüst der Retrieval-Taxonomien aus den Praxisfällen

Erfahrungen aus Industrie und Wissenschaft zeigen, dass diese Aufgaben von rein technologisch getriebenen Ansätzen wie Softwaresystemen zur automatischen Klassifikation von Inhalten derzeit nicht zufriedenstellend gelöst werden (s. [Andrews 2003a], [Warzecha 2001], [Hagen 2000], [Verity 2003], [Christ 2002, 133]).

Fordert das Retrieval-Leistungsverzeichnis der Technik Strategieplanung (s. Abschnitt 5.4.1) die Bereitstellung von attributbasierter Suche oder taxonomieorientierter Navigation, so müssen die Projektverantwortlichen eine Retrieval-Taxonomie erarbeiten und bereitstellen. Diese besteht aus Attributen und Werten, die dem Benutzer bei einer attributbasierten Suche zur Verfügung stehen oder dem begrifflichen Ordnungssystem, das ein Portalanwender bei einer taxonomieorientierten Navigation verwendet.



Attribut	Wertebereich	Bemerkung
Author	- offen -	alle Mitarbeiter des Instituts
Year	- offen -	Jahreszahlen
Chair	IWI 1, IWI 2, IWI 3, IWI 4	Lehrstuhl
Competence Center	• IWI 1: BPM, DWS, BAI, EMBE, DW2, AIM, RMA • IWI 2: BN, I-NET, TCC, PRO, ECC, PRO-BM, BKM, eBN, PSI, CRM, CRIM, iBN, BN2, M-Lab, TI, CKM • IWI 3: E-Learning, Knowledge Networks, Learning Center, Knowledge Source • IWI 4: CKM, C>K>P, IIM	Bezeichnung aller 32 abgeschlossenen oder laufenden Kompetenzzentren; aufgeführt nach Lehrstühlen
Topic	Ausschnitt für IWI 4 / C>K>P: Administration, Enterprise Application Integration, Enterprise Information Retrieval, IT-Marketing, M-Commerce, M-Payment, Methode, Portal, Privacy, Real-time Management, Sicherheit, System, Terminologiemanagement, Web Services, Wirtschaftsinformatik (allgemein)	Thematische Schwerpunkte aller 32 abgeschlossenen oder laufenden Kompetenzzentren; insgesamt 450 Begriffe

Abbildung 5-11: Attribute und Werte der Retrieval-Taxonomie in der attributbasierten Suche bei IWI-HSG

Für den dauerhaften Einsatz einer Information Retrieval-Lösung reicht es nicht aus, eine Taxonomie einmalig erstellt zu haben. Die Entstehung, Veränderung und im Laufe der Zeit wechselnde Verwendung von Begriffen stellt eine weitere Herausforderung dar. Die Prozessplanung berücksichtigt daher sowohl den initialen Aufbau als auch die Pflege der Retrieval-Taxonomie (s. [Kremer et al. 2003a], [Kremer/Riempp 2001]).

Die bereits beim Retrieval-Prozess eingeführten organisatorischen Rollen (s. Abschnitt 5.5.1) werden ergänzt bzw. erweitert um:

- *Autor* oder *Content Manager* als Ersteller von Informationsobjekten und Pflegeverantwortlicher der Taxonomie aus Sicht der Klassifikation von Informationsobjekten,
- *Applikationsverantwortlicher* mit Datenhoheit über eine zu integrierende Applikation und verantwortlich für die Pflege der Taxonomie aus Sicht der Anwendung,
- *Retrieval-Projektleiter* mit Verantwortung für die Durchführung, Koordination und Abstimmung des Retrieval-Projekts,
- *Review-Team* zur Prüfung und Freigabe der Retrieval-Taxonomie,
- *Terminologe* als Verantwortlicher für Aufbau und Pflege der Taxonomie,
- *Search Manager* für Einrichtung, Betrieb und Optimierung von Suchmaschinen,
- *Portal-Architekt* als Gestalter von Struktur, Layout und Navigation von Portalen,
- *Nutzer* bzw. *Anwendervertreter* als Endanwender von Retrieval-Funktionen und als repräsentativer Vertreter im Retrieval-Projekt.

Bei der *Entwicklung einer Retrieval-Taxonomie* werden die in einer Organisation vorhandenen Informationen in Gruppen (taxonomische Dimensionen) eingeteilt. So ist, je nach Unternehmen, eine Gliederung nach Produkten, Dienstleistungen, Prozessen, Kunden oder anderen Kriterien möglich. Bei einem Kreditinstitut wäre in einer Dienstleistungs-Dimension z.B. das Attribut ‚Vermögensbildung' mit den möglichen Werten ‚Fondsparen', ‚Lebensversicherung' und ‚Bausparen' denkbar. Das Projektteam erfasst in Workshops mit entsprechenden Fachvertretern die taxonomischen Dimensionen, Attribute und Werte, da diese in ihrem jeweiligen Bereich die grösste Kenntnis der relevanten Begriffe und Begriffsbeziehungen aufweisen. Um Fehler im begrifflichen Ordnungssystem zu vermeiden, prüft der Terminologe die vorgeschlagenen Attribute und Werte. Er behebt dabei auch mögliche Begriffsdefekte, z.B. durch die Behandlung von Synonymen [s. Brenner 1988].

[Ortner 1997] unterscheidet fünf Arten möglicher Defekte und gibt Handlungsanweisungen zur Behebung (s. Tabelle 5-12).

Begriffsdefekt	Beschreibung	Handlungsanweisung
Synonym	Unterschiedliche Bezeichnungen mit identischer Bedeutung (z.B. Programm und Software)	Kontrolle durch Synonymwörterbuch
Homonym	Gleiche Bezeichnung mit unterschiedlicher Bedeutung (z.B. Programm als Software und Programm als politische Richtlinie)	Beseitigen durch eindeutige Definition
Äquipollenz	Die Bezeichnungen beschreiben ein Objekt aus unterschiedlichen Blickwinkeln (z.B. Lagerbestand als mengenmässige und Warenkonto als wertmässige Rechnung des Artikelbestands eines Unternehmens)	Aufdecken und bei zu großer Überlappung beseitigen
Vagheit	Mangelnde Abgrenzung von Begriffen (z.B. die beiden Begriffe Entwicklung und Programmierung)	Klären durch Definition und Kontextbezug

Tabelle 5-12: Begriffsdefekte nach Ortner

Das Review-Team prüft die konsolidierte Taxonomie und gibt diese frei. Den Abschluss des erstmaligen Aufbaus einer Taxonomie bildet die Bereitstellung der gewonnen Begriffe und Relationen für die jeweilige Zielgruppe, z.B. über eine Schlüsseldatenbank zur Klassifizierung von elektronischen Dokumenten. Search Manager, Portal-Architekt und Applikationsverantwortliche übernehmen die bereitgestellte Taxonomie in ihren Anwendungen, z.B. durch Anbindung ihrer Applikationen an die neu erstellte Schlüsseldatenbank oder durch einen Import des Klassifikationsgerüsts.

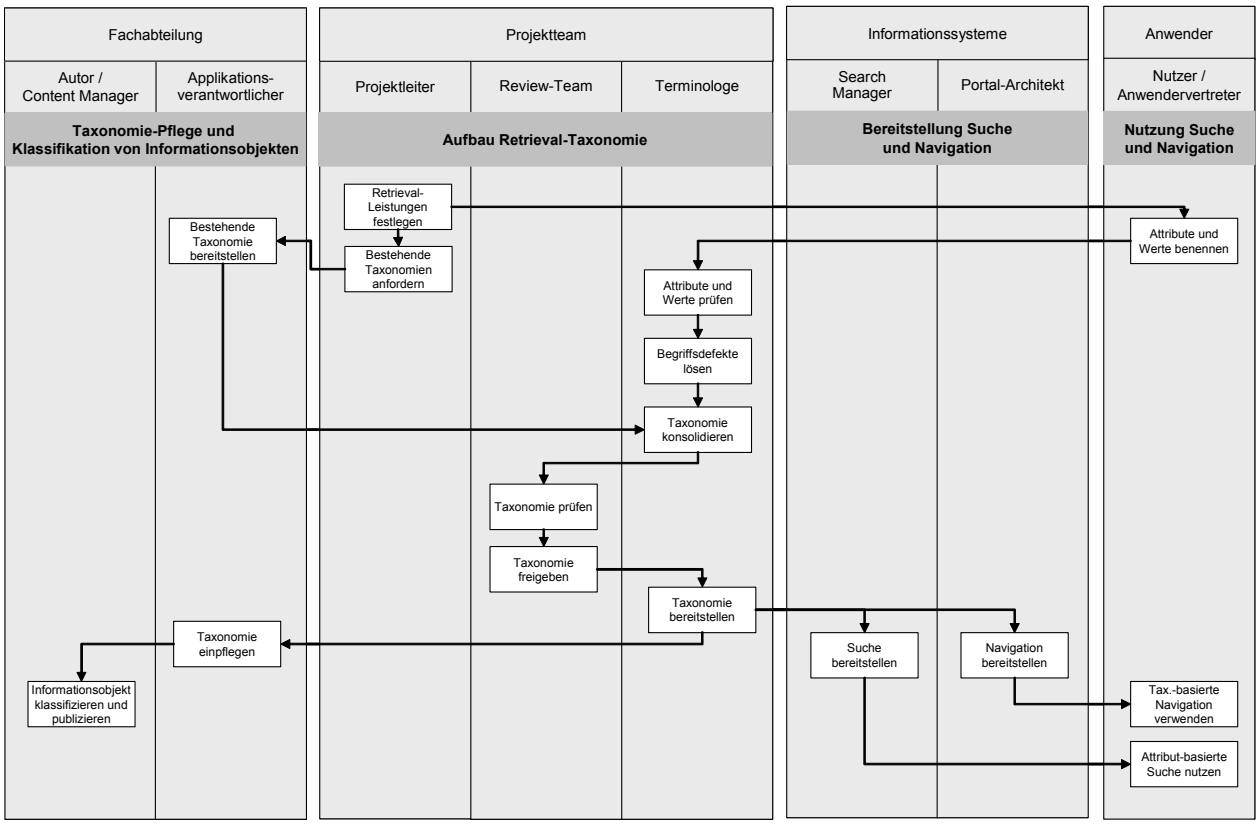


Abbildung 5-12: Aufgabenkettendiagramm für den Aufbau einer Retrieval-Taxonomie

Die Suche nach existierenden Branchen- oder Industriestandards erleichtert den Aufbau von Taxonomien. Normungsinstitute, Industrie- und Handelskammern oder auch grosse Unternehmen der jeweiligen Branche bieten eine gute Ausgangsbasis. Innerhalb eines Unternehmens können hilfreiche Informationen auch oft aus ERP-Systemen herangezogen werden. Diese bieten beispielsweise eine bewährte Klassifikation und Benennung von Kunden, Produkten, Regionen oder Organisationseinheiten.

Die *Pflege einer Retrieval-Taxonomie* geht von einem Änderungsbedarf im laufenden Betrieb aus. Dieser kann von den Autoren oder Nutzern der Informationsobjekte initiiert werden. Abbildung 5-13 zeigt die entsprechenden Rollen und Aufgaben.

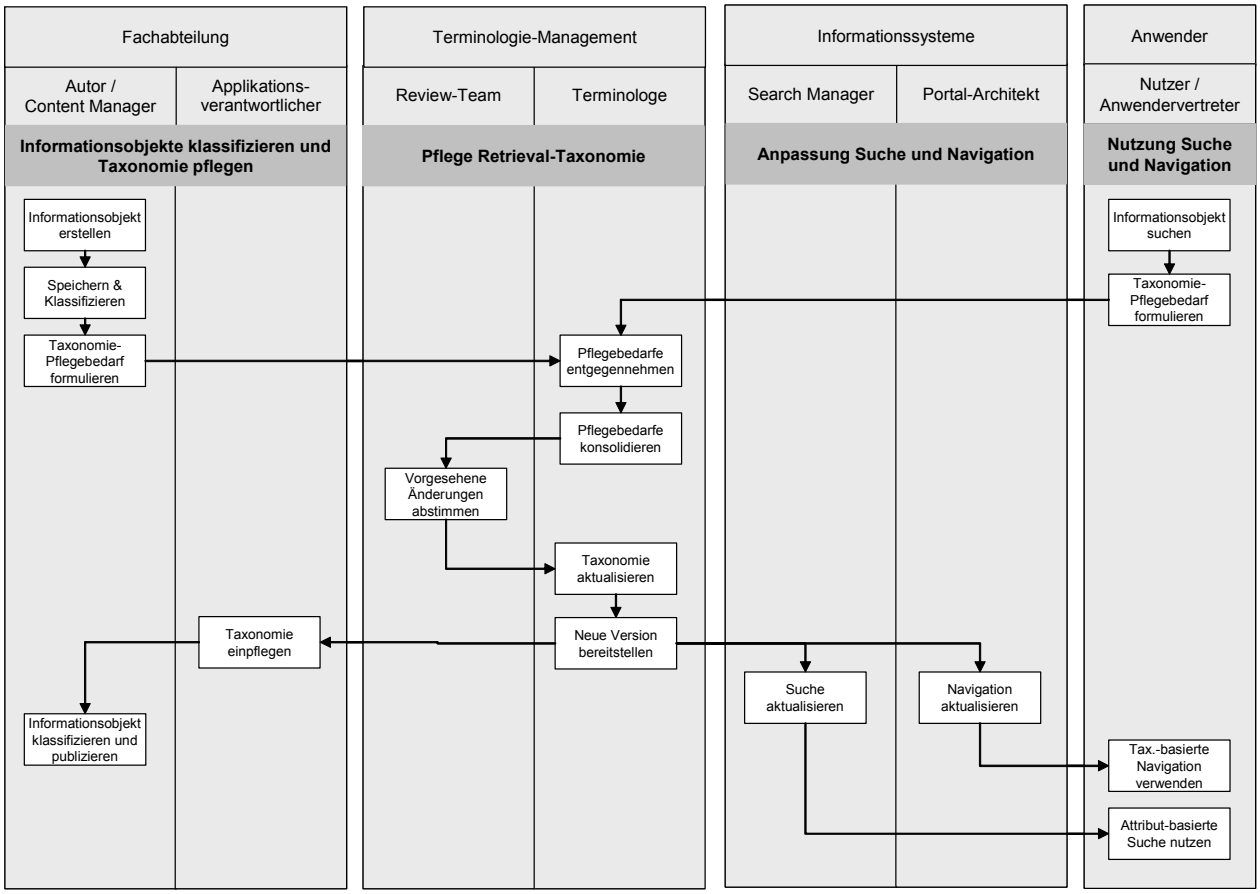


Abbildung 5-13: Aufgabenkettendiagramm für Pflege einer Retrieval-Taxonomie

> *Das Attribut ‚Topic' kennzeichnet bei IWI-HSG den thematischen Schwerpunkt eines Dokuments. Die Werte, z.B. ‚Portal' oder ‚Wissensmanagement', werden zur Klassifizierung und attributbasierten Suche verwendet. Zum Zeitpunkt der Erstellung dieser Arbeit verwaltet das Institut ca. 350 Topics in einer zentralen Datenbank (Parameter-Database), mit denen über 11'000 Informationsobjekte attributiert sind.*
>
> *Sowohl die Pflege der Topics als auch die Klassifizierung von Dokumenten erfolgt dezentral durch die Mitarbeiter der Kompetenzzentren. Pro Kompetenzzentrum*

> *nimmt ein Taxonomy-Manager die Erstellung, Änderung oder Deaktivierung von Topics in der Parameter-Database vor. Ein Mitarbeiter kann jederzeit per E-Mail einen Änderungsbedarf an seinen zuständigen Taxonomy-Manager kommunizieren.*
>
> *Die Änderung eines Topics durch einen Taxonomy-Manager führt nicht zu einer Re-Klassifizierung eines bestehenden Dokuments. Das Institut plant, der Suchmaschine über einen Export dieses kontrollierten Vokabulars die veränderten Topics als Synonyme zur Verfügung zu stellen. Eine Suchanfrage nach dem Topic ‚Wissensmanagement' stellt dem Anwender über einen Thesaurus dann auch Dokumente über die aktualisierte Suche zur Verfügung, die mit dem ursprünglichen Begriff ‚Wissensmanagement' klassifiziert worden sind.*

Die Projektverantwortlichen halten als Teilergebnisse der Prozessplanung den Retrieval-Prozess, die Retrieval-Taxonomie, sowie die für Aufbau und Pflege der Retrieval-Taxonomie notwendigen Prozesse fest.

5.5.3. Benutzer und Berechtigungen pflegen

In den meisten Unternehmen mit gewachsenen Infrastrukturen existiert eine Vielzahl an Applikationen mit unterschiedlichen Mechanismen für die Verwaltung von Benutzern und deren Berechtigungen. Information-Retrieval-Lösungen müssen diese Sicherheitsmechanismen von Portal, Suchmaschine und zu integrierenden Content-Applikationen organisatorisch und technologisch aufeinander abstimmen:

- *Autorisierung und Berechtigungsverwaltungen*. Bei der Autorisierung prüft eine Anwendung, ob ein Benutzer zur Durchführung einer bestimmten Aktion berechtigt ist (s. [Raepple 2001, 5] sowie Grundlagen in Abschnitt 2.2.2). Diese Prüfung, die so genannte Zugangskontrolle, basiert auf der Durchsetzung von Zugriffsrechten, die an eine Kombination aus Subjekt (z.B. Benutzer oder Applikation) und Objekt (z.B. Informationsobjekt, Applikation oder Gruppe von Objekten) gekoppelt sind (s. [Oppliger 1998, 55], [Ahuja 1996, 182]). In einer Suchergebnisliste dürfen beispielsweise ausschliesslich Informationsobjekte aus einer indizierten Anwendung aufgeführt werden, für die ein eindeutig identifizierter Benutzer auch beim direkten Zugriff auf die Applikation leseberechtigt ist. Dies setzt voraus, dass die Suchmaschine die Berechtigungsverwaltung bzw. Zugriffsrechte der Content-Applikation abbilden kann.
- *Authentifizierung und Benutzerverwaltungen*. Die Authentifizierung einer Anwendung identifiziert einen Benutzer, z.B. durch Eingabe von Benutzername und Passwort. Der Authentifizierungsmechanismus prüft deren Echtheit durch

Zugriff auf eine Benutzerverwaltung, die sämtliche Aktivitäten zum Anlegen, Ändern und Sperren von Benutzerkonten zusammenfassen [s. Puschmann 2003, 99]. Erfordern sowohl Portal, Suchmaschine als auch Content-Applikationen eine Authentifizierung und verfügen sie dazu über eigene Benutzerverwaltungen, so ist die Haltung und Pflege der Benutzerdaten redundant. Die Benutzer erwarten jedoch nach dem erstmaligen Login in ein Portal die automatisierte Authentifizierung gegenüber allen anderen Anwendungen (so genanntes Single Sign On, SSO). Die Authentifizierung eines Benutzers beim Login in ein Portal muss daher von Suchmaschine und Content-Applikationen transparent übernommen werden.

In vielen Unternehmen obliegt die Pflege von Benutzerdaten und Vergabe von Berechtigungen den Administratoren der jeweiligen Content-Applikationen [s. Zwerger/Schneider 2002, 128-129]. Die vorliegende Methode schlägt zur Reduktion dieser redundanten Datenpflege die Realisierung eines zentralen Benutzerverzeichnisses vor. Auf dieses greifen die im Portal integrierten Applikationen entweder direkt zu oder sie synchronisieren relevante Daten. Mit dem Aufbau eines zentralen Benutzerverzeichnisses sind jedoch auch Herausforderungen verbunden:

Herausforderung 1: Wahl eines zentralen Benutzerverzeichnisses

Der internationale OSI-Standard X.500 dient seit 1988 als Plattform für verteilte Verzeichnisdienste. Das Ziel dabei ist die Vernetzung von Computersystemen verschiedener Hersteller über eine Integration bzw. Synchronisation verschiedener, heterogener Verzeichnisse, so dass deren Anwender hersteller- und systemübergreifend Kommunikationsdienste nutzen können [ISO 1988]. Als Quasi-Standard für die Übertragung von Benutzerdaten hat sich überdies das Lightweight Directory Access Protocol (LDAP) durchgesetzt [s. Zwerger/Schneider 2002, 129]. Der spätere Systementwurf zeigt die technische Umsetzung auf.

Herausforderung 2: Zuordnung elektronischer Identitäten

Bei einer dezentralen Verwaltung von Benutzerdaten existieren in den verschiedenen Applikationen für denselben Anwender meist verschiedene Benutzernamen, Passwörter oder digitale Zertifikate [s. Zwerger/Schneider 2002, 131]. Zentrale Benutzerverzeichnisse stellen Mechanismen bereit, diese unterschiedlichen elektronischen Identitäten eines Benutzers zusammenzuführen (vgl. [Ant 2003], [Waveset 2001]).

Im Rahmen der Prozessplanung müssen die Projektverantwortlichen die Applikationsbeschreibungen (SOLL) der Technik Strategieplanung ergänzen. Sie geben pro Applikation die möglichen Authentifizierungsmechanismen und die Bezeichnung des zugehörigen Benutzerverzeichnisses an. In einer Tabelle ‚User

Mapping' ordnen sie den zukünftigen Anwendern der Retrieval-Lösung die in den Anwendungen verwendeten Benutzernamen zu. Tabelle 5-13 zeigt dies exemplarisch für einige Mitarbeiter bei IWI-HSG.

Mitarbeiter	Benutzername in Anwendung …				
	Abstract Datenbank	End-note	‚Publikationen' auf Dateiserver	Recherche-datenbanken	Organisations-handbuch
Walter Brenner	walter brenner	-	Wbrenner	WBR	- digitales Zertifikat -
Henning Gebert	henning gebert	-	Hgebert	99634875	- digitales Zertifikat -
Lutz Kolbe	lutz kolbe	-	Lkolbe	LKO	- digitales Zertifikat -
Stefan Kremer	stefan kremer	-	Skremer	00646737	- digitales Zertifikat -
Ragnar Schierholz	ragnar schierholz	-	Rschierholz	02670867	- digitales Zertifikat -

Tabelle 5-13: ‚User Mapping' bei IWI-HSG

Die Erfassung der zugehörigen Passwörter ist nicht notwendig. Diese können von entsprechenden Softwarelösungen synchronisiert oder im Rahmen eines ‚Employee Self Service' von den Mitarbeitern beim erstmaligen Zugriff auf die IR-Lösung selbständig angegeben werden [s. Zwerger/Schneider 2002, 131].

Herausforderung 3: Abgleich von Benutzergruppen

In der betrieblichen Praxis werden Berechtigungen nicht über eine direkte Zuordnung von Anwendern zu Informationsobjekten vorgenommen. Benutzerverwaltungen fassen zur Vereinfachung der Administration einzelne Anwender zu Nutzergruppen zusammen, die dann entsprechende Berechtigungen aufweisen (s. [Balzert 2001, 907], [Kaiser 2000, 227], [Alt et al. 2002, 74]). Die Ergebnisse der Technik Potenzialanalyse (s. Abschnitt 5.4.2) liefert den Projektverantwortlichen für die Abbildung existierender Zugriffsrechte folgende Informationen:

- Die Applikationsbeschreibungen geben einen Überblick über Zugriffs-beschränkungen von Informationsobjekten auf bestimmte *Anwendergruppen*.
- In den Prozessanforderungen sind die Informationsbedürfnisse den *Retrieval-Benutzergruppen* zugeordnet.

Über das Merkmal ‚Informationsbedarf' ordnen die Projektmitarbeiter die in den Berechtigungsverwaltungen vorhandenen Anwendergruppen den Retrieval-Benutzergruppen zu. Abbildung 5-14 stellt die möglichen Fälle dar und gibt Handlungsanweisungen:

- A) Der Informationsbedarf einer Retrieval-Benutzergruppe (P1) deckt sich 1:1 mit der Zugriffsberechtigung einer Anwendergruppe (A1) auf relevante Informations-objekte (IO_{Ax}). Im zentralen Benutzerverzeichnis besteht die Retrieval-Gruppe aus der Anwendergruppe (P1=A1).

- B) Der Informationsbedarf einer Retrieval-Benutzergruppe (P2) kann durch Verbindung mehrere Anwendergruppen (A3, B2) und deren Zugriff auf relevante Informationsobjekte (IO_{Ax}, IO_{Ay}, IO_{Az}, IO_{Bx}, IO_{By}) gedeckt werden. Die Retrieval-Gruppe im Benutzerverzeichnis setzt sich aus den beiden Anwendergruppen zusammen (P2=A3+B2).

- C) Der Informationsbedarf einer Retrieval-Benutzergruppe (P3) kann weder durch direkte Zuordnung noch durch Kombination existierender Anwendergruppen gedeckt werden (IO_{Bz}). Diese Benutzergruppe (P3) erfordert die Einrichtung einer neuen Anwendergruppe (B4) in der betroffenen Applikation. In Folge ist sie wie in A) zu behandeln.

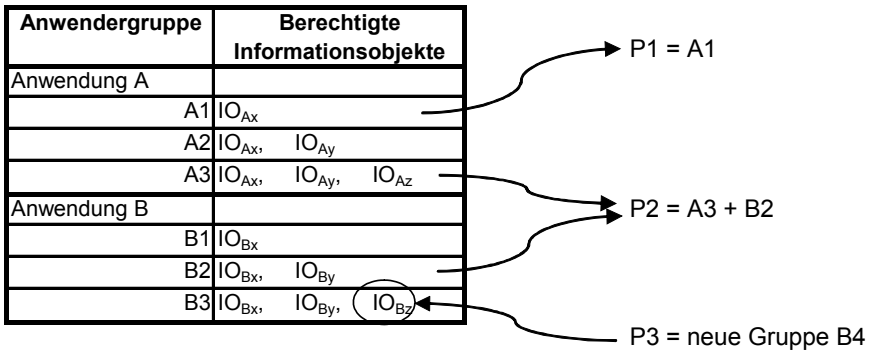


Abbildung 5-14: Benutzergruppenabgleich

Das Aufgabenkettendiagramm in Abbildung 5-15 stellt exemplarisch die Abläufe für den Aufbau einer zentralen Benutzerverwaltung dar. Die bereits bei der Retrieval-Taxonomie eingeführten organisatorischen Rollen (s. Abschnitt 5.5.2) werden ergänzt um:

- *Entwickler bzw. Administrator* für die Gestaltung und Verwaltung von Content-Applikation, Portal und Suchmaschine;

- *Sicherheitsplaner* als Gestalter für Benutzer- und Berechtigungspflege;

- *Directory Manager* als Verantwortlicher für die Bereitstellung und Pflege eines zentralen Benutzerverzeichnisses.

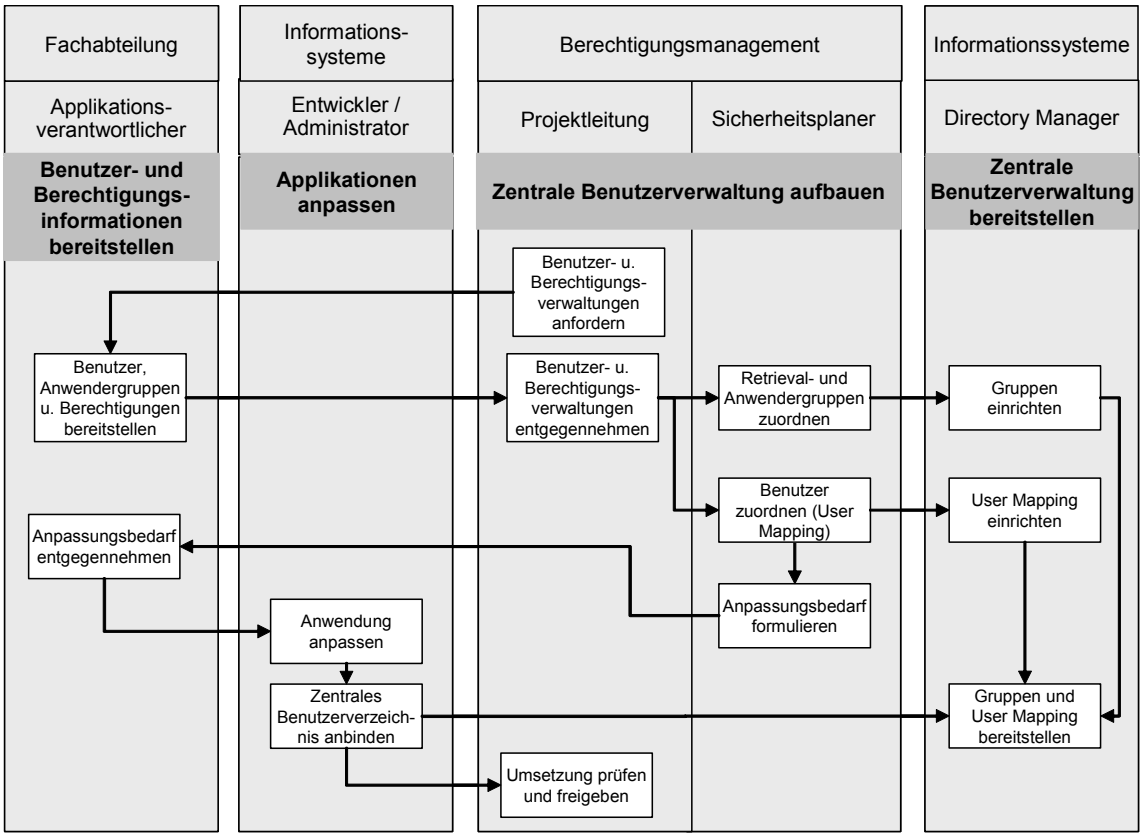


Abbildung 5-15: Aufbau einer zentralen Benutzerverwaltung

Nach dem initialen Aufbau der zentralen Benutzerverwaltung, z.B. im Rahmen eines Retrieval-Projekts, erfolgt die Kommunikation eines Änderungsbedarfs im laufenden Betrieb direkt an den Directory Manager.

Die Tabellen für das ‚User Mapping' und den Abgleich von Benutzergruppen sowie die zugehörigen Unterstützungsprozesse bilden die Teilergebnisse der vorliegenden Aktivität. Die technische Umsetzung der Ergebnisse der Prozessplanung stellt der Systementwurf in Abschnitt 5.6 dar.

5.5.4. Anwendungen integrieren

Eine Suchmaschine greift zur Bereitstellung ihrer Funktionen wie Indizierung, Klassifizierung, Informationssuche und -anzeige über eine Schnittstelle (Connector, Adapter, Gateway) auf die zu integrierenden Content-Applikationen zu (s. Abbildung 3-17, S. 63). Art und Umfang der Konnektoren müssen dabei mit den Ein- und Ausgabeschnittstellen der Applikation korrespondieren. Die Umsetzung der Retrieval-Taxonomie erfordert beispielsweise den Zugriff auf die Strukturinformationen einer Applikation. Auch ein Abgleich der Benutzergruppen kann nur erfolgen, wenn die Suchmaschine über einen Konnektor auf die Benutzerverwaltung der Anwendung zugreifen kann.

