

Entwurf und Entwicklung eines semantisch gestützten Dokumentenmanagementsystems für eine große dezentrale Organisation

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Software Engineering & Internet Computing

eingereicht von

Johann Grabner

Matrikelnummer 0226102

an der
Fakultät für Informatik der Technischen Universität Wien

Betreuung
Betreuer: Thomas Grechenig

Wien, 27.09.2011

(Unterschrift Verfasser/in)

(Unterschrift Betreuer/in)



Entwurf und Entwicklung eines semantisch gestützten Dokumentenmanagementsystems für eine große dezentrale Organisation

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Software Engineering & Internet Computing

eingereicht von

Johann Grabner

0226102

ausgeführt am

Institut für Rechnergestützte Automation

Forschungsgruppe Industrial Software

der Fakultät für Informatik der Technischen Universität Wien

Betreuung:

Betreuer: Thomas Grechenig

Wien, 27.09.2011

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benützt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am

Name

Danksagung

An dieser Stelle möchte ich meinen persönlichen Dank an alle Personen aussprechen, die mir die Durchführung dieser Diplomarbeit ermöglicht haben.

Mein spezieller Dank geht an Univ.-Prof. Dipl.-Ing. Dr. techn. Thomas Grechenig für die Bereitstellung einer praxisnahen Aufgabe, die mir eine persönliche und fachliche Weiterentwicklung ermöglicht hat.

Des Weiteren bedanke ich mich bei allen beteiligten Mitarbeitern des Projekts „TU Wien Informations-Systeme und Services“ und der Organisationseinheiten „Organisation und Koordinaten“ und „Gebäude und Technik“ der Technischen Universität Wien für ihre Kooperation.

Kurzfassung

Das World Wide Web befindet sich im Übergang vom Web 2.0 zum Semantic Web, und die Berücksichtigung der Semantik von Informationen bei der Verarbeitung gewinnt zunehmend an Bedeutung. Im Rahmen der vorliegenden Arbeit werden die Grundlagen des Semantic Webs diskutiert und die aktuelle wissenschaftliche Auseinandersetzung mit dem Bereich des Dokumentenmanagements dargestellt.

Im praktischen Teil dieser Arbeit wird für die Organisationseinheiten „Organisation und Koordination“ und „Gebäude und Technik“ der Technischen Universität Wien ein Dokumentenmanagementsystem in Form einer semantischen Web-Applikation entworfen und als Prototyp implementiert. Anhand der Bewältigung der Aufgabestellung aus der Praxis wird gezeigt, welche Möglichkeiten sich beim Einsatz semantischer Technologien im Dokumentenmanagement eröffnen. Das abgebildete Wissen aus dem Anwendungsbereich liefert die Basis und den Kontext für die automatisierte Verarbeitung von Daten. Die Arbeit resultiert in der Entwicklung eines Systems zur kollaborativen Beschlagwortung von Dokumenten, das den Benutzer bei der Ablage und Suche mit fachlichen Empfehlungen assistiert.

Keywords: *Dokumentenmanagementsystem, Semantisches Web, Softwareentwicklung, Web-Applikation, Folksonomie*

Abstract

The transition from Web 2.0 into the Semantic Web represents the currently ongoing step in the evolution of the World Wide Web and the consideration of semantics in information processing gets more important. This master thesis discusses the foundations of the Semantic Web and describes the latest scientific research activities in the area of document management.

The practical part of this work contains the design and prototyping of a semantic web-based document management system for the organisational units “Organisation und Koordination” and “Gebäude und Technik” of the Vienna University of Technology. Specifically, this work illustrates the use of semantic technologies to improve document management in a real business scenario. The processing of information is based on the knowledge and context of the application domain. The result is a collaborative tagging system for documents that assists its users by providing semantically meaningful recommendations of tags for categorization and navigation.

Keywords: *Document Management System, Semantic Web, Software Engineering, Web Application, Folksonomy*

Inhaltsverzeichnis

Inhaltsverzeichnis.....	I
Abbildungsverzeichnis	III
Tabellenverzeichnis	V
1 Einleitung	1
1.1 Problemstellung	1
1.2 Motivation	2
1.3 Zielsetzung	2
1.4 Aufbau der Arbeit.....	4
2 Grundlagen	6
2.1 Dokumentenmanagement	7
2.1.1 Leistungsspektrum aktueller Systeme	8
2.1.2 Wissenschaftliche Aktivitäten	10
2.2 Web 2.0	11
2.3 Semantisches Web.....	15
2.4 Software- und Web-Engineering.....	18
3 Methodik	23
3.1 Entwicklungsprozess	23
3.2 Dokumentation.....	25
4 Anforderungsanalyse	28
4.1 Ausgangszustand	28
4.1.1 Kanzlei.....	28
4.1.2 Gebäude und Technik	32
4.2 Probleme der Anwender	32
4.3 Anforderungen	39
4.3.1 Funktionale Anforderungen	41
4.3.2 Nicht funktionale Anforderungen	46
5 Entwurf.....	47
5.1 Systemkonzept	47
5.1.1 Vereinheitlichung individueller Beschreibungen	49
5.1.2 Unterstützung von Homonymen und Synonymen	49
5.1.3 Folksonomy mit semantischen Erweiterungen	49
5.1.4 Vorschläge bei Navigation und Schlagwortvergabe	51
5.2 Software-Architektur	52
5.2.1 Szenarien-Sicht	53
5.2.2 Logische Sicht	66

5.2.3	Prozesssicht	76
5.2.4	Entwicklungssicht	77
5.2.5	Physische Sicht	79
5.3	Datenmigration	80
6	Realisierung	85
6.1	Eingesetzte Technologien.....	85
6.1.1	Ruby on Rails	86
6.1.2	MySQL.....	88
6.1.3	Yahoo! User Interface Library.....	88
6.2	Benutzerschnittstelle.....	89
6.3	Suchfunktion	100
6.4	Optimierung	102
7	Zusammenfassung	107
	Literaturverzeichnis	110
	Anhang.....	115
	Glossar	115
	Kanzlei.....	115
	Gebäude und Technik	117

Abbildungsverzeichnis

Abbildung 1: Schlagwortwolke „Web 2.0“ aus [Ange05].....	11
Abbildung 2: Projektmanagement mit Redmine.....	27
Abbildung 3: Schematischer Aufbau einer Geschäftszahl	29
Abbildung 4: Textversion der KIS-Hauptmaske	31
Abbildung 5: KIS-Hauptmaske in grafischer Oberfläche	31
Abbildung 6: Anwendungsfalldiagramm für Akten.....	56
Abbildung 7: Anwendungsfalldiagramm für Schlagwörter.....	57
Abbildung 8: Anwendungsfalldiagramm der Kanzlei.....	57
Abbildung 9: Anwendungsfalldiagramm der „Gebäude und Technik“	57
Abbildung 10: Klassendiagramm „Triplestore“	66
Abbildung 11: Inhaltliche Wissensabbildung im Triplestore.....	69
Abbildung 12: Räumliche und zeitliche Wissensabbildung im Triplestore.....	70
Abbildung 13: Klassendiagramm „Triplestore und relationale Tabellen“	72
Abbildung 14: Klassendiagramm „Versionierung von Entitäten“	72
Abbildung 15: Klassendiagramm „Datenbankschema“.....	73
Abbildung 16: Sequenzdiagramm „Netzwerkcommunication“	77
Abbildung 17: Komponentendiagramm des Systems	78
Abbildung 18: Verteilungsdiagramm „TISS-Infrastruktur“	80
Abbildung 19: Beispiel für die Extraktion von Schlagworten.....	84
Abbildung 20: Startseite mit schematischen Layout.....	90
Abbildung 21: Startseite der Applikation	91
Abbildung 22: Angezeigte Informationen in der Eingabevervollständigung	91
Abbildung 23: Beispiel für Vervollständigung während der Eingabe	92
Abbildung 24: Schlagwort einer Liste hinzufügen.....	92
Abbildung 25: Schlagwort aus Liste entfernen.....	92
Abbildung 26: Formular für ein neues Schriftstück.....	93
Abbildung 27: Ausgefülltes Formular für ein neues Schriftstück.....	93
Abbildung 28: Formular zur Bearbeitung eines Schriftstücks.....	94
Abbildung 29: Optionale Anzeige des Datensatzes aus dem Altsystem	95
Abbildung 30: Wahl des Überbegriffs eines neuen Schlagwortes	95
Abbildung 31: Ausgefülltes Formular für ein neues Schlagwort.....	96
Abbildung 32: Formular zur Bearbeitung eines Schlagwortes.....	96
Abbildung 33: Suche nach Bearbeitungsschritten	97
Abbildung 34: Übersicht über laufende Fristen	97
Abbildung 35: Formular zur Umlagerung von Akten	98
Abbildung 36: Formular zur Erstellung von Berichten.....	98
Abbildung 37: Persönliche und gemeinsame Notizen	99
Abbildung 38: Protokoll der Benutzeraktionen	99

Abbildung 39: Suche nach Schlagworten	100
Abbildung 40: Suche nach Überbegriff	100
Abbildung 41: Suche nach Schriftstückattributen	101
Abbildung 42: Suche nach Text	101
Abbildung 43: Suche nach Zeitraum	101
Abbildung 44: Suche nach Geschäftszahl	102
Abbildung 45: Optionen der Eingabevervollständigung bei der Suche	102
Abbildung 46: Klassendiagramm „Optimierungstabellen“	104

Tabellenverzeichnis

Tabelle 1: Unterschiede zwischen DMS und RMS.....	8
Tabelle 2: Prägnante Webseiten als Beispiele für Web 2.0.....	15
Tabelle 3: Wertschätzungen der agilen Softwareentwicklung.....	19
Tabelle 4: Anwenderprobleme inklusive Herkunft nach Organisationseinheit...	39
Tabelle 5: Zusammenhang der adressierten Probleme und Anforderungen.....	40
Tabelle 6: Anforderungen inklusive Herkunft nach Organisationseinheit.....	45
Tabelle 7: Bedeutung der in Diagrammen verwendeten Begriffe.....	53
Tabelle 8: Allgemeine Anforderungen und Anwendungsfälle.....	54
Tabelle 9: Anforderungen der Kanzlei und Anwendungsfälle.....	55
Tabelle 10: Anforderungen der „Gebäude und Technik“ und Anwendungsfälle	55
Tabelle 11: Anwendungsfälle inklusive Herkunft nach Organisationseinheit.....	65
Tabelle 12: In Klassendiagrammen verwendete Datentypen.....	66
Tabelle 13: Verwendete Elemente der DCMI Metadata Terms.....	68
Tabelle 14: Diagrammnotation für Wissensabbildung im Triplestore.....	68
Tabelle 15: Bedeutung von mehrfach verwendeten Spaltennamen.....	74
Tabelle 16: Bedeutung der Datenbankspalten.....	76
Tabelle 17: Beispiele für Ablagebereiche des Kanzleinformationssystems.....	81

1 Einleitung

Die Kanzlei der Technischen Universität Wien ist zuständig für die Erfassung der per Post eingehenden Dokumente, die weitere Verteilung dieser Dokumente an die adressierten Organisationseinheiten und die anschließende Archivierung dieser Dokumente für einen Zeitraum von zehn Jahren. Zur Bewältigung dieser Aufgabe wird ein bereits jahrzehntelang im Einsatz befindliches Softwaresystem „KIS“ (Kanzleinformationssystem) verwendet. Dieses System befindet sich am Ende seines Lebenszyklus, weil es heutigen Anforderungen nicht mehr gerecht wird und Probleme mit der Verarbeitung der im Laufe der Jahre angesammelten Datenmenge auftreten.

In der Organisationseinheit „Gebäude und Technik“ besteht ebenfalls der Bedarf an einem System zur Verwaltung anfallender Dokumente. Die Aufgabe der „Gebäude und Technik“ ist die Bereitstellung und Instandhaltung der räumlichen und technischen universitären Infrastruktur. In der „Gebäude und Technik“ wurden bisher anfallende Dokumente nach organisatorischen Konventionen elektronisch oder in Papierform manuell abgelegt. Angesichts des zunehmenden Datenvolumens bleibt der entstehende Aufwand auf Dauer nicht tragbar.

Das Projekt „TU Wien Informations-Systeme und Services“¹ kurz „TISS“ der Technischen Universität Wien hat sich als Ziel gesetzt, ein einheitliches Gesamtsystem für die Anforderungen aus der Lehre, Forschung und Administration bereitzustellen. Das Dokumentenmanagement stellt als mehrfach wiederkehrende Anforderung einen Teilbereich des TISS-Projekts dar.

1.1 Problemstellung

Für die Kanzlei und die Organisationseinheit „Gebäude und Technik“ der Technischen Universität ist ein Prototyp eines Dokumentenmanagementsystems als Web-Applikation zu entwerfen und umzusetzen. Der Tätigkeitsbereich umfasst alle im Projekt auftretenden Aufgaben des Software- und Web-Engineerings, beginnend bei der laufenden Erfassung von Anforderungen, über die einzelnen Iterationen der Entwicklung hinweg, bis hin zur abschließenden Übernahme des Prototyps durch die Anwender.

Bei der Entwicklung des Prototyps müssen die unterschiedlichen Aspekte der beteiligten Organisationseinheiten erkannt und berücksichtigt werden. Die individuellen Bedürfnisse der Beteiligten sollen zusammen mit dem Wissen über die jeweiligen Anwendungsbereiche die Grundlage für den Prototyp bilden. Eine besondere Sorgfalt ist bei dem Design des Prototyps für die Kanzlei an den Tag zu legen, um das über Jahrzehnte angesammelte Anwenderwissen nicht durch grundlegende Differenzen in der Verwen-

¹ siehe <https://tiss.tuwien.ac.at/> und http://www.zid.tuwien.ac.at/ueber_tiss/

derung zu entwerfen. Der Prototyp soll als eine sinnvolle Erweiterung und Verbesserung bereits bekannter Funktionen wahrgenommen werden. Die „Gebäude und Technik“ hingegen legt speziellen Wert auf eine Integration in ihre bestehende Softwarelandschaft über entsprechende Schnittstellen wie zum Beispiel den Import von Daten aus SAP. Die Daten der Kanzlei und der „Gebäude und Technik“ müssen in den Datenbestand des Prototyps migriert werden. Hierzu ist die Konzeption einer Migrationsstrategie notwendig.

Die theoretischen Grundlagen dieser Arbeit sind in folgenden Bereichen zu finden:

- Dokumentenmanagement
- Semantisches Web und Web 2.0
- Software- und Web-Engineering

Neben der Ausarbeitung des aktuellen Wissensstandes in den zu Grunde liegenden Themenbereichen erfordert die Aufgabestellung die praktische kombinierte Anwendung unterschiedlicher Ingenieursdisziplinen aus dem Bereich des Software-Engineerings. Zur Anwendung kommen Methoden aus der agilen Software-Entwicklung, dem Requirement-Engineering und der Qualitätssicherung.

1.2 Motivation

Die wissenschaftliche Herausforderung liegt in der Konzeption und Umsetzung eines innovativen und semantischen Dokumentenmanagementsystems, das die Fähigkeiten von gebräuchlichen, auf Textvergleich basierten Suchmaschinen übersteigt. Ein weiterer Fokus liegt auf der intensiven Einbindung der Anwender mit ihren Vorkenntnissen im Entwicklungsprozess, um den Erwartungen an ein neues Dokumentenmanagementsystem gerecht werden zu können. Diese Schwerpunkte folgen dem semantischen Trend im Web und bilden zusammen mit der Anforderung, aktuelle Web-Technologien als Benutzerschnittstelle und zur Repräsentation von Informationen einzusetzen, eine zeitgemäße Aufgabenstellung.

Ein offener Entwicklungsprozess mit intensivem Anwenderkontakt und technischen Herausforderungen stellt neben dem praktischen Nutzen für die Anwender ein anspruchsvolles und attraktives Projekt dar. Die Berücksichtigung von Wissen aus dem Anwendungsbereich ermöglicht potenzielle Innovationen im Bereich des Dokumentenmanagements.

1.3 Zielsetzung

Das Ergebnis dieser Diplomarbeit soll ein in der Praxis verwendbarer Prototyp eines Dokumentenmanagementsystems sein, das die Semantik des Anwendungsbereichs mit einbezieht. Unter der Einbeziehung der Semantik wird verstanden, dass im System die Bedeutung und der Zusammenhang von Begriffen aus dem Anwendungsbereich be-

kannt sind. Um ein Beispiel zu nennen, würde im System mit dem Begriff „Fakultät“ die universitäre Organisationsstruktur in Verbindung gebracht. Durch die Kenntnis des Anwendungsbereichs werden dem System neue Möglichkeiten zur Assistenz des Benutzers bei seinen Tätigkeiten erschlossen.

Die Berücksichtigung der zu Grunde liegenden semantischen Zusammenhänge von Begriffen soll sich für die Anwender bei der Suche und Erfassung von Dokumenten in einer permanenten Unterstützung durch vom System vorgeschlagene inhaltliche und kontextabhängige Empfehlungen äußern. Das System soll diese Empfehlungen zum einen aus dem Wissen des Anwendungsbereichs und zum anderen aus der Verwendungsweise des Dokumentenmanagementsystems durch die Anwender ableiten. Für eine natürliche Benutzung des Systems ist auf sprachliche Besonderheiten wie gebräuchliche Abkürzungen, die Mehrdeutigkeit eines Wortes und Worte mit gleicher oder ähnlicher Bedeutung Rücksicht zu nehmen.

Im Kontext der Problemstellung werden folgende grundlegende Ansprüche von der Leitung des Projekts „TISS“ an ein neues Dokumentenmanagementsystem gestellt:

- *Umgang mit Dokumenten in elektronischer Form und Papierform*
Aus Gründen der Flexibilität und Vereinbarkeit mit bisherigen Ablagen muss das Dokumentenmanagementsystem elektronische Dokumente, Dokumente in Papierform und in beiden Formen vorliegende Dokumente verarbeiten können.
- *Einfaches Auffinden von Dokumenten*
Die Qualität eines Dokumentenmanagementsystems wird neben der Ablagefunktion primär durch seine Suchfunktion bestimmt. Die Entwicklung einer modernen, fortschrittlichen Suchmaschine für Dokumente wird als zentraler Punkt der Aufgabenstellung verstanden.
- *Aufhebung der Limitationen des Kanzleiinformationssystems*
Durch das vorangeschrittene Alter des Kanzleiinformationssystems enthält es Limitationen wie die Beschränkung der Anzahl angezeigter Suchergebnisse. Solche Einschränkungen sind aufzuheben.
- *Kompatibilität zu bestehenden Arbeitsabläufen*
Eine neue Software muss die Anforderungen der Kanzlei und der Organisationseinheit „Gebäude und Technik“ der Technischen Universität Wien abdecken, ohne für die etablierten Arbeitsabläufe inkompatibel zu werden.

Abschließend anzuführen ist der Bedarf einer detaillierten Vorgehensweise zur Übernahme der Daten aus dem Kanzleiinformationssystem.

1.4 Aufbau der Arbeit

Nach der Einleitung in diesem ersten Kapitel wird ein Überblick über die grundlegenden Themen dieser Diplomarbeit gegeben. Das zweite Kapitel beschäftigt sich speziell mit dem Dokumentenmanagement und erklärt die Grundideen von Semantic Web und Web 2.0. Ebenfalls enthalten sind für diese Diplomarbeit relevante Einblicke in das Software- und Web-Engineering. Wichtige Begrifflichkeiten werden erläutert und der Bezug zum theoretischen Hintergrundwissen wird hergestellt. Kapitel 2 liefert die notwendigen Grundlagen zum Verständnis der weiteren Abschnitte.

In Kapitel 3 wird die Vorgehensweise zur Erreichung der Zielsetzung beschrieben. Der Entwicklungsprozess wird durch eine begründete Selektion von bestehenden Methoden aus der agilen Softwareentwicklung zusammengestellt.

Kapitel 4 dokumentiert die Ergebnisse der vorgenommenen Anforderungsanalyse. Nach der Dokumentation des Ausgangszustands wird auf die bestehenden Probleme der Anwender im Dokumentenmanagement eingegangen. Das Kapitel endet mit der Beschreibung der erfassten Anforderungen.

Das ganzheitliche Systemkonzept für den Prototyp des Dokumentenmanagementsystems ist Thema des fünften Kapitels. Ausgehend von der Arbeitsweise und den Anforderungen der beiden Anwendergruppen werden vorgenommene Designentscheidungen offen gelegt und begründet. Als Resultat wird ein vollständiges und technologieunabhängiges Systemkonzept präsentiert, das die Anforderungen aus der Praxis erfüllen soll. Anhand von Diagrammen wird die erstellte Software-Architektur illustriert und dokumentiert. Die Definition einer Vorgehensweise für die Übernahme der Daten aus dem Altsystem der Kanzlei schließt Kapitel 5 ab.

Kapitel 6 setzt sich mit der softwaretechnischen Umsetzung des im vorangegangenen Kapitel festgehaltenen Konzepts auseinander. Es wird Aufschluss über die verwendeten Technologien gegeben und der Einsatz dieser wird gerechtfertigt. Markante Merkmale der technologischen Basis werden auf den Punkt gebracht. Um die Funktion des Systems zu veranschaulichen, werden die zentralen Elemente der Benutzerschnittstelle gezeigt und erläutert. Dabei wird aus einem Standpunkt der Anwendung ein Einblick in die essentielle Funktionalität der Software gewährt.

Zum Abschluss werden die Ergebnisse dieser Arbeit im siebten Kapitel zusammengefasst. Ein Resümee wird über den Projektverlauf und die erarbeiteten Erkenntnisse gezogen. Um die Erreichung der Zielsetzung beurteilen zu können, werden die Ergebnisse mit den gesetzten Zielen verglichen. Dieses Kapitel beinhaltet eine kritische Reflexion der gewählten Arbeitsweisen und versucht aus der Arbeit gelernte Aspekte hervorzuheben. Zum einen werden aus der Praxis gewonnene Erkenntnisse angeführt, zum anderen

die Vor- und Nachteile der Herangehensweise aufgezeigt. Zusätzlich wird auf zukünftige weiterführende Tätigkeiten und Erweiterungsmöglichkeiten eingegangen. Die wissenschaftliche und praktische Relevanz der behandelten Themen wird vermittelt.

2 Grundlagen

Dieses Kapitel fasst für diese Arbeit grundlegendes Wissen zusammen und dient als thematische Einführung für die folgenden weiteren Kapitel.

Begonnen wird mit der Betrachtung des Themas „Dokumentenmanagement“ im Fokus dieser Arbeit. Nach einer Begriffsdefinition werden zu diesem Zweck die Eigenschaften eines Dokumentenmanagementsystems beschrieben. Zur Abgrenzung des Themas werden die typischen Unterschiede zwischen einem elektronischen Dokumentenmanagementsystem und einer im Englischen als „Electronic Records Management System“ bezeichneten Applikationsart genannt. Danach wird das Leistungsspektrum erhältlicher Dokumentenmanagementsysteme beschrieben und die aktuelle wissenschaftliche Auseinandersetzung mit dem Dokumentenmanagement dargestellt.

Die Entwicklung einer zeitgemäßen Web-Applikation erfordert die Auseinandersetzung mit dem Web in seiner aktuellen Form. Des Weiteren ist die kollaborative Kategorisierung von Inhalten durch eine Gemeinschaft von Benutzern eine im Web häufig anzutreffende Systematik, die Ähnlichkeiten zu der Ablage von Dokumenten im Dokumentenmanagement aufweist. Das semantische Web thematisiert die Berücksichtigung der Bedeutung von Informationen. Aus diesen genannten Gründen sind die beiden nächsten Abschnitte dem Web 2.0 und dem semantischen Web gewidmet.

Der erste dem Web gewidmete Abschnitt verfolgt das Ziel eine pragmatische Definition für den Begriff „Web 2.0“ zu liefern. Die Priorität liegt dabei auf der Beschreibung der dem Web 2.0 zugeschriebenen Veränderungen in der Gebrauchsweise durch die Benutzer, anstatt einer subjektiven Abgrenzung des Begriffs. Ein Überblick über typische Arten von Websites inklusive Beispielen zeigt die praktische Bedeutung von Web 2.0 im täglichen Gebrauch. Eine Zusammenfassung der zur Realisierung eingesetzten Schlüsseltechnologien vervollständigt den Abschnitt.

Anschließend wird das semantische Web als Thema behandelt. Ähnlich wie im vorangegangenen Abschnitt wird nach einer einleitenden Begriffsdefinition das Konzept des semantischen Webs anhand seiner Prinzipien erklärt und auf Technologien zur Umsetzung eingegangen.

Vollendet wird dieses Kapitel im letzten Abschnitt mit einer zeitgemäßen Definition für den Begriff „Software Engineering“. Zusätzlich wird, nachdem der Leitsatz der agilen Softwareentwicklung inhaltlich wiedergegeben wurde, die Beziehung zwischen Software und Web Engineering geklärt.

2.1 Dokumentenmanagement

In [GöSc08] wird „Dokumentenmanagement“ als umfassender Begriff für alle Prozesse, Abläufe und Verantwortlichkeiten, die mit der Administration von Dokumenten zusammenhängen, beschrieben. Die zitierte Quelle führt die im Umgang mit Dokumenten anfallenden Aufgaben der folgenden Auflistung als inhaltliche Schwerpunkte des Dokumentenmanagements an:

- Kennzeichnung und Beschreibung
- Verwaltung der Versionen und Historien
- Ablage und Archivierung
- Verteilung und Umlauf im Rahmen von Geschäftsprozessen
- Suche
- Schutz vor Verfälschung, Missbrauch und Vernichtung
- Langfristige Sicherstellung des Zugriffs und der Lesbarkeit
- Regelung der Lebenszyklen inklusive Vernichtung
- Regelung der Verantwortlichkeiten für Inhalt und Verwaltung

Elektronische Dokumentenmanagementsysteme bilden ein Werkzeug zur Unterstützung der Anwender bei der Erledigung der soeben genannten Tätigkeiten. Laut [GuSe02] ist ein Dokumentenmanagementsystem zuständig für diese Aufgaben:

- Erfassung von Informationen in elektronischer Form
- Konvertierung von Informationen zur Archivierung in ein geeignetes Format
- Erfassung von Metadaten zu Dokumenten zur späteren Wiederauffindung
- Sichere Ablage auf geeigneten Speichermedien
- Wiederauffindung von gespeicherten Informationen
- Bereitstellung der Daten für den Anwender

Dokumentenmanagementsysteme beinhalten in der Praxis oftmals Funktionen aus dem „Records Management“, deshalb ist für das Begriffsverständnis eine sorgfältige Abgrenzung hilfreich. In der frei beziehbaren Spezifikation „Model Requirements for the Management of Electronic Records“² kurz „MoReq“ der Europäischen Kommission werden Unterscheidungsmerkmale für Systeme aus den Bereichen Dokumentenmanagement und „Records Management“ genannt. Ein Hauptunterscheidungsmerkmal ist der unterschiedliche Fokus der beiden Systemarten. Während Systeme im Bereich des Dokumentenmanagements primär die im alltäglichen Umgang mit Dokumenten anfallenden Aufgaben erfüllen sollen, liegt der Schwerpunkte beim „Records Management“ auf einer sicheren Aufbewahrung von Akten.

² siehe <http://moreq.info> und http://ec.europa.eu/transparency/archival_policy/moreq/

Im Inhalt der Spezifikation sind weitere konkrete Unterscheidungskriterien angeführt, welche diese Aussage widerspiegeln. Die Gegenüberstellung in Tabelle 1 beinhaltet einige dieser Kriterien, um die Unterschiede zwischen den beiden verwandten Systemarten zu verdeutlichen. In der Praxis sind in Implementierungen der beiden beschriebenen Systemarten häufig Funktionsteile der anderen inkludiert.

Charakteristische Merkmale	
Dokumentenmanagementsystem (DMS)	„Records Management“-System (RMS)
Dokumente dürfen verändert werden.	Akten werden vor Änderungen geschützt.
Dokumente können in verschiedenen Versionen vorhanden sein.	Eine Akte existiert im System nur in ihrer endgültigen Version.
Dokumente können von ihren Eigentümern gelöscht werden.	Die Löschung von Akten erfolgt nach vorgegebenen Prozessen und ist streng kontrolliert.
Die Struktur der Dokumentenablage kann unter Kontrolle der Benutzer sein.	Die Klassifizierung von Akten erfolgt nach zentral vorgegebenen Richtlinien und ist unter Kontrolle einer administrativen Einheit.

Tabelle 1: Unterschiede zwischen DMS und RMS

2.1.1 Leistungsspektrum aktueller Systeme

Das Leistungsspektrum von frei und kommerziell erhältlichen Dokumentenmanagementsystemen ist breit gefächert. Das Produktangebot beginnt bei Software für Einzelpersonen und erstreckt sich bis hin zu Lösungen für international tätige Organisationen. Neben den Produkten mit einem allgemeinen Zugang zum Dokumentenmanagement existieren auch Systeme mit Spezialisierung auf eine Branche. Dieser Abschnitt fasst zusammen, welche besonderen Leistungen von aktuellen Dokumentenmanagementsystemen angeboten werden, ohne sich auf ein konkretes Produkt zu beschränken. Es wird anhand von Beispielen gezeigt, wie sich die Erfüllung der im Abschnitt 2.1 angeführten Aufgaben eines Dokumentenmanagementsystems in der Praxis äußert. Die folgende Aufzählung beinhaltet eine Auswahl von nennenswerten Funktionen der verfügbaren Systeme:

- *Automatisierte Digitalisierung*
Bei der elektronischen Erfassung von Dokumenten aus Papier wird der Benutzer beim Scanvorgang unterstützt. Vom System bereitgestellte Deckblätter werden beim Scannen zur automatisierten Ablage und Zuteilung eines Dokuments beigelegt. Für die Indizierung wird der Text im Dokument automatisch erkannt
- *Einbindung von Druckaufträgen*
An Drucker gesendete Aufträge werden überwacht und vollautomatisch als Dokumente im System abgelegt. Im Druckdatenstrom vorhandene Attribute werden extrahiert und als Metadaten mit den Dokumenten verknüpft.

- *Einbindung von E-Mails*
Die Kommunikation per E-Mail wird erfasst und für berechtigte Personen einsehbar. Auf diesem Weg werden gesetzliche Anforderungen an die Dokumentation der elektronischen Geschäftskommunikation erfüllt.
- *Einbindung von Ordnern im Dateisystem*
Im Dateisystem vorhandene Ordner werden für den automatischen Import neuer Dokumente überwacht.
- *Integration in Office-Anwendungen*
In Office-Anwendungen gespeicherte Dokumente werden importiert und können für die weitere Bearbeitung wieder in der Applikation geöffnet werden. Zusätzlich zur Ablage des Originaldokuments wird die Datei in ein Format zur Langzeitspeicherung konvertiert.
- *Markierungen und Kommentare*
Verwaltete Dokumente können mit Markierungen, Notizen und Kommentaren versehen werden. Danach sind diese Dokumente wahlweise inklusive oder exklusive der beigefügten Informationen beziehbar.
- *Sicherheit*
Berechtigungen werden eingesetzt um Inhalte vor unbefugten Zugriff zu schützen. Digitale Signaturen werden verwendet um die Authentizität und Echtheit von Dokumenten zu gewährleisten.
- *Unterstützung der Kollaboration*
Das gemeinschaftliche Arbeiten wird durch bereitgestellte Funktionen gefördert. Aufgabenlisten und Kalender ermöglichen eine zeitliche Planung und Koordination von Tätigkeiten. Dokumente können zu einem Zeitpunkt gemeinsam von mehreren Personen bearbeitet werden.
- *Volltextsuche für Attribute und Inhalte*
Inhalte und Attribute von Dokumenten werden für die Suche indiziert. Bei der Angabe von Suchkriterien können Platzhalter und logische beziehungsweise relationale Operatoren verwendet werden. Begriffe für die Suche im Inhalt der Dokumente können mit Kriterien für Attribute kombiniert werden.
- *Workflows*
Die Abwicklung betrieblicher Prozesse wird unterstützt. Dokumente werden automatisch an die zuständigen Stellen zur Bearbeitung beziehungsweise Bestätigung verteilt und falls notwendig weitergeleitet.

2.1.2 Wissenschaftliche Aktivitäten

Dieser Abschnitt stellt einen Auszug der aktuellen wissenschaftlichen Auseinandersetzung mit dem Bereich des Dokumentenmanagements dar und soll über den Inhalt dieser Arbeit hinweg als Einstieg für eine vertiefende wissenschaftliche Lektüre dienen. Ausgewählt wurden Arbeiten, die im Kontext der Entwicklung eines Dokumentenmanagementsystems gesehen werden. Die bei der Recherche ausgewählten wissenschaftlichen Arbeiten werden in der folgenden Aufzählung referenziert und nach deren Zugang zum Thema kategorisiert.

- *Persönliches Dokumentenmanagement*
In [BoJa05] werden die Resultate einer zweijährigen Studie über das persönliche Dokumentenmanagement der teilnehmenden Probanden präsentiert und Folgerungen für das Design eines Dokumentenmanagementsystems gezogen. Strategien für das Management persönlicher Dokumente werden in [HeSr11] identifiziert, um eine Hilfestellung für die Entwicklung neuer Werkzeuge zu bieten.
- *Systemarchitektur*
Eine auf den Technologien RDF und OWL basierende Architektur für Dokumentenmanagementsysteme wird in [MeKu04] vorgestellt. In [KoFu07] ist das Thema die Architektur eines Systems, das handschriftliche Notizen auf gedruckten Dokumenten digitalisiert und in Folge elektronisch abrufbar macht. Eine Systemarchitektur nach diversen Standards der „International Organization for Standardization“³ kurz „ISO“ wird in [SuGe08] beschrieben.
- *Benutzerinteraktion und Benutzeroberflächen*
Die Interaktionen des Benutzers mit einem System, in dem der Zugriff auf Dokumente primär nach räumlichen Kriterien erfolgt, werden in [EtMa06] behandelt. Der vorgestellte Prototyp interpretiert Ortsangaben in natürlicher Sprache und soll in der nächsten Version auf eine kartographische Repräsentation der Ergebnisse setzen. In [Deng06] wird die Benutzeroberfläche eines Systems vorgestellt, das Dokumente in mehreren gleichzeitig angezeigten Verzeichnisansichten kategorisiert. Diese Verzeichnisansichten werden dynamisch erstellt und basieren auf dem gemeinsamen Vorkommen von Begriffen in Ordnern und Dateien. Die Autoren von [GiSh10] illustrieren eine Dokumentensuche mit Facetten und Vorschlägen für den Einsatz in Unternehmen und Organisationen. Bei der Verwendung von Facetten werden Informationen in mehrere voneinander unabhängige Mengen von Kategorien eingeteilt.

³ siehe <http://www.iso.org/>

- *Sicherheit*

In [SaPr06] wird eine Methode zur Erkennung von Benutzern aufgezeigt, die in einem Dokumentenmanagementsystem unter der Verwendung eines Benutzerkontos einer anderen Person die Inhalte von Dokumenten böswillig manipulieren wollen. Bei der Überwachung der Benutzer bei ihren Tätigkeiten wird ein Verhaltensprofil erstellt. Zeigt der Benutzer ein für sein Profil atypisches Verhalten, kann ein möglicher Angriff auf das System vorliegen.

2.2 Web 2.0

Unter Web 2.0 wird laut [Muru07] die zweite Phase in der Evolution des World Wide Webs verstanden. Dieser Evolutionsschritt beschränkt sich nicht nur auf technologische Weiterentwicklungen. Veränderungen in der Verwendungsart des Webs sind ebenfalls ein wesentlicher Bestandteil.

Soziale Interaktion und Kollaboration bilden Schwerpunkte von Webseiten, die unter dem Begriff „Web 2.0“ einzuordnen sind. Das Fehlen einer allgemein anerkannten, eindeutigen und präzisen Begriffsdefinition wird in [Muru07] durch die Vielschichtigkeit der mit Web 2.0 in Verbindung gebrachten Themen begründet.

Abbildung 1 aus [Ange05] zeigt diese Vielfalt und beinhaltet Schlagwörter für konzeptionelle Schwerpunkte inklusive der Namen repräsentativer Websites.

