



ulm university universität
uulm

Fakultät für Ingenieurwissenschaften und Informatik
Institut für Datenbanken und Informationssysteme

Vom fortschrittlichen Prozess-Management-System zum intelligenten Enterprise-Content-Management-System

Diplomarbeit

von
Marian Lindenlaub
Januar 2010

1. Gutachter: Prof. Dr. Manfred Reichert
2. Gutachter: PD Dr. Stefanie Rinderle-Ma
Betreuer: Dipl.-Inf. (U) Ulrich Kreher

Kurzfassung

Das enorme Informationsaufkommen und die steigende Relevanz von Informationen, sowie das damit verbundene Wissen als Produktionsfaktor in heutigen Unternehmen verlangen nach einer effizienten Verwaltung von Inhalten. In den letzten Jahren ist mit Enterprise-Content-Management (ECM) ein neuer Begriff entstanden, hinter dem sich ein Konzept verbirgt, mit dem alle unstrukturiert vorhandenen Informationen, wie Texte, Bilder, Grafiken oder Dokumente unternehmensweit verwaltet werden. Dabei wird der gesamte Lebenszyklus von der Erfassung, Bearbeitung und Nutzung über die Speicherung und Archivierung hin zur kontrollierten Vernichtung berücksichtigt. Die unterschiedlichen Aspekte reichen dabei von der Nutzung der Informationen durch einzelne Benutzer über die gemeinsame Bearbeitung bis zur Einbindung in unternehmensweite Geschäftsprozesse, um alle Informationen zur richtigen Zeit den entsprechenden Bearbeitern zur Verfügung zu stellen.

In dieser Arbeit wird das ECM-Konzept detailliert vorgestellt, um zu klären, welche Bereiche und Funktionen mit ECM adressiert werden. Da der Ansatz von ECM noch relativ jung ist, werden noch nicht alle Funktionen gleich gut von am Markt befindlichen Systemen unterstützt. Insbesondere ist die Prozessverwaltung häufig unausgereift. Um zu untersuchen, inwiefern dieser umfassende Ansatz von existierenden Systemen heutzutage realisiert wird, werden dazu einige verbreitete Dokumenten- und Content-Management-Systeme (DMS bzw. CMS) evaluiert.

Auf Basis der Evaluation werden anschließend Anforderungen an eine Integration von Systemen aus den Bereichen Dokumenten-, Content- und Prozessmanagement erarbeitet, um das umfassende ECM-Konzept durch Kombination dieser Systeme umzusetzen. Die ermittelten Anforderungen bilden im weiteren Verlauf der Arbeit die Grundlage für eine allgemeine konzeptionelle Integration der Systeme. Die Konzeption beschreibt, wie das Repository eines DMS als zentraler Speicherort für Dokumente und weitere Informationen, die meist in Form unstrukturierter Daten vorliegen, genutzt werden kann. Diese Informationen können dadurch zentral verwaltet, aufbewahrt und anderen Systemen bereitgestellt werden. Insbesondere geht die Konzeption auf die Aspekte der Einbindung dieser Informationen in Prozesse ein. Weiter wird eine Integration der Organisationsstrukturen und die Präsentation von Funktionen aus beiden Systemen über die Client-Komponenten vorgestellt. Der letzte Teil der Arbeit überträgt die allgemeinen Anforderungen und die Konzeption auf das Prozess-Management-System ADEPT2 [DAG⁺06]. Dazu werden die Erweiterungen für eine Integration von Inhalten eines DMS in den Datenfluss von ADEPT2 vorgestellt. Außerdem wird erläutert, wie die Organisationsstrukturen von ADEPT2 und einem DMS integriert werden können, um darüber Zugriffe und Berechtigungen von Benutzern steuern zu können.

Inhaltsverzeichnis

1. Einleitung	9
1.1. Motivation	9
1.2. Aufgabenstellung	11
1.3. Aufbau der Arbeit	11
2. Enterprise-Content-Management	13
2.1. Begriffsbestimmungen	13
2.1.1. Content	13
2.1.2. Dokument	14
2.1.3. Record	15
2.1.4. Dokumentenlebenszyklus	16
2.2. Merkmale des Enterprise-Content-Management	17
2.3. Komponenten von Enterprise-Content-Management	19
2.3.1. Capture	20
2.3.2. Manage	25
2.3.2.1. Dokumenten-Management	26
2.3.2.2. Records-Management	29
2.3.2.3. Web-Content-Management	32
2.3.2.4. Collaboration	34
2.3.2.5. Workflow/Business Process Management	36
2.3.3. Store	39
2.3.4. Preserve	42
2.3.5. Deliver	46
2.4. Enterprise-Content-Management als integratives Konzept	48
2.5. Zusammenfassung	49
3. Evaluierung von Dokumenten- und Content-Management-Systemen	51
3.1. Evaluierungskriterien	52
3.2. Alfresco Labs 3.0	56
3.3. Microsoft Office SharePoint Server 2007	64
3.4. ELOprofessional 6	72
3.5. DocuWare 5.1 ENTERPRISE	79
3.6. EASY ENTERPRISE.x 3.0	85
3.7. Funktionaler Vergleich	96
3.8. Zusammenfassung	99

4. Anforderungen an die Integration von DMS, CMS und PMS	101
4.1. Prozess-Management-Systeme	102
4.2. Dokumente in Geschäftsprozessen	103
4.3. Unterschiedliche Sichtweisen von DMS und PMS	104
4.4. Abgrenzung der Funktionsbereiche	105
4.5. Schnittstellen	107
4.6. Records Management als gemeinsame Basis für DMS und PMS	109
4.7. Anforderungen	110
4.7.1. MoReq2	111
4.7.1.1. Klassifikationsschema	111
4.7.1.2. Kontrolle und Sicherheit	113
4.7.1.3. Aufbewahrung und Vernichtung	114
4.7.1.4. Dokumentenmanagement	115
4.7.2. Integration von Dokumenten in Prozesse	116
4.7.3. Archivierung von Prozessdaten	116
4.7.4. Organisationsintegration	117
4.7.5. Clientintegration	118
4.8. Zusammenfassung	118
5. Konzeption der Integration von DMS und PMS	119
5.1. Datenintegration	119
5.1.1. Gemeinsames Repository	120
5.1.2. Schemata	120
5.1.3. Referenzen	123
5.1.4. Prozessintegration	124
5.2. Organisationsintegration	127
5.2.1. Verzeichnisdienste	128
5.2.2. Integration der Benutzerverwaltung	132
5.2.3. Zugriffsberechtigungen	133
5.2.4. Authentifizierung	135
5.2.4.1. Getrennte Authentifizierung	135
5.2.4.2. Single Credential	135
5.2.4.3. Single Sign On	136
5.3. Clientintegration	140
5.3.1. Benutzerpräsentation	141
5.3.2. Integration der Rich-Clients	141
5.3.3. Integration der Web-Clients	142
5.4. Zusammenfassung	142
6. ADEPT2-Metamodell	145
6.1. Kontrollfluss	145
6.2. Datenfluss	148
6.3. Korrektheit	150
6.4. Organisationsmodell	150

7. Integration von ADEPT2 mit einem DMS	153
7.1. Datenintegration	153
7.1.1. Schemata	153
7.1.2. Aufbau von Schemata	154
7.1.3. Abbildung von DMS-Inhalten auf strukturierte Datentypen . . .	155
7.1.4. Erweiterungen für Schemata	156
7.1.5. Realisierung der Zugriffe auf DMS-Inhalte	157
7.1.6. Erstellen der Schemata	159
7.1.7. Verwalten von Schemata	161
7.1.8. Zugriffsberechtigungen ermitteln und zuweisen	162
7.2. Organisationsintegration	163
7.2.1. Getrennte Benutzerverwaltungen	163
7.2.2. Synchronisation über Verzeichnisdienst	164
7.2.3. Authentifizierung mit Single Credential und Single Sign On . .	165
7.3. Clientintegration	169
7.3.1. Rich-Client	169
7.3.2. Web-Client	173
7.4. Zusammenfassung	174
8. Zusammenfassung und Ausblick	177
8.1. Zusammenfassung	177
8.2. Ausblick	179
A. Abkürzungsverzeichnis	183
Abbildungsverzeichnis	187
Tabellenverzeichnis	189
Literaturverzeichnis	191
Ehrenwörtliche Erklärung	201

1. Einleitung

Bei der Eroberung des Weltraums sind zwei Probleme zu lösen: die Schwerkraft und der Papierkrieg. Mit der Schwerkraft wären wir fertig geworden.

Wernher von Braun (1912 - 1977)

1.1. Motivation

Wirtschaft und Verwaltung werden zunehmend komplexer. Durch den immer stärker werdenden Konkurrenzkampf auf nationalen und internationalen Märkten müssen unter anderem immer größere Informationsmengen verarbeitet werden. Um schnelle und fundierte Entscheidungen treffen, begründen und dokumentieren zu können, ist ein schneller Zugriff auf Informationen zu jeder Zeit und an fast jedem Ort erforderlich. Die rasche Entwicklung und der Preisverfall von EDV¹-Systemen und das Vordringen von Informationstechnologien in nahezu alle Bereiche hat dazu geführt, dass in immer kürzeren Zeitabständen mehr Daten und Informationen generiert werden. Immer schneller ablaufende Geschäftsprozesse führen dazu, dass immense Datenmengen anfallen. In gleichem Maße nimmt jedoch die gezielte Verfügbarkeit dieser Daten und Informationen ab [GSSZ02].

Gleichzeitig wird das in einem Unternehmen vorhandene *Know-how* zu einem wesentlichen Wettbewerbsfaktor, der in den letzten Jahren enorm an Bedeutung gewonnen hat. Weiterhin haben sich die Arbeitsweisen in den Unternehmen in den letzten Jahrzehnten stark gewandelt. So herrschen in heutigen Unternehmen neben der Einzelarbeit, Gruppenarbeit, auch zunehmend projektorientierte und prozessorientierte Arbeiten vor [Tie02]. All diese unterschiedlichen Tätigkeiten erfordern es, dass die beteiligten Personen stets mit den aktuell benötigten Informationen versorgt werden. Außer den Informationen selbst, sollen diese auch optimiert für die jeweiligen Tätigkeiten zur Verfügung gestellt werden. So ist es bei Gruppenarbeit beispielsweise erforderlich, dass alle daran beteiligten Personen Zugriff auf alle gemeinsam genutzten Informationen haben, während bei der Durchführung von einzelnen Prozessschritten nur die Informationen bereitgestellt werden, die für die Durchführung der einzelnen Aufgaben erforderlich sind.

¹Elektronische Datenverarbeitung

Der größte Teil dieser Informationen liegt dabei in Form von elektronischen Dokumenten, E-Mails, Web-Portalen oder sogar in Papierdokumenten vor, die in Aktenschränken aufbewahrt werden. Das Ziel ist daher, den Informationsfluss in Unternehmen durch ein ganzheitliches Informationsmanagement zu optimieren, mit dem Informationen schneller, umfassender und effizienter verarbeitet und bereitgestellt werden können. Hierzu zählen kurzfristige Zugriffe auf vorhandene Informationen, die Verknüpfung zu Vorgängen, die Erfassung von Informationen, die Speicherung, Suche und die Verteilung an die richtigen Personen [ZGB05]. Zusätzlich kommen speziell in den letzten Jahren viele regulatorische Anforderungen hinzu, die festlegen, wie mit diesen Informationen umzugehen ist, wie lange und in welchem Format sie aufzubewahren sind [NOS⁺09].

Mit den Veränderungen des Marktes haben sich auch die Begrifflichkeiten geändert. *Archivierung*, *Schriftgutverwaltung*, *Dokumenten-Management* (DM²) und als neueste Wortschöpfungen *Information-Lifecycle-Management* (ILM³), *Enterprise-Content-Management* (ECM⁴) und die jeweils entsprechenden Abkürzungen sind Beispiele für eine Vielzahl an Begriffen, die manchmal zur Klärung, häufig aber auch eher zur Verwirrung beitragen [ZGB05].

ECM stellt dabei den bislang umfassendsten Ansatz zum Umgang mit vor allem unstrukturiert gespeicherten Informationen dar, die in Dokumenten, E-Mails, Webinhalten, usw. vorhanden sind. ECM berücksichtigt den kompletten Lebenszyklus. Dieser beginnt mit der Erfassung elektronischer, sowie analog vorliegender und erstellter Informationen und reicht bis zur Ablage und richtliniengesteuerten Aufbewahrung. Während der Nutzung und Bearbeitung dieser Informationen im Verlauf des Lebenszyklus integriert ECM verschiedene Aspekte, die die unterschiedlichen Arbeitsweisen und Organisationsstrukturen in Unternehmen unterstützen. So werden die Aspekte der Bearbeitung durch einzelne Benutzer über Dokumenten-Management-Funktionen berücksichtigt, sowie die Bereitstellung und Bearbeitung durch Kollaborationsfunktionen zur gemeinsamen Bearbeitung und Koordination. Weiter werden auch die Aspekte der Integration und Durchführung von rechnergestützten Geschäftsprozessen, sowie die Verteilung und entsprechende Aufbereitung von Inhalten berücksichtigt – seien es Webinhalte, die über Webserver im Internet bereitgestellt werden, interne Webportale, mit denen Benutzer individuell zugeschnittene Ansichten und Funktionen erhalten oder die Verteilung über verschiedene Distributionskanäle im Rahmen des Output-Management.

Wege und Möglichkeiten eines effizienten Umgangs mit sämtlichen in einem Unternehmen vorhandenen Informationen sind daher in der heutigen Informationsgesellschaft ein entscheidender Faktor. Durch den ganzheitlichen Ansatz von ECM sollen dabei die strategischen und technologischen Komponenten identifiziert werden.

²Dokumenten-Management

³Information Lifecycle Management

⁴Enterprise Content Management

1.2. Aufgabenstellung

Ziel dieser Arbeit ist die Integration eines vollwertigen PMS⁵ mit einem umfassenden DMS⁶, um damit die im ECM-Konzept beschriebenen Funktionen zu einer Gesamtlösung zusammenzuführen.

Dazu ist zunächst das hinter dem Begriff ECM steckende Konzept zu erläutern und dem Leser einen Einblick in den heterogenen Funktionsumfang zu bieten. Dabei ist eine detaillierte Untersuchung der unterschiedlichen Funktionsbereiche, sowie eine Abgrenzung der einzelnen funktionalen Komponenten notwendig.

Für die technische Beschreibung und die Untersuchung des Funktionsumfangs existierender Systeme, sollen dazu aktuelle CMS⁷ und DMS evaluiert werden. Mit der Evaluation soll ermittelt werden, wie gut die jeweiligen Funktionsbereiche von diesen Systemen bereits unterstützt werden. Insbesondere die Unterschiedlichkeit der existierenden Systeme ist dabei zu berücksichtigen. Aufgrund des Umfangs von ECM ist es nicht verwunderlich, dass nicht alle Funktionen gleich gut unterstützt werden. Insbesondere die Prozessverwaltung ist häufig nur rudimentär vorhanden.

Aufbauend auf den Ergebnissen der Evaluation sollen Anforderungen an eine Integration der verschiedenen Systeme (CMS, DMS und PMS) formuliert werden. Dazu sind die verschiedenen funktionalen Komponenten abzugrenzen, die in den Systemen teilweise redundant vorhanden sind. Basierend auf den Anforderungen, ist ein Konzept zur Integration der Systeme zu erarbeiten. Dabei ist das Prozessmodell in geeigneter Weise mit einem DMS zu verbinden, etwa welche Daten sind Prozessdaten, welche Daten werden vom DMS verwaltet und wie spiegelt sich dies in Prozessmodellen wider.

Abschließend soll das allgemeine Konzept auf das PMS ADEPT2 übertragen werden. Hierzu ist im Rahmen der Arbeit eine konzeptionelle Kopplung von ADEPT2 mit einem DMS zu erarbeiten.

1.3. Aufbau der Arbeit

In Kapitel 2 wird zunächst eine Einführung in grundlegende Begriffe im Zusammenhang mit der vorliegenden Arbeit gegeben. Es wird erläutert, was unter dem Begriff ECM zu verstehen ist und wie er als ein umfassendes Konzept der Informationsverwaltung interpretiert werden kann. Dazu werden die von ECM eingeschlossenen Teilbereiche näher beschrieben. In Kapitel 3 werden existierende Dokumenten- und Content-Management-Systeme untersucht. Die Evaluation dient einer Betrachtung des Stands

⁵Prozess-Management-System

⁶Dokumenten-Management-System

⁷Content Management System

der Technik und des Fortschritts aktueller Systeme bei ihrer Umsetzung des ECM-Konzepts. Basierend auf diesen Ergebnissen werden in Kapitel 4 Anforderungen an eine Integration von CMS, DMS und PMS zur Erfüllung des vollständigen ECM-Umfangs definiert. Dazu wird eine Einführung in die Funktionsweise von PMS gegeben und erläutert, wie sich diese Systeme von den Dokumenten- und Content-Management-Systemen unterscheiden. Kapitel 5 beschäftigt sich mit einer allgemeinen Konzeption zur Integration eines DMS mit einem PMS. Die Konzeption betrachtet die Aspekte der Integration von Inhalten und der Bereitstellung dieser Inhalte in einem zentralen Repository, die Einbindung der jeweiligen Organisations- und Benutzerverwaltungen und stellt exemplarisch dar, wie die Systeme dem Benutzer gegenüber über integrierte Client-Funktionen bereitgestellt werden können. In Kapitel 6 wird ein kurzer Überblick über das PMS ADEPT2 gegeben, bevor in Kapitel 7 die allgemeine Konzeption der Integration eines DMS mit einem PMS auf ADEPT2 angewendet wird. Abschließend werden in Kapitel 8 die Ergebnisse der Arbeit zusammengefasst und in einem Ausblick weitere Aspekte und Entwicklungen betrachtet, die zukünftig weiterverfolgt werden können.

2. Enterprise-Content-Management

Dieses Kapitel erläutert, was unter dem Begriff ECM verstanden werden kann, welchen Funktionsumfang ECM besitzt und welche Technologien darin eingeschlossen sind. Dazu werden zuerst einige Grundbegriffe aus dem Umfeld von ECM eingeführt. Anschließend wird der ECM-Begriff allgemein erläutert. Darauf aufbauend werden anhand eines funktionalen Modells die verschiedenen Bereiche detailliert betrachtet, die ECM umfasst.

2.1. Begriffsbestimmungen

Zu Beginn werden die grundlegenden Begriffe *Content*, *Dokument*, *Record* und *Dokumentenlebenszyklus* genauer beschrieben. Diese Begriffe werden im Zusammenhang mit ECM benötigt und werden in der Literatur sehr vielfältig definiert und verwendet. Sie werden für diese Arbeit genutzt, um ein einheitliches Verständnis der Begriffe zu schaffen.

2.1.1. Content

Der Begriff *Content* (engl. Inhalt) beinhaltet alles, was an inhaltlichen Informationen oder Nutzdaten in EDV-Systemen vorgehalten wird, und ist daher nur schwer zu fassen. Content kann als logische „Informationseinheit“ auf atomarer Ebene verstanden werden, der in DV-Systemen gespeichert, verwaltet und immer wieder in anderen Zusammenhängen und mit anderen Medien genutzt werden kann [GSMK04]. Häufig wird der Content-Begriff auch mit den Begriffen *Daten*, *Information* und *Wissen* definiert. Unter Daten wird eine Kette von elementaren Symbolen verstanden, die einer vereinbarten Syntax folgen. Wenn durch diese Daten eine Zustandsänderung beim Empfänger (Mensch, DV-System, usw.) hervorgerufen wird, kann man von Information sprechen [DP98, TH06]. Dabei ist der Kontextbezug wichtig. Information kann für verschiedene Menschen unterschiedliche Bedeutung haben. Der Begriff Wissen kann als Ansammlung von in einem System (Mensch, Bibliothek, EDV-System, usw.) gespeicherten Informationen, die als Bestandteil des individuellen Bewusstseins durch Verknüpfungen von interpretierten und subjektiv bewerteten Informationen verstanden werden [DP98].

Ausgehend von diesem Konzept lässt sich Content somit als (austauschbares) Informationspaket beschreiben, das zwischen den Begriffen Information und Wissen einzuordnen ist. Über den Nutzwert der Information macht der Contentbegriff keine Aussage, da der Wert durch die subjektiven Merkmale des Nutzers geprägt wird [KJOF01].

Content lässt sich weiterhin in verschiedene Klassen einordnen [Kam03b]:

- Strukturierter Content sind Daten, die in einem standardisierten Layout aus datenbankgestützten Systemen oder standardisierten Dateiformaten bereitgestellt werden (z. B. formatierte Datensätze aus einer Datenbank oder XML¹-Dateien).
- Schwach strukturierter Content sind Informationen und Dokumente, die zum Teil Layout und Meta-Daten mit sich tragen, jedoch nicht standardisiert sind (z. B. Textverarbeitungsdateien in einem proprietären Format).
- Unstrukturierter Content besteht aus beliebigen Informationsobjekten, deren Inhalt nicht direkt erschlossen werden kann und keine Trennung von Inhalt und Layout vorhanden ist (z. B. Bilder, Video, Sprache).

2.1.2. Dokument

In der Literatur finden sich vielfältige Definitionsansätze zu dem Begriff Dokument. So wird ein Dokument als „Oberbegriff für alle in Papierform vorliegenden Informationen“ in [BL94] bezeichnet, während in [Pfa95] alles darunter verstanden wird „was sich irgendwie digital auf Computern speichern lässt bzw. irgendwann in Zukunft speichern lassen wird“. Kampffmeyer geht auf die unterschiedliche Bedeutung im deutschen und englischen Sprachgebrauch ein und berücksichtigt sowohl physische, als auch elektronisch gespeicherte Dokumente [Kam03a]. In [GSSZ02] und [GSMK04] werden Dokumente anhand ihrer Eigenschaften charakterisiert, elektronische bzw. digitalisierte physische Dokumente werden in die beiden Klassen der NCI²- und CI³-Dokumente unterteilt. Die Unterscheidung erfolgt dadurch, dass ein EDV-System bei NCI-Dokumenten keine Möglichkeit besitzt, auf die Inhalte des Dokuments zuzugreifen. Im Gegensatz dazu stehen CI-Dokumente, deren Inhalte in kodierter Form vorliegen und von einem EDV-System gelesen und verarbeitet werden können.

Anders betrachtet kann ein Dokument als Ansammlung von Content-Einheiten angesehen werden, wobei ein Dokument aus ein oder mehreren Content-Einheiten zusammengesetzt ist und den Informationsträger beinhaltet. Der Informationsträger kann dabei ein physisches Medium, wie Papier oder Mikrofilm sein oder aber eine Datei, die in einem EDV-System gespeichert ist. Bei physischen Informationsträgern sind Layou-

¹Extensible Markup Language

²Non-Coded Information

³Coded Information

Informationen damit ebenfalls direkt festgelegt, wohingegen bei Dateien nicht immer die Layoutinformationen mit enthalten sind (z. B. XML-Dateien). Dadurch lässt sich ein Dokument in die Bereiche des unstrukturierten oder schwach strukturierten Content einordnen. Jedoch hängt diese Einordnung bei elektronischen Dokumenten auch stark vom verwendeten Dateiformat ab. Insbesondere durch die Verwendung von XML als Speicherformat für Dokumente, erhalten diese einen deutlich gesteigerten Strukturierungsgrad. Weiterhin kann bei XML auch eine besser Trennung zwischen Struktur und Layout vorgenommen werden, als dies bei anderen Speicherformate, wie z. B. dem DOC-Format von MS Word möglich ist. Eine Abgrenzung zwischen Content und Dokument ist graphisch in Abbildung 2.1 veranschaulicht.

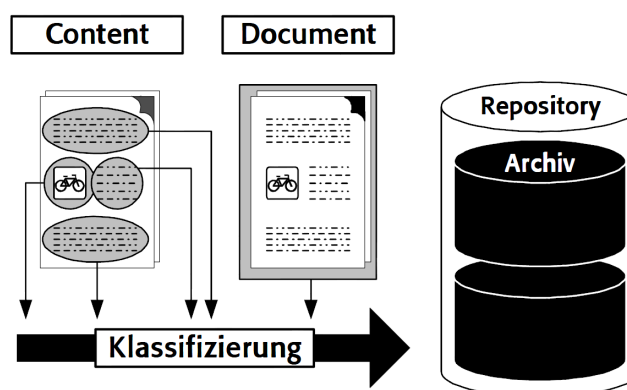


Abbildung 2.1.: Abgrenzung von Content und Dokument [MV02]

Für verschiedene Ausprägungen eines Dokuments wird der Begriff *Rendition* verwendet, um klarzustellen, dass der Inhalt bei allen Ausprägungen gleich ist, lediglich die Darstellung oder der Informationsträger abgewandelt sind.

2.1.3. Record

Records sind nach Definition der *ISO 15489 Information und Dokumentation - Schriftgutverwaltung* [ISO02]: „Informationen, die erzeugt, empfangen und bewahrt werden, um als Nachweis zu dienen, und Informationen, die einer Organisation oder Person dem Nachkommen rechtlicher Verpflichtungen oder bei der Durchführung von Geschäftsprozessen dienen“.

Damit kann ein Record ein beliebiger Inhaltstyp sein, der sich auf die Geschäftstätigkeit eines Unternehmens bezieht. Art des Mediums oder weitere Eigenschaften sind dafür nicht relevant. Damit können Records aus einer oder mehreren Informationseinheiten bestehen, die selbst Content oder ein Dokument sind. Beispielsweise würden sowohl eine Rechnung, als auch ein Buchungssatz einen Record darstellen. Eine Rech-

nung würde dabei jedoch in die Kategorie eines Dokuments fallen, während ein Buchungssatz der Kategorie Content zuzuordnen wäre. Eine Bedienungsanleitung für ein Gerät ist zwar ein Dokument, aber kein Record, da die Bedienung eines Geräts nicht einer geschäftlichen Handlung oder dem Nachweis bestimmter Sachverhalte dient.

2.1.4. Dokumentenlebenszyklus

Der Dokumentenlebenszyklus beschreibt die Stationen, die ein Dokument von der Entstehung bis zur Vernichtung durchläuft. Eine schematische Darstellung des Dokumentenlebenszyklus ist in Abbildung 2.2 zu finden.