Die meisten Suchmaschinenhersteller bieten Konnektoren für den Zugriff über HTTP, Dateisysteme (Windows- und UNIX-Plattformen), Open Database Connectivity (ODBC als standardisierte Schnittstelle für den Zugriff auf relationale Datenbanksysteme) oder weitere proprietäre Schnittstellen z.B. für Lotus Notes, Microsoft Exchange, Documentum, Opentext oder Oracle. Erfüllen diese Schnittstellen alle bisherigen Anforderungen aus Strategie- und Prozessplanung, so ist eine Anpassung der Content-Applikation nicht notwendig. Die Praxisfälle zeigen jedoch, dass zur Anbindung von Anwendungen an eine Portalsuchmaschine immer mehr oder minder umfangreiche Modifikationen notwendig sind:

- Der Abgleich der Benutzergruppen in der Prozessplanung erforderte bei Zumtobel Staff die Einrichtung einer weiteren Anwendergruppe im Content-Management-System (CMS). Zur Umsetzung mussten die Applikationsverantwortlichen das Datenmodell des CMS anpassen.

- Das McB-Portal erlaubt, bewährte Strukturierungen einzelner Projekte wiederzuverwenden um beispielsweise neue Projekte auf Basis dieser SOLL-Struktur anzulegen. Auch für bereits angelegte Projekte kann eine Aktualisierung auf Basis eines so genannten ‚Blueprints' erfolgen. Dieser Mechanismus musste in der führenden Content-Applikation zusätzlich programmiert werden.

- Die Applikationsbeschreibungen der Potenzialanalyse wiesen sowohl bei IMG als auch bei IWI-HSG eine Vielzahl heterogener Informationsobjekte, Strukturen, Formate und Benutzerverwaltungen auf. Um die in der Strategieplanung geforderten Leistungen zu realisieren, entschieden die jeweiligen Projekt-verantwortlichen, die bisherigen Inhalte durch eine vollständige Neuentwicklung zu integrieren.

Auf Basis der Retrieval-Taxonomie und des Berechtigungskonzeptes entscheiden die Projektverantwortlichen über einen Anpassungsbedarf bei der Anwendungsintegration. Sie vergleichen dazu die Applikations- und Schnittstellenbeschreibungen (IST) der Technik Potenzialanalyse (s. Abschnitt 5.3.3) mit den Anforderungen der Retrieval-Taxonomie und des Berechtigungskonzepts der vorliegenden Prozessplanung. Die Schnittstelle einer Applikation muss die zur Umsetzung der Retrieval-Taxonomie erforderlichen Strukturinformationen und die zur Implementierung des Berechtigungskonzepts benötigten Zugriffsinformationen bereitstellen (s. Tabelle 5-14, linke Grafik). Diese müssen zusätzlich von einem Suchmaschinen-Connector abgebildet werden können. Stehen entsprechende Leistungsmerkmale bei der Applikationsschnittstelle oder beim Konnektor nicht zur Verfügung, muss eine Anpassung erfolgen. Berücksichtigen Applikationsschnittstelle oder Konnektor im ungünstigsten Fall (‚Worst Case') ausschliesslich den reinen Inhalt eines Informationsobjekts (s. Tabelle 5-14, rechte Grafik), so muss die bereitstellende Applikation über Anpassungen die Struktur- und Zugriffsinformationen dem

ursprünglichen Inhalt vor der Bereitstellung durch die Ausgabeschnittstelle hinzufügen (‚Komposition'). Die Applikationsverantwortlichen müssen in diesem Fall auch die Suchmaschine anpassen. Entsprechende Mechanismen zerlegen die mit den Inhalten übertragenen Struktur- und Zugriffsinformationen nach der Indizierung (‚Dekomposition').

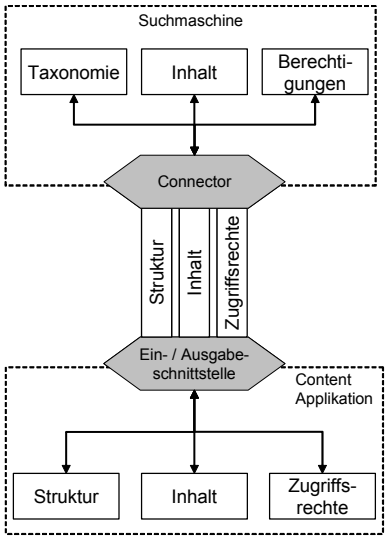
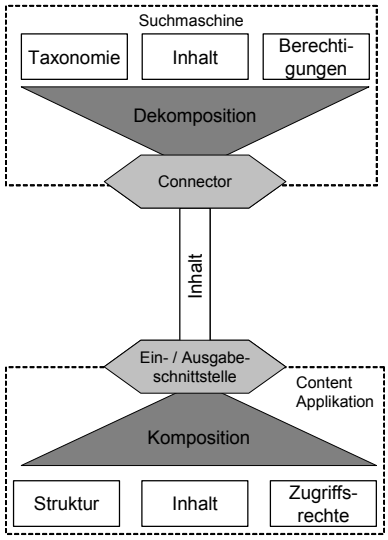
Abbildung von Inhalt, Struktur und Zugriffsrechten	
durch den Konnektor der Suchmaschine	**durch die Ein- / Ausgabeschnittstelle der Anwendung**
	

Tabelle 5-14: Abbildung von Inhalt, Struktur und Zugriffsrechten

Das Aufgabenkettendiagramm in Abbildung 5-16 stellt exemplarisch die Abläufe für die Anwendungsintegration dar. Zusätzlich zu den bei der Retrieval-Taxonomie und dem Berechtigungskonzept vorgeschlagenen Rollen (s. Abschnitt 5.5.2 und 5.5.3) sind folgende Mitwirkende notwendig:

- *Retrieval-Architekt* für die Konzeption des integrierten Informationssystems,
- *Suchmaschinenentwickler* für die Gestaltung der Suchmaschine.

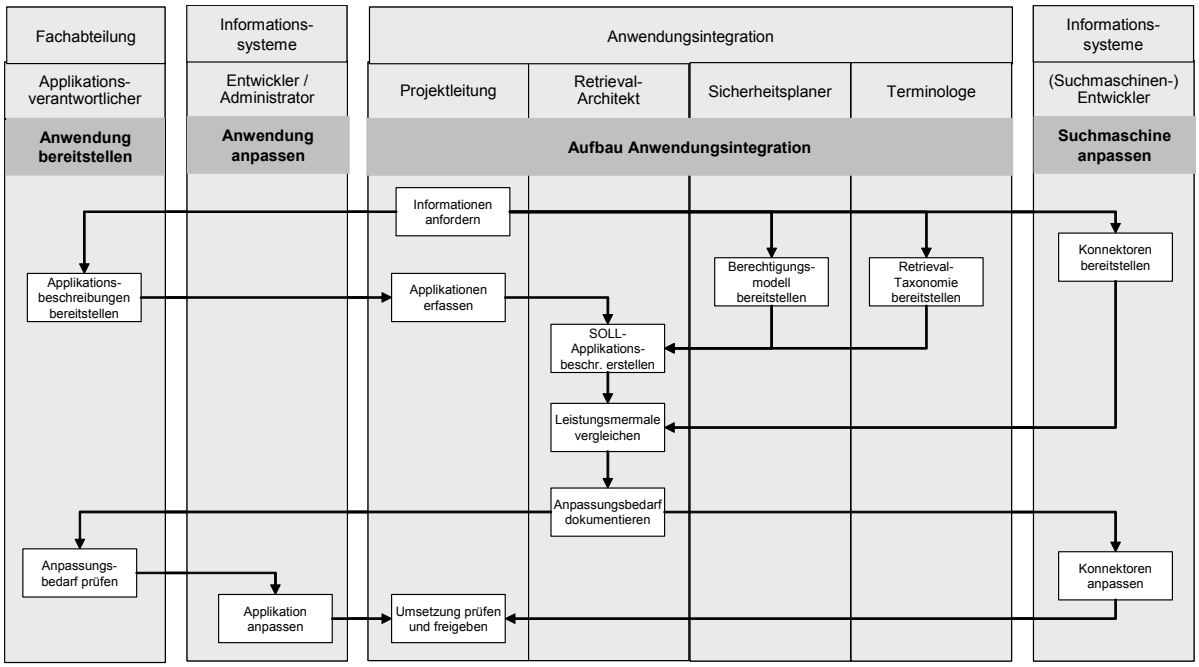


Abbildung 5-16: Aufgabenkettendiagramm zum Aufbau der Anwendungsintegration

Nach der initialen Anwendungsintegration im Rahmen eines Projekts erfolgt die Koordination der weiteren Pflege durch den Retrieval-Architekten. Die Integration neuer Anwendungen sollte in einem neuen Projekt erfolgen.

Die Teilergebnisse der vorliegenden Aktivität sind Applikationsbeschreibungen (SOLL), dokumentierte Anpassungsbedarfe an Content-Applikationen und Suchmaschine sowie das Aufgabenkettendiagramm für den Ablauf der Anwendungsintegration. Die technische Umsetzung der Ergebnisse stellt der folgende Systementwurf dar.

5.6. Systementwurf

Kurzbeschreibung der Technik: Im Systementwurf setzen die Projektmitarbeiter die fachlichen Ergebnisse von Strategie- und Prozessplanung konzeptionell auf Systemebene um. Sie spezifizieren Systemfunktionen, gestalten Anwendungsschnittstellen und implementieren Aspekte der Zugriffssicherheit.

Die Ergebnisse bilden den technischen Systementwurf, der Vorgaben und Rahmenbedingungen für die technische Umsetzung der in den vorangegangenen Phasen entwickelten Lösung gibt.

Ziele: Der Systementwurf hat zum Ziel:

- die grafische Benutzerschnittstelle sowie Funktionen für Suche und Navigation zu spezifizieren,
- Authentifizierungs- und Autorisierungsmechanismen der beteiligten Applikationen aufeinander abzustimmen,
- Integrationstechnologien für die Bereitstellung von Inhalten zu detaillieren und
- die technische Umsetzung von Indizierung und Klassifikation festzulegen.

Voraussetzungen (Input): Der Systementwurf benötigt die IS-Architektur (SOLL) der Technik Strategieplanung (s. Abschnitt 5.4) sowie den Retrieval-Prozess, den Benutzergruppenabgleich, das ‚User Mapping', die Retrieval-Taxonomie, die verfeinerten Applikations- und Schnittstellenbeschreibungen (SOLL) und die Abläufe zur Anwendungsintegration der Technik Prozessplanung (s. Abschnitt 5.5).

Ergebnisse (Output): Die Ergebnisse der Technik Systementwurf sind Benutzerschnittstellenbeschreibungen (einfache Suche, erweiterte Suche, Suchergebnisanzeige, Navigation), Funktionsbeschreibungen für Single Sign On, für Autorisierungsmechanismen, für Indizierungs-, Integrations- und Klassifikationsverfahren und ein Datenmodell für Information Retrieval.

Rollen: Die Technik Systementwurf wird unter Berücksichtigung der in Tabelle 5-15 beschriebenen Rollen angewendet.

Stufe	Initiatoren	Anwender (Mitarbeiter, Kunden)
Entscheider	Führungsverantwortlicher, Prozessverantwortlicher	Fachbereichsverantwortlicher
Verantwortlicher	Projektleiter	Fachbereichsvertreter, Applikationsverantwortlicher
Unterstützer	Projektteam, Search Manager, Portalarchitekt, Retrieval-Architekt, Anwendungsentwickler, Administrator, Sicherheitsplaner, Directory Manager, Marketing, Suchmaschinenentwickler	Fachbereichsmitarbeiter, Autor

Tabelle 5-15: Rollen der Technik Systementwurf

5.6.1. Funktionen definieren

Das Unterstützungspotenzial eines Informationssystems hängt wesentlich von seiner Funktionalität ab [s. Kruchten 2000, 195]. Die im Rahmen dieser Arbeit analysierten Portalarchitekturen (s. Kapitel 3) nennen dabei weitestgehend identische Gestaltungselemente bzw. Funktionen, z.B. Suche, Authentifizierung, Autorisierung, Benutzerverwaltung, Integration, Indizierung und Klassifikation.

Da Suchmaschinen nach dem Broker-Prinzip nicht zum Aufbau einer dynamischen, begriffsbasierten Navigationsstruktur geeignet sind (s. Tabelle 2-3, S. 25), fokussiert das vorliegende Dissertationsprojekt auf die Verwendung von Indexer-Suchmaschinen. Sie stellen den Anwendern mit Suche und Navigation zwei alternative Wege zur Erschliessung relevanter Informationen bereit.

Die folgenden Abschnitte spezifizieren diese Funktionen. Sie berücksichtigen dabei Applikationen und Datensammlungen als weitere relevante Gestaltungselemente des Systementwurfs, die auf Basis der Anforderungen von Business Engineering und Knowledge Management abgeleitet wurden (s. Abschnitt 2.5) und im integrierten Metamodell für Information Retrieval (s. Abschnitt 3.4) dargestellt sind.

5.6.2. Suche

Suchanfrage

Bei einer Suchanfrage formulieren die Anwender ihren Informationsbedarf in der grafischen Benutzerschnittstelle des Portals (Graphical User Interface, GUI), z.B. durch Eingabe eines Suchbegriffs in ein Dialogfeld. Abhängig von den in der Strategieplanung festgelegten Retrieval-Leistungen (Volltextsuche, attributbasierte Suche) müssen die Projektinitiatoren unterschiedliche Ein- und Ausgabemasken spezifizieren.

Abbildung 5-17: Einfache Suche im Internet-Portal von IWI-HSG

Die einfache Suche (‚Simple Search') bildet die Benutzerschnittstelle für eine Volltextsuche. Bei der einfachen Suche geben die Anwender in einem meist einzeiligen Eingabefeld ihren Suchbegriff ein. Der Benutzer übermittelt seine Suchanfrage zur Bearbeitung durch die Suchmaschine mit der ENTER-Taste oder durch Betätigung einer Schaltfläche. Die einfache Suche ist i.d.R. auf allen Portalseiten verfügbar.

Das Ergebnis der einfachen Suche ist eine Suchergebnisliste mit Hinweisen auf Informationsobjekte, die z.B. den gesuchten Begriff enthalten. Sie wird von der Suchmaschine auf Basis von Suchindizes erstellt und im Portal angezeigt (s. Grundlagen in Abschnitt 2.3.3).

Die erweiterte Suche (‚Advanced Search') bietet dem Anwender vielfältigere Möglichkeiten, die Suchanfrage und -anzeige seinen Bedürfnissen anzupassen. Sie bildet die grafische Benutzerschnittstelle für eine attributbasierte Suche. Grundlage zur Gestaltung der erweiterten Suche ist die Retrieval-Taxonomie aus der Technik Prozessplanung (s. Abschnitt 5.5.2). Aus ihr leiten die Projektmitarbeiter Attribute und Wertebereiche für die Eingabefelder ab. Abbildung 5-18 zeigt die erweiterte Suche im Intranet-Portal von IWI-HSG. Der Anwender kann die Suchanfrage z.B. auf den inhaltlichen Schwerpunkt (‚Topic'), den publizierenden Lehrstuhl oder das verantwortliche Kompetenzzentrum einschränken. Diese drei Attribute und die jeweiligen Wertebereiche sind Bestandteil der institutsweiten Retrieval-Taxonomie (s. Fallstudie IWI-HSG in Abschnitt 4.4).

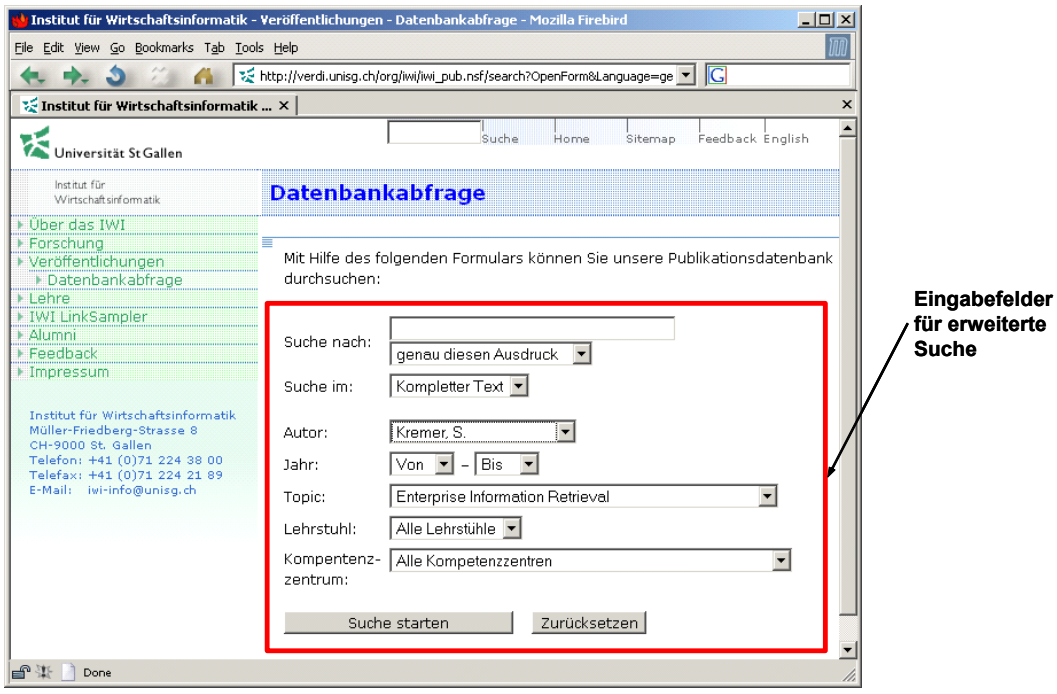


Abbildung 5-18: Erweiterte Suche im Internet-Portal von IWI-HSG

Die Projektverantwortlichen legen in den Ergebnisdokumenten ‚Benutzerschnittstelle - Einfache Suche' und ‚Benutzerschnittstelle - Erweiterte Suche' folgende Details fest:

- Layout und Position für das Eingabefeld zur einfachen Suche,
- Schaltflächen zum Aufruf einer erweiterten Suche,
- Eingabefelder und Gestaltung der erweiterten Suche.

Suchergebnisanzeige

Die Suchergebnisanzeige enthält, meist in Listenform, die Resultate der Suchanfrage des Benutzers. Die Reihenfolge wird implizit durch die interne Relevanzgewichtung der Suchmaschine, durch Angaben bei der Programmierung oder explizit durch Benutzereinstellungen vorgegeben, z.B. durch die Angabe der Sortierreihenfolge in der erweiterten Suche [s. Ferber 2003, 302].

Die Angaben zu den Suchergebnissen helfen den Benutzern, den Kontext der gefundenen Inhalte zu identifizieren. Ist es für die Anwender beispielsweise relevant, Suchergebnisse nach ihren Autoren zu differenzieren, so muss dieses Detail aus den Strukturinformationen der Inhalte (z.B. HTML-Meta-Tags) extrahiert und für jedes Suchergebnis ausgewiesen werden.

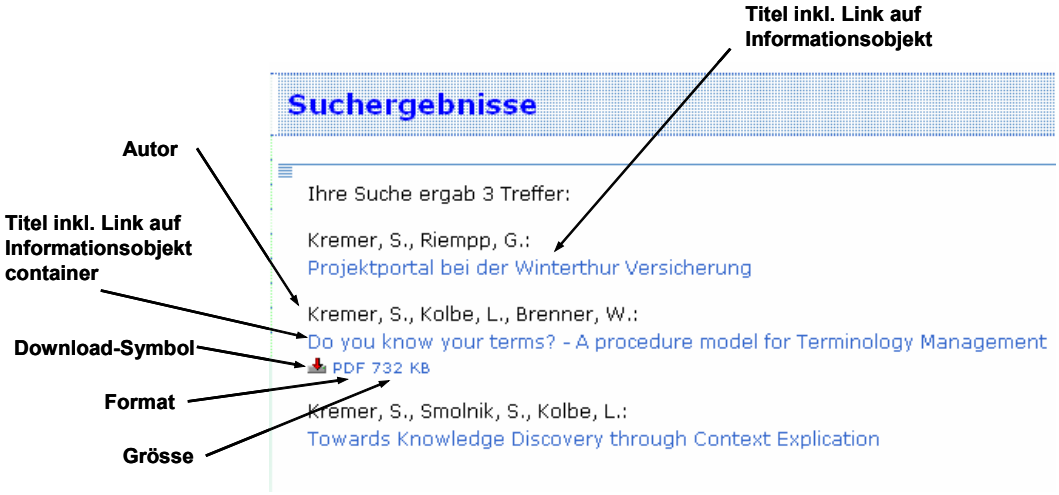


Abbildung 5-19: Suchergebnisliste im Intranet-Portal von IWI-HSG

In der Praxis erstellen die Anwender Informationsobjekte nicht ausschliesslich im Format der im Portal integrierten Content-Applikationen. Inhalte werden z.B. weiterhin auf Basis von Office-Applikationen als Word-, Powerpoint- oder PDF-Dateien erzeugt und in die Content-Applikation eingebunden. Dabei dient das Informationsobjekt der Content-Applikation als so genannter Informationsobjektcontainer (Container) für das eingebundene Informationsobjekt (s. Tabelle 5-16). Diese Verschachtelung ist eingefügten Dateianhängen in E-Mails vergleichbar.

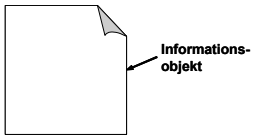
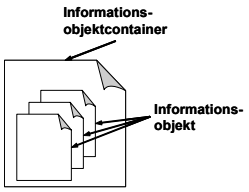
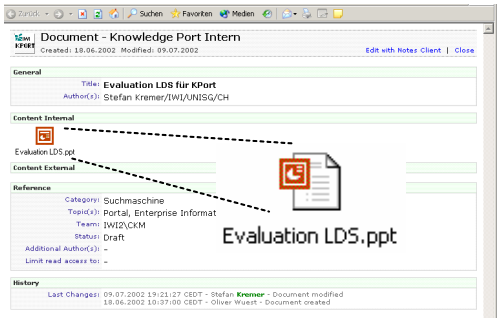
	Funktionsweise	Beispiel
Singuläres Informationsobjekt	 • Informationsobjekt besitzt eindeutiges Format, z.B. HTML • Metadaten sind im Informationsobjekt enthalten, z.B. als HTML-Meta-Tags	 HTML-Dokument, angezeigt im Web-Browser
Verschachteltes Informationsobjekt	 • Informationsobjektcontainer besitzt eindeutiges Format, z.B. HTML • Enthaltene Informationsobjekte besitzen eindeutige Formate, z.B. MS Powerpoint • Metadaten können in allen Informationsobjekten vorhanden sein	 HTML-Dokument mit verschachtelter Powerpoint-Datei, angezeigt im Web-Browser

Tabelle 5-16: Singuläre und verschachtelte Informationsobjekte

Die Portalsuchmaschine indiziert alle in einer Content-Applikation enthaltenen textbasierten Inhalte, d.h. sowohl ein eingebundenes Informationsobjekt als auch den zugehörigen Informationsobjektcontainer. Beide können dabei als Einträge in einer Suchergebnisliste vorkommen, wenn z.B. ein Suchbegriff sowohl im Informationsobjekt als auch im Text des Containers enthalten ist. Die Projektverantwortlichen müssen für die Spezifikation der Suchergebnisliste folgende Möglichkeiten unterscheiden:

- *Suchbegriff im Informationsobjekt.* Der Eintrag in der Suchergebnisliste weist direkt auf das Informationsobjekt, z.B. eine HTML-Seite mit dem Text ‚Geschäftsprozess'. Das Portal zeigt dem Anwender bei Auswahl des Treffers die HTML-Seite im Portal an. Die HTML-Meta-Tags liefern die beschreibenden Attribute.

- *Suchbegriff im Container*. Der Treffer in der Ergebnisliste zeigt auf ein Informationsobjekt, welches wiederum Informationsobjekte beinhaltet. Dies kann beispielsweise eine HTML-Seite sein, in der ein Word-Dokument eingefügt ist. In der HTML-Seite (Container) ist der Suchbegriff ‚Geschäftsprozess' enthalten, im Word-Dokument nicht. Das Portal stellt dem Portalbenutzer die HTML-Seite dar. In diesem Fall agiert der Informationsobjektcontainer wie ein Informationsobjekt. Die Suchergebnisliste enthält die Inhalte der HTML-Meta-Tags des Informationsobjekts.

- *Suchbegriff im Container und im Informationsobjekt.* Sowohl eine HTML-Seite als auch ein dort eingefügtes Word-Dokument enthalten den Suchbegriff ‚Geschäftsprozess'. In diesem Fall weist der Eintrag der Suchergebnisliste entweder auf das Informationsobjekt (Word-Dokument) oder auf den Container (HTML-Seite). Alternativ kann eine Suchergebnisliste auch mehrere Links pro Treffer enthalten, die sowohl auf einen Container als auch auf das darin enthaltene Informationsobjekt weisen. Da in der Praxis die Qualität von Office-Metadaten eher gering ist, sollten die Projektverantwortlichen zur Beschreibung von Suchergebnissen die Meta-Tags des HTML-Dokuments verwenden.

Die Informationsobjekte der im Intranet-Portal von IWI-HSG eingebundenen Publikationsdatenbank nehmen maximal ein weiteres Informationsobjekt auf. Die Suchergebnisliste (s. Abbildung 5-20, linke Seite) bietet bei solchen Treffern neben einem Link auf den Informationsobjektcontainer (s. Abbildung 5-20, obere rechte Grafik) eine direkte Verknüpfung zum enthaltenen Informationsobjekt (s. Abbildung 5-20, untere rechte Grafik).

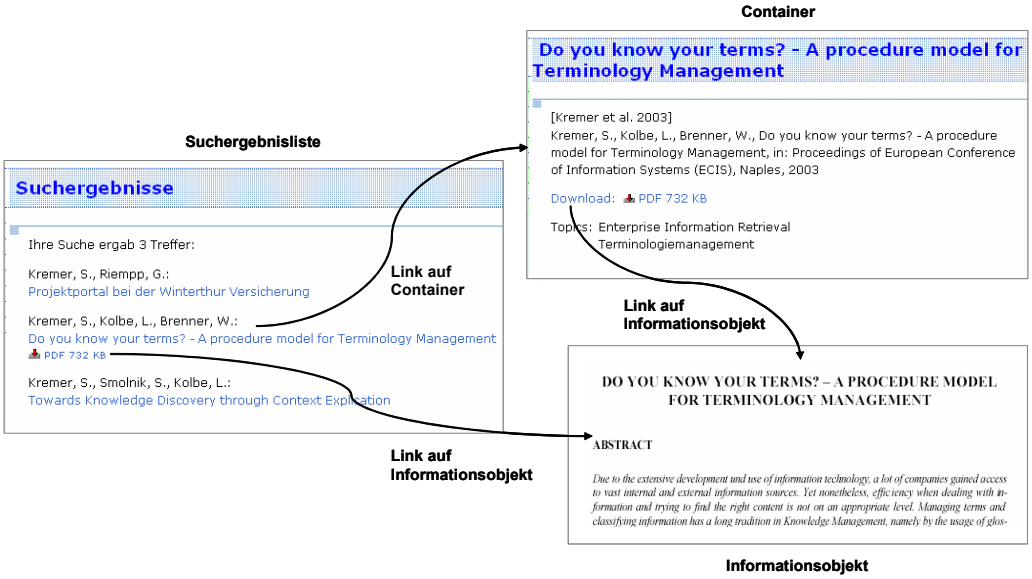
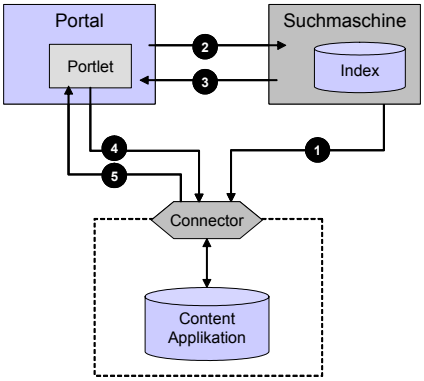
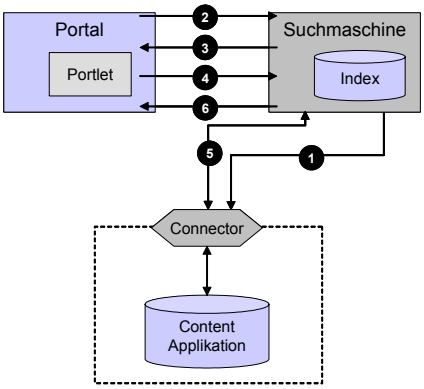


Abbildung 5-20: Verweise auf Informationsobjekte bei IWI-HSG

Der Portalbenutzer kann z.B. aufgrund der angegebenen Dateigrösse entscheiden, ob er dieses Informationsobjekt direkt aufrufen und ggf. längere Wartezeiten für die Übertragung grösserer Dateien in Kauf nehmen will.

Zusätzlich zu der Entscheidung, ‚wann' Informationsobjekte angezeigt werden, müssen die Verantwortlichen im Systementwurf festlegen ‚wie' diese dem Portalbenutzer präsentiert werden.

Dabei wählen die Projektverantwortlichen aus den in Tabelle 5-17 skizzierten Anzeigemechanismen aus.

Funktionsweise	Charakteristika
Portlet für Content-Applikation (‚Portlet')	
 • (1) Die Suchmaschine indiziert die Informationsobjekte der Content-Applikation und baut einen Index auf. • (2) Der Portalbenutzer stellt eine Suchanfrage. • (3) Die Suchmaschine liefert die Treffer zurück und zeigt die Suchergebnisliste im Portal an. • (4) Der Benutzer wählt einen Suchergebnislisteneintrag aus, den die Content-Applikation im Portlet anzeigt (5).	• Ein auf die inhaltsliefernde Applikation abgestimmtes Portlet zeigt das Informationsobjekt im Portal an. • Die Content-Applikation übernimmt nach der Auswahl eines Treffers aus der Suchergebnisliste die Anzeige. • Die Leistungsmerkmale der Content-Applikation und des Portals limitieren die Performance, z.B. für die Zeitdauer von der Auswahl eines Suchergebnisses bis zur Anzeige im Portal. • Suchbegriffe können im angezeigten Inhalt nicht hervorgehoben werden.
Portlet für Suchmaschine (‚Parsed')	
 • (1) – (4) analog zu ‚Portlet für Content-Applikation'. • (5) Die Suchmaschine fragt bei der Content-Applikation das gewählte Informationsobjekt an. • (6) Die Suchmaschine verarbeitet das Informationsobjekt und zeigt es im Portlet an.	• Die Inhalte der Content-Applikation werden ausschliesslich durch die Suchmaschine in einem Portlet zur Anzeige gebracht. • Die Suchmaschine kann die Inhalte vor der Anzeige aufbereiten, z.B. durch Veränderung des Layouts, Hervorheben von Suchbegriffen oder die Einblendung einer Leiste für die direkte Navigation zwischen den Treffern. • Die Performance ist sowohl durch die Leistungsmerkmale der Content-Applikation als auch die der Suchmaschine limitiert.

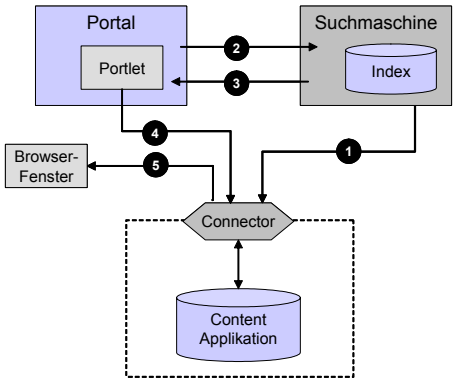
Anzeige in separatem Browser-Fenster (‚Bypassed')	
 • (1) – (4) analog zu ‚Portlet für Content-Applikation'. • (5) Die Content-Applikation zeigt das gewünschte Informationsobjekt in einem separaten Browser-Fenster an.	• Die Inhalte werden durch den direkten Aufruf des Informationsobjekts durch die Content-Applikation in einem separaten Browser-Fenster angezeigt. • Eine Manipulation der Inhalte wie das Hervorheben von Suchbegriffen oder das Einblenden von Navigationsleisten ist nicht möglich. • Die Performance ist ausschliesslich durch die Leistungsmerkmale der Content-Applikation begrenzt.