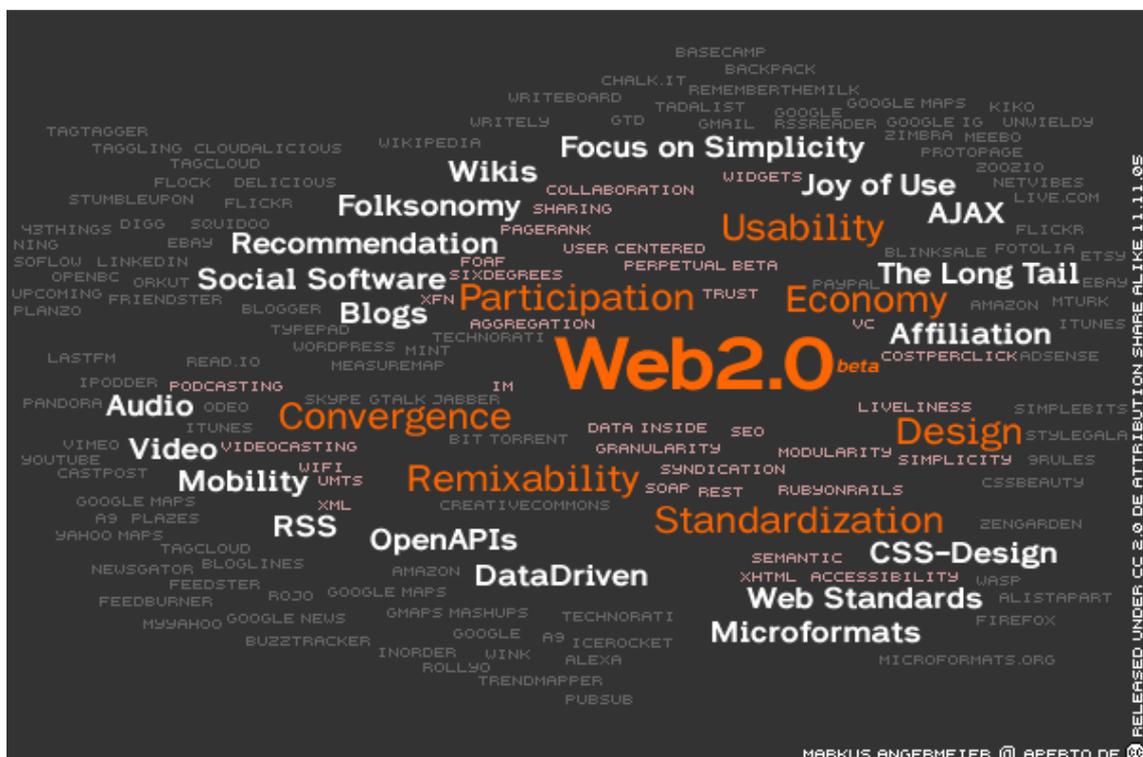


Abbildung 1: Schlagwortwolke „Web 2.0“ aus [Ange05]

Als pragmatische Definition wird in [MiRo06] unter Web 2.0 die Menge der Applikationen, Websites und Unternehmen verstanden, die gemeinsam dem Begriff die Bedeutung verleihen. Um den Bezug zur Praxis herzustellen, werden nennenswerte Arten von Websites aufgelistet, die laut [CaDa09], [Muru07], [PiFo08] und [VaPa09] mit Web 2.0 in Verbindung gebracht werden:

- *Wikis*
Unter Wikis versteht man Webseiten, deren Inhalt sich von den Benutzern über in der Seite integrierte Bedienelemente ändern lässt. Auf diesem Weg können Anwender gemeinsam an der Erstellung gewünschter Inhalte in Form von Artikeln arbeiten. Wikis stellen eine einfache Methode dar, Wissen und Erfahrungen festzuhalten und auszutauschen.
- *Blogs*
Das Wort „Blog“ ist eine Kurzform für den Begriff „Weblog“. Die Betreiber eines Blogs schreiben nach eigenem Ermessen Beiträge in Tagebuchform und stellen diese der Öffentlichkeit oder einer bestimmten Gruppe von Personen zur Verfügung. Verfasste Beiträge werden chronologisch, beginnend beim aktuellsten Eintrag, aufgelistet. Sehr häufig können Besucher eines Blogs ihre Meinungen zu einem Beitrag als Kommentar abgeben. In diesem Fall wird der Blog zu einem Werkzeug für den Austausch von Meinungen und Gedanken.
- *Soziale Netzwerke*
Soziale Netzwerke sind Websites, die soziale Verbindungen unterschiedlicher Arten zwischen Personen abbilden. Je nach Fokus des Netzwerks stehen geschäftliche, private oder generelle Kontakte im Mittelpunkt. Benutzer können nach der Registrierung in einem sozialen Netz ihr persönliches Profil bearbeiten und Kontakte zu anderen Benutzern knüpfen. Soziale Netzwerke unterstützen beim Austausch von Informationen und Finden neuer Bekanntschaften.
- *Web-Applikationen*
Die Funktionen einer Web-Applikation sind dem Benutzer über einen Browser zugänglich. Der Browser übernimmt die Interaktion mit dem Benutzer und die Darstellung der Benutzerschnittstelle, während ein Webserver für die Abarbeitung der Programmlogik zuständig ist. Mit Hilfe des Browsers werden Anfragen an den Webserver gestellt. Der Webserver bearbeitet diese und sendet dem Browser eine anzuzeigende Webseite als Antwort. Aktuelle Web-Applikationen unterscheiden sich in der Verwendung und Funktionalität heutzutage kaum von lokal auf einem Computer installierten Programmen.

- *Mashups*
Kombiniert und greift eine Website auf Dienste und Funktionen verschiedener Websites zurück, spricht man von einem Mashup. Ein Mashup aggregiert bestehende Inhalte in Form einer eigenständigen Internetpräsenz und kann diese um zusätzliche Funktionalitäten erweitern. Aufbauend auf den Funktionen der ausgewählten Bestandteile erschließen sich auf der Anwendungsebene neue Einsatzgebiete.

Häufig kommen bei der Realisierung der aufgezählten Arten von Websites folgende Technologien zum Einsatz:

- *XML*
Unter der Abkürzung XML verbirgt sich die Bezeichnung „Extensible Markup Language“. Mit XML werden hierarchisch strukturierte Daten in für Menschen leserlicher Textform abgebildet. Der grundlegende Baustein in XML zur Strukturierung der Daten wird mit dem englischen Wort „Tag“ bezeichnet. Der Anfang und das Ende eines Tags können weitere, hierarchisch untergeordnete Tags beziehungsweise Textdaten umschließen. Des Weiteren ist es möglich, Werte für Attribute eines Tags zu definieren. Die Besonderheit von XML ist, dass die Namen der Tags und Attribute frei vergeben werden können. Folgenderweise können wie in [BoBr99] angeführt, auf Basis der Syntax von XML neue Sprachen oder Formate für den Datenaustausch zwischen verschiedenen Computersystemen entworfen werden.
- *AJAX*
Die Abkürzung „AJAX“ steht für „Asynchronous JavaScript and XML“. AJAX ist eine elementare Methode zur Realisierung von Websites und Web-Applikationen im Web 2.0. Beim Surfen im Web stellt der Benutzer über seinen Browser durch die Eingabe einer WWW-Adresse oder Klicken eines Links auf einer Webseite eine Anfrage an einen Webserver. Der Web-Server bearbeitet diese Anfrage und antwortet mit der aufgerufenen Webseite. Beim Einsatz von AJAX enthalten Webseiten JavaScript-Quelltext, der Anfragen an einen Webserver absetzt und dessen Antworten auch verarbeiten kann. Dieser Datenaustausch innerhalb des Skripts wird als asynchron bezeichnet, weil er außerhalb der beschriebenen, herkömmlichen Zyklen des Ladens einer Webseite stattfindet. Mit Hilfe asynchroner Anfragen können Inhalte einer Webseite geändert werden, ohne diese neu laden zu müssen. Über diese Vorgehensweise lassen sich im Web Benutzerschnittstellen realisieren, die in der Verwendung mit lokal auf einem Computer installierten Anwendungen vergleichbar sind.

- *RSS*
„Really Simple Syndication“ kurz RSS, ist ein auf der Syntax von XML basierendes Format zur Abbildung einer Menge von Nachrichten. RSS-Dateien werden verwendet, um die Inhalte einer Webseite in Form von mehreren Einträgen zusammenzufassen. Stellt eine Webseite eine RSS-Datei zur Verfügung, spricht man von einem RSS Feed. Diese Feeds können mit aktuellen Browsern oder mit speziell dafür vorgesehenen Applikationen abonniert werden. Werden neue Nachrichten in die abonnierten Feeds eingespeist, sind diese in den Applikationen einsehbar, ohne dass der Benutzer die jeweiligen Webseiten besuchen muss.
- *Web-Services*
Nach der Definition in [BoHa09] ist ein Web Service ein Softwaresystem zur interoperablen Interaktion von Computersystemen über ein Netzwerk. Die Beschreibung der Schnittstellen für die Interaktion liegt in einem maschinell verarbeitbaren Format vor. Web Services bauen auf grundsätzlichen, generell verfügbaren Web-Technologien auf und verwenden diese zum Nachrichtentransport. Anders betrachtet könnte man Web Services als eine technische Manifestation der Grundgedanken von Web 2.0 in Richtung sozialer Interaktion und Kollaboration auffassen. Services werden der Öffentlichkeit oder einer ausgewählten Entwicklergruppe auf einer technischen Ebene zur Verfügung gestellt, um weitere Services auf einer höheren Abstraktionsebene zu erstellen oder eigene Anwendungen aus der Komposition von Services zu erschaffen.

Mit dem Web 2.0 assoziierte Websites bedienen sich zur Beschreibung von Inhalten oftmals einer „Folksonomy“. Der englische Begriff der „Folksonomy“ bezeichnet eine Systematik zur kollaborativen Kategorisierung von Inhalten per Zuweisung von Schlagworten durch die Benutzergemeinschaft. Jedem Anwender steht es dabei frei, die Liste von Schlagworten einer zu beschreibenden Ressource zu editieren. Benutzer, Schlagwörter und zu beschreibende Ressourcen werden in [SpMo07] als die Grundelemente eines Systems zur Beschlagwortung verstanden. Die Vergabe von Schlagworten wird mit dem englischen Fachbegriff „Tagging“ bezeichnet. Wenn diese Tätigkeit gemeinsam von einer Gruppe durchgeführt wird, spricht man von „Social Tagging“. Die durch das Social Tagging entstehende Ordnung ist in [ScRa06] durch die kollektive Intelligenz einer größeren Gruppe begründet. Laut [MoSa06] stellt das Ausnützen der kollektiven Intelligenz eine Grundlage für den Erfolg von Web 2.0 dar.

Tabelle 2 enthält zum Abschluss dieses Abschnitts eine Aufzählung von prägnanten Websites aus dem Web 2.0. Es wurde eine grobe Kategorisierung der Webpräsenzen aufgrund ihrer Art vorgenommen.

Art	Beispiele
Wiki	http://www.wikipedia.org/
Multimedia (Video, Fotos)	http://www.youtube.com/ http://www.flickr.com/
Soziale Netzwerke	http://www.facebook.com/
Blog, Microblog	http://www.blogger.com/ http://twitter.com/
Bookmarking	http://digg.com/ http://delicious.com/
Web-Applikationen	http://docs.google.com/ http://maps.yahoo.com/ http://mail.google.com/

Tabelle 2: Pragnante Webseiten als Beispiele fur Web 2.0

Anhand der soeben angefuhrten Beispiele soll die praktische Bedeutung und die Idee von Web 2.0 vermittelt werden. Eine Bevorzugung der in der Aufzahlung angefuhrten Webauftritte gegenuber nicht genannten Alternativen ist nicht beabsichtigt.

2.3 Semantisches Web

Die Berucksichtigung der Bedeutung von Informationen bei der Verarbeitung durch Computer ist die Essenz im Konzept des semantischen Webs. Durch die Miteinbeziehung der Bedeutung wird der Maschine ermoglicht, Zusammenhange von Informationen zu erkennen und in einer Art weiterzuverarbeiten, die dem Benutzer vorbehalten ware.

In [BeHe01] wird das semantische Web als eine Erweiterung des Webs bezeichnet, in der jede Information eine fur den Computer klar definierte Bedeutung hat. Dies soll den Anwendern ermoglichen, mit Computern auf einer hoheren Abstraktionsebene zusammenzuarbeiten. Anstatt ein Werkzeug zu sein, das den Anweisungen seines Benutzers folgt, ohne deren ubergeordnetes Ziel zu kennen, konnten Computer in der Verwendung die Rolle eines Assistenten bei der Aufgabebewaltigung einnehmen. Inhaltliche Schwerpunkte, die beschreiben, wie ein gewunschtes Ergebnis auszusehen hat, rucken in den Vordergrund, wahrend die Spezifikation des Weges zur Zielerreichung an Bedeutung verliert. Als mogliches Einsatzszenario werden Software-Agenten genannt. In diesem Kontext sind Software-Agenten Programme, die mit Hilfe des semantischen Webs Losungen fur benutzerspezifizierte Ziele durch das Sammeln, Filtern und Verarbeiten von Informationen bereitstellen konnen. Zur Realisierung des semantischen Webs ist in [BeHe01] von folgenden Prinzipien die Rede:

- *Strukturierte Information*

Zur Ermoglichung einer maschinellen Verarbeitung von Informationen mussen diese in einem definierten Format strukturiert werden. Um die Bedeutung einer Information erschließen zu konnen, muss man sie erst interpretieren. Eine kor-

rekte Interpretation erfordert jedoch auch die Kenntnis des Kontexts, in dem die Information eingebettet ist. Anders formuliert sind kontextlose Informationen ohne Bedeutung. Eine wohldefinierte Strukturierung von Information in einem spezifizierten Format alleine führt folglich nicht zum Verständnis der Bedeutung, dafür aber zu einer Basis für den Datenaustausch zwischen Maschinen.

- *Wissen*

Durch die Vernetzung von Informationen lässt sich Wissen abbilden. Wissen in dieser Auffassung ist eine Menge von Aussagen, die Beziehungen zwischen einzelnen Entitäten in einem Kontext beschreiben. Über diese Aussagen kann die Bedeutung der in einer Informationsstruktur abgebildeten Konzepte beschrieben werden. Kurz gesagt, verleihen die getroffenen Aussagen der strukturierten Information ihre Bedeutung. Die Grundelemente einer Aussage sind Subjekt, Prädikat und Objekt. Das Prädikat gibt an, in welcher Beziehung Subjekt und Objekt miteinander stehen. Ein in der Informationsstruktur vorkommendes Konzept kann in einer Aussage als Subjekt oder Objekt verwendet werden. In einer Menge von mehreren Aussagen, werden die Konzepte durch die über Prädikate hergestellten Verknüpfungen Teil eines Informationsnetzes. Das Problem des korrekten Verständnisses der semantischen Zusammenhänge von in unterschiedlichen Kontexten entstandenen Wissensrepräsentationen bleibt jedoch ungelöst.

- *Ontologie*

Aus unterschiedlichen Kontexten stammende Wissensrepräsentationen in einen gemeinsamen Kontext einzubetten, ist in der Informatik ein Verwendungszweck einer Ontologie. Zum anderen können in einer Ontologie auch Regeln enthalten sein, um logische Schlussfolgerungen automatisiert aus dem vorhandenen Wissensstand ableiten zu können. Ontologien beschreiben die Relationen zwischen verwendeten Begriffen und bilden eine Brücke für den Transfer von Wissen zwischen unterschiedlichen Domänen. Durch die Verwendung von Klassen werden Objekte nach ihrer Art unterschieden. Gemeinsamkeiten zwischen Klassen sind durch die Weitergabe von Eigenschaften per Vererbung ausdrückbar.

Die soeben aufgezählten Prinzipien nehmen in folgenden, dem semantischen Web zugeschriebenen Technologien Gestalt an. Die bereits behandelte Sprache „XML“ wurde für die Abbildung und den Austausch von strukturierten Informationen konzipiert und eignet sich durch seine weitläufige Verbreitung im Web als naheliegende syntaktische Grundlage. Die Wissensrepräsentation findet per „RDF“ statt und Ontologien werden mit „OWL“ formuliert. Die kommende Aufzählung fasst die beiden genannten Technologien zusammen, welche im Detail in [BrCa07] und [GaKr11] nachgelesen werden können:

- *RDF*

Das mit dem englischen Namen bezeichnete „Resource Description Framework“⁴, abgekürzt „RDF“, kann zur Beschreibung von Entitäten beliebiger Art eingesetzt werden. Die per RDF beschriebenen Entitäten werden als Ressourcen bezeichnet. Webseiten, Dokumente, Orte, Personen und so weiter können Ressourcen sein, um einige Beispiele zu nennen. Ressourcen werden durch einen eindeutigen Bezeichner identifiziert. Für gewöhnlich wird die Webadresse als Bezeichner verwendet, wenn eine im Web vorliegende Ressource Gegenstand der Beschreibung ist. Ressourcen werden durch die Auflistung ihrer Eigenschaften beschrieben. Neben der Anführung des Attributs muss die Art der Eigenschaft angegeben werden, um auszudrücken, wie die Ressource mit ihrem Attribut in Verbindung steht. Literale, also Zeichenketten, und Ressourcen können den Platz eines Attributs einnehmen. Zusammengefasst ist der Inhalt einer RDF-Datei eine Sammlung von Aussagen, in denen die Ressource das Subjekt, die Attributart das Prädikat und das Attribut das Objekt darstellt.

- *OWL*

Die Abkürzung „OWL“ steht für den Namen „Web Ontology Language“⁵. Das Einsatzgebiet der Sprache „OWL“ ist die Spezifikation von Ontologien. Die wesentlichen Grundelemente der Sprache sind Klassen, Eigenschaften und Instanzen. Eigenschaften beschreiben, wie zwei Instanzen, die Angehörige von Klassen sind, in Beziehung stehen. Die Struktur einer Aussage in OWL entspricht der bereits in RDF verwendeten Kombination von Subjekt, Prädikat und Objekt. In OWL getroffene Aussagen werden auch als Axiome bezeichnet und gelten immer als wahr. Über die Menge von Axiomen können logische Schlussfolgerungen getroffen werden. Die angewandte Logik folgt dem Grundsatz, der mit dem englischen Fachbegriff bezeichneten „Open World Assumption“. Dies bedeutet, dass beim Ziehen logischer Schlüsse Aussagen, die nicht als wahr abgeleitet werden können, nicht zwingend falsch sein müssen. Dieser Grundsatz geht davon aus, dass nie das gesamte Wissen bekannt ist. Das primäre Format für den Austausch per OWL abgebildeter Ontologien im semantischen Web ist RDF. Für OWL-Dokumente stehen auch andere syntaktische Ausdrucksformen zur Verfügung. Eine Entscheidungsgrundlage für die Wahl eines bestimmten Formats wird durch dessen Eignung für das jeweilige Anwendungsgebiet geliefert.

Die durch den Tripel aus Subjekt, Prädikat und Objekt gebildete Kombination ist im semantischen Web eine grundlegende und universelle Form zum Treffen von Aussagen. Bei realen Anwendungen führt die Abbildung komplexer Sachverhalte schnell zu einer

⁴ siehe <http://www.w3.org/RDF/> und <http://www.w3.org/standards/techs/rdf>

⁵ siehe <http://www.w3.org/standards/techs/owl>

großen Menge von Aussagen beziehungsweise Tripel. Um in diesen Fällen eine ausreichende Performance gewährleisten zu können, ist der Einsatz von Datenbanksystemen unumgänglich.

Bei großen Mengen von Tripeln, kommt eine sogenannter „Triplestore“ zum Einsatz. Ein Triplestore ist ein Datenbanksystem, das auf die Verarbeitung von in Form von Tripeln vorliegenden Informationen spezialisiert ist. Die aus Tripeln aufgebauten Datenstrukturen entsprechen einem Graph. In diesem Graph bilden die als Subjekte und Objekte verwendeten Ressourcen die Knoten, während die Prädikate den Kanten entsprechen. Die effiziente Abfrage von Graphen in relationalen Datenbanken gestaltet sich als problematisch und schafft so den Bedarf an spezialisierten Lösungen. Die Implementierung eines Triplestores erfolgt entweder durch den Aufbau auf ein bestehendes relationales Datenbankmanagementsystem oder durch eine komplette Neuentwicklung. Umfangreiche Graphen mit hoher Geschwindigkeit verarbeiten zu können, ist eine technologische Basis für die Realisierung des semantischen Webs.

2.4 Software- und Web-Engineering

Nach [Somm10] umfasst die Ingenieursdisziplin „Software Engineering“ alle Aspekte der Herstellung von Software. Diese Definition schließt alle Aktivitäten begonnen bei der Spezifikation bis hin zur Wartung der Software nach der Inbetriebnahme ein. Aufgrund der allgemeinen Natur dieser Begriffsdefinition ist eine genauere Betrachtung für eine vollständige Erklärung der Bedeutung notwendig.

Laut [ZuGr04] existiert für „Software Engineering“ eine Reihe von Definitionen. Auch wenn die im zuletzt referenzierten Werk als Beispiel angeführten Definitionen teilweise aus anderen Zugangsweisen zum Thema entstanden sind, enthalten sie in impliziter oder auch expliziter Weise diese wesentlichen Gemeinsamkeiten mit der zuerst genannten Definition:

- *Ingenieursdisziplin*
Bei der Ausübung als Ingenieursdisziplin ist ein systematischer Ansatz an den Tag zu legen. Eine systematische Herangehensweise ist zum einen wiederholbar und ermöglicht zum anderen eine Evaluierung des durchgeführten Prozesses.
- *Umfassender Tätigkeitsbereich*
Software Engineering beschränkt sich nicht auf die Entwicklung von Software, denn die im gesamten Lebenszyklus von Software anfallenden Tätigkeiten sind Teil des Fachgebiets.

Neben der methodischen und technischen Sichtweise spielen auch wissenschaftliche und organisatorische Aspekte im Software Engineering eine Rolle. Nach [Somm10]

sind das Management von Softwareprojekten und die wissenschaftliche Auseinandersetzung mit dem Fachgebiet weitere untrennbare Bestandteile im Begriffsverständnis.

Für eine zeitgemäße Auffassung des Software Engineerings muss jedoch noch eine zusätzliche Komponente berücksichtigt werden. In [Boeh06] wird ein Blick auf die Vergangenheit und Zukunft des Software Engineerings geworfen. Der Quelle nach lag im 20. Jahrhundert der Fokus auf der Verbesserung der Prozesse und Werkzeuge. Gegen Ende des Jahrtausends gewannen die Benutzer von Software zunehmend an Bedeutung. Neben der erhöhten Bedeutung der Benutzbarkeit von Software wanderte die Orientierung an den Bedürfnissen und Wünschen des Abnehmers in den Vordergrund. Zusammengefasst liefert das Verständnis von Software Engineering als Dienstleistung und Verpflichtung gegenüber dem Abnehmer den letzten Baustein für eine moderne Definition.

Laut [HiCo01] ist es bei der Entwicklung von Software notwendig geworden, flexibel auf sich ändernde Rahmenbedingungen und Anforderungen reagieren zu können. Durch diese Flexibilität sollen mit Planabweichungen einhergehende Aufwände so gering wie möglich gehalten werden. Genauer betrachtet ist dies seitens des Softwareherstellers ebenfalls ein Ausdruck der Bestrebung mit größerer Sorgfalt auf die Bedürfnisse des Auftraggebers einzugehen. Methoden zur Umsetzung der geforderten Flexibilität werden unter dem Begriff der „agilen Softwareentwicklung“ zusammengefasst.

In [HiCo01] wird der Leitsatz⁶ der agilen Softwareentwicklung formuliert. Dieser Leitsatz beinhaltet die Aussage, dass der linken Spalte gegenüber der rechten Spalte derselben Zeile aus Tabelle 3 mehr Bedeutung zugemessen werden soll.

Bedeutung	
mehr	<i>gegenüber</i>
	weniger
Individuen und Interaktionen	Prozesse und Werkzeuge
Funktionierende Software	Ausführliche Dokumentation
Zusammenarbeit mit Kunde	Vertragsverhandlungen
Reagieren auf Änderungen	Befolgung des Plans

Tabelle 3: Wertschätzungen der agilen Softwareentwicklung

Die in [Glog10] beschriebene Projektmanagementmethode „Scrum“ ist ein repräsentatives Beispiel für einen agilen Ansatz aus der Praxis. Diese Methode setzt zur iterativen Entwicklung eines Softwaresystems auf ein selbstorganisiertes Team mit überschaubarer Größe von etwa sieben Personen. Bei der Anwendung von „Scrum“ werden folgende Rollen vertreten, die gemeinsam ein sogenanntes „Scrum Team“ bilden:

⁶ Manifesto for Agile Software Development, siehe <http://www.agilemanifesto.org/>

- *„Product Owner“*
Der „Product Owner“ vertritt die Interessen des Abnehmers und stellt den geschäftlichen Nutzen des Produkts sicher. Vom Inhaber dieser Rolle werden gewünschte Funktionen aus der Sicht des Anwenders beschrieben und priorisiert. Die daraus resultierende Auflistung von umzusetzenden Funktionen wird als „Product Backlog“ bezeichnet.
- *„Scrum Master“*
Der „Scrum Master“ sorgt für die Einhaltung des Entwicklungsprozesses und ermöglicht den Mitgliedern des Teams sich frei von Ablenkungen ihren Tätigkeiten widmen zu können.
- *„Team“*
Das Team ist für die technische Umsetzung der im „Product Backlog“ angeführten Funktionen und die Schätzung des Entwicklungsaufwands zuständig.

Vor einer Entwicklungsiteration, die in „Scrum“ als „Sprint“ bezeichnet wird, legt das Team fest, welche Funktionen als Nächstes umgesetzt werden. Während eines Sprints werden keine Änderungen an den zur Umsetzung ausgewählten Produktfunktionen zugelassen. Am Ende einer Iteration wird das Ergebnis des Sprints überprüft und der „Product Backlog“ an geänderte oder neue Anforderungen angepasst. Die im Sprint gesammelten Erfahrungen werden berücksichtigt, um den Prozess zu verbessern und im Idealfall eine Steigerung der Produktivität zu erzielen.

Als zweites Beispiel für einen agilen Entwicklungsprozess wird „Extreme Programming“⁷ abgekürzt „XP“ genannt. Laut [Hans10] ist „Extreme Programming“ ein iterativer Entwicklungsprozess, der durch die strikte Einhaltung einer Kombination von vordefinierten Regeln und Praktiken charakterisiert wird. Die soeben zitierte Quelle gibt an, dass ein Teil dieser Regeln und Praktiken keine Neuerung darstellt und längst zu den Methoden der modernen Softwareentwicklung zählt. Als Beispiele für diese allgemein akzeptierten Vorgehensweisen werden genannt:

- Iterativ inkrementelle Entwicklung
- Planung der Iterationen und Zwischenergebnisse
- Einfaches Design
- Automatisierte Tests ab Projektbeginn

Das Verständnis von „Extreme Programming“ ist in den laut [Hans10] restlichen Regeln und Praktiken zu finden, deren Wirksamkeit umstritten ist. In der folgenden Aufzählung werden diese zusammengefasst:

⁷ siehe <http://www.extremeprogramming.org/> und <http://xprogramming.com/>

- *Erstellen von User Stories durch den Kunden*
Der Kunde soll Anforderungen als „User Stories“ formulieren. Eine „User Story“ beschreibt in der Sprache des Kunden mit wenigen Sätzen eine Anforderung.
- *Paarweises Programmieren*
Beim paarweisen Programmieren arbeiten zwei Personen zusammen an einem Computer. Während ein Programmierer den Quelltext verfasst, tauscht er sich mit seinem Partner über die Details der Entwicklung aus.
- *Quelltext als gemeinsames Eigentum*
Unter dem gemeinsamen Eigentum von Quelltext wird verstanden, dass jeder Entwickler berechtigt und befähigt ist, den Quelltext anderer Entwickler zu ändern.
- *Kontinuierliche Integration von Quelltext*
Von Paaren erstellter Quelltext wird kontinuierlich in den Quelltext des lauffähigen Systems integriert, damit das Ergebnis den aktuellen Stand der Entwicklung repräsentiert.
- *Einbeziehung des Kunden im Projektteam*
Der Kunde wird als Mitglied des Projektteams betrachtet und soll wenn möglich permanent als Ansprechpartner für eine persönliche Kommunikation zur Verfügung stehen.

Der verbleibende Teil dieses Abschnitts befasst sich mit dem Thema „Web-Engineering“. Man spricht von Web-Engineering, wenn im Software Engineering Web-Applikationen im Mittelpunkt der Betrachtung stehen. Laut [KaPr06] kann Web-Engineering als ein eigenständiger Zweig des Software Engineerings angesehen werden, denn die speziellen Eigenschaften von Web-Applikationen erfordern die Anpassung und Konzeption von bekannten und neuen Vorgehensweisen in der Entwicklung.

In [MeMo10] werden drei Bereiche identifiziert, in denen sich die Entwicklung von Web-Applikationen maßgeblich von der Entwicklung konventioneller Software unterscheidet:

- *Entwickler*
Von den bei der Entwicklung beteiligten Personen werden breiter gefächerte Kompetenzen gefordert. Für den Inhalt und das grafische Design müssen häufig Personen nahtlos in den Entwicklungsprozess eingebunden werden, die tendenziell über keine Kenntnisse der technischen Umsetzung verfügen.

- *Plattform*
Web-Applikationen laufen seitens des Clients innerhalb eines Browsers und benötigen Netzwerkzugriff auf den Server. Die Navigation erfolgt über Hyperlinks, die einzelne Webseiten miteinander verknüpfen oder auf andere Dokumente verweisen.
- *Zielgruppe*
Die Orientierung an den Bedürfnissen einer großen Menge an potenziellen Benutzern ist schwieriger als an einer überschaubaren Benutzergruppe. Einer ansprechenden visuellen Präsentation wird eine höhere Bedeutung zugemessen.

3 Methodik

In diesem Kapitel wird der zur Anwendung kommende Entwicklungsprozess definiert.

Die Entwicklung der Software erfolgt im Rahmen agiler Entwicklungsmethoden. Begründet durch die Tatsachen, dass für die Entwicklung nur die eigene Arbeitsleistung zur Verfügung steht und agile Methoden oftmals ein Projektteam voraussetzen, kommt ein eigener auf die Ausgangssituation angepasster Entwicklungsprozess zum Einsatz.

Mit dem Projekt „TISS“ als Vorbild stützt sich dieser Prozess auf Bestandteile der beiden laut [WiMe10] am weitesten verbreiteten Methoden der agilen Softwareentwicklung:

- „Scrum“
- „Extreme Programming“ (abgekürzt als „XP“)

Im Kontext dieses Projekts sinnvolle Praktiken der beiden aufgezählten Elemente werden übernommen und falls notwendig angepasst. Die Durchführung einer Anpassung der Methoden auf das Projekt wird durch die in [HrRu09] angeführte Annahme begründet, dass keine universelle Methode für alle Einsatzszenarien existiert. Unter dieser Annahme wurde die unter dem Namen „Crystal“ zusammengefasste Familie agiler Entwicklungsmethoden konzipiert.

Beim „Crystal“-Ansatz aus [Cock04] wird ein Projekt in drei Skalen eingeteilt. Die Einteilung erfolgt nach der Anzahl der beteiligten Personen, dem maximalen Schweregrad von auftretenden Fehlern und der Wichtigkeit der Zielsetzung. Jede Kombination von drei konkreten Werten verweist auf eine Methode, die als Basis für den Entwicklungsprozess dient. Die für dieses Projekt nach „Crystal“ vorgeschlagene Vorgehensweise „Crystal Clear“ ist laut [HrRu09] sehr ähnlich zu „Extreme Programming“ und wird deshalb nicht gesondert behandelt.

3.1 *Entwicklungsprozess*

Bevor mit der Prozessdefinition begonnen wird, werden Rahmenbedingungen festgelegt, denen der Prozess gerecht werden muss:

- Die von verschiedenen agilen Entwicklungsmethoden gemeinsam gestellte Forderung schnell Ergebnisse vorweisen zu können, ist eine Bedingung die der zur Anwendung kommende Prozess zu erfüllen hat.

- Der Prozess soll die Interaktion mit den Anwendern und deren Feedback in den Vordergrund stellen. Das zu einem Iterationsergebnis gesammelte Feedback soll laufend in die Entwicklung einfließen.

Nach der soeben durchgeführten Festlegung der Rahmenbedingungen wird nun mit der Prozessdefinition fortgefahren. Zu Beginn wird der Ausgangszustand analysiert und dokumentiert um sich in den Fachbereich einzuarbeiten. Wie in [HuJa11] zu lesen, werden Interviews mit den Anspruchspersonen zur Erfassung der Anforderungen durchgeführt. Für eine gemeinsame Gesprächsbasis werden während den Interviews die von den Anspruchspersonen verwendeten Begriffe aus dem Fachbereich erfragt. Die auf diesem Weg gesammelten Begriffe sind im Glossar am Ende dieser Arbeit zu finden. Nach dem Vorbild der in [HuJa11] angeführten Quellen für Anforderungen werden des Weiteren diese Punkte berücksichtigt:

- *Bestehende Dokumentation*
Relevante Dokumentation aus dem Fachbereich wird bei der Anforderungsanalyse beachtet. Das Studium der bestehenden Dokumentation hilft außerdem beim Erfassen der Ausgangslage.
- *Erfahrungen der Anwender*
Den alltäglichen Problemen der Anwender im Fachbereich wird Gehör geschenkt. Die Anwender werden gezielt nach Problemen befragt, welche sie bei ihrer Arbeit zu bewältigen haben.
- *Prototyp*
Die Zwischenergebnisse der Entwicklung werden laufend den Anwendern präsentiert. Die Reaktionen auf den Prototyp fließen gegebenenfalls als weitere Anforderungen in den Prozess ein.

In Anlehnung an das in [RuCh08] und [Barn10] beschriebene „Usability Testing“, bei dem Personen bei der Verwendung eines Systems zur Beurteilung dessen Benutzbarkeit beobachtet werden, werden die Mitarbeiter der Kanzlei bei der Arbeit mit dem Kanzleinformationssystem beobachtet. Für eine nachträgliche Analyse werden mit der freien Software „CamStudio“⁸ zur Bildschirmaufnahme Videos dieser Sitzungen angefertigt. Diese Maßnahme dient zur Aufdeckung von nicht erkannten Anforderungen beziehungsweise Problemen. Als weiterer Nutzen wird das Verständnis des Anwendungsbereichs vertieft. Basierend auf den aus genannten Quellen gesammelten Anforderungen wird ein Systemkonzept vor dem Start der Implementierung erstellt.

⁸ Siehe <http://camstudio.org/>

Bei der Implementierung wird in kurzen Entwicklungsiterationen gearbeitet, die in einem Resultat enden, das für die Anwender sichtbar ist. Unter der Sichtbarkeit für die Anwender wird in diesem Kontext verstanden, dass die Benutzeroberfläche neue Funktionen bietet oder eine Änderung im Systemverhalten erkennbar ist.

Vergleichbar mit der Vorgehensweise laut [Schw04] bei „Scrum“ wird zu Beginn jeder Iteration festgelegt, welche Elemente aus der Liste der Anforderungen abgearbeitet werden sollen. Nach dem Ende einer Iteration werden die Ergebnisse mit den Anwendern besprochen, Verbesserungsvorschläge entgegen genommen und die Liste der Anforderungen aktualisiert. Als Ergebnis der ersten Iteration ist ein Prototyp der Benutzeroberfläche ohne Funktion umzusetzen.

Die in „Scrum“ empfohlene Dauer einer Iteration von 30 Tagen wird als unverbindlicher Richtwert übernommen. Kürzere Laufzeiten sind erlaubt, solange ein für den Benutzer sichtbares Ergebnis erzielt wird. Längere Iterationsphasen werden nur im Ausnahmefall akzeptiert, falls die Erfüllung einer einzelnen Anforderung mehr Zeit als ein Monat in Anspruch nimmt und bei der Implementierung keine vorzeigbaren Zwischenergebnisse erzielbar sind.

Folgende in [HrRu09] angeführte Praktiken aus dem „Extreme Programming“ kommen bei der Implementierung zum Einsatz:

- *Einfaches Design*
Design zum Selbstzweck wird vermieden, indem es so einfach wie möglich gehalten wird. Der Versuch alle Eventualitäten im Design vorherzusehen ist nicht zielführend.
- *Refactoring*
Der Quelltext wird laufend verbessert und wenn möglich auch vereinfacht. Ein entsprechender Anteil der Entwicklungszeit muss für die Verbesserungen berücksichtigt werden.
- *Programmierstandards*
Bei der Programmierung werden einheitliche Standards und Konventionen befolgt. Die Verwendung bestehender Standards und Konventionen als Vorlage erleichtert eigene projektinterne Definitionen.

3.2 Dokumentation

Die Dokumentation der Software-Architektur erfolgt nach dem 4+1 Sichtenmodell aus [Kruc95]. Zur Erstellung einer umfassenden Beschreibung wird die Architektur mit

Diagrammen und Texten aus verschiedenen Standpunkten dokumentiert. Weitere Details zum 4+1 Sichtenmodell sind im Abschnitt 5.2 nachlesbar.

Als Diagrammnotation wird die „Unified Modeling Language“⁹ abgekürzt „UML“ verwendet. Diese Arbeit bedient sich zur Dokumentation der in der folgenden Aufzählung genannten Arten von UML-Diagrammen, welche im Detail in [ArNe05] und [MiHa06] nachlesbar sind.

- *Anwendungsfalldiagramm*
Diese Art von Diagramm zeigt, welche Funktionen ein System für welchen Interaktionspartner anbietet. Interaktionspartner werden Akteure genannt und können Personen oder auch andere Systeme sein. Das Diagramm setzt Anwendungsfälle, welche einen Teil der Funktionalität des Systems repräsentieren, untereinander und mit den Akteuren in Bezug.
- *Klassendiagramm*
Klassen werden im Klassendiagramm mit ihren Operationen, Attributen und Abhängigkeiten zueinander als Bausteine eines Softwaresystems gezeigt.
- *Komponentendiagramm*
Ein Komponentendiagramm stellt dar, welche Komponenten ein System beinhaltet und wie diese voneinander abhängen. Zur Darstellung der Abhängigkeit werden die von einer Komponente angebotenen Schnittstellen eingezeichnet.
- *Sequenzdiagramm*
Mit einem Sequenzdiagramm werden Abläufe und stattfindende Interaktionen zwischen unterschiedlichen Objekten in Form von Nachrichten dargestellt.
- *Verteilungsdiagramm*
Die Hard- und Software, welche für den Betrieb eines Systems notwendig ist, wird in einem Verteilungsdiagramm festgehalten.