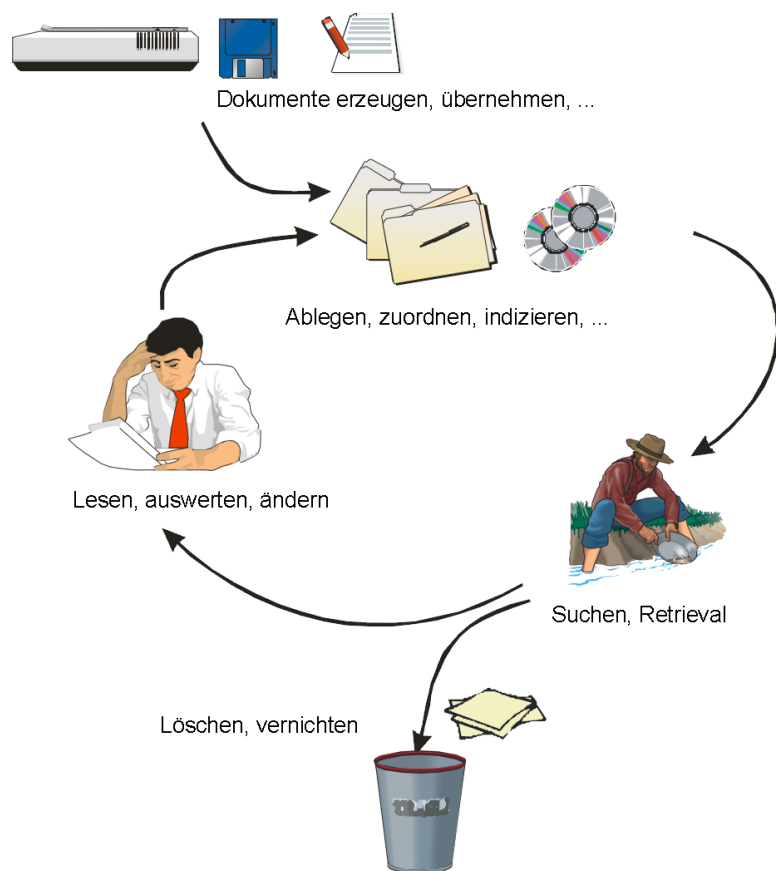


Abbildung 2.2.: Dokumentenlebenszyklus [GSMK04]

Der Zyklus beginnt mit der Erstellung bzw. Erfassung des Dokuments durch Erfassungskomponenten eines Dokumentenverwaltungssystems, etwa durch Kopieren einer Datei oder Einscannen. Nachdem das Dokument im Dokumentenverwaltungssystem gespeichert wurde, kann es ggf. noch bearbeitet oder ergänzt werden. Dabei entste-

hen neuere Versionen des Dokuments, oder es wird an weitere Benutzer zur Bearbeitung weitergeleitet. Wenn es in seiner endgültigen Fassung vorliegt, kann es (durch einen Freigabeprozess) veröffentlicht werden. Ab diesem Zeitpunkt wird das Dokument nicht mehr als lebendes Dokument betrachtet. In den meisten Fällen wird das Dokument in dieser freigegebenen Version nun archiviert und je nach Bestimmung für einen gewissen Zeitraum aufbewahrt. Während dieses Zeitraums wird nur noch lesend auf das Dokument zugegriffen. Nach Ablauf dieses Zeitraums kann bzw. muss das Dokument vernichtet werden [ZGB05, GSMK04].

2.2. Merkmale des Enterprise-Content-Management

Mit ECM entstand in den letzten Jahren ein neuer und in der Praxis vielfach verwendeter Begriff. Die hinter dem Begriff steckende Idee ist, die beiden Strömungen der DMS und CMS zusammenzuführen. Daher wird zunächst eine kurze Einführung in die Entstehung von DMS und CMS gegeben.

Dokumenten-Management-Systeme werden seit etwa 20 Jahren entwickelt und haben ihren Ursprung im *Information-Retrieval*, das damals primär im Bibliothekswesen eingesetzt wurde [ZGB05]. Im Laufe der Zeit kamen weitere funktionale Anforderungen an die Systeme hinzu, um die Probleme der wachsenden Papierflut in den Unternehmen beherrschbarer zu machen. Dazu wurden sie um Funktionen zur elektronischen Archivierung und der Erfassung der Eingangspost erweitert. Neben der Flut an Papierdokumenten kommen inzwischen eine zunehmende Zahl an digital erzeugten Dokumenten hinzu. Durch die Komplexität der Dokumentenstrukturen, sowie der veränderten Erstellungs- und Bearbeitungsprozesse, die möglichst ohne Medienbrüche erfolgen sollen, haben sich die Anforderungen erweitert. Sie haben sich dahingehend erweitert, dass ein DMS den gesamten Lebenszyklus von physischen und elektronischen Dokumenten von der Entstehung über die Veröffentlichung hin zur Archivierung und Vernichtung abdecken soll. Außerdem sollen zusätzlich auch eine Reihe regulatorischer Anforderungen berücksichtigt werden [ZGB05]. Zum Beispiel sind dies gesetzliche Rahmenbedingungen oder firmeninterne Richtlinien zum Umgang und der Aufbewahrung von Informationen.

Neben den DMS haben sich in den 90er Jahren CMS entwickelt, die darauf abzielen, Inhalte effizient zu erstellen, zu publizieren und zu verwalten. Sie werden häufig zur Gestaltung von Webinhalten im Intra- oder Internet verwendet und werden daher oft synonym als WCM⁴ bezeichnet [Boi04, Nak02]. Zunächst ging es bei diesen Systemen darum, umfangreiche Webseiten einheitlich zu erstellen, die mit einfachen Web-Editoren kaum noch zu verwalten und kontrollieren sind. Weiterhin wachsen die Anforderungen, dynamisch generierte Seiten mit aktuellen und personalisierten Inhalten aus Da-

⁴Web Content Management

tenbanken und anderen Informationssystemen zu erzeugen und dabei differenzierte Sichten und Zugriffsrechte, sowie Zugriffsstatistiken über die Nutzung bereitzustellen. Inzwischen sollen WCM auch dazu genutzt werden, um ganze Portale anzubieten, in die verschiedene Inhalte und Systeme integriert werden können, um dem Benutzer eine einfache Sicht auf verschiedene Inhalte zu ermöglichen [GSSZ02].

Die Überschneidungen zwischen DMS und CMS sind sehr groß. Beide Systeme beschäftigen sich mit der Erstellung und Verwaltung von Inhalten. DMS mit dem Schwerpunkt ganze Dokumente zu verwalten, zu archivieren und damit auch die Papierberge zu reduzieren, die sich in den Unternehmen finden. CMS hingegen sind eher darauf ausgerichtet die Inhalte in kleineren Content-Einheiten an mehreren Stellen in anderen Kontexten zu präsentieren oder eine gemeinsame Bearbeitung von Inhalten zu ermöglichen. Da das Internet auch immer mehr an Bedeutung für Unternehmen als Kommunikationsplattform mit Kunden und Partnern oder zur Informationsversorgung von Mitarbeitern auf Geschäftsreisen wird, wächst das Bedürfnis, die in DMS und CMS gespeicherten Informationen auch in beiden Systemen gleichermaßen nutzen zu können. Beispielsweise fallen durch die Integration von E-Commerce-Anwendungen wie Webshops auch Dokumente im CMS an, die zusammen mit den Dokumenten im DMS verwaltet und archiviert werden sollen. Sowohl die Bestellungen, die über einen Onlineshop getätigt werden, als auch Bestellungen, die über andere Kanäle z. B. telefonisch oder über Fax eingegangen sind, müssen an einer Stelle zusammenlaufen. Die daraus resultierenden Dokumente wie Auftragsbestätigungen, Lieferscheine und Rechnungen müssen dabei zentral im Archiv des DMS abgelegt werden und dem Kunden ggf. über einen persönlichen Bereich über die Webseite zur Verfügung gestellt werden. Häufig sind in größeren Unternehmen darüber hinaus auch mehrere DM-, CM⁵- oder WF⁶-Systeme zu finden, die lediglich die Bedürfnisse einzelner Abteilungen abdecken [Fra08].

Unternehmen sehen sich heutzutage mit den Problemen der wachsende Flut an elektronischen Daten konfrontiert und werden damit vor große Herausforderungen gestellt [MPHS06]. Folgende Aussagen von Analysten verdeutlichen, dass der Bedarf an umfassenden DMS und CMS in Unternehmen sehr hoch ist und die heute am Markt befindlichen Systeme diese Anforderungen nur unzureichend erfüllen:

- 80 % der Geschäftsdokumente liegen in unstrukturierter Form (häufig in Papierform) vor (IDC [GRC⁺07])
- Dokumenten-Missmanagement kostet 40-60 % der Arbeitszeit, 20-45 % der Gehaltskosten und 12-15 % des Unternehmensumsatzes (Gartner [SC05])
- 3 % aller Akten gehen komplett verloren. Im Durchschnitt kostet es 100 € eine verlorene Akte wieder herzustellen [OHE99]

⁵Content Management

⁶Workflow

- ca. 15-35 % der Arbeitszeit in Büros und Verwaltungen wird mit der Suche nach Dokumenten und Informationen verbracht und ca. 25 % der Dokumente werden nicht gefunden (IDC [FS04])
- Ein Dokument wird durchschnittlich 9-11 mal kopiert (Gartner [WLKC01])

An dieser Stelle setzt ECM mit einem umfassenden Ansatz an, um unstrukturierte Informationen einheitlich zu verwalten und bereitzustellen. Durch die Bezeichnung Content-Management wird versucht, die Einschränkungen des dokumentenorientierten Begriffs von DMS auszuräumen, da man dynamisch generierte Webseiten oder Auslagerungen aus ERP⁷-Anwendungen nur schwerlich unter dem Begriff Dokument einordnen würde. Dennoch müssen auch diese Informationen sicher abgespeichert und archiviert werden. Mit dem Begriff Enterprise soll deutlich gemacht werden, dass mit ECM ein unternehmensweites Repository bereitgestellt wird, in dem alle in schwach oder unstrukturierter Form vorliegenden Informationen eines Unternehmens gespeichert und verwaltet werden. Dadurch wird es zur zentralen Informationsquelle für Benutzer. Außerdem dient ECM auch als Content-Dienst für unterschiedliche Quell- und Zielsysteme dienen [ZGB05]. Darüber hinaus wird die Integration der Informationen in Geschäftsprozesse angestrebt. Zudem sollen sich die gespeicherten Informationen in unterschiedliche Medien integrieren lassen und über unterschiedliche Kommunikationskanäle verteilt werden können.

Um den Funktionsumfang von ECM konkreter zu bestimmen, liefert der Branchenverband AIIM⁸ eine allgemein anerkannte Definition [Ass08]:

Enterprise Content Management (ECM) is the strategies, methods and tools used to capture, manage, store, preserve, and deliver content and documents related to organizational processes. ECM tools and strategies allow the management of an organization's unstructured information, wherever that information exists.

ECM schließt herkömmliche Technologien der Bereiche Dokumenten-Management, Datenerfassung, Workflow oder Archivierung mit ein. ECM führt damit die verschiedenen Bereiche zusammen, die sich mit der Erstellung oder Bearbeitung von unstrukturiert gespeicherten Informationen beschäftigen.

2.3. Komponenten von Enterprise-Content-Management

Anhand der ECM-Definition der AIIM lässt sich aus den genannten Begriffen ein funktionales Modell entwickeln, das die dafür notwendigen Technologien und Komponenten beschreibt. Dieses Modell orientiert sich an den fünf Leitbegriffen der Definition

⁷Enterprise Resource Planning

⁸Association for Information and Image Management

der AIIM. Diese Komponenten müssen dabei nicht in einem einzelnen System abgebildet werden, sondern können auch auf mehrere Systeme verteilt werden. Sie werden häufig nur als Dienste in Anlehnung an eine SOA⁹ beschrieben [Kam03b]. Zur Veranschaulichung werden sie in verschiedene Kategorien eingeordnet, die sich am Dokumentenlebenszyklus orientieren.

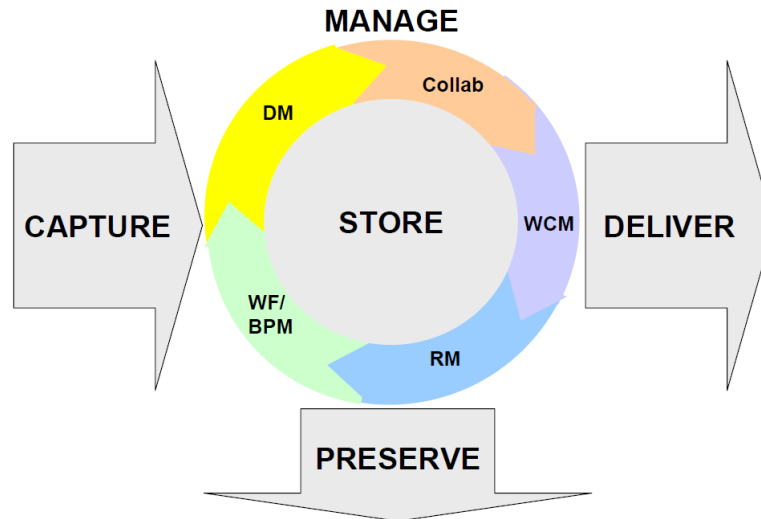


Abbildung 2.3.: ECM-Modell [KSP07]

Die Einordnung erfolgt in die Bereiche Capture (Erfassung), Store (Speicherung), Preserve (Archivierung), Deliver (Ausgabe) und Manage (Verwaltung). Im Bereich Manage werden die steuernden Komponenten zusammengefasst, die der Bearbeitung und Verteilung im Dokumentenlebenszyklus entsprechen. Der dynamische Teil des Lebenszyklus wird durch Dokumenten-Management (DM), Web-Content-Management (WCM), Collaboration (Collab) und Workflow (WF) mit Business Process Management (BPM) abgebildet, während Records-Management (RM) dazu dient die Informationen nach ihrer Erstellung und Bearbeitung sicher zu archivieren und kontrolliert am Ende der Aufbewahrungsfrist zu vernichten.

2.3.1. Capture

Zu Beginn des Dokumentenlebenszyklus steht die Erfassung von analog oder elektronisch vorliegenden Informationen. Der Bereich Capture umfasst dafür Funktionen und Komponenten zur Erfassung, Aufbereitung und Verarbeitung. Durch die sehr unterschiedlichen Medien und Datenquellen, in denen die Informationen vorliegen ergeben sich unterschiedliche Erfassungs-, Attributierungs- und Prüfschritte.

⁹Service-orientierte-Architektur

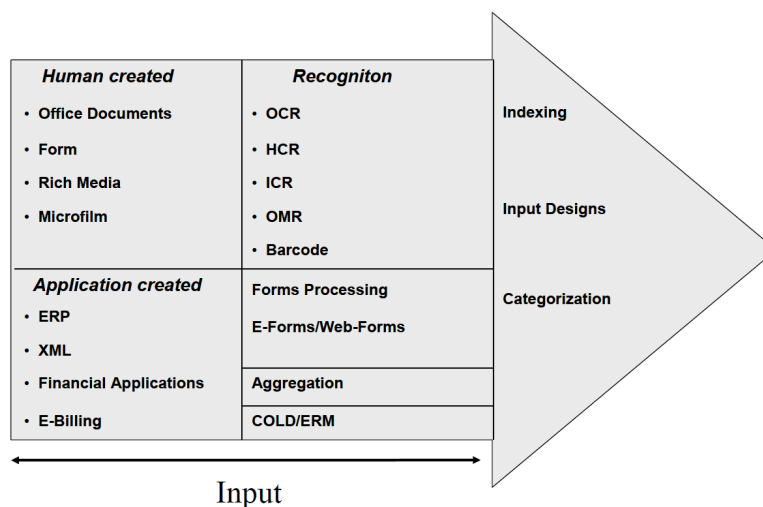


Abbildung 2.4.: ECM-Capture [KSP07]

Bei der Erfassung lässt sich grundsätzlich zwischen manuell und maschinell erzeugten bzw. erfassten Informationen unterscheiden. Bei manuell erzeugten bzw. erfassten Informationen können diese in allen Formen vorliegen, z. B. Papierdokumente, mit Anwendungsprogrammen erstellte Dokumente, E-Mails, digitale und analoge Bilder, Tonaufzeichnungen bis hin zum Mikrofilm. Bei der maschinellen bzw. automatisierten Erfassung liegen die Daten bereits elektronisch in einem anderen System, z. B. einem ERP-System oder in Form von EDI¹⁰- oder XML-Dokumenten vor und können damit durch geeignete Schnittstellen und anhand der vorliegenden Datenstrukturen automatisiert erfasst werden.

Erfassung von NCI-Dokumenten

Nach wie vor liegen immer noch eine sehr große Zahl von Dokumenten in Papierform vor. So zeigt eine 2009 von Deutsche Bank Research [May09] durchgeführte Studie, dass trotz steigender Tendenz zur elektronischen Rechnung, immer noch 95 % aller Rechnungen in Europa in Papierform verschickt werden. Um diese auf Papier gespeicherten Informationen elektronisch nutzen zu können, müssen sie zunächst digitalisiert werden. Die gängigste Variante hierbei ist der Einsatz von Scannern, um die Papierdokumente in ein digitales Bild zu transformieren. Scanner stehen für unterschiedliche Qualitätsanforderungen, Dokumentenformate und Erfassungsvolumen zur Verfügung. Details zu Erfassungsgeräten finden sich in [ZGB05], [GSSZ02] und [GSMK04].

¹⁰Electronic Data Interchange

Verarbeitung erfasster Informationen

Zur Verarbeitung von erfassten Informationen werden verschiedene Technologien zur Erkennung eingesetzt, um etwa die Inhalte aus den erstellten Bilddateien beim Scannen zu extrahieren oder Dokumente zu Vorgängen zuzuordnen. Im Folgenden werden die wichtigsten Technologien hierfür vorgestellt.

OCR¹¹ ist eine Technik zur *Texterkennung* gedruckter Schriftzeichen durch ein EDV-System. OCR wird vor allem dazu verwendet, um NCI-Dokumente in CI-Dokumente umzuwandeln, sodass die gespeicherten Informationen direkt von einem Computer interpretiert und verarbeitet werden können. Eingescannte Papierdokumente oder von einer Faxsoftware empfangene Faxe liegen i. d. R. zunächst als Bilddatei vor. Grundsätzlich werden mit OCR bearbeitete Dokumente als NCI-Dokumente betrachtet. Mit Hilfe von OCR können die erkannten Zeichenketten direkt im System verarbeitet und genutzt werden.

Außerdem existieren weitere Erkennungstechniken bzw. Erweiterungen zur Erhöhung der Erkennungsqualität wie HCR¹² zur Handschrifterkennung. Mit ICR¹³ werden zusätzliche Daten und logische Zusammenhänge in den Erkennungsprozess einbezogen, z. B. Referenzlisten oder die Einbeziehung vorhandener Datenquellen. Die Erkennung von Markierungen wie sie beim Ankreuzen eines Fragenkatalogs z. B. in der Marktforschung genutzt werden erfolgt mittels OMR¹⁴. IDR¹⁵ schließlich dient der automatischen Erkennung von Dokumenten, mit der eine automatische Klassifizierung von Dokumenten in bestimmte Kategorien realisiert werden kann. Dies erfordert jedoch einen hohen Trainingsaufwand der Software.

Barcodes dienen dazu, Papierdokumente eindeutig zu identifizieren, um diese Dokumente im System Vorgängen oder Akten zuzuordnen. Barcodes eignen sich auch gut für selbst erstellte Dokumente, die von anderen Personen weiter bearbeitet werden, z. B. ein Antrag, der zu einem Kunden geschickt und dort ausgefüllt wird. Nach der Rücksendung lässt sich das Formular durch Scannen des Barcodes eindeutig diesem Vorgang direkt zuordnen.

Formularerfassung wird in die beiden Bereiche Vordruckverarbeitung und elektronische Formulare unterteilt. Bei den Vordrucken werden die Formulare zunächst auf Papier gedruckt und nach dem Ausfüllen dann mit Hilfe von Scannern digitalisiert. Da das Formularlayout bekannt ist, können durch entsprechende Vorlagen die Inhalte des Formulars gezielt verarbeitet werden. Häufig werden für die vorgedruckten Linien und Texte bestimmte Farben verwendet, die von der Scannersoftware ausgeblendet

¹¹Optical Character Recognition

¹²Handprint Character Recognition

¹³Intelligent Character Recognition

¹⁴Optical Mark Recognition

¹⁵Intelligent Document Recognition

werden können. Hierbei kommen die bereits vorgestellten Erkennungstechniken zur Inhaltserschließung zum Einsatz. Bei elektronischen Formularen kann der Inhalt direkt verarbeitet werden, sofern die Struktur und das Layout der Formulare dem Erfassungssystem bekannt sind.

COLD¹⁶ sind Verfahren zur automatisierten Verarbeitung von Ausgabedatenströmen. Der COLD-Begriff entstand in den 80er Jahren mit den Anfängen der elektronischen Archivierung. Hier hatte man begonnen die erzeugten Dokumente nicht mehr in Papierform oder auf Microfiche, sondern digital auf optischen Medien abzulegen. Damit war es damals möglich auf Archivdaten über den Computer zuzugreifen, um die Zeiten für Suche und Transport drastisch zu reduzieren. Die Archivierung erfolgt hierbei i. d. R. über einen Druckertreiber, der die Informationen aus den Anwendungen entgegennimmt und statt auf einem Drucker auszugeben von einem COLD-Archivserver entgegengenommen wird. Daten können dabei in ein archivgeeignetes Langzeitformat wie PDF/A oder TIFF gewandelt werden. Während des Imports der Daten können außerdem Indexdaten aus den Dokumenten extrahiert werden, die die Basis für die spätere Suche nach den Dokumenten bilden.

ERM¹⁷ ist eher im englischsprachigen Raum geläufig und dient ebenfalls der Aufbereitung der Daten mit bestimmten Strukturinformationen, so dass sie unabhängig vom verwendeten Quellsystem an die Store- oder Preserve-Komponenten weitergereicht werden können.

Erfassung von CI-Dokumenten

Wenn CI-Dokumente direkt in einer Anwendung erstellt werden, liegt es nahe, diese direkt über die Capture-Komponenten abzuspeichern und die der Anwendung bekannten Attribute dabei automatisch mit dem Dokument abzuspeichern. Weitere Attribute können über entsprechende Eingabemasken durch den Benutzer ergänzt werden. Hierzu ist es erforderlich, dass die Anwendungen die Schnittstellen kennen, die für die Ablage notwendig sind. Für eine durchgängige Integration müssen dazu alle dokumentenerzeugenden und dokumentennutzenden Anwendungen an die ECM-Komponenten angebunden werden. Bei Altanwendungen, die keine Ankopplung an die ECM-Komponenten erlauben, können transparente Schnittstellen, wie etwa eine Netzwerkfreigabe, verwendet werden, über die Dokumente direkt geöffnet und abgelegt werden. Die Angabe von Attributen der Dokumente ist dabei i. d. R. nicht möglich. Bislang sind in diesem Bereich kaum standardisierte Schnittstellen vorhanden, über die Anwendungen Dokumente abspeichern können.

¹⁶Computer Output on Laserdisc

¹⁷Enterprise Reports Management

Indexieren erfasster Informationen

Bei der Indexierung werden Attribute und Metadaten der erfassten Dokumente definiert. Dabei ist zu berücksichtigen welche Dokumenttypen bei der Erfassung auftreten können und welche Attribute diese Dokumente kennzeichnen. Ein weiterer wesentlicher Faktor ist, welche Metadaten notwendig oder interessant sind, um ein Dokument möglichst eindeutig und vollständig zu beschreiben. Bei der Vergabe der Indexdaten kann zwischen freiem und gebundenem Indexieren unterschieden werden. Beim freien Indexieren werden die Bezeichnungen vom Bearbeiter selbst gewählt und dem Dokument zugeordnet. Beim gebundenen Indexieren hingegen wird vorher ein Index mit Schlagwörtern (Thesaurus) erstellt. Dies kann z. B. *Bestellung*, *Lieferschein* oder *Rechnung* sein, um den Dokumenttyp zu beschreiben. Der Thesaurus muss möglichst vollständig sein, um alle auftretenden Beschreibungen abzudecken. Bei der Indexierungsgenauigkeit muss zwischen dem Aufwand, der bei der Indexierung anfällt, und den Suchmöglichkeiten nach dem Dokument abgewogen werden. Hierbei sind auch Geschäftsprozesse und allgemeine Vorgehensweisen der Mitarbeiter zu betrachten, da die Indexierung die Grundlage für die Auswahlmöglichkeiten bei Suchanfragen, sowie die Einordnung in ein Ordnungssystem festlegt [Gau03].

Die erfassten Attribute der Dokumente werden typischerweise in einer eigenen Datenbank zusammen mit einem Verweis auf das eigentliche Dokument gespeichert. Dabei kann auch eine automatische Erkennung von Dokumentenattributen angewendet werden, um die Erfassungsvorgänge zu beschleunigen und damit auch Kosten zu reduzieren. Um die Attribute erkennen zu können, muss die Struktur des Dokuments im Voraus bekannt sein, wie dies bei Rechnungen eines Lieferanten oder Formularen der Fall ist. Dabei wird der Software über eine Vorlage mitgeteilt, in welchen Bereichen bestimmte Merkmale, wie z. B. eine Rechnungsnummer oder Anschrift des Lieferanten, vorhanden sein müssen, um das Dokument zuordnen zu können. Dadurch muss nicht das gesamte Dokument interpretiert werden, sondern nur die relevanten Bereiche, die benötigt werden. Zusätzlich können für diese Teile des Dokuments Wertebereiche eingeschränkt werden, was zu einer Verbesserung der Erkennungsqualität führt. Beispielsweise kann im Feld Postleitzahl eine 0 leicht als o gedeutet werden. Durch eine Einschränkung des Wertebereichs für dieses Feld kann die Software sicherstellen, dass das erkannte Zeichen eine 0 sein muss.