Tabelle 5-17: Anzeigemechanismen

Für die Suchergebnisanzeige legen die Projektverantwortlichen im Dokument ‚Benutzerschnittstelle - Suchergebnisliste' folgende Systemspezifika fest:

- Bestandteile der Suchergebnisliste,
- Verhalten von Links (Informationsobjekt, Container, beides),
- Anzeigemechanismus (‚Portlet', ‚Parsed', ‚Bypassed').

5.6.3. Navigation

Für die Gestaltung der Navigationskomponenten eines Portals kann das Layout-Paradigma aus Abbildung 5-21 zugrunde gelegt werden (vgl. [Bauer 2001, 57], [Rosenfeld/Morville 2002, 106]).

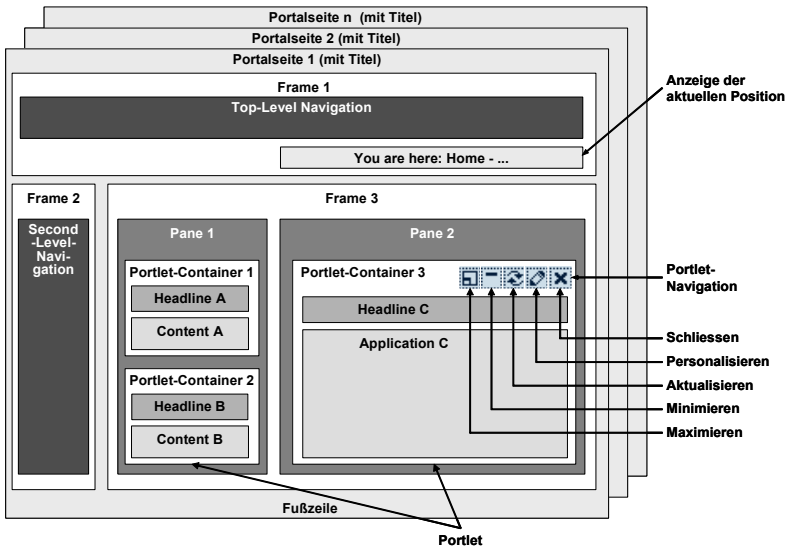


Abbildung 5-21: Grundlegendes Layout-Paradigma in Portalen (in Anlehnung an [Puschmann 2003, 81])

Das Portal-Layout besteht aus folgenden Elementen:

- Anwender steuern über die Einträge der *Top-Level-Navigation* einzelne Portalseiten an, die bestimmte Applikationen und Inhalte enthalten. Die Top-Level-Navigation ist die primäre Orientierungsfunktion im Portal. Sie kann beispielsweise nach Geschäftsbereichen eines Finanzdienstleisters, nach Produktgruppen eines Industrieunternehmens oder nach Prozessen eines Beratungsunternehmens strukturiert sein.

- Die *Second-Level-Navigation* ergänzt die Einträge der Top-Level-Navigation durch eine verfeinerte oder alternative Struktur. Top- und Second-Level-Navigation können für alle Anwender gleich oder je nach Rolle und Benutzer individuell sein. Die Navigationsleisten im SAP Enterprise Portal sind beispielsweise von der Benutzerrolle abhängig. Die Top-Level-Navigation beeinflusst in diesem Fall auch die Second-Level-Navigation. Im Beispiel des Beratungsunternehmens kann die Second-Level-Navigation mit Subprozessen die angebotenen Prozesse der Top-Level-Navigation verfeinern.

- *Portlets* zeigen im Frontend Inhalte und Funktionen aus den Backend-Applikationen an. Sie stellen z.B. eine Eingabemaske für Suchanfragen oder die Liste mit Suchergebnissen dar, die von der Portalsuchmaschine bearbeitet werden. Auf einer Portalseite werden vom Portaladministrator zentral oder rollenspezifisch vorgegebene sowie vom Benutzer je nach Zugriffsberechtigung zusätzlich ausgewählte Portlets angezeigt. Die Oberfläche und Funktionen einzelner Portlets können je nach Berechtigung eines Anwenders personalisiert werden.

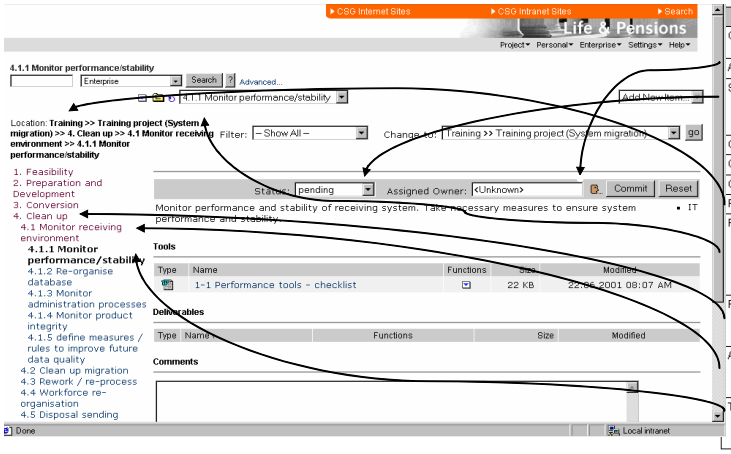
- Die *Portlet-Navigation* bietet dem Benutzer je nach Zugriffsberechtigung verschiedene Bedienungsfunktionen innerhalb eines Portlets wie ‚Schliessen' und ‚Personalisieren'.

- Eine *Headline* bezeichnet als Überschrift die Inhalte (*Content*) und Funktionen (*Application*) eines Portlets.

- Lokalisationsanzeigen (‚*You are here*'-Leisten) helfen den Anwendern bei der Bestimmung ihrer augenblicklichen Position in der Navigationsstruktur.

Mit dieser feingranularen Aufteilung der Elemente von Portalseiten definiert das Projektteam die inhaltsorientierte Navigation im Portal. Dabei spezifiziert es ausschliesslich die Bestandteile, die für die Informationsfindung durch eine taxonomieorientierte Navigation relevant sind. Als Grundlage dient hier die Retrieval-Taxonomie der Technik Prozessplanung (s. Abschnitt 5.5.2).

Die Attribute ‚Project', ‚Process', ‚Phase', ‚Activity' und ‚Task' der Retrieval-Taxonomie (Abbildung 5-22, rechte Seite) bilden die hierarchische First-Level Navigation im McB-Portal der Winterthur Versicherungen (Abbildung 5-22, linke Seite). Diese ist, entgegen der Darstellung des allgemeinen Layout-Paradigmas in Abbildung 5-21, nicht im oberen Bereich, sondern aufgrund technischer Vorgaben des verwendeten Systems Opentext Livelink auf der linken Bildschirmseite angeordnet.



Attribut	Wertebereich	Bemerkung
Competence	Actuary, Business Organisation, IT, Operations, Legal, Finance / Accounting, Testing, Data quality, Conversion	Gruppen von Spezialisten
Assigned to	-	Zuständiger Bearbeiter
Status	pending, in process, completed, not applicable	Bezeichnung unterschiedlicher Fertigstellungszustände von Projekten, Phasen, Aktivitäten oder Aufgaben
Created by	-	Ersteller eines Dokuments
Creation Date	-	Erstelldatum eines Dokuments
Country	-	Länderbezeichnung
Project	z.B. Migration-UK-Lifesystem	Projektbezeichnung
Process	Transparency, System Migration, Re-Engineering, Knowledge Management, Country relationship management, European Operational Council, McB strategy development, People management, Process and quality management, McB Project management, Performance measurement	Bezeichnung der Kernprozesse der McB-Initiative
Phase	Ausschnitt für Migration: 1. Feasability, 2. Preparation & Development, 3. Conversion, 4. Cleanup	Bezeichnung der Phasen der McB-Initiative
Activity	Ausschnitt für 4. Cleanup: 4.1 Monitor receiving environment, 4.2 Reorganize workforces, and 4.3. Dispose of the sending system	Bezeichnung der Aktivitäten der McB-Initiative
Task	Ausschnitt für 4.3 Dispose of the sending system: 4.3.1 Switch off the system, 4.3.2 Dispose of hardware, 4.3.3 Dispose of software	Bezeichnung der Aufgaben der McB-Initiative

Abbildung 5-22: Attribute und Werte der Retrieval-Taxonomie in der Navigation bei Winterthur Versicherungen

Neben den Anforderungen der Anwender an die Navigation unterliegt das Layout auch den technischen Möglichkeiten und Restriktionen der verwendeten Portalplattform. Auch Marketing-Abteilungen oder andere im Unternehmen für Fragen des Corporate Design zuständige Bereiche stellen u.U. Anforderungen an die grafische Gestaltung eines Portals.

Müssen die Retrieval-Verantwortlichen sowohl eine attributbasierte Suche als auch eine taxonomieorientierte Navigation realisieren, so sollten beide Retrieval-Leistungen aufeinander abgestimmt sein.

Die primäre Navigation (Top-Level-Navigation) im Winterthur-Portal ist nach Prozessen orientiert (s. Fallstudie in Abschnitt 4.3). Abbildung 5-23 zeigt im linken Bereich der oberen Grafik die Navigationseinträge für die Projektphase ‚4. Clean up' mit ‚4.1 Monitor receiving environment` und ‚4.1.1 Monitor performance/stability'. Die im rechten Bereich der oberen Grafik dargestellte sekundäre Navigation bietet Auswahlmöglichkeiten für das relevante Teilprojekt (‚Filter'), den Bearbeitungszustand (‚Status') sowie den zuständigen Bearbeiter (‚Assigned Owner'). Für die gezeigte Auswahl (‚Filter: - Show All' und ‚Status: pending' und ‚Assigned Owner: <Unknown>') ist beispielsweise das

Informationsobjekt ‚1-1 Performance tools – checklist` verfügbar. Auch die Maske für die erweiterte Suche (untere Grafik) stellt die Attribute ‚Filter', ‚Status' und ‚Assigned Owner' für das Information Retrieval im Portal zur Verfügung.

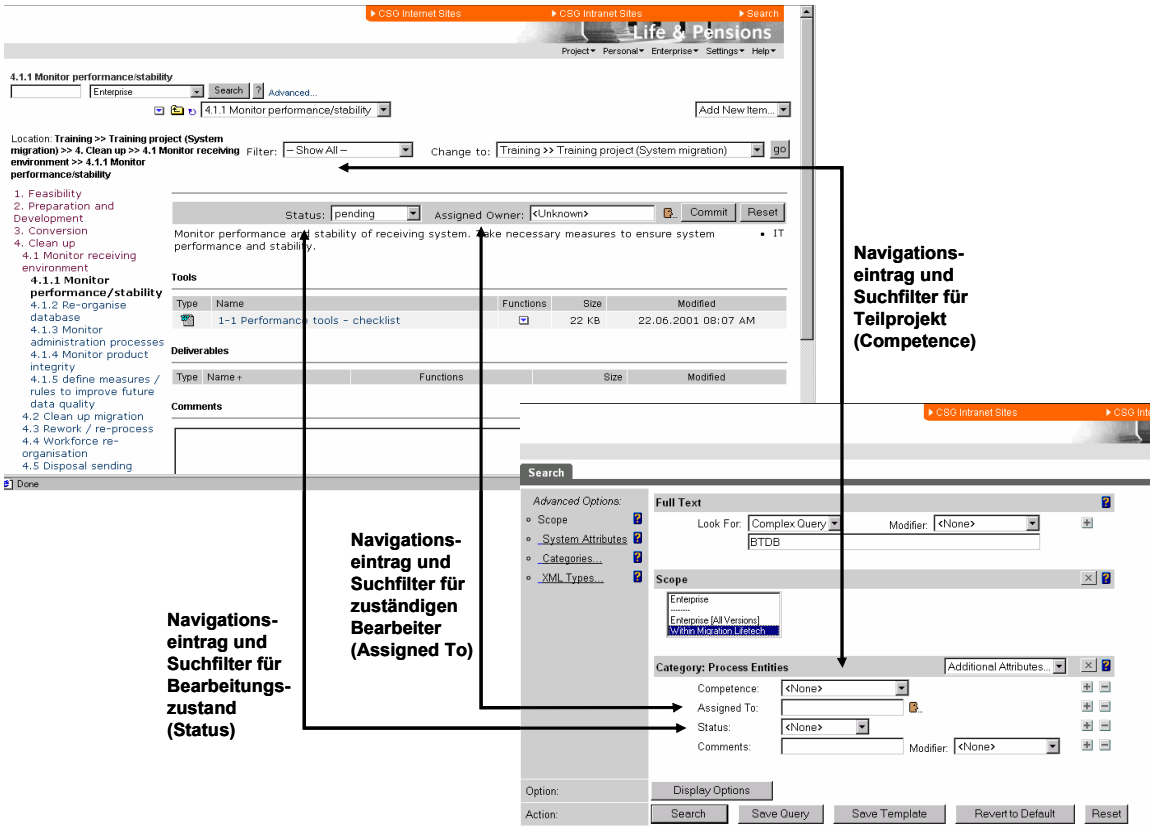


Abbildung 5-23: Abstimmung von Navigation und Suche im Winterthur-Portal

Wählt ein Anwender in der erweiterten Suche die oben genannten Werte zu diesen Attributen aus, so listet die Suchmaschine die ‚Performance tools – checklist' in der Suchergebnisliste auf. Diese Abstimmung von Navigation und Suche erleichtert den Anwendern des McB-Portals das Finden von Inhalten. Die Nutzung der erweiterten Suche setzt jedoch auch eine gewisse Erfahrung mit der Navigation des Portals und der Strukturierung der Inhalte voraus.

Im Dokument ‚Benutzerschnittstelle – Navigation' halten die Projektverantwortlichen die Bestandteile der Attribute und Wertebereiche fest, anhand derer Benutzer in den Inhalten navigieren können. Eine visuelle Designstudie ergänzt diese Entwurfsaktivität und dient der späteren Programmierung der grafischen Benutzerschnittstelle als Vorgabe.

5.6.4. Sicherheit

Authentifizierung und Single Sign On

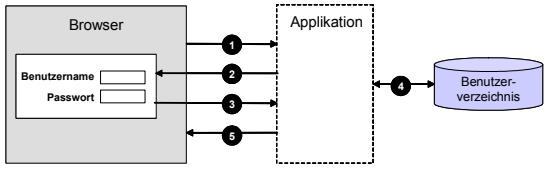
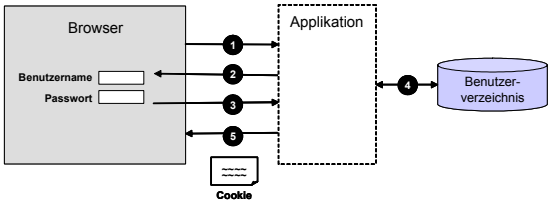
Die Authentifizierung einer Anwendung identifiziert einen Benutzer und prüft dessen Echtheit durch Zugriff auf ihre Benutzerverwaltung [s. Raepple 2001, 5]. Da eine Information-Retrieval-Lösung jedoch nicht aus einer einzelnen Applikation besteht, verlangt das Zusammenspiel von Suchmaschine, Portal und integrierten Content-Applikationen eine Integration der vorhandenen Authentifizierungsmechanismen. Single Sign On (SSO) ermöglicht darüber hinaus, dass ein einmal am Portal angemeldeter Benutzer auf die Suchmaschine und die integrierten Inhalte der Content-Applikationen zugreifen kann, ohne sich erneut authentifizieren zu müssen [vgl. Chinitz 2000, 34].

Die technische Umsetzung erfordert von den Projektverantwortlichen eine genaue Kenntnis der zu integrierenden Authentifizierungsmechanismen. Die beiden in der Praxis vorherrschenden Verfahren sind (s. [Raepple 2001, 134], [Zwerger/Schneider 2002, 153]):

- *Benutzername und Passwort*. Dieser Mechanismus authentifiziert den Benutzer über die Eingabe von Benutzername und Passwort. Die vom Benutzer eingegebenen Informationen werden gegen die in einem Benutzerverzeichnis hinterlegten Vergleichsinformationen verifiziert. Stimmen diese überein, ist der Benutzer erfolgreich authentifiziert. Ein Beispiel ist die Anmeldung mit Benutzernamen und Passwort an einem Portal über die weitverbreitete ‚Basic Authentication' des HTTP-Protokolls [s. Zwerger/Schneider 2002, 154].

- *Digitale Zertifikate*. Die Authentifizierung erfolgt durch den Vorweis eines elektronischen Benutzerzertifikats. Dieses verbindet die natürliche Identität einer Person oder eines Systems mit einer digitalen Identität [s. Merz 2002, 186]. Das digitale Zertifikat besteht i.d.R. aus einem öffentlichen und einem privaten elektronischen Schlüssel, die mathematisch voneinander abhängig sind, sowie weiteren Angaben über den Zertifikatsinhaber, wie z.B. Name, Adressinformationen und organisatorische Einordnung. Die Ausstellung und digitale Signatur des Zertifikats erfolgt durch Zertifizierungsinstanzen, die häufig auch als Certificate Authority (CA) oder Trust-Center bezeichnet werden. Grundsätzlich können dabei hard- und softwarebasierte Verfahren unterschieden werden. Während Software-Lösungen das Zertifikat des Portalbenutzers im Web-Browser oder in speziellen Clients verwalten und speichern, legen hardwarebasierte Lösungen dieses verschlüsselt auf Speicherkarten (Smart Cards, PC-Cards) oder Geräten für die USB-Schnittstelle (USB-Krypto-Tokens) ab [s. Zwerger/Schneider 2002, 205].

Tabelle 5-18 zeigt die Funktionsweise und Charakteristika der beiden Authentifizierungsverfahren. Die im Authentifizierungsprozess abgebildete generische ‚Applikation' steht dabei stellvertretend für ein Portal, eine Suchmaschine oder eine Content-Applikation.

Mechanismus	Authentifizierungsprozess	Charakteristika
Benutzername und Passwort		
HTTP Basic Authentication	 • (1) Der Benutzer ruft eine zugriffsgeschützte Applikation auf. • (2) Die Anwendung fordert den Benutzer über ein Eingabefenster zur Authentifizierung auf. • (3) Der Benutzer gibt seinen Benutzernamen und sein Passwort ein. • (4) Die Applikation prüft die eingegebenen Daten gegen die Informationen im Benutzerverzeichnis. • (5) Stimmen die Angaben, so gewährt die Anwendung dem Benutzer Zugang.	• Die Benutzerdaten werden mit jeder neuen Anfrage im HTTP-Header übertragen. • Das Passwort wird dabei unverschlüsselt übertragen, daher ist dieses Verfahren nur in Kombination mit einer zusätzlichen SSL-Verschlüsselung sicher.
HTTP Session Authentication	 • (1) Der Benutzer ruft eine zugriffsgeschützte Applikation auf. • (2) Die Anwendung fordert den Benutzer über ein HTML-Formular zur Authentifizierung auf. • (3) Der Benutzer gibt seinen Benutzernamen und sein Passwort ein. • (4) Die Applikation prüft die eingegebenen Daten gegen die Informationen im Benutzerverzeichnis. • (5) Stimmen die Angaben, so gewährt die Anwendung dem Benutzer Zugang. Sie generiert eine Session, die der Browser in Form eines Cookies oder als HTTP-Parameter bei jeder weiteren Anfrage überträgt.	• Die erste Anmeldung über das HTML-Formular sollte verschlüsselt über SSL erfolgen.

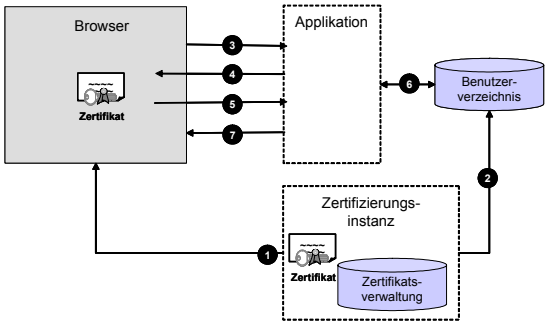
Digitale Zertifikate		
Public Key Infrastructure (PKI)	 • (1) Eine Zertifizierungsinstanz (Certificate Authority, CA) stellt ein digital signiertes Zertifikat aus, das beispielsweise im Browser des Anwenders hinterlegt wird. Das Zertifikat kann vom Anwender durch ein Passwort geschützt werden. • (2) Das Zertifikat wird in der Benutzerverwaltung der Applikation oder in einer für sie zugänglichen Benutzerverwaltung abgelegt. Die Applikation muss für die Authentifizierung auf Basis digitaler Zertifikate konfiguriert werden. • (3) Der Benutzer ruft die zugriffsgeschützte Applikation auf. • (4) Die Anwendung fordert den Browser zur Präsentation des digitalen Zertifikats auf. • (5) Der Browser präsentiert das digitale Benutzerzertifikat. • (6) Die Applikation verifiziert die Signatur des Zertifikats und optional dessen Gültigkeit gegen die eigene oder eine zentrale Benutzerverwaltung. • (7) Ist die Verifikation erfolgreich und das Zertifikat noch gültig, so gewährt die Anwendung dem Benutzer Zugang.	• Die Zertifikate können vom Unternehmen selbst oder einer kommerziellen Instanz, z.B. Verisign, ausgestellt werden. • Die digitalen Zertifikate enthalten Informationen zum Benutzer, zum Gültigkeitszeitraum, zur Zertifizierungsinstanz oder Attribute. • Bei der Verteilung der Zertifikate muss sicher-gestellt werden, dass ausschliesslich die Zertifikatsinhaber in den Besitz des Zertifikats gelangen. • Nutzen mehrere An-wender denselben Rechner, so ist die Ab-lage des Zertifikats im Browser zu unsicher. Zertifikate können dabei kopiert oder Passwörter unrechtmässig ange-eignet werden. Hier ist die Verwendung hardware-basierter Zertifikats-lösungen notwendig.

Tabelle 5-18: Authentifizierungsmechanismen

Die Projektverantwortlichen identifizieren mit den jeweiligen Applikations-verantwortlichen die in allen betroffenen Anwendungen vorhandenen Authenti-fizierungsmechanismen. Sie erweitern damit die Applikationsbeschreibungen (IST) der Technik Potenzialanalyse (s. Abschnitt 5.3.3).

Abhängig von den vorhandenen Authentifizierungsmechanismen ist die Umsetzung einer Single-Sign-On-Funktionalität mit vier grundlegenden Mechanismen möglich (vgl. [Carden 1999], [Zwerger/Schneider 2002, 69]):

- *Elektronisches Ticket*. Bei diesem Verfahren meldet sich ein Benutzer beim Portal an. Anschliessend erhält er ein elektronisches Ticket, das als nichtpersistenter Cookie im Browser abgelegt wird. Beim Zugriff auf alle weiteren Anwendungen wie Suchmaschine oder Content-Applikationen wird dieser Cookie übermittelt und ausgewertet. Dieses Verfahren setzt voraus, dass die angebundenen Applikationen das vom Portal oder einem zentralen Authentifizierungsserver ausgestellte Ticket akzeptieren. Weiterhin müssen Portal, Suchmaschine und Content-Applikationen unter derselben Domain im Intranet, Extranet oder Internet erreichbar sein.

Die Konfiguration eines elektronischen Tickets mit dem bei IWI-HSG evaluierten IBM Websphere Portal Server zeigt Abbildung 5-24. Ein Administrator konfiguriert in der Sicherheitsverwaltung (‚Security Center') des Portalservers die Merkmale zur Erstellung eines elektronischen Tickets. Die im Beispiel dargestellte Lightweight Third Party Authentication (LTPA) ist ein Standard von IBM, um ein Single Sign On zwischen verschiedenen Produkten dieses Herstellers zu ermöglichen. Im so genannten Token definiert der Administrator beispielsweise die Gültigkeitsdauer eines Tickets nach seiner Ausstellung (z.B. 120 Minuten) oder die betroffene Domain (z.B. iwi.unisg.ch). Nach der Generierung exportiert der Portaladministrator dieses Token in eine Datei und stellt diese den Administratoren der Content-Applikationen zur Verfügung. In der Systemverwaltung des Domino-Servers, der bei IWI-HSG die Content-Applikation beinhaltet, importiert ein Domino-Administrator das vom Portal ausgestellte Token.

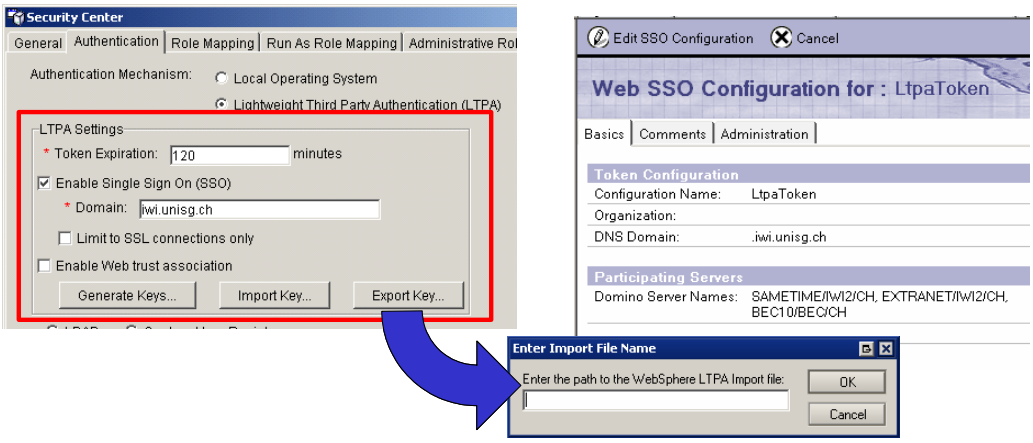


Abbildung 5-24: Gemeinsame Nutzung eines elektronischen Tickets durch Portal und Content-Applikationen bei IWI-HSG

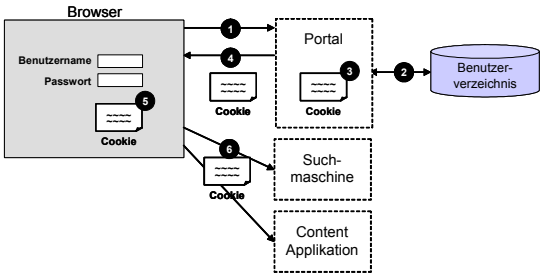
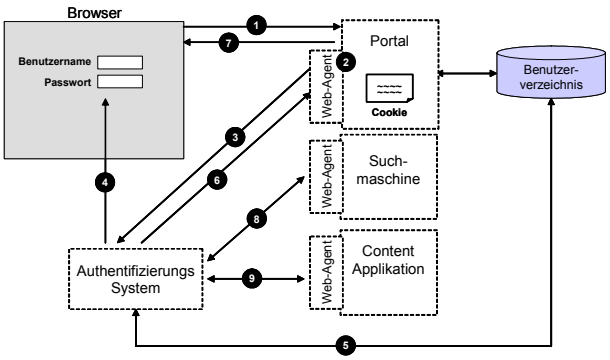
- *SSO-Client*. Bei diesem Verfahren werden auf den Rechnern der Benutzer zusätzliche Programme, so genannte SSO-Dienste (SSO-Clients), installiert. Diese erkennen selbständig Authentifizierungsaufforderungen verschiedener Applikationen. Der SSO-Client extrahiert daraufhin aus einem Unternehmensbenutzerverzeichnis die benötigten Anmeldeinformationen des jeweiligen Benutzers und leitet diese automatisch an die aufrufende Applikation weiter. Ein Beispiel für ein solches SSO-Verfahren ist das SecureLogin von Novell, das alle Benutzernamen und Passwörter zentral verwaltet.

- *Serverseitiges Agentensystem*. Bei diesem SSO-Verfahren werden zusätzliche Softwarekomponenten, so genannte Web-Agenten, bei den zu integrierenden Applikationen installiert. Diese übernehmen die Authentifizierungsprozesse aller beteiligten Applikationen. Ist ein Benutzer an einer Anwendung erfolgreich

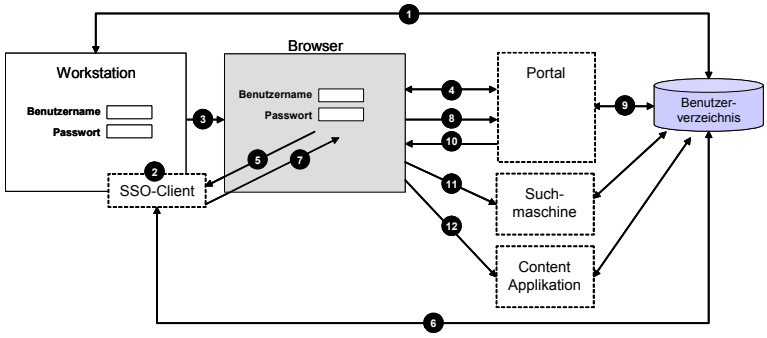
authentifiziert, so akzeptieren alle anderen Anwendungen diesen Zustand auch. Produkte wie IBM Tivoli Access Manager, Netegrity Siteminder oder Novell PolicyDirector liefern Agenten für eine Vielzahl von Systemen, z.B. SAP, IBM Lotus Notes und Windows-Betriebssysteme.

- *Zertifikatsbasierte Public-Key-Infrastruktur*. Ein weiteres SSO-Verfahren ist mit dem Einsatz digitaler Zertifikate realisierbar (s. Tabelle 5-18). Dabei werden die Zertifikate sowohl zur Anmeldung am Portal als auch zur Authentifizierung an den integrierten Applikationen eingesetzt. Es setzt voraus, dass alle beteiligten Applikationen eine Authentifizierung anhand von digitalen Zertifikaten unterstützen.