Die Liste der Anforderungen und zur Umsetzung notwendigen Tätigkeiten wird mit der Projektmanagementapplikation „Redmine“¹⁰ verwaltet. Abbildung 2 zeigt als Beispiel eine Projektübersicht als Gantt-Diagramm. Für die Bereitstellung der Projektdokumentation wird das applikationsinterne Wiki eingesetzt.

⁹ siehe <http://www.uml.org/>

¹⁰ siehe <http://www.redmine.org/>

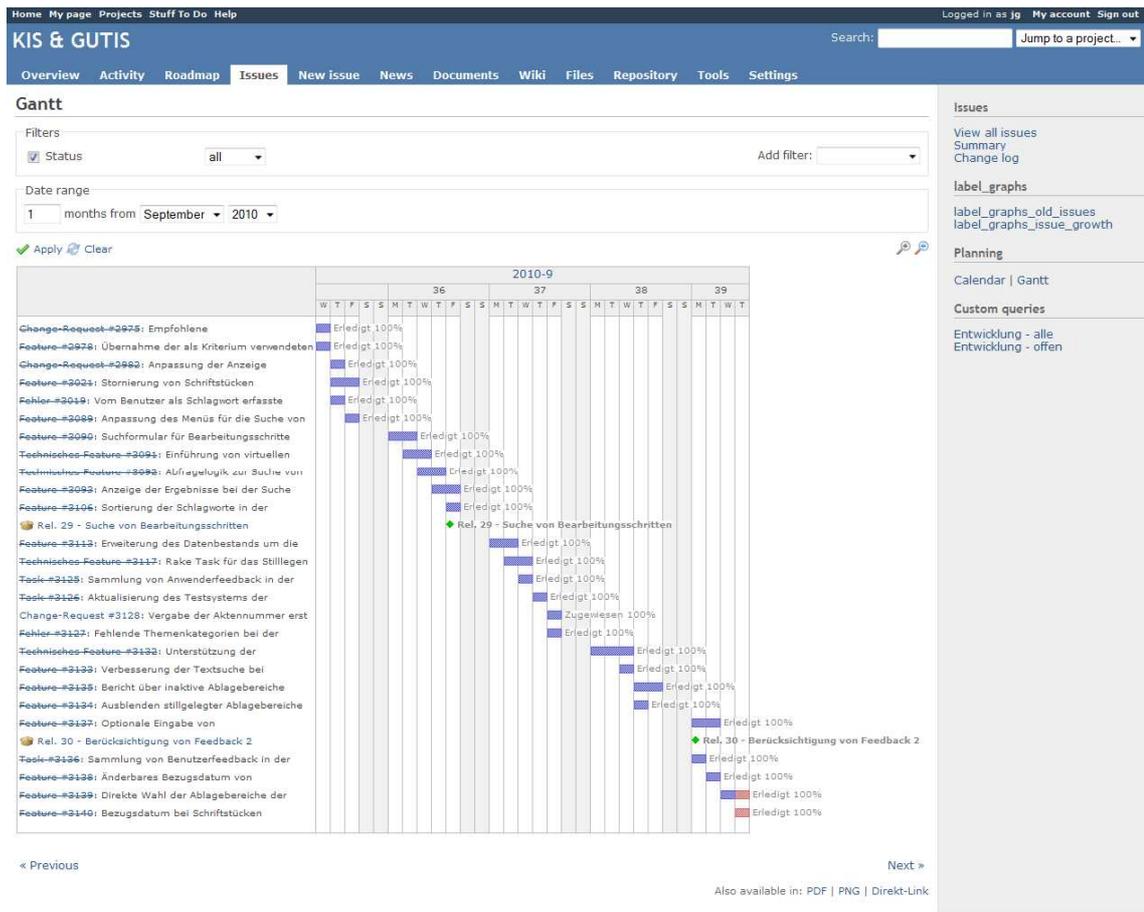


Abbildung 2: Projektmanagement mit Redmine

4 Anforderungsanalyse

Dieses Kapitel beschreibt das Ergebnis der begleitend zu den Entwicklungsiterationen durchgeführten Anforderungsanalyse.

Der erste Abschnitt dokumentiert den Ausgangszustand zu Beginn des Projekts. Es wird ein Blick auf den Aufgabenbereich der Kanzlei und das in Betrieb befindliche Kanzleinformationssystem geworfen. Die Verwaltung von Dokumenten in der „Gebäude und Technik“ sind ebenfalls ein Thema.

Im Anschluss werden die Probleme der Anwender beim Dokumentenmanagement im Detail betrachtet und kategorisiert. Die Problemlösung spielte eine maßgebliche Rolle bei der Anforderungserhebung. Die erhobenen Anforderungen sind folglich der Inhalt des dritten Abschnitts.

4.1 Ausgangszustand

In diesem Abschnitt und seinen Unterpunkten ist die im Rahmen der Anforderungsanalyse erfasste Ausgangslage dokumentiert. Der Ausgangszustand in der Kanzlei wurde im September und Oktober 2008 ermittelt. Die Erhebung in der „Gebäude und Technik“ erfolgte anschließend im November 2008.

4.1.1 Kanzlei

Das Kanzleinformationssystem, abgekürzt „KIS“, ist seit über einem Jahrzehnt in Betrieb. Die Mitarbeiter der Kanzlei erfassen per Post eingehende Dokumente als Schriftstücke im System und teilen diese den Organisationseinheiten der Universität zu. Unter einem Schriftstück wird in der Kanzlei, ein einzelnes Dokument oder auch eine Gruppe von zusammengehörigen Dokumenten zu einem Thema verstanden. Der hausinterne Transport von Dokumenten wird von einem Boten übernommen, der zweimal pro Tag die Kanzlei besucht und die Dokumente entgegennimmt. Im Normalfall werden Dokumente als Kopien verschickt, wenn das Original nicht erforderlich ist. Die Dokumente benötigen zirka einen halben Tag bis sie auf diesem Weg an ihrem Ziel angekommen sind.

Die Archivierung von Schriftstücken für einen Zeitraum von zehn Jahren fällt ebenfalls in den Zuständigkeitsbereich der Kanzlei. Die Aufbewahrung der Schriftstücke erfolgt in Aktenordnern. Als Dokument und somit Teil eines Schriftstücks zählen auch relevante E-Mails, die an die Kanzlei zur Archivierung weitergeleitet wurden. Die Mitarbeiter drucken diese E-Mails aus und legen sie in Papierform den Schriftstücken bei. Aus Platzgründen können in der Kanzlei nur Schriftstücke aufbewahrt werden, die im Zeitraum der ungefähr letzten fünf Jahre angefallen sind. Ältere Schriftstücke werden im Keller in dafür vorgesehenen Räumlichkeiten gelagert. Die Mitarbeiter der Kanzlei gaben während der Interviews an, dass etwa zwei bis drei Mal in der Woche der Weg in

den Keller notwendig sei. Aktenordner, deren Inhalt älter als zehn Jahre ist, werden an das Archiv der Universität übergeben.

Die Ablage von Schriftstücken im Kanzleiinformationssystem erfolgt nach langjährig etablierten formellen und informellen Konventionen der Mitarbeiter. Mit der Einführung des Kanzleiinformationssystems wurde ein Aktenplan aufgestellt, der die Methode zur Kategorisierung der Schriftstücke anhand formaler Richtlinien festhält. Der Begriff der „Geschäftszahl“, einer eindeutig einem Schriftstück zugeordneten Zeichenfolge, wurde eingeführt.

Die Geschäftszahl besteht aus vier aneinander gereihten Zahlen, von denen die ersten beiden eine semantische Kategorisierung des Schriftstücks in einer zweistufigen Themenhierarchie ausdrücken. Über die Geschäftszahl ist somit ersichtlich, zu welchem Thema der Inhalt eines Schriftstücks angehört. Abbildung 3 zeigt den schematischen Aufbau einer Geschäftszahl anhand eines Beispiels.

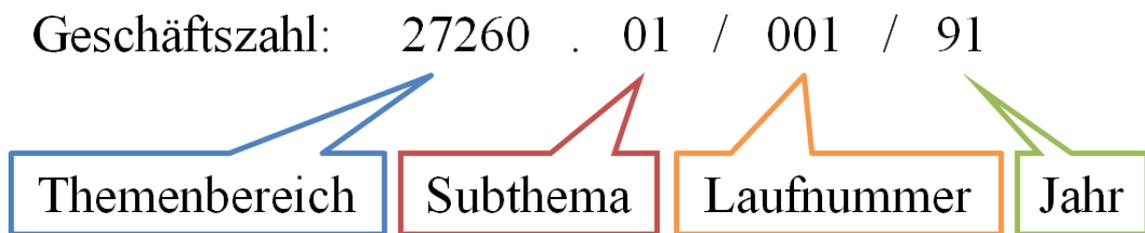


Abbildung 3: Schematischer Aufbau einer Geschäftszahl

Die anschließende Aufzählung beschreibt den Aufbau einer Geschäftszahl im Detail:

- *Themenbereich*
Die Nummer des Themenbereichs ordnet das Schriftstück in der obersten Gliederungsebene einem Element in der Auflistung aller Themen zu. Bei der Vergabe von Nummern an die Themenbereiche wurden in der Universitätsorganisation bereits bestehende Nummerierungen übernommen. Die Nummer für den Themenbereich einer Organisationseinheit entspricht der öffentlich bekannten numerischen Bezeichnung. Über die Zuordnung von Zahlenbereichen zu einer Gruppe von themenverwandten Inhalten, wurde eine übergeordnete Ebene der Gliederung implizit eingeführt.
- *Subthema*
Die Zuordnung zur zweiten Gliederungsebene erfolgt über die Nummer des Subthemas. Für allgemeine Schriftstücke, die keinem Unterthema angehören, ist diese Zahl gleich null. Das Subthema stellt eine weitere Unterteilung eines Themenbereichs dar. In mehreren Themenbereichen eines Zahlenbereichs wiederkehrende Unterthemen tragen im Normalfall auch gleiche Nummern.

- *Laufnummer*
Über die Laufnummer werden Schriftstücke innerhalb eines Themenbereichs, Subthemas und Jahres chronologisch durchgehend nummeriert. Die Nummerierung beginnt bei eins und endet bei 999. Sollten innerhalb eines Jahres 1000 oder mehr Schriftstücke anfallen, muss für den Überlauf eine neue Kombination von Themenbereich und Subthema akquiriert werden.
- *Jahr*
Die Jahreszahl gibt für gewöhnlich Aufschluss über das Jahr, in dem das Schriftstück angefallen ist. Die Ausnahme bilden Serien von zusammengehörigen Schriftstücken über den Zeitraum eines Jahres hinweg. In diesen Fällen können die Folgeschriftstücke bei Bedarf auch unter dem ersten Jahr abgelegt werden.

Zusätzlich zur inhaltlichen Kategorisierung eines Schriftstücks bei der Erfassung durch die Geschäftszahl wird der Inhalt in einem kurzen Beschreibungstext durch den Bearbeiter textuell konkretisiert.

Des Weiteren können mit dem Kanzleiinformationssystem vier vordefinierte Arten von Bearbeitungsschritten für Schriftstücke samt Zeitangaben protokolliert werden. Zu diesen Arten zählen:

- *Erstattung*
Bei der Erstattung wird das Eingehen eines Schriftstücks von einem Absender protokolliert. Dabei kann die vom Absender vergebene Schriftstücknummer eingegeben werden.
- *Zuteilung*
Ein Schriftstück kann einem Empfänger innerhalb der Universität zugeteilt werden. Diese Zuteilung geht wahlweise mit einer Rückgabefrist einher.
- *Expedition*
Unter der Expedition wird der Versand eines Schriftstücks an einen externen Empfänger verstanden. Als weitere Datenangabe kann neben dem Empfänger eine Antwortfrist angegeben werden.
- *Ablage*
Wird ein Schriftstück zur Aufbewahrung abgelegt, kann dies zusammen mit dem Ort der Ablage angegeben werden.

Die soeben erläuterten Bearbeitungsschritte sind optional und können im System beliebig aneinander gereiht werden. Die Logik der Abfolge wird nicht maschinell geprüft.

Zu Beginn war die Benutzerschnittstelle des Kanzleinformationssystems rein textbasiert. Abbildung 4 zeigt einen Auszug aus der Systemdokumentation in dem die Textversion der Hauptmaske zu sehen ist.

```

TUWIS-150           K I S - Kanzlei-Information           lFkt:           TUWIS-150
-----
Themenbereich  Lfd.Nr.  Jahr  Bearbeitung  Text
E              .
-----
Datum:           Inhalt:
-----
Index:           Art der Mitteilung:           Nummer:
                sonst. GZl.:
                bewilligt:
-----
      Nur eingeben, wenn Selektion von Bearbeitungsschritten gewünscht ist:
Absender/Empfänger           Datum
E:                           Fremdzahl:
Z:                           Rückgabe:
X:                           Frist:
A:

```

Abbildung 4: Textversion der KIS-Hauptmaske

Die Benutzerschnittstelle wurde später dann auf eine grafische Oberfläche umgestellt. In Abbildung 5 ist die aktualisierte Version der Benutzerschnittstelle ersichtliche.

Abbildung 5: KIS-Hauptmaske in grafischer Oberfläche

Neben der Kanzlei haben noch andere organisatorische Einheiten der Universität Zugriff auf das Kanzleinformationssystem. Darunter fallen Rektorat, Archiv und Studienabteilung.

4.1.2 Gebäude und Technik

In der Organisationseinheit „Gebäude und Technik“ werden anfallende Dokumente in einer gemeinsam verwendeten Verzeichnisstruktur abgelegt. Diese Ressource ist über das Netzwerk erreichbar. Die Kategorisierung in Ordner erfolgt auf der ersten Ebene nach der Person, die ihre Dokumente ablegt. Anwendungsspezifische Elemente wie zum Beispiel Gebäude, Projekte und Kostenstellen bilden in Unterordnern weitere Gliederungsebenen. Zur Regelung der individuellen Ablage durch die Mitarbeiter wurde anhand von abteilungsinternen Konventionen ein Standard geschaffen.

Zusätzlich zur elektronischen Ablage auf Netzwerkressourcen mit gemeinsamem Zugriff werden Dokumente in Papierform in den Büros der Mitarbeiter gelagert. Die Aufbewahrung erfolgt in herkömmlichen Aktenordnern und Aktenschränken. Die Inhalte der Akten eines Schrankes sind einem definierten gemeinsamen Bereich von Themen zuordenbar.

4.2 Probleme der Anwender

Nach der Erhebung des Ausgangszustands durch Interviews und die Analyse der Dokumentation des Kanzleinformationssystems wurden weiterführende persönliche Gespräche mit den Mitarbeitern der Kanzlei geführt. Mit diesen Besprechungen wurden mehrere Ziele verfolgt.

Zum einen wurde der wertvolle Erfahrungsschatz der Mitarbeiter unter der Berücksichtigung ihrer alltäglichen Probleme in die Anforderungserhebung eingebunden, zum anderen konnten sich die Mitarbeiter am Gestaltungsprozess mit ihren Wünschen und Erwartungen beteiligen. Des Weiteren wurde die Gelegenheit geschaffen, auf etwaige Befürchtungen der Anwender eingehen zu können. Es galt dabei Ängste vor einer langen Einarbeitungszeit, fehlenden Funktionen und weitreichenden Umstrukturierungen etablierter Arbeitsweisen durch eine neue Software abzubauen.

Außerdem willigten die Mitarbeiter ein, sich bei der Abarbeitung einiger typischer Geschäftsfälle mit dem Kanzleinformationssystem beobachten zu lassen. Die Tätigkeiten wurden während der Ausführung vom Bearbeiter kommentiert. Im Anschluss an die Demonstration fand die Behandlung aufgetretener Fragen statt. Im Rahmen dieser Sitzung konnte das Verständnis für den Anwendungsbereich vertieft und die Liste der Probleme um weitere Elemente ergänzt werden. Die resultierende Auflistung wurde in diesem Abschnitt dokumentiert.

In der „Gebäude und Technik“ entfiel der Schritt einer vergleichbaren detaillierten Problemerkhebung, weil erstens dort bisher keine auf das Dokumentenmanagement spezialisierte Software eingesetzt wurde und zweitens in der Organisationseinheit die Konzeption eines über das gemeinsame Verwenden von Netzwerkressourcen hinausgehenden strukturierten Prozesses erst begonnen wurde. Des Weiteren war zu Projektbeginn nur die Kanzlei als Interessent an dem Dokumentenmanagementsystem bekannt.

Auf Anfrage der „Gebäude und Technik“ wurde das in Entwicklung befindliche Konzept für die Weiterentwicklung des Kanzleinformationssystems vorgestellt. Das Konzept fand in der „Gebäude und Technik“ Zustimmung und bildete den Ausgangspunkt für eine weitere Zusammenarbeit. In der Folge wurde das Konzept auf die Bedürfnisse der Organisationseinheit adaptiert. Aus den erhaltenen Einblicken konnte die für die Kanzlei begonnene Auflistung von Anwenderproblemen um weitere die „Gebäude und Technik“ betreffende Einträge erweitert werden.

Für die Weiterverarbeitung wurden die erfassten Probleme nach ihren zu Grunde liegenden Anwendungsbereichen kategorisiert. Für jeden Anwendungsbereich wurde ein eindeutiges Kürzel vergeben, welches später in der Auflistung der Anforderungen wiederverwendet wird. Auf diesem Wege soll die Zuordnung zum jeweiligen Bereich anhand des Bezeichners erleichtert werden. Die Kürzel wurden hinter dem Bereichsnamen in Klammern angeführt. Für die Kategorisierung wurden diese Bereiche identifiziert:

- *Dokumentenverwaltung (DOC)*
In diesen Bereich fallen alle Schwierigkeiten, die unter dem generellen Umgang mit Dokumenten einzuordnen sind.
- *Suche (SEA)*
Diese Kategorie betrifft bei der Suche von erfassten Dokumenten auftretende Probleme.
- *Drucken (PRI)*
Nicht vorhandene Möglichkeiten, Abzüge von gewünschten Informationen per Ausdruck auf Papier bringen zu können, sind unter diesem Punkt einzuordnen.
- *Schnittstellen (INT)*
Zusätzliche Aufwände aufgrund fehlender oder unzureichender Schnittstellen sind Thema dieses Bereichs.
- *Befugnisse (PER)*
Problematiken bei der Regelung des Zugriffs auf Informationen bilden die fünfte und letzte Kategorie.

Es folgt die Beschreibung der in der Kanzlei aufgenommenen Anwenderprobleme kategorisiert nach den genannten Bereichen. Zur eindeutigen Identifizierung und späteren Referenzierung wurden die einzelnen Probleme nummeriert:

- *Dokumentenverwaltung*

Diese Kategorie enthält allgemeine Probleme der Kanzlei, die den generellen Umgang mit Dokumenten betreffen:

1. *Fehlende Unterstützung bei der Dokumentablage durch angepasste Software*

Die Ordnung wird mehr durch die Erfahrungen und etablierten Konventionen der Anwender als durch das System bestimmt.

2. *Keine themenübergreifende Zuordnung*

Schriftstücke können nur einem Themenkreis zugeordnet werden. Themenübergreifende Inhalte können nicht abgebildet werden. Ohne die von den Anwendern geschaffenen Konventionen zu kennen, können mehrere Zuordnungsmöglichkeiten plausibel erscheinen.

3. *Platzprobleme*

Die Lagerung der Akte der letzten zehn Jahre benötigt mehr Platz als im Büro vorhanden ist. Einige Schriftstücke bestehen aus sehr vielen Dokumenten und sind deshalb schwierig handzuhaben.

4. *Überlauf von Laufnummern*

Falls in einem Jahr zu einem Themenkreis zu viele Schriftstücke anfallen, muss ein neuer Themenkreis angelegt werden, weil die Laufnummer auf drei Stellen begrenzt ist. Dieses Problem tritt in der Kanzlei jedoch nicht auf.

5. *Geschäftszahlen für zukünftige Schriftstücke*

Es können Geschäftszahlen für Schreiben angefordert werden, die noch nicht existieren. In so einem Fall besteht die Schwierigkeit einen Inhaltstext anzugeben unter dem der Eintrag wiedergefunden werden kann.

6. *Freie Nummern für Themenbereiche*

Die Anwender führen eine handschriftliche Liste mit den zuletzt vergebenen Themenkreisen um über die nächste freie Nummer Bescheid zu wissen. Das System vergibt diese Nummern nicht automatisch.

7. *Nachvollziehbarkeit bei Inhaltstexten*

Der Inhaltstext bei einem Schriftstück ergibt sich durch nicht dokumentierte Konventionen der Anwender. Diese Konventionen legen fest, welche beschrei-

benden Attribute wie zum Beispiel Vertragsname, Paragraph, etc. angeführt werden sollen. Dieses Wissen lässt sich für neue Anwender oder Dritte schwer aus dem System herleiten.

8. *Aufbewahrungsort eines Schriftstücks*

Der Aufbewahrungsort eines Schriftstücks ist nicht direkt ersichtlich. Es müssen Rückschlüsse über die Reihenfolge der Bearbeitungsschritte gezogen werden. Der Aufbewahrungsort gibt Aufschluss über die für die Aufbewahrung zuständige Organisationseinheit und ist vom Begriff „Aufenthaltort“ zu unterscheiden. Der Aufenthaltort ist der Ort, an dem sich ein Schriftstück befindet.

9. *Behandlung von Stornierungen*

Stornierungen werden durchgeführt, indem dem Themenkreis- beziehungsweise Inhaltstext ein entsprechender Hinweis hinzugefügt wird. Bei Personalakten wird versucht lückenlos durchgängige Laufnummern zu vergeben, weil diese vom Personalbüro gewünscht werden. Daten von stornierten Schriftstücken werden mit Daten neuer Schriftstücke überschrieben, um die fortlaufende Nummerierung zu gewährleisten.

10. *Handhabung von Personalakten*

Bei Dokumenten für den Abgang von Personal gibt es jeweils ein Datum für das Ausscheiden und die Ablage. Es wird nur eines der beiden Daten im Kanzleiinformationssystem vermerkt. Dokumente für den Abgang von Personal sind nicht besonders gekennzeichnet. Wieder ausgehobene Personalakten sind ebenfalls nicht ersichtlich.

- *Suche*

Die folgenden Unterpunkte beschreiben Missstände bei der Abfrage von bereits im System erfassten Daten:

1. *Begrenzte Suchmöglichkeiten*

Es kann nicht in mehr als einem Themenbereich gesucht werden. Die Suche erfolgt entweder in einem oder in allen Bereichen. Folglich kann nicht ausschließlich in einer Gruppe zusammengehöriger Themenbereiche gesucht werden.

2. *Limitierte Anzahl der Suchergebnisse*

Die Anzahl der angezeigten Suchergebnisse ist zahlenmäßig beschränkt. In Folge dieser Einschränkung werden nicht immer alle Ergebnisse angezeigt.

3. *Nicht unterscheidbare Suchergebnisse*

Bei der Anzeige von Suchergebnissen kann es vorkommen, dass Einträge aufgrund gleichen Texts auf den ersten Blick nicht von einander unterscheidbar sind. Der richtige Eintrag kann in so einem Fall nur durch das nachfolgende Aufrufen der Detailansichten gefunden werden. Dieses Problem tritt zum Beispiel bei Personalakten auf, die unter dem Namen der Person zu finden sind.

4. *Suche mit ungenauen Datumsangaben*

Datumsangaben in eingehenden Dokumenten sind ungenau oder umfassen einen größeren Zeitraum (zum Beispiel „im September 2008“). Bei der Suche wird jedoch ein konkretes Datum erwartet.

- *Drucken*

Aufgrund der Ablage von Schriftstücken in Papierform besteht für die Anwender die Notwendigkeit Informationen in elektronischer Form den Akten als Ausdruck beizulegen. Das Drucken der gewünschten Informationen wird vom Kanzleinformationssystem nicht unterstützt.

1. *Beschriftung von Umschlägen per Hand*

Um zusammenhängende Dokumente zu gruppieren wird ein auf A4 gefaltetes Papier im Format A3 als Umschlag verwendet. Die Beschriftung dieser Umschläge erfolgt per Hand. Dies ist ein zusätzlicher Aufwand und führt eventuell zu Problemen mit der Lesbarkeit.

2. *Schriftstückdaten nicht ausdrückbar*

Schriftstückdaten und andere Informationen sind im bestehenden Kanzleinformationssystem nicht ausdrückbar. Bei der Notwendigkeit eines Ausdrucks werden Bildschirmabbilder angefertigt.

- *Schnittstellen*

Fehlende Schnittstellen zu anderen Systemen bilden zusätzliche Erschwernisse für die Benutzer:

1. *Projektdatenbank*

Einträge aus der Projektdatenbank müssen manuell übertragen werden. Das Übertragen von Informationen stellt eine fehleranfällige Tätigkeit dar. Werden Daten manuell dupliziert und danach in einem der Systeme geändert, stellt sich die Frage, wie diese Änderungen an die betreffenden Datensätze der anderen Systeme weitergeben werden sollen. Ohne einem beachtlichen Synchronisationsaufwand oder die Schaffung eines automatisierten Weges werden die Datenbestände der beteiligten Systeme widersprüchlich.

Nach der soeben durchgeführten Aufzählung für die Kanzlei ist die nachfolgende Liste den in der „Gebäude und Technik“ erkannten Problemen gewidmet.

- *Dokumentenverwaltung*

Wie bereits bei der oben vorgenommenen Auflistung für die Kanzlei enthält diese Kategorie allgemeine Probleme, die den generellen Umgang mit Dokumenten in der „Gebäude und Technik“ betreffen.

1. *Fehlende Unterstützung bei der Dokumentablage durch angepasste Software*

Ob die individuelle Ablage von Dokumenten der vorgeschlagenen Konvention entspricht, ist stark vom Wissenstand und der Arbeitsweise des Anwenders abhängig. Persönliche Einflüsse kommen stärker zu Geltung als das bei der Unterstützung der Strukturierung durch ein Softwaresystem der Fall wäre. Die Konsequenzen können sehr unterschiedliche Ablagestrukturen und eine verbesserungswürdige Wiederauffindbarkeit sein.

- *Suche*

Aus der Ablage in gemeinsam verwendeten Ordnern ergibt sich für das Wiederfinden von Dokumenten folgende Problematik:

1. *Begrenzte Suchmöglichkeiten*

Die Funktionalität der Dokumentsuche ist auf die vom Betriebssystem bereitgestellte Dateisuche limitiert. Das primäre Kriterium bei der Suche in einem Dateisystem ist der Dateiname. Deswegen sind gute Suchergebnisse stark von der Vergabe von sprechenden und einheitlichen Verzeichnis- und Dateinamen abhängig. Vom Betriebssystem unterstützte allgemeine Dateiattribute, wie zum Beispiel ein Änderungsdatum, können auf den Anwendungsfall spezialisierte Metadaten keinesfalls ersetzen.

- *Befugnisse*

Der Zugriff auf Informationen muss geregelt werden. Einerseits sind sensible Informationen vor dem unrechtmäßigen Zugriff zu schützen, andererseits erleichtert die Unsichtbarkeit irrelevanter Informationen das Arbeiten der Beteiligten.

1. *Mehraufwand bei der Konfiguration*

Die Regelung von Zugriffsberechtigungen auf Informationen nach ihrem Themengebiet gestaltet sich durch die individuell hierarchische Ablage sehr aufwendig und fehleranfällig. Berechtigungen müssen in der zweiten oder einer tieferen Gliederungsebene bei redundanten Verzeichnissen mehrfach konfiguriert werden.

2. *Keine anwendungsspezifische Berechtigungsgranularität*

Eine einzelne Datei entspricht der kleinsten logischen Einheit, der in einem Dateisystem Berechtigungen zugewiesen werden können. Die zum Einsatz kommenden elementaren Berechtigungsarten Lesen und Schreiben decken nur den Grundbedarf einer anwendungsspezifischen Zugriffskontrolle. Realistische Verwendungsszenarien inkludieren die Bearbeitungsrechte von Informationseinheiten mit feinerer Granularität wie zum Beispiel das Setzen einzelner Metadaten eines Dokuments.

Zwei der vier Unterpunkte in der soeben vorgenommenen Aufzählung decken sich mit bereits in der Problemliste der Kanzlei erkannten Punkten. Die Kategorie „Befugnisse“ bringt einzig und allein einen neuen Aspekt in die Sammlung der Anwendungsergebnisse.

Das Problem einer möglichst strukturierten Ablage neuer Dokumente ist in beiden Organisationseinheiten gegenwärtig, wenn es sich auch in anderen Formen zu erkennen gibt. Gleiches gilt für die Beschränkung von möglichen Suchkriterien. In beiden Herangehensweisen erfüllen die vorhandenen Möglichkeiten der Dokumentensuche die Bedürfnisse der Anwender nicht ausreichend.

Tabelle 4 fasst die in der Kanzlei und „Gebäude und Technik“ aufgenommen Probleme in einer gemeinsamen Liste mit fortlaufender Nummerierung zusammen. In der Auflistung ist jedes erkannte Problem nur einmal vertreten, selbst wenn es in beiden Organisationseinheiten auftritt. Die Herkunft eines Eintrags ist über seine Zuordnung ersichtlich.

Über den Bezeichner lässt sich ein Problem in den für die beiden Organisationseinheiten einzeln erstellten Listen auffinden. Diese Bezeichner sind aus dem Kürzel für den Anwendungsbereich und der Nummer des Eintrags in der jeweiligen Liste zusammengesetzt.

Problem		Herkunft	
Bezeichner	Titel	Kanzlei	Gebäude und Technik
DOC	Dokumentenverwaltung		
DOC.1	Fehlende Unterstützung bei der Dokumentenablage durch angepasste Software	x	
DOC.2	Keine themenübergreifende Zuordnung	x	
DOC.3	Platzprobleme	x	
DOC.4	Überlauf von Laufnummern	x	
DOC.5	Geschäftszahlen für zukünftige Schriftstücke	x	
DOC.6	Freie Nummern für Themenbereiche	x	
DOC.7	Nachvollziehbarkeit bei Inhaltstexten	x	
DOC.8	Aufbewahrungsort eines Schriftstücks	x	
DOC.9	Behandlung von Stornierungen	x	
DOC.10	Handhabung von Personalakten	x	
SEA	Suche		
SEA.1	Beschränkte Suchkriterien	x	
SEA.2	Limitierte Anzahl der Suchergebnisse	x	
SEA.3	Nicht unterscheidbare Suchergebnisse	x	
SEA.4	Suche mit ungenauen Datumsangaben	x	
PRI	Drucken		
PRI.1	Beschriftung von Umschlägen per Hand	x	
PRI.2	Schriftstückdaten nicht ausdrückbar	x	
INT	Schnittstellen		
INT.1	Projektdatenbank	x	
PER	Befugnisse		
PER.1	Mehraufwand bei der Konfiguration		x
PER.2	Keine anwendungsspezifische Berechtigungsgranularität		x

Tabelle 4: Anwenderprobleme inklusive Herkunft nach Organisationseinheit

4.3 Anforderungen

Die Liste der Anforderungen ist das Ergebnis mehrerer Einflüsse. Die Anforderungsanalyse wurde auf Basis der von der Führung des Projekts „TISS“ geforderten Vorgabe durchgeführt, mit den bestehenden Arbeitsabläufen der Anwender kompatibel zu bleiben. Das in Verwendung stehende Kanzleinformationssystem lieferte mit seiner bestehenden Funktionalität einen Teil der Anforderungen.

Ein weiterer Teil ergab sich aus den mit den Anwendern durchgeführten Interviews. Durch ihren in der Praxis gesammelten Erfahrungsschatz konnten sie ihnen bekannte Problembereiche gezielt ansprechen.

Vervollständigt wurden die Anforderungen durch Ergänzungen des Autors, welche ihren Ursprung in der Beobachtung der Anwender beim Arbeiten mit dem System und der

intensiven Auseinandersetzung mit den Verbesserungsmöglichkeiten im Fachgebiet hatten.

Die Begründung der vorgenommenen Ergänzungen ist im kommenden Abschnitt „Systemkonzept“ nachzulesen. Tabelle 5 zeigt, welche der im Anwendungsbereich erkannten Probleme von den Anforderungen bezüglich einer Lösung adressiert werden.

		Probleme																					
		DOC.1	DOC.2	DOC.3	DOC.4	DOC.5	DOC.6	DOC.7	DOC.8	DOC.9	DOC.10	SEA.1	SEA.2	SEA.3	SEA.4	PRI.1	PRI.2	INT.1	PER.1	PER.2			
		← Art der Anforderung	Fehlende Software-Unterstützung	Keine themenübergreifende Zuordnung	Platzprobleme	Überlauf von Laufnummern	Geschäftszahlen für zukünftige Schriftstücke	Freie Nummern für Themenbereiche	Nachvollziehbarkeit bei Inhaltstexten	Aufbewahrungsort eines Schriftstücks	Behandlung von Stornierungen	Handhabung von Personalakten	Begrenzte Suchmöglichkeiten	Limitierte Anzahl der Suchergebnisse	Nicht unterscheidbare Suchergebnisse	Suche mit ungenauen Datumsangaben	Beschriftung von Umschlägen per Hand	Schriftstückdaten nicht ausdrückbar	Projektdatenbank	Mehraufwand bei der Konfiguration	Keine spezifische Berechtigungsgranularität		
		Ursache für Problementdeckung →																					
		I	I	I	G	G	G	G	I	G	G	G	B	B	B	B	B	G	G	I	I		
Anforderungen	DOC.1	Dokumente in Papier und elektronischer Form	V	x	x																		
	DOC.2	Inhaltstexte als Schlagwortlisten	P		x				x														
	DOC.3	Einheitliche Inhaltstexte für Schriftstücke	P						x														
	DOC.4	Entwürfe von Schriftstückeinträgen	E				x																
	DOC.5	Automatische Vergabe von Laufnummern	E			x		x															
	DOC.6	Entwurf von Schriftstückbeschreibungen	P						x														
	DOC.7	Protokolle für Suche und Einsichtnahme	P										x										
	DOC.8	Stornierung von Schriftstücken	E								x												
	DOC.K.1	Beibehaltung bestehender Nummerierungen	V																				
	DOC.K.2	Aufbewahrungsort von Schriftstücken	E							x													
	DOC.K.3	Benachrichtigungen bei Fristen	E																				
	DOC.K.4	Referenzen auf Schriftstücke	A																				
	SEA.1	Inhaltliche Vernetzung	P										x			x							
	SEA.2	Schrittweises Verfeinern von Kriterien	P										x										
	SEA.3	Suche nach kombinierten Kriterien	A																				
	SEA.4	Unbeschränkte Anzahl an Suchergebnissen	E											x									
	SEA.5	Aufgelöste Themenkreise ausblenden	E										x										
	SEA.K.1	Suche nach Bearbeitungsschritten	A																				
	PRI.K.1	Komfortables Ausdrucken	E														x	x					
	PRI.K.2	Umschläge für Schriftstücke	E														x						
	REP.K.1	Rückstandsausweise	E																				
	REP.K.2	Ausgeschiedenes Personal	E									x											
	INT.K.1	Projektdatenbank	E																	x			
INT.G.1	SAP	E																					
PER.G.1	Rollenbasiertes Berechtigungssystem	E																			x		
PER.G.2	Aktionsbasierte Berechtigungsgranularität	E																				x	

Tabelle 5: Zusammenhang der adressierten Probleme und Anforderungen

4.3.1 Funktionale Anforderungen

Die folgende Auflistung enthält die gemeinsamen Anforderungen an die Funktionen des zu entwickelnden Systems, welche für beide Organisationseinheiten ihre Gültigkeit haben:

- *Dokumentenverwaltung*

1. *Dokumente in Papier und elektronischer Form*

Das System muss mit Dokumenten in Papier und elektronischer Form umgehen können. Ein Schriftstück kann Dokumente in beiden Formen beinhalten.

2. *Inhaltstexte als Schlagwortlisten*

Die Inhaltstexte für Schriftstücke sollen sich aus einer Liste von Schlagwörtern zusammensetzen.

3. *Einheitliche Inhaltstexte für Schriftstücke*

Der Anwender soll bei der Eingabe einheitlicher Inhaltstexte für Schriftstücke unterstützt werden.

4. *Entwürfe von Schriftstückeinträgen*

Schriftstücke sollen auch als Entwurf abgelegt werden können. Entwürfe sind aus der Anzeige der regulären Suchergebnisse ausgenommen und bleiben bis zur Freigabe beziehungsweise zum Verwurf in einem Arbeitsvorrat.

5. *Automatische Vergabe von Laufnummern*

Geschäftszahlen und Nummern für neue Themenkreise sollen vom System vorgeschlagen werden.

6. *Entwurf von Schriftstückbeschreibungen*

Beschreibungsdaten von Schriftstücken im Entwurfsstatus sollen jederzeit bearbeitbar sein.

7. *Protokolle für Suche und Einsichtnahme*

Zuletzt durchgeführte Suchanfragen und angesehene Schriftstücke sollen in einem Protokoll für den Schnellzugriff wieder aufrufbar sein.

8. *Stornierung von Schriftstücken*

Schriftstücke sollen stornierbar sein. Stornierte Schriftstücke werden standardmäßig in den Suchergebnissen nicht angezeigt. Auf Wunsch ist eine Einsichtnahme in stornierte Schriftstücke zu gestatten.