Bei stark textorientierten Dokumenten kann die Texterkennung auf den gesamten Text angewendet und so später direkt mit einer Textverarbeitungssoftware bearbeitet werden. Zusätzlich kann mit dem erkannten Text ein Volltextindex erstellt werden, der die Grundlage für eine spätere Volltextsuche darstellt, um das Dokument anhand einer inhaltlichen Übereinstimmung mit den gesuchten Begriffen wiederzufinden.

Abbildung 2.5 zeigt schematisch den Vorgang des Scannens, Attributierens und der Speicherung im System:

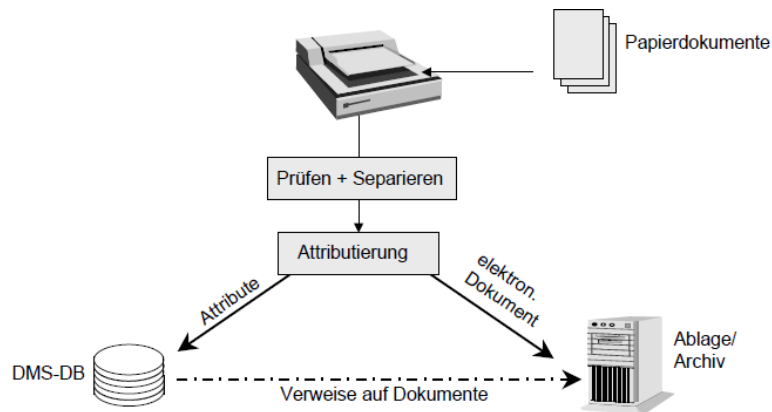


Abbildung 2.5.: Scannen und Indexieren [GSSZ02]

Die Komponenten für die inhaltliche Erschließung können in drei Bereiche unterteilt werden:

- **manuelle Indexierung**

Der einfachste Fall stellt die manuelle Vergabe von Indexattributen dar. Die Indexattribute sollen ein Dokument mit wenigen Attributen möglichst eindeutig beschreiben. Die Anzahl der verwendeten Attribute richtet sich an der Art des Dokuments und dem geplanten Zugriff auf das Dokument aus.

- **Vorlagen und Profile**

Um die Attributierung der Dokumente zu vereinheitlichen werden Vorlagen oder Profile für bestimmte Dokumenttypen erstellt. Somit wird sichergestellt, dass alle definierten Attribute eines Dokumenttyps mit Informationen versorgt werden. Hierfür werden i. d. R. Eingabemasken definiert und den Bearbeitern bei der Attributierung bereitgestellt.

- **automatische Klassifikation**

Auf Basis der vorhandenen Informationen in den Dokumenten selbst kann auch Software eingesetzt werden, um die Informationen aufgrund ihres Inhalts automatisiert zu klassifizieren und die notwendigen Indexdaten zu erstellen. Die Systeme arbeiten meist anhand vordefinierter Kriterien oder werten die Informationen selbstlernend aus. Dabei ist es notwendig, dass die Ergebnisse kontrolliert werden, um falsche Zuordnungen zu vermeiden.

2.3.2. Manage

Die Manage-Komponenten dienen der Verwaltung, Bearbeitung und Nutzung der Informationen. Die Verwaltung der Informationen erfolgt i. d. R. durch Nutzung von

Datenbanken, die die Beziehungen zwischen den Informationen und den Metadaten verwalten. Zusätzlich werden Berechtigungssysteme benötigt, über die der Zugriff auf die Informationen kontrolliert und der Schutz der Informationen sichergestellt werden kann. Ziel von ECM ist es, diese beiden Komponenten einmal übergreifend für alle Bereiche bereitzustellen.

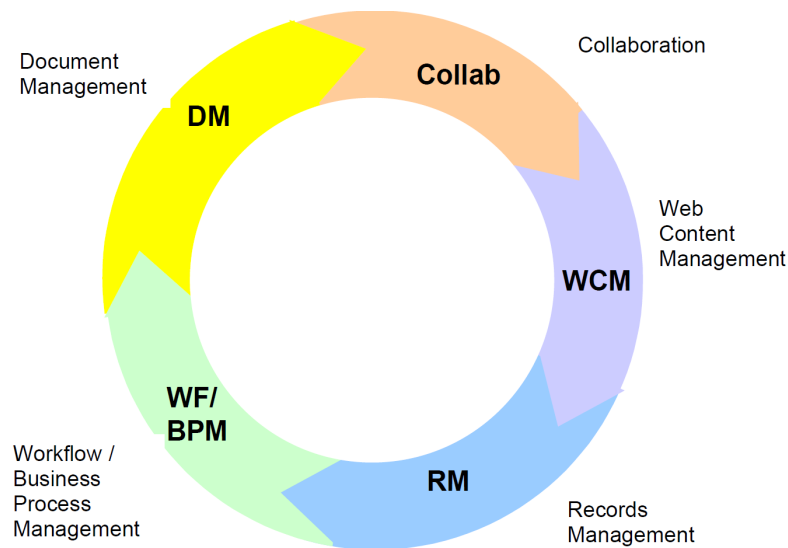


Abbildung 2.6.: ECM-Manage [Kam03b]

2.3.2.1. Dokumenten-Management

Der Bereich DM dient der Verwaltung, Kontrolle und Organisation von Dokumenten. Dabei soll der gesamten Lebenszyklus der Dokumente von der Erzeugung bis zur Langzeitarchivierung abgedeckt werden.

Die Hauptfunktionen des DM sind strukturierte Dokumentenablagen bereitzustellen, in welchen die Dokumente abgelegt werden und über die eine Navigation im Dokumentenbestand und die Suche nach Dokumenten realisiert wird. Weiterhin werden in diesem Bereich auch Zugriffsberechtigungen auf Dokumente festgelegt. Für die Absicherung der Lesbarkeit von Dokumenten über lange Zeit werden hierfür Viewer-Komponenten bereitgestellt, mit denen Dokumente angezeigt werden können, ohne die Quellenanwendung dafür auf dem Computer des Benutzers vorzuhalten [ZGB05].

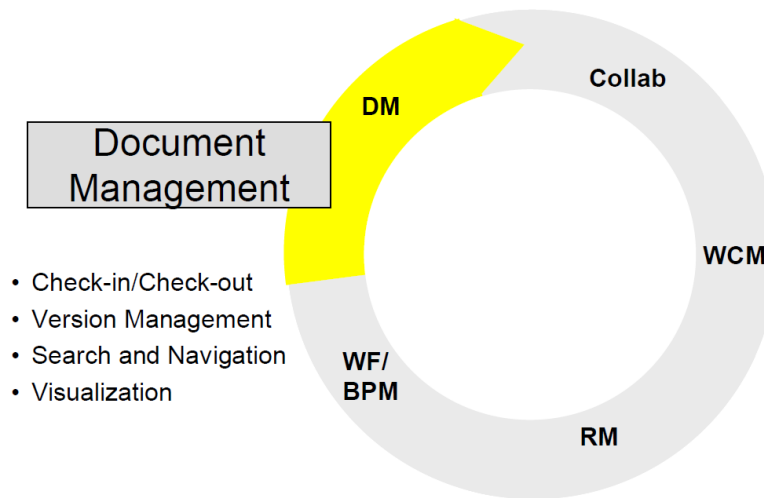


Abbildung 2.7.: Dokumenten-Management [KSP07]

Dokumentenverwaltung

Nachdem die Dokumente durch die Erfassungskomponenten abgelegt sind, können sie im weiteren Verlauf ihres Lebenszyklus bearbeitet und verändert werden. Änderungen können grundsätzlich am Dokument selbst, an seiner Struktur oder an seinen Metadaten vorgenommen werden. Im Folgenden werden typische Funktionen aus dem DM-Bereich zur Bearbeitung von Dokumenten vorgestellt.

Versionsverwaltung Bei der Erstellung oder Bearbeitung von Dokumenten entstehen verschiedene Versionen. Die DM-Komponenten halten die verschiedenen Versionen des Dokuments bereit und erlauben die Nachvollziehbarkeit von Änderungen. Bei der Versionierung existieren mehrere Konzepte wie lineare oder verzweigte Versionierung. Üblicherweise wird eine verzweigte Versionierung verwendet, die eine Erstellung von Haupt- und Nebenversionen erlaubt. Dadurch lassen sich unterschiedliche Änderungen besser dokumentieren, indem z. B. wichtige Änderungen oder Erweiterungen von Dokumenten mit einer Hauptversion und kleinere Änderungen durch eine Nebenversion gekennzeichnet werden.

Check-In und Check-Out stellen ein zentrales Konzept zur Kontrolle von gleichzeitigen Zugriffen auf Dokumente dar. Durch den Check-Out eines Dokuments wird sichergestellt, dass nur ein Bearbeiter Änderungen an einem Dokument vornehmen kann, indem es exklusiv für diesen zur Bearbeitung bereitgestellt wird. Weitere Bearbeiter können während dieser Zeit lediglich lesend auf das Dokument zugreifen. Durch einen Check-In werden die vorgenommenen Änderungen an dem Dokument wieder eingebracht. Dabei wird i. d. R. eine neue Version des Dokuments erstellt.

Rendition-Verwaltung Als Rendition wird die Erzeugung einer Kopie eines Dokuments in einem anderen Format bezeichnet. Die Nutzung dieser Funktion kann etwa zur Erstellung von schreibgeschützten Kopien verwendet werden, die an externe Empfänger versendet werden, können oder zur Transformation in ein anderes Dateiformat, um auf die Informationen ohne die Quellenwendung zugreifen zu können. Die Beziehung zum Originaldokument wird vom DMS sichergestellt.

Mobile Nutzung Durch die wachsende Mobilität von Mitarbeitern in Unternehmen sollen die Daten aus dem DM-Bereich auch unterwegs auf Reisen und unabhängig von einer Netzwerkverbindung zur Verfügung stehen. Dafür werden Funktionen bereitgestellt, die es ermöglichen, Teile des Datenbestands offline auf die Festplatte eines Notebooks oder auf optische Datenträger wie CD oder DVD auszulagern. Der Benutzer kann damit auch unterwegs auf alle offline verfügbar gemachten Daten zugreifen. Die Erstellung und Bearbeitung von Dokumenten kann ebenfalls unterstützt werden. Nach der Rückkehr an den Arbeitsplatz werden die neu erstellten Dokumente dann in den Datenbestand eingebracht. Bei Änderungen an bereits bestehenden Dokumenten muss ggf. organisatorisch eingegriffen werden, falls in dieser Zeit Änderungen eines anderen Benutzers an den gleichen Dokumenten vorgenommen wurden. Hierzu muss entschieden werden, welche Änderung eingebracht werden soll.

Virtuelle Aktenbildung Eine virtuelle Akte stellt eine Sicht auf zusammengehörige Dokumente bereit. Das dahinterliegende Konzept basiert darauf, dass Dokumente aus verschiedenen Speicherorten oder aus einem integrierten Drittsystem visuell als zusammengehörig dem Benutzer präsentiert werden. Dadurch können unterschiedliche Anforderungen an die Zusammenstellung der Informationen aus DM-Sicht realisiert werden.

Suche Die Recherchemöglichkeiten nach Dokumenten werden im wesentlichen von den Attributen, die bei der Erfassung angegeben wurden, dem Dokumentenformat und der Dokumentenindexierung vorgegeben. Für jeden Dokumenttyp steht i. d. R. eine eigene Suchmaske mit den passenden Attributfeldern zur Verfügung. Üblicherweise lassen sich die gesuchten Begriffe auch mit logischen Operationen miteinander verknüpfen. Wenn die DM-Komponenten das Dateiformat des Dokuments unterstützen, können für die Dokumente darüber hinaus Volltextindizes erstellt werden, um die Möglichkeiten einer Volltextsuche zur Verfügung zu stellen. Insbesondere die Volltextsuche führt bei einem großen Datenbestand dazu, dass die Relevanzrate der gefundenen Dokumente relativ niedrig ausfällt. Daher kommt einer genauen Attributierung und präzisen Indexierung eine große Bedeutung in Bezug auf die Qualität der Ergebnisse einer Suche zu. Je genauer die Indexierung mit der Recherche übereinstimmt, desto einfacher und schneller lassen sich auch die gewünschten Dokumente wiederfinden. Moderne Suchkomponenten erlauben auch die automatische Suche nach Synonymen oder Begriffen in anderen Sprachen. Beispielsweise würden „KFZ“, „car“ und „Fahrzeug“ gefunden werden, wenn nach dem Begriff „Auto“ gesucht wird.

Überschneidungen mit anderen Bereichen

Die Funktionen des DM überschneiden sich teilweise deutlich mit denen der anderen Manage-Komponenten. Besonders die RM¹⁸-Komponenten besitzen viele Überschneidungen mit DM, indem sie ebenfalls Funktionen zur Verwaltung von Dokumenten während des Lebenszyklus abdecken oder die Beziehungen von Dokumenten und Akten abbilden. Aber auch die enge Integration mit anderen Anwendungen wie Office-Programmen oder Groupware-Systemen lassen schnell die Grenzen zwischen DM und Collaboration verwischen.

2.3.2.2. Records-Management

Der Bereich RM umfasst Funktionen und Komponenten zur Verwaltung von Records, wichtigen aufbewahrungspflichtigen bzw. aufbewahrungswürdigen Informationen.

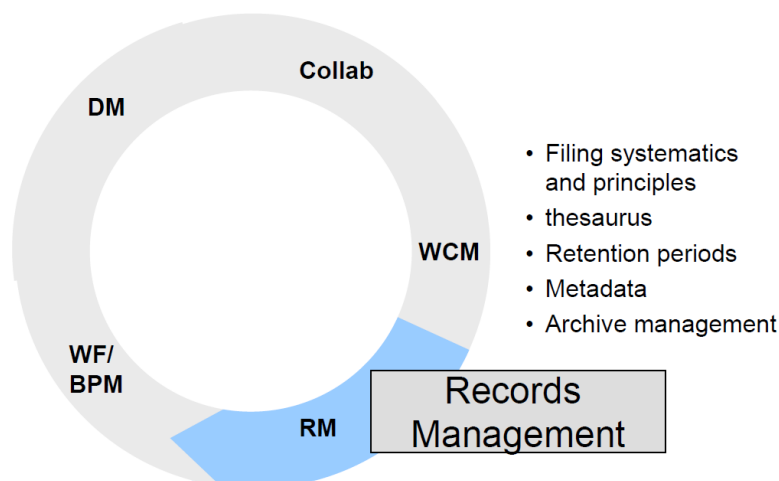


Abbildung 2.8.: Records-Management [KSP07]

Die Funktionen gehen über die Versionsverwaltung des Bereichs DM hinaus, indem sichergestellt wird, dass Änderungen an Dokumenten und Records mit den Benutzerinformationen und den vorgenommenen Änderungen zuverlässig protokolliert werden. Dadurch werden durch RM Nachweise zur Einhaltung von gesetzlichen Vorgaben, die etwa von Behörden oder durch Gesetze auferlegt werden, ermöglicht. Da die RM-Komponente diese Richtlinien übergreifend für alle Inhalte definiert, können auch Nachweise über den Stand von Webinhalten zu bestimmten Zeitpunkten erbracht werden.

¹⁸Records Management

Das RM hat seinen Ursprung in der Verwaltung von Papierakten und Archiven. Dabei wurden die Papierunterlagen kategorisiert, indexiert und mit dem genauen Standort der Papierakte in einer Datenbank gespeichert. Aufbewahrungsfristen werden anhand der Dokumentenart bestimmt und der Archivverwalter wird benachrichtigt, wenn eine Papierakte zur Aussonderung und Vernichtung ansteht. Die gleichen Konzepte dieser Verwaltung lassen sich auch auf elektronisch gespeicherte Informationen übertragen. Sie werden von diesem Funktionsbereich für umfassende Funktionen zur Steuerung und Protokollierung der Zugriffe und Verwendung der Dokumente genutzt. Zum Funktionsumfang der RM-Komponenten gehören:

Geordnete Ablage von Informationen in strukturierten Verzeichnissen und Aktenplänen: Dies umfasst die Umsetzung von Ordnungssystemen, in denen die Records aufbewahrt und verwaltet werden. Je nach Unternehmen oder Branche müssen hierdurch rechtliche Anforderungen an die Ablage von Unterlagen und die Dokumentation von Prozessen abgedeckt werden. So sind Aktenpläne für öffentliche Verwaltungen häufig vordefiniert und müssen im System gemäß den Vorgaben abgebildet werden. Für Unternehmen existieren i. d. R. keine konkreten Vorgaben. Hier können bereits vorhandene Ablagesysteme analysiert werden, um mögliche Ordnungssysteme zu entwickeln [Gau03]. Daneben existieren auch einige, meist länderspezifische Standards oder Spezifikationen, anhand derer Ordnungs- und Ablagesysteme aufgebaut werden können.

Definition von Recordschemata: In diesem Bereich wird der Aufbau der konkreten Records mit den zugehörigen Attributen, Metadaten und Dokumenten definiert, die in der Ablage als Records abgelegt werden. Dabei kann systemseitig festgelegt werden, welche Records eines bestimmten Schemas an welchen Stellen im Aktenplan gespeichert werden, um fehlerhaftes Einsortieren von Records zu vermeiden.

Thesaurus-gestützte Indexierung von Informationen: Durch einen Thesaurus werden Begriffe definiert, um die zu speichernden Informationen eindeutig zu indizieren. Damit soll erreicht werden, dass die Bezeichnungen innerhalb des Ablagesystems oder Recorderigenschaften einer einheitlichen Terminologie im ganzen Unternehmen folgen. Der Thesaurus muss in regelmäßigen Abständen auf Vollständigkeit hin überprüft werden. Häufig finden sich internationale oder branchenspezifische Standards für Bezeichnungen, die hier integriert werden können.

Definition und Verwaltung von Aufbewahrungs- und Vernichtungsfristen: Aufbewahrungs- und Vernichtungsfristen dienen der Einhaltung gesetzlicher oder unternehmensinterner Vorgaben bzgl. der Aufbewahrung von Records. Sie bilden die Basis für die Aussonderung von Records, die nicht mehr aufbewahrt werden sollen oder evtl. auch nicht mehr aufbewahrt werden dürfen. Letzteres tritt vor allem durch Anforderungen aus dem Datenschutzbereich auf. Die Aussonderung von Records muss über die RM-Komponenten genau protokolliert werden. Sie stellen Funktionen bereit, um die jeweiligen Bereiche im Ablagesystem autorisierten Personen zuzuordnen, die

für Aufbewahrung und Vernichtung verantwortlich sind. Die zur Aussonderung anstehenden Records werden über die Store- und Preserve-Komponenten physisch entfernt.

Schutz von Informationen gemäß ihrer Eigenschaften: Dieser Bereich deckt die Definition von Zugriffsberechtigungen und Änderungen an Records ab. Dabei genügt es nicht, dass Berechtigungen nur auf Verzeichnisse oder ganze Records im Ablagesystem vergeben werden, sondern die Berechtigungssysteme müssen Zugriffsberechtigungen detaillierter bis auf einzelne Inhaltskomponenten anhand inhaltlicher Kriterien unterstützen, etwa dass ein Sachbearbeiter nur Rechnungen mit einem Rechnungsbetrag unter 10.000 € einsehen darf.

Protokollierung sämtlicher Änderungen und Zugriffe: In diesem Bereich werden die Protokollierungen von Änderungen und Zugriffen definiert. Der Grad der Protokollierung hängt hier vor allem von den Inhalten der Records und gesetzlichen Vorgaben ab. Sie dienen vor allem der Nachweisbarkeit bestimmter Sachverhalte. Sie dienen beispielsweise dem Nachweis, dass Auflagen bzgl. des Datenschutzes beim Zugriff und der Verwendung von gespeicherten Informationen eingehalten werden.

Rechtliche Aspekte

Inzwischen greifen zahlreiche gesetzliche Vorschriften direkt in wesentliche Bereiche der Aufbewahrung und Unternehmensprozesse ein. So finden sich aufbewahrungsrelevante Vorschriften im HGB¹⁹, AO²⁰, GoB²¹, GoBS²² und den GDPdU²³. Außerdem existieren weitere internationale Gesetze und Regelungen wie SOX²⁴, die direkt wirksam sein können. Diese Regelungen gelten medienunabhängig und setzen teilweise voraus, dass die Daten im Originalformat aufzubewahren sind (E-Mails, elektronische Rechnungen). Dadurch müssen sich die im RM-Bereich definierten Richtlinien auch übergreifend auf alle Komponenten umsetzen lassen.

Normen und Spezifikationen

Im Bereich des RM existieren einige Normen und Spezifikationen, die mit Hilfe der RM-Komponenten umgesetzt werden können. Die wichtigsten Spezifikationen hierzu sind die *ISO 15489 Information und Dokumentation - Schriftgutverwaltung* [ISO02],

¹⁹Handelsgesetzbuch

²⁰Abgabenordnung

²¹Grundsätze ordnungsgemäßer Buchführung

²²Grundsätze ordnungsgemäßer DV-gestützter Buchführung

²³Grundsätze zum Datenzugriff und zur Prüfbarkeit digitaler Unterlagen

²⁴Sarbanes-Oxley Act

DoD5015.2²⁵ [DoD07] und die MoReq2²⁶ [Eur08a]. Diese Spezifikationen und Standards bieten eine gute Orientierung bei der Verwaltung von Records, und es empfiehlt sich die eigenen Records anhand dieser Vorgaben zu verwalten, falls dies gesetzliche Vorgaben nicht bereits erfordern [Ada07].

2.3.2.3. Web-Content-Management

In den letzten Jahren hat sich das Internet zu einem Medium entwickelt, über das immer mehr Informationen ausgetauscht werden. Firmenwebseiten bestehen heutzutage nicht mehr nur aus einer Selbstdarstellung des Unternehmens und Beschreibungen der angebotenen Produkte, sondern es werden weitergehende Informationen und Funktionen darüber bereitgestellt, wie Abfragen von Lieferterminen, der Produktionsstatus einer bestellten Ware oder die Reklamationsabwicklung.

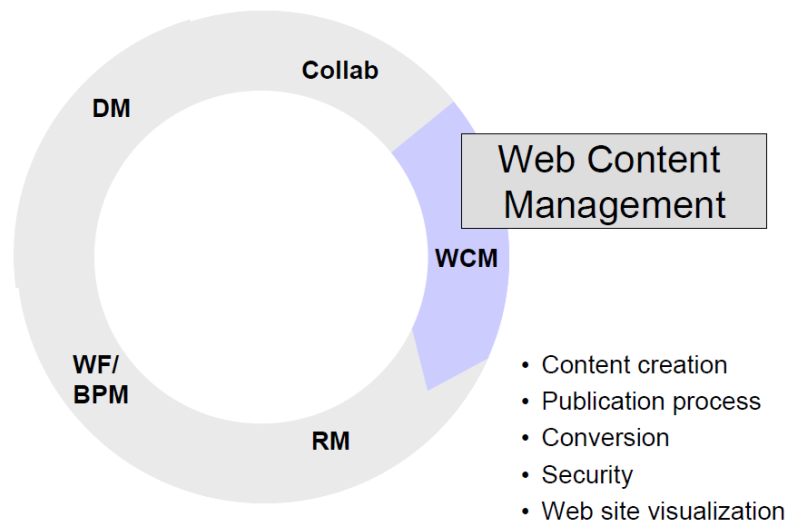


Abbildung 2.9.: Web-Content-Management [KSP07]

Für viele dieser Funktionen werden dafür Daten und Dokumente, wie Produktdatenblätter, Bedienungsanleitungen, Kundendaten oder Belege aus ERP-Systemen benötigt, um die entsprechenden Webinhalte bereitstellen zu können. Die meisten dieser Informationen liegen bereits in elektronischen Archiven oder in Dokumenten-Repositories vor. Daher ist es naheliegend im Rahmen von ECM, das als zentrales Repository für diese Informationen agiert, die Funktionen des WCM einzubeziehen, um die bereits vorhanden Informationen direkt in die Webinhalte, ohne Brüche in den Anwendungen, integrieren zu können. Eine Aktualisierung eines Dokuments könnte so auch di-

²⁵Department of Defense 5015.2-STD

²⁶Model Requirements for the Management of Electronic Documents and Records 2

rekt über die Webinhalte abgebildet werden, ohne dass eine manuelle Aktualisierung erforderlich wird.

Für die Bereitstellung von Inhalten in Web-Portalen, Intranet und im Internet müssen die Informationen allerdings in geeignete Formate transformiert werden, um sie über den Browser und ggf. gängige Zusatzprogramme, wie einen PDF-Reader anzeigen zu können. So ist den meisten Empfängern wenig geholfen, wenn sie eine technische Zeichnung eines Geräts im Dateiformat der CAD-Anwendung erhalten, welches sie nicht öffnen können. Außerdem ist es aus Sicht des Unternehmens nicht erwünscht, dass diese Daten veränderbar im Internet bereitgestellt werden.

Um Webinhalte erstellen und verwalten zu können, stellt die WCM-Komponente ein Redaktionssystem für die Gestaltung und Struktur von Webinhalten bereit. Im Redaktionssystem wird zunächst ein einheitliches Layout für die verschiedenen Webseiten definiert, in die die Inhalte dann eingebracht werden können. Um den Inhalt von Dokumenten auf der Webseite zu veröffentlichen oder die Dokumente in einem geeigneten Format bereitzustellen, werden dafür Transformationstechniken benötigt, die die Informationen aus Dokumenten extrahieren und in XML oder HTML²⁷ bzw. vom Quelldateiformat in das gewünschte Zielformat transformieren. Dies kann etwa bei Bildformaten notwendig sein, da heutige Digitalkameras durch ihre hohe Auflösung schnell in Bereiche von einigen MB pro Bild vordringen. Dies führt bei direkter Verwendung, etwa in einer Bildergalerie jedoch dazu, dass sehr große Datenmengen bei jedem Aufruf der Seite übertragen werden müssen und dadurch der Seitenaufbau für den Benutzer sehr langsam wird. Daher sollten diese Bilder automatisch komprimiert werden können und somit am besten eine zweite Rendition des Bildes anfertigen, die beim Aufruf der Seite angezeigt werden kann. Das hochaufgelöste Bild kann dann auf Anforderung dem Benutzer trotzdem angeboten werden.