Tabelle 5-19 zeigt die Funktionsweise und Charakteristika der beschriebenen SSO-Verfahren bei der Integration von Portal, Suchmaschine und Content-Applikationen:

Single-Sign-On-Verfahren	Charakteristika
Elektronisches Ticket	
 • (1) Der Anwender authentifiziert sich am Portal, z.B. durch Angabe von Benutzername und Passwort. • (2) Der Portal-Server verifiziert die Informationen, z.B. durch Vergleich mit den Benutzerdaten im Benutzerverzeichnis. • (3) Nach erfolgreicher Überprüfung generiert das Portal ein Ticket. • (4) Der Portal-Server übermittelt die Startseite des Portals und überträgt das Ticket in Form eines Cookies an den Browser. • (5) Der Browser speichert das Ticket als nichtpersistenten Cookie und überträgt es bei jeder weiteren Anfrage. • (6) Alle weiteren Applikationen wie Suchmaschine und Content-Applikationen verarbeiten beim Aufruf eines Benutzers das elektronische Ticket.	• Bereits der Besitz des Tickets reicht zur Authentifizierung. Daher sollte es ausschliesslich über eine SSL-verschlüsselte Verbindung übertragen werden. • Zur Umstellung ihrer Authentifizierungsfunktionen auf die Verwendung eines elektronischen Tickets müssen bestehende Applikationen u.U. angepasst werden. Die Hersteller von Ticket-Servern bieten hier Programmierbibliotheken. • Die Verwendung von Cookies muss erlaubt sein.
Serverseitiges Agentensystem	
 • (1) Der Benutzer ruft eine geschützte Portalseite auf. • (2) Das Portal ruft einen Web-Agenten zur Authentifizierung auf. • (3) Der Web-Agent leitet die Anfrage an das Authentifizierungs-System weiter. • (4) Das Authentifizierungs-System fordert den Benutzer mit einem Login-Formular zur Eingabe von Benutzernamen und Passwort auf. • (5) Das Authentifizierungs-System authentifiziert den Benutzer gegen das Benutzerverzeichnis. • (6) Die erfolgreiche Authentifizierung leitet das System an den Web-Agenten weiter. • (7) Das Portal zeigt die gewünschte Seite an. • (8/9) Die Authentifizierung gegen alle weiteren Systeme verläuft analog, wobei die Web-Agenten einen einmal authentifizierten Benutzer erkennen und nicht erneut auf das Benutzerverzeichnis zugreifen.	• Das Verfahren erfordert die Installation von Softwarekomponenten auf Portal-, Suchmaschinen- und Content-Servern. • Die Hersteller bieten eine Vielzahl vorkonfigurierter Agenten für verschiedenste Applikationen. • Müssen Agenten für Applikationen bereitgestellt werden, für die kein Agent existiert, so kann eine Programmierschnittstelle (Application Programming Interface, API) des Herstellers genutzt werden.

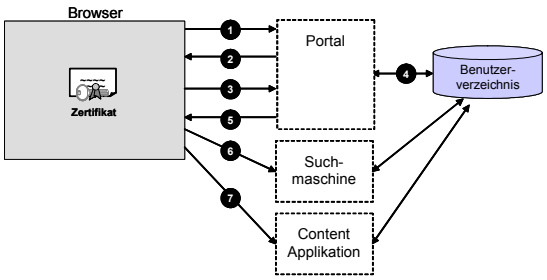
SSO-Client



- (1) Der Benutzer meldet sich an seinem Rechner (Workstation) an. Die Anmeldeinformationen werden im Benutzerverzeichnis verifiziert.
- (2) Nach erfolgreicher Anmeldung wird auf dem Rechner ein SSO-Dienst gestartet.
- (3/4) Der Benutzer startet seinen Browser und ruft eine geschützte Portalseite auf. Das Portal sendet dem Browser ein Login-Formular und wartet auf die Authentifizierungsangaben des Benutzers.
- (5/6) Die Anfrage des Portals nach Benutzernamen und Passwort wird vom SSO-Dienst erkannt und über eine verschlüsselte Verbindung werden der Portalbenutzername und das Passwort aus dem Unternehmens-Benutzerverzeichnis extrahiert.
- (7/8) Der SSO-Client trägt die Benutzerdaten automatisch in das Login-Formular des Portals ein. Dabei übernimmt das Portal die Anmeldedaten genauso, als hätte ein Benutzer diese direkt eingegeben.
- (9) Das Portal überprüft die Authentifizierungsinformationen gegen das Unternehmens-Benutzerverzeichnis.
- (10) Nach erfolgreicher Überprüfung übermittelt der Portal-Server die gewünschte Seite.
- (11/12) Erfordert auch der Zugriff auf Suchmaschine und Content-Applikationen eine Authentifizierung, so werden diese für das jeweilige System analog der Schritte (4) bis (10) vom SSO-Dienst ermittelt.

- Das Verfahren eignet sich, wenn bereits ein Benutzerverzeichnis mit entsprechenden Leistungsmerkmalen im Unternehmen vorhanden ist, z.B. das eDirectory von Novell.
- Der SSO-Dienst erfordert eine Installation auf den Rechnern der Benutzer.
- Der SSO-Dienst erkennt eine Vielzahl von Eingabeaufforderungen, z.B. auch von Mainframe-Systemen.
- Das Verfahren kann im Bedarfsfall mit einem elektronischen Ticket kombiniert werden.
- Der Einsatz ist auch in stark heterogenen Systemumgebungen möglich.

Zertifikatsbasierte Public-Key-Infrastruktur



(1) Der Benutzer ruft eine zugriffsgeschützte Applikation auf.

(2) Die Anwendung fordert den Browser zur Präsentation des digitalen Zertifikats auf.

(3) Der Browser präsentiert das digitale Benutzerzertifikat.

(4) Die Applikation verifiziert die digitale Signatur des Zertifikats und optional dessen Gültigkeit gegen die eigene oder eine zentrale Benutzerverwaltung.

(5) Ist die Verifikation erfolgreich und das Zertifikat noch gültig, so gewährt die Anwendung dem Benutzer Zugang.

(6/7) Die Authentifizierung gegen alle weiteren Systeme verläuft analog.

- Alle beteiligten Anwendungen müssen eine Authentifizierung anhand von digitalen Zertifikaten unterstützen.
- Der Einsatz ist auch in stark heterogenen Systemumgebungen möglich.

Tabelle 5-19: Überblick über verschiedene SSO-Verfahren

Die Projektverantwortlichen erarbeiten mit den Applikationsverantwortlichen und dem Sicherheitsplaner mögliche Lösungsansätze. Sie berücksichtigen dabei bestehende Applikationsbeschreibungen (IST) der Technik Potenzialanalyse und stellen die dort erfassten Authentifizierungsmechanismen den Charakteristika der in Tabelle 5-19 dargestellten Verfahren gegenüber. Die abschliessende Einigung auf ein zu implementierendes SSO-Verfahren induziert i.d.R. Anpassungsbedarf bei den bestehenden Applikationen. Diesen ergänzt das Projektteam in den Applikationsbeschreibungen (SOLL).

Benutzerverwaltung

In der Benutzerverwaltung bzw. dem Benutzerverzeichnis sind sämtliche Aufgaben zusammengefasst, die mit dem Anlegen, Modifizieren und Sperren von elektronischen Identitäten zusammenhängen. Um ein Single Sign On bereitzustellen und die redundante Pflege von Benutzerdaten für Portal, Suchmaschine und Content-Applikationen zu reduzieren, schlägt die Prozessplanung die Konsolidierung aller Benutzerverwaltungen in einem zentralen Verzeichnis auf Basis des Lightweight Directory Access Protocol (LDAP) vor (s. Abschnitt 5.5.3). Das ‚User Mapping' (s. Tabelle 5-13, S. 154) liefert dazu bereits die Zuordnung unterschiedlicher elektronischer Identitäten zu einem Benutzer. Diese organisatorischen Ergebnisse setzen die Systemverantwortlichen wie folgt um:

- *Bereitstellung des zentralen Benutzerverzeichnisses*. Die Projektverantwortlichen verwenden, falls im Unternehmen vorhanden, ein bestehenden LDAP-Verzeichnis. Auch das Benutzerverzeichnis einer existierenden Anwendung kann als LDAP-Verzeichnis dienen, wenn der Anpassungsaufwand gering ist. Das proprietäre Benutzerverzeichnis ‚Domino Directory' der Groupware-Plattform IBM Lotus Notes erlaubt beispielsweise durch die Umstellung eines Konfigurationsdokuments den alternativen Zugriff über LDAP. Weitere Produkte für die Bereitstellung eines zentralen Benutzerverzeichnisses auf LDAP-Basis sind z.B. IBM Policy Director, Novell eDirectory, Oracle Internet Directory oder das Apple Open Directory.

- *Substitution der Authentifizierungsmechanismen*. Die Projektverantwortlichen prüfen, ob die zu integrierenden Anwendungen die Substitution des eigenen Benutzerverzeichnisses durch ein zentrales LDAP-Verzeichnis zulassen. Im einfachsten Fall erlaubt die Anwendung einem Administrator in einem Konfigurationsmenü, die Angabe eines zentralen LDAP-Verzeichnisses vorzunehmen. Eine aufwändigere Alternative ist die Verwendung spezieller Programmierschnittstellen (Application Programming Interface, API), mit denen der generische Authentifizierungsmechanismus einer Anwendung auf die Verwendung eines zentralen LDAP-Verzeichnisses umgestellt werden kann. [SUN

2004] bietet dazu mit den Java Authentication and Authorization Services (JAAS) beispielsweise einen offenen Standard.

Im Rahmen der Softwareevaluation nutzen die Projektmitarbeiter bei IWI-HSG ein vorhandenes Lotus Notes / Domino Benutzerverzeichnis (‚Domino Directory'), um die Benutzerverwaltung der Suchmaschine Verity Information Server über LDAP abzubilden. Abbildung 5-25 zeigt die Konfiguration im ‚Domino Directory' und im Administrationsdialog der Verity-Suchmaschine.

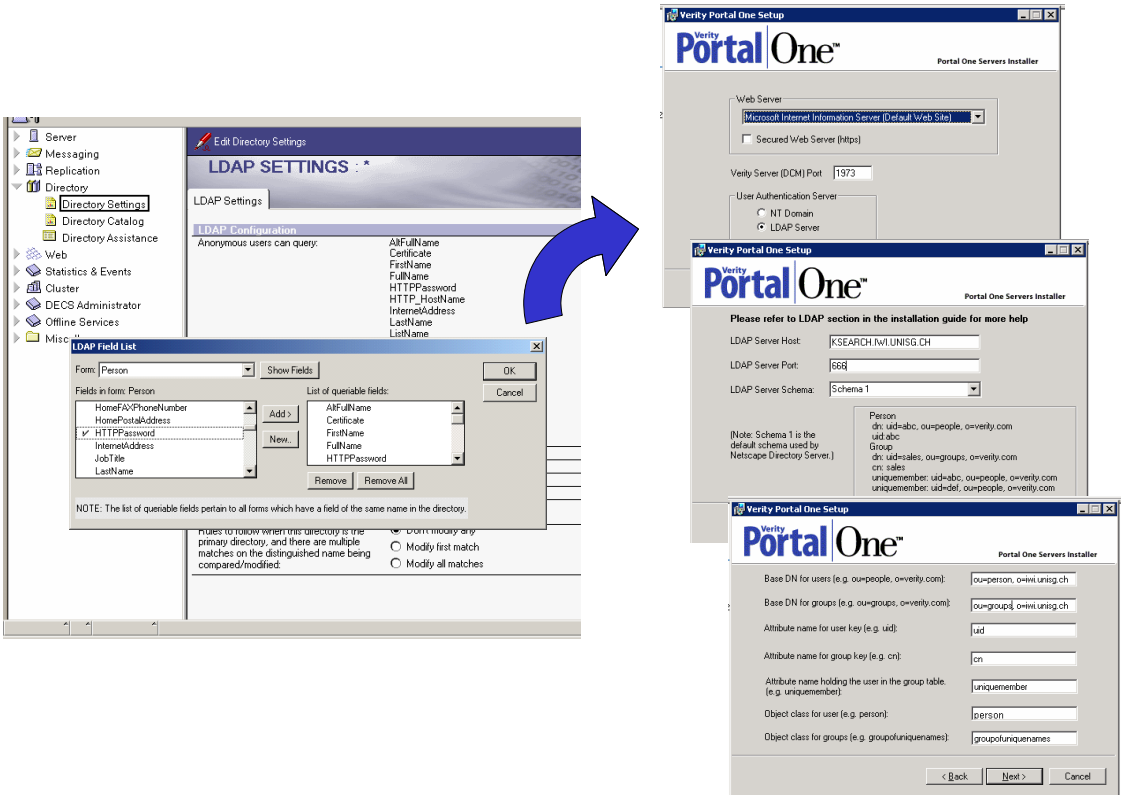


Abbildung 5-25: Nutzung eines bestehenden, zentralen LDAP-Verzeichnisses für die Benutzerverwaltung der Verity-Suchmaschine (Evaluation bei IWI-HSG)

Bei der Konfiguration zeigten sich besondere Herausforderungen, da die vom Domino Directory bereitgestellten Angaben nicht mit der vom Verity-Server benötigten Benutzerdatenstruktur, dem so genannten LDAP-Schema, übereinstimmten. Durch eine manuelle Korrektur konnten die beiden unterschiedlichen LDAP-Schemata aufeinander abgestimmt werden.

- *User Mapping umsetzen*. Über ein User Mapping unterschiedlicher Benutzernamen setzen die Systemverantwortlichen einen einheitlichen Zugriffsmechanismus um, ohne die dezentralen Systeme erst zusammenführen zu müssen. Die Ablage und Pflege dieser Zuordnungstabellen unterstützen beispielsweise SUN Java System Identity Server, Microsoft Active Directory Services, Novells eDirectory,

Computer Associates' eTrust Admin, Waveset Lighthouse oder IBM Tivoli Identity Manager.

Autorisierung

Bei der Autorisierung prüft eine Anwendung, ob ein Benutzer zur Durchführung einer bestimmten Aktion berechtigt ist. Analog zur Authentifizierung verlangt auch die Autorisierung bei Information-Retrieval-Lösungen eine Abstimmung der beteiligten Systeme wie Portal, Suchmaschine und Content-Applikationen. Dokumente, auf deren Inhalt ein Anwender keinen Zugriff hat, dürfen beispielsweise nicht in einer Suchergebnisliste erscheinen.

Die Prüfung der Autorisierung erfolgt wahlweise bei der Indizierung oder der Suchanfrage. Beide Fälle gehen davon aus, dass die zu indizierenden Dokumente Zugriffsbeschränkungen aufweisen:

- *Autorisierungsprüfung bei der Indizierung*. Die Zugriffsbeschränkungen für Benutzer oder Gruppen werden von der Suchmaschine bei der Indizierung ausgelesen und im Index gespeichert. Sucht ein Anwender, so listet die Suchmaschine ausschliesslich Inhalte aus dem Index auf, für die der Benutzer aufgrund der im Index gespeicherten Informationen zugriffsberechtigt ist. Ändern die Applikationsverantwortlichen die Zugriffsberechtigungen zwischen dem Zeitpunkt der Indizierung und der Suchanfrage eines Benutzers, so kann die Suchergebnisliste Dokumente anzeigen, auf die der Benutzer de facto keine Zugriffsrechte mehr besitzt.
- *Autorisierungsprüfung bei der Suchanfrage*. Die Suchmaschine indiziert Dokumente ohne Berücksichtigung von bestehenden Zugriffsbeschränkungen. Erst bei der Suchanfrage eines Anwenders fragt die Suchmaschine beim Aufbau der Suchergebnisliste für jedes einzelne Dokument bei der betreffenden Content-Applikation an, ob der Anfragende zugriffsberechtigt ist. Nur bei bestehender Autorisierung zum Zeitpunkt der Benutzung erscheint ein Treffer in der Suchergebnisliste. Dieser Vorteil ist jedoch mit hohen Anforderungen an das Antwortzeitverhalten der Anwendungen und das Netzwerk verbunden, da die Suchmaschine für jedes potentielle Suchergebnis die aktuellen Zugriffs-beschränkungen bei der Applikation nachfragt.

Tabelle 5-20 fasst die Funktionsweise und Charakteristika zusammen. Sie dient den Projektverantwortlichen zur Auswahl eines geeigneten Verfahrens.

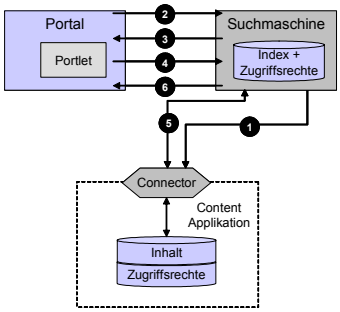
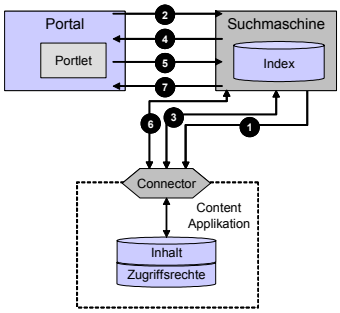
Autorisierungsprüfung	Charakteristika
Bei der Indizierung	
 • (1) Die Suchmaschine indiziert die Inhalte der Content-Applikation. Sie legt diese im Index ab und vermerkt die Zugriffsrechte pro Dokument. • (2) Der Benutzer stellt eine Suchanfrage. • (3) Die Suchmaschine liefert dem Benutzer eine Suchergebnisliste gemäss den gespeicherten Zugriffsrechten. • (4) Der Benutzer ruft ein Suchergebnis auf. • (5) Die Suchmaschine fragt das Dokument bei der Content-Applikation nach und zeigt es dem Benutzer im Portal an (6).	• Die Zugriffsrechte sind im Index der Suchmaschine für jedes Informationsobjekt gespeichert. • Das Verfahren setzt entsprechende Schnittstellen (Konnektoren) der Suchmaschine voraus, die diese Zugriffsrechte bei der Indizierung auslesen können. • Vor der Anzeige der Suchergebnisse erfolgt keine Überprüfung, ob die im Index gespeicherten Zugriffsrechte von denen im operativen System abweichen. • Die Suchergebnisliste wird vergleichsweise schnell angezeigt, kann jedoch im schlimmsten Fall Einträge aufweisen, auf die der Benutzer keinen Zugriff hat.
Bei der Suchanfrage	
 • (1) Die Suchmaschine indiziert die Inhalte der Content-Applikation. • (2) Der Benutzer stellt eine Suchanfrage. • (3) Die Suchmaschine prüft für die Einträge im Index bei der Content-Applikation nach, ob der aktuelle Benutzer zugriffsberechtigt ist. • (4) Die Suchmaschine liefert dem Benutzer eine Suchergebnisliste gemäss den in der Content-Applikation hinterlegten Zugriffsrechten. • (5) Der Benutzer ruft ein Suchergebnis auf. • (6) Die Suchmaschine fragt das gewünschte Dokument bei der Content-Applikation nach und zeigt es dem Benutzer an (7).	• Vor der Anzeige der Suchergebnisliste wird für jeden Treffer geprüft, ob der aktuelle Benutzer zugriffsberechtigt ist. • Die Suchmaschine fragt dazu die aktuelle Zugriffsberechtigung des Benutzers in den Content-Applikationen ab, aus denen die Inhalte stammen. • Das Verfahren setzt entsprechende Konnektoren der Suchmaschine voraus, die vor der Anzeige eine solche Filterung vornehmen können. • Der Aufbau der Suchergebnisliste kann vergleichsweise lange dauern.

Tabelle 5-20: Verfahren zur Autorisierungsprüfung von Suchergebnissen

Die Pflege der Zugriffsrechte erfolgt über Benutzergruppen, die zur Vereinfachung der Administration einzelne Anwender zusammenfassen. Im Ergebnisdokument ‚Benutzergruppenabgleich' der Technik Prozessplanung haben die Projektverantwortlichen die Zugriffsgruppen der Content-Applikationen bereits den Berechtigungsgruppen der Suchmaschine zugeordnet (s. Abschnitt 5.5.3). Die Abbildung dieses Konzepts durch die Portalsuchmaschine kann auf unterschiedlichen Ebenen erfolgen:

- *Autorisierung auf Suchmaschinenebene*. Bei diesem Verfahren steht die Portalsuche nur der Benutzergruppe zur Verfügung, die Zugriff auf die geschützten Content-Applikationen hat. Alle anderen Anwender sind von der Verwendung der Suche ausgeschlossen. Daher ist die Autorisierung auf Suchmaschinenebene nur dann sinnvoll, wenn alle anderen dargestellten Verfahren versagen.

- *Autorisierung auf Indexebene*. Unterschiedliche Zugriffsrechte werden durch separate Suchindizes abgebildet [vgl. Bach 2003, 268]. Je nach Zugriffsbeschränkung müssen pro Applikation ein oder mehrere Indizes erstellt werden. Im ungünstigsten Fall (‚Worst Case') muss jeder Benutzergruppe ein eigener Index zur Verfügung gestellt werden, z.B. durch die mehrfache Indizierung einer Anwendung unter Verwendung unterschiedlicher Benutzerkennungen (Identifications, IDs). Die Benutzer sind dabei Vertreter der jeweiligen Berechtigungsgruppe. In der Praxis wird dafür häufig eine spezielle Benutzerkennung, die so genannte Dummy-ID für die Suchmaschine bereitgestellt.

Die Zugriffsrechte des Content-Management-Systems (CMS) und des Produktkatalogs (PDB) sind bei Zumtobel Staff landes- und sprachspezifisch. So stehen den Anwendern des deutschsprachigen Customer Portals in Österreich andere Inhalte zur Verfügung als den französischsprachigen Benutzern des Schweizer-Portals. Die Suchmaschine TREX berücksichtigt diese Zugriffsrechte durch die Autorisierung auf Indexebene.

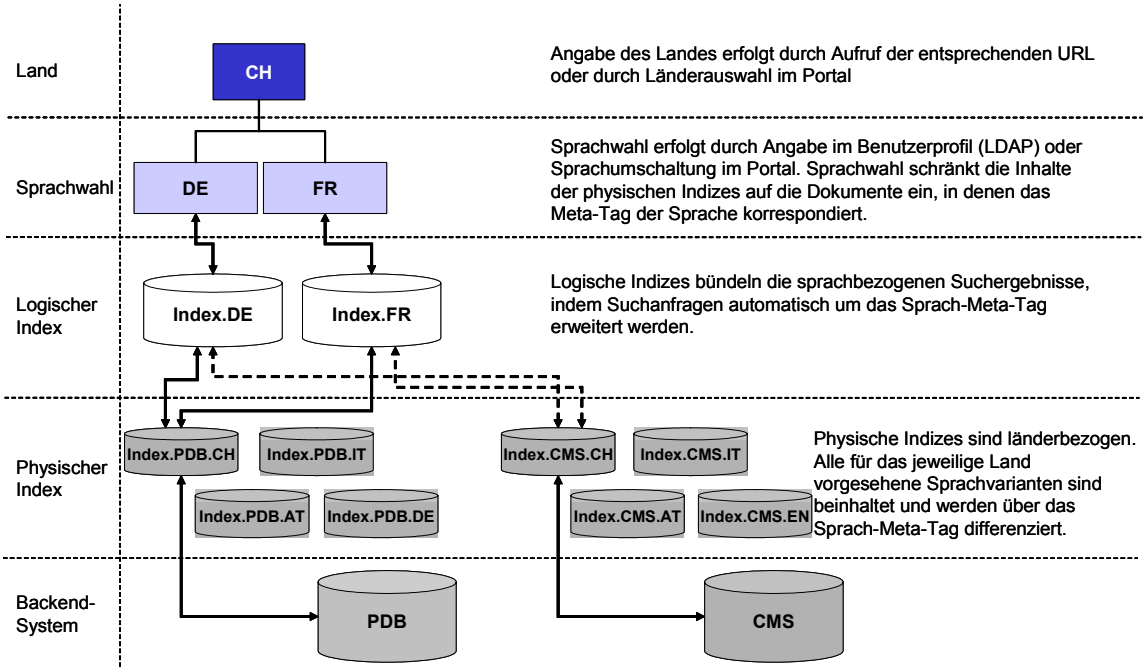


Abbildung 5-26: Landes- und sprachspezifische Zuordnung von Suchmaschinenindizes im Zumtobel Staff Portal

- *Autorisierung auf Informationsobjektebene*. Die Zugriffsrechte werden bei der Indizierung für jedes Informationsobjekt einzeln abgebildet [vgl. Bach 2003, 268]. Dies setzt jedoch voraus, dass die Zugriffsrechte bei der Erstellung im Dokument gespeichert worden sind.

Zugriffsbeschränkungen können bei den Content-Applikationen von IWI-HSG im Bedarfsfall für jedes einzelne Dokument vergeben werden.

Die Abbildung zeigt, wie die Gruppe ‚IWI 2 CKM' beim Erstellen des Dokuments für den Lesezugriff autorisiert wird.

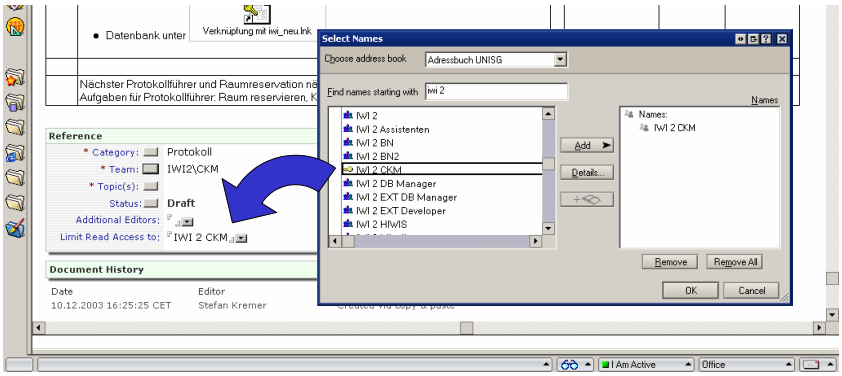


Abbildung 5-27: Autorisierung auf Informationsobjektebene bei IWI-HSG

Die Portalsuche ‚Domain Search' berücksichtigt diese Zugriffsbeschränkung bei der Indizierung. Personen, die nicht in der Gruppe enthalten sind, finden das Dokument bei einer Suchanfrage nicht.

Autorisierungsebene	Charakteristika
Suchmaschine	• Die gesamte Suchfunktionalität steht ausschliesslich einer speziellen Benutzergruppe zur Verfügung.
Index	• Unterschiedlichen Benutzergruppen stehen separate Indizes zur Verfügung. • Das Verfahren setzt voraus, das Inhalte aus Content-Applikationen gemäss der Zugriffsberechtigungen der Benutzergruppen in separate Indizes eingeteilt werden können. • Inhalte in den verschiedenen Indizes sind u.U. redundant und erhöhen das Datenvolumen in den Indizes. Dies kann einerseits die Antwortzeiten bei einer Suchanfrage verlängern. Andererseits ist oft die Erweiterung der bestehenden Server-Hardware notwendig (mehr Plattenspeicher etc.).
Informationsobjekt	• Den Benutzern bzw. Benutzergruppen stehen alle Inhalte gemäss ihrer Zugriffsrechte in einem Index zur Verfügung. • Das Verfahren erfordert die Zuordnung von Benutzern bzw. Benutzergruppen zu Informationsobjekten.

Tabelle 5-21: Autorisierungsebenen von Suchmaschinen

Analog zur Autorisierungsprüfung entscheiden die Projektverantwortlichen gemeinsam mit den Applikationsverantwortlichen und dem Sicherheitsplaner anhand der Funktionsweise und Charakteristika der Tabelle 5-21 über die Umsetzung der Autorisierungsebene. Sie halten die Ergebnisse im Dokument ‚Autorisierungsprüfung und Autorisierungsebenen' fest.

5.6.5. Integration

Information Retrieval fokussiert mit dem Einsatz von Portalsuchmaschinen auf die Integration von Daten aus anderen Applikationen. Suchmaschinen greifen dabei über definierte technische Schnittstellen (Konnektoren) auf relevante Inhalte der benötigten Anwendungen zu [vgl. Linthicum 2000, 1-20]. Diese Integration schafft einerseits technische und organisatorische Abhängigkeiten, erfordert daher die Festlegung von Integrationsstandards [s. Gebert 2003, 174]. Anderseits beeinflussen die gewünschten oder notwendigen Integrationsstandards den späteren Softwareentscheid.

Die Projektverantwortlichen für Information Retrieval müssen bei der Festlegung von Integrationsstandards unterschiedliche Dimensionen berücksichtigen:

- *Integrationsreichweite (Innerbetrieblich, Überbetrieblich)*. Information Retrieval in Portalen ist nicht auf den Einsatz innerhalb eines Unternehmens begrenzt. Der vorliegende technische Systementwurf fokussiert auf Beispiele der innerbetrieblichen Integration, ohne a priori zwischenbetriebliche Ansätze auszugrenzen.

- *Integrationsebene (Strategie, Prozess, System)*. Integration muss Gestaltungselemente auf Strategie-, Prozess- und Systemebene betrachten [s. Österle et al. 1996, 3]. Information Retrieval in Portalen umfasst, im Sinne dieser Arbeit, alle drei Ebenen (s. Gestaltungselemente des integrierten Metamodells in Abschnitt 3.4). Die Technik Systementwurf setzt die aus den vorherigen Phasen abgeleiteten strategischen Ziele und Prozessanforderungen auf Systemebene um.

- *Integrationshoheit (lesender Zugriff, schreibender Zugriff)*. Suchmaschinen verursachen ausschliesslich lesenden Zugriff auf die angebundenen Content-Applikationen. Obwohl sie bei der Integration einzig als Nachfrager auftreten und keine Integrationshoheit besitzen, zeigen die Praxisbeispiele (s. Kapitel 4), dass in Retrieval-Projekten dabei jedoch durchaus auch die inhaltsliefernden Content-Applikationen angepasst werden müssen.

- *Integrationsbereich (Backend, Frontend)*. Portale stellen personalisierte Inhalte und Funktionen in einem Frontend bereit. Der Nutzen für den Anwender ist jedoch die Backend-Integration der im Frontend angebotenen Leistungen [s. Fleisch/Österle 2001a, 28]. Information Retrieval muss daher elektronische Inhalte

sowohl im Frontend (s. Systementwurf für Suchanfrage und Suchergebnisanzeige in Abschnitt 5.6.2) als auch im Backend integrieren (s. Gestaltungselemente des Metamodells in Abschnitt 3.4).

- *Integrationsgegenstand (Prozess, Objekt, Daten)*. Information Retrieval integriert bei der Indizierung von Inhalten die Objekte und Daten bestehender Anwendungen. Damit adressiert Information Retrieval Qualitätssteigerungen, Zeitersparnis oder Kostenreduktion in informationsintensiven Prozessen (s. Abschnitt 3.4.1), sieht jedoch Prozesse selbst nicht als Integrationsgegenstand.
- *Integrationsformate (HTML, XML) und -schnittstellen (HTTP, HTTPS etc.)*. Die unterschiedlichen Formate und Ausgabeschnittstellen der inhaltsliefernden Applikationen müssen beim Aufbau einer Information-Retrieval-Lösung integriert werden. Sie verursachen den grössten technischen Aufwand.

Gemäss den oben genannten Ausprägungen steht bei IR-Projekten der lesende Zugriff auf Informationsobjekte und Daten unterschiedlicher Content-Applikationen in verschiedenen Formaten über Backend-Schnittstellen im Vordergrund. Diese Content-Applikationen sind im betrieblichen Umfeld aufgrund ihrer Entstehung häufig heterogen. Der Systementwurf hat bereits die Integration heterogener Benutzer- und Berechtigungsverwaltungen behandelt (vgl. technische Spezifikation für Authentifizierung und Single Sign On, S. 171 sowie Autorisierung, S. 180). Durch die für Information Retrieval relevanten, jedoch u.U. stark unterschiedlichen Informationsobjektformate, verfügbaren technischen Schnittstellen, und Datenstrukturen der Anwendungen entstehen weitere Integrationsherausforderungen [vgl. Alt et al. 2002, 80]:

Integrationsherausforderung 1: Unterschiedliche Speicher- und Ausgabeformate

Sowohl Anwender als auch Portalsuchmaschinen greifen beim Information Retrieval ausschliesslich lesend auf die Inhalte von Content-Applikationen zu. Während interne Speicherformate der Anwendungen durchaus unterschiedlich sein können, stehen für die Verantwortlichen in IR-Projekten beim Systementwurf die möglichen Ausgabeformate im Vordergrund. Kommerzielle Suchmaschinen können dabei bis zu mehr als 250 unterschiedliche Formate lesen und indizieren. Die Suchmaschine K2 der Firma Verity berücksichtigt beispielsweise neben HTML und XML auch Microsoft Office Dokumente, Adobe PDF, ZIP-Dateien oder Framemaker-Dokumente.