- *Suche*

1. *Inhaltliche Vernetzung*

Inhaltlich verwandte Themenkreise beziehungsweise Schriftstücke sollen miteinander vernetzt werden, sodass das Springen zwischen ähnlichen und verwandten Inhalten möglich wird. Die Suche soll so ein „Stöbern“ im Inhalt ermöglichen.

2. *Schrittweises Verfeinern von Kriterien*

Die Suchkriterien sollen schrittweise auf Basis des Suchergebnisses verfeinert werden können um die Ergebnisse weiter einzuschränken.

3. *Suche nach kombinierten Kriterien*

Unterschiedliche Suchkriterien wie zum Beispiel Inhalt, Bearbeitungsschritte, usw. sollen miteinander kombinierbar sein.

4. *Unbeschränkte Anzahl an Suchergebnissen*

Die Anzahl der Suchergebnisse beliebiger Abfragen soll unbeschränkt sein.

5. *Aufgelöste Themenkreise ausblenden*

Bei der Suche sollen aufgelöste Themenkreise ausblendbar sein.

Die vorangegangene Liste wird durch diese Anforderungen der Kanzlei erweitert:

- *Dokumentenverwaltung (Kanzlei)*

1. *Beibehaltung bestehender Nummerierungen*

Bestehende Nummern wie Geschäftszahlen, Themenkreise und Schriftstücke sollen wie bisher verwendet werden können.

2. *Aufbewahrungsort von Schriftstücken*

Der Aufbewahrungsort von Schriftstücken soll im System ersichtlich sein. Bei der Übergabe der Zuständigkeit für die Aufbewahrung an eine andere Organisationseinheit soll der Aufbewahrungsort für ein oder mehrere Schriftstücke aktualisierbar sein.

3. *Benachrichtigungen bei Fristen*

Der Anwender soll beim Ablaufen von an Schriftstücke gebundene Fristen benachrichtigt werden.

4. *Referenzen auf Schriftstücke*

Nachfolgende beziehungsweise themenverwandte Schriftstücke sollen in den Beschreibungsdaten eines Schriftstücks referenzierbar sein.

- *Suche (Kanzlei)*

1. *Suche nach Bearbeitungsschritten*

Schriftstücke sollen über eine Suche nach Bearbeitungsschritten auffindbar sein.

- *Berichte (Kanzlei)*

1. *Rückstandsausweise¹¹*

Auf Wunsch soll eine Liste der ausständigen Schriftstücke inklusive dazugehöriger Aufenthaltsorte und Fristen erstellbar sein.

2. *Ausgeschiedenes Personal*

Ein Bericht über das ausgeschiedene Personal für ein gewünschtes Jahr soll erstellt werden können. Diese Informationen sind aus den Personalakten beziehbar.

- *Drucken (Kanzlei)*

1. *Komfortables Ausdrucken*

Angezeigte Informationen wie zum Beispiel Schriftstückdaten sollen einfach auszudrucken sein.

2. *Umschläge für Schriftstücke*

Am Umschlag für Schriftstücke anzubringende Daten sollen ausdrückbar sein.

- *Schnittstellen (Kanzlei)*

1. *Projektdatenbank*

Eine Schnittstelle zur Projektdatenbank soll einen automatischen Abgleich und Transfer von Projektdaten zwischen den Systemen ermöglichen.

Die Anforderungen der „Gebäude und Technik“ vervollständigen die Liste aller erfassten Anforderungen:

¹¹ Wiederanforderung der Kanzlei eines versandten Schriftstücks nach Ablauf einer Frist

- *Schnittstellen („Gebäude und Technik“)*
 1. *SAP*

Projektdateien aus dem in Verwendung stehenden SAP-System sollen in das Dokumentenmanagement eingebunden werden.

- *Befugnisse („Gebäude und Technik“)*
 1. *Rollenbasiertes Berechtigungssystem*

Berechtigungen sollen an definierbare Rollen vergeben werden können. Die Verwaltung der Befugnisse erfolgt durch die Zuordnung dieser Rollen zu Personen.

 2. *Aktionsbasierte Berechtigungsgranularität*

Neben der Berücksichtigung elementarer Lese- und Schreibrechte sollen Berechtigungen auf Ebene der vom Benutzer im System durchführbaren Aktionen vergeben werden können. Die Grundlage hierfür wird aus den Anwendungsfällen geliefert.

Die soeben durchgeführten Auflistungen von Anforderungen werden in Tabelle 6 für eine einfache weitere Bezugnahme kategorisiert und mit fortlaufender Nummerierung zusammengefasst.

Es kommen wiederholt die im Abschnitt 4.2 bereits verwendeten Anwendungsbereiche zur Kategorisierung zum Einsatz. Diese Liste der Anwendungsbereiche wird für die Anforderungen um den folgenden Eintrag erweitert:

- *Berichte (REP)*

Unter diesen Bereich fallen Anforderungen zur Erstellung von Auswertungen und Berichten aus dem im System vorhandenen Datenbestand.

Der eindeutige Bezeichner eines Tabelleneintrags enthält zusätzlich den Anfangsbuchstaben der Organisationseinheit, falls es sich um keine für Kanzlei und „Gebäude und Technik“ gemeinsam gültige Anforderung handelt. Es folgt die soeben genannte Tabelle:

Anforderungen		Herkunft	
Bezeichner	Titel	Kanzlei	Gebäude und Technik
DOC	Dokumentenverwaltung		x
DOC.1	Dokumente in Papier und elektronischer Form		x
DOC.2	Inhaltstexte als Schlagwortlisten		x
DOC.3	Einheitliche Inhaltstexte für Schriftstücke		x
DOC.4	Entwürfe von Schriftstückeinträgen		x
DOC.5	Automatische Vergabe von Laufnummern		x
DOC.6	Entwurf von Schriftstückbeschreibungen		x
DOC.7	Protokolle für Suche und Einsichtnahme		x
DOC.8	Stornierung von Schriftstücken		x
DOC.K	Dokumentenverwaltung (Kanzlei)	x	
DOC.K.1	Beibehaltung bestehender Nummerierungen	x	
DOC.K.2	Aufbewahrungsort von Schriftstücken	x	
DOC.K.3	Benachrichtigungen bei Fristen	x	
DOC.K.4	Referenzen auf Schriftstücke	x	
SEA	Suche		x
SEA.1	Inhaltliche Vernetzung		x
SEA.2	Schrittweises Verfeinern von Kriterien		x
SEA.3	Suche nach kombinierten Kriterien		x
SEA.4	Unbeschränkte Anzahl an Suchergebnissen		x
SEA.5	Aufgelöste Themenkreise ausblenden		x
SEA.K	Suche (Kanzlei)	x	
SEA.K.1	Suche nach Bearbeitungsschritten	x	
PRI.K	Drucken (Kanzlei)	x	
PRI.K.1	Komfortables Ausdrucken	x	
PRI.K.2	Umschläge für Schriftstücke	x	
REP.K	Berichte (Kanzlei)	x	
REP.K.1	Rückstandsausweise	x	
REP.K.2	Ausgeschiedenes Personal	x	
INT.K	Schnittstellen (Kanzlei)	x	
INT.K.1	Projektdatenbank	x	
INT.G	Schnittstellen („Gebäude und Technik“)		x
INT.G.1	SAP		x
PER.G	Befugnisse („Gebäude und Technik“)		x
PER.G.1	Rollenbasiertes Berechtigungssystem		x
PER.G.2	Aktionsbasierte Berechtigungsgranularität		x

Tabelle 6: Anforderungen inklusive Herkunft nach Organisationseinheit

4.3.2 Nicht funktionale Anforderungen

Neben den Anforderungen an die Funktionalität muss das System auch andere Qualitätskriterien erfüllen. Die Annahme des Systems durch die Anwender hängt maßgeblich von diesen nicht funktionalen Anforderungen ab, die aus dem von der Leitung des Projekts „TISS“ gestellten Anspruch des einfachen Auffindens von Dokumenten gefolgert wurden:

- *Suchleistung*

Die Ergebnisse von an das System gestellten Suchanfragen sind im Zeitrahmen von für Web-Anwendungen üblichen Antwortzeiten zur retournieren. Als konkreter Wert werden im Normalfall ein bis zwei Sekunden für typische Suchanfragen angenommen.

- *Schutz der Privatsphäre der Anwender*

Mit über Anwender gesammelten Daten hat das System einen sorgfältigen und diskreten Umgang zu pflegen. Falls vom System Daten über den Anwender gesammelt werden, die laut Anforderungen nicht zwingend für den Systembetrieb notwendig sind, ist dem Anwender die Möglichkeit bereitzustellen, diese Daten zu löschen. Darunter fallen gespeicherte Suchvorgänge und Auflistungen der zuletzt aufgerufenen Akten, um zwei Beispiele für solche Daten zu nennen.

5 Entwurf

In diesem Kapitel wird ein ganzheitliches Systemkonzept zur Umsetzung des Dokumentenmanagementsystems vorgestellt.

Im Anschluss an die Anforderungsanalyse im vorigen Kapitel wird das hinter der Entwicklung stehende Konzept beschrieben und begründet. Die Designentscheidungen werden auf eine nachvollziehbare Weise offen gelegt und kommentiert.

Dem Systemkonzept wird im Abschnitt „Software-Architektur“ eine konkrete Form verliehen. Anhand von Diagrammen und begleitenden Texten wird die Architektur des Systems dokumentiert.

Vervollständigt wird dieses Kapitel durch die Definition einer Vorgehensweise für die Migration der Daten aus dem Altsystem. Das Verfahren zur Übernahme wird schrittweise aufgelistet und erklärt.

5.1 Systemkonzept

Der Ausgangszustand, die Probleme der Anwender und die erfassten Anforderungen bilden neben den Vorgaben der Projektführung den Ausgangspunkt für die Erstellung eines ganzheitlichen Systemkonzepts. Die Kompatibilität zu bestehenden Arbeitsweisen zu wahren, stellt eine Rahmenbedingung bei der Konzeption dar. Die Suchfunktion als primäre Einflussgröße auf die Güte eines Dokumentenmanagementsystems zu verstehen, liefert eine grundlegende Annahme für weitere Überlegungen.

Als naheliegender Lösungsansatz wurde die Beibehaltung der im Kanzleisystem vorherrschenden Beschreibungsmethode von Akten per Text und Einordnung in eine Themenhierarchie in Betracht gezogen. Durch die Aufhebung bestehender technischer Limitationen und die Ausdehnung der Suche eines Begriffs über mehrere beschreibende Eigenschaften hinweg, könnte den Anwendern die Arbeit erleichtert werden. Bisher wird im Kanzleiinformationssystem ein Suchbegriff nur mit den Werten der Eigenschaft verglichen unter der die Eingabe stattfand. Um dies anhand eines Beispiels zu erklären, findet man bei der Suche nach der Eingabe eines Inhaltstexts nur Einträge deren Inhaltstext mit dem Suchkriterium übereinstimmt. Einträge, deren Themenbereich oder Unterthema dem Kriterium entsprechen, bleiben jedoch unentdeckt. Nach der Erweiterung des Suchverhaltens würden mit der bisherigen Methode unentdeckte Einträge ebenfalls in den Ergebnissen aufscheinen. Dieser Ansatz zeichnet sich zwar durch seine Simplizität aus, aber adressiert kaum die Probleme der Anwender. Die Konventionen der Anwender bei der Ablage und Formulierung von Inhaltstexten bleiben im gleichen Umfang Erfordernisse für Ordnung und Auffindbarkeit im System.

Dies wird begründet durch die auf Textvergleichen basierende Suchfunktion. Bei dieser Suche wird in den Daten nach dem Vorkommen von Zeichenfolgen aus dem Suchkriterium Ausschau gehalten. Je nach Algorithmus können auch ähnliche Zeichenfolgen, die nicht vollkommen mit dem gesuchten Text übereinstimmen, in den Ergebnissen enthalten sein. In natürlichen Sprachen existieren unterschiedliche Ausdrucksmöglichkeiten. Dies beginnt bereits auf der Ebene eines einzelnen Worts und setzt sich beim Satzbau fort. Ein Wort kann mehrere unterschiedliche Bedeutungen haben. Mehrere Wörter können dieselbe oder eine ähnliche Bedeutung haben. In den beiden soeben beschriebenen Fällen von Wörtern spricht man von Homonymen und Synonymen.

Ist der Suchbegriff bei der Textsuche ein Homonym, sind Teile des Suchergebnisses unerwünscht, wenn nur bestimmte Bedeutungen gefragt sind. Sind alle inhaltlich relevanten Suchergebnisse für einen Begriff erwünscht, erfordert eine Textsuche die Kenntnis aller verwendeten Synonyme dieses Begriffs. Hinzu kommt, dass die Verwendung von Abkürzungen weitere Schwierigkeiten bei der maschinellen Verarbeitung mit sich bringt. Ohne die Miteinbeziehung der Bedeutung und des Kontexts kann nicht auf das Wort geschlossen werden, das sich hinter einer Abkürzung verbirgt. Wenn ein Wort und dessen Abkürzung nicht mit den gleichen Zeichen beginnen, kann unter der Abkürzung das zugehörige Wort mittels Textvergleich nicht mehr gefunden werden.

Eine gleichgestellte Verwendung von mehreren Worten in einem Suchkriterium bringt eine weitere Problematik mit sich. Nicht alle Wörter aus dem Kriterium sollten zur Ausschließung von Suchergebnissen verwendet werden. Zum Beispiel mag es nicht in allen Anwendungen sinnvoll erscheinen, das Fehlen eines Artikels oder Bindeworts als hinreichende Bedingung für den Ausschluss aus den Suchergebnissen heranzuziehen.

Mit entsprechendem Aufwand können Wörter, die nicht maßgeblich an der Bedeutung einer textuellen Inhaltsbeschreibung beteiligt sind, aus dem Suchkriterium ausgefiltert werden. Soll diese Filterung über das Aussortieren einer Liste von bereits bekannten Worten hinausgehen, wird die Miteinbeziehung der Bedeutung in der Verarbeitung unerlässlich.

Für das Systemkonzept wurde resümierend aus den soeben wiedergegebenen Überlegungen zur Textsuche die Schlussfolgerung gezogen, dass für eine wesentliche Verbesserung der Suche die Grenzen zeichenbasierter Textvergleiche überschritten werden müssen. Zur Erreichung dieses Ziels wird auf die Miteinbeziehung der Semantik von Begriffen bei der Suche gesetzt. Die nächsten Abschnitte beschreiben die Idee für die Umsetzung eines Dokumentenmanagementsystems unter diesen Voraussetzungen.

5.1.1 Vereinheitlichung individueller Beschreibungen

Die verschiedenen Ausdrucksmöglichkeiten bei der Verwendung von natürlicher Sprache und die hinzukommenden unterschiedlichen Betrachtungsweisen der Benutzer wirken einer Vereinheitlichung von inhaltlichen Beschreibungstexten entgegen. In einem Satzteil können mehrere miteinander verknüpfte und relevante Begriffe vorkommen, die eventuell auch aus mehreren Wörtern zusammengesetzt sein können. Eine verlässliche maschinelle Extraktion der unterschiedlichen Begriffe wird unmöglich, wenn die Maschine die natürliche Sprache nicht in ihrer Bedeutung verarbeiten kann.

Die Freiheiten bei der textuellen Beschreibung führen zu komplexen Zusammenhängen zwischen den verwendeten Begriffen. Diese Komplexität spricht gegen eine für den Benutzer verständliche Umsetzung eines Dokumentenmanagementsystems.

Zur Vereinfachung bietet sich die Durchführung der Inhaltsbeschreibung mit einer Auflistung von Schlagworten an. Dieses Konzept hat sich im Web 2.0 bereits bewährt und kommt häufig zum Einsatz. Der Gebrauch von Schlagworten stellt zwar keine Garantie für die Vereinheitlichung der Inhaltstexte dar, schafft aber eine bessere Voraussetzung durch die Reduktion der Komplexität der Sprache.

5.1.2 Unterstützung von Homonymen und Synonymen

Wenn auch nicht immer eine genaue Übereinstimmung von Zeichenfolgen erforderlich ist, haben textvergleichende Algorithmen bei der Qualität der Suchergebnisse ihre Grenzen. Semantische Relationen zwischen Begriffen mit unterschiedlicher Schreibweise werden nicht erkannt und folglich nicht im Ergebnis berücksichtigt. Durch die Unterstützung von Homonymen, Synonymen und Abkürzungen im Dokumentenmanagementsystem werden neue Möglichkeiten zur Assistenz des Benutzers erschlossen. Verwendet der Anwender bei der Eingabe ein Homonym, kann das System nach der gewünschten Bedeutung fragen. Des Weiteren kann die Zuordnung von alternativen Schreibweisen zum jeweiligen Begriff automatisch vom System vorgenommen werden.

In der Verwendung bei der Suche bedeutet dies für den Benutzer, dass auch Abkürzungen und Synonyme anstatt des eigentlichen Begriffs angegeben werden können, um zum selben Ergebnis zu gelangen.

5.1.3 Folksonomy mit semantischen Erweiterungen

Für die Akzeptanz der Applikation ist wichtig, dass die Anwender während der täglichen Verwendung des Systems ihre Freiheiten bei der Beschlagwortung behalten. Der Einsatz einer zentral kontrollierten Taxonomie von Begriffen kommt aufgrund der geforderten Dynamik nicht in Frage. Diese Entscheidung wird gestützt durch den von der Projektleitung von „TISS“ gestellten Anspruch, kompatibel mit bestehenden Arbeitsweisen zu bleiben. Es wird auf die kollektive Intelligenz der Anwender vertraut, anstatt

sie durch das System einzuschränken. In [MoSa06] wird der Erfolg von Web 2.0 durch die Nutzung der kollektiven Intelligenz begründet. Als logische Konsequenz wurde der Schwerpunkt im Konzept darauf gesetzt, wie man Anwender bei ihrer Arbeit unterstützen kann.

Die Miteinbeziehung von Semantik in die Verarbeitung ermöglicht dem System die Unterstützung des Anwenders mit inhaltlichen Vorschlägen. Es wird die Annahme getroffen, dass Anwender bei der Kategorisierung eher auf vom System vorgeschlagene Schlagwörter zurückgreifen, als Neueingaben zu tätigen. Diese Annahme wird begründet durch die Tatsache, dass Neueingaben gegenüber der Auswahl existierender Schlagwörter mehr Aufwand darstellen.

Ausgehend von dieser Annahme wirkt sich die Qualität der vom System gelieferten semantischen Vorschläge maßgeblich auf die Vereinheitlichung der verwendeten Schlagwörter aus. Als grundlegendes Designelement kommen Vorschläge nicht nur bei der Kategorisierung sondern auch bei der Navigation durch den Datenbestand zum Einsatz.

Zur Bereitstellung qualitativer Vorschläge für die Kategorisierung wurde neben der Behandlung von Homonymen und Synonymen in vergleichbarer Weise zu [MaCa10] das sprachliche Konzept eines Überbegriffs eingeführt. Ermöglicht durch die Grenzen der Anwendungsdomäne, kann ein Metamodell auf Basis der Relationen zwischen den verwendeten Überbegriffen erstellt werden. Für einen Überbegriff wird im weiteren Text auch der linguistische Fachbegriff „Hyperonym“ verwendet.

Bei der Konstruktion dieses Metamodells¹² wurde die Entscheidung getroffen, Verbindungen zwischen Überbegriffen nur mit einer einzigen semantischen Relationsart abzubilden. Diese gewählte Art der Relation drückt aus, dass ein Überbegriff dem anderen untergeordnet wird oder mit anderen Worten ausgedrückt, ein Teil des anderen Überbegriffs ist. In der Linguistik wird diese Relation, die etwas als einen Teil von etwas anderen ausweist, als Meronymie bezeichnet. Worte, die zu einem anderen Wort in dieser Relation stehen, werden Meronyme genannt.

Der Fokus im Metamodell liegt primär auf der Aussage über die Existenz einer Relation zwischen zwei Überbegriffen anstatt der detaillierten Beschreibung der Relationsart. Für den Anwender bedeutet dies, dass ihm die Interpretation der Art der Relation zwischen seinem eingegebenen Schlagwort und dem vom System vorgeschlagenen relevanten Schlagwörtern überlassen bleibt. Es wird auf die Assoziationen der Anwender zu den gezeigten Kombinationen von unterschiedlichen Schlagwörtern vertraut.

¹² siehe Metaebene 1 in Abbildung 11 und Abbildung 12

Die vereinfachte, im Metamodell abgebildete Ontologie des Anwendungsbereichs ist die Basis für semantische Erweiterungen. Nach dem Vorbild des Metamodells wird ein Modell¹³ konstruiert, in dem Begriffe aus dem Anwendungsbereich, deren Hyperonyme im Metamodell vorkommen, in Relation gesetzt werden. Bei der Schlagwortsuche werden bei der Eingabe einer Schlagwortkombination nicht nur Ergebnisse gefunden, deren Schlagwörter exakt übereinstimmen, sondern auch Kombinationen deren Schlagwörter den gesuchten untergeordnet sind. Ein Schlagwort entspricht bei der Suche also einem semantischen Bedeutungsraum, der die im System vorhandenen semantisch relevanten Schlagworte umfasst.

5.1.4 Vorschläge bei Navigation und Schlagwortvergabe

Bei jedem Suchergebnis wird der Benutzer durch Navigationsvorschläge unterstützt. Die Navigationsvorschläge beinhalten neben den Schlagwörtern, die zusammen mit den Schlagwörtern des Suchkriteriums vergeben wurden, aus dem Anwendungswissen abgeleitete Vorschläge zu weiteren Einschränkung der Ergebnisse. Zur Erstellung dieser Vorschläge werden die Meronyme der eingegebenen Schlagwörter herangezogen. Ein Meronym bezeichnet einen Teil des von einem anderen Wort bezeichneten Ganzen. Das schrittweise Hinzufügen weiterer Meronyme ermöglicht dem Anwender die gestellte Suchanfrage auf einer semantischen Basis bis zum gewünschten Resultat anzupassen.

Vorschläge werden auch bei der Schlagworteingabe und Beschlagwortung von Akten eingesetzt. Bei der Eingabe eines neuen Schlagworts werden dem Benutzer nach Auswahl eines Hyperonyms vom System Meronyme vorgeschlagen, mit denen das eingegebene Schlagwort in Bezug gesetzt werden kann. Über das Metamodell wird in Erfahrung gebracht, welche Schlagwörter für eine Relation in Frage kommen.

Zur Beschlagwortung einer Akte schlägt das System einzelne Schlagwortkombinationen, beginnend bei der häufigsten, sortiert nach der Anzahl des Auftretens, vor. Editiert der Benutzer die Liste der vergebenen Schlagworte, werden die empfohlenen Kombinationen vom System auf die Eingabe angepasst. Anhand der vorgeschlagenen Kombinationen wird dem Benutzer vermittelt, unter welchen Kombinationen bereits Akten abgelegt wurden. So hilft das System dem Benutzer Akten unter dem gewünschten Ablagebereich einzuordnen. Zusammengefasst wird der Anwender durch die Verwertung des abgebildeten Anwendungswissens bei allen Tätigkeiten im System durch Vorschläge unterstützt.

Für die Strukturierung der Anzeige von Hyperonymen werden drei semantische Achsen herangezogen. Die Unterteilung erfolgt nach den Kategorien Zeit, Ort und Inhalt. In der Navigation werden die vorgeschlagenen Schlagwörter nach ihren Hyperonymen und

¹³ siehe Metaebene 0 in Abbildung 11 und Abbildung 12

deren semantischen Achsen gruppiert angezeigt. Je nach der Anzahl der mit dem Suchkriterium zusammenhängenden Schlagwörter wird eine passende Anzeigestrategie für die Vorschläge gewählt. Bei weniger als sechs Einträgen werden die Schlagwörter hintereinander aufgelistet. Ein Kombinationslistenfeld wird für die Anzeige von bis zu hundert Schlagwörtern verwendet. Mehr als hundert zusammenhängende Schlagwörter werden durch ein Textfeld mit Eingabevervollständigung und einer Liste der fünf häufigsten mit dem Suchkriterium auftretenden Schlagwörter repräsentiert.

5.2 Software-Architektur

Dieser Abschnitt widmet sich mit seinen Unterpunkten der Dokumentation der Software-Architektur. Die Gliederung dieses Abschnitts folgt dem in [Kruc95] beschriebenen 4+1 Sichtenmodell.

Diesem Modell nach wird eine Software-Architektur in ihrer Dokumentation aus vier grundlegenden Standpunkten beschrieben. Diese Punkte sind ein Teil der folgenden Aufzählung. Die englischen Originalbezeichnungen aus der zitierten Referenz wurden in Klammern hinter den übersetzten Bezeichnungen angeführt, um eine eindeutige Zuordnung sicherzustellen.

- *Logische Sicht (Logical View)*
Die logische Sicht stellt die Abstraktion des Systems in Objekte oder Klassen dar. Die Funktionalität des Systems für den Endbenutzer bildet den Gegenstand der Betrachtung.
- *Prozesssicht (Process View)*
Die Prozesssicht beschreibt die Interaktionen zwischen simultan ablaufenden Prozessen im System.
- *Entwicklungssicht (Development View)*
Die Unterteilung des Systems in Komponenten mit definierten Schnittstellen ist Thema der Entwicklungssicht.
- *Physische Sicht (Physical View)*
Die physische Sicht gibt Aufschluss über die Verteilung des Systems auf die einzusetzende Hardware.

Die Anwendungsszenarien des Systems bilden einen Ausgangspunkt für die Konzeption der vier angeführten Sichten und stellen die zusätzliche Sichtweise dar. Anhand der Szenarien kann das Zusammenspiel der vier Sichten zur Überprüfung der Architektur veranschaulicht und getestet werden.

Tabelle 7 beinhaltet die Bedeutungen der englischen Begriffe, die in den Diagrammen und Texten dieses Abschnitts verwendet wurden.

Begriff	Bedeutung
Document	Dokument
Note	Notiz
Permission	Befugnis
Record	Akte, Schriftstück
Record Action	Bearbeitungsschritt
Resource	Ressource (Akte, Schlagwort)
Role	Rolle
Search	Suche
Tag	Schlagwort
User	Benutzer

Tabelle 7: Bedeutung der in Diagrammen verwendeten Begriffe

5.2.1 Szenarien-Sicht

Dieser Abschnitt dokumentiert die Einsatzszenarien des geplanten Systems. Begonnen wird mit der Darstellung der Abhängigkeit der Anwendungsfälle von den Anforderungen in den nächsten drei Tabellen.

Eine eingetragene Abhängigkeit bedeutet, dass die Implementierung des betroffenen Anwendungsfalles die jeweilige Anforderung oder zumindest einen Teil davon abdecken wird. Nach den Tabellen folgen Anwendungsfalldiagramme zur grafischen Illustration der Systemfunktion. Die textuelle Beschreibung der durch die Anwendungsfälle repräsentierten Funktionen vervollständigt die Dokumentation.

Zur Kategorisierung der Anwendungsfälle kamen wiederholt die im Abschnitt 4.3.1 angeführten Anwendungsbereiche zum Einsatz. Die Vergabe der eindeutigen Bezeichner erfolgte nach der in Tabelle 6 angewandten Vorgehensweise. Einige Anwendungsfälle ließen sich nicht eindeutig einem der bestehenden Bereiche zuordnen, weil mehr als eine Kategorie zutrifft. Für diese Fälle wurden diese beiden zusätzlichen Kategorien eingeführt:

- *Schlagworte (TAG)*
Anwendungsfälle, die den Umgang mit Schlagwörtern bei der Suche oder Beschreibung von Akten betreffen, fallen in diesen Bereich. Die Verwaltung von Schlagworten ist ebenfalls inkludiert.
- *Empfehlungen (SUG)*
Wenn dem Benutzer ermöglicht wird, bei seinen Eingaben auf Empfehlungen des Systems zurückzugreifen, werden die Anwendungsfälle unter diesem Bereich eingeordnet.

Die soeben angeführten Bereiche sind inhaltlich als ein Teil der beiden bestehenden Kategorien „Dokumentenverwaltung“ und „Suche“ zu sehen.

Tabelle 8, die erste der drei Tabellen, zeigt die Verbindungen zwischen den für beide Organisationseinheiten geltenden Anforderungen und den Anwendungsfällen.

		Anforderungen												
		DOC.1	DOC.2	DOC.3	DOC.4	DOC.5	DOC.6	DOC.7	DOC.8	SEA.1	SEA.2	SEA.3	SEA.4	SEA.5
		Dokumente in Papier und elektronischer Form	Inhaltstexte als Schlagwortlisten	Einheitliche Inhaltstexte für Schriftstücke	Entwürfe von Schriftstückeinträgen	Automatische Vergabe von Laufnummern	Entwurf von Schriftstückbeschreibungen	Protokolle für Suche und Einsichtnahme	Stornierung von Schriftstücken	Inhaltliche Vernetzung	Schrittweises Verfeinern von Kriterien	Suche nach kombinierten Kriterien	Unbeschränkte Anzahl an Suchergebnissen	Aufgelöste Themenkreise ausblenden
Anwendungsfälle	DOC.1	Akt erfassen	x	x	x			x						
	DOC.2	Akt als Entwurf erfassen	x	x		x	x							
	DOC.3	Aktentwurf freigeben	x		x									
	DOC.4	Akt bearbeiten	x	x	x	x		x						
	DOC.5	Akt stornieren	x			x			x					
	DOC.6	Ablagebereich wählen	x				x							
	DOC.7	Dokumente verwalten	x											
	DOC.8	Dokumentversionen verwalten	x											
	DOC.9	Dokument hochladen	x											
	DOC.10	Dokument runterladen	x											
	SEA.1	Akt suchen		x					x		x	x	x	x
	TAG.1	Schlagwortliste verwalten		x	x			x			x	x		
	TAG.2	Schlagwort hinzufügen		x	x			x			x	x		
	TAG.3	Schlagwort entfernen		x				x			x			
	TAG.4	Schlagwörter verwalten		x							x			
	TAG.5	Schlagwort erfassen		x							x			
	TAG.6	Schlagwort bearbeiten		x							x			
	TAG.7	Titel definieren		x										
	TAG.8	Bezeichner definieren		x										
	TAG.9	Überbegriff wählen		x							x			
TAG.10	Semantische Relation setzen		x							x				
TAG.11	Synonyme definieren		x							x				
SUG.1	Empfehlung wählen		x	x						x				

Tabelle 8: Allgemeine Anforderungen und Anwendungsfälle

Nach den allgemeinen Anforderungen beinhaltet Tabelle 9 die speziellen Anforderungen der Kanzlei.

			Anforderungen									
			DOC.K.1	DOC.K.2	DOC.K.3	DOC.K.4	SEA.K.1	PRJ.K.1	PRJ.K.2	REP.K.1	REP.K.2	INT.K.1
			Beibehaltung bestehender Nummerierungen	Aufbewahrungsort von Schriftstücken	Benachrichtigungen bei Fristen	Referenzen auf Schriftstücke	Suche nach Bearbeitungsschritten	Komfortables Ausdrucken	Umschläge für Schriftstücke	Rückstandsausweise	Ausgeschiedenes Personal	Projektdatenbank
Anwendungsfälle	DOC.1	Akt erfassen	x									x
	DOC.2	Akt als Entwurf erfassen	x									
	DOC.4	Akt bearbeiten		x	x		x					
	DOC.K.1	Akten dem Archiv zuteilen		x								
	DOC.K.2	Bearbeitungsschritte verwalten			x				x			
	DOC.K.3	Bearbeitungsschritt erfassen			x				x			
	DOC.K.4	Bearbeitungsschritt bearbeiten			x				x			
	DOC.K.5	Frist setzen			x				x			
	DOC.K.6	Frist als erfüllt erklären			X				x			
	SEA.1	Akt suchen		x		x	x	x				
	TAG.5	Schlagwort erfassen										x
	REP.K.1	Bericht erstellen			x			x	x	x	x	
	REP.K.2	Laufende Fristen anzeigen			x					x		

Tabelle 9: Anforderungen der Kanzlei und Anwendungsfälle

Um die Dokumentation der Abhängigkeiten abzuschließen, enthält Tabelle 10 die Anforderungen der „Gebäude und Technik“.

			Anforderungen		
			PER.G.1	PER.G.2	INT.G.1
			Rollenbasiertes Berechtigungssystem	Aktionsbasierte Berechtigungsgranularität	SAP
Anwendungsfälle	DOC.1	Akt erfassen			x
	TAG.5	Schlagwort erfassen			x
	PER.G.1	Rollen verwalten	x	x	
	PER.G.2	Rollen an Personen vergeben	x		
	PER.G.3	Berechtigungen an Rollen vergeben		x	

Tabelle 10: Anforderungen der „Gebäude und Technik“ und Anwendungsfälle

Abbildung 7 beschäftigt sich mit Anwendungsfällen für Schlagwörter im Detail.

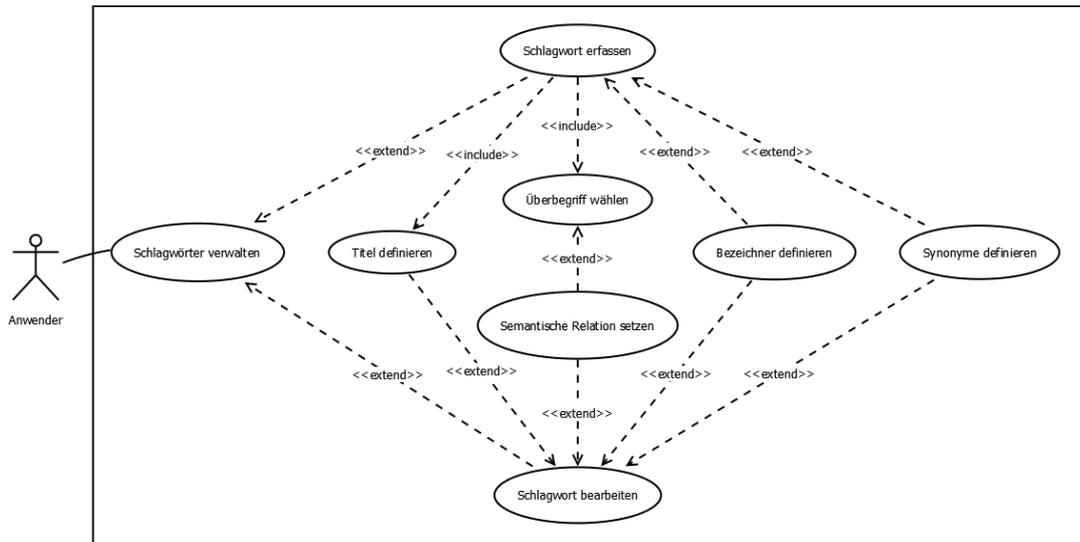


Abbildung 7: Anwendungsfalldiagramm für Schlagwörter

In Abbildung 8 sind auf die Kanzlei beschränkte Anwendungsfälle enthalten.

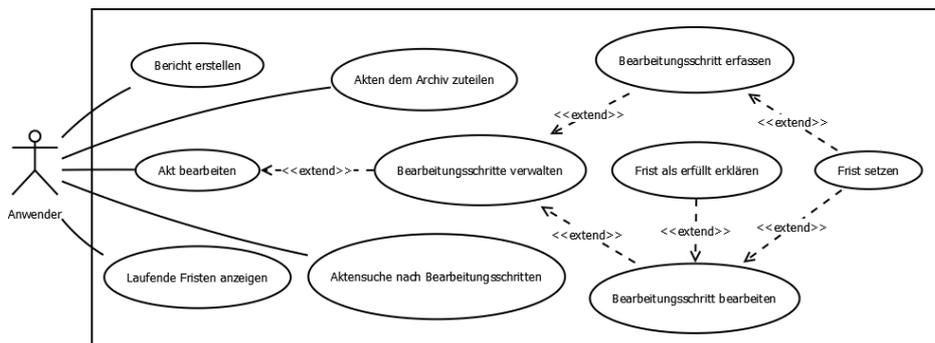


Abbildung 8: Anwendungsfalldiagramm der Kanzlei

Die exklusiven Anwendungsfälle der „Gebäude und Technik“ sind in Abbildung 9 vertreten.

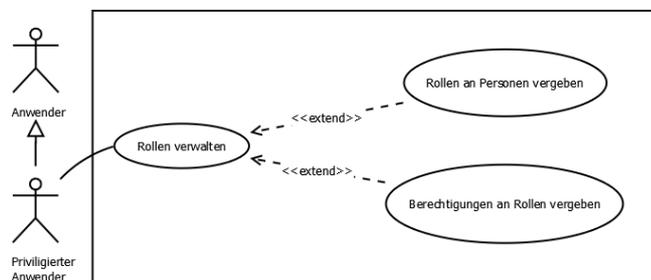


Abbildung 9: Anwendungsfalldiagramm der „Gebäude und Technik“

Es folgt eine Liste mit Beschreibungen der gemeinsamen Anwendungsfälle von Kanzlei und „Gebäude und Technik“.

- *Dokumentenverwaltung*

1. *Akt erfassen*

Zur Erfassung eines Aktes weist der Benutzer dem neuen Schriftstück ein oder mehrere Schlagworte zu und wählt einen Ablagebereich. Bei diesen Dateneingaben können Empfehlungen des Systems gewählt werden. Zu Beginn werden jeweils drei Empfehlungen für Schlagwortvergabe und Ablagebereich vorgeschlagen. Die Anzahl der Empfehlungen kann schrittweise um drei weitere erhöht werden. Der gewählte Ablagebereich wird vom System zur Generierung des eindeutigen Bezeichners für den Akt verwendet. Aus diesem Grund lässt sich der Ablagebereich nach der Erfassung nicht mehr modifizieren. Vor der Bestätigung der Erfassung kann eine beliebige Anzahl an Dokumenten für das Hochladen ausgewählt werden. Sind unter den zugeteilten Schlagworten beim Abschluss der Erfassung dem System unbekannt Elemente enthalten, werden diese automatisch erfasst.