Da Webinhalte öffentlich zugänglich sind und dadurch auch rechtsverbindliche Aussagen an Dritte gegeben werden, ist es erforderlich, dass Kontrollfunktionen zur Veröffentlichung bereitgestellt werden. Hierfür werden WF-Funktionen benötigt, auf die in Abschnitt 2.3.2.5 eingegangen wird. Mit ihnen lassen sich Abläufe definieren, mit denen Inhalte genehmigt und veröffentlicht werden können. Darüber hinaus ermöglichen diese Abläufe, dass dokumentiert wird, wer für die Erstellung, Genehmigung und Veröffentlichung der Inhalte verantwortlich war. WCM muss auch besonders die sichere Trennung zwischen öffentlichen und nicht-öffentlichen Inhalten gewährleisten.

Portale

Web-Portale sind heutzutage zu einem weit verbreiteten Werkzeug für Integrationsaufgaben geworden. Sie dienen dem Ziel, Informationen aus verschiedenen Quellen

²⁷Hypertext Markup Language

zusammenfassend dem Benutzer zu präsentieren. Dies können außer in Dokumenten gespeicherte unstrukturierten Informationen aus dem ECM-Bereich selbst, auch weitere Daten aus anderen Systemen z. B. CRM²⁸- oder ERP-Systeme sein. Weiterhin bieten Portale den Benutzern die Möglichkeit, mit einer Anmeldung mit den benutzerspezifischen Rechten auf verschiedene Systeme und Daten zuzugreifen, sodass mehrfache Anmeldungen an diversen Systemen mit ggf. unterschiedlichen Authentifizierungen entfallen.

Durch Anwendungsintegrationen oder eingebettete Komponenten aus Groupwareanwendungen wie E-Mail, Kalender, Chats oder gemeinsamen Arbeitsbereichen können Portale außerdem die Bereitstellung der Arbeitsumgebung für die Benutzer auf Administrationsseite vereinfachen, da Anwendungen nicht bis an den Computer des Benutzers ausgerollt werden müssen. Damit können Funktionen aus dem Bereich Collaboration übernommen werden (s. Abschnitt 2.3.2.4). Durch die Verwendung von Internet-technologien lassen sich die Portale auch relativ einfach durch Firewalls außerhalb des Unternehmensnetzwerkes zugänglich machen, so dass Benutzer auch von unterwegs mit ihrer gewohnten Arbeitsumgebung und mit Livedaten versorgt werden können. Für den Benutzer können auch unterschiedliche Portale, je nach Arbeitssituation, angeboten werden, wie z. B. Wissens- oder Weiterbildungsportale.

Da die Inhalte bei größeren Portalen von vielen Personen bearbeitet und erstellt werden, bedarf es Funktionen zur kontrollierten Erstellung, Bearbeitung und Veröffentlichung. Die dafür notwendigen Funktionen sind eine einfache und intuitive Oberfläche zur Erstellung von Inhalten und Layoutvorgaben für Webinhalte, Funktionen zur Strukturierung und Auszeichnung der XML- bzw. HTML-Inhalte, Integration in DMS-Funktionen wie Check-In/Check-Out, Versionierung, Berechtigungen und die Integration und Verlinkung auf vorhandene Informationen wie z. B. Dokumente aus dem DM-Bereich. Weiterhin wird die Integration in WF-Funktionen für die Erfassung, Genehmigung, Freigabe, Archivierung oder den Rückruf von Informationen auf der Webseite benötigt [GSSZ02].

2.3.2.4. Collaboration

Der Collab²⁹-Bereich soll vor allem die Arbeit in Gruppen unterstützen und verbessern.

Speziell in den letzten Jahren haben Collaboration-Funktionen durch die rasante Verbreitung von schnellen Internetverbindungen und der direkten Anbindung von mobilen Benutzern und Niederlassungen bei größeren Unternehmen eine deutlich gestiegene Bedeutung erfahren. Dies ist darauf zurückzuführen, dass im Zuge der Globalisierung heutzutage viel häufiger weltweit verteilte Teams gemeinsam an Projekten

²⁸Customer Relationship Management

²⁹Collaboration

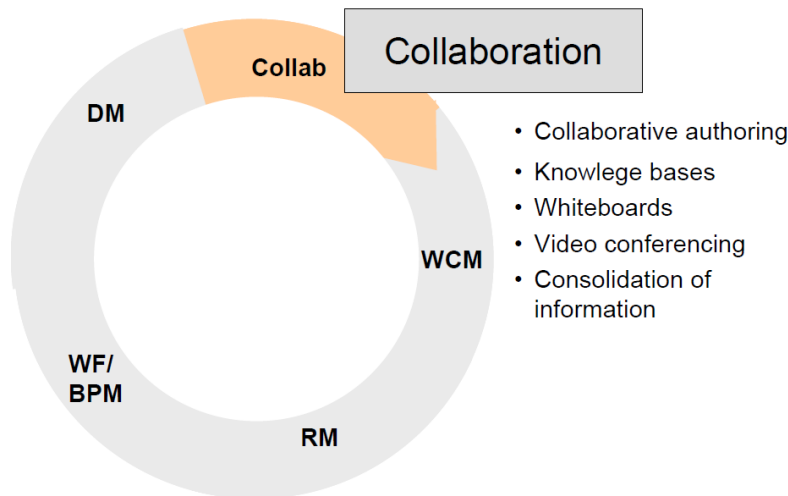


Abbildung 2.10.: Collaboration [KSP07]

arbeiten und die Zeit- und Kostenaufwendungen für Geschäftsreisen reduziert werden sollen.

Dazu gehören die Bereitstellung von gemeinsamen Arbeitsbereichen zum gleichzeitigen und kontrollierten Bearbeiten und Erstellen von Informationen, Informationssammlungen wie Knowledgebases, Skill-Datenbanken, Informationsressourcen und Hintergrundinformationen. Weiterhin umfasst der Bereich Verwaltungskomponenten wie Whiteboards zur Ideensammlung und Diskussionsforen zum Ideenaustausch von Projektmitgliedern oder Anwendungen wie Instant Messaging oder Videokonferenzen zur direkten Kommunikation von Mitarbeitern untereinander, sowie Koordinationskomponenten zur Termin- und Projektplanung mit integrierten Projektmanagementfunktionen.

Diese Funktionen erfordern einen direkten Zugriff auf Dokumente aus dem DMS und Koordinationsfunktionen, die meist in Groupwareanwendungen zu finden sind und integriert werden können. So können beispielsweise Termine, die in gemeinsam genutzten Terminkalendern zur Planung festgelegt werden, direkt in die in Groupwaresystemen vorhandenen persönlichen Kalender der Benutzer übertragen werden. Darüber hinaus sollen häufig auch externe Mitarbeiter oder Partner in die Projektarbeit eingebunden werden. Dadurch ergeben sich weitere Anforderungen an möglichst geringe Änderungen auf Clientseite, da es nicht praktikabel ist, Software auf Rechnern zu installieren, die nicht zum eigenen Unternehmen gehören. Aus diesen beiden Anforderungen bietet sich hierfür der Einsatz von Portalen an, in die die oben genannten Collab-Funktionen eingebracht werden, statt eigene Programme für diesen Bereich zu entwickeln.

2.3.2.5. Workflow/Business Process Management

Der Bereich WF/BPM³⁰ verfolgt das Ziel, Geschäftsprozesse und Abläufe systemseitig zu unterstützen, indem Geschäftsprozesse komplett oder teilweise automatisiert werden und dabei Dokumente, Informationen oder Aufgaben von einem Bearbeiter zum nächsten weitergereicht werden [Fis07].

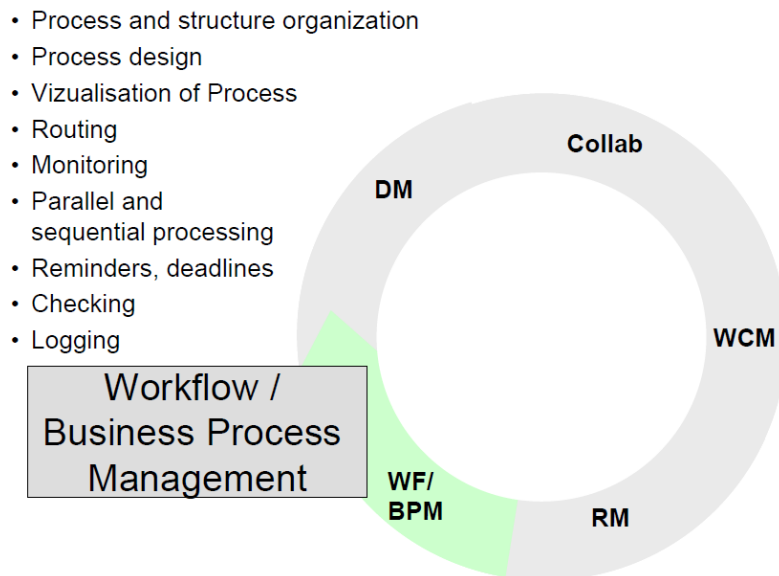


Abbildung 2.11.: Worklow/BPM [KSP07]

Workflows können in zwei Kategorien eingeteilt werden, die sich in der Art der Strukturierung und Häufigkeit ihrer Ausführung unterscheiden [Kam03b]:

Adhoc-Workflow: Adhoc-Workflows besitzen keine fest definierten Abläufe. Der Anwender definiert zur Laufzeit den Prozess, den Ablauf und die erforderlichen Daten. Üblicherweise können die Abläufe im Verlauf des Prozesses auch durch die weiteren beteiligten Benutzer abgeändert werden.

Production-Workflow: Auf Basis vordefinierter Abläufe werden Prozesse definiert und anhand der Vorlage ausgeführt und gesteuert. Ausnahmen bei diesen Workflows sind ausschließlich über vordefinierte Abweichungen möglich.

Die Abläufe der Prozesse werden zunächst im System modelliert. Die Ausführung kann über eine Workflow-Engine oder eine Workflow-Anwendung mit eigenständigem Client, realisiert werden. Die Workflow-Engine sorgt für die Ausführung der einzelnen Prozessschritte. Außerdem stellt sie sicher, dass die Daten, die zur Ausführung

³⁰Business Process Management

der Prozessschritte benötigt werden, den Benutzern bei der Ausführung der einzelnen Schritte, zur Verfügung gestellt werden. Gartner definiert den Funktionsumfang der WF-Komponenten im Rahmen von ECM wie folgt [SBC⁺08]: „minimale Anforderungen sind einfache Dokumentenprüfungs- und Freigabeprozesse und weitergehend noch graphische Editoren für die Modellierung, sowie serielles und paralleles Routing von Dokumenten“.

Im Zusammenhang mit DM stellt WF Funktionen bereit, um dokumentenbezogene Prozesse zu unterstützen. Eine wichtige Aufgabe in dokumentenorientierten Prozessen kommt der automatisierten Verteilung von Dokumenten an die entsprechenden Empfänger zu. Dies beginnt bei der Bereitstellung der Dokumente nach der Erfassung in elektronische Posteingangskörbe der Benutzer. Je nach Art des Dokuments schließen sich bestimmte Genehmigungs- und Freigabeprozesse an. So müssen etwa beim Erfassen von Eingangsrechnungen die einzelnen Positionen auf der Rechnung von verschiedenen Mitarbeitern auf Korrektheit überprüft werden, ob alle Postionen geliefert und in einwandfreiem Zustand angekommen sind. Außerdem ist zu prüfen, ob alle Positionen mit dem korrekten Preis berechnet wurden und ggf. Sonderkonditionen oder die Zahlungsbedingungen korrekt sind. Erst wenn alle Punkte von den jeweils zuständigen Mitarbeitern bestätigt wurden, kann die Rechnung zur Zahlung an die Buchhaltungsabteilung weitergegeben werden, die mit der Zahlung und Buchung der Rechnung den Prozess abschließt.

Bei der Erstellung von Dokumenten schließt sich häufig ein Freigabeprozess an, bei dem das Dokument an einen oder mehrere Benutzer weitergeleitet werden muss, um eine Freigabe für einen bestimmten Versionsstand zu erhalten. Gegebenenfalls sind Rückmeldungen an den Ersteller zu senden, um Änderungen an dem Dokument vorzunehmen, damit das Dokument freigegeben werden kann. Die Bearbeitung von Anträgen oder Verträgen erfordern ebenfalls Prozesse, um die Bearbeitung durchzuführen. Bei jedem Bearbeitungsschritt müssen dem Bearbeiter alle notwendigen Dokumente und Informationen für die Bearbeitung zur Verfügung gestellt werden. Dabei muss auch der Ablauf nachvollziehbar bleiben, um den Verlauf von Prozessen analysieren und ggf. optimieren zu können.

Viele Geschäftsprozesse sind jedoch nicht dokumentenorientiert, wie etwa ein Bestellprozess. In diesen Prozessen dienen Dokumente lediglich als begleitende Informationsträger, oder sie lösen den Beginn bestimmter Prozesse aus. So dient bei Eingang eines Bestellformulars das Dokument nur als Prozessauslöser und stellt die Daten zu den bestellten Produkten und die Kundendaten zur Verfügung. Die hauptsächlichen Aktivitäten finden jedoch in anderen Systemen, z. B. ERP-Systemen statt, die den Bestellprozess oder die Produktion bzw. Bestellung der Artikel verwalten.

Zum Bereich der Production-Workflows gehören folgende Funktionen:

Definition und Ausführung von Prozessen: Die Prozessabläufe werden zunächst im System mit Hilfe eines Editors festgelegt. Dazu gehört die Festlegung von sequenziell oder parallel auszuführenden Prozessschritten. An Entscheidungspunkten, die den weiteren Ablauf von Prozessen beeinflussen, müssen Verzweigungen und die zugehörigen Entscheidungsregeln definiert werden, bei welchen Ereignissen oder Bedingungen welcher Zweig ausgewählt wird. Wiederkehrende Prozessschritte werden über Schleifen realisiert.

Außerdem muss definiert werden, welche Informationen zur Ausführung von Aktivitäten benötigt werden und wie diese Informationen vom System bereitgestellt werden können. Bei den einzelnen Prozessschritten muss hinterlegt werden, unter welchen Bedingungen ein Prozessschritt als abgeschlossen betrachtet werden kann. Ebenso muss sichergestellt werden, dass ein Prozessschritt erst ausgeführt werden kann, wenn alle Informationen, die zur Ausführung benötigt werden, auch vorhanden sind und vom System bereitgestellt werden können [ZGB05].

Verwaltungsfunktionen: Um Prozesse durchzuführen, müssen die klassischen Verwaltungsfunktionen wie Wiedervorlage, Fristen, Delegation, Kontrolle von Bearbeitungszuständen, sowie die Ergebnisse der Bearbeitung vom System unterstützt und dem Benutzer die entsprechenden Informationen bereitgestellt werden.

Überwachungs- und Auswertungsfunktionen: Sie ermöglichen die Überwachung von laufenden Prozessen. Der Verlauf von Prozessen wird visualisiert, womit jederzeit der Zustand analysiert werden kann, in dem sich Prozesse befinden. Über diese Funktionen kann das System auch Probleme und Fehler melden, die im Verlauf von Prozessen aufgetreten sind und die ggf. ein Eingreifen erfordern. Zum Umfang gehören auch Funktionen, um bestimmte Zeitüberschreitungen bei der Bearbeitung zu melden, und somit ein Eskalationsmanagement zu ermöglichen. Weiterhin stehen in diesem Bereich auch Funktionen zur Auswertung von abgeschlossenen und noch laufenden Prozessen zur Verfügung, um Statistiken und Daten zur Prozessoptimierung zu erhalten.

BPM: BPM geht über die Ansprüche von WF hinaus und zielt darauf ab, die Aufgaben komplett aus einer prozessorientierten Sicht durchzuführen. Dabei wird eine vollständige Integration aller dafür notwendigen Daten und Anwendungen angestrebt. Die Daten und Anwendungen obliegen der Kontrolle des BPM-Systems, das die Prozesse steuert. Zu BPM gehören Prozess- und Datenkontrolle durch den BPM-Server, vollständige Integration der zur Ausführung der Prozesse notwendigen Anwendungen innerhalb der Prozesse. Zur Ausführung werden bestimmte Anwendungsprogramme in die Prozessausführung integriert. Diese Anwendungen müssen dann automatisch vom System gestartet und beendet werden. Weiterhin müssen auch Daten mit den Programmen über Schnittstellen ausgetauscht werden, damit die Informationen im Prozess an die weiteren Prozessschritte übergeben werden können. Neben dem Ablauf der Prozesse muss in dieser Phase auch definiert werden welche Benutzer die jeweiligen Schritte ausführen und bearbeiten können und dürfen.

Im Unterschied zu WF umfasst BPM damit den kompletten Lebenszyklus von Prozessen zu unterstützen und ermöglicht eine nahtlose Dokumentationskette aller prozessrelevanter Daten. WF stellt einzelne Funktionen zur Unterstützung von Abläufen bereit. Diese sind häufig in andere Produkte wie DMS oder ERP-Systeme in die Oberflächen integriert und erlauben meist nur relativ starre und kurze Abläufe, die nicht immer in einen umfangreicheren Geschäftsprozess integriert sind.

2.3.3. Store

Die Store-Komponenten stellen Funktionen zur Speicherung von lebenden Dokumenten, Daten und Informationen zur Verfügung, die nicht oder noch nicht archivierungswürdig sind. Die Entscheidung, ob und welche Informationen archiviert werden, werden über die Manage-Komponenten gesteuert. Die von der AIIM genannten Speicherkomponenten lassen sich in die drei Bereiche Library-Services als Informationsverwaltungsdienste, Repositories als Speicherorte und Technologies als unterschiedliche Speichertechnologien unterteilen.

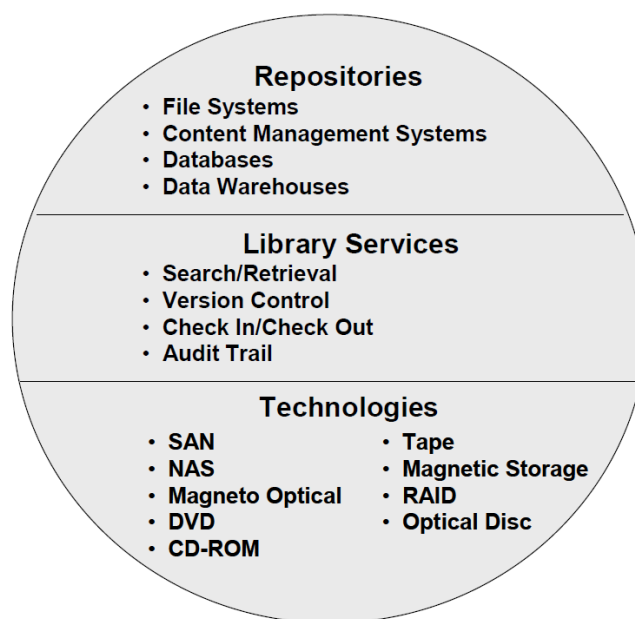


Abbildung 2.12.: ECM-Store [KSP07]

Library-Services

Die Library-Services stellen die systemnahen Komponenten, mit denen der Zugriff und die Speicherung der Informationen für die in den Abschnitten 2.3.1 und 2.3.2 vorge-

stellten Capture- und Manage-Komponenten realisiert werden. Sie verwalten sowohl die Speicherorte der dynamischen Ablage (Store), als auch derer im Langzeitarchiv (Preserve). Der genutzte Speicherort wird anhand der Eigenschaften und Klassifizierung der Informationen bestimmt. Durch die Integration der Manage-Komponenten werden durch die Library-Services die notwendigen Such- und Recherchefunktionen bereitgestellt.

Da die Manage-Datenbanken keine Kenntnis über die physischen Speicherorte besitzen, werden diese Verweise durch die Library-Services verwaltet. Dadurch wird eine transparente Verlagerung der Daten an einen anderen physischen Speicherort realisierbar. Typischerweise werden die Speicherorte in Onlinespeicher (Informationen im direkten Zugriff), Nearlinespeicher (Informationen, die sich in indirektem Zugriff befinden, wie DVDs, die von einem Roboter zunächst in ein Leselaufwerk eingelegt werden müssen) und Offlinespeicher (Informationen, die nicht im direkten Zugriff durch das System stehen, z. B. außer Haus gelagerte Bänder) eingeteilt. Weitere Aufgaben der Library-Services sind das Erstellen und Speichern von Protokolldaten des Zugriffs und Änderungen an den gespeicherten Informationen in Form von Audit-Trails³¹ und die Bereitstellung des Versionsmanagement zur Kontrolle der unterschiedlichen Stände und der Check-In-/Check-Out-Funktionen zur Kontrolle gleichzeitiger Zugriffe.

Repositories

Die zum Einsatz kommenden Datenspeichersysteme, die i. d. R. kombiniert eingesetzt werden, sind das vom Betriebssystem bereitgestellte Dateisystem, spezielle Content-Speichersysteme mit proprietären Speichertechniken, Datenbanken sowie Data Warehouses.

Dateisysteme werden primär zur Zwischenspeicherung von temporären Daten oder Caches verwendet. So werden etwa beim Massenimport von Datensätzen die Daten zunächst im Dateisystem zwischengespeichert, bevor die Daten in die eigentlichen Speichersysteme importiert werden. ECM verfolgt das Ziel, das Dateisystem als Speicherort von Informationen, etwa Freigaben, in denen Dateien abgelegt werden, abzulösen oder zumindest zu entlasten. Außerdem werden im Dateisystem typischerweise die Dokumente (ohne Metadaten) als Dateien oder in speziellen Containerformaten gespeichert, die später ggf. leichter an Langzeitarchivsysteme ausgelagert werden können. Beispielsweise können dies Containerdateiformate in einer bestimmten Dateigröße zur gezielten Auslagerung auf andere Speichermedien, wie CD oder DVD sein. Außerdem können sie zur einfacheren Organisation oder zur Einhaltung von Richtlinien genutzt werden, etwa alle E-Mails des Unternehmens eines Jahres. So lassen sich nach Ablauf der Aufbewahrungsfrist die zu löschenden Daten einfach ermitteln und entsorgen.

³¹ Audit-Trail ist ein Verfahren von Software-Systemen, bei dem Benutzer und ihre Aktivitäten über einen definierten Zeitraum überwacht und protokolliert werden.

Datenbanken werden vorrangig dazu verwendet die Zusatzinformationen zu den eigentlichen Daten zu speichern, auf die häufig und vor allem sehr schnell zugegriffen werden soll. Dies können Metadaten, Suchindizes oder Zugriffsberechtigungen auf die Daten sein. Zumeist werden hierfür Kombinationen aus relationalen Datenbanken, z. B. für Zugriffsinformationen, Versionsstände, festgelegte Attribute von Dokumenten und Volltextdatenbanken für die Speicherung von Volltextindizes verwendet. Sie können auch zur Speicherung der eigentlichen Daten verwendet werden. Große Datenmengen in Form von Binärdaten werden meistens nicht direkt in Datenbanken, sondern im Dateisystem oder speziellen Containerformaten abgelegt, um die relationale Datenbank klein und performant zu halten.

Data Warehouses sind komplexe datenbankbasierte Informationsverwaltungssysteme, die zur Aggregation von Daten aus unterschiedlichen Systemen und Datenquellen dienen. Häufig werden sie für Analysezwecke im betriebswirtschaftlichen Bereich zur Entscheidungshilfe genutzt.

Speichertechnologien

Der Bereich Speichertechnologien beschreibt die zur Speicherung der Informationen eingesetzten Speichermedien. Hier kommen je nach Anwendungszweck und Systemumgebung sehr unterschiedliche Technologien zum Einsatz.

Magnetische Onlinespeicher mit festplattenbasierten RAID³²-Systemen [PGK88] kommen heutzutage typischerweise in Form von direkt an den Server angeschlossenen Festplatten als DAS³³, direkt in Netzwerken zugänglich gemachten Speichersystemen als NAS³⁴ oder SAN³⁵-Systeme zum Einsatz, die als Speicherinfrastruktur bereitgestellt werden. Der Einsatz der jeweiligen Technologie hängt hier sehr stark von der bereits vorhandenen Infrastruktur und den Anforderungen an Speicherbedarf, Zugriffszeit, Datenreplikation und Verfügbarkeit der Daten ab.

Magnetbänder bieten insbesondere in Kombination mit Bandrobotern hohe Speicherkapazitäten zu niedrigen Kosten. Da auf Magnetbänder nur sequentiell zugegriffen werden kann, bieten sie sich vor allem für die Speicherung von großen Datenmengen an, auf die selten zugegriffen wird. Magnetbänder existieren sowohl als wiederbeschreibbare Medien oder als WORM³⁶-Medien. Dadurch können mit einer Technologie sowohl die Anforderungen an Unveränderbarkeit der Daten für die Archivierung auf dem Medium, als auch an wiederbeschreibbare Medien z. B. für Backupzwecke erfüllt werden.

³²Redundant Array of Inexpensive Disks

³³Direct Attached Storage

³⁴Network Attached Storage

³⁵Storage Area Network

³⁶Write Once Read Many

Optische Speicher dienen in erster Linie zur Speicherung von Archivdaten. Durch die Verwendung von WORM-Medien, wie CD-R oder DVD-R werden die Anforderungen an Unveränderbarkeit der Daten auf dem Medium für Archivierungszwecke erfüllt. Durch relativ kurze Zugriffszeiten durch einen wahlfreien Zugriff und den Einsatz von Jukeboxen können die Zugriffszeiten für die Daten relativ gering gehalten werden und damit für selten genutzte Daten durchaus eingesetzt werden.

Durch den starken Preisverfall und die erreichbaren Kapazitäten von Festplatten in den letzten Jahren werden Magnetbänder oder optische Speichermedien heutzutage kaum noch für die Speicherung von lebenden Dokumenten und Informationen eingesetzt.

2.3.4. Preserve

Der Bereich der Archivierung dient der sicheren und langfristigen, statischen und unveränderbaren Aufbewahrung und Sicherung der Informationen. Die Richtlinien zu Aufbewahrungszeiten und Zugriffen auf diese Daten werden im Bereich des RM festgelegt.