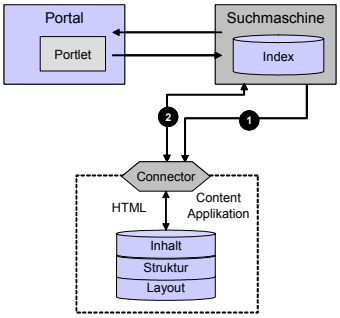
Gemäss [Finkelstein/Aiken 2000, 313] sind die Hypertext Markup Language (HTML) und die Exensible Markup Language (XML) die für Suchmaschinen wichtigsten Ausgabeformate:

- Betriebliche Anwendungen geben ihre Inhalte, unabhängig vom internen Speicherformat, vielfach im HTML- und XML-Format aus. Anwender greifen auf

diese HTML-Inhalte direkt oder auf XML-Inhalte nach einer automatischen Konvertierung (z.B. von XML zu HTML) mit einem Web-Browser zu.

- Kommerzielle Portalplattformen bieten standardisierte Schnittstellen zu HTML und XML, z.B. zur Anzeige von Inhalten einer Content-Applikation in einem Portlet [s. Puschmann 2003, 131].
- HTML und XML implementieren elektronische Verweise (Links) zu weiteren Informationsobjekten. Suchmaschinen nutzen diese Verknüpfungen zwischen Informationsobjekten für die automatische Indizierung.

Tabelle 5-22 zeigt den grundlegende Aufbau sowie die für Information Retrieval relevanten Charakteristika der Ausgabeformate HTML und XML:

Ausgabeformat	Charakteristika
HTML	
 • (1) Die Suchmaschine indiziert Inhalte und Strukturinformationen der HTML-Datei. • (2) Die in der HTML-Datei definierten Layout-Informationen werden bei der Anzeige des Inhalts verwendet.	• Inhalt, Struktur und Layout eines Informationsobjekts sind in HTML miteinander verbunden. • Inhaltsbeschreibende Strukturinformationen können in HTML-Meta-Tags hinterlegt und von Suchmaschinen für eine attributbasierte Suche berücksichtigt werden. • Bei der Indizierung von HTML-Dateien verweisen die enthaltenen Links die Indizierungskomponente der Suchmaschine (Crawler, Spider, Indexer) auf weitere Informationsobjekte.

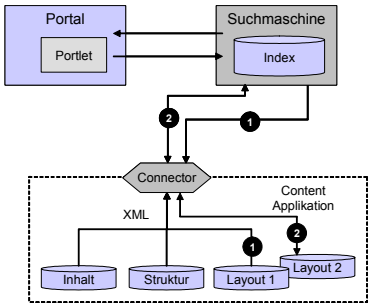
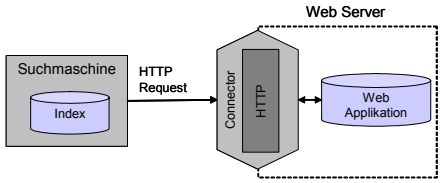
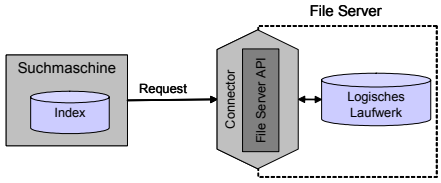
XML	
 • Eine Dokumententypdefinition (DTD bzw. XML Schema) bestimmt für eine XML-Datei gültige Elemente, Attribute sowie deren hierarchische Struktur. • Die XML-Datei verwaltet den eigentlichen Inhalt eines Informationsobjekts. Er ist durch die in der DTD bzw. im XML-Schema definierten Elemente (Tags) strukturiert. • Erst zum Zeitpunkt der Informationsanzeige werden die XML-Inhalte mit Layoutanweisungen versehen, die in so genannten Stylesheets abgelegt sind. Die in der Praxis am häufigsten verwendeten Sprachen zur Definition von Stylesheets sind Cascading Style Sheets (CSS) und die Extensible Stylesheet Language Tranformation (XSLT). • Die Trennung von Inhalt und Layout erlaubt unterschiedliche Anzeigemechanismen bei der Indizierung (1) und Anzeige (2) von Inhalten.	• Inhalt, Struktur und Layout sind in XML-Dateien voneinander getrennt [vgl. Merz 2002, 201]. • Inhalte lassen sich in beliebig kleine Strukturinformationen zerlegen und semantisch aufbereiten [vgl. Puschmann 2003, 76]. • Bei der Indizierung von XML-Dateien kann die Suchmaschine weitere Inhalte indizieren, indem sie den im Dokument enthaltenen Links ‚folgt'. • Für die Informationsanzeige müssen Inhalt, Struktur und Layout zusammengefügt und dem Benutzer präsentiert werden.

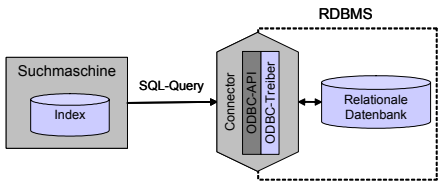
Tabelle 5-22: Aufbau und Charakteristika von HTML und XML

Die Projektverantwortlichen verfeinern die Applikationsbeschreibungen (IST) aus der Technik Potenzialanalyse (s. Abschnitt 5.3.3), in der sie bereits Ausgabeformate der Anwendungen erfasst haben. Sie vermerken, ob die Ausgabe von Inhalten als HTML- oder XML-Datei möglich ist. Pro Informationsobjekt dokumentieren sie, ob es als Container fungiert, d.h. Informationsobjekte eines anderen Formats enthält oder auf solche verweist (s. Tabelle 5-16, S. 164).

Integrationsherausforderung 2: Verschiedene Ausgabeschnittstellen

Analog zu den Ausgabeformaten analysieren die IR-Verantwortlichen die Ausgabeschnittstellen der Content-Applikationen. Diese Schnittstellen können Spezifika aufweisen, welche die Möglichkeiten zur Integration der Inhalte durch eine Suchmaschine limitieren. Tabelle 5-23 zeigt vier unterschiedliche Ausgabeschnittstellen, über die eine Portalsuchmaschine unter Verwendung ihrer Konnektoren auf Inhalte der Content-Applikationen zugreifen kann. Die Charakteristika umfassen auch spezifische Limitationen:

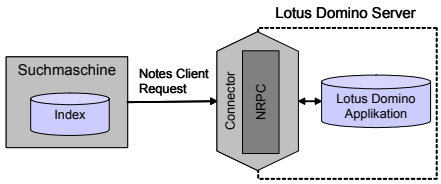
Konnektor	Charakteristika
HTTP(S)	
 • Hypertext Transfer Protocol (HTTP) ist ein zustandsloses Protokoll für die Übertragung von Hypermedia-Informationen. Es wird von Web-Browsern verwendet, um auf Web-Server zuzugreifen. • Die Suchmaschine stellt als Client eine Anfrage (HTTP-Request) an den Web Server. • Der Web Server sendet das durch die von der Suchmaschine übermittelte URL bezeichnete Dokument als Antwort (Response) zurück. • Bezeichnet die URL ein Programm, das dynamisch Daten erzeugt, z.B. ein Servlet, dann werden die Inhalte der Programmausgabe als Antwort zurückgeschickt [vgl. Balzert 1996, 956]. • Die Suchmaschine indiziert die als Antwort übermittelten Inhalte. • HTTPS stellt durch eine ergänzende Verschlüsselung über Secure Sockets Layer (SSL) eine gesicherte HTTP-Verbindung her.	• Die Informationsobjektformate sind durch die HTTP-Schnittstelle nicht limitiert, da HTTP beliebige Daten überträgt, z.B. eine HTML-Datei, ein Word-Dokument oder eine PDF-Datei. • Die Indizierung von Strukturinformationen ist möglich, sofern diese in HTML-Meta-Tags hinterlegt sind. • Zugriffsbeschränkungen verwaltet der Web Server, z.B. auf Verzeichnis- oder Informationsobjektebene. Greift die Suchmaschine auf diese Inhalte zu, muss sie sich gegenüber dem Web Server authentifizieren. • Authentifzierungsangaben können im HTTP-Konnektor der Suchmaschine durch Angabe von Benutzernamen und Passwort hinterlegt werden. Verlangt der Web-Server eine Authentifizierung, werden diese Daten automatisch übertragen. Das Verfahren ist auf die Angabe eines gültigen Benutzernamens beschränkt. Die automatische Übermittlung funktioniert nur, wenn der Web Server eine HTTP Basic Authentication anfordert (s. Tabelle 5-18, S.173).
Dateisystem / Fileserver	
 • Ein Dateisystem ist ein Bestandteil eines Betriebssystems, der Dateien in hierarchisch organisierten Verzeichnissen verwaltet. • Die Suchmaschine greift über eine dateisystemspezifische Programmierschnittstelle (Application Programming Interface, API) auf das Dateisystem eines logischen Laufwerks zu. Sie liest dabei die im Dateisystem enthaltenen Verzeichnisse und Dateien aus. • Der File Server übermittelt die von der Suchmaschine angeforderten Verzeichnisse und Dateien. • Die Suchmaschine indiziert die Dateien des Dateisystems.	• Marktübliche Suchmaschinen decken eine Vielzahl von Dateisystemen ab, z.B. das New Technology File System (NTFS) von Windows NT/2000/XP und UNIX-Systeme. • Die Suchmaschine kann sich im Bedarfsfall gegen die Benutzerverwaltung des Dateisystems authentifizieren, muss dafür jedoch einen dem Dateisystem entsprechenden Konnektor aufweisen. • Das Informationsobjektformat hängt nicht von der Dateisystem-Schnittstelle ab. • Sind Strukturinformationen in den indizierten Dateien vorhanden, so können diese indiziert werden. Aufgrund der Vielfalt der möglichen Dateiformate kann die Auswahl der Attribute nicht beeinflusst werden. Die Suchmaschine indiziert diese automatisch. • Die Anzeige eines Suchergebnisses erfolgt im HTML-Format. Die Suchmaschine muss alle Dokumente bei der Anzeige im Portal in HTML konvertieren.

<table>
<tr><th colspan="2">ODBC</th></tr>
<tr>
<td>


Open Database Connectivity (ODBC) ist eine standardisierte Schnittstelle für den Zugriff auf relationale Datenbanksysteme.
Ursprünglich von Microsoft spezifiziert, mittlerweile betriebssystemübergreifender De-facto-Standard spezifiziert ODBC ein API, das Funktionen zum Öffnen und Verwalten einer Datenbank-Verbindung, zum Senden von Anfragen und zur Auswertung der gelieferten Daten zur Verfügung stellt [s. Balzert 1996, 1017-1019].
Die Suchmaschine stellt über ODBC-Befehle eine Anfrage an die Datenbank.
Die Datenbank liefert die angefragten Inhalte, die von der Suchmaschine indiziert werden.

</td>
<td>

Über entsprechende ODBC-Treiber der Hersteller kann eine Vielzahl von Datenbanken angesprochen werden, z.B. Microsoft Access, MySQL oder Oracle Datenbanken.
Die Suchmaschine kann sich im Bedarfsfall gegen die Benutzerverwaltung der Datenbank authentifizieren. Dabei kann lediglich ein gültiger Benutzername übermittelt werden.
Die Anzeige eines Suchergebnisses erfolgt im HTML-Format. Die Suchmaschine muss relevante Datenbankinhalte bei der Anzeige im Portal in HTML konvertieren.

</td>
</tr>
<tr><th colspan="2">Lotus Notes</th></tr>
<tr>
<td>


Von IBM / Lotus spezifiziert, stellt der Notes Remote Procedure Call (NRPC) eine proprietäre Schnittstelle für den Zugriff auf Lotus Domino Applikationen bereit.
Die Suchmaschine stellt einen Notes Remote Procedure Call an den Domino Server.
Der Domino Server übermittelt die angeforderten Inhalte, die von der Suchmaschine indiziert werden.
Lotus Domino Server erlauben neben dem Zugriff über NRPC auch den Datenbankzugriff über HTTP(S) oder ODBC.

</td>
<td>

Der Zugriff über NRPC erfordert neben einem Notes-Connector auch einen installierten Notes Client oder entsprechende Laufzeitbibliotheken auf der Suchmaschine.
Die Suchmaschine benötigt für den Zugriff auf Inhalte eine gültige Notes-ID.
Die Indizierung setzt eine Zugriffsberechtigung auf die zu indizierenden Dokumente voraus.
Die Informationsobjektformate sind durch den Notes-Connector nicht limitiert, da NRPC beliebige Daten überträgt, z.B. ein Notes Dokument, ein Word-Dokument oder eine PDF-Datei.
Indizierung von Strukturinformationen erfolgt durch das Auslesen von Feldern eines Notes-Dokuments.

</td>
</tr>
</table>

Tabelle 5-23: Funktionsweise und Charakteristika unterschiedlicher Anwendungsschnittstellen

Viele Suchmaschinenhersteller berücksichtigen auch weitere proprietäre Schnittstellen z.B. über Konnektoren für Microsoft Exchange, Documentum, Opentext oder Oracle.

Die Projektmitarbeiter ergänzen die bereits verfeinerte Applikationsbeschreibung aus der Technik Systementwurf um eine Liste der für ihr Einsatzgebiet notwendigen Konnektoren zur Anbindung der Content-Applikationen. Sind Zugriffsbeschränkungen auf die Inhalte vorhanden, so halten die Applikationsverantwortlichen diese für jede Schnittstelle unter Angabe von Benutzerverzeichnis, Berechtigungsgruppen bzw. Berechtigungskonzept fest.

Integrationsherausforderung 3: Divergierende Datenstrukturen

Unterschiedliche Datenmodelle führen i.d.R. zu unterschiedlicher Bezeichnung und Bedeutung der auszutauschenden Daten (s. [Österle 1996, 13], [Puschmann 2003, 140]). Informationsobjekte mit inhaltlich gleicher Bedeutung werden häufig in verschiedenen Content-Applikationen mit unterschiedlichen Datenstrukturen gespeichert. So sind zum Beispiel in Content-Applikation A die Autorennamen im Feld ‚Authors' und der Titel des Informationsobjekts im Feld ‚Title', in Content-Applikation B in ‚Autoren' bzw. ‚Titel', in Content-Applikation C in ‚Creator' bzw. ‚Subject' usw. gespeichert. Komplexere Suchanfragen, die auf einer attributbasierten Suche des Anwenders basieren (s. erweiterte Suche, S. 161) oder auch der strukturierte Aufbau einer Suchergebnisliste (s. Suchergebnisanzeige, S. 163) setzen jedoch ein einheitliches Retrieval-Datenmodell voraus.

Die Projektmitarbeiter vervollständigen in Zusammenarbeit mit den Applikationsverantwortlichen die Applikationsbeschreibung (IST) aus der Technik Systementwurf um eine Darstellung der jeweiligen Datenmodelle der Content-Applikation.

Die summarische Betrachtung der drei Integrationsherausforderungen (unterschiedliche Ausgabeformate, verschiedene Ausgabeschnittstellen, divergierende Datenstrukturen) lässt den IR-Verantwortlichen zwei grundsätzliche Entscheidungsalternativen für die Wahl eines Integrationsvorgehens:

- *Physische Integration.* Inhalte aus Content-Applikationen sind physisch integriert, wenn sie in einer gemeinsamen Datensammlung abgelegt sind. Ein solcher Zustand lässt sich entweder erzielen, wenn vor der Implementierung der Information-Retrieval-Lösung keine einschlägigen Vorläufersysteme existieren, oder wenn alle relevanten Altinhalte einmalig in eine neue Content-Applikation migriert werden. Die physische Integration ist auch unter Berücksichtigung der Aufwendungen für Standardisierung und Migration durchweg empfehlenswert, da hier standardisierte Datenstrukturen sowie eine überschaubare Anzahl von Ausgabeformaten und Ausgabeschnittstellen entstehen.

> *Die Ablage auf dem Fileserver von IWI-HSG war gekennzeichnet durch eine uneinheitliche Ablagestruktur, eine Vielzahl ‚alter' Dokumentversionen sowie eine sehr niedrige Qualität der Metadaten. Für die strukturierte Navigation und Suche im Portal hätten diese zuvor manuell nachbearbeitet werden müssen. Die Projektverantwortlichen entschieden daher, erhaltenswerte Dokumente in eine neue Content-Applikation zu migrieren. Sie reduzierten damit die Anzahl der Schnittstellen zwischen der Suchmaschine und den Inhalten auf einen Anwendungstyp.*

- *Logische Integration.* Um verschiedenartig strukturierte Informationsobjekte in getrennten Datensammlungen über eine Portalsuche erschliessbar zu machen, müssen diese zunächst auf eine gemeinsame Datenstruktur, das so genannte Retrieval-Datenmodell zusammengeführt werden. Ausgehend von einer Datenstruktur je Informationsobjekttyp erstellen die Systemverantwortlichen für jede Content-Applikation so genannte Mappings. Diese bilden relevante Datenmodellbestandteile der Anwendungen auf die SOLL-Struktur der Suchmaschine ab (vgl. [Jansen 2000, 186], [Riehm 1997, 16], [Alt et al. 2002, 105]).

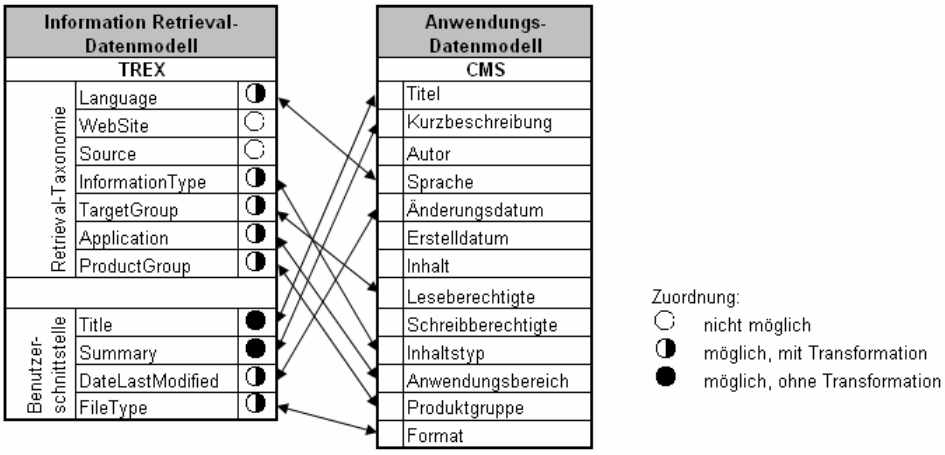
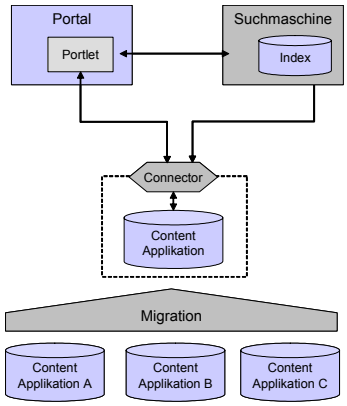


Abbildung 5-28: Schemaintegration

Tabelle 5-24 beschreibt die Unterschiede der beiden Verfahren. Die Charakteristika dienen den Projektverantwortlichen als Entscheidungsgrundlage.

Integrationsverfahren	Charakteristika
Physische Integration	
	• Die ursprünglichen Content-Applikationen existieren nach der Migration in eine neue führende Content-Applikation nicht mehr. • Die neue Content-Applikation verfügt über eine einheitliche Struktur, identische Informationsobjekt-formate und eine gemeinsame Benutzerverwaltung.

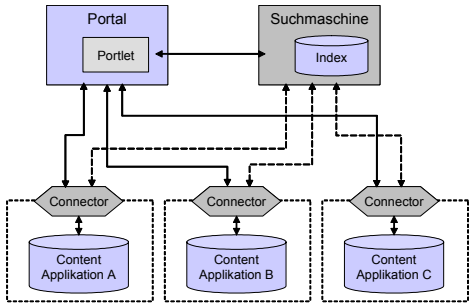
Logische Integration	
	• Die ursprünglichen Content-Applikationen mit eigenen Strukturen, Informationsobjektformaten und Benutzerverwaltungen bleiben bestehen. • Die spezifischen Konnektoren müssen pro Content-Applikation aufgebaut und betrieben werden.

Tabelle 5-24: Datenfluss, Charakteristika und Praxisfälle unterschiedlicher Integrationskonzepte

Die Integrationsherausforderungen und Entscheidungsalternativen dienen den Projektinitiatoren zur Abstimmung mit den Betreibern der Content-Applikationen und als Grundlage für die spätere Softwareauswahl. Die Projektverantwortlichen verfeinern und ergänzen im Rahmen des Systementwurfs die bisherige Applikationsbeschreibung.

5.6.6. Indizierung

Indizierung (auch Indexierung) bezeichnet die Zuordnung von inhaltsbeschreibenden Begriffen, so genannten Deskriptoren, zu einem Dokument [s. Nohr 2000, 19]. Die Deskriptoren repräsentieren den Inhalt eines Dokuments und dienen zu dessen Erschliessung, z.B. durch Navigation in oder Suche nach Begriffen. Die automatische Indexierung von Content-Applikationen durch Suchmaschinen unterscheidet folgende Verfahren (s. [Baeza-Yates/Ribeiro-Neto 1999, 191], [Anderson/Pérez-Carballo 2001]):

- *A) Volltextindizierung*. Die Suchmaschine indiziert die in Informationsobjekten enthaltenen Textbestandteile. Strukturinformationen bleiben hingegen unberücksichtigt. Der Index enthält die indizierten Begriffe mit zugehörigen Verweisen auf die jeweiligen Fundorte. Die Retrieval-Leistung ist bei dieser Vorgehensweise auf eine reine Volltextsuche beschränkt.

- *B) Attributbasierte Indizierung*. Die Suchmaschine extrahiert Inhalte in Kombination mit ihren Strukturelementen, z.B. Feldinhalte von Datensätzen, Meta-Tags aus HTML-Dokumenten oder Attribute in XML-Dateien. Diese Vorgehensweise erlaubt durch die Bereitstellung einer erweiterten Suche (s. S. 162) eine im Vergleich zu A) sehr viel genauere Definition der Suchanfrage. Für die attributbasierte Indizierung benötigt die Suchmaschine detaillierte Informationen über die Datenstruktur aller zu indizierenden Quellen. Unterschiedliche Datenstrukturen müssen in den Konnektoren der Suchmaschine

über eine Schemaintegration (s. Abbildung 5-28) abgeglichen werden, die relevante Bestandteile auf ein einheitliches Retrieval-Datenmodell abbildet.

Tabelle 5-25 illustriert den grundsätzlichen Ablauf zur Vorbereitung einer attributbasierten Indizierung am Beispiel des Verity Information Servers im Rahmen einer Evaluation bei IWI-HSG.

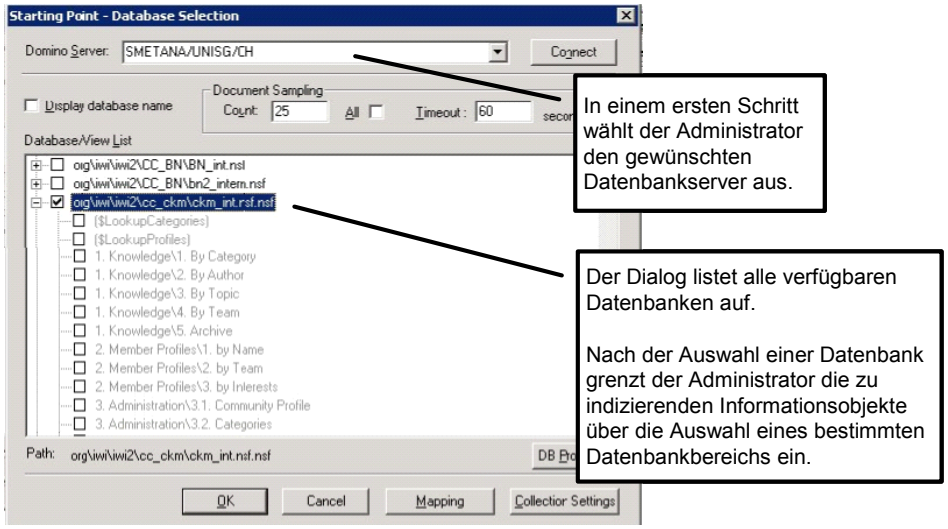
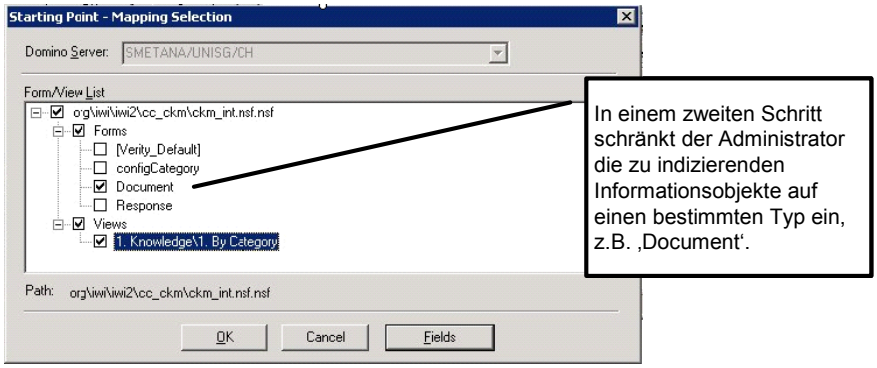
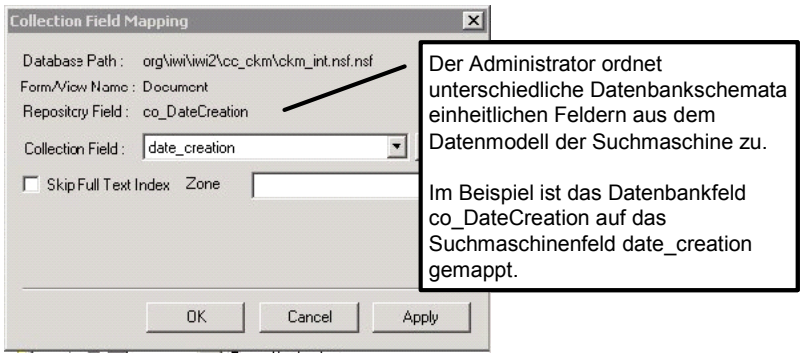
Schritt	Einstellungen im Verity Information Server
1. Anlegen einer neuen Content-Applikation über den Lotus-Notes-Connector	
2. Spezifikation der zu indizierenden Dokumente über deren Eigenschaften	
3. Zuordnung von Feldern der zu indizierenden Applikationen zu einheitlichen Feldern aus dem Datenmodell der Suchmaschine (so genannte Schemaintegration)	

Tabelle 5-25: Attributbasierte Indizierung mit dem Verity Information Server (Evaluation bei IWI-HSG)

Der Aufwand zur Implementierung und Pflege einer attributbasierten Indizierung ist im Vergleich zur Volltextindizierung erheblich höher. Die Projektverantwortlichen

entscheiden daher anhand der folgenden Checkliste, ob eine attributbasierte Indizierung notwendig ist.

	Fragestellung	Technik	Ergebnis
Notwendigkeit für attributbasierte Indizierung	Ist die Bereitstellung einer attributbasierten Suche oder einer taxonomieorientierten Navigation gefordert?	Strategieplanung	Retrieval-Leistungen (s. S. 133)
	Beinhaltet die Retrieval-Taxonomie Attribute, anhand derer die Benutzer Suchen oder Navigieren sollen?	Prozessplanung	Retrieval-Taxonomie (s. S. 147)
	Ist eine erweiterte Suche in Strukturinformationen der Content-Applikationen geplant?	Systementwurf	Erweiterte Suchanfrage (s. S. 161)
	Sollen die inhaltsbeschreibenden Attribute, die bei der Anzeige von Suchergebnissen für jeden Treffer angezeigt werden aus den Strukturinformationen der indizierten Content-Applikationen ausgewählt und zugeordnet werden können?	Systementwurf	Suchergebnisanzeige (s. S. 163)
	Erfordert die Berücksichtigung von Lese-beschränkungen der Content-Applikationen die Abbildung dieses Zugriffskonzepts in den Attributen der indizierten Inhalte?	Systementwurf	Autorisierungsebenen (s. S. 183)

Tabelle 5-26: Checkliste für die Auswahl eines Indizierungsverfahrens

Die Systemverantwortlichen halten die Entscheidung für ein Indizierungsverfahren in der Kriterienliste für die Suchmaschinenauswahl fest.

5.6.7. Klassifikation

Klassifikation dient im Information Retrieval dazu, Informationsobjekte systematisch zu ordnen [s. Ferber 2003, 47]. Die so genannte Taxonomie, ein hierarchisches Klassifikationsschema (Klassifikationssystem), ermöglicht einem Portalbenutzer dabei die strukturierte Erschliessung einer grossen Menge von Informationsobjekten aus heterogenen Quellen anhand homogener Klassen oder Kategorien (s. [Gaus 2003], [Bailey 1994], [Logan 2001]). Die Navigation im Produktkatalog der Firma Zumtobel Staff mit seinen Produktklassen (s. Fallstudie in Abschnitt 4.1) oder die kategorienbezogene Suche im Portal der Winterthur Versicherungen (s. Fallstudie in Abschnitt 4.3) dienen als Beispiele.

Die Retrieval-Taxonomie der Technik Prozessplanung (s. Abschnitt 5.5) gibt den Systemverantwortlichen bereits das Klassifikationssystem bzw. die zu implementierenden Klassen vor. Die zu klassifizierenden Objekte der an die Suchmaschine angebundenen Content-Applikationen müssen dabei Eigenschaften aufweisen, anhand derer sie von Menschen oder Systemen in Klassen eingeordnet werden können. Auf Systemebene stehen dazu unterschiedliche Klassifikationsverfahren zur Verfügung:

- *A) Manuelle Klassifikation*. Der Ersteller eines Dokuments ordnet manuell bestehende oder neue Inhalte einer oder mehrere Klassen zu. Die Angabe der Klassenzugehörigkeit kann dabei im Dokument selbst gespeichert werden, z.B.

über Meta-Tags, erfordert dann aber eine attributbasierte Indizierung (s. S. 192). Alternativ können die jeweiligen Klassen Informationen verwalten, welche Dokumente zu ihnen gehören.

Auf Basis der von weltweit 20'000 freiwilligen Editoren gepflegten Taxonomie des 'Open Directory Project' sind im Index der Internet-Suchmaschine Google ca. 1,5 Mio. Internetseiten (Informationsobjekte) in über 460'000 Kategorien abgelegt [DMOZ 2002]. Ausgehend von 16 Hauptkategorien listet Google die Startseite der Firma Zumtobel Staff beispielsweise in der Hauptkategorie ‚Wirtschaft' und den Unterkategorien ‚Elektrotechnik und Leuchten > Lampen und Leuchten' auf.