2. *Akt als Entwurf erfassen*

Bei der Erfassung eines Aktes als Entwurf muss zwingend ein Ablagebereich gewählt werden, um dem System die Generation eines eindeutigen Bezeichners zur Identifikation zu ermöglichen. Die aus dem Anwendungsfall „Akt erfassen“ übrigen Schritte sind wahlweise durchführbar. Entwürfe werden bei der Suche von Akten nicht aufgefunden.

3. *Aktentwurf freigeben*

Der Benutzer kann einen Entwurf freigeben, vorausgesetzt der Entwurf erfüllt die Kriterien zur Erfassung eines Aktes. Die Freigabe eines Entwurfs ist inhaltlich mit der Erfassung eines neuen Aktes gleichzusetzen.

4. *Akt bearbeiten*

Um zur Bearbeitung eines Aktes zu gelangen, muss der Akt in den Suchergebnissen gewählt werden. Zur Auswahl eines Entwurfs für die Bearbeitung existiert eine von den Suchergebnissen getrennte Ansicht. Für Entwürfe ist die Bearbeitung auf Änderungen der vergebenen Schlagworte und das Hinzufügen von angehängten Dateien beschränkt. Für bereits vollständig erfasste Akten können bei der Bearbeitung weitere Eigenschaften editiert werden. Zu diesen Eigenschaften zählen Referenzen auf vorhergehende beziehungsweise nachfolgende Akten, die Art der im Akt enthaltenen Dokumente und der Aufbewahrungsort. In der Kanzlei kann der Benutzer Bearbeitungsschritte für den Akt eingeben. Zusätzlich zum Hochladen neuer Dokumente können auch neue Versionen bereits existierender Dokumente angelegt werden.

5. *Akt stornieren*

Bei der Stornierung vom Benutzer wird ein Akt als storniert markiert. Diese Markierung erfolgt im System per Vergabe des Schlagwortes „storniert“. Wird ein Aktentwurf storniert, verliert dieser seinen Entwurfsstatus.

6. *Ablagebereich wählen*

Die Wahl eines Ablagebereichs durch den Benutzer erfolgt von oben nach unten in der Themenhierarchie. Themenkategorie, Themenbereich und Unterthema bilden die drei Ebenen der Hierarchie. Die Auswahl eines Elements macht die untergeordneten Elemente der nächsten Ebene selektierbar. In diesem Punkt gibt es Unterschiede zwischen Kanzlei und „Gebäude und Technik“. In der Kanzlei ist eine Auswahl der Themenkategorie auf oberster Ebene optional, um zu bestehenden Ablagemethoden kompatibel zu bleiben. Aus diesem Grund kann direkt bei der Wahl eines Themenbereichs, die verpflichtend für eine gültige Eingabe ist, begonnen werden. Ob ein Unterthema auszuwählen ist, bleibt ebenfalls dem Benutzer überlassen. Ein für die „Gebäude und Technik“ gültiger Ablagebereich erfordert hingegen immer eine Themenkategorie und einen Themenbereich.

7. *Dokumente verwalten*

Der Benutzer kann mit der Verwaltung von Dokumenten assoziierte Aufgaben erledigen. Darunter wird primär das Hoch- und Herunterladen von Dokumenten verstanden. Zudem wird dem Benutzer das Wissen bereitgestellt, von wem welche Datei wann hochgeladen wurde. Initial werden die drei zuletzt hochgeladenen Dokumente sortiert nach absteigender Aktualität aufgelistet. Auf Wunsch des Benutzers wird die Anzeige wahlweise um drei weitere oder alle angehängten Dokumente erweitert.

8. *Dokumentversionen verwalten*

Dieser Anwendungsfall ist eine Spezialisierung des gerade eben beschriebenen Anwendungsfalls „Dokumente verwalten“. Die Funktionen decken sich mit der Ausnahme, dass die Versionen eines Dokuments den Gegenstand der Betrachtung liefern.

9. *Dokument hochladen*

Zu einem Akt können vom Benutzer eine beliebige Anzahl an Dokumenten hochgeladen werden. Die Auswahl von Dateien erfolgt über Dateieingabefelder im Formular. Beim Speichern des Akts findet der Transfer der Dateien statt.

10. *Dokument herunterladen*

Durch einen Klick auf ein in der Auflistung der Dokumente angezeigtes Element startet der Benutzer das Herunterladen der Datei.

- *Suche*

1. *Akt suchen*

Die Suche nach Akten räumt dem Benutzer die Möglichkeit ein, drei verschiedene Arten von Kriterien einzusetzen. Die erste Methode zur Suche ist die Angabe einer Liste von Schlagwörtern. Bei übereinstimmender Kombination von Schlagwörtern mit dem Kriterium scheint der Akt in der Ergebnisliste auf. Bei der Ermittlung der Übereinstimmung kommen die im Abschnitt 5.1 behandelten Prinzipien zum Einsatz. Dies bedeutet, dass Überbegriffe und Meronyme in aus der natürlichen Sprache gewohnter Weise verwendet werden. Die Verwendung eines Überbegriffs im Kriterium führt zur Übereinstimmung mit allen Schlagwörtern, die diesen Überbegriff haben. Ein gesuchtes Schlagwort stimmt auch mit seinen Meronymen überein. Zweitens besteht die Möglichkeit bei der Eingabe eines Textfragments, das keinem dem System bekannten Schlagwort entspricht, mit einer Volltextsuche die Beschlagwortungen der Akte nach einer Übereinstimmung zu durchsuchen. Als dritte und letzte Art ein Kriterium anzugeben, kann ein Teil der Aktennummer zur Einschränkung der Ergebnisse verwendet werden. Es wird auch eine Funktion bereitgestellt, um bei einem über die Aktennummer exklusiven Zugriff direkt zur Bearbeitungsansicht eines Aktes zu gelangen. Die drei beschriebenen Arten von Kriterien können bei der Suche miteinander kombiniert werden. Suchergebnisse werden standardmäßig absteigend nach Erfassungsdatum sortiert und mit zehn Akten pro Seite angezeigt. Der Benutzer kann auf die Sortierung und angezeigte Anzahl pro Seite Einfluss nehmen. Für einen Akt in der Ergebnisliste werden Aktennummer, Schlagwörter, Erfassungszeitpunkt und Erfasser angezeigt. Per Klick auf eine Aktennummer gelangt der Anwender zur Bearbeitungsseite des korrespondierenden Aktes. Jedes Suchresultat wird dem Benutzer zusammen mit einer Navigation zur Verfeinerung der Suchanfrage präsentiert. Ein Klick auf eines der in den Ergebnissen oder Navigation aufgelisteten Schlagwörter, fügt es zu den Suchkriterien hinzu.

- *Schlagworte*

1. *Schlagwortliste verwalten*

Die Verwaltung einer Liste von Schlagwörtern durch den Anwender stellt einen mehrfach eingesetzten Funktionsbaustein der Software dar und kommt bei der Suche und der Vergabe von Schlagwörtern an einen Akt zum Einsatz. Das Sys-

tem schlägt für die Eingabe des Benutzers maximal zehn passende Schlagworte zur Vervollständigung vor. Jedes Schlagwort wird zusammen mit seinem Überbegriff und falls vorhanden dem Bezeichner präsentiert, um eine bessere Unterscheidung zu ermöglichen. Der Bezeichner ist eine vom Benutzer frei zuordenbare Zeichenkette zur besseren Identifikation. Der Benutzer kann sich jederzeit per Mausklick für einen der Vorschläge entscheiden.

2. *Schlagwort hinzufügen*

Die Bestätigung eines gewählten Schlagworts per Eingabetaste oder Klick fügt es in die Liste ein.

3. *Schlagwort entfernen*

Jedes in der Liste befindliche Schlagwort kann per Klick wieder entfernt werden.

4. *Schlagwörter verwalten*

Um ein bestehendes Schlagwort bearbeiten zu können, muss es vorher ausgewählt werden. Die Auswahl wird über die bei der Schlagwortverwaltung bereits beschriebene Eingabevervollständigung vorgenommen. Auf diesem Weg sucht der Benutzer nach einem existierenden Schlagwort, bevor er sich entschließt ein neues zu erstellen.

5. *Schlagwort erfassen*

Der Benutzer kann ein neues Schlagwort zum Datenbestand des Systems hinzufügen. Als Eingaben sind für die Erfassung Titel und ein Überbegriff zwingend notwendig. Weitere optionale Eingaben sind Bezeichner und alternative Titel.

6. *Schlagwort bearbeiten*

Die Bearbeitung von Schlagwörtern unterliegt gegenüber der Erfassung einer zusätzlichen Restriktion. Der Überbegriff und die semantische Relationen von in Verwendung stehenden Schlagwörtern sind nicht mehr zu editieren.

7. *Titel definieren*

Die Angabe eines Titels ist für der Erfassung oder Bearbeitung eines Schlagwortes unerlässlich. Der Benutzer füllt den Titel per Textfeld im Formular aus.

8. *Bezeichner definieren*

Unter den Schlagworten mit einem gemeinsamen Überbegriff muss der Bezeichner eindeutig sein. Die Angabe des Bezeichners erfolgt ebenfalls per Textfeld im Formular.

9. *Überbegriff wählen*

Im System vorhandene Überbegriffe können mittels Kombinationsfeld gewählt werden. Andererseits kann der Benutzer aber auch einen eigenen Überbegriff anführen. Die Wahl eines bestehenden Überbegriffs öffnet dem Anwender in Folge die Möglichkeit eine semantische Relation anzugeben.

10. *Semantische Relation setzen*

Das System leitet aus dem abgebildeten Anwendungswissen und dem gewählten Überbegriff für eine semantische Relation in Frage kommende Schlagwörter ab. Der Benutzer setzt die semantische Relation, indem er eines dieser Schlagwörter auswählt.

11. *Synonyme definieren*

Vom Benutzer können Synonyme in Form von alternativen Titeln in beliebiger Anzahl eingegeben werden. Zur Bereitstellung dieser Funktion wird auf Wunsch des Anwenders ein zusätzliches Textfeld für einen weiteren Eintrag angezeigt.

- *Empfehlungen*

1. *Empfehlung wählen*

Empfehlungen unterstützen in drei verschiedenen Arten den Benutzer bei der Dateneingabe. Bei der Erfassung eines Aktes werden anfänglich die jeweils drei am häufigsten vergebenen Schlagwortkombinationen und Ablagebereiche für die Auswahl angezeigt. Vorgeschlagene Schlagwortkombinationen können bei einer späteren Bearbeitung eines Aktes wieder abgerufen werden. Mit einem einzelnen Klick können Schlagwortkombinationen als Ganzes in die Schlagwortliste übernommen werden. Aber auch das Hinzufügen eines einzelnen Schlagwortes aus den Vorschlägen ist möglich. Bei Änderung der Schlagwortliste oder des Ablagebereichs aktualisieren sich auch die Listen mit den Empfehlungen. Die in der Navigation gezeigten Schlagwörter zur Einschränkung der Suche stellen die dritte Art von Empfehlungen dar. In der Navigation ausgewählte Schlagwörter werden zum Suchkriterium hinzugefügt.

Folgende Anwendungsfälle werden exklusiv der Kanzlei zugeordnet:

- *Dokumentenverwaltung (Kanzlei)*

1. *Akten dem Archiv zuteilen*

Nach dem Ablauf der Aufbewahrungsfrist eines Aktes in der Kanzlei wird dieser an das Archiv übergeben. Diese Übergabe wird im System durch die Ak-

tualisierung des Aufbewahrungsorts vollzogen. Hierzu ist vom Benutzer die Menge von zu archivierenden Akten per Angabe eines Zeitraums auszuwählen. Für Einzelfälle kann die Übergabe auch manuell durch Umsetzen des Aufbewahrungsorts eines Aktes durchgeführt werden.

2. *Bearbeitungsschritte verwalten*

Der Benutzer kann bei der Bearbeitung eines Akts mehrere neue Bearbeitungsschritte erfassen und die bestehenden bearbeiten oder entfernen. Bevor das System die Eingabefelder für eine neue Eingabe zur Verfügung stellt, muss der Benutzer die Art des Bearbeitungsschritts in einem Kombinationsfeld wählen. Zwecks Übersichtlichkeit wird nochmals die Strategie aus dem Anwendungsfall „Dokumente verwalten“ zu Anzeige der Bearbeitungsschritte angewandt.

3. *Bearbeitungsschritt erfassen*

Bei der Eingabe eines neuen Bearbeitungsschritts protokolliert das System selbstständig den Erfassungszeitpunkt. Die Arten und Attribute der vier Bearbeitungsschritte „Erstattung“, „Zuteilung“, „Expedition“ und „Ablage“ sind aus Abschnitt 4.1.1 zu entnehmen.

4. *Bearbeitungsschritt bearbeiten*

Der Benutzer kann die für einen Bearbeitungsschritt eingetragenen Daten nachträglich ohne Restriktionen ändern.

5. *Frist setzen*

Durch die Eingabe eines Datums an dem die Frist zu Ende geht, verknüpft der Benutzer eine zeitliche Bedingung mit dem Bearbeitungsschritt.

6. *Frist als erfüllt erklären*

Der Benutzer kann den Status einer Frist für die Bearbeitungsschritte „Zuteilung“ und „Expedition“ per Kontrollkästchen auf den Zustand „erfüllt“ setzen.

- *Suche (Kanzlei)*

1. *Aktensuche nach Bearbeitungsschritten*

Die Benutzer in der Kanzlei haben die Möglichkeit Akten nach ihren Bearbeitungsschritten aufzufinden. In einer erweiterten Suche formuliert der Benutzer das Kriterium für einen gesuchten Bearbeitungsschritt. Als Ergebnis dieser Suche werden Akten aufgelistet, die übereinstimmende Bearbeitungsschritte beinhalten.

- *Berichte (Kanzlei)*

1. *Bericht erstellen*

In der Kanzlei können Berichte in den Ausgabeformaten HTML und PDF erstellt werden. Eine Zusammenfassung laufender Fristen und die Auflistung der Personalausritte in einem gewählten Zeitraum kann als Bericht erstellt werden.

2. *Laufende Fristen anzeigen*

Für einen einfachen Überblick kann der Benutzer sich eine Auflistung der momentan unerfüllten Fristen anzeigen lassen. Diese Liste beinhaltet eine Tabelle der betroffenen Akten inklusive dem Ende der Fristen. Der Benutzer kann in dieser Ansicht eine oder mehrere Fristen als erfüllt ausweisen.

Diese aufgezählten Anwendungsfälle sind nur in der „Gebäude und Technik“ relevant:

- *Befugnisse („Gebäude und Technik“)*

1. *Rollen verwalten*

Privilegierte Benutzer regeln die Autorisierung über die Verwaltung der Beziehungen zwischen Personen, Rollen und Berechtigungen. Die beschriebenen Anwendungsfälle sind im System als Berechtigungen abgebildet. Rollen können nach Belieben erstellt, bearbeitet oder entfernt werden.

2. *Rollen an Personen vergeben*

Eine Person erhält bei der Rollenvergabe alle mit der Rolle verknüpften Berechtigungen. Es können auch mehrere Rollen einer Person zugewiesen werden.

3. *Berechtigungen an Rollen vergeben*

Funktionen der Mitarbeiter werden im System durch Rollen abgebildet. Eine Funktion entspricht einer Menge von an eine Rolle vergebenen Berechtigungen.

Zum Abschluss listet Tabelle 11 alle in diesem Abschnitt vorgekommenen Anwendungsfälle, sortiert nach Bezeichner, auf.

Anwendungsfälle		Herkunft	
Bezeichner	Titel	Kanzlei	Gebäude und Technik
DOC	Dokumentenverwaltung		x
DOC.1	Akt erfassen		x
DOC.2	Akt als Entwurf erfassen		x
DOC.3	Aktentwurf freigeben		x
DOC.4	Akt bearbeiten		x
DOC.5	Akt stornieren		x
DOC.6	Ablagebereich wählen		x
DOC.7	Dokumente verwalten		x
DOC.8	Dokumentversionen verwalten		x
DOC.9	Dokument hochladen		x
DOC.10	Dokument herunterladen		x
DOC.K	Dokumentenverwaltung (Kanzlei)	x	
DOC.K.1	Akten dem Archiv zuteilen	x	
DOC.K.2	Bearbeitungsschritte verwalten	x	
DOC.K.3	Bearbeitungsschritte erfassen	x	
DOC.K.4	Bearbeitungsschritt bearbeiten	x	
DOC.K.5	Frist setzen	x	
DOC.K.6	Frist als erfüllt erklären	x	
SEA	Suche		x
SEA.1	Akt suchen		x
SEA.K	Suche (Kanzlei)	x	
SEA.K.1	Aktensuche nach Bearbeitungsschritten	x	
TAG	Schlagworte		x
TAG.1	Schlagwortliste verwalten		x
TAG.2	Schlagwort hinzufügen		x
TAG.3	Schlagwort entfernen		x
TAG.4	Schlagwörter verwalten		x
TAG.5	Schlagwort erfassen		x
TAG.6	Schlagwort bearbeiten		x
TAG.7	Titel definieren		x
TAG.8	Bezeichner definieren		x
TAG.9	Überbegriff wählen		x
TAG.10	Semantische Relation setzen		x
TAG.11	Synonyme definieren		x
SUG	Empfehlungen		x
SUG.1	Empfehlung wählen		x
REP.K	Berichte (Kanzlei)	x	
REP.K.1	Bericht erstellen	x	
REP.K.2	Laufende Fristen anzeigen	x	
PER.G	Befugnisse („Gebäude und Technik“)		x
PER.G.1	Rollen verwalten		x
PER.G.2	Rollen an Personen vergeben		x
PER.G.3	Berechtigungen an Rollen vergeben		x

Tabelle 11: Anwendungsfälle inklusive Herkunft nach Organisationseinheit

5.2.2 Logische Sicht

In der logischen Sicht wird die Funktionalität des Systems in Klassen und Objekte abstrahiert. Zur grafischen Illustration werden aus UML Klassen- und Objektdiagramme eingesetzt.

Tabelle 12 beinhaltet eine Übersicht über die in den Klassendiagrammen verwendeten Datentypen. Aufgrund der zwischen verschiedenen Datenbanksystemen bei Namen und Arten von Datentypen vorherrschenden Unterschiede, wurden eigene Datentypen definiert. Die Bedeutung der eigens gewählten Datentypen ist ebenfalls aus der Tabelle zu entnehmen.

Datentyp	Bedeutung
Integer	Ganzzahl
String	Zeichenkette mit begrenzter Länge
Text	Zeichenkette mit beliebiger Länge
Time	Datum mit Uhrzeit
Date	Datum
Binary	Binärdaten

Tabelle 12: In Klassendiagrammen verwendete Datentypen

Zur Ermöglichung der flexiblen Abbildung von semantischen Beziehungen wurde als Datenstruktur der im semantischen Web gebräuchliche Triplestore gewählt. Gegenüber einem konventionellen relationalem Datenbankschema hat diese Datenstruktur den Vorteil, dass sie um neue Entitäten und Relationen erweitert werden kann, ohne dass eine Anpassung des Datenbankschemas notwendig wird. In einer herkömmlichen Datenbank wird für gewöhnlich eine neue Tabelle für die Abbildung einer neuen Entität hinzugefügt. Im Gegenzug erhöht die Strukturierung von Informationen in einem Triplestore als Graph die Komplexität der notwendigen Datenbankabfragen.

Auf den Punkt gebracht, wird bei der Entwicklung eine erhöhte Komplexität für eine Steigerung der Flexibilität in Kauf genommen. Die im Projekt „TISS“ bereitgestellte Infrastruktur gibt die Verwendung einer relationalen Datenbank vor. Abbildung 10 zeigt das Klassendiagramm für die Realisierung des Triplestores innerhalb eines relationalen Datenbanksystems.

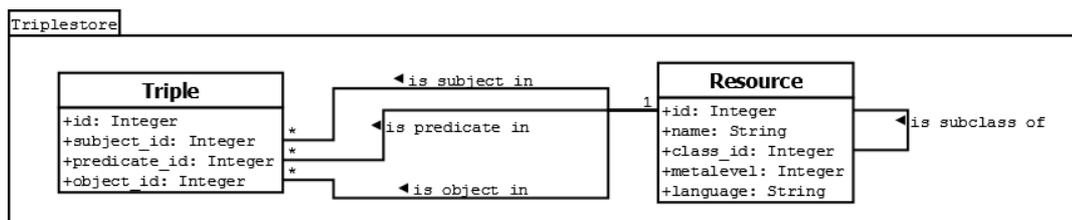


Abbildung 10: Klassendiagramm „Triplestore“

Die Tabelle „Triple“ stellt die Beziehungen zwischen den vorhandenen Datensätzen der Tabelle „Resource“ her. In der Tabelle „Resource“ werden Systemressourcen, Schriftstücke und Schlagworte samt Literale eingetragen.

Bei Schriftstücken und Schlagworten wird als Name der Ressource der relative Pfad verwendet, unter dem die Ressource per Browser aufgerufen werden kann. Bei Literalen ist die Zeichenkette selbst gleichzeitig der Name der Ressource. Titel, Bezeichner und Synonyme von Schlagworten werden als Literale gespeichert.

Jede Ressource wird einer von drei Metaebenen zugeteilt und hat eine andere Ressource als Klasse. Die nächste Aufzählung geht genauer auf die Kategorisierung in Metaebenen ein:

- Null
Bleibt der Wert für die Metaebene ungesetzt, handelt es sich um Ressourcen, die für den Betrieb des Systems notwendig sind. Zur Beschreibung anderer Ressourcen wurden die DCMI¹⁴ Metadata Terms in den Datenbestand aufgenommen.
- 1
Das Metamodell zum Wissen aus dem Anwendungsbereich wird auf der ersten Metaebene spezifiziert. Hierzu werden die Beziehungen zwischen den Überbegriffen der Schlagworte im Triplestore abgebildet. Technisch gesehen, sind Überbegriffe Schlagworte mit einer Zuordnung zu ersten Metaebene.
- 0
Auf der Metaebene 0 werden anhand des Metamodells konkrete Aussagen über die Schlagworte der Anwendungsdomäne getroffen. In ihrer Gesamtheit bilden diese Aussagen ein Modell des Anwendungsbereichs. Das Metamodell gibt an, welche Relationen zwischen Schlagworten gesetzt werden dürfen und dient zur Validierung des Modells.

Für die Beschreibung von Ressourcen wird nur eine Teilmenge der DCMI Metadata Terms verwendet. Bei der Auswahl der geeigneten Klassen und Eigenschaften wurde darauf geachtet, das Beschreibungsschema so einfach wie möglich zu halten.

Tabelle 13 zeigt welche Elemente der DCMI Metadata Terms eingesetzt werden:

¹⁴ Dublin Core Metadata Initiative, siehe <http://dublincore.org/>

Element	Verwendung
Klassen	
dcterms:Location	Ortsangaben
dcterms:PeriodOfTime	Zeitangaben
Eigenschaften	
dcterms:title	Titel eines Schlagwortes
dcterms:identifier	Optionaler Bezeichner eines Schlagwortes
dcterms:alternative	Optionales Synonym eines Schlagwortes
dcterms:hasPart	Meronyme eines Schlagwortes
dcterms:references	Referenzen auf ein Schlagwort
dcterms:subject	Zuordnung eines Schlagwortes zu einem Schriftstück

Tabelle 13: Verwendete Elemente der DCMI Metadata Terms

Innerhalb eines Tripels wird „`rdfs:subClassOf`“ als Prädikat verwendet, um eine Ressource als Orts- oder Zeitangabe auszuweisen. Die betroffene Ressource und die entsprechende Klasse aus Tabelle 13 bilden Subjekt und Objekt des Tripels. Die Einteilung in Orts- und Zeitangaben beschränkt sich exklusiv auf Überbegriffe und wird somit nur in der ersten Metaebene vorgenommen. Ob ein Schlagwort sich auf etwas Zeitliches oder Räumliches bezieht, wird daher aus der Zuteilung des Überbegriffs geschlossen. Die des Weiteren in der Tabelle angeführten Eigenschaften nehmen bei der Beschreibung von Ressourcen im Tripel die Stelle des Prädikats ein.

Die beiden nächsten Abbildungen zeigen den Zusammenhang zwischen Modell und Metamodell anhand eines Auszugs des abgebildeten Wissens aus dem Anwendungsbe-
reich. Aus Platzgründen wurden die inhaltliche, räumliche und zeitliche Ebene der Wissensabbildung auf Abbildung 11 und Abbildung 12 aufgeteilt. Tabelle 14 erklärt die in den Abbildungen verwendete Notation, welche auf der Notation für Klassen- und Objektdiagramme in UML basiert.

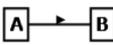
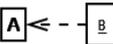
Notationselement	Bedeutung
	Schlagwortüberbegriff „A“
	Schlagwort „B“
	„B“ ist ein Meronym von „A“
	„B“ ist eine Instanz von „A“

Tabelle 14: Diagrammnotation für Wissensabbildung im Triplestore

Die inhaltliche Ebene der Wissensabbildung aus Abbildung 11 stellt im Wesentlichen ein Modell der Organisationsstruktur der Technischen Universität Wien dar:

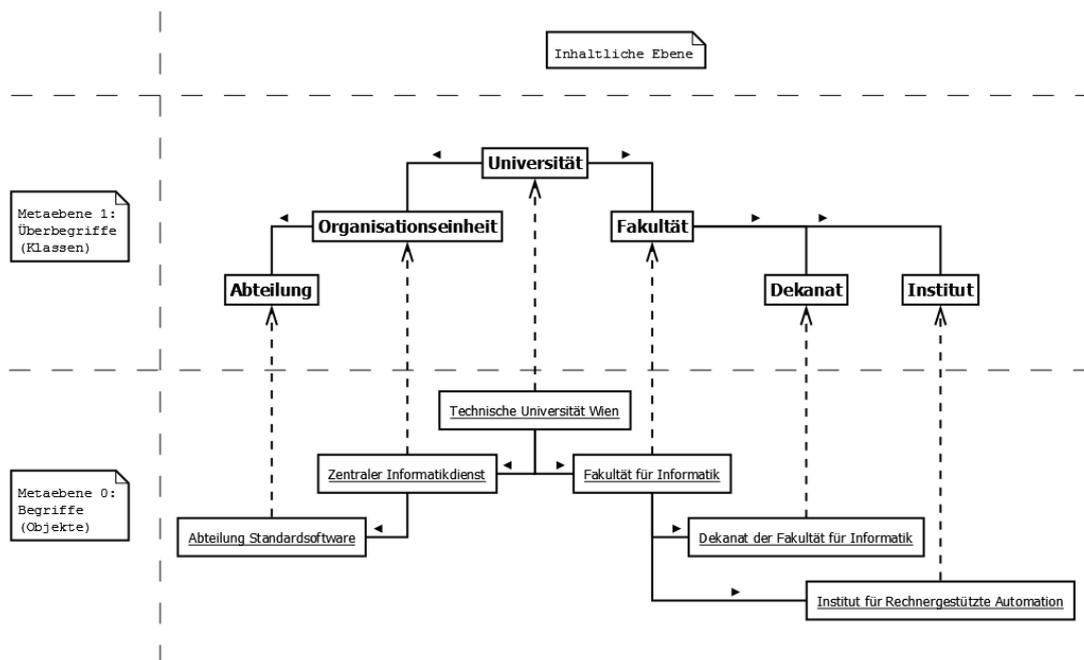


Abbildung 11: Inhaltliche Wissensabbildung im Triplestore

Auf der inhaltlichen Ebene existieren noch weitere Überbegriffe mit untergeordneten Schlagworten. Aufgrund nicht vorhandener Relationen wurde auf die triviale Abbildung dieser Elemente im vorigen Diagramm verzichtet:

- *Ablagebereiche (Kategorie, Thema, Unterthema)*
Schlagworte, die Schriftstücke in die dreistufige Hierarchie zur Ablage einordnen, sind in der Applikation unter dem Überbegriff „Ablagebereich“ bekannt. Die Ebene des Ablagebereichs wird ausgedrückt, indem eines der Worte „Kategorie“, „Thema“ und „Unterthema“ dem Überbegriff in Klammern beigefügt wird.
- *Aufbewahrungsort*
Die Zuweisung von Schriftstücken zu Organisationseinheiten, die für die Aufbewahrung zuständig sind, wird durch Schlagworte mit dem Überbegriff „Aufbewahrungsort“ vorgenommen.
- *Bearbeitungsschritt*
Unter diesem Überbegriff werden die Bearbeitungsschritte der Kanzlei in Form von Schlagworten im System eingetragen.
- *Dokumentenart*
Die im Anwendungsbereich vorkommenden Arten von Dokumenten werden ebenfalls als Schlagworte im System abgelegt. Die Schlagworte tragen den Überbegriff „Dokumentenart“.

- *Inhaltsbeschreibung*
Textuelle Zusammenfassungen des Inhalts eines Schriftstücks werden dem Überbegriff „Inhaltsbeschreibung“ zugeordnet.
- *Person*
Personen, werden in der Applikation als solche über diesen Überbegriff deklariert.

Über die Neuerfassung eines Schlagworts kann ein Anwender die Menge der Begriffe zu einem Überbegriff jederzeit ergänzen. Auswahlfelder in der Applikation werden in Folge von den Benutzern erweiterbar, wenn die Auswahlmöglichkeiten vom Überbegriff abhängig sind.

Abbildung 12 erweitert die Wissensabbildung um eine zeitliche und räumliche Dimension:

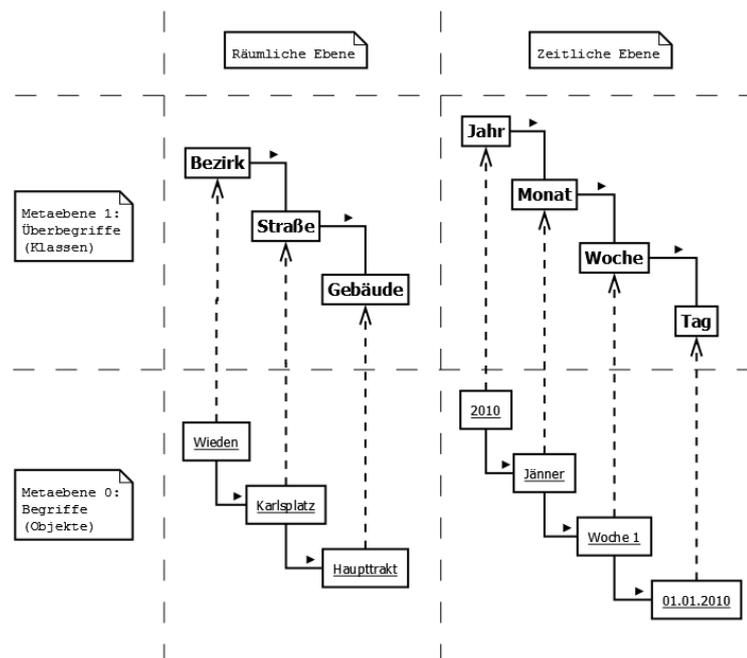


Abbildung 12: Räumliche und zeitliche Wissensabbildung im Triplestore

Für die zeitliche Ebene ist anzumerken, dass die Begriffe aus der Abbildung keine realen Schlagworte im System repräsentieren. Eine Abbildung aller in der Applikation vorkommenden Zeitangaben wäre nicht praktikabel und hätte einen negativen Einfluss auf das Leistungsverhalten und den Speicherverbrauch des Systems. Durch eine Mustererkennung auf Basis der regional gängigen Schreibweisen werden Zeitangaben als virtuelle Schlagworte unterstützt. Für den Anwender ist bei der Verwendung kein Unterschied zwischen virtuellen und gewöhnlichen Schlagworten erkennbar. Lediglich die Verarbeitung hinter den Kulissen des Systems stellt eine Abweichung dar.

Bei virtuellen Schlagworten werden nur Überbegriffe im Datenbestand abgelegt. Zusätzlich sind dem System Regeln für das Format und folglich die Erkennung von Instanzen eines virtuellen Schlagworts bekannt. Falls ein Suchbegriff noch keinem Überbegriff zugeordnet wurde oder ein Überbegriff für eine Zeitangabe von seinem Titel abweicht, findet eine Mustererkennung statt. Bei Übereinstimmung mit einem bekannten Muster handelt es sich um ein virtuelles Schlagwort. Letztendlich werden virtuelle Schlagworte bei der Weiterverarbeitung in Abfragekriterien für die Datenbank übersetzt.

Die Grundlage für den Datenbestand im Triplestore wird aus den aufgezählten Quellen bezogen. Das Setzen von semantischen Relationen zwischen den Ressourcen vervollständigt den Datenbestand.

- *Adressbuch*¹⁵ *der Technischen Universität Wien*
Aus dem Adressbuch werden die Organisationseinheiten, Fakultäten, Dekanate und Institute der Universität übernommen.
- *Website*¹⁶ *der „Gebäude und Technik“*
Von dieser Website werden die Adressen der Gebäude der Universität bezogen.

Für eine vollständige Abbildung aller Attribute der Ressourcen im vorgestellten Triplestore ist ein relationales Datenbanksystem aufgrund der Abfragekomplexität nicht die beste Wahl. Wenn die Anzahl der Attribute im Vorhinein bekannt und nicht dynamisch wie im Triplestore festgelegt wird, ist die Speicherung in einer Tabelle mit fixer Spaltenanzahl effizienter. Einzelne Datensätze können in einem relationalen Datenbanksystem sehr effizient ausgelesen werden, wenn die Tabelle nach dem Suchkriterium indiziert wurde. Aus den genannten Gründen kommt eine Hybridlösung zum Einsatz. Zusätzliche Attribute der im Triplestore eingebundenen Ressourcen werden in konventionellen relationalen Tabellen untergebracht, wenn diese Attribute aus maximal einem Wert bestehen. Abbildung 13 zeigt die Anbindung der Ressourcen aus dem Triplestore an die relationalen Tabellen.

¹⁵ siehe <http://tiss.tuwien.ac.at/>

¹⁶ siehe <http://www.gut.tuwien.ac.at/>

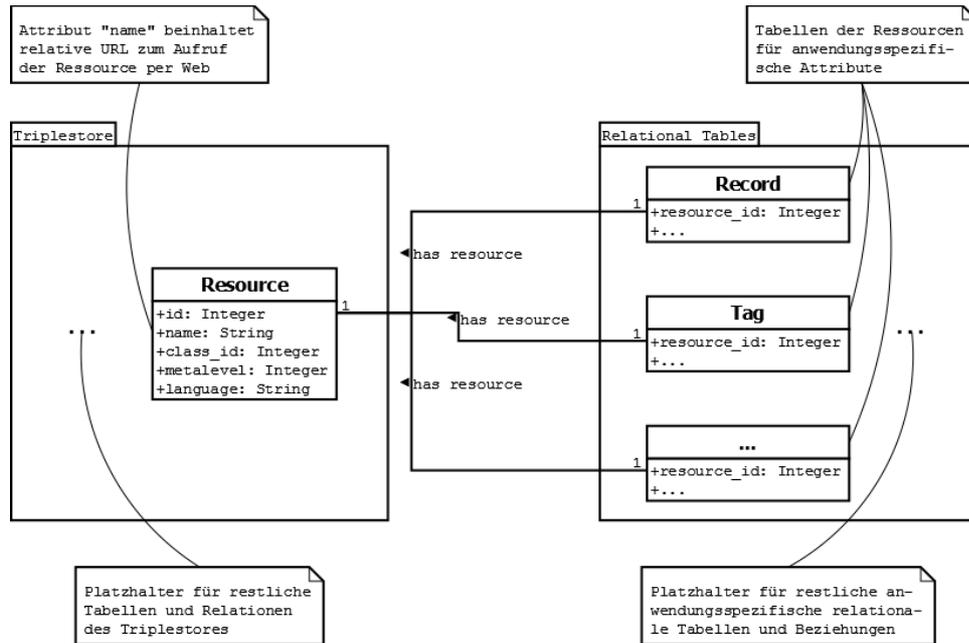


Abbildung 13: Klassendiagramm „Triplestore und relationale Tabellen“

In dieser Applikation ist es unerlässlich, dass Änderungen von Schriftstücken und Schlagworten im Verlauf der Zeit nachvollziehbar bleiben. Das System muss die Möglichkeit offen halten, dass Inhalte mit Bezeichnungen aus der Vergangenheit aufgefunden werden können. Abbildung 14 zeigt wie diese Anforderung in der Datenbank abgebildet wird.