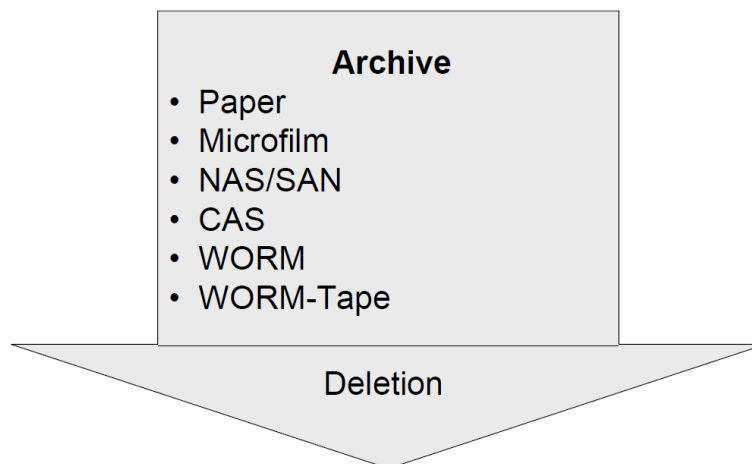


Abbildung 2.13.: ECM-Preserve [KSP07]

Die Archivierungskomponenten regeln die technischen Aspekte, um die Geräte und die entsprechenden Medien für die Archivierung transparent für die anderen Komponenten zur Verfügung zu stellen. Hierzu gehören Medien, die zum einen eine hohe Lebenserwartung und sichere Datenspeicherung über Jahre hinweg, und zum anderen die Unveränderlichkeit der gespeicherten Informationen auf Speichermedienebene gewährleisten.

Optische Medien werden hauptsächlich in diesen Bereichen in Form von CD-R oder DVD-R eingesetzt, die aufgrund ihrer technischen Beschaffenheit lediglich einmal beschreibbar sind. Für den Nearlinezugriff lassen sich hierfür Jukeboxen einsetzen, die das Wechseln von Medien durch einen Roboter gesteuert, regeln. Die Lebensdauer von optischen Medien wird von den Herstellern mit etwa 30-100 Jahren angegeben [All09].

Magnetbänder sind seit einigen Jahren auch in speziellen Versionen verfügbar, die die WORM-Eigenschaften sicherstellen. Dies wird bei LTO über die Firmware des Bandlaufwerks in Kombination mit dem in jedem LTO-Band vorhanden Cartridge-Memory-Chip realisiert, der seit LTO-3 einen WORM-Marker erhalten hat. Der Einsatz von Magnetbändern resultiert daraus, dass Bandlaufwerke und entsprechende Roboter für das automatische Wechseln der Medien in den meisten mittleren und größeren Unternehmen bereits vorhanden sind und für Datensicherungszwecke genutzt werden. Damit können die Geräte für Backup- und Archivierungszwecke genutzt werden. Außerdem erreichen Magnetbänder deutlich höhere Kapazitäten (bis zu 1,6 TB), als optische Medien (bis zu 30 GB) und eignen sich daher bei großen Archiven besser. Magnetbänder besitzen jedoch eine kürzere Lebenserwartung in Bezug auf die Lesbarkeit der gespeicherten Informationen. Daher wird ein häufigeres Umkopieren der Datenbestände erforderlich, was jedoch durch die fortschreitende technische Entwicklung und die damit verbundene Kapazitätssteigerung meistens dazu führt, dass dadurch die Anzahl der Medien reduziert werden kann.

Content Archive Systems Mit dem Preisverfall von Festplatten und der stark gestiegenen Kapazitäten sind seit wenigen Jahren auch festplattenbasierte Archivsysteme verfügbar, die z. B. von IBM, NetApp oder EMC angeboten werden (s. Abbildung 2.14). Die Ausfallsicherheit wird dabei durch geeignete RAID-Systeme und Datenreplikation realisiert. Für Nachweise zur Unveränderlichkeit der archivierten Daten werden Prüfmechanismen wie Hashverfahren eingesetzt, die die Integrität der Daten gewährleisten oder zumindest Änderungen erkennbar machen [CLR01]. Der physische Speicherort wird bei diesen Systemen häufig transparent für das System gehalten, indem z. B. bei CAS³⁷-Systemen der Zugriff auf die gespeicherten Informationen nicht über die Speicheradresse auf dem System, sondern über den Inhalt der Informationen erfolgt. Wenn ein neues Dokument gespeichert wird, berechnet die Speichereinheit des CAS-Systems einen Hash-Wert der Datei und gibt diesen als Inhaltsadresse an das angebundene System zurück. Der gebildete Hash-Wert dient zur Überprüfung, dass ein identischer Inhalt nicht ein zweites Mal gespeichert wird. CAS-Systeme arbeiten besonders effizient mit Daten, die nur selten geändert werden. Sie können außerdem den Zugriff auf Dokumente beschleunigen und stellen sicher, dass das gefundene Dokument mit dem Original übereinstimmt. CAS-Systeme sind vor allem für die Speicherung von sehr großen Dokumentenmengen (bis in den Petabytebereich) mit hohen Zugriffsraten und der Forderung nach kurzen Antwortzeiten geeignet [EMC08].

³⁷Content Addressed Storage



Abbildung 2.14.: IBM DR 550 Archiv-System [IBM09b]

NAS und SAN-Systeme lassen sich ebenfalls ähnlich wie die CAS-Systeme mit geeigneter Software um die geforderten WORM-Eigenschaften erweitern. Der Einsatz dieser Systeme zur Archivierung begründet sich meist darauf, dass solche Systeme bereits für die Speicherung von veränderbaren Daten und anderen Systemen im Einsatz sind und somit das vorhandene Wissen bei der Administration der Systeme genutzt werden kann.

Mikrofilm und Papier können nach wie vor auch zur Archivierung von Informationen genutzt werden. Für elektronische Informationen werden sie heute jedoch eher weniger in Betracht gezogen, da die auf diesen Medien gespeicherten Informationen nicht maschinell auswertbar und die Zugriffsmöglichkeiten ebenfalls beschränkt sind. Außerdem sehen inzwischen einige Gesetze vor, dass originär digitale Unterlagen auch im Originalformat elektronisch aufbewahrt werden müssen, z. B. im Bereich der E-Mailarchivierung nach GDPdU.

Migration

Ein wichtiger Bereich, der bei der elektronischen Archivierung zu beachten ist, ist die Migration auf andere, neuere Speichertechnologien. Die Aufbewahrungszeiten in Archiven variieren sehr stark, abhängig von individuellen Aufbewahrungszeiten und ge-

setzunglichen Anforderungen. Beispielsweise sind bei Daten aus dem medizinischen Bereich oft Aufbewahrungsfristen von bis zu 30 Jahren einzuhalten oder in speziellen Bereichen, wie etwa Grundbuchämtern, ohne zeitliche Begrenzung. Insbesondere das letztgenannte Beispiel macht deutlich, dass eine Migration der Daten auf andere Datenträger und Speichertechnologien beachtet werden muss. Dabei müssen die Archivierungskomponenten in der Lage sein, die Daten für die anderen Komponenten transparent und sicher auf andere Medien zu übertragen. Durch die technische Entwicklung ist es hierbei auch notwendig, die Unterstützung an neue Speichertechnologien und Geräte anzupassen. Der Migrationsprozess ist kontinuierlich in regelmäßigen Abständen durchzuführen.

Hierarchisches Speicher-Management

HSM³⁸ beschreibt eine Speicherstrategie, um die Informationen kosteneffizient entlang ihres Lebenszyklus auf den jeweils günstigsten Medien zu speichern. Dadurch können Anforderungen an Aufbewahrungsfristen, erforderlichen Zugriffsgeschwindigkeiten, gesetzlichen Anforderungen oder vertraglich zugesicherte Zugriffszeiten durch die geeignete Speicherung auf den entsprechenden Medien erfüllt werden. So werden neue Informationen bei der Speicherung zunächst auf dem schnellsten Medium gespeichert und bei Erreichen einer gewissen Speicherdauer von vergleichsweise teuren Festplatten in SAN- oder NAS-Systemen auf kostengünstigere Medien, wie Band oder optische Medien migriert. Da die Zugriffshäufigkeit mit zunehmendem Alter der Daten auch abnimmt, können längere Zugriffszeiten hier leichter toleriert werden.

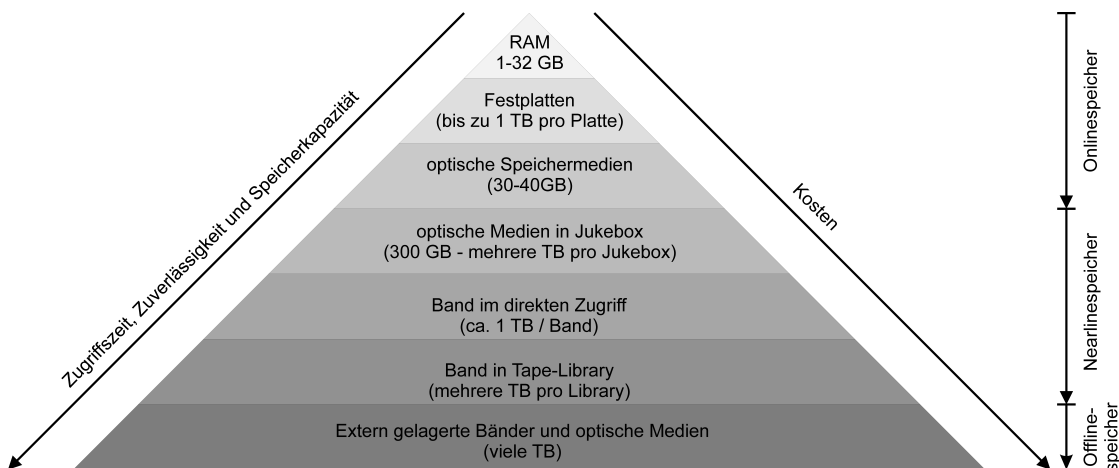


Abbildung 2.15.: Hierarchisches Speicher-Management nach [GSSZ02]

³⁸Hierarchisches Speicher-Management

Eine Übersicht über die Speichermedien, die in einem HSM-System vorhanden sind, sind in Abbildung 2.15 veranschaulicht und in drei Bereiche unterteilt. Onlinespeicher werden für sehr schnelle Zugriffe auf den Hauptspeicher und schnelle Magnetplatten verwendet. Nearlinespeicher kommen für Speichersysteme zum Einsatz, auf die nicht direkt zugegriffen werden kann, die jedoch mit Hilfe einer vorherigen Ansteuerung eines Roboters online gesetzt werden können. Offlinespeicher stellen das untere Ende der Hierarchie dar und bestehen aus extern gelagerten Bändern oder optischen Speicherplatten. Die Kosten nehmen dabei beim Abstieg in der Hierarchie ab, während Zugriffszeiten, Zuverlässigkeit und die Speicherkapazitäten zunehmen.

2.3.5. Deliver

Die Deliver-Komponenten dienen der Bereitstellung der Informationen für die Komponenten aus den Bereichen Manage, Store und Preserve. Weiterhin dienen sie der Übergabe von Daten an weitere Systeme, etwa der Informationsübergabe mit Datenträgeraustausch oder der Erzeugung von formatierten Ausgabedateien für den Druck. Ebenso gehören Funktionen zur Konvertierung oder Kompression in diesen Bereich. Der Bereich Deliver ist daher häufig auch unter dem Begriff *Output Management* zu finden.

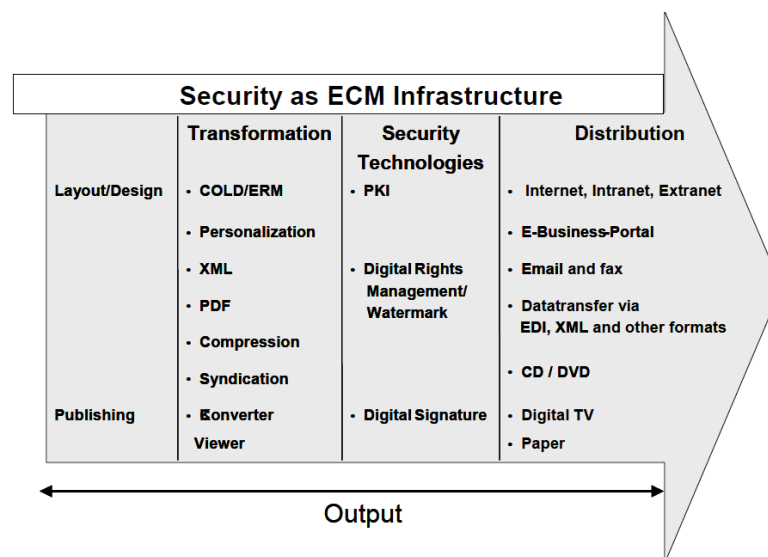


Abbildung 2.16.: ECM-Deliver [KSP07]

Die Deliver-Komponenten umfassen die drei Bereiche Transformationstechnologien, Sicherheitstechnologien und Distribution:

Transformationstechnologien

Die Transformationstechnologien stellen die technischen Komponenten für Transformationen im Hintergrund bereit. Sie werden i. d. R. als Dienste für die anderen ECM-Komponenten bereitgestellt und sollten immer verlustfrei, kontrolliert und nachvollziehbar ablaufen.

COLD/ERM stellt im Zusammenhang mit Deliver die Aufbereitung von Daten für Verteilung, Druck und der Übergabe an ein Archiv bereit. Typische Beispiele hierfür sind individualisierte Kundenanschriften oder der gesammelte Druck mehrerer Briefe, die an einen Kunden gesendet werden soll, um zu vermeiden, dass der Empfänger an einem Tag mehrere Briefe erhält.

Personalisierung dient dazu, die einzelnen Komponenten und Anwendungen auf die Bedürfnisse des Benutzers anzupassen. Dies steigert die Produktivität und Benutzerfreundlichkeit der Anwendungen. Besonders stark wird dies im Bereich der Portale deutlich, in denen ganze Anwendungen individuell unter einer Oberfläche und Anmeldung für den Benutzer zusammengestellt werden können.

Kompression wird verwendet, um den Speicherplatzbedarf von Daten zu reduzieren. Gängige Verfahren hierfür sind die Verwendung von ZIP-Formaten als verlustfreie Kompression für alle Arten an Informationen oder JPEG (verlustbehaftet) für große Bilddateien [Sal06, Mia99].

Syndication beschreibt die Bereitstellung von Inhalten in unterschiedlichen Formaten, Umfang und Aufbereitungen. Durch Syndication kann der gleiche Inhalt an mehreren Stellen für verschiedene Anwendungszwecke genutzt werden. Dadurch wird eine fehlerträchtige redundante Datenspeicherung vermieden.

Konverter und Viewer dienen der Transformationen von Informationen, um einheitliche Formate zu erzeugen und verwenden zu können. Viewer-Komponenten dienen dazu Inhalte in verschiedenen Formaten bereitzustellen, ohne dass eine spezielle Anwendung benötigt wird, die das Dateiformat kennt.

Sicherheitstechnologien

Sicherheitstechnologien werden typischerweise als Querschnittsfunktionen für alle ECM-Komponenten bereitgestellt. Elektronische Signaturen werden etwa bei der Erfassung, Änderungen an Inhalten und beim Versand eingesetzt. Dazu gehören PKI³⁹ als Basis für den Einsatz digitaler Signaturen, die die Schlüssel- und Zertifikatsverwaltung, sowie die Prüfung der Authentizität von Signaturen übernehmen. Um die Inte-

³⁹Public Key Infrastructure

grität der übermittelten Nachricht und die Authentizität des Absender sicherzustellen, werden in Europa drei Arten der elektronischen Signatur (einfache, fortgeschrittene und qualifizierte) eingesetzt. In Deutschland kann lediglich die qualifizierte elektronische Signatur nach SigG⁴⁰ und BGB⁴¹ die gesetzlich vorgeschriebene Schriftform ersetzen.

Distribution

Die vorgestellten Technologien dienen dazu, die unterschiedlichen Inhalte mit ECM kontrolliert und nutzungsorientiert dem entsprechenden Empfänger und Bearbeiter bereitzustellen. Die Bereitstellung kann dabei aktiv, z. B. durch E-Mail, Rundschreiben oder passiv durch Publikation auf einer Webseite geschehen. Hierzu können verschiedene Kommunikationskanäle wie Portale, E-Mail und Fax, mobile Geräte oder Papier zum Einsatz kommen. Die Aufgaben der Deliver-Komponenten bestehen darin, die Informationen für den Empfänger optimiert aufbereitet zur Verfügung zu stellen und die Nutzung möglichst zu kontrollieren.

2.4. Enterprise-Content-Management als integratives Konzept

Mit ECM ist ein Konzept entstanden, das die ganzheitliche Verwaltung von Informationen im Unternehmen zum Ziel hat. Dazu sollen Inhalte, die in unstrukturierter Form als Dokumente oder Contentobjekte vorliegen, gespeichert, verwaltet und anderen Anwendungen zur Verfügung gestellt werden. Diese sind für bis zu 90% aller Daten eines Unternehmens verantwortlich [Jen05]. Darüber hinaus dient es auch der Archivierung von Informationen aus anderen Systemen, wie z. B. ERP oder CRM-Systemen und führt damit auch die bislang von den Anwendungen selbst verwalteten Daten mit dem Content und den Dokumenten zusammen und macht sie darüber hinaus übergreifend nutzbar. Dadurch sollen mit ECM die Einschränkungen der heutzutage vorherrschenden „Inselanwendungen“ überwunden werden, die ihre Daten jeweils separat speichern. Dies umfasst auch die Herausforderungen, unternehmenskritische Inhalte zu identifizieren und zu verwalten, etwa geschäftsrelevante E-Mails in persönlichen Postfächern, und die Integration von CM mit weiteren unternehmensweit eingesetzten Anwendungen wie ERP-, CRM, PDM⁴²-Systeme, sowie die Bereitstellung von Suchwerkzeugen [PM05, AII06]. Dafür können die Funktionen als Dienste bereitgestellt werden, die von den Anwendungen genutzt werden können, ohne die Funktionen mehrfach und redundant implementieren und warten zu müssen. Damit wird mit

⁴⁰Signaturgesetz

⁴¹Bürgerliches Gesetzbuch

⁴²Produktdatenmanagement

ECM auch wirklich ein unternehmensweites Repository entstehen, das strukturierte, schwachstrukturierte und unstrukturierte Information zusammenführt. Dadurch entfällt außerdem die redundante Speicherung der Informationen, die durch die Verwaltung durch mehrere Anwendungen zwangsläufig entsteht. Um dies umzusetzen, werden jedoch einheitliche Integrationsdienste und standardisierte Schnittstellen benötigt. Damit kann Anwendungen der Zugriff auf die Informationen ermöglicht werden, ohne dass individuelle und meist aufwändige, sowie fehleranfällige Integrationen und Anpassungen benötigt werden. Aktuelle Fragen wie: wo lässt man die Abertausenden von E-Mails, wie geht man mit elektronisch signierten Handelsbriefen um, wie werden Informationen aus dem unorganisierten Dateisystem überführt, wie kann kontrolliert werden, dass keine Informationen verloren gehen oder nicht bearbeitet werden, können mit der Umsetzung des ECM-Konzepts als zentralen Bestandteil der Infrastruktur beantwortet werden [Kam03a].

2.5. Zusammenfassung

In diesem Kapitel wurden zunächst die Grundbegriffe Content, Dokument, Record und Dokumentenlebenszyklus vorgestellt, die in der anschließenden Vorstellung des ECM-Konzepts von Bedeutung sind. Das ECM-Konzept wurde anhand der Definition der AIIM als ein 5-Komponenten Modell vorgestellt, das die einzelnen Funktionsbereiche umfasst. Dabei wurde der Bereich der Erfassung und Erschließung von Informationen aus verschiedenen Datenquellen betrachtet. Die Verwaltungskomponenten Dokumenten-Management, Records-Management, Web-Content-Management, Collaboration und Workflow/BPM, übernehmen zusammen die Verwaltung und Steuerung von Content und Dokumenten während der Erstellung, Bearbeitung und Verteilung im Lebenszyklus der Informationen. Weiter wurden Komponenten zur Speicherung und Organisation der Informationen und damit in Zusammenhang stehende Speichertechnologien vorgestellt. Anschließend wurde auf die Aspekte der elektronischen Langzeitarchivierung und der kontrollierten Vernichtung von Informationen am Ende des Lebenszyklus eingegangen, bevor im Bereich Deliver Komponenten zur Transformation und Distribution von Informationen über verschiedene Kanäle erläutert wurde. Abschließend wurde die Positionierung und die Bedeutung von ECM als integratives Konzept in der Infrastruktur moderner IT-Landschaften betrachtet.

3. Evaluierung von Dokumenten- und Content-Management-Systemen

In diesem Kapitel wird ein Überblick über am Markt befindliche DMS und CMS gegeben. Da der Markt für diese Systeme sehr groß ist und hier hunderte von Anbietern ihre Produkte bewerben, gestaltet es sich schwierig, einen ganzheitlichen Überblick über den Funktionsumfang der Systeme und den Stand der Technik in diesem Bereich zu geben. Mit Alfresco und dem Microsoft Office Sharepoint Server 2007 wurden zunächst zwei in diesem Bereich neuere Produkte ausgewählt. Sie haben nicht den herkömmlichen Archivsystemhintergrund, der den meisten DMS-Anbietern zugrunde liegt. Für die weiteren Systeme fiel die Auswahl anhand einer Studie des VOI¹ zu eingesetzten DMS in Deutschland (Abbildung: 3.1) auf EASY ENTERPRISE.x mit der weitesten Verbreitung in Deutschland, DocuWare 5 ENTERPRISE und ELOprofessional 6 [BHW07].

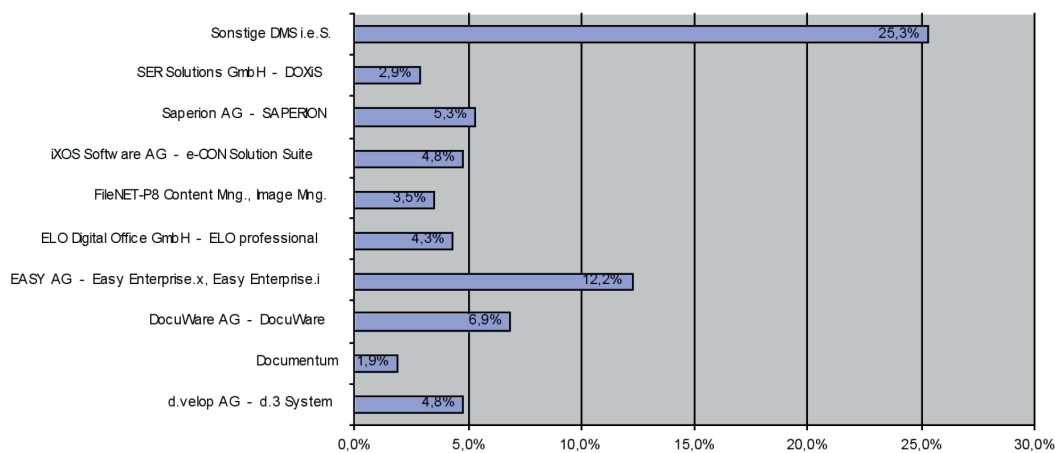


Abbildung 3.1.: VOI-Studie zur Untersuchung des Marktpotentials 2007 [BHW07]

Um den Funktionsumfang bewerten und vergleichen zu können, werden zunächst die Evaluierungskriterien vorgestellt, nach denen die Systeme untersucht werden. Die Grundlage dazu bildet das vorhergehende Kapitel 2 mit dem beschriebenen Funktionsumfang zu ECM. Die Betrachtung der Funktionsbereiche der jeweiligen Systemen

¹Verband Organisations- und Informationssysteme e.V.

orientiert sich am Aufbau von Kapitel 2. Sie beschreiben, wie die Funktionsbereiche von den Systemen umgesetzt werden. Zusätzlich werden Aufbau und Architektur der Systeme zu Beginn kurz präsentiert. Es wird darauf eingegangen, welche Technologien für die Implementierung der Systeme verwendet werden. Außerdem wird dabei beschrieben, welche Schnittstellen die Systeme bereitstellen. Am Ende des Kapitels werden die Systeme in einer Gegenüberstellung in ihrem Funktionsumfang verglichen, und es wird aufgezeigt, wo sie sich in ihrem Funktionsumfang unterscheiden und welche Bereiche jeweils abgedeckt werden.

3.1. Evaluierungskriterien

Dieser Abschnitt erläutert die Evaluierungskriterien nach denen die ausgewählten Systeme untersucht werden. Die Systeme müssen den Benutzer bei der Erstellung bzw. Erfassung, Änderung, Prüfung und Verteilung von Dokumenten und Content unterstützen. Hierzu werden die dazu bereitgestellten Funktionen auf die verschiedenen Funktionsbereichen hin untersucht.

Capture

Im Bereich Capture werden Funktionen zur Erfassung und Verarbeitung von Content und Dokumenten untersucht. Dazu werden die Möglichkeiten der Erfassung von Papierdokumenten betrachtet und wie die Erkennung und Extraktion von Inhalten mit Hilfe von Texterkennungsfunktionen, wie OCR oder die Zuordnung von Dokumenten über Barcodes realisiert wird. Diese Funktionen stellen einen wesentlichen Punkt bei der Erfassung dar, da in Papierform vorliegende Informationen nach wie vor einen großen Teil der Kommunikation und des Informationsaustauschs ausmachen. Außerdem werden Funktionen zur Erfassung von COLD-Daten untersucht. Hierzu sollten Schnittstellen vorhanden sein, um vor allem ältere Anwendungen und den Masseneimport dieser Daten zu ermöglichen.

Im Bereich der Erfassung von CI-Dokumenten werden Integrationsmöglichkeiten mit Anwendungen betrachtet. Dies umfasst vor allem Integrationen in typische Office-Anwendungen wie MS Office, mit denen ein Großteil der Dokumente in Unternehmen erstellt und bearbeitet wird. Weiterhin werden Erfassungsmöglichkeiten über virtuelle Drucker oder durch eine Bereitstellung über eine Netzwerkfreigabe behandelt. Diese Funktionen sind ebenfalls relevant für die Erfassung, da hierüber wichtige Schnittstellen für dokumentenerzeugende Anwendungen bereitgestellt werden müssen, um neben den Papierdokumenten auch die digital erzeugten Dokumente ohne Medienbrüche im System verwalten und nutzen zu können.