Abbildung 5-29: Ausschnitt aus der Taxonomie der Internet-Suchmaschine Google

- *B) Automatische Klassifikation*. Im Gegensatz zu A) erstellt die Suchmaschine bei der automatischen Klassifikation selbständig Klassen sowie Bezeichnungen und ordnet volltextindizierte Dokumente den Klassen zu. Die automatische Klassifikation arbeitet nach einem so genannten ‚Black-Box'-Verfahren, d.h. die Klassifikationsregeln sind weder für Anwender noch für Administratoren ersichtlich. Während Klassenbezeichnungen von Administratoren leicht geändert werden können, erfordert die Kalibrierung der Klassenzugehörigkeit ein zeitaufwändiges Training über Lernmengen (ausgewählte Beispieldokumente), die manuell angelegt und zugeordnet werden müssen [vgl. Gilchrist 2001]. Verity K2 Enterprise, IBM Lotus Discovery Server und Discovery der Firma 80-20 Software sind Beispiele für automatische Klassifizierer.

Der ‚K-map Editor' ist das Administrationswerkzeug zur Modifikation der vom IBM Lotus Discovery Server automatisch erstellten Klassifikation. Abbildung 5-30 zeigt im linken Fenster die Baumstruktur der Taxonomie von IWI-HSG. Die Klassenbezeichnungen sind in diesem Laborversuch bereits manuell nachbearbeitet

worden. Ein Administrator kann die Dokumente der jeweiligen Kategorie, dargestellt im rechten Fenster, in eine oder mehrere andere Kategorien im linken Fenster per Drag&Drop verschieben oder auch Löschen.

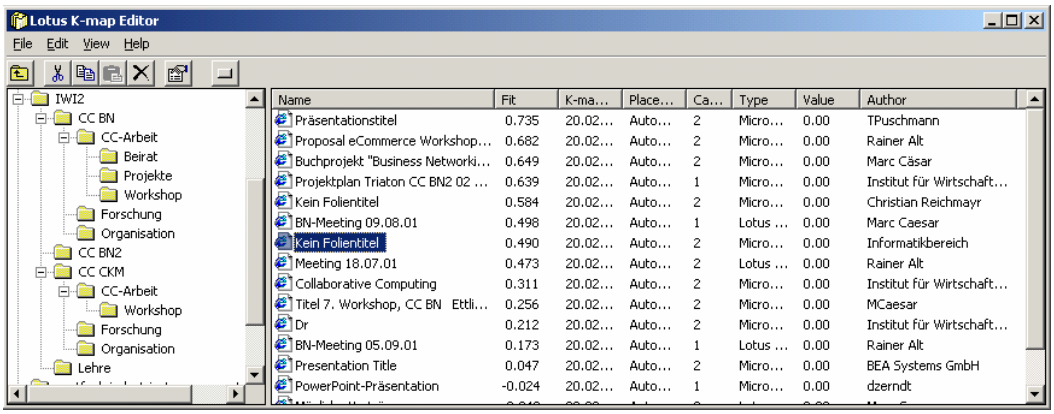


Abbildung 5-30: ‚K-map Editor' des IBM Lotus Discovery Server

Durch die Verschiebung von Dokumenten in andere Kategorien passt der Discovery Server seine Klassifikationsregeln an. Bei einer Reklassifizierung vorhandener Inhalte oder beim Hinzufügen neuer Dokumente in den Suchmaschinenindex finden die veränderten Regeln Anwendung.

- *C) Semi-automatische Verfahren*. Sie erlauben eine manuelle Kontrolle und Steuerung der automatischen Klassifikation per Einsicht und Definition von Regeln zur Klassenbildung durch einen Administrator. Diese Regeln führen z.B. pro Klasse auf, welche Begriffe oder Begriffskombinationen in einem indizierten Dokument enthalten sein müssen, damit es zu dieser Klasse gehört. Auch komplexere Operationen, wie der Ausschluss bestimmter Begriffe (NOT), Angaben zu Begriffspositionen innerhalb eines Dokuments (NEAR) oder zu Attribut/Wert-Kombinationen (z.B. Produktgruppe=‚Anbau- und Pendelleuchten'). Letztere setzt, analog zu A), eine attributbasierte Indizierung voraus.

Mit dem ‚eClassifier' bietet IBM ein, zu diesem Zeitpunkt nicht kommerziell vertriebenes Administrationswerkzeug zur Manipulation der automatischen Klassifikation an. Abbildung 5-31 zeigt die Anpassung der Klassifikationsregel bei IWI-HSG für die Kategorie ‚Supply Chain Management' über Wortfilter auf Volltextbasis.

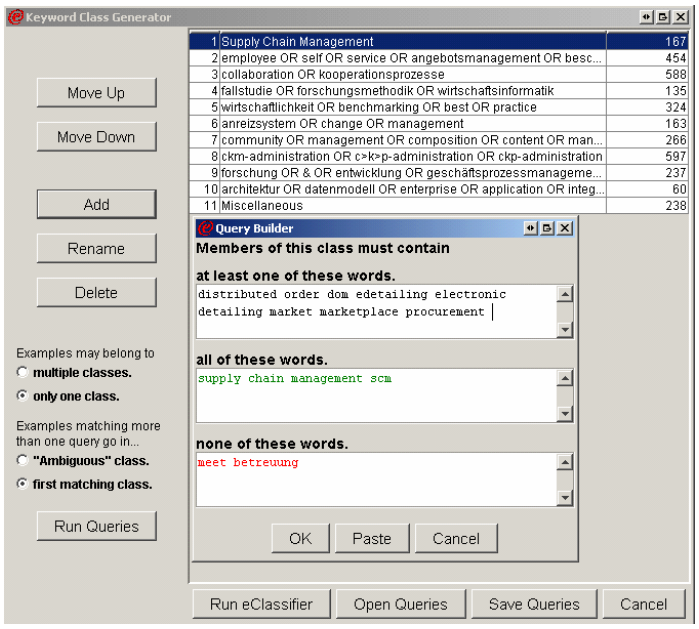


Abbildung 5-31: ‚eClassifier' des IBM Lotus Discovery Server

Während A) das Expertenwissen der Anwender nutzt, ist es bei grossen Informationsbeständen jedoch auch mit Kosten für die manuelle Pflege verbunden und verlangsamt den Prozess der Informationsbereitstellung für die Benutzer [vgl. Murray 2001]. Es ist dennoch geeignet, wenn andere Verfahren nicht anwendbar sind oder eine möglicherweise fehlerhafte Kategorisierung eines automatischen Klassifizierers nicht akzeptabel ist.

Der rein automatische Ansatz von B) bringt nicht immer zufriedenstellende Ergebnisse hervor (s. [Andrews 2003a], [Warzecha 2001], [Hagen 2000], [Verity 2003], [Christ 2002, 133]). [Riempp 2003, 229] und [Gebert 2003, 261-266] beschreiben Laborversuche, bei denen die automatische Extraktion der auf Basis einer manuell erstellten Taxonomie klassifizierten Inhalte deutliche Abweichungen zwischen der ursprünglichen Taxonomie und dem Extrakt aufweist. Das Verfahren ist dann geeignet, wenn bestehende Inhalte nicht klassifiziert sind bzw. nicht mit vertretbarem Aufwand klassifiziert werden können, den Benutzern aber dennoch Dokumentenklassen zur Navigation und Suche angeboten werden sollen. Aufgrund der beschriebenen Ungenauigkeiten der Klassifikationsergebnisse sollte dieses Verfahren ausschliesslich für den unternehmensinternen Einsatz verwendet werden.

Die Leistungsfähigkeit von Klassifikationssystemen mit dem Expertenwissen der Anwender kombiniert C) und erreicht in der Regel exaktere Ergebnisse als vollautomatische Verfahren (s. [Gebert 2003, 261-266], [Raghavan 2002]). Auch hier ist die nachträgliche Anpassung eines falsch klassifizierten Dokumentenbestandes mit Aufwand verbunden [s. Christ 2002, 138], dieser ist jedoch geringer als bei B). Das Verfahren ist gut geeignet, wenn die zu integrierenden Content-Applikationen strukturierte Inhalte aufweisen, die jedoch in unterschiedlichen Datenstrukturen abgebildet sind und eine Schemaintegration erfordern (s. S. 191). Die regelbasierte

Zuordnung heterogener Datenfelder der Anwendungen zu einem einheitlichen Retrieval-Datenmodell stellt sicher, dass ein konsistentes Klassifikationsgerüst ohne Anpassung der Datenmodelle der Content-Applikationen bereitgestellt werden kann.

Folgende Ergebnisse des vorliegenden Methodenvorschlags unterstützen die Prüfung der Systemverantwortlichen, ob eine manuelle Klassifikation der Informationsobjekte notwendig bzw. realisierbar ist:

	Fragestellung	Technik	Ergebnis
Notwendigkeit	Benötigen die Anwender zur Unterstützung ihrer Prozesse eine attributbasierte Suche?	Strategieplanung	Retrieval-Leistungen (s. S. 133)
	Sind Angaben bezüglich Zugriffsbeschränkungen in den Dokumenten der Content-Applikationen bisher nicht vorhanden oder können diese im Berechtigungskonzept der Suchmaschine nicht abgebildet werden?	Systementwurf	Autorisierungsmechanismen (s. S. 183)
Realisierbarkeit	Sieht der Aufbau und Betrieb der Retrieval-Taxonomie im Unternehmen entsprechende Rollen vor?	Prozessplanung	Pflegeprozesse (s. S. 145)
	Lassen die Verantwortlichen der zu integrierenden Content-Applikationen Modifikationen zu?	Systementwurf	Integrationsverfahren (s. S. 184)
	Erlauben Formate und Datenstrukturen der Anwendungen eine Klassifikation?	Systementwurf	Integrationsverfahren (s. S. 184)

Tabelle 5-27: Prüfung von Notwendigkeit und Realisierbarkeit einer manuellen Klassifikation

Zur manuellen Klassifikation von Inhalten müssen die Ersteller oder Pflegeverantwortlichen aus ihren gewohnten Applikationen auf eine einheitliche Klassifikation zugreifen können. Die Verwaltung der Klassifikationsmerkmale sollte daher in einem zentralen System erfolgen. Dezentrale Pflegbarkeit, Berechtigungen, Änderungen des Klassifikationsverwaltungssystems und auch Integration bestehender Klassifikationen von Content-Applikationen sind bereits durch die Prozessplanung geregelt (s. Abschnitt 5.5).

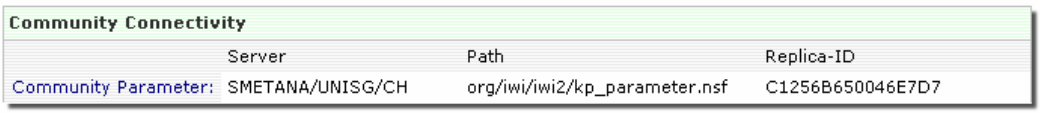
Aus Systemsicht müssen die Verantwortlichen in Abhängigkeit vom benötigten Funktionsumfang und den möglichen technischen Schnittstellen entscheiden, ob als zentrales Klassifikationsverwaltungssystem:

- eine bestehende Content-Applikation als führendes System genutzt werden kann,
- eine neue Applikation implementiert werden muss oder
- diese Rolle der Suchmaschine zugeschrieben werden soll.

Den Zugriff von Content-Applikationen auf das Klassifikationsverwaltungssystem halten die Verantwortlichen im Systementwurf über Schnittstellenbeschreibungen (SOLL) fest.

> *Die Information-Retrieval-Lösung bei IWI-HSG verwendet zur zentralen Verwaltung der Klassifikationsmerkmale eine zu diesem Zweck implementierte*

Applikation. Diese so genannte ‚Taxonomy-Database' basiert, wie die weiteren Portalkomponenten, auf der Groupwareplattform Lotus Notes. Sämtliche in die Portalsuche integrierten Content-Applikationen greifen zur Referenzierung von Dokumenten über eine Schnittstellendefinition (‚Community Parameter') auf diese Anwendung zu.



Community Connectivity			
	Server	Path	Replica-ID
Community Parameter:	SMETANA/UNISG/CH	org/iwi/iwi2/kp_parameter.nsf	C1256B650046E7D7

Abbildung 5-32: Schnittstellendefinition für den Zugriff auf das zentrale Klassifikationsverwaltungssystem bei IWI-HSG

5.7. Konsolidierung und abschliessende Schritte

Die folgenden Abschnitte stellen keine eigene Technik im eigentlichen Sinne des Vorgehensmodells dar. Sie skizzieren jedoch die weitere Vorgehensweise vom Abschluss des Systementwurfs bis zur Bereitstellung des operativen Systems. Aufgrund der grossen Varianz hinsichtlich der technischen, organisatorischen, kulturellen und finanzwirtschaftlichen Unterschiede bei der eigentlichen Implementierung können die Schritte ‚Software evaluieren', ‚System implementieren' und ‚Organisation entwickeln' nur schwer generalisiert werden. Die vorliegende Arbeit greift jedoch spezifische Aspekte, die den erfolgreichen Abschluss eines Information Retrieval Projekts unterstützen.

5.7.1. Software evaluieren

Zur Abschätzung der Anschaffungskosten und zur Beurteilung der Umsetzbarkeit der groben fachlichen Anforderungen haben die Projektverantwortlichen in der Technik Strategieplanung (s. Abschnitt 5.4.3) bereits eine Softwarevorauswahl durchgeführt. Die so genannte ‚Short list' grenzt die Realisierungsalternativen auf die Produkte einiger weniger Suchmaschinenhersteller ein. Die Auflistung der Leistungsmerkmale und Kosten von 3 bis 7 Anbietern erlaubt in der Strategieplanung einen positiven Projektentscheid. Die endgültige Auswahl ist jedoch erst möglich, wenn auch die Prozessplanung und der Systementwurf abgeschlossen sind.

Tabelle 5-28 liefert den Projektverantwortlichen Kriterien für ein fundiertes Auswahlverfahren. Die Checkliste lehnt sich an die Funktionskataloge verschiedener Softwarehersteller an und wurde beispielsweise in den vorliegenden Praxisfällen (s. Kapitel 4) herangezogen. Sie enthält neben funktionalen auch die für eine Suchmaschinenauswahl relevanten nicht-funktionalen Kriterien. Die dritte Tabellenspalte verweist für jedes funktionale Kriterium auf relevante Abschnitte dieser Arbeit. Die dort aufgeführten Techniken und Ergebnisse dienen den Projektverantwortlichen bei der Softwareauswahl als ‚Nachschlagewerk'.

Anforderung	Beschreibung	Technik & Ergebnis
Indizierung & Crawler		
Web Crawler	Indizierung von Quellen über HTTP-Schnittstelle	*Potenzialanalyse:* Applikationen erfassen (s. S. 128) *Strategieplanung:* Architekturübersicht erstellen (s. S. 136) *Prozessplanung:* Anwendungen integrieren (s. S. 156) *Systementwurf:* Integration (s. S. 184) Indizierung (s. S. 192)
File Server Crawler	Indizierung von Dateisysteminhalten	
Crawler für lokales Dateisystem	Indizierung des lokalen Dateisystems vom Portal oder Suchmaschine	
Lotus Notes Crawler	Indizierung von Lotus Notes Datenbanken	
Attachments	Indizierung von Dateianhängen	
Komprimierte Dokumente	Indizierung von ZIP-Dateien	
MS Office Dokumente	Indizierung von Word-, Excel- und Powerpoint-Dateien	
Adobe PDF Dokumente	Indizierung von PDF-Dateien	
Architektur & Integration		
Betriebssystem	Zur Installation der Suchmaschine benötigtes Betriebssystem	*Strategieplanung:* Architekturübersicht erstellen (s. S. 136) *Prozessplanung:* Anwendungen integrieren (s. S. 156) *Systementwurf:* Integration (s. S. 184)
Erforderliche Softwarekomponenten	Für Installation und Betrieb der Suchmaschine benötigte Komponenten, z.B. Web-Server	
Integration in bestehendes Portal	Integrationstechnologien zur Einbindung von Such- und Navigationsfunktionen in eine bestehende Portalplattform	
LDAP-Integration	Option der Suchmaschine, ein zentrales LDAP-Verzeichnis zur Benutzerverwaltung anzubinden	*Systementwurf:* Authentifizierung und SSO (s. S. 171) Benutzerverwaltung (s. S. 178) Autorisierung (s. S. 180)
Single Sign On	Mechanismen der Suchmaschine, eine bestehende Benutzerauthentifizierung zu übernehmen	
Berücksichtigung existierender Zugriffsbeschränkungen	Möglichkeiten der Suchmaschine, Berechtigungskonzepte der indizierten Content-Applikationen abzubilden	
Content Organisation		
Automatische Taxonomie	(hierarchische) Klassifikation kann automatisch erstellt und für attributbasierte Suche bereitgestellt werden	*Potenzialanalyse:* Handlungsoptionen entwickeln (s. S. 130) *Strategieplanung:* Leistungen festlegen (s. S. 133) *Prozessplanung:* Retrieval-Taxonomie bereitstellen (s. S. 147) *Systementwurf:* Suchanfrage (s. S. 161) Navigation (s. S. 167) Indizierung (s. S. 192) Klassifikation (s. S. 194)
Manuelle Taxonomie	(hierarchische) Klassifikation kann manuell erstellt und für attributbasierte Suche bereitgestellt werden	

Suchanfrage		
Volltextsuche	Suche nach Inhalten, in denen ein oder mehrere Suchbegriffe an beliebiger Stelle im Dokument vorkommen	*Potenzialanalyse:* Handlungsoptionen entwickeln (s. S. 130) *Strategieplanung:* Leistungen festlegen (s. S. 133) *Systementwurf:* Suchanfrage (s. S. 161) Indizierung (s. S. 192)
Attributbasierte Suche	Gezielte Suche auf einzelnen Attributen / Feldern der angebunden Content-Applikationen	
Logische Operatoren	Verknüpfung von mehreren Suchbegriffen über UND, ODER, NICHT	*Prozessplanung:* Retrieval-Prozess definieren (s. S. 146)
Platzhalter (Wildcards)	Suche mit Platzhaltern, z.B. Licht* für Suche nach Begriffen, die mit ‚Licht' beginnen	
Phrasensuche	Suche nach ganzen Sätzen	
Stemming	Automatische Rückbildung eines Suchbegriffs auf seinen Wortstamm, z.B. automatische Anzeige von Suchergebnisse für ‚Leuchte' bei einer Suchanfrage nach ‚Leuchten'	
Verfeinerung von Suchanfragen	Verfeinerung der Suchanfrage nach Anzeige der Suchergebnisse (erneute Suche in Suchergebnissen)	
Thesaurus	Suche nach Synonymen	
Mehrsprachige Suchergebnisse	Suchanfrage in der Sprache des Anwenders liefert auch andersprachige Suchergebnisse	
Editierbarer Thesaurus	Synonymlisten in der Suchmaschine organisatorisch pflegbar	
Suchergebnisse		
Suchanfragemaske	Möglichkeiten zur Anpassung	*Systementwurf:* Suchanfrage (s. S. 161)
Suchergebnisliste	Möglichkeiten zur Anpassung	*Systementwurf:* Suchergebnisanzeige (s. S. 163)
Gruppierung von Suchergebnissen	Automatische Gruppierung von gleichartigen Suchergebnissen	*Prozessplanung:* Retrieval-Prozess definieren (s. S. 146)
Sortierreihenfolge	Rangfolge der Suchergebnisse in der Liste	
Automatische Zusammenfassung (Auto Summary)	Zusammenfassung des gefundenen Dokuments durch die Suchmaschine	
Anzeige verbundener Dokumente (Related Documents)	Anzeige von ähnlichen Dokumenten zu einem Suchergebnis.	
Hervorheben von Suchbegriffen (Term Highlighting)	Hervorheben des Suchbegriffs bzw. der Suchbegriffe	*Systementwurf:* Suchergebnisanzeige (s. S. 163)

Performance & Skalierbarkeit		
Antwortzeiten	Zeitdauer von der Suchanfrage bis zur Anzeige der Suchergebnisliste	*Prozessplanung:* Retrieval-Prozess definieren (s. S. 146)
Maximale Anzahl gleichzeitiger Suchanfragen (Concurrent User)	z.B. 50 Concurrend User gegen einen Index von 2 x 20.000 Dokumenten	
Lastverteilung (Load Balancing)	Automatische Lastverteilung zwischen mehreren Suchmaschinen	- Herstellerspezifisch -
Client Typen		
Web Browser	Zugriff und Suche über einen Web Browser möglich.	*Prozessplanung:* Retrieval-Prozess definieren (s. S. 146)
Windows Client		
Java Client		
Verfügbarkeit & Support		
Verfügbarkeit, Produkthistorie	Produkt ist stabil verfügbar.	- Herstellerspezifisch -
Release-Planung		
Referenzinstallationen	Anbieter verfügt über Referenzinstallationen in vergleichbarer Größe.	
Support	Unterstützung des Anbieters bei Problemfällen Anbieter leistet telefonische Unterstützung (24x7)	

Tabelle 5-28: Kriterienliste für Suchmaschinenauswahl

5.7.2. System implementieren

In der Systemimplementierung setzt das Projektteam den technischen Systementwurf um. Sie installieren die auf Basis der funktionalen und nicht-funktionalen Anforderungen ausgewählte Suchmaschine und binden sie in das Portal ein. Gemäss den festgelegten Schnittstellen integrieren sie die relevanten Content-Applikationen. Nach erfolgreichen Tests geben die Projektverantwortlichen die neue Retrieval-Lösung für den Produktiveinsatz frei. Das Ergebnis der Systemimplementierung ist ein umgesetztes, abgenommenes und in den Regelbetrieb übergebenes Produktivsystem. Die abschliessenden Schritte der Implementierung sind:

- *Suchmaschine installieren und konfigurieren*. Hierbei wird die angeschaffte Software auf der bereitgestellten Hardware installiert. In der Praxis unterstützen entsprechend versierte Mitarbeiter des Suchmaschinenherstellers häufig diesen Schritt. Die Einbindung der Suchmaschine in das Unternehmensnetzwerk müssen die Projektverantwortlichen frühzeitig bei ihren Netzwerkadministratoren beantragen. Vor der eigentlichen Konfiguration der Suchmaschine sollte nach deren Installation eine vollständige Sicherung durchgeführt werden. Dieses Backup

nimmt zwar Zeit in Anspruch, kann im Bedarfsfall aber jederzeit zurückgespielt werden und eine Neuinstallation beschleunigen.

- *Anwendungen anpassen*. Die Anwendungsentwickler passen die beteiligten Content-Applikationen, das zentrale Benutzerverzeichnis und das Portal gemäss den Vorgaben Applikationsbeschreibung (SOLL) und Schnittstellenbeschreibung (SOLL) der Technik Systementwurf (s. Abschnitt 5.6) an. Trotz entsprechend genauer Erfassung und Vorgabe der Veränderungsbedarfe sollten die Projektverantwortlichen zusätzliche Zeiten für die Abstimmung mit den Anwendungsentwicklern einplanen. Eine zeitgleiche Einführung von mehreren neuen Anwendungen, z.B. Suchmaschine und Content-Applikation, sollten die Projektverantwortlichen vermeiden. Ein Information-Retrieval-Projekt sollte erst initiiert werden, wenn neu lizensierte Anwendungen implementiert sind.

- *Indizes anlegen*. Die erstmalige Indexierung von Inhalten kann sehr zeitaufwändig sein, u.U. sogar einen ganzen Tag beanspruchen. In dieser Zeit können und sollten keine anderen Arbeiten an der Suchmaschine vorgenommen werden. Um Erfahrungswerte zu sammeln, sollte das Projektteam am Anfang kleinere Testdatenbestände indizieren oder die Indizierung über Nacht laufen lassen. Die Projektverantwortlichen müssen die Zeitpunkte der Indizierung mit den Applikations- und Netzwerkverantwortlichen im Unternehmen koordinieren, da sowohl die Anwendungen als auch das Netzwerk durch den anfänglich intensiven Zugriff der Suchmaschine auf die Content-Applikationen erheblich belastet werden. Die initialen, aber auch die fortlaufenden, z.B. täglichen Indizierungsvorgänge, sollten in Zeiten durchgeführt werden, in denen die Anwendungen nicht intensiv genutzt werden. Zeitgleiche Backup- oder Datenreplikationsläufe müssen vermieden werden.

- *Evaluation mit Pilotanwendern durchführen*. Die Freischaltung einer neuen Information-Retrieval-Lösung sollte während der Implementierung in evolutionären Schritten erfolgen. Ausgewählten Anwendergruppen, so genannten Pilotanwendern oder ‚Power Usern' stellen die Projektverantwortlichen einzelne Funktionen in einem frühen Implementierungsstadium auf einem Testsystem zur Verfügung. Zeitgleich zu der Einarbeitung von Rückmeldungen können die Projektverantwortlichen diese Anwender auch für Lasttests des neuen Systems einbinden.

- *System- und Abnahmetests absolvieren*. Die einzelnen Systembestandteile sind bei der Installation und Konfiguration der Suchmaschine bzw. bei der Anpassung der Anwendungen getestet worden. Der System- und Abnahmetest überprüft das gesamte neue Informationssystem. Dabei stellen die Projektverantwortlichen nicht

nur sicher, dass alle Anforderungen korrekt erfüllt sind. Der System- und Abnahmetest beinhaltet insbesondere auch Startup- und Shutdown-Prozeduren, Backup- und Recovery-Verfahren und Prüfungen des Antwortzeitverhaltens.

- *System freischalten und in den Regelbetrieb übergeben*. Die Übergabe in den Regelbetrieb leitet den Abschluss des Projekts ein. Das Projektteam muss dazu beispielsweise auch eine Benutzer- und Betreiberdokumentation erstellen sowie Entwicklungs- und Administrationswerkzeuge übergeben.

- *Nutzungsmessung durchführen*. In der Nutzungsmessung beurteilen die Verantwortlichen den Projekterfolg auf Basis der erreichten Ziele. Grundlage dazu ist beispielsweise die Verfolgung, Analyse und Auswertung des Nutzerverhaltens mit Hilfe von so genannten Logfiles, Cookies oder Session-IDs [s. Merz 2002, 520]).

- *Veränderungen kommunizieren*. Die Projektverantwortlichen bereiten den Produktivbetrieb durch Schulungsmassnahmen und die Kommunikation der Veränderungen vor (s. [Gebert 2003, 207], [Thiesse 2001, 178]. Diese projektbegleitenden Aktivitäten beeinflussen die organisatorischen und kulturellen Rahmenbedingungen und sind eine wesentliche Voraussetzung für die umfassende Nutzung eines neuen Systems [s. Österle et al. 1992, 313].

5.8. Dokumentationsmodell

Das Dokumentationsmodell stellt die im Methodenvorschlag verwendeten Ergebnisdokumente und deren Beziehungen zum integrierten Metamodell für Information Retrieval in Portalen dar.

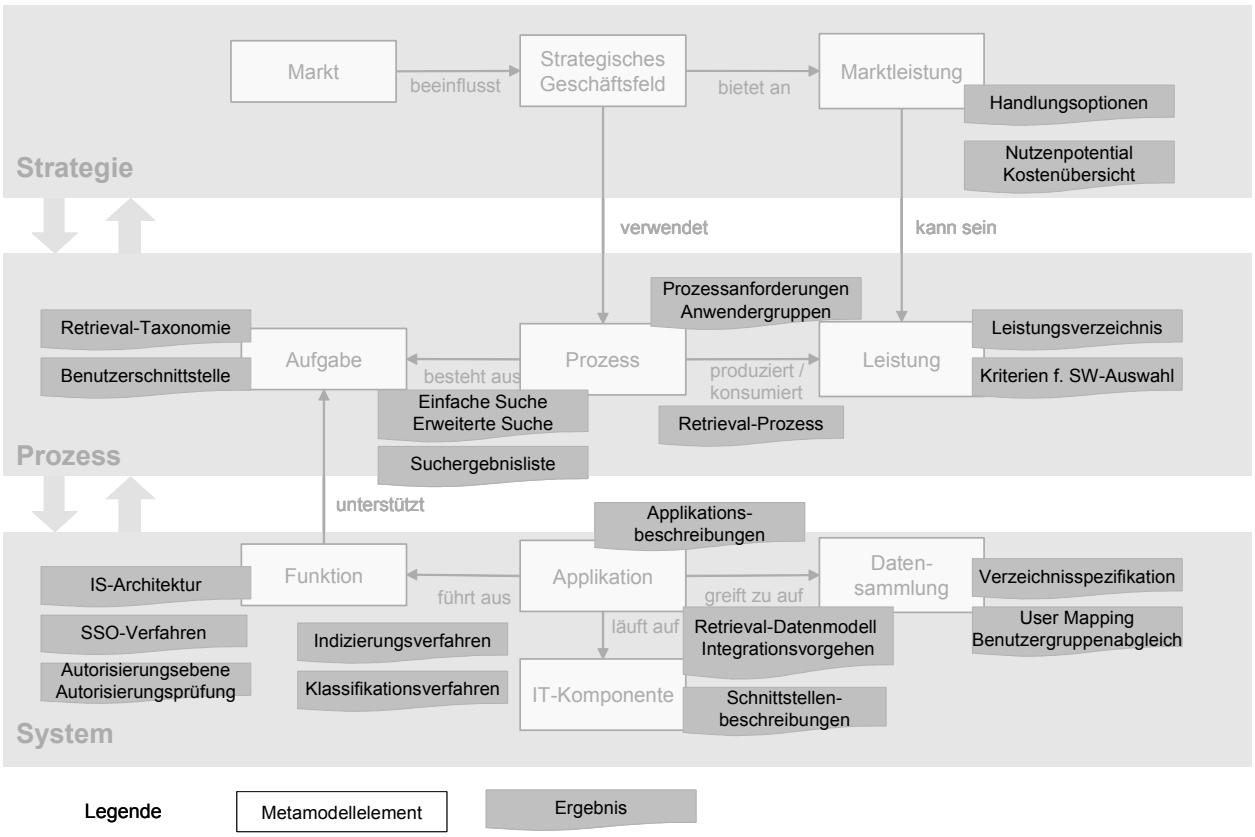


Abbildung 5-33: Einordnung der Ergebnisdokumente in das Business-Engineering-Metamodell

Auf Beispiele verzichtet das Dokumentationsmodell. Diese sind in den jeweiligen Beschreibungen der Techniken enthalten.

6. Zusammenfassung und Ausblick

Ziel dieser Arbeit ist die Entwicklung einer Methode zur Einführung von Information Retrieval in Portalen. Die Zusammenfassung stellt die wesentlichen Ergebnisse des vorliegenden Dissertationsprojekts den in Kapitel 1.3 und Kapitel 2.5 formulierten Zielsetzungen gegenüber. Den Abschluss des vorliegenden Kapitels und auch dieser Arbeit skizzieren Ansätze zur möglichen Weiterentwicklung, welche die Ergebnisse der Dissertation reflektieren und einen Ausblick auf zukünftige Entwicklungen geben.

6.1. Ergebnisse der Arbeit

Forschungsergebnisse des Business Engineering verfolgen die duale Zielsetzung, wissenschaftliche Fundierung mit praktischer Anwendbarkeit zu verbinden. Die Ergebnisse der vorliegenden Arbeit versuchen diesem Anspruch zu genügen, indem der vorgestellte Methodenvorschlag sowohl auf dem aus sieben wissenschaftlich dokumentierten Portal- und Information-Retrieval-Methoden abgeleiteten Metamodell als auch auf den Erkenntnissen aus vier Praxisfällen basiert.