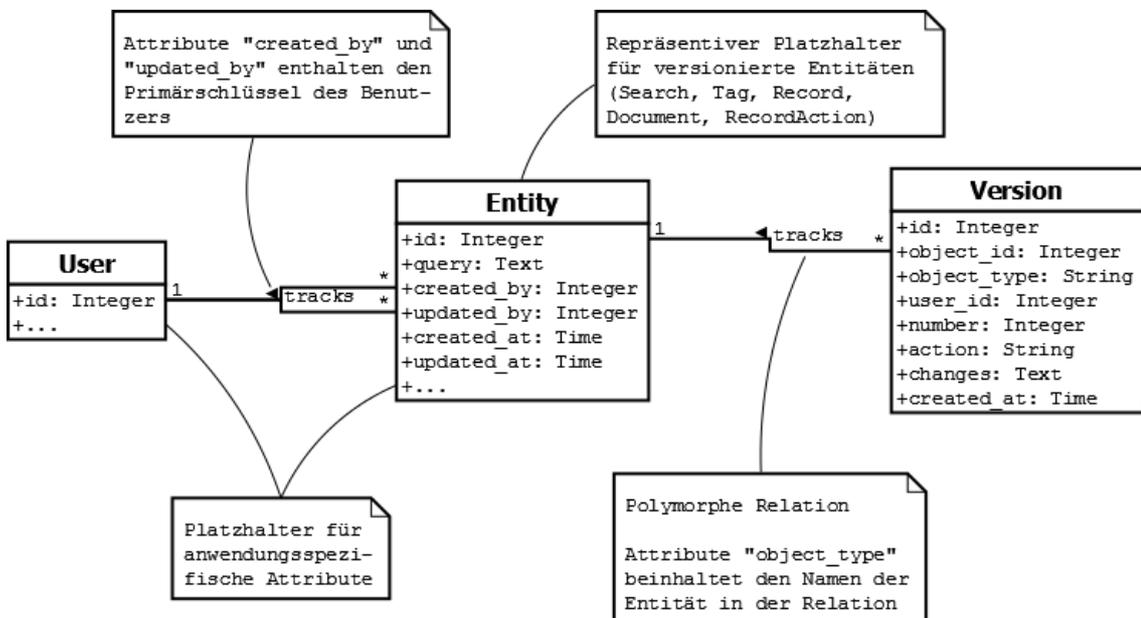


Abbildung 14: Klassendiagramm „Versionierung von Entitäten“

Die Klasse „Entity“ aus dem Diagramm ist als Platzhalter für alle Entitäten im Datenbankschema zu verstehen, deren Änderungen nachvollziehbar sein müssen. Zu diesen Entitäten werden Schriftstücke, Dokumente, Schlagworte und Suchen gezählt. In der

Tabelle „Version“ wird bei Änderung einer überwachten Entität, ein Datensatz zur vorigen Version erstellt. Dabei werden die bei der Aktualisierung des Datensatzes vorgenommenen Spaltenänderungen als Text in der Spalte „changes“ gespeichert.

Begründet durch die Verwendung des Triplestores, reicht die Tabelle „Version“ alleine nicht aus, um alle Attributänderungen zu erfassen. Die Änderungen von exklusiv im Triplestore abgebildeten Attributen werden in URL-Kodierung in der Spalte „query“ abgelegt.

Abbildung 15 zeigt das gesamte Datenbankschema auf einen Blick. Um diesen Abschnitt zu vollenden, werden im Anschluss an die Abbildung die Bedeutungen der Tabellenspalten kommentiert.

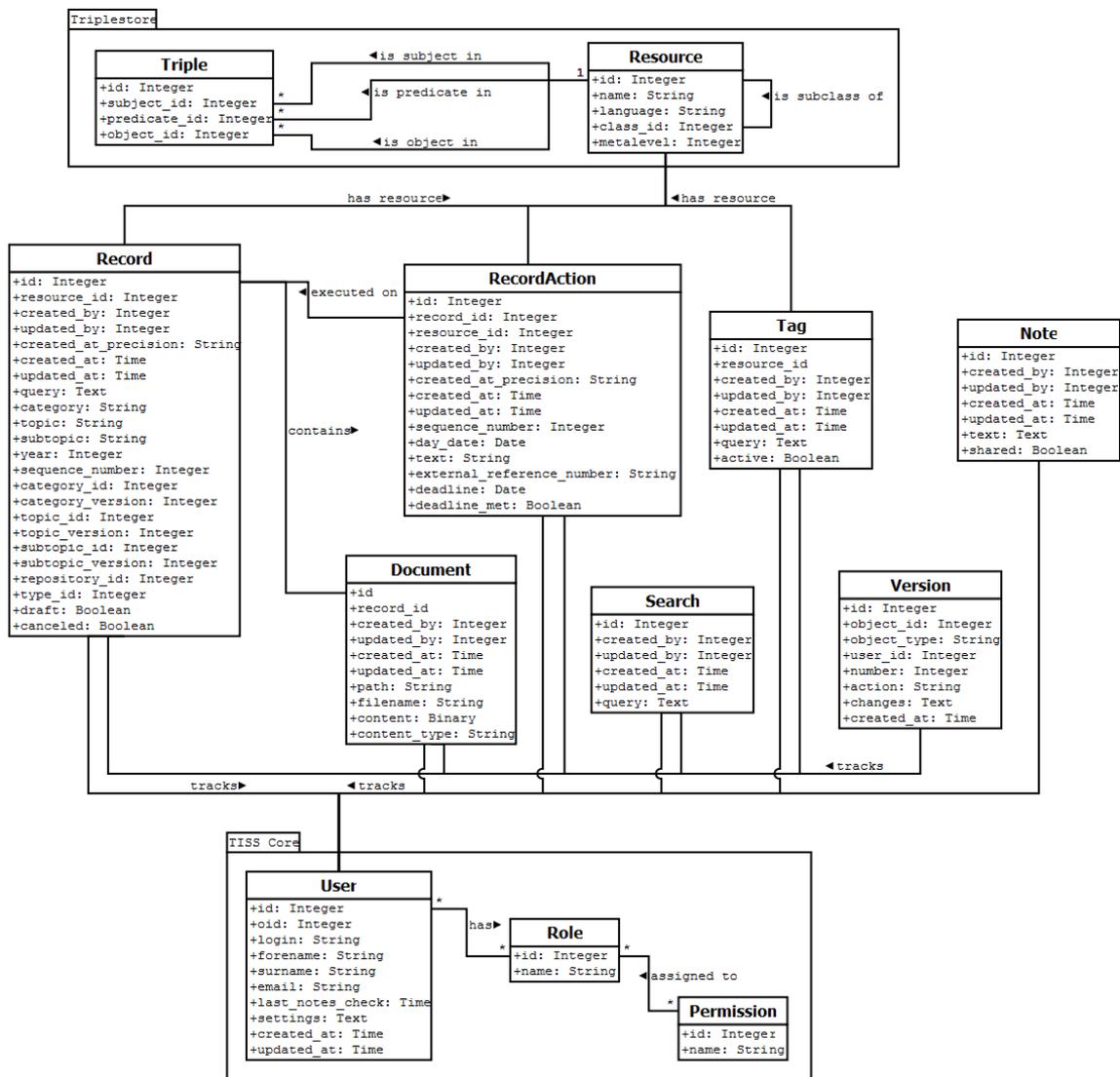


Abbildung 15: Klassendiagramm „Datenbankschema“

Die Daten der dem Paket „TISS Core“ zugeordneten Tabellen werden per Webservice aus dem TISS-Kernsystem¹⁷ bezogen.

Bei der Namensgebung von Tabellenspalten werden folgende allgemeine Konventionen eingehalten:

- *Primärschlüssel tragen den Namen „id“.*
Als Primärschlüssel kommen ausschließlich Surrogatschlüssel zum Einsatz. Ein Surrogatschlüssel ist ein künstlich erzeugter, meistens fortlaufender Zahlenwert, der einem Datensatz zugeordnet wird.
- *Fremdschlüssel enden im Namen mit „_id“.*
Bei Eindeutigkeit der Relation kann zusätzlich der Namen der referenzierten Entität als Präfix verwendet werden. Fremdschlüssel zur Zuordnung von Datensatzänderungen zu einem Bearbeiter stellen mit der Endung „_by“ eine Ausnahme dar.

Tabelle 15 fasst die Bedeutung der Spalten zusammen, die namentlich in mehr als einer Tabelle gegenwärtig sind.

Spaltenname	Bedeutung
created_by	Primärschlüssel des Erstellers aus der Tabelle „User“
updated_by	Primärschlüssel des Bearbeiters aus der Tabelle „User“
created_at	Erstellungszeitpunkt des Datensatzes
created_at_precision	Präzision der Zeitangabe des Erstellungszeitpunkts <ul style="list-style-type: none"> • „D“ ... Datum • „T“ ... Datum mit Uhrzeit
updated_at	Aktualisierungszeitpunkt des Datensatzes
query	Attributänderungen im Triplestore in URL-Kodierung

Tabelle 15: Bedeutung von mehrfach verwendeten Spaltennamen

Die Bedeutungen der restlichen Spalten mit Erklärungsbedarf sind in Tabelle 16 nachzulesen. Die angeführten Referenzen des Objekts „Record“ beziehen sich auf die Ressourcen der Schlagworte mit dem genannten Überbegriff.

Spaltenname	Bedeutung
Document	
path	Dateipfad bei Ablage der Datei im Dateisystem
filename	Dateiname
content	Binärdaten bei Ablage der Datei in der Datenbank
content_type	Internet Media Type der Datei

¹⁷ siehe <http://tiss.tuwien.ac.at/>

Note	
text	Text der Notiz
shared	Status der Sichtbarkeit der Notiz
Record	
category	Kürzel des zugeordneten Ablagebereichs erster Ebene
topic	Kürzel des zugeordneten Ablagebereichs zweiter Ebene
subtopic	Kürzel des zugeordneten Ablagebereichs dritter Ebene
year	Jahreszahl der Erfassung
sequence_number	Laufnummer des Schriftstücks
category_id	Referenz auf den zugeteilten Ablagebereich erster Ebene
category_version	Versionsnummer des zugeteilten Ablagebereichs erster Ebene zum Zeitpunkt der Schriftstück erfassung
topic_id	Referenz auf den zugeteilten Ablagebereich zweiter Ebene
topic_version	Versionsnummer des zugeteilten Ablagebereichs zweiter Ebene zum Zeitpunkt der Schriftstück erfassung
subtopic_id	Referenz auf den zugeteilten Ablagebereich dritter Ebene
subtopic_version	Versionsnummer des zugeteilten Ablagebereichs dritter Ebene zum Zeitpunkt der Schriftstück erfassung
repository_id	Referenz auf den Aufbewahrungsort
type_id	Referenz auf die Dokumentenart
draft	Status als Entwurf <ul style="list-style-type: none"> • „true“ ... Entwurf • „false“ ... Schriftstück
canceled	Status der Gültigkeit <ul style="list-style-type: none"> • „true“ ... storniert • „false“ ... gültig
RecordAction	
sequence_number	Laufnummer des Bearbeitungsschritts
day_date	Bezugsdatum
text	Beschreibungstext
external_reference_number	Externe Geschäftszahl
deadline	Frist für die Erfüllung des Bearbeitungsschritts
deadline_met	Status der Frist <ul style="list-style-type: none"> • „true“ ... erfüllt • „false“ ... nicht erfüllt
Resource	
name	Relative URL oder Literal
language	Sprachcode nach ISO 639-1
class_id	Klasse der Ressource
metalevel	Metaebene der Ressource
Tag	
active	Status der Verwendung <ul style="list-style-type: none"> • „true“ ... in Verwendung • „false“ ... stillgelegt
User	
oid	Eindeutige Nummer des Benutzers im TISS-Kernsystem
last_notes_check	Zeitpunkt der letzten Einsicht in die Notizen

settings	Anwendungsspezifische Einstellungen im Textformat
Version	
object_id	Primärschlüssel des versionierten Objekts
object_type	Name der versionierten Entität (zum Beispiel „Record“)
number	Versionsnummer
action	Art der auf dem Datensatz durchgeführten Aktion <ul style="list-style-type: none"> • „create“ ... Datensatz wurde erstellt • „read“ ... Datensatz wurde gelesen • „update“ ... Datensatz wurde aktualisiert • „delete“ ... Datensatz wurde gelöscht
changes	Änderungen des Datensatzes im Textformat

Tabelle 16: Bedeutung der Datenbankspalten

5.2.3 Prozesssicht

Der generelle Ablauf der Kommunikation zwischen Browser und Web Server wird in Abbildung 16 dargestellt.

Ruft der Benutzer eine Webseite der Applikation auf, wird vom Browser eine Anfrage an den Webserver gesendet. Die Anfrage wird von der am Web Server laufenden Applikation verarbeitet und mit dem Ergebnis beantwortet. Der Quelltext der resultierenden Webseite wird an den Browser zur Anzeige übertragen. Innerhalb dieses Quelltexts befinden sich Skriptelemente, die vom Browser ausgeführt werden.

Zur Erweiterung der Funktion der Benutzeroberfläche wird die Ausführung von Skriptteilen an Eingaben des Benutzers gebunden. Innerhalb dieser Skriptteile werden asynchrone Anfragen an den Webserver geschickt, falls weitere Informationen zur Anpassung der Benutzeroberfläche erforderlich sind. Nach Eintreffen der Antwort vom Webserver wird der Quelltext der Webseite vom Skript entsprechend transformiert.

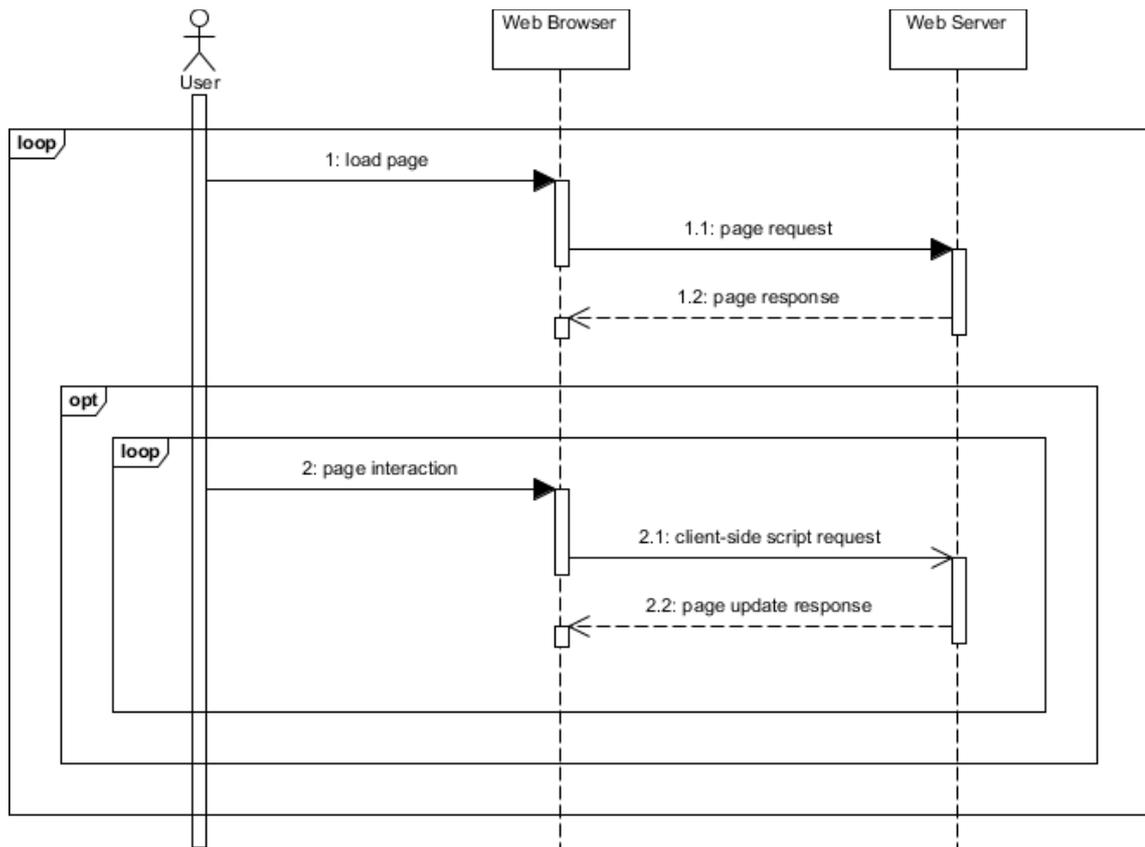


Abbildung 16: Sequenzdiagramm „Netzwerkkommunikation“

5.2.4 Entwicklungssicht

Abbildung 17 zeigt, aus welchen Komponenten das System zusammengesetzt ist. Des Weiteren sind die Schnittstellen, die von einer Komponente angeboten oder verwendet werden, ersichtlich.

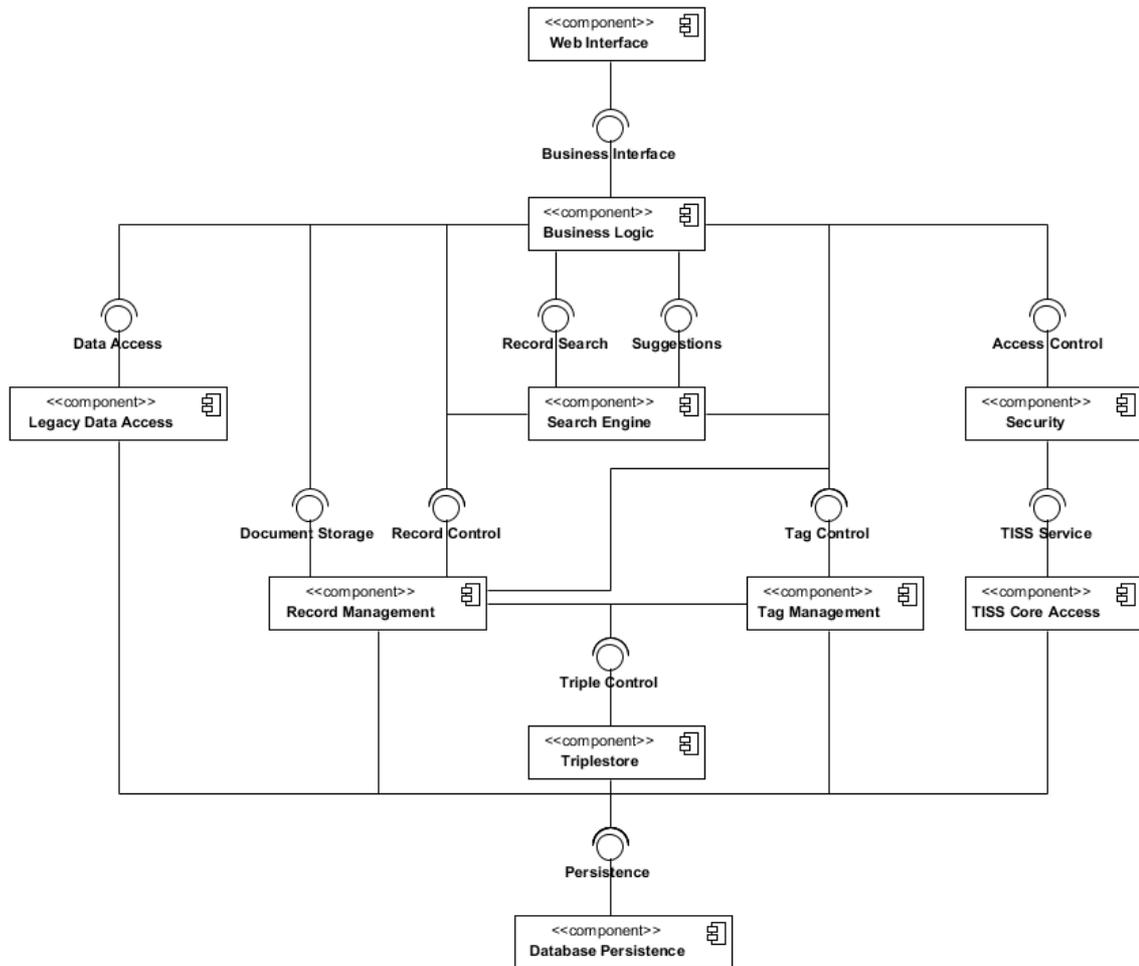


Abbildung 17: Komponentendiagramm des Systems

Die Komponente „Web Interface“ umfasst alle technischen Artefakte die zur Bereitstellung der Benutzeroberfläche innerhalb eines Browsers notwendig sind. Darunter werden auch in Webseiten inkludierte Skripte zur Erweiterung der Benutzerschnittstelle gezählt.

Vom Anwender über die Benutzeroberfläche gestellte Anfragen werden von der Komponente „Business Logic“ verarbeitet. Diese Komponente bedient sich der anderen Komponenten um die Logik der Geschäftsprozesse abzubilden. Die Geschäftslogik prüft die Befugnisse des Benutzers, bevor die entsprechenden Systemfunktionen ausgeführt werden.

Die Zugriffskontrolle stützt sich auf die Komponente „Security“, welche die Informationen zur Aufrechterhaltung der Sicherheit aus dem TISS-Hauptsystem bezieht. Der Zugriff auf die Daten aus dem TISS-Hauptsystem wird von der Komponente „TISS Core Access“ realisiert.

Für den Umgang mit den Daten aus dem Altsystem ist die Komponente „Legacy Data Access“ zuständig. Neben dem Lesezugriff wird auch die Datenmigration unterstützt.

Die Komponente „Search Engine“ realisiert die Algorithmen der Suche und liefert Empfehlungen für die Navigation, die Schlagwortvergabe und die Eingabevollständigung. Um die Suchergebnisse bereitstellen zu können, wird auf die Komponenten „Record Management“ und „Tag Management“ zugegriffen. Diese beiden Komponenten stellen den Umgang mit Akten beziehungsweise Schlagworten in Form von Funktionsbausteinen für die Geschäftslogik zur Verfügung. Als Zusatz beinhaltet die Komponente „Record Management“ eine Schnittstelle zur Verwaltung der in einem Akt abgelegten Dokumente.

Die Speicherung und Abfrage des Informationsgraphen ist Aufgabe der Komponente „Triplestore“. Akten und Schlagworte werden nach ihren Beziehungen untereinander vernetzt.

Die Komponente „Database Persistence“ kümmert sich um den Transfer der Informationen aus der Datenbank in Objekte für die applikationsinterne Verwendung. Bei Bedarf wird die Datenbank wieder an die Informationen in den Objekten angeglichen.

5.2.5 Physische Sicht

Mit dem Verteilungsdiagramm in Abbildung 18 wird die im TISS-Projekt zum Einsatz kommende technische Infrastruktur beschrieben.

Von der Infrastruktur werden Web-Applikationen in den Sprachen Ruby und Java unterstützt. Zwei Server stellen eine Reihe von virtuellen Maschinen mit je einer Ausführungsumgebung für eine der beiden Sprachen zur Verfügung. Nach dem Pfad einer aufgerufenen Webseite wird entschieden, an welche Gruppe von Applikationsservern die Anfrage weitergeleitet wird. Nachdem die Entscheidung getroffen wurde, ob für die Verarbeitung der Anfrage eine Applikation in Java oder Ruby zuständig ist, wird eine Applikationsserverinstanz für die Abarbeitung ausgewählt. Dabei wird die Last auf die vorhandenen Instanzen verteilt.

Im Falle eines Ausfalls des primären Datenbankservers kann die sekundäre Instanz die Arbeit übernehmen. Der Datenbestand der sekundären Instanz wird während des regulären Betriebs aktuell gehalten.

Für die Ablage von Dateien steht ein gemeinsam genutzter dedizierter Server bereit.

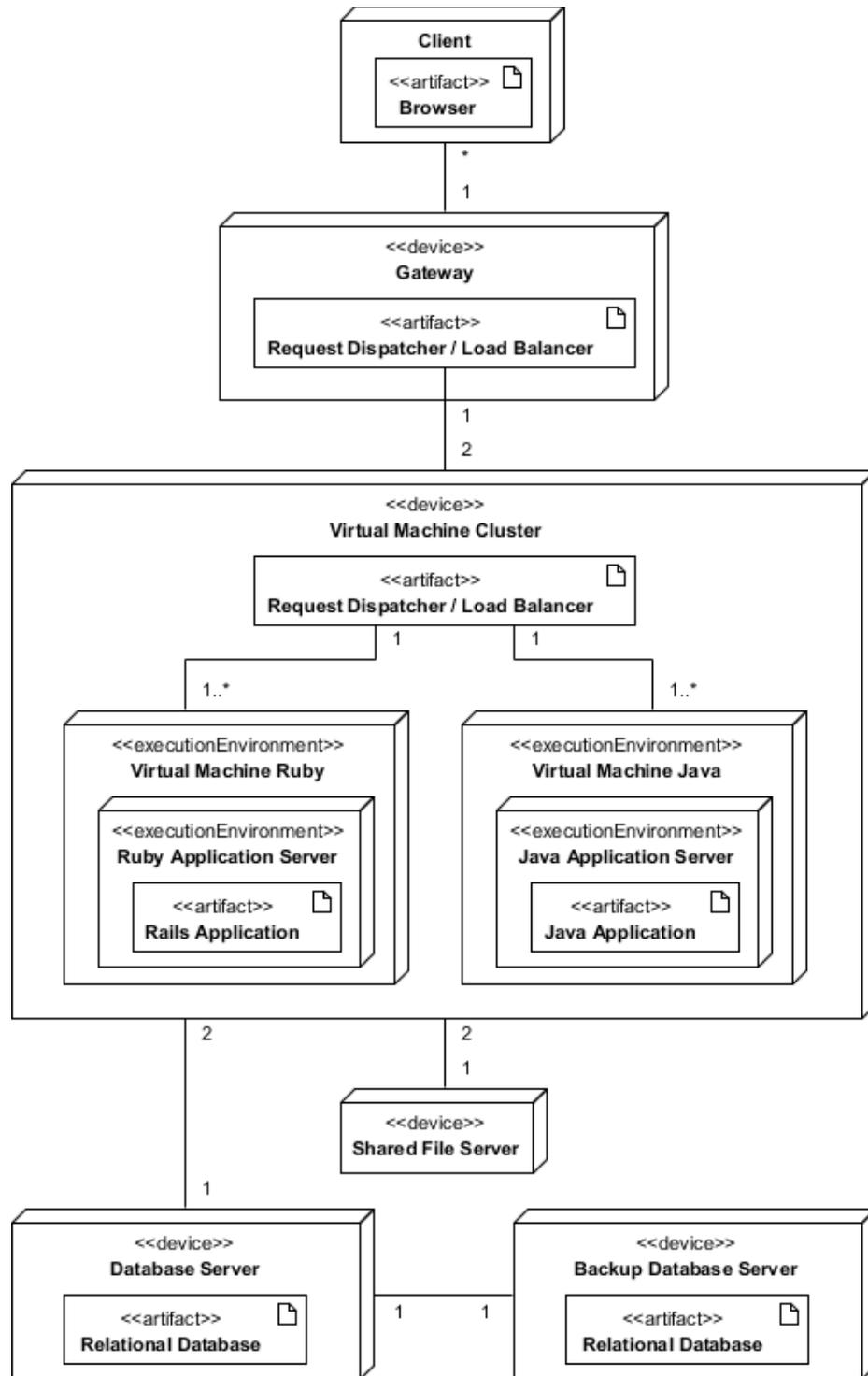


Abbildung 18: Verteilungsdiagramm „TISS-Infrastruktur“

5.3 Datenmigration

Um eine Ablöse des in Betrieb stehenden Kanzleinformationssystems möglich zu machen, muss das nachfolgende System den vorhandenen Datenbestand vollständig übernehmen. In diesem Abschnitt ist die eigens für die Migration der Daten entworfene Vorgangsweise festgehalten.

Das Hauptproblem bei der Übernahme der Daten besteht darin, aus den beschreibenden Attributen eines Schriftstücks des Altsystems eine Liste von passenden Schlagworten anzufertigen. Die Schlagwörter dieser konstruierten Liste werden bei der Migration dem Schriftstück zugewiesen. Vor der Beschreibung des Verfahrens zur Datenübernahme wird ein Blick auf die Attribute eines Schriftstücks geworfen, die sich zur Generierung von Schlagworten eignen. Weitere Details zu den in der folgenden Aufzählung genannten Attributen sind im Abschnitt 4.1.1 zu finden:

- *Ablagebereich*
Die Nummer für den Themenbereich und das Subthema eines Schriftstücks formen in Kombination einen Ablagebereich. Der Name des Ablagebereichs wird im Kanzleinformationssystem von den Anwendern eingetragen.
- *Inhaltsbeschreibung*
Die Inhaltsbeschreibung eines Schriftstücks liegt als frei formulierter Text mit einer Beschränkung der Länge auf 150 Zeichen vor.

Tabelle 17 liefert einige Beispiele für Namen von Ablagebereichen. Als Beispiele wurden besonders problematische Fälle ausgewählt, um die Schwierigkeiten und Grenzen bei der Extraktion von Schlagworten zu vermitteln. Ein Ablagebereich, der einem anderen untergeordnet ist, kann in seinem Namen den Beschreibungstext des übergeordneten Ablagebereichs enthalten. Unterschiede bei der Schreibweise, den verwendeten Abkürzungen und den gesetzten Trennzeichen schließen neben Tippfehlern die Zuordnung zu einem einzelnen gemeinsamen Begriff durch eine Maschine ohne Sprachverständnis aus.

Nummer		Name
Thema	Sub-thema	
10	03	Universitätsverwaltung, Kanzleiordnung
10	04	Rechnungsabschlüsse der Universitätsverwaltung
10	11	Universitätsverwaltung - Werkverträge
183	00	Institut für Rechnergestützte Automation
183	02	Protokolle der Institutskonferenz
183	03	Institut für Rechnergestützte Automation, Wahl des Institutsvorstandes
183	05	Institut für Rechnergestützte Automation, Institutsordnung
183	12	Institut für Rechnergestützte Automation - Anträge gem. § 15 Abs. 2 FOG
183	26	Institut für Rechnergestützte Automation, Verträge nach § 26
184	00	Institut für Informationssysteme
184	02	Institut für Informationssysteme, Protokolle d. Institutskonferenz
184	03	Wahl des Institutsvorstandes
184	05	Institut für Informationssysteme - Institutsordnung
184	12	Institut f. Informationssysteme - Verträge gemäß § 15 (2) FOG
184	26	Inst. f. Informationssysteme, Verträge nach § 26

Tabelle 17: Beispiele für Ablagebereiche des Kanzleinformationssystems

Wie bereits angekündigt, wird als nächstes das Verfahren zur Übernahme der Daten beschrieben. Die Datenmigration umfasst folgende Schritte:

- *Aufbereitung der Daten*

Dieser optionale Schritt soll zur Verbesserung der Qualität der migrierten Daten beitragen, indem die Zuordnung unterschiedlicher Schreibweisen zu einem gemeinsamen Schlagwort verbessert wird und bei der Dateneingabe entstandene Fehler behoben werden. Eine vollständige Umsetzung der genannten Maßnahmen gestaltet sich in Anbetracht des vorhandenen Datenvolumens als sehr aufwendig. Deswegen liegt es im Ermessen des Bearbeiters, der die Migration durchführt, zu entscheiden, wie genau eine sinnvolle Umsetzung der Maßnahmen auszusehen hat.

1. *Behandlung von Abkürzungen*

Abkürzungen sollten durch ihre ausgeschriebene Bedeutung ersetzt werden. Alternativ kann auch eine Liste der verwendeten Abkürzungen inklusive deren Bedeutung für die Berücksichtigung bei der Migration erstellt werden.

2. *Angleichung der Formulierungen*

Wenn der Name eines untergeordneten Ablagebereichs stark vom übergeordneten Thema abweicht, können die Formulierungen aneinander angeglichen werden.

3. *Rechtschreibprüfung*

Die Durchführung einer Rechtschreibprüfung mit einem Textverarbeitungsprogramm ist der Weg mit dem geringsten Aufwand zur Verbesserung der Qualität des Datenbestands.

4. *Behebung von Eingabefehlern*

Durch eine unzureichende Überprüfung der Eingaben und die Verwendung von unpassenden Datentypen besteht im Altsystem die Möglichkeit, dass fehlerhafte Eingaben in den Datenbestand aufgenommen wurden. Daten, die außerhalb des gültigen oder sinnvollen Wertebereichs liegen, sollten aufgefunden und korrigiert werden.

- *Maschinelle Datenmigration*

Die Informationen der im Kanzleiinformationssystem abgelegten Schriftstücke werden maschinell in das neue System übertragen. Zur Migration eines einzelnen Schriftstücks werden diese Arbeitsschritte von der Maschine ausgeführt:

1. *Schlagwörter ermitteln und übernehmen*

Der Ablagebereich eines Schriftstücks bildet die Grundlage für die Ermittlung der Schlagwörter. Befindet sich der Ablagebereich in der zweiten Ebene der Themenhierarchie, wird der übergeordnete Ablagebereich in die Ermittlung einbezogen. Ablagebereiche werden je nach ihrer Hierarchieebene unter den Überbegriffen „Ablagebereich (Thema)“ oder „Ablagebereich (Unterthema)“ als Schlagwörter für die Verwendung bei der Suche migriert. Falls der Name des untergeordneten Ablagebereichs mit dem Text seines übergeordneten Ablagebereichs beginnt, wird nur der noch nicht vorgekommene Textteil zur Schlagwortbildung herangezogen. Aus dem Namen eines Ablagebereichs werden maximal zwei Schlagwörter durch die Berücksichtigung der in den Daten gebräuchlichen Trennzeichen extrahiert. Sind die Schlagwörter dem System noch nicht bekannt, werden sie unter dem allgemeinen Überbegriff „Thema“ übernommen, falls keine Zuordnung zu einem spezifischen Überbegriff maschinell durchführbar ist. Die in Textform vorliegende inhaltliche Beschreibung eines Schriftstücks aus dem Altsystem wird unter dem Überbegriff „Inhaltsbeschreibung“ als Schlagwort migriert. Abbildung 19 zeigt die in diesem Abschnitt beschriebene Vorgangsweise anhand eines Beispiels.

2. *Schriftstück übernehmen*

Nach der Ermittlung der Schlagwörter folgt die Übertragung der Daten des Schriftstücks. Die Attribute des Schriftstücks werden übernommen und falls notwendig in den passenden Datentyp konvertiert. Für den Fall, dass eine Konvertierung nicht durchführbar ist, wird eine Meldung mit einem Verweis auf den problematischen Datensatz ausgegeben. Im neuen System eingeführte Attribute werden mit den vorhandenen Daten berechnet und gesetzt. Letztendlich werden die ermittelten Schlagwörter dem Schriftstück zugewiesen.

3. *Bearbeitungsschritte übernehmen*

Zur Vervollständigung der Migration eines Schriftstücks werden im letzten Schritt dessen Bearbeitungsschritte übernommen. Analog zum vorigen Arbeitsschritt werden gegebenenfalls Datentypen konvertiert und Attribute ergänzt.

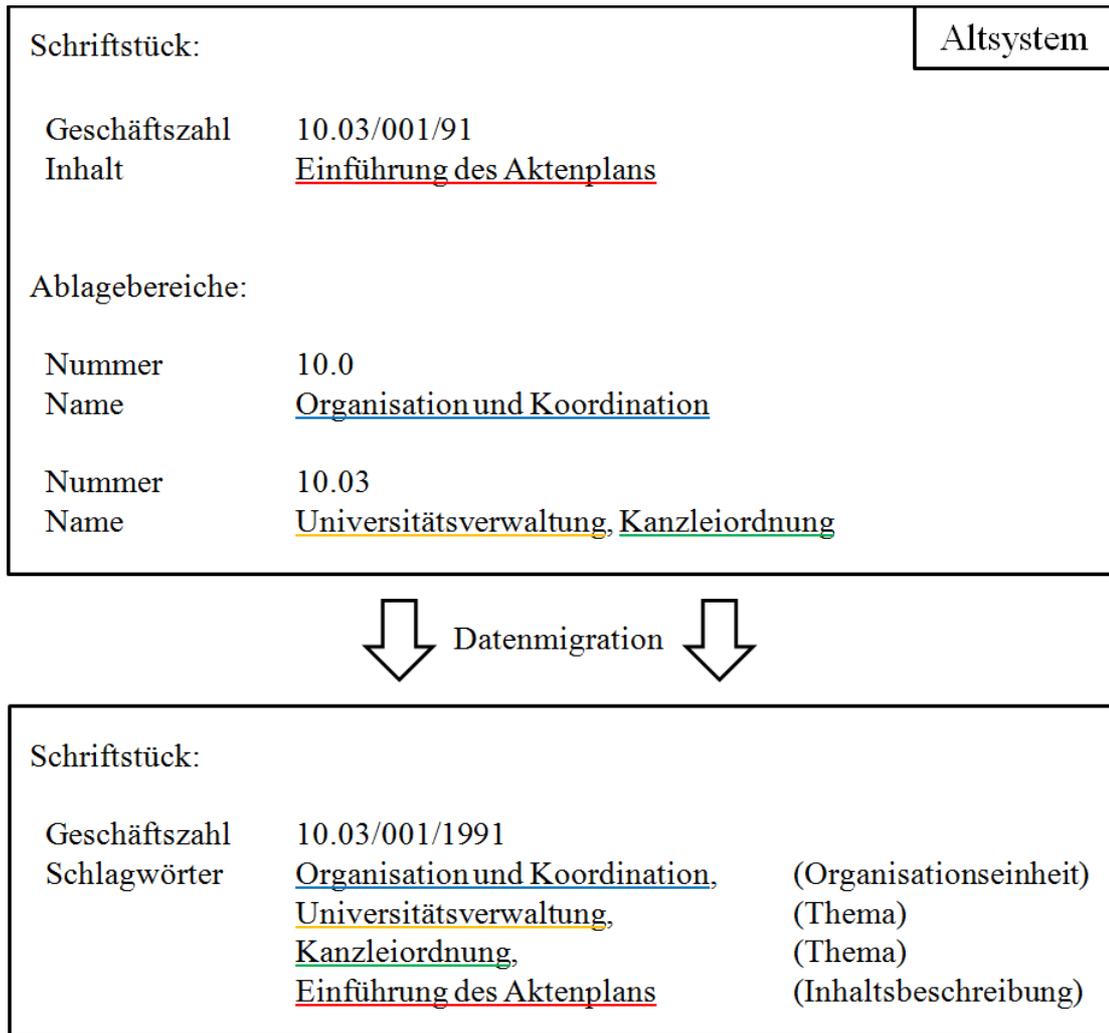


Abbildung 19: Beispiel für die Extraktion von Schlagworten

Abschließend anzumerken ist, dass eine Kopie des finalen Datenbestands des Kanzleiinformationssystems in der Datenbank des neuen Systems abgelegt wird. Diese Entscheidung wurde aus mehreren Gründen getroffen. Zum einen besteht durch diese Maßnahme die Möglichkeit den originalen Datensatz eines migrierten Schriftstücks anzuzeigen. Dem Benutzer wird so ein zusätzliches Gefühl von Sicherheit vermittelt, weil keine Daten bei der Migration verloren gehen können. Über den Vergleich zum originalen Datensatz wird die Arbeitsweise der Migration nachvollziehbar und Fehler kommen schneller ans Tageslicht. Zum anderen können nach der Migration entdeckte Fehler auch im Nachhinein korrigiert werden. Fehlende Informationen können im Rahmen einer aufbauenden Migration ergänzt werden.