Darauf aufbauend sind die Informationen bei der Erfassung mit Metadaten und Attributen zu versehen. Hierfür werden die Indexierungskomponenten für NCI- und CI-Dokumente betrachtet. Die einfachste Variante stellt die manuelle Indexierung dar. Weitergehende Funktionen sind die Erfassung über Vorlagen, mit denen Funktionen zur Formularverarbeitung umgesetzt werden. Als umfangreich kann eine automatisierte Klassifikation von Inhalten angesehen werden, die vor allem in Bereichen mit hohem Erfassungsvolumen einen hohen Grad der Automatisierung ermöglicht. Hierbei stellen die manuelle Erfassung und die Erfassung über Vorlagen grundlegende relevante Funktionen für die Verarbeitung erfasster Informationen dar, während Funktionen zur automatischen Klassifikation nicht zwingend für jeden Einsatzzweck erforderlich sind.

Dokumenten-Management

Im Bereich DM werden die grundlegenden DM-Funktionen Check-In/Check-Out, Versionierung, Renditionverwaltung, Suche und eine strukturierte Ablage von Dokumenten geprüft, die zwingend im Bereich der Dokumentenverwaltung vorhanden sein müssen. Hierzu sind Funktionen zur flexiblen Nutzung und Organisation von Dokumenten, sowie Aggregationsmöglichkeiten, wie die Erstellung von virtuellen Mappen wichtige Punkte. Neben der Dokumentenverwaltung sind die Such- bzw. Recherchefunktionen eine wichtige Anwendung von DM. Hierzu werden die Funktionen auf die Erstellung individueller Suchanfragen hin untersucht. Dies umfasst die Erstellung individueller Suchmasken, um das Auffinden von Informationen gezielt zu unterstützen, sowie die Funktionen der übergreifenden Volltextsuche für Anfragen, bei denen der Ablageort nicht bekannt ist. Den Suchfunktionen kommt dabei ebenfalls eine hohe Bedeutung zu, da sie einen wesentlichen Beitrag zur Nutzung der gespeicherten Informationen beitragen. Weiterhin wird untersucht wie Funktionen aus den angrenzenden Bereichen Capture, Store und die weiteren Manage-Bereiche integriert sind.

Records-Management

Die RM-Komponenten werden bzgl. ihrer Funktionen zur Erstellung von Aktenplänen, der Einrichtung von Schemata für Records und Dokumente, das Anlegen von Aufbewahrungs- und Vernichtungsfristen und die Protokollierung sämtlicher Änderungen an Records analysiert. Zusätzlich werden weitergehende Möglichkeiten zu inhaltsbezogenen Berechtigungskonzepten und der Erstellung eines Thesaurus zur einheitlichen Verwendung von Begriffen vorgestellt.

Web-Content-Management

Im Bereich WCM werden die Möglichkeiten der Erstellung und Verwaltung von Webinhalten betrachtet. Hierzu sind vor allem die Funktionen von Redaktionssystemen relevant, mit denen die Veröffentlichungen der Webinhalte umgesetzt werden. Dabei werden Integrationen von Inhalten aus den anderen untersuchten Bereichen, wie die direkte Integration von Dokumenten und Records bewertet. Außerdem werden Möglichkeiten der Portalbereitstellung bzw. -integration untersucht, um hierüber Integrationen mit anderen Systemen webbasiert zu ermöglichen. Hierbei gilt die Integration in andere Portale als besonders relevant.

Collaboration

Im Bereich Collaboration werden Funktionen zur gemeinsamen Bearbeitung von Inhalten und zur Koordination behandelt. Hierzu sind Funktionen zur Bereitstellung von gemeinsamen Arbeitsbereichen, Knowledge Bases, gemeinsam nutzbarer Terminkalender zur Ressourcenplanung und Kommunikationsmöglichkeiten wichtig. Ebenso sind die Integrationen von Anwendungen zur Bearbeitung relevant.

Workflow/BPM

Im Bereich WF werden die drei verschiedenen Workflow-Arten Adhoc-Workflow, Production-Workflow und BPM untersucht. Dabei wird auf die Modellierungs- und Ausführungsaspekte der Workflow-Komponenten eingegangen.

Adhoc-Workflows dienen der Erstellung kleinerer Abläufe, die direkt durch den Benutzer erstellt werden. Im Bereich Production-Workflow werden Aspekte der Modellierung und Ausführung von Workflows evaluiert, mit denen sich vorgegebene Geschäftsprozesse abbilden lassen. Hierzu werden graphische Modellierungswerkzeuge und die Funktionen zur Ausführung der Workflows untersucht. Bei beiden genannten Workflow-Arten ist außerdem von Bedeutung, dass Inhalte aus den Bereichen DM, WCM und RM integriert werden können.

Im Bereich BPM werden darüber hinaus Integrationsmöglichkeiten von Drittanwendungen in die Prozesse untersucht. Hierzu ist es notwendig, dass neben den Inhalten der erwähnten Komponenten auch Daten aus anderen Quellen über die Workflow-Engine integriert werden können bzw. Daten an andere Systeme übergeben werden, die in die Prozessausführung involviert sind. Weiterhin sind in diesem Bereich Monitoring-Funktionen zur Analyse von Prozessen erforderlich.

Store

Im Bereich Store werden die technischen Komponenten zur Speicherung von Inhalten mit Hilfe von Datenbanken, Dateisystemen und speziellen Speichersystemen, untersucht. Hierbei sind die Integrationsmöglichkeiten von unterschiedlichen Speichersystemen relevant, um Funktionen wie HSM zu ermöglichen, um hierüber die Daten gemäß Anforderungen an Speicherkosten oder Zugriffszeiten optimiert speichern zu können. Besonders wichtig ist hierbei die transparente Bereitstellung von Inhalten für die Manage-Komponenten, unabhängig vom physischen Speicherort oder Speichermedium.

Preserve

Der Bereich Preserve geht auf die technischen Aspekte der Archivierung ein. Hier steht vor allem die Unterstützung geeigneter Langzeitarchivmedien- und Geräte im Vordergrund. Außerdem wird eine Integration in die Funktionen der Store-Komponenten benötigt, um auch elektronisch archivierte Daten transparent über die Manage-Komponenten zur Verfügung zu stellen. Außerdem sind die Möglichkeiten der technischen Umsetzung von RM-Richtlinien relevant, um hierüber die flexible Speicherung, unterschiedliche Aufbewahrungs- und Zugriffsanforderungen durchgängig zu ermöglichen.

Deliver

Im Bereich Deliver werden Funktionen zur Transformation und Schnittstellen zur Weitergabe von Daten an andere Systeme und die Bereitstellung von Sicherheitsfunktionen untersucht. Die Transformationsfunktionen müssen hierbei Dienste für die Manage-Komponenten bereitstellen, um z. B. Dokumente in andere Formate zu konvertieren. Außerdem besitzen diese Funktionen eine hohe Relevanz für die Aufbereitung von Daten, um diese in bestimmten Formaten an andere Systeme weiterzureichen. Diese Funktionen sind auch bedeutend für die Weitergabe an weitere Systeme im Rahmen des Output-Management. Hierzu sollten Schnittstellen angeboten werden, über die weitere Systeme auf Inhalte zugreifen können und Exportfunktionen für Daten in bestimmte Formate z. B. COLD unterstützt werden.

Im Bereich Sicherheitsfunktionen wird untersucht, inwiefern sich Zugriffe auf Inhalte z. B. durch Verschlüsselungstechniken absichern lassen. Außerdem sind Funktionen zur digitalen Signatur von Inhalten wichtig, mit denen die Authentizität sichergestellt werden kann, die u. a. durch rechtliche Anforderungen vorausgesetzt werden. Weiter-

gehende Funktionen können hierbei die Unterstützung von digitalen Mediennutzungsrechten wie DRM² sein.

3.2. Alfresco Labs 3.0

Alfresco wurde im Jahr 2005 als Open-Source-Projekt gestartet. Im Gegensatz zu den meisten Open-Source-CMS legt es den Fokus nicht auf Web-Content-Management, sondern auf Dokumenten-Management. Es wird allgemein als führendes DMS im Open-Source-Bereich betrachtet.

Aufbau und Architektur

Die Alfresco Infrastruktur-Komponenten sind konsequent auf Basis weiterer Open-Source-Software, etwa Apache Lucene [Apa09], Hibernate [Red09] oder Spring [VMw09] aufgebaut [Pot08]. Abbildung 3.2 zeigt eine Übersicht der Architektur. Alfresco selbst läuft in einem Servlet-Container, in dem Web-Applikationen die Funktionen des Systems bereitstellen. Auf der linken Seite sind die verfügbaren Schnittstellen (CIFS³, WebDAV⁴, FTP⁵ und SMTP⁶) zu sehen, über die Zugriffe der Client-Komponenten realisiert werden. Die umfangreichen Schnittstellen und API⁷s, die in eigene Anwendungen integriert werden können, sind auf der rechten Seite zu sehen. Auch besitzt Alfresco als erster Anbieter hier eine teilweise implementierte CMIS⁸-Schnittstelle. CMIS soll dazu dienen eine standardisierte API bereitzustellen, die unabhängig vom konkreten Aufbau des Repositories ist. Alfresco besitzt keinen eigenen Client, sondern realisiert die Zugriffe vollständig über einen Web-Client.

Die Benutzerverwaltung erfolgt über die integrierten Funktionen oder über eine Anbindung an LDAP⁹-Verzeichnisse, über die auch SSO¹⁰ realisiert werden kann.

²Digital Rights Management

³Common Internet File System

⁴Web-based Distributed Authoring and Versioning

⁵File Transfer Protocol

⁶Simple Mail Transfer Protocol

⁷Application Programming Interface

⁸Content Management Interoperability Services

⁹Lightweight Directory Access Protocol

¹⁰Single Sign On

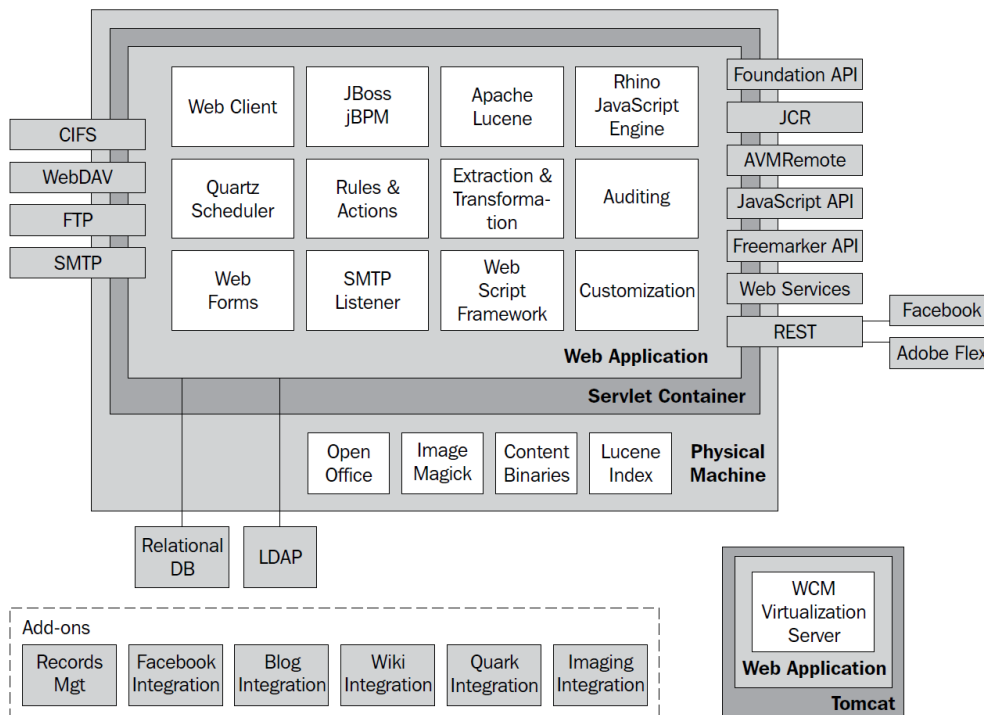


Abbildung 3.2.: Alfresco-Architektur [Pot08])

Capture

Alfresco bietet selbst keine Komponenten für die Einbindung von Scannern oder OCR-Funktionen, um NCI-Dokumente zu digitalisieren. Eine lose Integration von Scannern kann über die integrierte CIFS-Schnittstelle realisiert werden, indem die vom Scanner erzeugten Bilddateien hierüber im Alfresco-Repository abgelegt werden. Ebenso fehlen Komponenten zur Inhaltserkennung von NCI-Dokumenten. Um diese Lücke zu schließen, existieren mehrere Integrationen von Drittanbietern wie Kofax [Kof09], eCopy [eCo09] oder Captiva [EMC09], die Alfresco um einen Scan-Client und OCR-Funktionen erweitern. Abbildung 3.3 zeigt eine Übersicht der Erfassungsarchitektur.

Der Bereich COLD/ERM wird ebenso nicht direkt von Alfresco bereitgestellt, sondern muss über Drittanbieter realisiert werden. Als Beispiel kann hier die SMARTi COLD/ERM-Erweiterung genannt werden. Sie nimmt Ausgabedaten durch den Druckertreiber entgegen und legt die Daten gemäß den Formatierungsanweisungen im Alfresco-Repository mit den zugehörigen Metadaten aus den Datenströmen in der Indexeddatenbank ab [fil09].

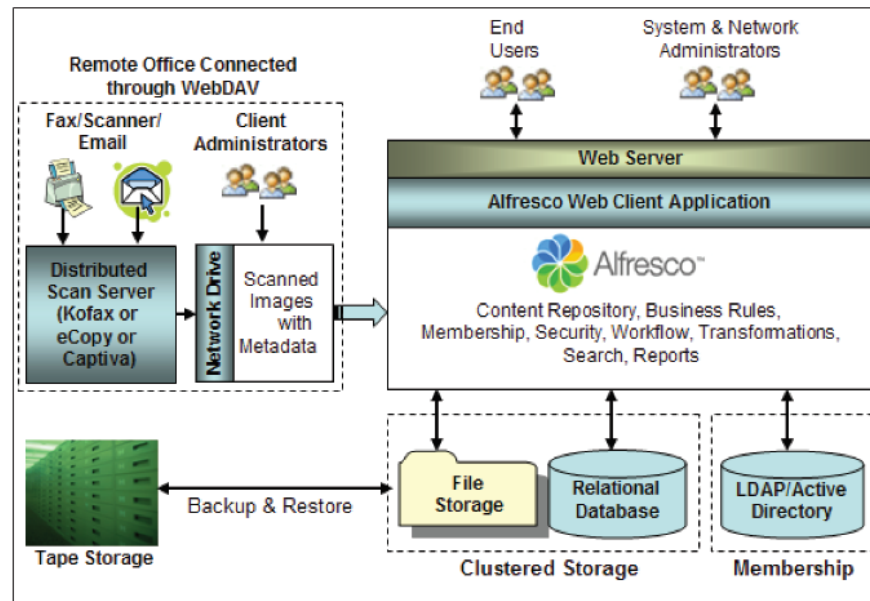


Abbildung 3.3.: Architektur der Erfassungskomponenten [Sha06]

Die Erfassung von CI-Dokumenten wird durch Integrationen in MS Office und Open Office in Form von Plugins realisiert. Die Plugins erweitern die beiden Anwendungen um Funktionen zum direkten Speichern und Öffnen von Dokumenten über die WebDAV-Schnittstelle. Attribute und Metadaten können den Dokumenten beim Speichervorgang zugeordnet werden. Dabei werden die Standardmetadaten von Office-Dokumenten (Autor, Dateigröße, Erstellungsdatum, usw.) automatisch aus den Dateien übernommen. Weiteren Anwendungen kann der Zugriff auf Inhalte aus dem Alfresco-Repository über eine transparente CIFS-Schnittstelle ermöglicht werden.

Die Komponenten für die inhaltliche Erschließung beschränken sich auf manuelle Indexierung, sowie der Erstellung eigener Erkennungsprofile, um bei der Erfassung die definierten Attribute mit Inhalten zu versorgen. Eigene Formulare können durch Javascript ergänzt werden, um Attribute hinzuzufügen oder abhängig vom Speicherort, durch Regeln und Skripte automatisiert, Attributwerte zu füllen [Pot08].

Dokumenten-Management

Die DMS-Funktionen von Alfresco werden über den Web-Client, FTP, WebDAV, CIFS-Netzwerkfreigaben, oder über Plugins in MS Office und Open Office bereitgestellt.

Die Dokumentenablage erfolgt über Verzeichnisstrukturen, in Alfresco *Spaces* genannt, die hierarchisch angeordnet werden können und die Speicherung von Dokumenten, At-

tribute, Versionierung und Zugriffsberechtigungen verwalten. Eigenschaften der Verzeichnisse können an Unterverzeichnisse vererbt werden. Neben der hierarchischen Anordnung können Verzeichnisinhalte über Views aggregiert werden. Damit lassen sich z. B. alle einem Projekt zugeordneten Dokumente zusammenfassen. Als Besonderheit erweitert Alfresco, die über die CIFS-Schnittstelle bereitgestellten *Spaces* um Verknüpfungen, mit denen Check-In-/Check-Out-Funktionen, durch Drag'n'Drop der Dokumente auf die Verknüpfungen angeboten werden.

Zu jedem in Alfresco gespeicherten Informationsobjekt können Metadaten definiert werden, die in der Datenbank abgelegt werden. Es wird dabei zwischen systeminternen Metadaten, wie Berechtigungen oder interne ID, und frei definierbaren Metadaten unterschieden, die die Inhalte beschreiben und für Suchanfragen verwendet werden. Für gängige Dateiformate wie PDF oder DOC können Standardmetadaten wie Autor oder Erstellungsdatum automatisch aus den Dateien extrahiert werden.

Die Inhalte in Alfresco können neben der Strukturierung durch Verzeichnisse und Views in Kategorien beliebiger Schachtelungstiefe eingeteilt werden. Kategorien können nur von Administratoren angelegt werden und werden den Benutzern dadurch als Ablagestruktur vorgegeben. Durch diese Zuordnung von Inhalten zu Kategorien ergibt sich eine weitere Möglichkeit, um Suchanfragen über Kategorien durchzuführen.

Records-Management

Seit August 2009 besitzt das Alfresco RM-Addon eine DoD5015.2-Zertifizierung. Da die DoD5015.2-Anforderungen jedoch relativ oberflächlich gehalten sind, lassen sich mit den integrierten Funktionen nur grundlegende Anforderungen erfüllen.

Für die Erstellung einer strukturierten Ablage können Aktenpläne mit Hilfe von *File-Plans* angelegt und verwaltet werden. Sie ermöglichen die Klassifikation durch Ablage der Informationen in hierarchischen Verzeichnissen, die Bezeichnung von Dateinamen, Records-Kategorien und Metadaten durch den Einsatz von Javascript. Für die Erstellung von Aufbewahrungs- und Löschrufen werden sog. *Business Rules* genutzt, mit denen sich Regeln zur Klassifikation, Versionierung, Benachrichtigungen oder Transformationen von Dokumentenformaten realisieren lassen. Weitere Funktionen können durch Javascript ergänzt werden. Bei Ablauf von Aufbewahrungsfristen können die Inhalte automatisch aus Ansichten gelöscht werden. Records sind dadurch für Benutzer nicht mehr zugänglich. Über spezielle Rollen ist die Vernichtung der Records zu prüfen und durchzuführen.

Wenn in einem File-Plan organisierte Dokumente als Record klassifiziert werden, müssen alle im Record-Schema definierten Record-Eigenschaften angegeben werden. Dies wird über automatisch generierte Formulare realisiert, die auf der jeweiligen Records-Definition basieren. Über Audit-Trails können Inhaltzugriffe und Änderungen um-

fassend protokolliert werden. Hierbei können alle relevanten Informationen zur Aufzeichnung konfiguriert werden und in Protokolldateien gesichert werden.

Web-Content-Management

Der WCM-Bereich von Alfresco kam 2007 als Erweiterung des Kernprodukts hinzu und dient explizit der Erstellung von Webseiten. Die Funktionen *Business Rules*, *Audit-Trails*, dateibasierte Berechtigungen, Volltextindizierung und WebDAV stehen für den WCM-Bereich nicht zur Verfügung. Die Inhalte werden auch nicht in *Spaces* organisiert.

Webinhalte werden mit Web Forms erstellt und in XML-Dateien gespeichert und können über Transformationsfunktionen in andere Formate wie PDF, RSS¹¹ oder HTML transformiert werden [Pot08]. Jeder WCM-Benutzer erhält eine Testumgebung in Form einer Sandbox, um zu verhindern, dass unerwünschte Änderungen sich direkt auf die aktuell veröffentlichte Webseite auswirken. Hierzu wird eine virtuelle Kopie der kompletten Webseite für jeden WCM-Benutzer erstellt. Die Sandbox enthält damit immer die komplette Webseite zzgl. der Änderungen des jeweiligen Benutzers. Um Änderungen oder neue Inhalte überprüfen zu können werden Vorschaufunktionen angeboten, die auch als Beurteilungsgrundlage für Freigabeprozesse zur Einbringung von Änderungen oder neuen Inhalten in die veröffentlichte Webseite dienen. Vor jeder Änderung an der echten Webseite werden alle WCM-Daten durch einen Snapshot gesichert, auf den jederzeit mittels Rollback zurück gesprungen werden kann. Alle Webprojekte, Sandboxes und *WCM-Workflows* besitzen jeweils ihren eigenen Speicherbereich [Pot08].

Alfresco lässt sich in Webportale integrieren oder kann selbst ein Portal bereitstellen. Hierbei werden die J2EE¹²-basierten Portale JBoss Portal [JBo09] oder Liferay Portal [Lif09] unterstützt. Abbildung 3.4 zeigt schematisch, wie Alfresco durch die Verwendung von Portlets entweder über dessen SOAP- oder REST¹³-Schnittstelle in ein Liferay-Portal eingebunden werden kann. Dort wird außerdem die Integration mit dem SSO-Dienst Yale CAS dargestellt [Jas09]. Die Authentifizierung erfolgt hierbei über einen LDAP-Verzeichnisdienst, über den die Benutzereingaben validiert werden. Die Authentifizierung erfolgt dabei transparent über Kerberos-Tickets im Hintergrund und ermöglicht somit eine SSO-Anmeldung.

Collaboration

Die Collaboration-Funktionen sind unter der Bezeichnung *Alfresco Share* zusammengefasst und werden über einen Apache Tomcat-Server bereitgestellt [Pot08, Sha06].

¹¹Really Simple Syndication oder auch Rich Site Summary

¹²Java Platform Enterprise Edition

¹³Representational State Transfer

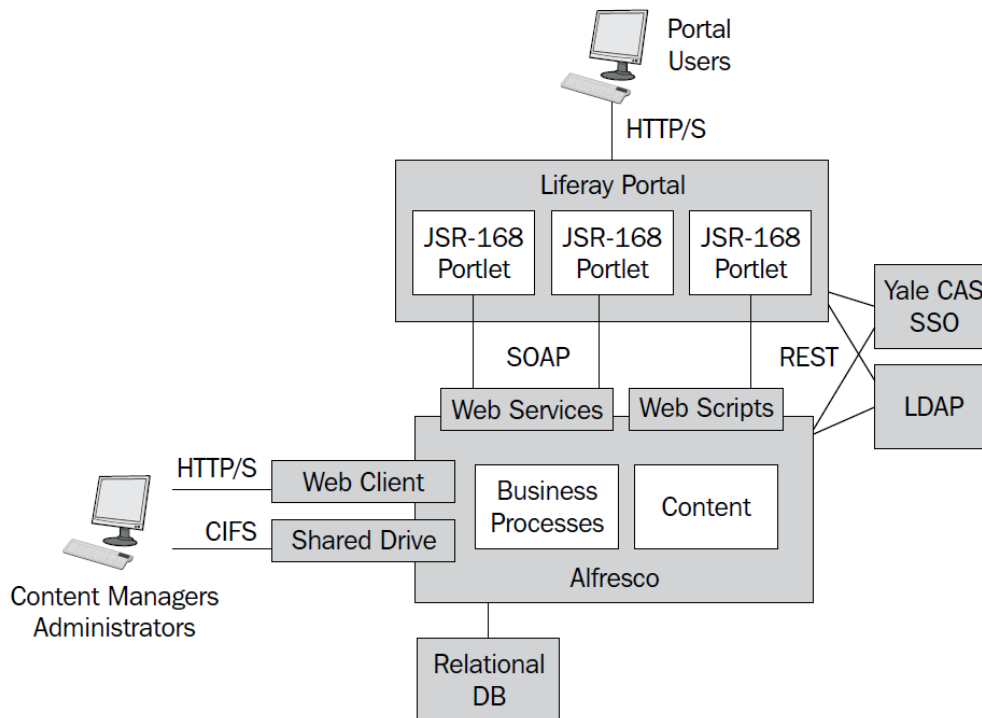


Abbildung 3.4.: Portalintegration von Alfresco und Liferay [Pot08]

Die Kernelemente für diesen Bereich bilden bei Alfresco spezielle *Sites*, die gemeinsam nutzbare Arbeitsbereiche zur Unterstützung der Gruppenarbeit oder für Projektgruppen bieten. Den Mitgliedern dieser *Sites* stehen neben dem Zugriff auf DM-Funktionen, gemeinsame Kalender, Blogs, Wikis und Diskussionsforen zur Verfügung. Allerdings können die Inhalte dieser *Sites* nicht archiviert oder als abgeschlossen und damit unveränderbar markiert werden, was z. B. den Abschluss eines Projekts kennzeichnen könnte. Dadurch können die *Sites* durch Einstellung vieler Inhalte unübersichtlich werden, und die Dokumentation von abgeschlossenen Arbeiten ist nur eingeschränkt möglich.

Für die Kommunikation der *Sites-Mitglieder* stehen keine Instant-Messaging oder E-Mail-Funktionen zur Verfügung, so dass die direkte Kommunikation nicht unterstützt wird und ein Rückgriff auf andere Systeme erforderlich wird. Weitere Einschränkungen existieren auch bei den gemeinsamen Kalendern, da es beispielsweise nicht möglich ist, bestimmte Personen zur Terminabstimmung auszuwählen, sondern immer alle Mitglieder einer *Site* eine Anfrage für eine Terminabstimmung erhalten.