Ergebnisse des Grundlagenkapitels

Der dieser Arbeit zugrundeliegende Forschungsrahmen des *Business Engineering* (BE) adressiert die ingenieurmässige Vorgehensweise bei der projektorientierten Umsetzung von Potenzialen der Informationstechnologie. Die resultierenden Lösungen auf Strategie-, Prozess- und Systemebene orientieren sich dabei konsequent an den Anforderungen der Kunden eines Unternehmens im Informationszeitalter. Als zentrale Schnittstelle zwischen Unternehmen im Informationszeitalter und ihren Kunden, Lieferanten und Mitarbeitern integrieren *Portale* die Inhalte und Funktionen heterogener unternehmensinterner und externer Anwendungen. Navigations- und Suchmechanismen von *Information-Retrieval-Systemen (Suchmaschinen)* ermöglichen diesen Zielgruppen eine effiziente Erschliessung der für die Durchführung ihrer Abläufe (*Kundenprozesse*) relevanten Informationen. Die Vielzahl der in der betrieblichen Praxis anzutreffenden heterogenen, semi- oder unstrukturierter Informationsquellen und teilweise zugriffsgeschützten inhaltsführenden Anwendungen (Content-Applikationen) verlangt eine systematische Integration. Der Einbezug des *Wissens* eines Informationserstellers bei der Indexierung verbessert die Erschliessbarkeit von Inhalten, z.B. durch Klassifikation mit inhalts- oder kontextbeschreibenden Attributen (Metadaten). Die Verbindung der Einzelbeiträge dieser in den Grundlagen dargestellten Themenschwerpunkte ergibt die in Tabelle 3-1 zusammengefassten Konsequenzen für eine Information-Retrieval-Methode.

Ergebnisse des Methodenvergleichs

Eine umfassende, an den Anforderungen von Geschäftsprozessen ausgerichtete Betrachtung von Information-Retrieval-Lösungen in Portalen im Sinne der Problemstellung dieser Arbeit ist bisher in der Literatur nicht zu finden. Die Darstellung und Bewertung sieben ausgewählter, wissenschaftlich dokumentierter Ansätze zeigt jedoch, dass einzelne Bestandteile einer Information-Retrieval-Methode in den verschiedenen Forschungsansätzen bereits Verwendung finden. Trotz unterschiedlicher Zielsetzung erlauben diese Ansätze die Ableitung und Ergänzung relevanter Gestaltungselemente. Als Ergebnis dieser Analyse fasst die vorliegende Arbeit die wesentlichen Gestaltungselemente in einem Metamodell für Information Retrieval in Portalen zusammen. Das Metamodell bildet folgende Erkenntnisse ab:

- Information-Retrieval-Lösungen sind nicht selbst wertschöpfend, sondern definieren ihr *Unterstützungspotenzial* durch die Verbesserung in den unterstützten Geschäftsprozessen.
- *Retrieval-Leistungen* werden von Geschäftsprozessen in Form von Volltextsuche, attributbasierter Suche, taxonomieorientierter Navigation oder als Kombination konsumiert.
- Die Bereitstellung dieser Leistungen erfordert *Aufbau- und Betriebsprozesse,* wie die Pflege eines kontrollierten Vokabulars für eine Stichwortsuche, die Bereitstellung eines Suchformulars oder den Aufbau eines Navigationsgerüstes.
- Erforderliche Aufbau- und Betriebsprozesse lassen sich aus den Retrieval-Leistungen direkt ableiten. Für eine *Volltextsuche* sind beispielsweise der Aufbau und die Pflege eines entsprechenden Suchindex notwendig. Retrieval-Leistungen wie attributbasierte Suche oder taxonomieorientierte Navigation, die den Anwendern einen definierten Ordnungsrahmen zur Informationsidentifikation zur Verfügung stellen, erfordern hingegen die Erstellung und Pflege eines Klassifikationsgerüstes für Inhalte (Retrieval-Taxonomie). Die Definition von Abläufen für die Integration bestehender Anwendungen, existierender Zugriffs-berechtigungen und bereitzustellender Benutzerschnittstellen ist bei allen drei Retrieval-Leistungen unverzichtbar.
- Die *Integration inhaltsführender Anwendungen* (Content-Applikationen) umfasst die Analyse von Inhalten, deren Struktur sowie deren Zugriffsbeschränkungen. Die technische Anbindung der Content-Applikationen an eine Suchmaschine benötigt Schnittstellen, die so genannten Konnektoren. Die Auswahl eines Konnektors, und damit einer Suchmaschine, hängt von den technischen Spezifika der im Unternehmen vorhandenen Content-Applikationen sowie den bereitzustellenden Retrieval-Leistungen ab.

Ergebnisse der Praxisfälle

Um die Relevanz des aus der wissenschaftlichen Literatur abgeleiteten Information-Retrieval-Metamodels zu überprüfen, untersucht diese Arbeit *vier unterschiedliche Lösungsansätze aus der betrieblichen Praxis*. Wesentliche Erkenntnisse sind:

- In keinem der Praxisbeispiele war das Information-Retrieval-Projekt von der Realisierung finanzieller Vorteile getrieben. Ausgangspunkt war vielmehr ein *zunehmender Leidensdruck innerhalb der Organisationen*, den Anwendern die Informationsidentifikation zu erleichtern.

- Die untersuchten Praxisfälle sehen das *Unterstützungspotenzial* von Information Retrieval in Portalen primär in der *Qualitätssteigerung und/oder Zeitersparnis* in den zu unterstützenden Prozessen. Kennzahlen zur Erfolgsmessung wurden nicht definiert, in allen Unternehmen jedoch bei Projektabschluss nachträglich als sinnvoll erachtet bzw. gewünscht.

- Eine gründliche *Erhebung der Anwenderanforderungen* entlang ihrer Strategie, Prozesse und Systeme war in allen Unternehmen erforderlich. Die Analyse der Anwenderbedürfnisse erfolgte über strukturierte Befragung von repräsentativen Anwendervertretern in Workshops oder durch Interviews mit einzelnen Führungskräften.

- In allen dargestellten Praxisfällen wurden die *vorhandenen Informationsstrukturen und zu integrierenden Anwendungen* bereits zu Projektbeginn analysiert. Die frühzeitige Abschätzung der Integrationsfähigkeit vorhandener Systeme anhand einiger weniger Merkmale ist charakteristisch für den Aufbau einer Information-Retrieval-Lösung. Verfügbare Ausgabeschnittstellen, enthaltene Dokumentenformate, gespeicherte Klassifikationsmerkmale oder verwendete Benutzerverzeichnisse beeinflussen die Auswahl der zur Verwendung kommenden Systeme wie Portal, Suchmaschine oder Content-Applikationen erheblich.

- Alle Ansätze klassifizieren die Portalinhalte bei der Erstellung, um eine konsistente Navigation sowie präzise Suchfunktionen bereitzustellen. Je umfangreicher das *Klassifikationsgerüst (Retrieval-Taxonomie)* ist, desto exaktere Suchanfragen können die Anwender formulieren. Jedoch steigert die Erstellung und Pflege einer Vielzahl von Attributen und Wertebereichen den Aufwand für die Betreiber.

- Je unbekannter dem Portalbetreiber die genauen *Informationsbedürfnisse der Zielgruppe* und deren ‚Begriffswelt' sind, umso schlanker fällt die Retrieval-Taxonomie aus. Weisen Unternehmen und Anwender eines Information-Retrieval-Systems hingegen ein gemeinsames Begriffsverständnis auf, z.B. beim Einsatz einer Suchmaschine in einem Mitarbeiterportal, so erlaubt die hohe Transparenz

dieser gemeinsamen ‚Begriffswelt' eine umfangreiche Retrieval-Taxonomie für die attributbasierte Suche und die taxonomieorientierte Navigation.

- *Attribute und Werte zur Klassifikation von Inhalten* können in vielen Fällen aus im Unternehmen vorhandenen Dokumenten (z.B. dem Dienstleistungsportfolio, der Aufbauorganisation oder einer Unternehmenspräsentationen) abgeleitet und aus bestehenden Systemen automatisiert übernommen werden (z.B. betriebswirtschaftlicher Standardsoftware, elektronischen Produktkatalogen oder CRM-Systemen).

- Die mit der Realisierung einer umfassenden Retrieval-Taxonomie verbundenen Erhebungs- und Abstimmungsaufwände führten in allen Praxisfällen dazu, dass in der finalen Lösung ausschliesslich die für das Unternehmen *wichtigsten inhaltscharakterisierenden Attribute* in der Retrieval-Taxonomie abgebildet waren.

- Die Praxisfälle zeigen, dass die Einführung einer Portalsuchmaschine *organisatorische Abläufe* verändert. Den Ausgangspunkt bildet der *Retrieval-Prozess des Endanwenders*, der die notwendigen Aktivitäten für den *Aufbau und Betrieb einer Retrieval-Taxonomie*, die *Integration bestehender Anwendungen* und die *Abbildung bestehender Benutzer- und Berechtigungsmodelle* bestimmt.

- Auf Systemebene wurde *Single Sign On* in allen Praxislösungen realisiert, bei dem ein am Portal authentifizierter Benutzer ohne weitere Login-Aufforderungen auch über die Suchmaschine auf zugriffsgeschützte Inhalte zugreifen kann. Die Unternehmen verwendeten dazu ein *zentrales Benutzerverzeichnis*, auf das alle beteiligten Anwendungen (Portal, Suchmaschine, Content-Applikationen) zur Authentifizierung von Benutzern gemeinsam zugreifen.

Ergebnisse des Methodenvorschlags

Unter Verwendung des Metamodells und der Erkenntnisse aus den vier Praxisfällen entwickelt die vorliegende Arbeit einen *Methodenvorschlag* im Forschungsrahmen des Business Engineering (Strategie, Prozess, Systeme), der die Erkenntnisse der wissenschaftlichen Ansätze mit den Herausforderungen und Restriktionen der betrieblichen Praxis verbindet. Die vier Techniken des Vorgehensmodells unterstützen die Analyse, Planung, Gestaltung und Implementierung einer Information-Retrieval-Lösung von der Anforderungsanalyse bis zur Umsetzung:

- Die *Potenzialanalyse* hilft Unternehmen, den ‚Leidensdruck' zur Einführung einer Portalsuchmaschine anhand spezifischer Schwachstellen bei der betrieblichen Aufgabenerfüllung zu konkretisieren (z.B. verteilte Inhalte erschliessen, multiple Sichten auf Inhalte abbilden, Inhalte anhand von Strukturinformationen erschliessen und heterogene Informationsbestände durchsuchen). Kritische

Erfolgsfaktoren und Führungsgrössen ermöglichen die Operationalisierung der Zielsetzung und eine Erfolgsmessung nach Projektabschluss (z.B. in den Bereichen Qualitätssteigerung, Zeitersparnis und Kostenreduktion). Die Verantwortlichen dokumentieren in der Potenzialanalyse die Anforderungen kundenorientierter Prozesse an eine Portalsuche. Neben den Einsatzmöglichkeiten identifizieren sie aber auch mögliche technische Restriktionen durch die frühzeitige Analyse der Integrationsfähigkeit vorhandener Systeme (z.B. hinsichtlich der Dokumentenformate, Ausgabeschnittstellen, Struktur- und Klassifikationsmerkmale und Zugriffsbeschränkungen). Die Integration einer Vielzahl heterogener Systeme, die Abbildung unterschiedlicher Benutzer- und Berechtigungskonzepte sowie die Bereitstellung anwendungsübergreifender Navigationsstrukturen (taxonomieorientierte Navigation) und attributbasierter Suche treiben die Kosten einer Information-Retrieval-Lösung in die Höhe. Die beiden Letzteren sind nur sinnvoll, wenn die Informationsbedürfnisse der adressierten Zielgruppe genau bekannt sind, z.B. beim unternehmensinternen Einsatz.

- Im Rahmen der *Strategieplanung* legen die Projektverantwortlichen die zukünftige Prozessunterstützung durch die Information-Retrieval-Lösung fest (Volltextsuche, attributbasierte Suche, taxonomieorientierte Navigation). Auf Basis der erfassten Applikationen skizzieren sie die zu implementierende Systemarchitektur. Die Projektverantwortlichen bewerten die Nutzenaspekte und schätzen die Kosten der einzuführenden Lösung ab. Die Anschaffungs- und Betriebskosten für die Suchmaschinensoftware ermitteln die Projektverantwortlichen dabei durch die Grobbewertung unterschiedlicher Realisierungsvarianten in einem Softwareauswahlverfahren. Ein finaler Produktentscheid erfolgt jedoch erst nach Durchführung der Prozessplanung und Abschluss des Systementwurfs.

- In der *Prozessplanung* definieren die Verantwortlichen Abläufe, mit denen die Retrieval-Leistungen der Strategieplanung organisatorisch umgesetzt werden sollen. Sie skizzieren den zukünftigen *Retrieval-Prozess* und leiten aus diesem Aktivitäten für den *Aufbau und die Pflege einer Retrieval-Taxonomie*, zur *Integration von Applikationen* und zur Entwicklung *eines Benutzer- und Berechtigungsmodells* für die Portalsuche ab.

- Im *Systementwurf* setzen die Projektmitarbeiter die fachlichen Ergebnisse von Strategie- und Prozessplanung konzeptionell auf Systemebene um. Sie spezifizieren Systemfunktionen, gestalten Anwendungsschnittstellen und implementieren Aspekte der Zugriffssicherheit.

Ein Rollen- und Dokumentationsmodell ergänzt die Techniken des Vorgehensmodells und gewährleistet die Vollständigkeit der Methode in der Diktion des Method Engineering. Die Integration vertiefender Beispiele und Auszüge aus originären Ergebnisdokumenten der Fallstudien unterstützt das Verständnis der Entwurfstechniken und erweitert gleichzeitig die Darstellung der vier Praxisfälle.

6.2. Ansätze zur Weiterentwicklung

Die vorliegende Arbeit konnte aus der Bewertung sieben wissenschaftlich dokumentierter Ansätze und den Erkenntnissen aus vier Praxisfällen eine Methode zur Einführung von Information Retrieval in Portalen entwickeln. Die vorgeschlagenen Gestaltungselemente auf Strategie-, Prozess- und Systemebene bilden die Methode umfassend, jedoch keineswegs abschliessend ab.

Der Autor versteht die Ergebnisse dieser Arbeit als Ausgangspunkt für weitere wissenschaftliche Forschungsaktivitäten und Umsetzungen in der Praxis. Folgende Aspekte bieten Ansatzpunkte zur Weiterentwicklung:

- *Durchgängige und umfassende Anwendung der Methode*. Obwohl sich wesentliche Bestandteile der Methode in den vier Praxisfällen bewährt haben, muss von einem Methodenvorschlag gesprochen werden, da eine durchgängige und umfassende Anwendung noch aussteht. Die Erkenntnisse aus der vollständigen Anwendung der Methode in ihrer vorliegenden Form in weiteren betrieblichen Praxisfällen müssen der Validierung des vorliegenden Ansatzes dienen. Der Methodenvorschlag kann durch die Identifikation spezifischer Lösungsmuster für einzelne Branchen, Unternehmen oder Prozesse erweitert und verfeinert werden. Weitere Erkenntnisse liefern Untersuchungen kooperativer Information-Retrieval-Lösungen, die Informationsbestände verschiedener Unternehmen in einer Portalsuche integrieren.

- *Modularisierung und Standardisierung von Information-Retrieval-Komponenten*. Die Integration von Portal, Suchmaschine und Content-Applikationen ist in der betrieblichen Praxis derzeit von individuellen und projektspezifischen Lösungen geprägt. Eine Modularisierung und Standardisierung der beteiligten Information-Retrieval-Komponenten wie Authentifizierung, Autorisierung, Indexierung, Klassifizierung, Suche oder Navigation kann den Integrationsaufwand zwischen den beteiligten Anwendungen reduzieren. WebServices übernehmen hier aus betriebswirtschaftlicher Sicht klar abgrenzbare, hoch standardisierbare Aufgaben aus Prozessen, die zeit- und/oder transaktionsbasiert verrechenbar und in die Applikationswelt eines Unternehmens mit den bestehenden Systemen integrierbar sind (s. [Österle 2002a, 32], [Gisolfi 2001]). Aus technischer Sicht zielen WebServices auf die Kommunikation lose gekoppelter Softwarekomponenten [s. Sleeper/Robins 2001, 7]. Auf der Basis technischer Protokoll- und

Kommunikationsstandards dienen sie als Grundlage einer unternehmens- und anwendungsübergreifenden automatisierten System-zu-System-Kommunikation. Durch ihre modulare und standardisierte Funktionsweise ermöglichen WebServices einem Unternehmen, eindeutig abgrenzbare Bestandteile oder sogar die gesamte Portalsuche an einen externen Dienstleister auszulagern. Die Betreiber der Internetsuchmaschine Google bieten beispielsweise einen solchen WebService an, der die Indizierung und Suche unternehmenseigener Informationsbestände mit der externen Suchmaschine Google erlaubt [s. Google 2003].

- *Trennung von Inhalten und Klassifikation*. In den vier untersuchten Praxisfällen und dem Methodenvorschlag bildet die manuelle Klassifikation von Inhalten bei ihrer Erstellung ein zentrales Merkmal jeder Information-Retrieval-Lösung. Die Klassifikation von Inhalten anhand der Attribute und Werte einer Retrieval-Taxonomie bildet die Grundlage für eine attributbasierte Suche und eine taxonomieorientierte Navigation. Den Vorteilen der gezielten Erschliessung von verteilten, heterogenen und grossen Informationsbeständen anhand dieser Klassifikationsmerkmale steht jedoch der Nachteil bei der Veränderung von Begriffen gegenüber. Unternehmen müssen entscheiden, ob sie bestehende Inhalte erneut klassifizieren, wenn die betrieblichen Anforderungen die Aufnahme neuer Begriffe oder das Ersetzen bestehender Begriffe in der Retrieval-Taxonomie notwendig machen. Eine mögliche Lösung bietet hier der auf der ISO 13250 basierende Ansatz für so genannte Topic Maps (s. [ISO 2002], [Smolnik/Nastansky 2002], [Rath 2003]). Topic Maps ermöglichen eine Suche und Navigation anhand von Begriffen (‚Topics') einer Retrieval-Taxonomie, ohne dass diese in den Informationsobjekten als Text direkt enthalten sein müssen. Stattdessen verweisen sie über Zuordnungsregeln (‚Rules'), z.B. eine Volltextsuchanfrage, auf relevante Inhalte (‚Occurences') und trennen dadurch den Inhalt von der eigentlichen Klassifikation. Topic Maps erweitern darüber hinaus die bei bisherigen Portalsuchmaschinen angezeigte lineare Suchergebnisliste um die Darstellung von Zusammenhängen (‚Associations') zwischen einzelnen Topics, z.B. in Form eines so genannten Hyperbolic Tree (s. Abbildung 6-1).

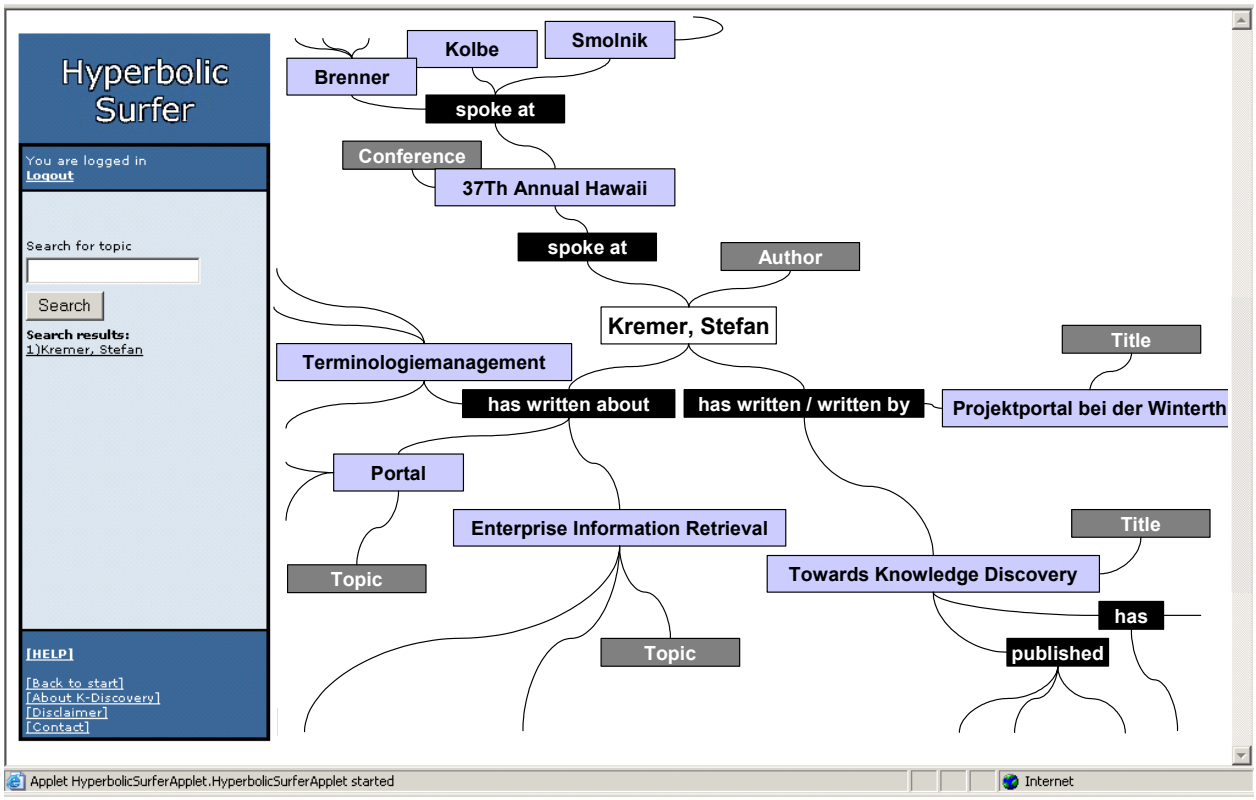


Abbildung 6-1: Topic Map bei IWI-HSG

Mit der weiteren Anwendung des vorliegenden Methodenvorschlags, der Nutzung von WebServices und dem Einsatz von Topic Maps können Wissenschaftler gemeinsam mit Unternehmen neue Forschungsgebiete erschliessen und Ergebnisse erarbeiten.

Das vorliegende Dissertationsprojekt hat deutlich gemacht, dass eine Weiterentwicklung der bereits vorhandenen Lösungsansätze zur Unterstützung eines Unternehmens beim Übergang aus dem Industrie- in das Informationszeitalter unverzichtbar ist. Hierbei setzt jedoch das Gelingen einer Transformation immer eine Kombination aus methodisch fundierten Vorgehensweisen und gezieltem Einsatz leistungsstarker Informationstechnologien voraus.

Literaturverzeichnis

[Abecker et al. 2002]
Abecker, A., Hinkelmann, K., Maus, H., Müller, H.J., Geschäftsprozessorientiertes Wissensmanagement - Effektive Wissensnutzung bei der Planung und Umsetzung von Geschäftsprozessen, Springer, Berlin, 2002

[Ahuja 1996]
Ahuja, V., Network and Internet Security, Academic Press, Chestnut Hill, 1996

[Alt et al. 2002]
Alt, R., Cäsar, M., Leser, F., Puschmann, T., Reichmayr, C., Österle, H., Methode zur Entwicklung von Prozessportalen, Arbeitsbericht, Institut für Wirtschaftsinformatik, Universität St.Gallen, 2002

[Alt et al. 2003]
Alt, R., Cäsar, M., Leser, F., Puschmann, T., Reichmayr, C., Zurmühlen, R., Methode zur Entwicklung von Prozessportalen, in: Alt, R., Österle, H. (Hrsg.), Real-Time Business: Lösungen, Bausteine und Potentiale des Business Networking, Springer, Berlin etc., 2003, S. 257-278

[Altenhofen et al. 2002]
Altenhofen, C., Stanisic-Petrovic, M., Junker, M., Kieninger, T., Hofmann, H.R., Dokumentenverwaltungs-Systeme - erwünscht, aber noch zu teuer, in: Wissensmanagement, Vol. 4 (2002) Ausgabe 3, S. 32-35

[Anderson/Pérez-Carballo 2001]
Anderson, J.D., Pérez-Carballo, J., The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part II: Machine indexing, and the allocation of human versus machine effort, in: Information Processing & Management, Vol. 37 (2001) Issue 2, S. 255-277

[Andrews 2003a]
Andrews, W., Search and Taxonomy Converge for Information Retrieval, Report, GartnerGroup, Stamford, 2003

[Andrews 2003b]
Andrews, W., Visionaries Invade the 2003 Search Engine Magic Quadrant, Report, GartnerGroup, Stamford, 2003

[Ant 2003]
Ant, A., Password Management, Single Sign-On, and Authentication Management Infrastructure Products, Report, GartnerGroup, Stamford, 2003

[Applehans et al. 1998]
Applehans, W., Globe, A., Laugero, G., Managing Knowledge: A Practical Web-Based Approach, Addison Wesley Publishing Company, Reading, MA, 1998

[Bach 2003]
Bach, V., Architektur rollenzentrierter Unternehmensportale, Habilitation, Institut für Wirtschaftsinformatik, Universität St.Gallen, St.Gallen, 2003

[Bach et al. 2000]
Bach, V., Österle, H., Vogler, P., Business Knowledge Management in der Praxis, Springer, Berlin, 2000

[Baeza-Yates/Ribeiro-Neto 1999]
Baeza-Yates, R., Ribeiro-Neto, B., Modern Information Retrieval, Addison Wesley, New York, 1999

[Baeza-Yates/Schäuble 2002]
Baeza-Yates, R., Schäuble, P., Retrieving Information: A Discipline with a Tradition, in: UPGRADE (The European Online Magazine for the IT-Professional), III (2002) 3, S. 3-4

[Bailey 1994]
Bailey, K.D., Typologies and taxonomies: an introduction to classification techniques, 3rd Edition, Sage, Thousand Oaks u.a., 1994

[Balzert 1996]
Balzert, H., Lehrbuch der Software-Technik, Band 1 (Software-Entwicklung), Spektrum Akademischer Verlag, Heidelberg, Berlin, Oxford, 1996

[Balzert 2001]
Balzert, H., Lehrbuch der Sotware-Technik, Band 1 (Software-Entwicklung), 2. Auflage, Spektrum Akademischer Verlag, Heidelberg, 2001

[Bauer 2001]
Bauer, H., Unternehmensportale - Geschäftsmodelle, Design, Technologien, Galileo Press, Bonn, 2001

[Bayles 1998]
Bayles, D., Extranets: Building the Business-to-Business Web, Prentice Hall, Upper Saddle River, 1998

[Beall/Hodges 2002]
Beall, S., Hodges, R., Storage & Retrieval: Software Comparison Columns, Report, GartnerGroup, Stamford, 2002

[Berners-Lee 1997]
Berners-Lee, T., Web architecture: Metadata Tim Berners-Lee, 1997, W3C, http://www.w3.org/DesignIssues/Metadata, 14.05.2003

[Blessing 2001]
Blessing, D., Content Management für das Business Engineering - Fallbeispiele, Modelle und Anwendungen für das Wissensmanagement bei Beratungsunternehmen, Dissertation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2001

[Brecht 1999]
Brecht, L., Unternehmensführung im Zeitalter des Internets, Habilitation, Institut für Wirtschaftsinformatik, Universität St.Gallen, St. Gallen, 1999

[Brenner 1995]
Brenner, C., Techniken und Metamodell des Business Engineering, Dissertation, Institut für Wirtschaftsinformatik, Universität St.Gallen, St.Gallen, 1995

[Brenner 1988]
Brenner, W., Entwürfe betrieblicher Datenelemente - Ein Weg zur Integration von Informationssystemen, Springer, Berlin, 1988

[Brenner 1993]
Brenner, W., Konzepte des Informationssystemmanagements, 1. Auflage, Physika-Verlag, Heidelberg, 1993

[Brenner et al. 2002]
Brenner, W., Zarnekow, R., Pörtig, F., Entwicklungstendenzen im Informationsmanagement, in: Österle, H., Winter, R. (Hrsg.), Business Engineering, Springer, Berlin, 2002, S. 148-168

[Brewer 2002]
Brewer, E.A., The Consumer Side of Search, in: Communications of the ACM, 45 (2002) 9, S. 40-41

[Brin/Page 1998]
Brin, S., Page, L., The Anatomy of a Large-Scale Hypertextual Web Search Engine, in: WWW7 / Computer Networks 30(1-7) (1998), S. 107-117

[Britzelmaier 1999]
Britzelmaier, B., Informationsverarbeitungs-Controlling - Ein datenorientierter Ansatz, B. G. Teubner, Stuttgart, 1999

[Budin 1990]
Budin, G., Scientifc Knowledge Structures, in: Proceedings of the Terminology and Knowledge Engineering 1990, Indeks Verlag, Frankfurt, 1990, S. 77-83

[Bullinger et al. 2002]
Bullinger, H.-J., Eberhardt, C.-T., Gurzki, T., Hinderer, H., Marktübersicht: Portal Software für Business-, Enterprise-Portale und E-Collaboration, Frauenhofer IRB Verlag, Stuttgart, 2002

[Bullinger et al. 1997]
Bullinger, H.-J., Wörner, K., Prieto, J., Wissensmanagement heute - Daten, Fakten, Trends, Fraunhofer-Institut für Arbeitswirtschaft und Organisation (IAO), Stuttgart, 1997

[Bullinger et al. 1998]
Bullinger, H.-J., Wörner, K., Prieto, J., Wissensmanagement - Modelle und Strategien für die Praxis, in: Bürgel, H.D. (Hrsg.), Wissensmanagement - Schritte zum intelligenten Unternehmen, Springer, Berlin, 1998

[Büren et al. 2003]
Büren, A., Kolbe, L.M., Brenner, W., Content Management zur Realisierung von Customer Knowledge Performance – Konzeption und Umsetzung bei einem Finanzdienstleister, in: Uhr, W., Esswein, W., Schoop, E. (Hrsg.), Wirtschaftsinformatik 2003, (2003), Physika-Verlag, Heidelberg, 2003, S. 519-540

[Carden 1999]
Carden, P., The New Face of Single Sign-On, in: Network Computing, Vol. 10 (1999) Issue 6, S. 32-42

[Cathro 1997]
Cathro, W., Metadata: An Overview, http://www.nla.gov.au/nla/staffpaper/cathro3.html, 23.05.2003

[Chinitz 2000]
Chinitz, J., Single Sign On - Is it really possible?, in: Access Control Systems and Methodology, Vol. 4 (2000)

[Christ 2002]
Christ, O., Eine Architektur für das Content-Management - Strategien, Prozessmodelle, Softwarelösungen und Projektszenarien, Dissertation, Universität St. Gallen, St. Gallen, 2002

[Collins 2001]
Collins, H., Corporate Portals: Revolutionizing Information Access to Increase Productivity and Drive the Bottom Line, Amacom, New York, 2001

[Davenport/Marchand 2001]
Davenport, T.H., Marchand, D., Is KM just good information management?, in: Financial Times Limited, 25 (2001) 4, S. 2

[Davenport/Prusak 1998]
Davenport, T.H., Prusak, L., Working Knowledge: How Organizations Manage What They Know, Harvard Business School Press, Boston, 1998

[Davydov 2001]
Davydov, M.M., Corporate Portals and e-Business Integration, McGraw-Hill, New York, 2001

[Delphigroup 2002]
Delphigroup, Perspectives on Information Retrieval, Report, 2002