6 Realisierung

Dieses Kapitel widmet sich den Details der softwaretechnischen Umsetzung des im vorigen Kapitel vorgestellten Systemkonzepts.

Begonnen wird mit der Beschreibung der Besonderheiten der eingesetzten Technologien. Die Designprinzipien des Web-Frameworks „Ruby on Rails“ werden zusammengefasst und zusätzlich werden die Unterschiede zu anderen Frameworks angeführt. Der Einsatz von „MySQL“ als Datenbank und der „Yahoo! User Interface Library“ als JavaScript und CSS-Bibliothek wird durch die Aufzählung der für die Entwicklung positiven Eigenschaften dieser Technologien begründet.

Der darauf folgende Abschnitt zeigt die wesentlichen Merkmale der entworfenen Benutzerschnittstelle. Ausgehend von der Erklärung der auf der Startseite der Applikation angezeigten Elemente werden die für die grundsätzliche Bedienung notwendigen Formulare und Eingabefelder behandelt. Darunter fallen zum Beispiel die Eingabevervollständigung für Schlagworte, die Verwaltung einer Schlagwortliste und die Formulare zur Erfassung von Schriftstücken und Schlagworten.

Danach wird die Suchfunktion der Applikation anhand von Beispielen erklärt. Für jede vom System unterstützte Art von Suchkriterium wurde ein repräsentatives Beispiel konstruiert. Die Suche nach Schlagwohierarchien, Überbegriffen, Schriftstückattributen, Erfassungszeiträumen, Geschäftszahlen und vorkommenden Texten wird von den Beispielen abgedeckt.

Zum Abschluss wird auf Optimierungsmaßnahmen eingegangen, die zur Erreichung der laut Abschnitt 4.3.2 gewünschten Systemleistung notwendig waren. Als Einstieg in das Thema werden allgemeine Maßnahmen zur Verbesserung der Systemleistung betrachtet. Die Beschreibung der praktischen Umsetzung dieser Maßnahmen knüpft inhaltlich an den allgemeinen Teil an. Die Verwendung der zwecks Optimierung erstellten Tabellen wird dokumentiert, indem dem Leser die Bedeutung der Tabellenspalten nähergebracht wird.

6.1 *Eingesetzte Technologien*

Dieser Abschnitt widmet sich mit seinen Unterpunkten den verwendeten Technologien „Ruby on Rails“, „MySQL“ und „Yahoo! User Interface Library“. Jedem aufgezählten Element ist ein einzelner Abschnitt gewidmet, in dem die prägnanten Merkmale der Technologie mit Relevanz für das Projekt zusammengefasst wurden. Bei der Auswahl wurde vor allem auf Flexibilität, Verfügbarkeit und Komfort bei der Entwicklung geachtet.

Aufgrund der zufriedenstellenden Unterstützung von „Ruby on Rails“ kam „Netbeans“¹⁸ als integrierte Entwicklungsumgebung zum Einsatz. Zur Bereitstellung der Infrastruktur auf den Entwicklungssystemen wurde die benötigte Software mit dem aggregierten Softwarepaket „BitNami RubyStack“¹⁹ installiert, das „MySQL“, „Ruby“, „Ruby on Rails“ und den Webserver „Apache“ umfasst.

6.1.1 Ruby on Rails

„Ruby on Rails“²⁰ kurz „Rails“ ist ein quelloffenes Framework auf Basis der Programmiersprache „Ruby“²¹ zur Entwicklung von Web-Applikationen, das nach folgenden Prinzipien entworfen wurde:

- *DRY (Don't Repeat Yourself)*
Wiederholungen von Programmcode sind zu vermeiden. Tritt derselbe oder ein ähnlicher Quelltext an mehreren Stellen in der Applikation auf, erschwert dies die Lesbarkeit und Wartung. Im Falle einer Änderung muss meist jedes Auftreten des Quelltexts angepasst werden. Es solches Vorgehen führt in Folge des Öfftens zu Fehlern. Deswegen wird empfohlen, wiederkehrende Programmteile zu identifizieren und als wiederverwendbare Funktionsbausteine umzusetzen.
- *Konvention vor Konfiguration*
Innerhalb des Frameworks werden Annahmen für den zu erwartenden Normalfall getroffen, um den Aufwand der Konfiguration so gering wie möglich zu halten. Die Notwendigkeit der Konfiguration durch das Editieren von Dateien oder Übergabe von zusätzlichen Parametern an Funktionen wird so für den Standardfall minimiert. Häufig an Web-Applikationen gestellte Anforderungen können durch diese Unterstützung des Frameworks einfach in der Programmierung erfüllt werden.
- *REST (Representational State Transfer)*
Das Ansprechen an Steuern von Ressourcen der Applikation über die standardisierten Verben „GET“, „POST“, „PUT“ und „DELETE“ des HTTP-Protokolls wird als Architekturstil für Web Services von Rails favorisiert und entsprechend unterstützt.

In der Architektur folgt Rails wie auch viele andere Web Frameworks dem MVC-Paradigma. MVC steht für die Trennung des Quelltexts nach Funktionsart in die drei im Englischen als „Model“, „View“ und „Controller“ bezeichneten Schichten. Dem „Mo-

¹⁸ siehe <http://netbeans.org/>

¹⁹ siehe <http://bitnami.org/stack/rubystack>

²⁰ siehe <http://rubyonrails.org/>

²¹ siehe <http://www.ruby-lang.org/>

del“ zugeordnete Programmteile beinhalten die Geschäftslogik der Applikation mit der Abbildung der dazugehörigen Daten. Die Schicht „View“ ist für die Präsentation der Daten in der Benutzerschnittstelle zuständig. Im Falle einer Web-Applikation versteht sich darunter die Produktion des Quelltexts der Webseite. Unter den Bereich „Controller“ fällt die Steuerung der Anwendung bei der Verarbeitung der Benutzereingaben. Bei einer Web-Applikation äußert sich dies in der Entgegennahme der vom Browser gesendeten Anfragen, dem Aufruf der zur Verarbeitung notwendigen Geschäftslogik und der Weitergabe der Daten an die Anzeigelogik zur Erstellung der auszugebenden Webseite.

Die folgende Aufzählung schenkt einigen Besonderheiten von Rails Beachtung, die in der Dokumentation²² des Frameworks nachgelesen werden können:

- *Automatisation*
Häufig auftretende Tätigkeiten sind im Framework automatisiert und können per Kommandozeile aufgerufen werden. Per Kommando werden zum Beispiel neue Dateien mit Quelltext als Vorlage generiert, von der Applikation benötigte Abhängigkeiten installiert oder Testdaten in die Datenbank geladen. Zur Automatisation projektspezifischer Vorgänge können neue Kommandos hinzugefügt werden.
- *Migrationen*
Inkrementelle Änderungen an der Struktur der Datenbank werden in mehrere einzelne Dateien gruppiert und Migrationen genannt. Jede Migration enthält die Logik zur Durchführung und Rückgängigmachung ihrer Änderungen. So kann die Struktur der Datenbank jederzeit an eine gewünschte Version angeglichen werden. Auf diesem Weg können Änderungen der Entwicklung komfortabel in Produktion gesetzt werden.
- *Konsole*
Die Konsole bietet einen interaktiven Modus der Befehlsausführung innerhalb der Programmumgebung der Web-Applikation. Per Kommandozeile können durch die Eingabe von Ruby-Quelltext die Funktionen des Frameworks und die der entwickelten Komponenten angesprochen werden. Erwartungsgemäß wird das Ergebnis der Auswertung eines eingegebenen Befehls textuell ausgegeben. Die Einsatzgebiete der Konsole gestalten sich vielseitig. Beispielsweise sind das Testen und die Analyse von kleinen Programmteilen neben der Ausführung einmaliger Anweisungen zur Wartung denkbar.

²² siehe <http://rubyonrails.org/documentation> und <http://guides.rubyonrails.org/>

Durch den Funktionsumfang des Frameworks und die Vielzahl an verfügbaren Plugins und umfangreichen Bibliotheken der Sprache „Ruby“ eignet sich Rails vorzüglich zur Entwicklung von Prototypen. Durch die Einbindung bereits vorhandener Komponenten lassen sich erste Ergebnisse vergleichsweise schnell erzielen. Der Einsatz von Rails im Projekt „TISS“ ist ein weiterer Grund das Framework zu verwenden, weil die benötigte technische Infrastruktur bereits zur Verfügung steht.

6.1.2 MySQL

„MySQL“²³ ist ein relationales Datenbankmanagementsystem mit offenem Quelltext. Laut den Angaben der offiziellen Website ist MySQL die am weitesten verbreitete Datenbank. Die freie Verfügbarkeit und einfache Installation auf einem Arbeitsplatz sind gute Argumente für den Einsatz in der Entwicklung.

Das Programm „MySQL Workbench“²⁴, die Weiterentwicklung der früher unter dem Namen „MySQL GUI Tools“ bekannten Applikationen, ist ebenfalls von der offiziellen Website beziehbar. Dieses Werkzeug unterstützt den Anwender mit einer grafischen Benutzerschnittstelle bei dem Design, der Abfrage und Administration der Datenbank. Das Vorhandensein eines geeigneten Werkzeugs um die Tätigkeiten bei der Entwicklung, wie zum Beispiel das Sichern und Wiederherstellen eines Datenbankzustands, zu vereinfachen ist sekundär ein weiterer Grund für die eingesetzte Technologie.

6.1.3 Yahoo! User Interface Library

Die „Yahoo! User Interface Library“²⁵ abgekürzt mit „YUI“ ist eine frei verwendbare JavaScript- und CSS-Bibliothek. Diese Bibliothek zeichnet sich neben ihren umfassenden Funktionen durch eine qualitativ hochwertige Dokumentation aus.

Im Funktionsumfang ist die Begründung für den Einsatz dieser Technologie im Rahmen dieser Arbeit zu finden. Die nachfolgende Aufzählung hebt die Komponenten der Bibliothek mit einer besonderen Bedeutung für die Implementierung hervor:

- *CSS (Cascading Style Sheets)*
Die enthaltenen CSS-Dateien liefern bei Einbindung einen Ausgangspunkt für das visuelle Design der Website. Das Aussehen von in verschiedenen Browsern unterschiedlich angezeigten HTML-Elementen wird durch die Bibliothek angeglichen. Die per CSS unterstützte optische Einteilung einer Webseite in Zeilen und Spalten, beschleunigt und vereinfacht die Realisierung eines Layouts.

²³ siehe <http://mysql.com/>

²⁴ siehe <http://mysql.com/products/workbench/>

²⁵ siehe <http://developer.yahoo.com/yui/>

- *Steuerelemente*
Die standardmäßigen HTML-Eingabeelemente werden durch eine Reihe von erweiterten Steuerelementen wie zum Beispiel Baumansichten, Kalender und Tabellen ergänzt. Für die Umsetzung der Applikation benötigte Textfelder mit Eingabevervollständigung lassen sich auf einfachen Weg bewerkstelligen.
- *Module*
Die Entwicklung mit JavaScript wird durch die bereitgestellten Module immens erleichtert. Tätigkeiten einer niedrigen Abstraktionsebene wie zum Beispiel die Kommunikation mit dem Webserver bei asynchronen Anfragen werden durch die Bibliothek bereits abgedeckt. In der Programmierung ermöglicht dies eine Verlagerung des Schwerpunkts auf die Anwendungslogik.

Zusammengefasst gibt der Einsatz der vorgestellten Bibliothek bei der Implementierung Entwicklungszeit frei, die stattdessen in die Verbesserung der Applikation investiert werden kann. Die Erstellung eines auf einem Raster basierenden Layouts mit YUI erfordert keine tiefgründigen CSS-Kenntnisse und eignet sich somit ideal für einen Prototyp.

6.2 Benutzerschnittstelle

Dieser Abschnitt beschreibt die wesentlichen Elemente der Benutzerschnittstelle. Begonnen wird mit der in Abbildung 21 gezeigten Startseite der Applikation. Nach einer ersten Phase mit dem in Abbildung 20 dargestellten schematischen Layout wurde das Design an das derzeitige vorherrschende Entwicklungsdesign der TISS-Applikation angepasst. Die Startseite der Applikation setzt sich aus folgenden Komponenten zusammen:

- *Menü*
Der Benutzer kann die Funktionen der Applikation über die links angezeigten Menüeinträge anwählen.
- *Suchkriterien*
Die Elemente zur Eingabe von Suchkriterien werden mittig zwischen Menü und Navigationsleiste angezeigt. In die Liste der Suchkriterien können Schlagworte, Überbegriffe, Textfragmente und Zeitangaben aufgenommen werden. Zusätzlich können Merkmale der Geschäftszahl spezifiziert werden.

- *Inhaltsbereich*

Unter den Eingabefeldern für die Suchkriterien befindet sich der Inhaltsbereich, in dem Eingabeformulare und Suchergebnisse angezeigt werden. Lange Listen werden auf mehrere Seiten aufgeteilt mit der Möglichkeit zwischen den Seiten zu blättern und die Anzeigeparameter wie Einträge pro Seite und die Sortierkriterien zu verändern.

- *Navigation*

Die rechts platzierte Leiste zeigt eine vom Inhaltsbereich kontextabhängige Navigation. Auf der Startseite wird eine Auswahl der meistverwendeten Schlagworte als Einstiegspunkt für weitere Suchanfragen präsentiert. Während der Suche werden Vorschläge zur Verfeinerung der eingegebenen Suchkriterien aufgelistet. Bei Formlaren wird auf in Bezug stehende Inhalte verwiesen.

Suche Schriftstück erfassen Schlagwort erfassen Berichte Fristen Erweiterte Suche	Johann Grabner Notizen Entwürfe Protokoll Abmelden
<div style="border: 1px solid black; padding: 5px;"> <div style="display: flex; justify-content: space-between;"> <div> <p>tiss21 Suchbegriffe <input type="text"/></p> <p>Geschäftszahl <input type="text"/> . <input type="text"/> . <input type="text"/> / <input type="text"/> / <input type="text"/></p> <p><input type="button" value="Suche"/></p> </div> <div style="text-align: right;"> </div> </div> </div>	
<p>Willkommen im Testsystem</p> <p>tiss21 ist ein in Entwicklung befindliches Dokumentenmanagementsystem (DMS) für die Kanzlei der TU Wien. Dieses System wird im Rahmen des TISS Projekts entwickelt. Die Zielsetzung von tiss21 ist die Berücksichtigung der inhaltlichen Zusammenhänge zwischen den zu verwaltenden Dokumenten und dem Anwendungsbereich. Konkret bedeutet das für Sie, dass Sie vom System bei Ihrer Arbeit laufend mit Vorschlägen für verwandte Themen und Sucheinschränkungen unterstützt werden. Zum Beispiel werden Ihnen nach Eingabe des Begriffs "Karlsplatz" die am Karlsplatz stehenden Gebäude und weitere verwandte Themen vorgeschlagen.</p> <p>Bei Fragen, Problemen und Kritikpunkten stehe ich Ihnen gerne als Kontaktperson zur Verfügung. Ich entwickle das System persönlich und bin bemüht Ihre Anforderungen und Ihr Anwendungswissen laufend in die Entwicklung einzubringen. Wenn Sie mir etwas mitteilen wollen oder Fragen haben, können Sie jederzeit von dem unten angeführten Feedback-Link johann.grabner@inso.tuwien.ac.at Gebrauch machen und mir eine E-Mail senden.</p> <p>Das System tiss21 ist in einer frühen Phase von mir online gestellt worden, um Ihnen die Möglichkeit zur Mitgestaltung und Beeinflussung der Entwicklung zu geben. Es sind noch nicht alle Funktionalitäten enthalten, dennoch gibt es bereits eine Vorschau auf Funktionen, die Sie erwarten. Die entsprechenden Seiten wurden mit dem Wort "Vorschau" versehen. Das System wird schrittweise in kurzen Intervallen von mir erweitert, Details finden Sie nachfolgend unter Neuigkeiten.</p> <p>Damit ich in der Entwicklung flexibel auf Ihre Bedürfnisse reagieren kann, wird in dieser Testphase nur der Browser Firefox in der aktuellen Version 3 unterstützt. Falls Sie einen anderen Browser verwenden und Anzeigeprobleme auftreten, können Sie mir diese gerne für eine spätere Behebung mitteilen, aber verwenden Sie vorerst bitte Firefox. Ich empfehle die tragbare Version von Firefox, die nach Entpackung ohne Installation lauffähig ist.</p> <p>Ich freue mich auf eine persönliche Zusammenarbeit.</p> <p>Mit freundlichen Grüßen</p> <p>Johann Grabner</p> <p>Neuigkeiten</p> <ul style="list-style-type: none"> • 11. Jänner 2010: Der Datenbestand der Kanzlei vom Dezember 2009 wurde in das Testsystem übernommen. • 4. Jänner 2010: Eine Zusammenfassung der bei Schriftstücken laufenden Fristen ist einsehbar. • 23. Dezember 2009: Bearbeitungsschritte von Schriftstücken können eingegeben und editiert werden. <p>weitere Neuigkeiten</p>	<p>Navigation</p> <p>Thema</p> <p>Inhaltsbeschreibung </p> <p>Anfrage, Überleitung storno, STORNO Anrechnung von Ruhegenussvordienstszeiten, ...</p> <p>Thema </p> <p>Zulassung zum Studium / Allgemeine Studienangelegenheiten, Anerkennung von Prüfungen, Rückerstattung des Studienbeitrages, Nachbesetzung von Planstellen, Schülerfreifahrt, ...</p> <p>Unterthema </p> <p>Verträge nach § 27, Ausschreibungen (PA II), Entinventarisierung, in den Ferien, Protokolle, ...</p>
Feedback an johann.grabner@inso.tuwien.ac.at Silk-Icons von famfamfam	

Abbildung 20: Startseite mit schematischen Layout

Abbildung 21: Startseite der Applikation

In Abbildung 22 steht der Aufbau der Eingabevervollständigung im Mittelpunkt. Passend zur Eingabe werden Schlagworte vorgeschlagen und unter dem Eingabefeld angezeigt. Unter der fett formatierten Übereinstimmung mit dem Suchbegriff werden weiterführende Informationen zum Schlagwort angeführt. Darunter fällt der Überbegriff und abhängig von der Übereinstimmung das Kürzel beziehungsweise der Titel. Die zusätzlich gezeigten Informationen dienen einerseits zur eindeutigen Identifikation des Schlagworts und machen andererseits neue Benutzer mit den Daten vertraut.

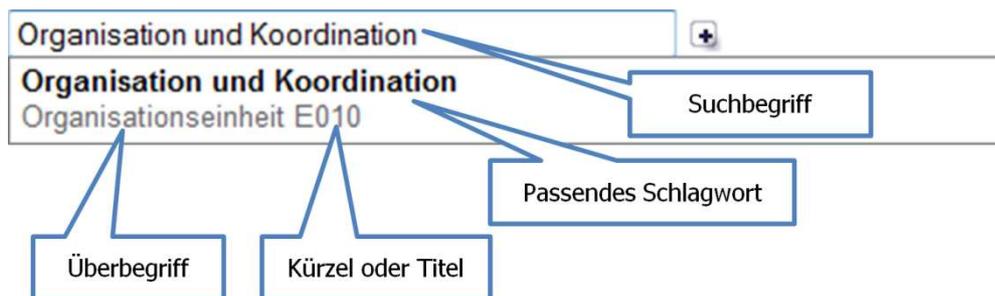


Abbildung 22: Angezeigte Informationen in der Eingabevervollständigung

Abbildung 23 liefert ein Beispiel aus der Anwendung für die Begleitung der Benutzereingabe durch Empfehlungen der Eingabevervollständigung. Die Vorschläge werden laufend an die Veränderungen des Suchbegriffs angepasst.

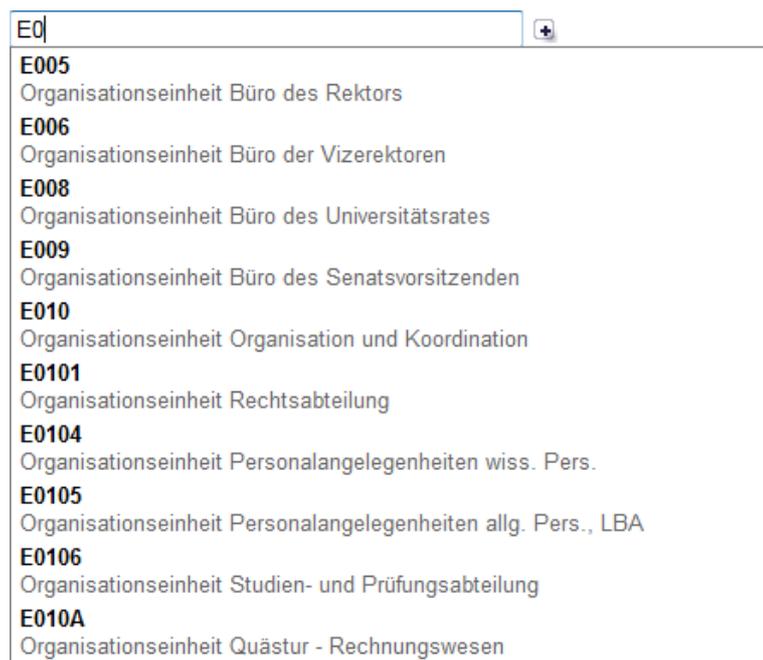


Abbildung 23: Beispiel für Vervollständigung während der Eingabe

Die Abbildungen 19 und 20 zeigen, wie eine Liste von Schlagworten über die Benutzerschnittstelle der Applikation verwaltet wird. Diese Listen kommen bei der Eingabe von Suchbegriffen und der Vergabe von Schlagworten zum Einsatz. Nach der Auswahl eines Schlagworts mit Hilfe der Eingabevervollständigung wird es per Mausklick auf das Symbol „+“ oder per Eingabetaste zur Liste hinzugefügt.

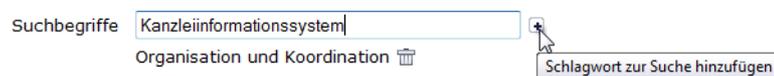


Abbildung 24: Schlagwort einer Liste hinzufügen

Mit einem Mausklick auf das hinten angestellte Symbol „☒“ lässt sich ein Element aus der Liste wieder entfernen.

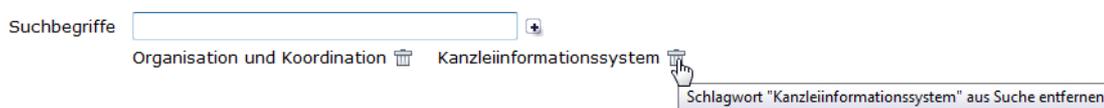


Abbildung 25: Schlagwort aus Liste entfernen

Bei der Erfassung eines neuen Schriftstücks wird von dem in Abbildung 26 abgebildeten Formular ausgegangen. Der Benutzer wird bei der Zuweisung der Schlagworte und des Ablagebereichs laufend durch Vorschläge vom System unterstützt. Dies bedeutet, dass bei Änderung der in Bezug stehenden Formulardaten auch die Vorschläge aktualisiert werden. Anstatt die Eingabe über die Felder selbst durchzuführen, kann der Benutzer jederzeit einen Vorschlag beziehungsweise einen Teil daraus übernehmen. Die vollständige Übernahme eines Vorschlags geschieht über das Symbol „+“.

Schriftstück erfassen

Hinweis: Pflichtfelder sind mit einem Stern (*) gekennzeichnet.

Inhalt

Beschlagwortung*

Oft zusammen vergebene Schlagwörter

- Institut für Rechnergestützte Automation (E183)
- Designing Comfort (DECO) , Institutsgebäude (GE)
- Establishing Security (ESSE) , Institutsgebäude (GE)

Ablagebereich* Thema

Vorgeschlagene Ablagebereiche
keine Empfehlungen

Dokumente

[weiteres Dokument anhängen](#)

Abbildung 26: Formular für ein neues Schriftstück

Nachdem zumindest ein Schlagwort vergeben und ein Ablagebereich gewählt wurde, kann das Schriftstück erfasst werden. Dem System unbekannte Schlagwörter werden automatisch unter dem allgemeinen Überbegriff „Thema“ angelegt und können im Nachhinein editiert werden. Aus dem gewählten Ablagebereich wird die eindeutige Zeichenfolge zur Identifikation abgeleitet. Die Angabe von hochzuladenden Dokumenten gestaltet sich als optional. Abbildung 27 beinhaltet ein Beispiel für ein ausgefülltes Formular zur Neuerfassung eines Schriftstücks.

Schriftstück erfassen

Hinweis: Pflichtfelder sind mit einem Stern (*) gekennzeichnet.

Inhalt

Beschlagwortung* Institut für Rechnergestützte Automation (E183)
Building Systems Diplomarbeit

Oft zusammen vergebene Schlagwörter

- Establishing Security (ESSE) , Institutsgebäude (GE)
- Designing Comfort (DECO) , Institutsgebäude (GE)
- Institut für Rechnergestützte Automation (E183)

Ablagebereich* Thema

Vorgeschlagene Ablagebereiche

- 183.00: Institut für Rechnergestützte Automation

Dokumente

[weiteres Dokument anhängen](#)

Abbildung 27: Ausgefülltes Formular für ein neues Schriftstück

Nach der Erfassung werden weitere anwendungsspezifische Eingabefelder für die Verwaltung des Schriftstücks angezeigt. Abbildung 28 hält die Bearbeitungsansicht für den

Datensatz aus der vorigen Abbildung fest. Im Gegensatz zum gewählten Ablagebereich bleibt die Beschlagwortung veränderbar. Vorhergehende und nachfolgende Schriftstücke können über die Angabe der Geschäftszahl referenziert werden. Im System als Schlagworte bekannte Dokumentenarten und Aufbewahrungsorte sind dem Schriftstück zuordenbar.

Des Weiteren können mit dem Schriftstück durchgeführte Bearbeitungsschritte der Kanzlei vermerkt werden. Nach einem Klick auf das Symbol „“ werden Links zu älteren Versionen des hochgeladenen Dokuments aufgelistet. Bei Bedarf kann vom Benutzer nach der Auflistung eine neue Version des Dokuments hochgeladen werden.

Schriftstück 183.00/004/2010

Daten	
Geschäftszahl	183.00/004/2010
Erfassungsdatum	13.07.2010
Vorzahlen	<input data-bbox="518 869 534 891" type="button" value="+"/>
Nachzahlen	<input data-bbox="518 898 534 920" type="button" value="+"/>
Dokumentenart	<input data-bbox="518 927 922 949" type="text" value=""/>
Aufbewahrungsort	Kanzlei <input data-bbox="598 965 614 987" type="button" value="v"/>
Bearbeitungsschritte	neuer Bearbeitungsschritt <input data-bbox="742 994 885 1016" type="text" value=""/> <input data-bbox="901 1003 917 1025" type="button" value="+"/>

Inhalt	
Beschlagwortung*	<input data-bbox="518 1099 901 1122" type="text" value=""/> <input data-bbox="917 1108 933 1131" type="button" value="+"/> Building Systems (BUSY) <input data-bbox="1173 1108 1189 1131" type="button" value="x"/> Diplomarbeit <input data-bbox="1300 1108 1316 1131" type="button" value="x"/> Empfehlungen anzeigen
Ablagebereich*	183.00: Institut für Rechnergestützte Automation
Dokumente	13.07.2010, 17:40: Diplomarbeit.pdf <input data-bbox="837 1243 853 1265" type="button" value="i"/> weiteres Dokument anhängen

Abbildung 28: Formular zur Bearbeitung eines Schriftstücks

Falls ein Schriftstück zur Bearbeitung aufgerufen wird, das aus dem Altsystem übernommen wurde, wird unter dem Eingabeareal „Inhalt“ aus Abbildung 28 der obere Teil aus Abbildung 29 angezeigt. Per Mausklick kann der zugehörige Datensatz aus dem Altsystem angezeigt werden.



Abbildung 29: Optionale Anzeige des Datensatzes aus dem Altsystem

Die beiden nächsten Abbildungen befassen sich mit der Eingabe von Schlagworten. Bei der manuellen Erfassung eines neuen Schlagwortes durch den Benutzer sind die Angabe eines Titels und die Wahl eines Überbegriffs zwingend erforderlich. Abbildung 30 zeigt den Zustand des Eingabeformulars unmittelbar vor der Wahl des Überbegriffs.

Abbildung 30: Wahl des Überbegriffs eines neuen Schlagwortes

Sobald sich der Anwender für einen Überbegriff entschieden hat, bietet das System die Möglichkeit einen Bezug zu bereits bestehenden Schlagworten herzustellen. Dies ist in Abbildung 31 ersichtlich, in der das begonnene Beispiel zu Ende geführt wird.

Schlagwort erfassen

Hinweis: Pflichtfelder sind mit einem Stern (*) gekennzeichnet.

Daten

Titel*

Überbegriff*

Kürzel

Weitere Titel
[weitere Titel hinzufügen](#)

Zusätzliche Angaben

Institut

Abbildung 31: Ausgefülltes Formular für ein neues Schlagwort

Im System erfasste Schlagwörter können nachträglich editiert werden. In Abbildung 32 ist das Formular zur Bearbeitung eines Schlagworts zu sehen. Bei der Bearbeitung ist es möglich, das betroffene Schlagwort stillzulegen. Stillgelegte Schlagwörter werden standardmäßig bei Textfeldern mit Eingabevervollständigung nicht mehr angezeigt.

Schlagwort "Building Systems"

Hinweis: Pflichtfelder sind mit einem Stern (*) gekennzeichnet.

Daten

Titel*

Überbegriff*

Kürzel

Weitere Titel
[weitere Titel hinzufügen](#)

Verknüpfung

Institut

Verwendung

Status Schlagwort stillgelegt
Stillgelegte Schlagwörter werden standardmäßig bei der Suche ausgeblendet.

Abbildung 32: Formular zur Bearbeitung eines Schlagworts

Die Ansicht in Abbildung 33 wird zum Auffinden von Akten nach Bearbeitungsschritten verwendet. Nachdem die Art des Bearbeitungsschritts ausgewählt wurde, können weitere Übereinstimmungskriterien angegeben werden. Vertrauliche Daten innerhalb der Suchergebnisse wurden in der Abbildung unkenntlich gemacht.

Bearbeitungsschritt

Datum Empfänger Frist

Suchergebnisse 1 bis 5 von 5

Bearbeitungsschritte pro Seite, nach sortiert

74808.01/001/2011
...
 05. Januar 2011, X : PA 1

70001.01/001/2011
...
 05. Januar 2011, X : PA 2

74809.01/001/2011
...
 05. Januar 2011, X : PA 1

74347.01/001/2011
...
 05. Januar 2011, X : PA 1

74604.01/001/2011
...
 05. Januar 2011, X : PA 2

Abbildung 33: Suche nach Bearbeitungsschritten

Mit dem Aufruf der in Abbildung 34 dargestellten Ansicht erhält der Benutzer eine Auflistung der momentan für Akten relevanten Fristen. Um abgelaufene Fristen optisch hervorzuheben, wird mit rotem Text gearbeitet.

Aktuelle Fristen (Stand 12.01.2011)

<input type="checkbox"/>	Geschäftszahl	Bearbeitungsschritt	Empfänger	Fristende	Status	Aktion
<input type="checkbox"/>	6210.00/010/2010	zugeteilt	<input type="text"/>	10.01.2011	Frist abgelaufen	✓
<input type="checkbox"/>	15005.00/002/2010	expediert	<input type="text"/>	25.01.2011	Frist läuft	✓

Status "Frist erfüllt"

Abbildung 34: Übersicht über laufende Fristen

Die Umlagerung einer Menge von Akten kann mit dem in Abbildung 35 gezeigten Formular durchgeführt werden. Der Benutzer kann mit Kriterien spezifizieren welche Akten umzulagern sind. Bei der Umlagerung wird der Aufbewahrungsort der Akten umgesetzt.

Schriftstücke umlagern

Umlagerung

Art Archivierung (von der Kanzlei in das Archiv)

Kriterien

Geschäftszahl . / /

Erfassungszeitraum von , bis

Abbildung 35: Formular zur Umlagerung von Akten

In Abbildung 36 ist das Formular zur Erstellung von Berichten zu sehen. Nach der Auswahl eines Berichtstyps und dem Ausgabeformat kann der Benutzer weitere Parameter für die Auswertung angeben.

Bericht erstellen

Bericht

Typ*

Ausgabeformat* Webseite Excel PDF

Abfrage

Zeitraum Beginn Ende

Sortierung

Sortieren nach ,

Dann nach ,

[weitere Sortierebene hinzufügen](#)

Abbildung 36: Formular zur Erstellung von Berichten

Benutzer können persönliche und gemeinsame Notizen anlegen, wie in Abbildung 37 gezeigt wird. Eine persönliche Notiz ist nur für den Verfasser sichtbar, während gemeinsame Notizen für alle anderen Benutzer zu sehen sind.

Notizen

Persönliche Notizen

Notiz 10.01.2011, 14:32

Gemeinsame Notizen

Notiz 10.01.2011, 14:38
 Johann Grabner

Abbildung 37: Persönliche und gemeinsame Notizen

Alle Suchvorgänge und Benutzeraktionen auf Akten werden vom System protokolliert. In der Ansicht aus Abbildung 38 kann der Benutzer die vom System über ihn gesammelten Informationen einsehen. Gespeicherte Suchvorgänge können vom Benutzer gelöscht werden.

Protokoll

Statistik

Zuletzt aufgerufen 15005.00/002/2010
 Schlagwörter: [Anlageninformationen - TV - 200, 200-Punkte-Gewinnvergleich](#)
 15008.00/001/2011
 Schlagwörter: [In-Flussnahme der letzten Werte in der Formel \(Schrittweise\)](#)
 15008.00/080/2010
 Schlagwörter: [In-Flussnahme, Wechselkursrate \(Kauf\)](#)
[weitere Schriftstücke anzeigen](#)

Zuletzt erfasst 20054.00/003/2011
 Schlagwörter: [Anwendung der Buchführung, Lohn Eintrag](#)
 13050.00/001/2011
 Schlagwörter: [Interaktive Multimediale, Anlagen für die 11. Monatsrechnung](#)
 49610.54/001/2011
 Schlagwörter: [Lernhilfe, Programmier, IT](#)
[weitere Schriftstücke anzeigen](#)

Zuletzt aktualisiert 6210.00/010/2010
 Schlagwörter: [IT - 200, 200-Punkte-Gewinnvergleich](#)
 11018.00/001/2011
 Schlagwörter: [Anlageninformationen, Anlagen für die 11. Monatsrechnung](#)
 30002.00/013/2010
 Schlagwörter: [Anlageninformationen, Anlagen für die 11. Monatsrechnung](#)
[weitere Schriftstücke anzeigen](#)

Suchanfragen von 05.01.2011 bis 12.01.2011

Suche Mittwoch, 12. Januar 2011, 13:21 Uhr
 Aktennummer: *.183.00/*/2010

Suche Mittwoch, 12. Januar 2011, 13:18 Uhr
 Schlagwörter: Juli 2010

Suche Mittwoch, 12. Januar 2011, 13:11 Uhr
 Schlagwörter: System %arbeit

Suche Mittwoch, 12. Januar 2011, 12:52 Uhr
 Schlagwörter: Kanzlei

Suche Dienstag, 11. Januar 2011, 15:04 Uhr
 Schlagwörter: Forschungsgruppe

Zurück 1 2 3 Vorwärts

Abbildung 38: Protokoll der Benutzeraktionen

6.3 Suchfunktion

In diesem Abschnitt wird die Suchfunktion der Anwendung anhand von Beispielen erklärt. In einer Serie von Abbildungen wird jeweils ein Suchkriterium mit einem passenden Ergebnis angeführt. Die roten Markierungen in den Abbildungen kennzeichnen den Grund für die Übereinstimmung mit dem Suchkriterium. Die angeführten Arten von Suchkriterien können beliebig miteinander kombiniert werden und schließen sich nicht bei der gemeinsamen Verwendung in einer Suchanfrage aus.

Wird ein Schlagwort bei der Suche verwendet, enthält das Suchergebnis die Schriftstücke an die dieses oder ein dem Suchbegriff untergeordnetes Schlagwort verteilt wurde. In Abbildung 39 wird dieses Verhalten anhand eines Beispiels demonstriert. Das Suchergebnis resultiert aus der Tatsache, dass die Forschungsgruppe „Building Systems“ ein Teil des gesuchten Instituts ist.