Zur Koordination stehen einfache Workflow-Funktionen, wie ein automatisches Dokumentenrouting innerhalb der *Site-Mitglieder* zur Verfügung. Diese können um eigene Workflows erweitert werden. Planung und Belegung von Ressourcen, wie Besprechungsräume oder Prüfungen auf Anwesenheit bestimmter Personen, sowie die

gemeinsame Nutzung von Anwendungen (Desktop- oder Application-Sharing) oder Whiteboards sind nicht vorhanden.

Workflow

Alfresco unterscheidet im WF-Bereich verschiedene Workflows, die sich in *Simple-* und *Advanced-Workflows* unterteilen. Mit *Simple-Workflows* lassen sich einfache, dokumentenorientierte Workflows für Dokumentenrouting- und Freigabeprozesse umsetzen. Für die Erstellung eines *Simple-Workflow* werden die Benutzer für die Ausführung der einzelnen Aktivitäten mit Hilfe eines Assistenten ausgewählt, und die Aktivitäten werden mit *Content-Items* und *Alfresco Spaces* verknüpft. Dies dient dazu, bei der Ausführung einzelner Aktivitäten die Content-Items von einem *Alfresco Space* an einen anderen zu verschieben und damit anderen Benutzern zur weiteren Bearbeitung zugänglich zu machen. Die Berechtigungen zur Ausführung werden über die Zugriffsrechte auf die Verzeichnisse bestimmt.

Advanced-Workflows dienen der Bereitstellung aufgabenbezogener und komplexerer Workflows, die bei Alfresco über eine Integration der JBPM¹⁴-Workflow-Engine realisiert werden. Mit *Advanced-Workflows* können neben den sequenziellen Prozessen der *Simple-Workflows* auch parallele Bearbeitungen und Verzweigungen definiert werden. Außerdem wird der Datenfluss nicht über die *Spaces* realisiert, an die die Contentobjekte verschoben werden, sondern über die Workflow-Engine gesteuert, und es können automatisierte Aktionen hinterlegt werden. Zusätzlich können Benachrichtigungen per E-Mail eingerichtet werden, die den Benutzer über zur Ausführung vorliegende Prozessschritte informieren.

Mit *Adhoc-Workflows* können ein oder mehrere Dokumente an einen oder mehrere Benutzer zur Bearbeitung gesendet werden. Die *Adhoc-Workflows* werden bei den Benutzern in die Aufgabenliste gestellt, in der auch die anderen auszuführenden Prozessschritte angezeigt werden. Zur Bearbeitung der *Adhoc Tasks* können die Dokumente bearbeitet werden, falls der Empfänger Berechtigungen besitzt, um das Dokument zu ändern. Zum Beenden des *Adhoc-Workflows* kann der Empfänger den Status auf *Completed* setzen, um die Dokumente wieder an den Absender zurück zu leiten. Eine Weiterleitung von zugewiesenen *Adhoc-Workflows* an einen weiteren Benutzer ist außerdem möglich.

In Alfresco stehen zunächst lediglich die Einrichtung von *Adhoc-Workflows* und Freigabeprozessen über Assistenten zur Verfügung, die auf die JBPM-Workflow-Engine zugreifen. Hiermit lassen sich Freigabeprozesse definieren, indem die einzelnen Prozessschritte nacheinander angegeben werden. Den Benutzern wird eine Arbeitsliste bereitgestellt, in der alle zur Ausführung anstehenden Prozessschritte aufgelistet werden.

¹⁴JBoss Business Process Management

Weitergehende Prozesse können über eine Prozessdefinition in Form einer jPDL XML-Datei manuell oder mit Hilfe des graphischen Editors erstellt werden, der in Abbildung 3.5 dargestellt ist.

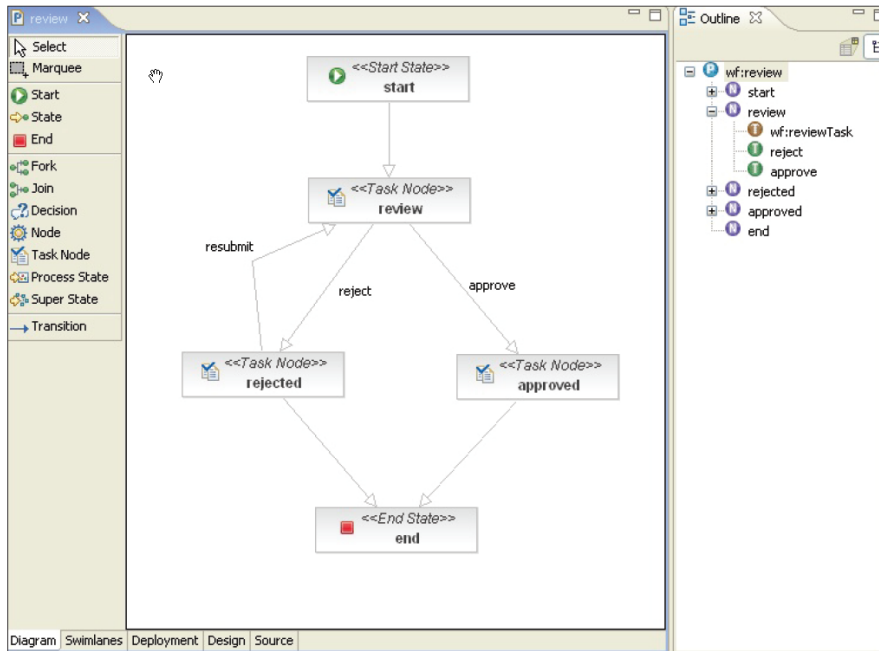


Abbildung 3.5.: JBPM Workflow-Editor [Sha06]

Store

Die Store-Komponenten in Alfresco bilden die technische Basis für die Speicherung. Die Inhalte werden bei Alfresco in einer relationalen Datenbank und dem Dateisystem abgelegt. Die Metadaten werden in der relationalen Datenbank, die Binärdateien im Dateisystem abgelegt. Für die Volltextindizierung kommt eine Apache Lucene-Datenbank zum Einsatz. Das Content-Repository stellt die Dienste für den Zugriff über die in Abbildung 3.2 beschriebenen Schnittstellen zur Verfügung. Für den Zugriff auf das Content-Repository mit anderen Programmen stehen Web Services und eine JCR¹⁵-Schnittstelle zur Verfügung. Außerdem bietet Alfresco eine Implementierung der relativ neuen CMIS-Schnittstelle. Alfresco unterstützt für die Ablage der Daten lediglich Onlinespeicher, die vom Betriebssystem bereitgestellt werden. Optische Medien oder Bandlaufwerke werden nicht unterstützt.

¹⁵Content Repository API for Java

Preserve

Alfresco bietet selbst keine direkte Unterstützung für den Bereich der Archivierung. Über die RM-Funktionen können jedoch Regeln zur Auslagerung der Informationen an spezielle Archivsysteme wie z. B. IBM DR-550 oder EMC Centera [Sha06] definiert werden. Eine Integration der E-Mail-Archivierung aus gängigen Groupwaresystemen ist ebenfalls nur über Drittanbieter möglich.

Deliver

Im Bereich Deliver stellt Alfresco Funktionen zu Transformationen bereit, die über *Business Rules* verwaltet werden. Die Transformationskomponenten unterstützen eine Reihe gängiger Dateiformate, wie PPT, HTML, PDF, JPEG, PNG, TXT, Flash, ODF, usw. sowohl für die Anzeige über den integrierten Viewer, als auch zur Transformation in andere Dateiformate. Zusätzlich bietet Alfresco auch die Möglichkeit Bilddateien automatisch anzupassen, etwa durch automatische Reduktion der Bildgröße, um sie für die Verwendung auf der Webseite optimiert bereitzustellen.

Für die Personalisierung bietet Alfresco von kleineren Änderungen, wie die Anpassung von Symbolen in den Bedienungselementen, eine individuelle Anpassung für jeden Benutzer im Bereich des Web-Clients bis hin zur komplett eigenen Entwicklung von individuellen Oberflächen. Jeder Benutzer kann den Aufbau des eigenen Web-Clients anhand von Vorlagen selbst gestalten, indem die einzelnen Komponenten innerhalb einer Vorlage ein- oder ausgeblendet bzw. an unterschiedliche Positionen verschoben werden können.

Im Bereich der Sicherheitstechnologien können die Zugriffe auf Alfresco durch den Einsatz von Standardwebservern über Zertifikate und SSL-Verbindungen abgesichert werden, um einen verschlüsselten Zugriff auf die Informationen zu erhalten. Digitale Signaturen und DRM werden von Alfresco direkt nicht unterstützt. Hierfür sind eigene Erweiterungen oder Module von Drittanbietern notwendig.

3.3. Microsoft Office SharePoint Server 2007

MOSS¹⁶ 2007 stellt die dritte Version des Sharepoint Servers von Microsoft dar. Die Ursprünge liegen in mehreren Microsoft-Produkten und Zukäufen anderer Software-Hersteller, die im Laufe der Zeit nach und nach als einzelne Produktlinien aufgegeben und in MOSS integriert wurden.

¹⁶Microsoft Office Sharepoint Server

Aufbau und Architektur

Die dienstorientierte MOSS 2007-Architektur folgt dem SOA-Konzept [Eng07]. Abbildung 3.6 zeigt einen Überblick über die MOSS 2007 zugrunde liegende Architektur. MOSS 2007 setzt als Grundlage ausschließlich auf das Portfolio von Microsoft, wie das ASP.NET 2.0-Framework, .NET-Framework und die Windows Workflow Foundation. Die unterstützten Betriebssysteme beschränken sich auf Windows Server 2003 und 2008 mit dem integrierten IIS-Webserver und den MS SQL-Server. Darüber sind grundlegende ECM-Komponenten abgebildet, die MOSS 2007 bereitstellt. Diese Funktionen entsprechen dem Umfang der von Microsoft als Windows Sharepoint Services angebotenen kostenlosen Variante. In MOSS 2007 sind gegenüber den Windows Sharepoint Services zusätzliche Komponenten mit erweitertem Funktionsumfang in den Bereichen Portal, Suche, Workflow, CM und Collaboration vorhanden. Die Zugriffe auf die Dienste und Inhalte werden vollständig über einen Web-Client realisiert.

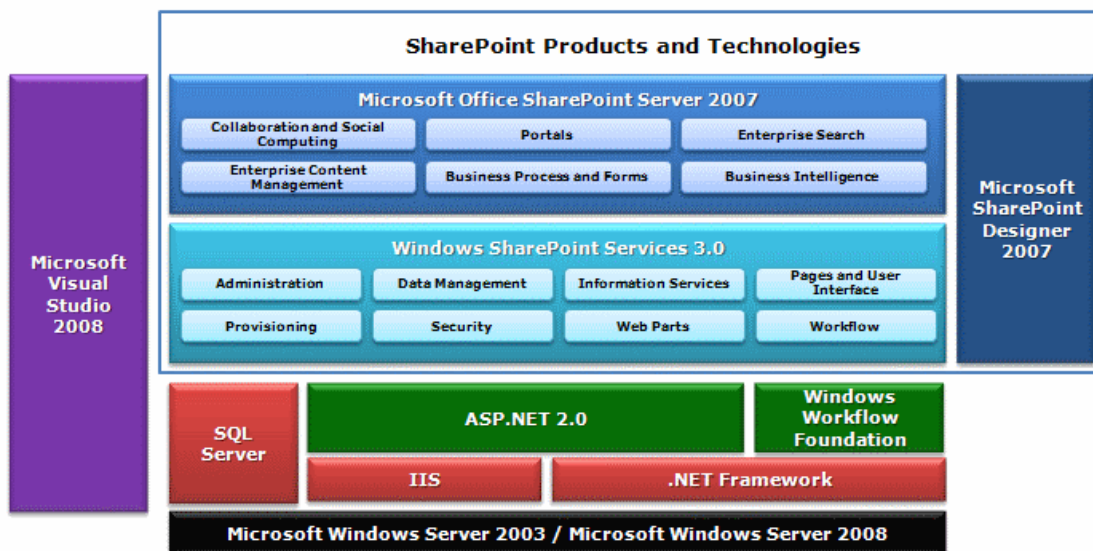


Abbildung 3.6.: MOSS 2007-Architektur [Mic09c]

MOSS 2007 besitzt keine eigenständige Benutzerdatenbank, in der Benutzer gespeichert werden. Die Benutzerverwaltung wird daher über Verzeichnisdienste integriert. Im Gegensatz zu früheren Sharepoint-Versionen kann MOSS 2007 neben der Integration in MS Active Directory in weitere LDAP-Verzeichnisdienste integriert werden. Außerdem existieren weitere Anbindungsmöglichkeiten an andere Benutzerverzeichnisse, wie die Windows-Benutzerdatenbank des Betriebssystems.

Capture

Der Capture-Bereich gestaltet sich ähnlich wie bei Alfresco und besitzt ebenfalls keine eigenen Funktionen zur Erfassung von Papierdokumenten oder COLD-Daten. Drittanbieter wie Kofax [Kof09] oder Captaris [EMC09] bieten jedoch auch hier Erweiterungen für diese Erfassungsbereiche an.

Die Erfassung von CI-Dokumenten wird durch die in MS Office bereits vorhandene Integration mit MOSS realisiert. Der Zugriff erfolgt über die WebDAV-Schnittstelle. Office-Dokumente können damit direkt aus einer Dokumentenbibliothek geöffnet und dort gespeichert werden. Attribute und Metadaten werden bei der Speicherung im Repository angegeben.

Die Komponenten für die inhaltliche Erschließung beschränken sich auf Funktionen zur manuellen Indexierung. Die zu einem Dokumententyp definierten Attribute und Metadaten müssen beim Einstellen von Dokumenten manuell ausgefüllt werden. Hierbei werden automatisch Formulare generiert, die auf den Attributen und Metadaten einer Dokumentenbibliothek basieren.

Über die *ENTERPRISE-Suche* können neben den in MOSS gespeicherten Inhalten weitere Datenquellen eingebunden werden. Dazu werden sog. *Crawler* eingesetzt, die externe Datenquellen durchsuchen und indizieren. Standardmäßig sind dies Webseiten, MS Exchange-Datenbanken, Netzwerkfreigaben oder SQL¹⁷-Datenbanken. Dabei werden auch verschiedene Sprachen der Inhalte erkannt, die in den Suchergebnissen berücksichtigt werden. So können je nach Sprache des angemeldeten Benutzers, Inhalte bevorzugt werden, die in der gleichen Sprache verfasst sind.

Dokumenten-Management

MOSS verwendet zur Speicherung von Dokumenten sog. Dokumentenbibliotheken, mit denen sich Verzeichnisstrukturen abbilden lassen, die über den webbasierten Portalzugriff von MOSS oder über die WebDAV-Schnittstelle des Windows Explorers zugreifbar sind. Dokumentenbibliotheken bieten verschiedene Ansichten für die gespeicherten Inhalte wie Listen oder eine vom Windows Explorer her bekannte Exploreransicht mit Verzeichnissen und Dateien. Mit Dokumentenvorlagen werden Dokumententypen mit den zugehörigen Metadaten definiert, die Dokumentenbibliotheken zugeordnet werden können. Veränderungen an Dokumenten können über eine integrierte Versionierung protokolliert werden. Die Versionierung kann individuell für jede Dokumentenbibliothek aktiviert oder deaktiviert werden. Benutzer können sich über sog. Abonnements über Änderungen an Dokumenten, Verzeichnissen oder gespeicherten

¹⁷Structured Query Language

Suchanfragen informieren lassen. Die Informationen können auch per E-Mail versendet werden.

In Verbindung mit MS Word 2007 können Dokumente als endgültig gekennzeichnet werden. Diese Funktion dient dazu, alle Bearbeitungsfunktionen für das Dokument zu entfernen, um eine weitere Bearbeitung zu verhindern. MOSS unterstützt die automatische Konvertierung von Dokumenten in andere Dateiformate, so können z. B. Word-Dokumente in HTML-Dokumente umgewandelt werden, um die Inhalte direkt auf einer Sharepoint-Webseite zu veröffentlichen.

Records-Management

Die RM-Funktionen ermöglichen Metadaten, Workflows und Policies für Dokumententypen zu definieren. Für die Abbildung von Aktenplänen werden ebenfalls Dokumentenbibliotheken verwendet, die in eine hierarchische Beziehung gesetzt werden können. Die Definition von Records erfolgt über Inhaltstypen, die Dokumente, E-Mails, Mediendaten und weitere Metadatenfelder umfassen.

Die Regeln zum Umgang mit Records werden über Informationsverwaltungsrichtlinien definiert, die mit Inhaltstypen verknüpft werden. Dazu erlaubt MOSS die Festlegung von Regeln im Ablagesystem, die Regeln für einzelne Dokumentenbibliotheken oder über ganze *Websitesammlungen* erlauben. Hierüber werden auch die Aufbewahrungsfristen der Records definiert. Nach Ablauf von Aufbewahrungsfristen können die Records dann in spezielle Archivwebseiten verschoben werden [MHS08]. Audit-Trails werden durch die automatischen Protokollierungen aller Änderungen umgesetzt. Zur Aufbereitung können die Protokolle dazu in MS Excel Berichte konvertiert werden. Durch ein Addon kann eine Erweiterung der RM-Funktionen eingeführt werden, mit denen eine zertifizierte DoD5015.2-Handhabung von Records ermöglicht wird.

Das RM-Konzept von MOSS 2007 deckt jedoch nur sehr grundlegende Funktionen für diesen Bereich ab. So existiert z. B. keine Unterscheidung zwischen lebenden und archivierten Records. Die Transformationen von Dokumenten sind auf MS Office-Formate beschränkt, die Reportfunktionen mit MS Excel sind i. d. R. nicht ausreichend gesichert, um weitergehende Anforderungen an eine sichere Protokollierung und entsprechende Nachweisfunktionen zu ermöglichen, außerdem fehlen Langzeitarchivierungsfunktionen [GSA08].

Web-Content-Management

Mit der Veröffentlichung von MOSS 2007 hat Microsoft das bis dahin eigenständige WCM-Produkt CMS 2002 eingestellt und die Funktionen in MOSS 2007 integriert und erweitert.

Die WCM-Funktionen werden über *Websitesammlungen* zur Verfügung gestellt. Hierzu werden verschiedene *Websites* erstellt, die hierarchisch in Beziehung gesetzt werden. Innerhalb einer *Websitesammlung* können verschiedene Vorlagen genutzt werden. Diese umfassen *Startwebsiteshierarchien*, *Zusammenarbeitsportale* und *Veröffentlichungsportale*. Letztere dienen der Veröffentlichung von Inhalten im Internet über *Veröffentlichungssites*. Diese benötigen eine *Mastersite*, die das Layout und die Bedienfunktionen für die Seite einheitlich vorgeben. Master- und Layoutseiten können über den Sharepoint-Designer oder Visual Studio erstellt oder verändert werden. Über *Site-Variations* können verschiedene Versionen von Webseiten verwaltet werden, die z. B. mehrere Sprachen oder unterschiedliche Aufbereitungen für mobile Geräte ermöglichen und damit die Konsistenz sicherstellen.

Zur Erstellung von Inhalten auf einer *Veröffentlichungssite*, können Dokumente mit Hilfe der MS Office-Programme erstellt werden und durch Konvertierungsfunktionen in HTML umgewandelt oder integrierte HTML-Editoren verwendet werden. Dokumente aus dem DM-Bereich können direkt in die Webseiten integriert werden. Hierzu werden auch Transformationsfunktionen bereitgestellt, die z. B. Inhalte eines Word-Dokuments in HTML zur direkten Anzeige über einen Browser transformieren.

Die Bereitstellung von Portalen gehört zum Kernbereich von MOSS. Die Funktionen des Portals werden über die eingangs erwähnten Zusammenarbeitsportale bereitgestellt. Ein Portal ist zunächst eine normale Webseite, die einen Web-Editor enthält. Weiterhin gehören Bibliotheken zum Speichern von Seiten, Vorlagen und wiederverwendbaren Elementen zu den Funktionen. Durch Vorlagen können automatisch generierte Hierarchien von Webseiten erstellt werden, die verschiedene Inhalte wie Nachrichten-, Dokumenten- oder Suchbereiche beinhalten. Für die Gestaltung von Portalinhalten lassen sich Benutzerprofile definieren, mit denen individuelle Seiten Benutzern bzw. Gruppen zugeteilt werden können. Weitere Anpassungen, wie persönliche Informationen und Dokumente, können darüber ebenfalls genutzt werden. Die Vorlagen werden zentral über die Administrationsseite des MOSS verwaltet und zugewiesen. In das Portal können die Funktionen Kalender, Aufgaben und E-Mail aus der Groupware MS Exchange direkt durch die Einbindung von *Webparts* integriert werden. *Webparts* dienen in MOSS der individuellen Anpassung und Erweiterung eigener Module für die Integration in ein Portal [MHS08]. Sie bilden das Pendant zu Portlets im Java-Bereich.

Collaboration

Die Collaboration-Funktionen in MOSS sind umfangreich. Durch den Einsatz der beschriebenen Portalfunktionen können in MOSS vielfältige Anforderungen an die Zusammenarbeit von Gruppen umgesetzt werden. Listen, Dokumentenbibliotheken, Kalender, Aufgabenlisten, Kontaktadressen, Wikis, usw. können sowohl für die eigene Nutzung, als auch für andere Mitglieder einer Gruppe freigegeben oder durch spezielle Teamwebsites ganzen Gruppen zur Verfügung gestellt werden. Durch die webbasierte

Nutzung von MOSS können diese auch relativ einfach über das Internet bereitgestellt werden, um z. B. externen Partnern oder mobilen Nutzern Zugriffe zu ermöglichen.

Durch die Integration mit MS Exchange können sämtliche E-Mail-, Kontakt- und Kalenderfunktionen der Groupware direkt in MOSS genutzt werden. Durch die enge MS Office-Integration können auch Instant-Messaging-Funktionen z. B. über den Groove-Client aus der Office-Suite genutzt werden [Mic07]. Zusätzlich zu Instant-Messaging kann Groove auch gemeinsam nutzbare Arbeitsbereiche bereitstellen. Gemeinsam bearbeitete Dokumente können per Drag'n'Drop in einen Arbeitsbereich gezogen werden, bei dem das Dokument dabei allen Teilnehmern der Gruppe zur Verfügung gestellt wird. Änderungen können direkt mit Dokumentenbibliotheken synchronisiert werden. Groove bietet außerdem die Möglichkeiten eines Desktop- oder Application-Sharing. Damit lassen sich auf einem Computer eines Gruppenmitglieds ausgeführte Programme, anderen Benutzern zugänglich machen, um gemeinsam daran zu arbeiten. Weitere Anwendungen können über die *Webparts* in MOSS ebenfalls auf Teamwebsites eingebunden werden, um weitere Daten aus anderen Quellsystemen einbinden und gemeinsam bearbeiten zu können.

Workflow

MOSS setzt mit seinen Workflow-Funktionen auf die Windows Workflow Foundation Services, die ein Framework und Werkzeuge für die Entwicklung innerhalb des .NET-Frameworks bereitstellen. Dazu werden die Workflows Foundation Services durch spezifische MOSS-Funktionen erweitert. Abbildung 3.7 zeigt die zugrunde liegende Architektur der Integration der MOSS-Funktionen mit der Workflow-Engine des .NET-Frameworks.

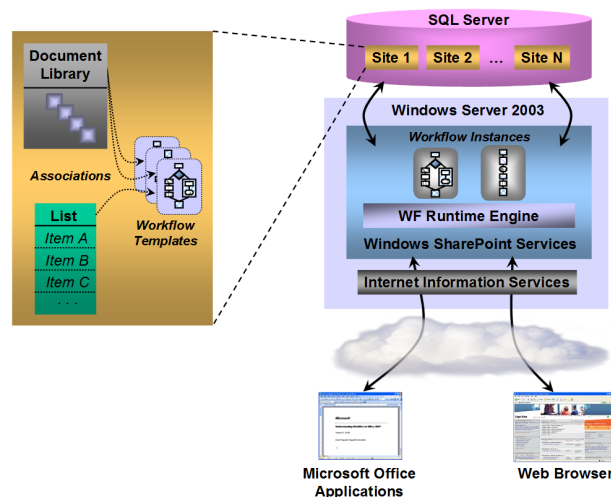


Abbildung 3.7.: Workflow Foundation und MOSS [Cha06]

Für die Erstellung von Workflows in MOSS werden zum einen der Sharepoint-Designer, und zum anderen Visual Studio 2005 Designer zur Verfügung gestellt. Der Sharepoint-Designer gibt Benutzern die Möglichkeit, eigene Workflows zu erstellen. Dies können z. B. einfache Freigabeprozesse für Dokumente sein, die innerhalb von MOSS ablaufen.

Der Visual Studio Designer hingegen zielt auf Entwickler ab, die mit einem graphischen Editor (Abbildung 3.8) zunächst Prozesse modellieren. Zusätzlich kann an den Aktivitäten eigener Programmcode hinterlegt werden, der beim Aufruf der Aktivitäten ausgeführt werden. Über den Visual Studio Designer sind somit auch umfangreichere Workflows realisierbar. Die Erstellung solcher Workflows erfordert jedoch ein tiefes technisches Verständnis und Programmierkenntnisse, wodurch diese Prozesse aufwendig in der Erstellung und fehlerträchtig sind. Außerdem sind Änderungen an Prozessen nur schwer einzubringen.