[Demarest 1997]
Demarest, M., Knowledge Management: An Introduction, http://www.noumenal.com/marc/km1.pdf, 09.04.2003

[Dey/Abowd 1999]
Dey, A.K., Abowd, G.D., Towards a Better Understanding of Context and Context-Awareness, College of Computing, Georgia Institute of Technology, Atlanta, GA, 1999

[Dias 2001]
Dias, C., Corporate Portals: a literature review of a new concept in information management, in: International Journal of Information Management, (2001) 21, S. 269-287

[DMOZ 2002]
DMOZ, The Open Directory Project, DMOZ.ORG, http://dmoz.org/, 22.11.2002

[Drucker 1994]
Drucker, P.F., The New Realities in Government and Politics/in Economics and Business/in Society and World View, HarperBusiness, 1994

[Enkel et al. 2000]
Enkel, E., Raimann, J., Seufert, A., Vassiliadis, S., Wicki, Y., Back, A., von Krogh, G., MERLIN - Materializing, Experience, Refining and Learning in Knowledge Networks, 2000

[Eppler et al. 1999]
Eppler, M., Seifried, P., Röpnack, A., Improving Knowledge Intensive Processes through an Enterprise Knowledge Medium, in: Proceedings of the ACM SIGCPR conference on computer personnel research, ACM Press, 1999, S. 222-230

[Fahey/Prusak 1998]
Fahey, L., Prusak, L., The eleven deadliest sins of Knowledge Management, in: California Management Review, 40 (1998) 3

[Fank/Trojan 2003]
Fank, M., Trojan, J., Knowledge Management in der Praxis - wohin entwicklen sich Intranets, http://www.ifem.org/kmpraxis_band3.htm, 26.11.2003

[Felber/Budin 1989]
Felber, H., Budin, G., Terminologie in Theorie und Praxis, Gunter Narr Verlag, Tübingen, 1989

[Ferber 2003]
Ferber, R., Information Retrieval - Suchmodelle und Data-Mining-Verfahren für Textsammlungen und das Web, dpunkt.verlag, Heidelberg, 2003

[Ferstl/Sinz 1998]
Ferstl, O.K., Sinz, E.J., Grundlagen der Wirtschaftsinformatik, Band 1, 3. Aufl., Oldenbourg, München, Wien, 1998

[Finkelstein/Aiken 2000]
Finkelstein, C., Aiken, P.G., Building Corporate Portals with XML, McGraw-Hill, New York, 2000

[Fleisch 2000]
Fleisch, E., Koordination in Netzwerkunternehmen - Prozessorientierung als Gestaltungsprinzip bei der Vernetzung von Unternehmen, Habilitation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2000

[Fleisch/Österle 2001a]
Fleisch, E., Österle, H., Das Tor zur IT-Welt: Thesen zum erfolgreichen Portaleinsatz, in: Computerwoche extra, (2001) 2, S. 28-31

[Fleisch/Österle 2001b]
Fleisch, E., Österle, H., Vom elektronischen Schaufenster zum Prozessportal - Sieben Thesen zur Gestaltung von erfolgreichen Internetportalen, in: io Management, 70 (2001) 4, S. 38-44

[Fleisch/Österle 2004]
Fleisch, E., Österle, H., Auf dem Weg zum Echtzeitunternehmen, in: Alt, R., Österle, H. (Hrsg.), Real-time Business. Lösungen, Potentiale und Herausforderungen des Business Networking, Springer, Berlin, 2004, S. 1-16

[Franken/Gadatsch 2002]
Franken, R., Gadatsch, A., Integriertes Knowledge Management. Konzepte, Methoden, Instrumente und Fallbeispiele, Vieweg, Braunschweig, 2002

[Fuhr 2002]
Fuhr, N., Information Retrieval - Skriptum zur Vorlesung, Fachbereich Informatik, Uni Dortmund, Dortmund, 2002

[Gaus 2003]
Gaus, W., Dokumentations- und Ordnungslehre, 4. Auflage, Springer, Heidelberg, 2003

[Gebert 2003]
Gebert, H., IT-gestütztes Kompetenzmanagement in kundenorientierten Geschäftsprozessen, Dissertation, Institut für Wirtschaftsinformatik, Universität St.Gallen, St.Gallen, 2003

[Gebert et al. 2003]
Gebert, H., Geib, M., Kolbe, L.M., Brenner, W., Knowledge-enabled Customer Relationship Management - Integrating Customer Relationship Management and Knowledge Management Concepts, in: Journal of Knowledge Management, 7 (2003) 5, S. 107-123

[Geib/Riempp 2002]
Geib, M., Riempp, G., Customer Knowledge Management - Wissen an der Schnittstelle zum Kunden effizient handhaben, in: Abecker, A., Hinkelmann, K., Maus, H., Müller, H.-J., Geschäftsprozessorientiertes Wissensmanagement - Effektive Wissensnutzung bei der Planung und Umsetzung von Geschäftsprozessen, Springer, Berlin et al., 2002, S. 393-417

[Gerick 2000]
Gerick, T., Recherchetechniken: Suchen und Finden sind zweierlei, in: Computerwoche, (2000) 7

[GI 2004]
GI, Gesellschaft für Informatik, Fachgruppe Information Retrieval - Definition und Abgrenzung von Information Retrieval, http://ls6-www.informatik.uni-dortmund.de/ir/fgir/mitgliedschaft/brochure2.html, 26.03.2004

[Gilchrist 2001]
Gilchrist, A., How automatic categorization can power knowledge management, in: Knowledge Management Review, 4 (2001) 2, S. 10-11

[Gillet 2001]
Gillet, F.E., Making Enterprise Portals pay, Report, Forrester Research Inc., 2001

[Gisolfi 2001]
Gisolfi, D., Web Services Architect, Part 1: An Introduction to Dynamic e-business, IBM developerWorks, http://www-106.ibm.com/developerworks/webservices/library/ws-arc1/, 05.12.2001

[Goldratt 1990]
Goldratt, E.M., The Haystack Syndrome: Shifting Information Out of the Data Ocean, North River Press, Croton-on-Hudson, NY, 1990

[Google 2003]
Google, Google Web APIs Reference, Google, http://www.google.com/apis/download.html, 28.08.2003

[Gordon/Pathak 1999]
Gordon, M., Pathak, P., Finding information on the World Wide Web - The retrieval effectiveness of search engines, in: Information Processing and Management, Vol. 35 (1999), S. 141-180

[Grochla 1982]
Grochla, E., Grundlagen der organisatorischen Gestaltung, Poeschel, Stuttgart, 1982

[Gronover 2003]
Gronover, S., Multi-Channel-Management - Konzepte, Techniken und Fallbeispiele aus dem Retailbereich der Finanzdienstleistungsbranche, Dissertation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2003

[Grover/Davenport 2001]
Grover, V., Davenport, T.H., General perspectives on knowledge management: Fostering a research agenda, in: Journal of Management Information Systems, 18 (2001) 1, S. 5-21

[Guarino 1992]
Guarino, N., Concepts, Attributes, and Arbitrary Relations: Some Linguistic and Ontological Criteria for Structuring Knowledge Bases, in: van de Reit, R.P., Meersman, R.A. (Hrsg.), Linguistic Instruments in Knowledge Engineering, North-Holland, 1992, S. 195 - 211

[Gurzki/Hinderer 2003]
Gurzki, T., Hinderer, H., Eine Referenzarchitektur für Software zur Realisierung von Unternehmensportalen, in: Reiner, U., Abecker, A., Staab, S., Stumme, G. (Hrsg.), Proceedings of the WM2003: Professionelles Wissensmanagement - Erfahrungen und Visionen, P-28 Bonner Köllen Verlag, Lecture Notes in Informatics, Luzern, 2003

[Gutzwiller 1994]
Gutzwiller, T., Das CC RIM-Referenzmodell für den Entwurf von betrieblichen, transaktionsorientierten Informationssystemen, Physica-Verlag, Heidelberg, 1994

[Hagen 2000]
Hagen, P., Must Search stink?, Report, Forrester, 2000

[Harris et al. 2003]
Harris, K., Caldwell, F., Linden, A., Knox, R.E., Logan, D., Taxonomy Creation - Bringing Order to Complexity, Report, GartnerGroup, Stamford, 2003

[Heisig/Vorbeck 2001]
Heisig, P., Vorbeck, J., Benchmarking Survey Results, in: Mertins, K., Heisig, P., Vorbeck, J. (Hrsg.), Knowledge Management - Best Practices in Europe, Springer, Berlin etc., 2001, S. 97-123

[Hellmuth 1997]
Hellmuth, T.W., Terminologiemanagement: Aspekte einer effizienten Kommunikation in der computergestützten Informationsverarbeitung, Dissertation, Universität von Konstanz, Konstanz, 1997

[Hess 1996]
Hess, T., Entwurf betreiblicher Prozesse - Grundlagen, Bestehende Methoden, Neue Ansätze, Dissertation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 1996

[Hesse 2002]
Hesse, W., Ontologie(n), in: Informatik Spektrum, 25 (2002) 6, S. 477-480

[Höller et al. 1998]
Höller, J., Pils, M., Zlabinger, R., Internet und Intranet, Springer, Berlin, 1998

[IBM 2003]
IBM, Guide to WebSphere Portal 4.2, IBM, ftp://ftp.software.ibm.com/software/websphere/portal/pdf/Guide-to-Websphere-Portal.pdf, 21.08.2003

[IMG 1997a]
IMG, PROMET BPR, Methodenhandbuch für den Entwurf von Geschäftsprozessen, Version 2.0, Information Management Group, St. Gallen, 1997

[IMG 1997b]
IMG, PROMET BPR: Method for Business Process Redesign, Release 2.0, IMG AG, St. Gallen, 1997

[ISO 1988]
ISO, Information Processing Systems - Open Systems Interconnection - The Directory. Draft International Standard ISO/IEC DIS 9594, International Standard Organization / International Electrotechn. Commission, 1988

[ISO 2002]
ISO, ISO/IEC 13250 - Topic Maps, International Organization For Standardization, Geneva, 2002

[Ives/Learmonth 1984]
Ives, B., Learmonth, G.P., The Information System as a Competitive Weapon, in: Communications of the ACM, 27 (1984) 12, S. 1193-1201

[IWI-HSG/IMG 2003]
IWI-HSG, IMG, Studie zum Einsatz und Nutzen von Portallösungen in der Schweiz, Report, IMG AG, St. Gallen, 2003

[Jansen 2000]
Jansen, C., Prozessunterstützung durch Wissensplattformen für Business Engineers, Dissertation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2000

[Kaiser 2000]
Kaiser, T.M., Methode zur Konzeption des Intranets, Dissertation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2000

[Kalakota/Robinson 2001]
Kalakota, R., Robinson, M., e-Business 2.0: roadmap for success, 2. Edition, Addison Wesley, Boston, 2001

[Kargl 2000]
Kargl, H., Management und Controlling von IV-Projekten, Oldenbourg, München, 2000

[Klemke 2000]
Klemke, R., Context Framework - an Open Approach to Enhance Organisational Memory Systems with Context Modelling Techniques, in: Proceedings of the The Third International Conference on Practical Aspects of Knowledge Management (PAKM 2000), CD-ROM, Basel, 2000

[Kofmann/Senge 1993]
Kofmann, F., Senge, P.M., Communities of Commitment: The Heart of Learning Organizations, in: Learning Organizations: Developing Cultures for Tomorrow's Workplace, (1993), S. 14-43

[Köhne et al. 2000]
Köhne, M., Enkel, E., Seufert, A., Back, A., von Krogh, G., Development Aspects and Supportive Factors in Knowledge Networks, Working Paper, University of St.Gallen, 2000

[Kolbe et al. 2003]
Kolbe, L.M., Österle, H., Brenner, W. (Hrsg.), Customer Knowledge Management - Kundenwissen erfolgreich einsetzen, Springer Verlag, Berlin Heidelberg, 2003

[KPMG 2001]
KPMG, Knowledge Management im Kontext von e-Business, Status quo und Perspektiven 2001, KPMG Consulting AG, 2001

[Kraus 2003]
Kraus, P., Kosten senken durch Wissensmanagement, in: Wissensmanagement, (2003) 4, S. 50-51

[Kremer et al. 2003a]
Kremer, S., Kolbe, L.M., Brenner, W., Do you know your terms? - A procedure model for Terminology Management, in: Proceedings of the European Conference on Information Systems, CD-ROM, Naples, 2003

[Kremer/Riempp 2001]
Kremer, S., Riempp, G., Babels Türme - Terminologiemanagement in B2E-Portalen, in: Computerwoche, 26 (2001) 30.06.2001, S. 59-60

[Kremer et al. 2003b]
Kremer, S., Smolnik, S., Kolbe, L., Towards Knowledge Discovery through Context Explication, in: Proceedings of the 37th Annual Hawaii International Conference On System Sciences, CD-ROM, Hawaii, 2003

[Kruchten 2000]
Kruchten, P., The Rational Unified Process, An Introduction, Addison-Wesley, Boston, 2000

[Krüger 1994]
Krüger, W., Organisation der Unternehmung, Kohlhammer, Stuttgart, 1994

[Latham 2001]
Latham, L., Web Content Management and Portals: Who's Doing What?, Report, GartnerGroup, Stamford, 2001

[Lee 1989]
Lee, A.S., A Scientific Methodology for MIS Case Studies, in: MIS Quarterly, (1989), S. 32-50

[Legner 1999]
Legner, C., Benchmarking informationssystemgestützter Geschäftsprozesse, Deutscher Universitätsverlag, Wiesbaden, 1999

[Leonard 1998]
Leonard, D., Wellsprings of Knowledge: Building and Sustaining the Sources of Innovation, Harvard Business School Press, 1998

[Linthicum 2000]
Linthicum, D.S., Enterprise Application Integration, Addison-Wesley, Upper Saddle River, NJ, 2000

[Logan 2001]
Logan, D., Understanding and Using Taxonomies, Report, GartnerGroup, Stamford, 2001

[Maier 2002]
Maier, R., Knowledge Management Systems - Information and Communication Technologies for Knowledge Management, Springer, Berlin etc., 2002

[McAdam/McCreedy 1999]
McAdam, R., McCreedy, S., A critical review of knowledge management models, in: The Learning Organization, 6 (1999) 3, S. 91-100

[Mertens 2003]
Mertens, P., Die Wirtschaftsinformatik auf dem Weg zur Unternehmensspitze - alte und neue Herausforderungen und Lösungsansätze, in: Uhr, W., Esswein, W., Schoop, E. (Hrsg.), Wirtschaftsinformatik 2003/Band I, Physica-Verlag, Heidelberg, 2003, S. 49-74

[Mertens et al. 1998]
Mertens, P., Bodendorf, F., König, W., Picot, A., Schumann, M., Grundzüge der Wirtschaftsinformatik, Springer, Berlin, 1998

[Merz 2002]
Merz, M., E-Commerce und E-Business, dpunkt Verlag, Heidelberg, 2002

[Miles/Huberman 1994]
Miles, M.B., Huberman, A.M., Qualitative Data Analysis: An Expanded Sourcebook, Sage Publications, Thousand Oaks, 1994

[Moriarty 1990]
Moriarty, T., Are you ready for a Repository, in: Database Programming & Design, Vol. 3 (1990) Issue 3, S. 61-71

[Murray 2001]
Murray, P., How smarter companies get results from KM, in: Financial Times, (2001) July 24, S. 11

[Nohr 2000]
Nohr, H., Automatische Verfahren der Dokumentenanalyse, in: Nohr, H. (Hrsg.), Wissensmanagement: Wie Unternehmen ihre wichtigste Ressource erschliessen und teilen, Business Village, Göttingen, 2000, S. 61-87

[Nohr 2003]
Nohr, H., Grundlagen der automatischen Indexierung: Ein Lehrbuch, Logos, Berlin, 2003

[Nonaka 1994]
Nonaka, I., A Dynamic Theory of Organizational Knowledge Creation, in: Organization Science: A Journal of the Institute of Management Sciences, 5 (1994) 1, S. 14-38

[Nonaka/Konno 1998]
Nonaka, I., Konno, N., The Concept of "Ba": Building a Foundation for Knowledge Creation, in: California Management Review, 40 (1998) 3, S. 40-54

[Nonaka/Takeuchi 1995]
Nonaka, I., Takeuchi, H., The Knowledge-Creating Company - How Japenese Companies Create the Dynamics of Innovation, Oxford University Press New York, New York, 1995

[North 1999]
North, K., Wissensorientierte Unternehmensführung - Wertschöpfung durch Wissen, Gabler, Wiesbaden, 1999

[Nusser 1998]
Nusser, S., Sicherheitskonzepte im WWW, Springer, Heidelberg, 1998

[Oppliger 1998]
Oppliger, R., Internet and Intranet Security, Artech House, Boston, 1998

[Ortner 1997]
Ortner, E., Methodenneutraler Fachentwurf: zu den Grundlagen anwendungsorientierter Informatik, B. G. Teubner Verlagsgesellschaft, Stuttgart, Leipzig, 1997

[Österle 1993]
Österle, H., Informationsmanagement 2000, in: Pressmar, D., Informationsmanagement, Gabler, Wiesbaden, 1993, S. 164-184

[Österle 1995]
Österle, H., Business Engineering: Prozess- und Systementwicklung, Band 1 (Entwurfstechniken), Springer, Berlin et al., 1995

[Österle 1996]
Österle, H., Integration: Schlüssel zur Informationsgesellschaft, in: Österle, H., Riehm, R., Vogler, P. (Hrsg.), Middleware: Grundlagen, Produkte und Anwendungsbeispiele für die Integration heterogener Welten, Vieweg, Braunschweig/Wiesbaden, 1996, S. 1-24

[Österle 1999]
Österle, H., Enterprise in the Information Age, in: Österle, H., Fleisch, E., Alt, R. (Hrsg.), Business Networking: Shaping Enterprise Relationships on the Internet, Springer, Berlin et al., 1999, S. 17-54

[Österle 2002a]
Österle, H., Geschäftsmodell des Informationszeitalters, in: Österle, H., Fleisch, E., Alt, R. (Hrsg.), Business Networking in der Praxis, Springer, Berlin, 2002, S. 17-38

[Österle 2002b]
Österle, H., Übergang zur Informationsgesellschaft (New Economy), in: Dubs, R., Euler, D., Rüegg-Stürm, J. (Hrsg.), Einführung in die Managementlehre, Verlag Paul Haupt, Bern, 2002, S. 329-349

[Österle/Blessing 2000]
Österle, H., Blessing, D., Business Engineering Model, in: Österle, H., Winter, R. (Hrsg.), Business Engineering: Auf dem Weg zum Unternehmen des Informationszeitalters, Springer, Berlin etc., 2000, S. 61-80

[Österle/Blessing 2003]
Österle, H., Blessing, D., Business Engineering Modell, in: Österle, H., Winter, R. (Hrsg.), Business Engineering, Springer, Berlin etc., 2003, S. 65-85

[Österle et al. 1992]
Österle, H., Brenner, W., Hilbers, K., Unternehmensführung und Informationssystem - Der Ansatz des St. Galler Informationssystem-Managements, Teubner, Stuttgart, 1992

[Österle et al. 1996]
Österle, H., Riehm, R., Vogler, P., Middleware - Grundlagen, Produkte und Anwendungsbeispiele für die Integration heterogener Welten, Verlag Vieweg, Wiesbaden, 1996

[Österle/Senger 2003]
Österle, H., Senger, E., Realtime Management - Fünf Fallstudien, Arbeitsbericht, Institut für Wirtschaftsinformatik, Universität St.Gallen, St. Gallen, 2003

[Österle/Winter 2003]
Österle, H., Winter, R., Business Engineering, in: Österle, H., Winter, R. (Hrsg.), Business Engineering, Springer, Berlin etc., 2003, S. 3-18

[Plumtree 2001]
Plumtree, An Overview of Corporate Portal Technology and Deployment - Survey Results From Organizations That Have Deployed Corporate Portals, Plumtree Software, San Francisco, 2001

[Polanyi 1966]
Polanyi, M., The Tacit Dimension, Routledge & Kegan Paul, Gloucester, 1966

[Porter 1984]
Porter, M.E., Wettbewerbsstrategie: Methoden zur Analyse von Branchen und Konkurrenten, Campus, Frankfurt, 1984

[Probst et al. 1999]
Probst, G.J.B., Raub, S., Romhardt, K., Wissen managen - Wie Unternehmen ihre wertvollste Ressource optimal nutzen, Gabler, Wiesbaden, 1999

[Puschmann 2003]
Puschmann, T., Collaboration Portale - Architektur, Integration, Umsetzung und Beispiele, Dissertation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2003

[Raepple 2001]
Raepple, M., Sicherheitskonzepte für das Internet - Grundlagen, Technologie und Lösungskonzepte für die kommerzielle Nutzung, dpunkt Verlag, Heidelberg, 2001

[Raghavan 2002]
Raghavan, P., Information Retrieval for Enterprise Content, in: UPGRADE (The European Online Magazine for the IT-Professional), III (2002) 3, S. 5-8

[Ramos 2002]
Ramos, L., Technology Overview: The Future of Enterprise Search, Report, Giga Information Group Inc., Cambridge, CA, 2002

[Rath 2003]
Rath, H., The Topic Maps Handbook, Empolis, Gütersloh, 2003

[Rehäuser/Krcmar 1996]
Rehäuser, J., Krcmar, H., Wissensmanagement in Unternehmen, in: Schreyögg, G., Conrad, P. (Hrsg.), Wissensmanagement, Managementforschung, Springer, Berlin, 1996, S. 1-40

[Reichmayr 2002]
Reichmayr, C., Collaboration und WebServices: Architekturen, Portale, Techniken und Beispiele, Dissertation, Universität St. Gallen, Difo-Druck, Bamberg, 2002

[Rich 1992]
Rich, P., The Organizational Taxonomy: Definition and Design, in: Academy of Management Review, 17 (1992) 4, S. 758-781

[Riehm 1997]
Riehm, R., Integration von heterogenen Applikationen, Dissertation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 1997

[Riempp 2003]
Riempp, G., Integrierte Wissensmanagement-Systeme in dienstleistungsorientierten Organisationen, Habilitation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2003

[Rijsbergen 1979]
Rijsbergen, C.J.K., Information Retrieval, 2nd, Butterworth, London, 1979

[Rodi 1992]
Rodi, F., Semiotik, in: Seiffert, H., Radnitzky, G. (Hrsg.), Handlexikon zur Wissenschaftstheorie, München, 1992, S. 296-302

[Röhricht/Schlögel 2001]
Röhricht, J., Schlögel, C., cBusiness - Erfolgreiche Internetstrategien durch Collaborative Business, Addison-Wesley, München, 2001

[Rosenfeld/Morville 2002]
Rosenfeld, L., Morville, P., Information Architecture for the World Wide Web - Designing large-scale websites, O'Reilly, Sebastopol, 2002

[Salton/McGill 1987]
Salton, G., McGill, M.J., Information Retrieval - Grundlegendes für Informationswissenschaftler, McGraw-Hill, Hamburg, New York, 1987

[Sanchez 1997]
Sanchez, R., Managing Articulated Knowledge in Competence-based Competition, in: Sanchez, R., Heene, A. (Hrsg.), Strategic Learning and Knowledge Management, Chichester, 1997, S. 163-187

[SAP 2003]
SAP, mySAP Enterprise Portal - SAP Enterprise Portal 6.0 - Technologie, SAP, https://websmp101.sap-ag.de/~form/sapnet?_FRAME=CONTAINER&_OBJECT=011000358700000644092003D, 25.08.2003

[Schackmann/Schü 2001]
Schackmann, J., Schü, J., Personalisierte Portale, in: Wirtschaftsinformatik, Vol. 43 (2001) 6, S. 623-625

[Scheer 1998]
Scheer, A.-W., ARIS - Vom Geschäftsprozess zum Anwendungssystem, 3. Aufl., Springer, Berlin etc., 1998

[Schreyögg 2001]
Schreyögg, G., Wissen, Wissenschaftstheorie und Wissensmanagement, in: Schreyögg, G. (Hrsg.), Wissen in Unternehmen: Konzepte, Massnahmen, Methoden, Erich Schmidt Verlag, Berlin, 2001, S. 3-17

[Schulze 2000]
Schulze, J., Prozessorientierte Einführungsmethode für das Customer Relationship Management, Dissertation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2000

[Schüppel 1996]
Schüppel, J., Wissensmanagement. Organisatorisches Lernen im Spannungsfeld von Wissens- und Lernbarrieren, Gabler, Wiesbaden, 1996

[Schwarz 2000]
Schwarz, J., Mass Customization von Prozessen durch Unternehmensportale, in: Information Management & Consulting, Vol. 15 (2000) Issue 2, S. 40-45

[Senge 1990]
Senge, P.M., The fifth Discipline - The art and practice of the learning organization, Random House, London, 1990

[Senger 2004]
Senger, E., Zum Stand der elektronischen Kooperation - Fallstudien, Muster und Handlungsoptionen, Dissertation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2004

[Shilakes/Tylman 1998]
Shilakes, C.C., Tylman, J., Enterprise Information Portal, Merrill Lynch & Co., New York, NY, 1998

[Sleeper/Robins 2001]
Sleeper, B., Robins, B., Defining Web Services, Report, The Stencil Group, www.stencilgroup.com, 2001

[Smolnik/Nastansky 2002]
Smolnik, S., Nastansky, L., K-Discovery: Using Topic Maps to identify Distributed Knowledge Structures in Groupware-based Organizational Memories, in: Proceedings of the Hicss 35, IEEE Computer Society Press, Big Island, Hawaii, 2002, S. CD-ROM, 10

[Sparck Jones/Willett 1997]
Sparck Jones, K., Willett, P., Readings in Information Retrieval, Morgan Kaufmann Publishers, San Francisco, 1997

[Staab 2002]
Staab, S., Wissensmanagement mit Ontologien und Metadaten, in: Informatik Spektrum, 25 (2002) 3, S. 194-209

[Stahlknecht/Hasenkamp 2002]
Stahlknecht, P., Hasenkamp, U., Einführung in die Wirtschaftsinformatik, Springer, Berlin, 2002

[Stake 1995]
Stake, R.E., The Art of Case Study Research, Sage Publications, London, 1995

[SUN 2004]
SUN, Java Authentication and Authorization Service (JAAS), SUN, http://java.sun.com/products/jaas/index.jsp, 11.03.2004

[Surmacz 2003]
Surmacz, J., Needle in a Haystack, CIO, 2003

[Sveiby 1997]
Sveiby, K.E., A Knowledge-based Theory of the Firm to guide Strategy Formulation, http://www.sveiby.com/Knowledgetheoryoffirm.htm, 12.03.2003

[Thiesse 2001]
Thiesse, F., Prozessorientiertes Wissensmanagement - Konzepte, Methode, Fallbeispiele, Dissertation, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, 2001

[Thommen 2002]
Thommen, J.P., Betriebswirtschaftslehre, Versus Verlag AG, Zürich, 2002

[Tidwell 2002]
Tidwell, D., Xslt, O'Reilly, Cambridge et al., 2002

[Ulrich 1984]
Ulrich, H., Die Betriebswirtschaftslehre als anwendungsorientierte Sozialwissenschaft, in: Ulrich, H., Dyllick, T., Probst, G. (Hrsg.), Management, Haupt, Bern, 1984, S. 170-195

[Ulrich/Fluri 1995]
Ulrich, P., Fluri, E., Management, Paul Haupt, Bern/Stuttgart, 1995

[Verity 2003]
Verity, Branchenspezifisches Klassifizieren von Content mit Verity und Factiva, Verity, http://www.verity.com/de/company/press/2003/de_031030.html, 10.12.2003

[von Krogh et al. 2000]
von Krogh, G., Ichijo, K., Nonaka, I., Enabling Knowledge Creation: How to Unlock the Mystery of Tacit Knowledge and Release the Power of Innovation, Oxford University Press, New York, 2000

[von Krogh/Roos 1995]
von Krogh, G., Roos, J., A perspective on knowledge, competence and strategy, in: Personnel Review, 24 (1995) 3, S. 56-76

[Warzecha 2001]
Warzecha, A., B2E Best Practices: Picking the low-hanging fruit, Meta Group, 2001

[Waveset 2001]
Waveset, T., Technology for successfull Access Management, Waveset Technologies Inc., Austin, 2001

[Wayne 1991]
Wayne, E., Terminology Update: Words versus Terms: Is there a Difference?, in: Standardization News, (1991) November 1991, S. 17

[Wenger 1997]
Wenger, E., Communities of Practice: Learning, Meaning, and Identity, Cambridge University Press, Cambridge, 1997

[White 2000]
White, M., Enterprise Information Portals, in: The Electronic Library, 18 (2000) 5, S. 354-362

[White 2002]
White, M., Behind the Firewall - Buying the Proper Search Solution for Your Intranet, EContent Magazine, www.econtentmag.com/Magazine/Columns/02/firewall2_02.htm, 30.08.2002

[Whyte 1991]
Whyte, W.F., Participatory Action Research, Sage Publications, Newbury Park etc., 1991

[Wiig 1995]
Wiig, K.M., Knowledge Management A Trilogy - Volume 3, Knowledge Management Methods: Practical Approaches to Managing Knowledge, Schema Press, Arlington, 1995

[Wiig 1999]
Wiig, K.M., Introducing Knowlege Management into the Enterprise, in: Liebowitz, J. (Hrsg.), Knowledge Management Handbook, CRC Press, Boca Raton, 1999, S. 1-41

[Yin 1994]
Yin, R.K., Case Study Research. Designs and Methods, SAGE Publications, London, 1994

[Zahn 1998]
Zahn, E., Wissen und Strategie, in: Bürgel, H.D. (Hrsg.), Wissensmanagement - Schritte zum intelligenten Unternehmen, Springer, Berlin/Heidelberg, 1998

[Zahn et al. 2000]
Zahn, E., Foschiani, S., Tilebein, M., Wissen und Strategiekompetenz, in: Hammann, P., Freiling, J. (Hrsg.), Die Ressourcen- und Kompetenzperspektive des Strategischen Managements, Deutscher Universitäts-Verlag, 2000, S. 47-64

[Zwerger/Schneider 2002]
Zwerger, F., Schneider, G., Sichere Unternehmensportale mit SAP, Galileo Press, Bonn, 2002

Lebenslauf

Persönliche Daten

Name, Vorname	Kremer, Stefan
Geboren am	30.07.1970
Geburtsort	Hilden, Deutschland
Familienstand	verheiratet, eine Tochter
Staatsangehörigkeit	Deutsch

Ausbildung und Wehrdienst

2000 – 2004	Doktorandenstudium an der Universität St.Gallen
1991 – 1999	Studium des Wirtschaftsingenieurwesens an der Universität Paderborn, Abschluss mit akademischem Grad Diplom-Wirtschaftsingenieur
1990 – 1991	Grundwehrdienst bei der Deutschen Bundeswehr
1987 – 1990	Kollegschule Kikweg in Düsseldorf, Bildungsgang Abitur / Wirtschaft
1981 – 1987	Gesamtschule Kikweg in Düsseldorf
1977 – 1981	Grundschule in Düsseldorf

Berufstätigkeit

2000 – 2004	Wissenschaftlicher Mitarbeiter am Institut für Wirtschaftsinformatik der Universität St.Gallen, Prof. Dr. Walter Brenner und Prof. Dr. Hubert Österle
1999 – 2000	Berater bei der PricewaterhouseCoopers Unternehmensberatung GmbH, Frankfurt