Suchbegriffe 

Fakultät für Informatik  Institut für Rechnergestützte Automation (E183) 

Geschäftszahl . / / 

↓ Suche ↓

183.00/004/2010
 Schlagwörter: Building Systems (BUSY), Diplomarbeit
 erfasst am Dienstag, 13. Juli 2010, 17:15 Uhr von Johann Grabner
 Aufbewahrungsort: Kanzlei

Abbildung 39: Suche nach Schlagworten

Bei der Suche können auch Überbegriffe verwendet werden, wie in Abbildung 40 gezeigt wird. In so einem Fall landen alle Schriftstücke im Ergebnis, an die ein Schlagwort mit dem gesuchten Überbegriff vergeben wurde.

Suchbegriffe 

Forschungsgruppe 

Geschäftszahl . / / 

↓ Suche ↓

183.00/004/2010
 Schlagwörter: Building Systems (BUSY), Diplomarbeit
 erfasst am Dienstag, 13. Juli 2010, 17:15 Uhr von Johann Grabner
 Aufbewahrungsort: Kanzlei

Abbildung 40: Suche nach Überbegriff

Spezielle Schlagworte, die einem Schriftstück als Attribute zugeteilt werden, sind ebenfalls in der Suche einsetzbar. Davon sind alle Schlagworte mit den Überbegriffen „Auf-

bewahrungsort“, „Dokumentenart“ und „Ablagebereich“ betroffen. Für Abbildung 41 wurde als Beispiel eine Suche nach dem Aufbewahrungsort „Kanzlei“ durchgeführt.

Suchbegriffe 

Kanzlei 

Geschäftszahl . / / 

↓ Suche ↓

183.00/004/2010
 Schlagwörter: Building Systems (BUSY), Diplomarbeit
 erfasst am Dienstag, 13. Juli 2010, 17:15 Uhr von Johann Grabner
 Aufbewahrungsort: Kanzlei

Abbildung 41: Suche nach Schriftstückattributen

Abgesehen von der Suche mit Schlagworten wird auch eine konventionelle Volltextsuche unterstützt. Abbildung 42 beinhaltet die Suche nach in der Beschlagwortung vorkommenden Textteilen. Der Platzhalter „%“ repräsentiert einen beliebigen Beginn des gesuchten Words.

Suchbegriffe 

System %arbeit 

Geschäftszahl . / / 

↓ Suche ↓

183.00/004/2010
 Schlagwörter: Building Systems (BUSY), Diplomarbeit
 erfasst am Dienstag, 13. Juli 2010, 17:15 Uhr von Johann Grabner
 Aufbewahrungsort: Kanzlei

Abbildung 42: Suche nach Text

Bei der Suche mit einem Zeitraum wird das Ergebnis auf alle Schriftstücke eingeschränkt, deren Erfassung in die angegebene Zeitspanne fällt. Das System erkennt Zeitangaben in regional gebräuchlichen Formaten. Ein Beispiel ist in Abbildung 43 zu finden.

Suchbegriffe 

Juli 2010 

Geschäftszahl . / / 

↓ Suche ↓

183.00/004/2010
 Schlagwörter: Building Systems (BUSY), Diplomarbeit
 erfasst am Dienstag, 13. Juli 2010, 17:15 Uhr von Johann Grabner
 Aufbewahrungsort: Kanzlei

Abbildung 43: Suche nach Zeitraum

Wird nach einer Geschäftszahl gesucht, setzt sich das Ergebnis aus allen Schriftstücken zusammen, deren Geschäftszahl mit den angegebenen Teilen der gesuchten Geschäftszahl übereinstimmt. Dies ist in Abbildung 44 erkennbar.

Suchbegriffe

Geschäftszahl . / /

↓ Suche ↓

183.00/004/2010
 Schlagwörter: Building Systems (BUSY), Diplomarbeit
 erfasst am Dienstag, 13. Juli 2010, 17:15 Uhr von Johann Grabner
 Aufbewahrungsort: Kanzlei

Abbildung 44: Suche nach Geschäftszahl

Abbildung 45 zeigt zum Abschluss, welche Optionen die Eingabevervollständigung für Schlagworte bei der Suche bietet. Standardmäßig werden stillgelegte Schlagworte und aus dem Altsystem übernommene Ablagebereiche nicht in den Resultaten der Eingabevervollständigung angezeigt. Migrierte Ablagebereiche werden normalerweise ausgeblendet, weil der Zugriff im neuen System primär über vergebene Schlagworte erfolgt. Auf Wunsch kann dieses Verhalten vom Benutzer angepasst werden.

Suchbegriffe

Stillgelegte Schlagworte anzeigen
 Ablagebereiche anzeigen

Geschäftszahl . / /

Suche [Optionen verbergen](#)

Abbildung 45: Optionen der Eingabevervollständigung bei der Suche

6.4 Optimierung

Zur Gewährleistung der Systemfunktion für das Datenvolumen der Kanzlei wurden Optimierungsmaßnahmen notwendig. Bisher sind im Kanzleinformationssystem über 200000 Schriftstücke eingetragen worden. Jährlich ist mit dem Anwachsen des Datenbestands um etwa 10000 Schriftstücke zu rechnen. Naiv formulierte Datenbankabfragen führen in Anbetracht des semantischen Graphs und der vorhandenen Datenmenge schnell zu inakzeptablen Ausführungszeiten. Das durch die Optimierung verfolgte Ziel ist, die Antwortzeiten der Web-Applikation im tolerierbaren Rahmen zu halten. Als tolerierbarer Zeitrahmen wurden für den Normalfall maximal zwei Sekunden angenommen. Der Schlüssel zur Erreichung der gewünschten Systemleistung liegt in den Maßnahmen der folgenden Aufzählung. Die Theorie zur ersten Maßnahme der Aufzählung wird in [KeEi06] behandelt. Die beiden anderen Maßnahmen wurden aus der Entwicklung des Systems geschlossen.

- *Daten indizieren*

Eine ordnungsgemäße Indizierung ist für einen effizienten Datenbankzugriff bei einem größeren Datenvolumen zwingend erforderlich. Wird eine Spalte einer Datenbanktabelle bei einer Abfrage als Sortierkriterium oder in einer Bedingung verwendet, bietet sich meist an, die betroffene Spalte zu indizieren. Wurde keine Indizierung vorgenommen, muss die Datenbank bei der Ausführung solcher Abfragen die ganze Tabelle durchsuchen. Bei Tabellen mit kleinerem Datenvolumen fällt der Aufwand des gesamten Lesens gegenüber dem Auffinden der gewünschten Datensätze über den Index noch nicht ins Gewicht. Der Wirkungsgrad der Indizierung erhöht sich jedoch schnell mit zunehmender Datenmenge. Die Wahl der richtigen Indizierung ist ebenfalls ausschlaggebend. Aus Effizienzgründen sind bei Datenfeldern zur Speicherung von Text datenbankspezifische Funktionen zur Volltextindizierung einzusetzen, wenn Abfragekriterien die Übereinstimmung einzelner oder mehrere Worte innerhalb eines Texts fordern. Sind mehrere Spalten gemeinsam der Gegenstand bedingter Abfragen mit eventuell sortiertem Ergebnis, ist ein Index in Erwägung zu ziehen, der diese Spalten als Gruppe umfasst.
- *Abfragen vereinfachen*

Die Reduzierung der Komplexität einer Datenbankabfrage wirkt sich dementsprechend positiv auf die benötigte Ausführungszeit aus. Im Idealfall wird die Abfrage von aus mehreren Tabellen stammenden Daten gänzlich vermieden. Die Replikation und Aggregation dieser Daten in einer einzelnen Tabelle mit Indizes für den optimierten Zugriff ermöglicht Abfragen mit reduzierter Komplexität. Des Weiteren können bei der Replikation Datenstrukturen, deren Abfrage mit hohem Aufwand verbunden ist, in eine Struktur transformiert werden, die eine effizientere Handhabung erlaubt. Die Zusammenfassung mehrerer Spalten in einem einzelnen Datenfeld vermindert nebenbei die Anzahl der betroffenen Felder in den Bedingungen der Abfragen. Erlaubt die Komplexität der Daten eine Transformation in numerische Werte nicht, kann die Ablage in einem Textformat erfolgen. Um von der Leistungsfähigkeit der Volltextsuche Gebrauch machen zu können, sollten mit Trennzeichen begrenzte Zeichenketten zur Abbildung der Information verwendet werden.
- *Abfragestrategien wählen*

Kompakte Abfrageergebnisse sind wichtig für die Leistung der Applikation. Ergebnismengen mit hoher Datensatzanzahl erhöhen den technischen Aufwand bei der Handhabung und Auswertung. Die maximale Anzahl von retournierten Datensätzen sollte deswegen bei jeder Abfrage zumindest limitiert werden. Eine Limitation der Ergebnisse führt ohne Sortierung nach absteigender Relevanz eventuell zum Verlust gewünschter Informationen. Mit Hilfe der Sammlung sta-

tistischer Werte über die Häufigkeit von Elementen wird es in der Anwendungslogik der Applikation möglich, eine passende Abfragestrategie zu wählen. Je nach dem zu erwartenden Umfang der Ergebnismenge wird die Genauigkeit und Detailliertheit der von der Applikation durchgeführten Abfragen skaliert. Eine Verfeinerung der Abfragen bei einer kleinen Ergebnismenge verbessert die Qualität des Ergebnisses. Im Gegenzug kommt bei umfangreichen Ergebnissen eine Vereinfachung der Abfragen der Systemleistung zugute.

Die redundante Haltung von Daten macht die Weitergabe von Änderungen der Originaldaten an den replizierten Datenbestand notwendig. Dies betrifft die datenbankinterne Indizierung, für die Abfrage transformierte Daten und gesammelte Statistiken. Die durch diese Maßnahmen bei der Abfrage erzielte Leistungssteigerung geht einher mit einem gesteigerten Aufwand beim Schreiben und Ändern von Datensätzen. In dieser Anwendung ist dieses Verhalten gewünscht, weil Datensätze bei weitem häufiger gelesen als geschrieben werden.

Zwecks Erreichung der gewünschten Leistungsvorgaben wurden die Tabellen aus Abbildung 46 in der Applikation eingeführt. Um die Funktion der eingeführten Tabellen zu erklären, werden diese in den nächsten Absätzen genauer betrachtet.

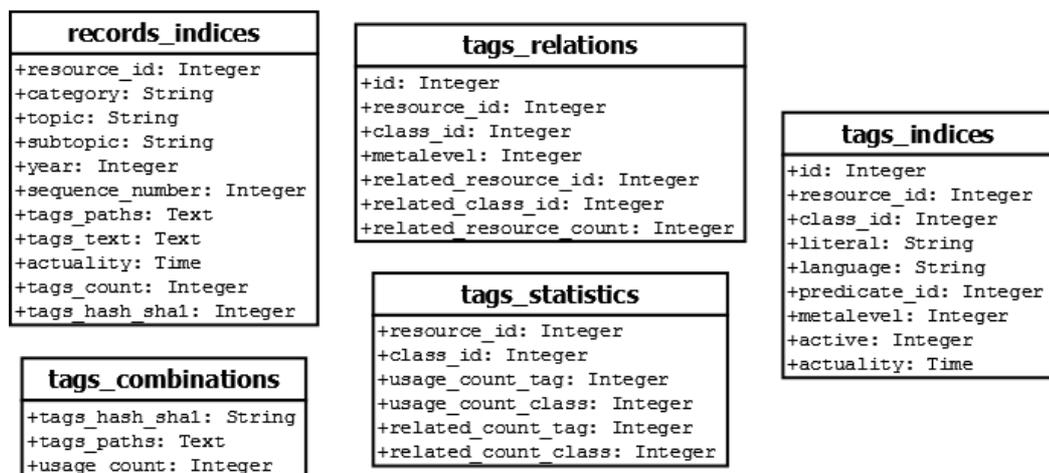


Abbildung 46: Klassendiagramm „Optimierungstabellen“

Die Tabelle „records_indices“ ermöglicht die effiziente Abfrage von Schriftstücken nach deren Attributen und zugeteilten Schlagworten. Ein Großteil der Spalten wurde aus der Schriftstücktabelle repliziert.

Die in der folgenden Aufzählung angeführten Spalten stellen Ausnahmen dar, weil sie aus dem Datenbestand errechnet werden müssen:

- „tags_paths“

Um bei der Suche nach einem vorhandenen Schlagwort semantisch untergeordnete Elemente effizient ins Ergebnis mit einbeziehen zu können, werden die ans Schriftstück vergebene Schlagworte in textuelle Pfadangaben umgewandelt. Der Pfad gibt an, welche Position ein Schlagwort in der semantischen Hierarchie einnimmt, weil alle übergeordneten Schlagworte inkludiert werden. Die durch einen Schrägstrich getrennten Primärschlüssel der mit den Schlagworten verbundenen Ressourcen ergeben aneinandergereiht die Pfadangabe als Zeichenkette. Die Primärschlüssel werden innerhalb des Pfads absteigend nach der Anzahl der Schlagworte in der Hierarchie zwischen dem referenzierten und gefragten Schlagwort gereiht. So können per Textsuche partielle oder vollständige Pfadangaben verwendet werden, um passende Schriftstücke aufzufinden.

- „tags_text“
Alle Bezeichner und Titel der an das Schriftstück vergebenen Schlagworte werden in dieser Spalte zu einer einzelnen Zeichenkette zusammengefügt. Auf diesem Weg wird eine konventionelle Suche nach Textteilen ermöglicht.
- „tags_count“
In dieser Spalte wird die Anzahl der dem Schriftstück zugeordneten Schlagworte abgelegt.
- „tags_hash_sha1“
Zur effizienten eindeutigen Identifikation der vergebenen Schlagwortkombination wird ein Hashwert berechnet, der in dieser Spalte gespeichert wird. Als Eingabewert für die Hashfunktion wird eine Zeichenkette aus den aufsteigend sortierten Primärschlüsseln der Ressourcen gebildet. In dieser Zeichenkette wird ein Beistrich als Trennzeichen zwischen den einzelnen Zahlenwerten verwendet.

Über die Tabelle „tags_combinations“ wird eine Rangliste von Schlagwortkombinationen nach Verwendungshäufigkeit abgebildet. Hierzu werden die Spalten „tags_hash_sha1“ und „tags_paths“ aus der Tabelle „records_indices“ übernommen und zur Speicherung einer Anzahl um die Spalte „usage_count“ ergänzt. Diese Anzahl repräsentiert, an wie viele Schriftstücke dieselbe Kombination von Schlagworten vergeben wurde.

Ein Datensatz der Tabelle „tags_relations“ beschreibt entweder, wie oft zwei bestimmte Schlagworte gemeinsam an ein Schriftstück vergeben wurden, oder wie oft die unter einem bestimmten Überbegriff fallenden Schlagworte mit den zu einem andere Überbegriff zugehörigen Schlagworten verwendet wurden. Ist der Wert von „metalevel“ gleich null trifft ersteres zu, während bei einem Wert von eins zweiteres der Fall ist.

Das Datenfeld „metalevel“ gibt somit an, auf welche Metaebene sich der Datensatz bezieht. Die Anzahl der aufgetretenen Beziehungen wird in der Spalte „related_resource_count“ abgelegt. In der Tabelle „tags_statistics“ werden zu jedem Schlagwort folgende Kennzahlen abgelegt:

- „usage_count_tag“
An wie viele Schriftstücke das Schlagwort verteilt wurde, wird von diesem Wert angegeben.
- „usage_count_class“
Dieser Wert repräsentiert die Anzahl der an Schriftstücke verteilten Schlagwörter, die das Schlagwort als Überbegriff haben.
- „related_count_tag“
Mit dieser Anzahl wird festgehalten mit wie vielen unterschiedlichen Schlagworten das Schlagwort gemeinsam in allen Beschlagwortungen vorgekommen ist.
- „related_count_class“
Die Anzahl der unterschiedlichen Überbegriffe der mit dem Schlagwort gemeinsam in allen Beschlagwortungen vergebenen Schlagwörter wird durch diese Spalte festgehalten.

Die Daten der beiden zuletzt aufgezählten Tabellen „tags_relations“ und „tags_statistics“ bilden eine wichtige Grundlage für die Berechnung von Navigationsvorschlägen zur Einschränkung des Suchergebnisses. Nach den vorhandenen Häufigkeitswerten wird entschieden, mit welcher Detailliertheit die Navigationsvorschläge erstellt werden können ohne unverträglich für die Systemleistung zu werden.

Das Ziel der letzten Optimierungstabelle „tags_indices“ ähnelt im Prinzip der ersten beschriebenen Tabelle „records_indices“. Anstatt von Schriftstücken sollen mit Hilfe dieser Tabelle Schlagworte nach einem Kriterium effizient auffindbar sein. Dies spielt eine wichtige Rolle für die Erstellung von Vorschlägen zur Eingabevervollständigung. In der textindizierten Spalte „literal“ wird die Zeichenkette abgelegt unter der das Schlagwort aufgefunden werden soll. Existieren mehrere Bezeichnungen für dasselbe Schlagwort, wird pro Bezeichnung ein eigener Datensatz in der Tabelle abgelegt.

7 Zusammenfassung

Die Miteinbeziehung der Semantik bei der Verarbeitung von Informationen eröffnet Softwaresystemen neue Möglichkeiten. Tätigkeiten, die bisher dem Anwender vorbehalten waren, weil sie ein Verständnis von Zusammenhängen und Bedeutungen erfordern, können von einer Maschine übernommen werden. Im Idealfall wird der Anwender auf diesem Weg von nicht relevanten Details technischer oder fachlicher Natur entlastet und kann sich in Folge besser auf die wesentlichen Teile seiner Arbeit konzentrieren. Kann das Ziel einer Produktivitätssteigerung erreicht werden, steht dem Einsatz einer semantischen Applikation im täglichen Geschäftsleben nichts mehr im Weg.

Im Bereich des Dokumentenmanagements kann in sehr hohem Ausmaß vom Einsatz semantischer Technologien profitiert werden. Dokumente schnell auffinden und einfach ablegen zu können sind Voraussetzungen für eine effiziente Verwaltung. In dieser Diplomarbeit wurde anhand der Entwicklung eines Prototyps gezeigt, welchen Nutzen semantische Technologien bei der Suche und Ablage von Dokumenten mit sich bringen können.

Das entwickelte System bietet Empfehlungen zur Vervollständigung von eingegebenen Suchbegriffen an und versorgt die Anwender während der Suche mit inhaltlich relevanten Vorschlägen zur weiteren Einschränkung der Ergebnismenge. Die einfache Verwaltung des Suchkriteriums durch Hinzufügen oder Entfernen einzelner Suchbegriffe ermöglicht in Kombination mit den Vorschlägen der Navigation ein Stöbern im Datenbestand. Diese Art zu suchen ist sehr vorteilhaft für Personen, die nicht mit der Ablage oder dem Anwendungsbereich vertraut sind. Bei der Vergabe von Schlagworten helfen inhaltliche Vorschläge, die gegen die Ablage eines Themas unter unterschiedlichen Begriffen vorbeugen.

Auf den ersten Blick wurden die angestrebten Ziele vom entwickelten Prototyp erreicht. Das System berücksichtigt bei der Suche und Ablage die Semantik von Begriffen, Dokumente können elektronisch abgelegt werden und die Limitationen des Kanzleiformationssystems wurden aufgehoben. Um eine präzise Aussage über die Effizienz des Systems tätigen zu können, ist eine weiterführende wissenschaftliche Auseinandersetzung mit dem Ergebnis notwendig.

Durch den gewählten Entwicklungsprozess wurde ein sehr gutes Kommunikationsverhältnis zu den Anwendern aufgebaut. Teilweise wurden Funktionen umgesetzt, die nicht zwingend für den Prototyp erforderlich waren. Als Konsequenz werden für zukünftige agile Entwicklungsprozesse eine verbesserte Einbindung des Projektmanagements und eine erhöhte Sorgfalt bei der Abgrenzung verfolgt. Durch eine gezielte Unterstützung des Projektmanagements können die Grenzen von Prototypen besser gewahrt werden.

In der Entwicklung konnten die ersten Ergebnisse nach kurzer Zeit vorgewiesen werden. Das Framework „Ruby on Rails“ eignet sich in Verbindung mit der „Yahoo! User Interface Library“ für die kurzfristige Gestaltung eines Prototypen. Solange man die Konventionen des Frameworks befolgt, können bei der Entwicklung von Web-Applikationen häufig wiederkehrende Aufgaben einfach gelöst werden. Bei komplexeren Aufgabenstellungen wandert dieser Vorteil sehr schnell in den Hintergrund. Bei der Formulierung komplizierter Datenbankabfragen stößt man schnell an die Grenzen des Frameworks.

In der Praxis müssen sich semantische Applikation in betrieblichen Organisationen erst durchsetzen. Im Rahmen dieses Projekts wurden Erkenntnisse gewonnen, die den zögerlichen wirtschaftlichen Einsatz semantischer Applikation erklären. Als erster Punkt wird das Fehlen semantischer Relationen in den bestehenden Datenständen angeführt. Für den Einsatz in einer semantischen Applikation müssen vorhandenen Datenmengen aufbereitet und mit zusätzlichen Beziehungen angereichert werden. Unabhängig davon, ob dieser Schritt von Menschen oder Maschinen durchgeführt wird, ist ein laufender Prozess zur Sicherstellung der Datenqualität erforderlich, um einen längerfristigen Nutzen erzielen zu können. Bei umfangreichen Datenbeständen ist dieser Prozess sehr aufwendig und kann möglicherweise sehr viele Probleme ans Tageslicht bringen. Wenn semantische Applikationen für die Benutzer einen Mehraufwand bei der Spezifikation von Relationen gegenüber der bisherigen Methode der Datenverarbeitung mit sich bringen, ist die Akzeptanz des Systems gefährdet. Des Weiteren wird Entwicklern von semantischen Technologien mehr abverlangt. Die Verlagerung der Komplexität aus der täglichen Arbeit der Benutzer in das Softwaresystem führt zu komplizierten Algorithmen, deren Entwicklung ein detailreiches Wissen über den Anwendungsbereich erfordert. Nebenbei ist die Lernkurve für die Verwendung semantischer Technologien steiler als bei gewöhnlichen relationalen Datenbanken. Das Fehlen der notwendigen technischen Infrastruktur ist ein weiterer Grund für die mäßige Anwendung. Die in den meisten Organisationen verfügbaren Technologien bieten keine gezielte Unterstützung für eine semantische Informationsverarbeitung. Wird auf die vorhandenen Technologien aufgebaut, muss ein beträchtlicher Anteil der Entwicklungszeit zur Schaffung der Projektinfrastruktur in der Infrastruktur eingesetzt werden. Für die Einführung neuer Technologien in die Infrastruktur der Organisation müssen zusätzliche Praktiken erlernt werden.

Zum Abschluss werden einige auf diese Diplomarbeit aufbauende Tätigkeiten in Ausblick gestellt. Eine weitere wissenschaftliche Auseinandersetzung mit den vorgestellten Mechanismen zur Ablage und Suche von Dokumenten liegt nahe. Die Wirksamkeit des Systems kann im Vergleich mit anderen Ansätzen evaluiert werden. Einen anderen Ausgangspunkt für zukünftige Erweiterungen am System bietet das Modell zur Abbildung des Wissens aus dem Anwendungsbereich. Die Menge der im Modell unterstütz-

ten Arten von semantischen Relationen kann ausgedehnt werden, um bei der Wissensabbildung einen erhöhten Detailgrad zu ermöglichen. Für spätere Versionen des Systems kann ein Austausch von semantischen Informationen über mehrere Instanzen der Applikation ins Auge gefasst werden. Ebenfalls denkbar ist eine über mehrere Applikationsinstanzen verteilte Abfrage zur Aggregation von Wissen.

Bevor semantische Technologien im Berufsleben allgegenwärtig werden können, müssen geeignete Werkzeuge für die Entwicklung und Administration zur Verfügung stehen. Wird die Eintrittsschwelle durch eine allgemeine Verfügbarkeit und der entsprechenden Verbreitung von Wissen gesenkt, wird es möglich, bei einem gerechtfertigten Aufwand Nutzen aus der semantischen Datenverarbeitung zu ziehen. In Anbetracht der steigenden Menge von in Organisationen anfallenden Daten wird eine Berücksichtigung der Semantik in Zukunft notwendig werden, um bei der Verwaltung wirtschaftlich bleiben zu können.

Literaturverzeichnis

- [Ange05] Markus Angermeier. The huge cloud lens bubble map web2.0. <http://kosmar.de/archives/2005/11/11/the-huge-cloud-lens-bubble-map-web20/> (01.03.2011). 2005.
- [ArNe05] Jim Arlow and Ila Neustadt. UML 2 and the Unified Process: Practical Object-Oriented Analysis and Design. Addison-Wesley Professional, 2005.
- [Barn10] Carol M. Barnum. Usability Testing Essentials: Ready, Set...Test! Morgan Kaufmann, 2010.
- [BeHe01] Tim Berners-Lee, James Hendler, and Ora Lassila. The Semantic Web. Scientific American, 284:34-43, 2001.
- [BoBr99] Jon Bosak and Tim Bray. XML and the Second-Generation Web. Scientific American, 280:89-93, 1999.
- [Boeh06] Barry Boehm. A view of 20th and 21st century software engineering. In Proceedings of the 28th International Conference on Software Engineering. 2006.
- [BoHa09] David Booth, Hugo Haas, Francis McCabe, Eric Newcomer, Michael Champion, Chris Ferris, and David Orchard. Web Services Architecture. <http://www.w3.org/TR/2004/NOTE-ws-arch-20040211/> (01.03.2011), 2004.
- [BoJa05] Olha Bondarenko and Ruud Janssen. Documents at hand: Learning from paper to improve digital technologies. In Proceedings of the SIGCHI conference on Human factors in computing systems, 2005.
- [BrCa07] Karin Koogan Breitman, Marco Antonio Casanova, and Walter Truszkowski. Semantic Web: Concepts, Technologies and Applications. Springer London, 2007.
- [CaDa09] Sven Casteleyn, Florian Daniel, Peter Dolog, and Maristella Matera. Engineering Web Applications. Springer Berlin Heidelberg, 2009.

- [Cock04] Alistair Cockburn. *Crystal Clear: A Human-Powered Methodology for Small Teams*. Addison-Wesley Professional, 2004.
- [Deng06] Andreas R. Dengel. Six thousand words about multi-perspective personal document management. In *Enterprise Distributed Object Computing Conference Workshops, 2006. EDOCW '06. 10th IEEE International, 2006*.
- [EtMa06] Patrick Etcheverry, Christophe Marquesuzaà, and Sandrine Corbineau. Designing suited interactions for a document management system handling localized documents. In *Proceedings of the 24th annual ACM international conference on Design of communication, 2006*.
- [GaKr11] Fabien L. Gandon, Reto Krummenacher, Sung-Kook Han, and Ioan Toma. *Handbook of Semantic Web Technologies*. Springer Berlin Heidelberg, 2011.
- [GiSh10] Andreas Girgensohn, Frank Shipman, Francine Chen, and Lynn Wilcox. Docubrowse: faceted searching, browsing, and recommendations in an enterprise context. In *Proceedings of the 15th international conference on Intelligent user interfaces, 2010*.
- [Glog10] Boris Gloger. Scrum. *Informatik-Spektrum*, 33:195-200, 2010.
- [GöSc08] Klaus Götzer, Ralf Schmale, Berthold Maier, and Torsten Komke. *Dokumenten-Management: Informationen im Unternehmen effizient nutzen*. dpunkt Verlag, 2008.
- [GuSe02] Jürgen Gulbins, Markus Seyfried, and Hans Strack-Zimmermann. *Dokumenten-Management: Vom Imaging zum Business-Dokument*. Springer, 2002.
- [Hans10] Eckhart Hanser. *Agile Prozesse: Von XP über Scrum bis MAP*. Springer Berlin Heidelberg, 2010.
- [HeSr11] Sarah Henderson and Ananth Srinivasan. Filing, piling & structuring: Strategies for personal document management. In *Proceedings of the 2011 44th Hawaii International Conference on System Sciences, 2011*.
- [HiCo01] Jim Highsmith and Alistair Cockburn. Agile Software Development: The Business of Innovation. *Computer*, 34:120-127, 2001.

- [HrRu09] Peter Hruschka, Chris Rupp, and Gernot Starke. *Agility kompakt: Tipps für erfolgreiche Systementwicklung*. Spektrum Akademischer Verlag, 2009.
- [HuJa11] Elizabeth Hull, Ken Jackson, and Jeremy Dick. *Requirements Engineering*. Springer London, 2011.
- [KaPr06] Gerti Kappel, Birgit Pröll, Siegfried Reich, and Werner Retschitzegger. *Web Engineering: The Discipline of Systematic Development of Web Applications*. John Wiley & Sons, 2006.
- [KeEi06] Alfons Kemper and André Eickler. *Datenbanksysteme: Eine Einführung*. Oldenbourg Wissenschaftsverlag, 2006.
- [KoFu07] Kosuke Konishi, Naohiro Furukawa, and Hisashi Ikeda. Data model and architecture of a paper-digital document management system. In *Proceedings of the 2007 ACM symposium on Document engineering*, 2007.
- [Kruc95] Philippe Kruchten. The 4+1 view model of architecture. *IEEE Software*, 12:42-50, 1995.
- [MaCa10] Murad Magableh, Antonio Cau, Hussein Zedan, and Martin Ward. Towards a multilingual semantic folksonomy. In *Proceedings of the IADIS International Conferences Collaborative Technologies 2010 and Web Based Communities 2010*, 2010.
- [MeKu04] Erica Meena, Ashwani Kumar, and Laurent Romary. An extensible framework for efficient document management using RDF and OWL. In *Proceedings of the Workshop on NLP and XML (NLPXML-2004): RDF/RDFS and OWL in Language Technology*, 2004.
- [MeMo10] Emilia Mendes and Nile Mosley, editors. *Web Engineering*. Springer, 2010.
- [MiHa06] Russ Miles and Kim Hamilton. *Learning UML 2.0*. O'Reilly Media, 2006.
- [MiRo06] David E. Millard and Martin Ross. Web 2.0: Hypertext by any other name? In *Proceedings of the Seventeenth ACM Conference on Hypertext and Hypermedia*, 2006.

- [MoSa06] Enrico Motta and Marta Sabou. Next Generation Semantic Web Applications. In *The Semantic Web - ASWC 2006*. Springer, 2006.
- [Muru07] San Murugesan. Understanding web 2.0. *IT Professional*, 9:34-41, 2007.
- [PiFo08] Marlon E. Pierce, Geoffrey C. Fox, Joshua Rosen, Siddharth Maini, and Jong Y. Choi. Social networking for scientists using tagging and shared bookmarks: a web 2.0 application. In *Collaborative Technologies and Systems, 2008. CTS 2008. International Symposium on, 2008*.
- [RuCh08] Jeffrey Rubin, Dana Chisnell, and Jared Spool. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. John Wiley & Sons, 2008.
- [SaPr06] Vidyardaman Sankaranarayanan, Suranjan Pramanik, and Shambhu Upadhyaya. Detecting masquerading users in a document management system. In *Communications, 2006. ICC '06. IEEE International Conference on, 2006*.
- [Schw04] Ken Schwaber. *Agile Project Management with Scrum*. Microsoft Press, 2004.
- [ScRa06] Michael Schuster and Dieter Rappold. Social Semantic Software - was soziale Dynamik im Semantic Web auslöst. In *Semantic Web*. Springer, 2006.
- [Somm10] Ian Sommerville. *Software Engineering*. Pearson Education, 2010.
- [SpMo07] Lucia Specia and Enrico Motta. Integrating Folksonomies with the Semantic Web. In *The Semantic Web: Research and Applications*. Springer, 2007.
- [SuGe08] Li Sui, Gengchen Shi, Ping Song, and Xingyu Yuan. Design and implementation of ISO document management system. In *Proceedings of the 2008 International Conference on Computer Science and Software Engineering - Volume 02, 2008*.
- [VaPa09] Francisco Valverde and Oscar Pastor. Facing the technological challenges of web 2.0: A ria model-driven engineering approach. In *Web Information Systems Engineering - WISE 2009*. Springer Berlin / Heidelberg, 2009.

- [WiMe10] Hans Wilhelm Wieczorrek and Peter Mertens. Management von IT-Projekten: Von der Planung zur Realisierung. Springer, 2010.
- [ZuGr04] Wolfgang Zuser, Thomas Grechenig, and Monika Köhle. Software Engineering. Mit UML und dem Unified Process. Pearson Studium, 2004.

Anhang

Glossar

Kanzlei

Akt / Akte

Ein oder mehrere Schriftstücke werden in einem Büroordner physisch abgelegt. Dieser Büroordner wird in der Kanzlei als Akte bezeichnet.

Aufenthaltsort

Als Aufenthaltsort wird der Ort bezeichnet, an dem sich ein Schriftstück befindet.

Aufbewahrungsort

Der Aufbewahrungsort gibt Aufschluss über die für die Aufbewahrung zuständige Organisationseinheit und ist vom Begriff „Aufenthaltsort“ zu unterscheiden.

Bearbeitungsschritt

Mit einem Schriftstück durchführbare protokollierte Aktionen. KIS unterstützt die Bearbeitungsschritte „abgelegt“, „erstattet“, „expediert“ und „zugeteilt“.

Bearbeitungsschritt „abgelegt“

Wird ein Schriftstück zur Aufbewahrung abgelegt, kann dies zusammen mit dem Ort der Ablage angegeben werden.

Bearbeitungsschritt „erstattet“

Bei der Erstattung wird das Eingehen eines Schriftstücks von einem Absender protokolliert. Dabei kann die vom Absender vergebene Fremdzahl des Schriftstücks eingegeben werden.

Bearbeitungsschritt „expediert“

Unter der Expedition wird der Versand eines Schriftstücks an einen externen Empfänger verstanden. Als weitere Datenangabe kann neben dem Empfänger eine Antwortfrist angegeben werden.

Bearbeitungsschritt „zugeteilt“

Ein Schriftstück kann einem Empfänger innerhalb der Universität zugeteilt werden. Diese Zuteilung geht wahlweise mit einer Rückgabefrist einher.

Dokument

Ein einzelnes eigenständiges Schreiben aus einem Schriftstück

Fremdzahl

Von einem Absender eines Schriftstücks vergebene eindeutige Bezeichnung zur externen Identifikation des Schriftstücks

Frist

Bei den Bearbeitungsschritten „expediert“ und „zugeteilt“ kann eine optionale Antwort- bzw. Rückgabefrist gesetzt werden.

Geschäftszahl

Eindeutige einem Schriftstück als Bezeichnung zugeteilte Zeichenfolge. Im Altsystem besteht die Geschäftszahl aus den für Thema, Unterthema, Laufnummer und der Jahreszahl der Erfassung angereihten Zeichenfolgen. Zur Vorbeugung von Namenskonflikten bei der Kurzbezeichnung von Themen wurde im neuen System eine übergeordnete optionale Themenkategorie eingeführt.

KIS

Abkürzung für „Kanzleiinformationssystem“

Laufnummer

Die Laufnummer nummeriert die zu einer Kombination aus Thema, Unterthema und Jahr erfassten Schriftstücke fortlaufend.

Nachzahl

Referenz eines Schriftstücks auf ein folgendes Schriftstück

Rückstandsausweis

Wiederanforderung eines versandten Schriftstücks nach Ablauf einer Frist

Schriftstück

Sammlung von zusammenhängenden Dokumenten zu einem Thema

Subthema

Synonym für „Unterthema“

Thema

Das Thema ist ein Teil der Geschäftszahl und bildet die zweite inhaltliche Gliederungsebene.

Themenbereich

Synonym für „Thema“

Themenkategorie

Im neuen System wurde zusätzlich zu den zwei bereits vorhandenen Gliederungsebenen Thema und Unterthema eine optional verwendbare übergeordnete Themenkategorie eingeführt. Aufgrund dieser Themenkategorie kann im Falle von überlagerten Bezeichnungen für Themen und Unterthemen eine eindeutige Unterscheidung getroffen werden. Die Themenkategorie ist ein Teil der Geschäftszahl und bildet die erste inhaltliche Gliederungsebene.

Themenkreis

Gruppierung von Schriftstücken nach Inhalt, setzt sich aus den beiden Zeichenfolgen (im Altsystem Zahlen) für Themenbereich und Unterthema zusammen

Unterthema

Das Unterthema ist ein Teil der Geschäftszahl und bildet die dritte inhaltliche Gliederungsebene.

Universitätsarchiv

Akten, deren Inhalt älter als zehn Jahre ist, werden zur Aufbewahrung von der Kanzlei an das Archiv der Universität übergeben.

Vorzahl

Referenz eines Schriftstücks auf ein vorhergehendes Schriftstück

Gebäude und Technik*Akt / Akte*

Sammlung von zusammenhängenden Dokumenten zu einem Thema, Synonym zum Begriff des Schriftstücks der Kanzlei

Aktennummer

Synonym für den Begriff „Geschäftszahl“ der Kanzlei. Eine Gliederung in Unterthemen auf der dritten Ebene findet in der GUT keine Anwendung.

GUT

Abkürzung für „Gebäude und Technik“

SAP-Projektstrukturplan

Hierarchische Gliederung von Projekten im SAP mit bis zu fünf Gliederungsebenen.