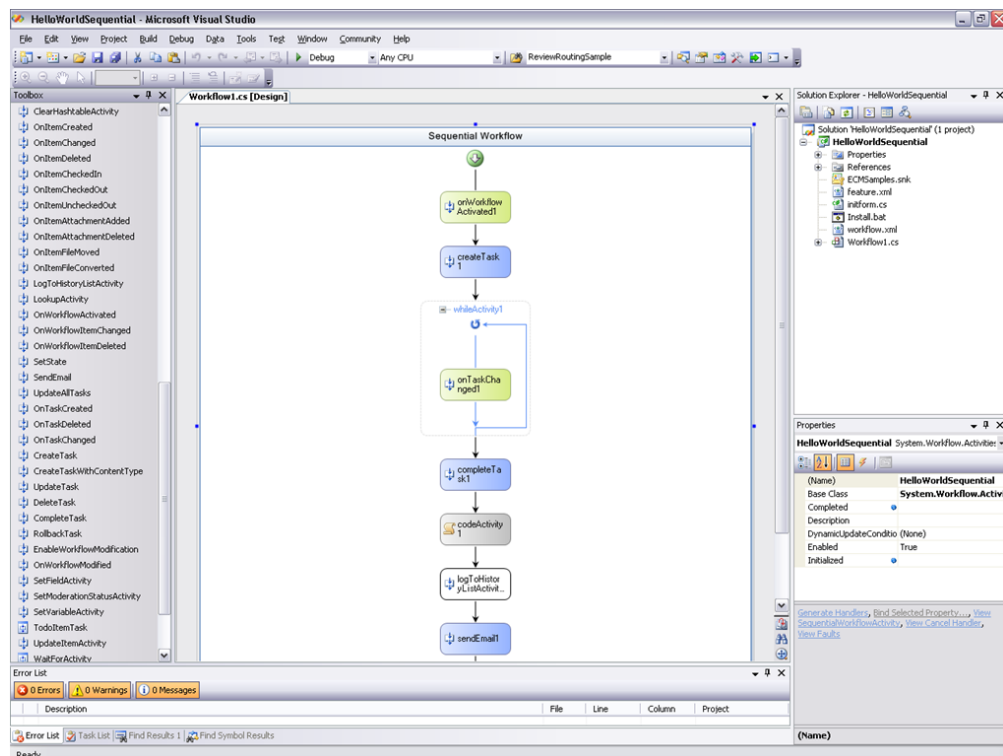


Abbildung 3.8.: Workflow Designer für Visual Studio 2005 [Cha06]

Neben der Erstellung eigener Workflows können auch bereits vordefinierte Workflows verwendet werden. Diese umfassen eine Dispositionsgenehmigung, um bei automatisiertem Löschen von Inhalten vorher eine Nachricht zu versenden, ob die Inhalte wirklich gelöscht werden sollen. Der sog. *Drei Status Workflow* mit Anfangs-, Zwischen- und Endzustand, dient dazu, abhängig von drei Zuständen, Nachrichten zu verschicken. Weiter existieren Feedback-, Signatur- und Genehmigungsworkflows, mit denen

Dokumente an einen Benutzer oder eine Gruppe von Benutzern gesendet werden kann, um Feedback, eine digitale Signatur oder eine Genehmigung des Dokuments anzufordern.

Durch die enge Integration mit MS Office lassen sich Workflows auch direkt in die Office-Programme integrieren, indem die Formulare, mit denen Benutzer Aktivitäten ausführen, direkt in den Programmen, statt im Browser angezeigt werden.

Tabelle 3.1 zeigt einige Unterschiede, die sich bei der Erstellung von Workflows in MOSS mit dem Sharepoint-Designer bzw. dem Visual Studio Designer ergeben.

Tabelle 3.1.: Unterschiede Sharepoint-Designer und Visual-Studio-Designer

Sharepoint-Designer 2007	Visual Studio 2005 Designer
Nur sequenzielle Workflows	sequenzielle und Statusworkflows
Aktivitäten auf MOSS-Funktionen beschränkt	Erstellung eigener Aktivitäten möglich
Keine Abweichungen in Workflows möglich	Abweichungen können in Workflow programmiert werden
Nur automatische Generierung von Formularen (anpassbar)	Erstellung eigener Formulare
Workflows werden auf Basis einer Liste erstellt und sind zur Entwurfszeit an die Daten gebunden	Workflow wird als Vorlage erstellt, die an mehreren Stellen verwendet werden kann
Keine Erweiterungen durch eigene Programmierung möglich	Eigene C# oder Visual Basic Programme können eingebunden werden

Store

Die Store-Komponenten stellen den Zugriff auf die in MOSS gespeicherten Inhalte über den Web-Client und die WebDAV-Schnittstelle zur Verfügung. Außerdem realisieren sie die Dokumentenverwaltungsfunktionen wie Check-In/Check-Out, die Versionsverwaltung und die Protokollierung von Änderungen für die Audit-Trails auf technischer Ebene. Sämtliche in MOSS gespeicherten Informationen werden inkl. der Indexdaten in einer SQL-Datenbank gespeichert. Daraus entstehen sehr große Datenbanken, die schnell auf viele TB anwachsen. Es werden keine Speicherbereiche außerhalb der Datenbank, wie etwa dem Dateisystem, optischen Laufwerken oder Bandlaufwerken unterstützt. Dadurch ist MOSS nur eingeschränkt tauglich für die Verwaltung großer Datenmengen, die im Rahmen von ECM (s. Kapitel 2.2) anfallen. Außerdem existieren einige Grenzwerte, die nicht überschritten werden sollten, um massive Performanceeinbußen zu vermeiden [Eng07, MHS08]. Dies sind z. B. 2000 Websites pro Website-sammlung, 2000 Listen pro Website, 2000 Elemente pro Ordner, 2000 Elemente pro An-

sicht. Diese Grenzwerte können in größeren Unternehmen schnell erreicht werden. Sie müssen bei der Planung berücksichtigt werden. Weiterhin kann durch Produkte von Drittanbietern eine Auslagerung der Inhalte aus der SQL-Datenbank vorgenommen werden.

Preserve

Da sämtliche gespeicherten Inhalte nur in der SQL-Datenbank von MOSS gespeichert werden, existiert auch keine Unterstützung für eine Langzeitarchivierung. Für diesen Bereich gibt es allerdings eine Vielzahl klassischer Archivsystemanbieter, die durch eigene Erweiterungen des MOSS versuchen, diese Lücke zu schließen. Typischerweise werden hierdurch die Dokumente aus der Sharepoint-Datenbank in das Archivsystem extrahiert und die eigentlichen Daten durch Verweise in das Archivsystem ersetzt.

Deliver

Die Transformationstechnologien in MOSS erfüllen nur teilweise die Anforderungen, die die AIIM an die Komponenten für ECM in diesem Bereich stellt. So ist keine Transformation von COLD-Daten vorhanden, Kompressionsverfahren zur Datenreduktion fehlen. Außerdem beschränkt sich die Bereitstellung von Inhalten in unterschiedlichen Formate auf MS Office-Formate. Im Bereich der Sicherheitstechnologien greift MOSS auf die Standardfunktionen des Windows Server-Betriebssystems zurück und unterstützt damit die Einbindung in eine PKI. DRM-Funktionen sind in MOSS enthalten und können auf alle Inhalte in Listen und Bibliotheken angewendet werden. Der Einsatz sollte jedoch gut geplant sein, da ansonsten sehr schnell Einschränkungen beim gemeinsamen Zugriff auf Inhalte auftreten können. Auch lassen sich Inhalte automatisch mit digitalen Signaturen zur Absicherung der Authentizität versehen.

3.4. ELOprofessional 6

ELOprofessional wurde im Jahr 1998 von der Firma ELO Digital Office GmbH eingeführt. Die Schwerpunkte der Software liegen auf dem Dokumenten-Management, Workflow-Management und der elektronischen Archivierung. Neben ELOprofessional werden noch ELOoffice für kleine Installationen und ELOenterprise angeboten. Seit Version 6.0 besitzen ELOenterprise und ELOprofessional die gleiche Architektur.

Aufbau und Architektur

ELOprofessional verwendet eine Java- und SOA-basierte Client-/Server-Architektur. Abbildung 3.9 zeigt einen Überblick über die Komponenten. Den Kern der Software bildet der ELO Index-Server, der den DM-Server und den Workflow-Server beinhaltet. Der Application-Server besteht aus einem Apache Tomcat, der die Funktionen über eine SOAP-Schnittstelle bereitstellt, über die die anderen ELO-Module, Clients oder Drittanwendungen mit dem Server kommunizieren können. Die Inhaltsspeicherung und -Zugriffe werden über die Serverbase-Komponenten koordiniert, die die jeweiligen Datenspeicher verwalten.

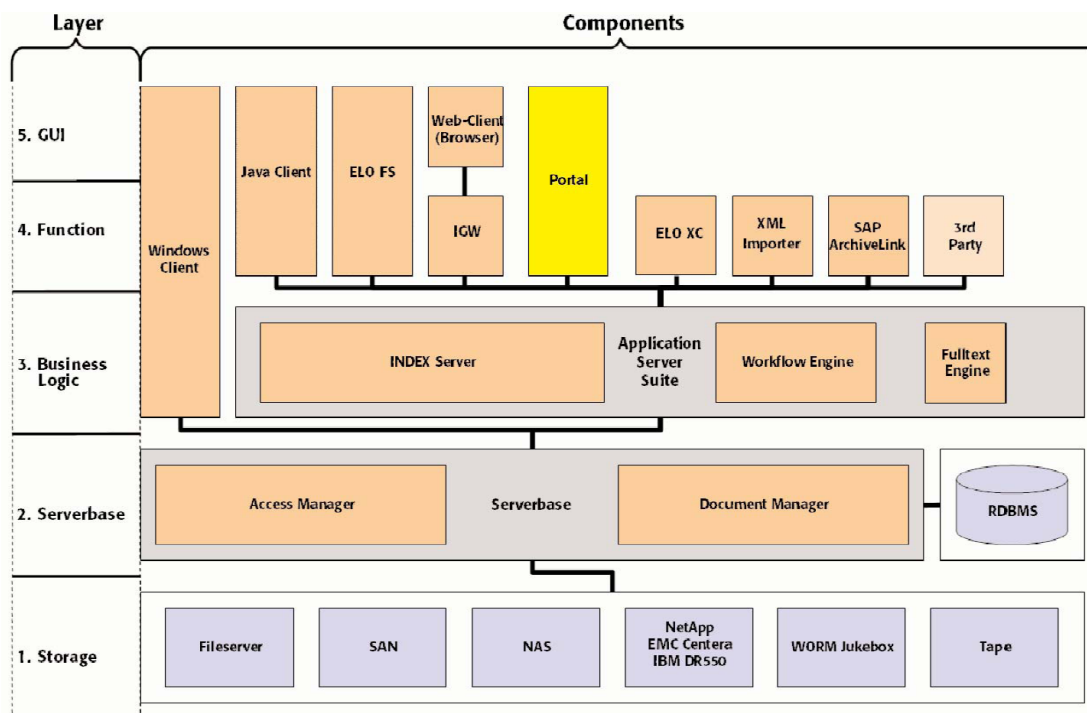


Abbildung 3.9.: Architektur von ELOprofessional [ELOb]

Die Benutzerverwaltung erfolgt bei ELOprofessional entweder über die integrierte Benutzerverwaltung oder durch eine Anbindung an MS Active Directory. Dabei können Benutzer und Gruppen aus diesem Verzeichnis in ELOprofessional importiert werden. Die Authentifizierung erfolgt durch den Verzeichnisdienst.

Capture

ELOprofessional unterstützt sowohl die Erfassung von NCI-, als auch die Erfassung von CI-Dokumenten. Für die Erfassung von Papierdokumenten steht ein integrierter Client für die Anbindung von Scannern zur Verfügung, mit dem einzelne Seiten oder mehrseitige Dokumente eingescannt werden können. Die so erfassten Dokumente können innerhalb des ELO-Clients nachträglich bearbeitet werden, um etwa Korrekturen wie leere Seiten Entfernen, Drehen von Seiten, Sortieren und Heften von Seiten zu zusammenhängenden mehrseitigen Dokumenten vorzunehmen. Die Dokumente werden dabei im TIFF-Dateiformat abgespeichert.

Mit der integrierten OCR-Texterkennung können die Inhalte der gescannten Dokumente in einen Volltextindex extrahiert werden, der in der SQL-Datenbank abgelegt wird. Metadaten von Dokumenten können mit den Werten aus dem Volltextindex versorgt werden. Eine automatische Zuordnung der Indexdaten aus dem Volltextindex zu den Attributwerten ist jedoch nicht möglich. Die Daten können lediglich beim Bearbeiten der Indexdaten aus einer Liste von erfassten Werten ausgewählt werden.

Um auf Dokumenten angebrachte Barcodes zu erfassen, mit denen Papierdokumente zugeordnet werden können, steht eine Komponente zur Erkennung von Barcodes zur Verfügung. ELOprofessional unterstützt eine manuelle Indexierung und eine Indexierung über Vorlagen, die in ELO angelegt werden und die Indexattribute der Dokumentenklasse definieren. Für die automatische Klassifikation und Erkennung von Dokumenten existiert das in Abbildung 3.10 dargestellte Zusatzmodul ELO DocXtractor für Eingangsdokumente und Formularerkennung.

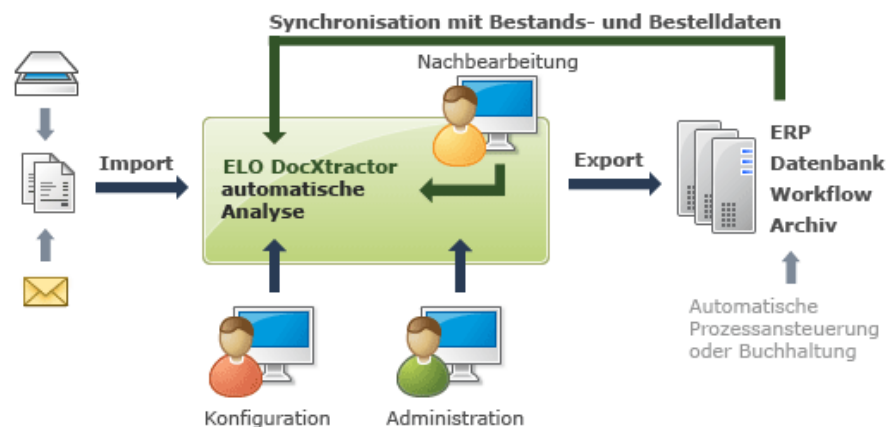


Abbildung 3.10.: ELO-DocXtractor [ELOa]

Dies wird mit Hilfe von Musterdokumenten realisiert, die bei der Erfassung von der DocXtractor-Komponente analysiert werden. Eingehende Dokumente werden einem

Dokumententyp zugeordnet, um die entsprechenden Indexdaten zu extrahieren. Dabei können die erkannten Informationen aus Dokumenten mit externen Datenquellen von angebundenen Systemen abgeglichen werden, um diese zu verknüpfen, etwa die Zuordnung einer eingehenden Rechnung zu einer Bestellung aus dem ERP-System.

Für die Erfassung von COLD-Daten stellt ELOprofessional eine Schnittstelle zur Verfügung, die eine Übernahme der Druckspooldateien, die Zerlegung in einzelne Dokumente, Abfrage von Indexinformationen und die Ablage der Indizes in der Indexdatenbank umfasst. Weiter werden Schnittstellen für diverse ERP-Systeme angeboten, über die Daten direkt archiviert werden können. Der ELO-Client umfasst außerdem einen virtuellen Archivdrucker, mit dem Daten aus beliebigen Anwendungen auf diesem Drucker in ELOprofessional abgelegt werden können.

Für die Erfassung aus Anwendungen besitzt ELOprofessional Plugins für MS Office. Damit lassen sich in Office erstellte Dokumente direkt in ELOprofessional abspeichern und mit den zugehörigen Metadaten versehen. Außerdem lassen sich Dokumente direkt aus den Office-Programmen suchen und öffnen.

Dokumenten-Management

ELOprofessional bietet für den Zugriff auf die DM-Funktionen drei Clients an: einen Windows-Client, Java-Client und einen Web-Client. Der Windows-Client bietet den vollen Funktionsumfang. Der Java-Client besitzt eine ggü. dem Windows-Client geänderte Oberfläche. Dadurch wird eine Navigation für die Benutzer erschwert, da die Funktionen an anderer Stelle präsentiert werden. Beim Funktionsumfang fehlen dem Java-Client Funktionen zur Benutzerverwaltung, Indizierung, Offlinearchivnutzung, Erzeugung und Anzeige von Berichten und Reports. Der Web-Client umfasst lediglich die wichtigsten Funktionen zum Dokumentenzugriff und zur Bearbeitung. Neben den speziellen Clients kann eine Erweiterung für den Zugriff über den Windows Explorer bereitgestellt werden. In dieser Erweiterung wird der Zugriff auf das ELOprofessional-Archiv über ein virtuelles Netzlaufwerk ermöglicht.

Die DM-Funktionen erfassen alle wichtigen Funktionen wie Check-In/Check-Out, Versionierung, Verwaltung von Zugriffsberechtigungen, Indexdatenverwaltung und eine strukturierte Ablage der Dokumente. Die Ablage wird durch hierarchisch angeordnete Verzeichnisse realisiert. Alle Dokumente lassen sich grundsätzlich in allen Verzeichnissen abspeichern oder verschieben, sofern die Berechtigungen des Benutzers ausreichend sind. Zusätzlich können Dokumente und Verzeichnisse als Verknüpfungen an beliebigen Stellen in der Ablagestruktur eingeblendet werden. Zur Anzeige der gespeicherten Dokumente steht ein integrierter Viewer zur Verfügung, der alle gängigen Dateiformate anzeigen kann. Für die gespeicherten Dokumente kann definiert werden, ob eine Versionierung aktiviert werden soll oder nicht. Wenn die Versionierung aktiviert ist, wird bei jedem Check-In eine neue Version automatisch erzeugt. Die

Versionsnummer kann individuell vergeben werden, und es steht ein Kommentarfeld zur Verfügung. Außerdem können die Dokumente als unveränderbar markiert werden. Eine Konvertierung in TIFF oder PDF steht für alle von ELOprofessional anzeigbaren Dateiformate zur Verfügung. Für die mobile Nutzung können Inhalte aus dem ELOprofessional-Server in eine auf dem Client installierte SQL-Datenbank repliziert werden, um Dokumente Offline im Zugriff zu haben. Dokumente können im Client über die Suchfunktion gefunden werden. Dazu kann eine auf den Indexattributen der Dokumente basierende Suchmaske oder Volltextsuche verwendet werden.

Records-Management

ELOprofessional deckt den Bereich RM im Wesentlichen über die Funktionen aus dem Bereich DM ab. Zusätzlich können Aktenpläne zur kontrollierten, hierarchischen Ablage von Records zentral erstellt und vorgegeben werden. Die Protokollierung von Änderungen an Dokumenten, die als Records klassifiziert werden, muss manuell aktiviert werden. Durch das Berechtigungsmodell ist es möglich, die Zugriffe auf einzelne Dokumente und Verzeichnisse zu definieren. Mit Hilfe von Stichwortlisten lässt sich eine thesaurusgestützte Verwendung von Metadaten realisieren. Aufbewahrungsrichtlinien von Records können über eine zugehörige Ablagemaske definiert oder individuell für einzelne Dokumente festgelegt werden.

Einschränkungen bestehen jedoch bei der Speicherung der Records, die nicht explizit in spezielle Archive oder auf spezielle physische Datenträger gesteuert werden kann. Somit lassen sich Anforderungen wie die sichere Vernichtung von bestimmten Dokumenten nur schwer oder gar nicht umsetzen, wenn diese Daten auf einem WORM-Medium mit anderen Daten zusammen gespeichert sind. Anforderungen an getrennte Datenhaltungen von Records lassen sich dadurch ebenfalls nicht realisieren.

Web-Content-Management

ELOprofessional besitzt in der evaluierten Version 6 keine Funktionen für den Bereich WCM. In Version 7 soll dieser Bereich abgedeckt werden.

Für die Portalintegration besitzt ELOprofessional eine Integration in MOSS. Bei der Integration in MOSS 2007 kann das komplette MOSS-Repository zur Inhaltsspeicherung ersetzt werden. Somit werden alle Inhalte statt in der SQL-Datenbank von MOSS im ELO-Archiv gespeichert. Damit wird auch eine Erweiterung der MOSS-Funktionen im Bereich Archivierung erreicht, in dem die Archivierungsfunktionen von ELOprofessional genutzt werden.

Collaboration

Collaboration-Funktionen finden sich in ELOprofessional kaum. So gibt es keine Funktionen zur gemeinsamen Bearbeitung von Dokumenten oder zur Terminplanung. Die einzige Funktion, um Kontakt zu einem anderen Benutzer aufzunehmen, ist eine einfache Nachrichtenübermittlungsfunktion, mit der eine Textmitteilung an einen anderen Benutzer verschickt werden kann.

Workflow

Der Bereich Workflow wird von ELOprofessional durch die integrierten *Adhoc-* und die *Standard-Workflows* abgedeckt. Mit Hilfe von *Adhoc-Workflows* kann ein Benutzer sequenzielle oder parallele Abläufe definieren, um Dokumente an einen oder mehrere andere Benutzer zu verteilen. Dabei kann ein *Freigabe-* oder ein *zur Kenntnisnahme Workflow* mit sequenzieller oder paralleler Zustellung des Dokuments gewählt werden. Nach Abschluss des Workflows kann optional ein Skript ausgeführt werden.

Die *Standard-Workflows* dienen der Definition und Ausführung von wiederkehrenden Prozessen. Dafür wird ein graphischer Editor bereitgestellt, mit dem die Workflows ohne Programmierkenntnisse graphisch modelliert werden können. Zusätzlich zu den sequenziellen und parallelen Ablaufstrukturen können auch Entscheidungsverzweigungen verwendet werden. Entscheidungen können automatisiert anhand von Indexdaten des Dokuments getroffen werden. In Abbildung 3.11 ist der *ELO-Workflow-Editor* mit einem Rechnungsfreigabeprozess abgebildet.

Weiterhin können Aufrufe von Funktionen in angebundenen Anwendungen während des Workflows ausgeführt werden (z. B. Ausführung einer SAP-Transaktion). Diese Aufrufe werden durch Skripte gesteuert, die an den Prozessschritten als Aktivitäten hinterlegt werden. Bei der Ausführung eines Prozessschritts kann der Verlauf des Workflows durch die Protokollierung nachvollzogen werden, indem angezeigt wird welche Benutzer im bisherigen Verlauf beteiligt waren bzw. welche Aktivitäten ausgeführt wurden. Bei Löschung eines Benutzers, der für die Ausführung von Prozessschritten definiert wurde, muss ein Nachfolger eingetragen werden, auf den die Workflows übertragen werden. Eine Vertreterregelung bei Abwesenheit eines Benutzers bietet die Möglichkeit, einen oder mehrere Stellvertreter anzugeben, an die die auszuführenden Prozessschritte zur Bearbeitung weitergeleitet werden.

Damit ermöglichen die Workflow-Komponenten von ELOprofessional dokumentenorientierte Workflows zur Verteilung und Freigabe von Dokumenten. Durch die Einbindung von Skripten können außerdem noch zusätzliche Aktionen in anderen Anwendungen initiiert werden.

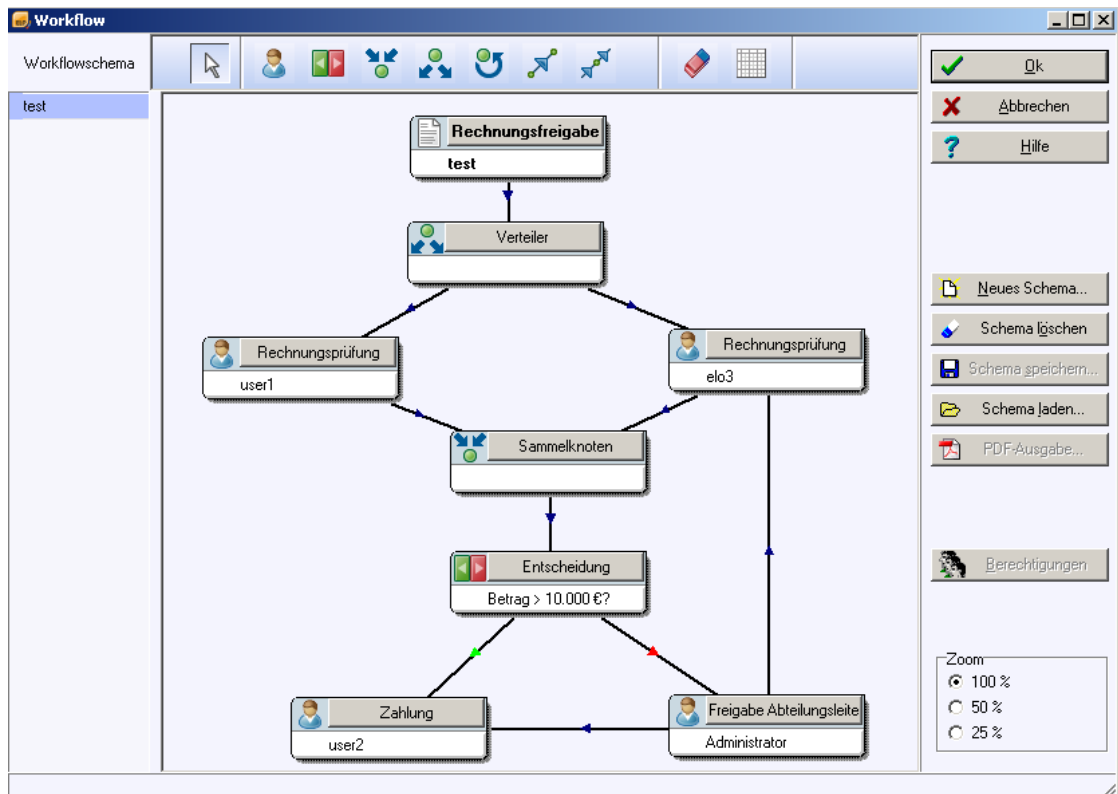


Abbildung 3.11.: Workflow-Editor

Store

Bei ELOprofessional sind die Bereiche Store und Preserve nicht explizit getrennt. Die Indexdaten werden bei ELOprofessional inkl. der Volltextindizes in einer relationalen Datenbank abgelegt. Hierzu können Datenbank-Server von Microsoft, Oracle oder IBM zum Einsatz kommen. Die Daten können im Dateisystem auf dem eigenen Server oder auf NAS- oder SAN-Systemen gespeichert werden. Optische Datenträger und Bandlaufwerke werden zur Speicherung insbesondere von Archivdaten ebenso unterstützt wie CAS-Archivserver von IBM oder EMC.

Preserve

Bei der Archivierung kopiert der ELOprofessional Backup Server die Daten in regelmäßigen Abständen im Hintergrund auf die entsprechenden Archivsysteme. Der Accessmanager von ELOprofessional kann die so ausgelagerten Daten beim Zugriff wieder direkt vom Archivsystem lesen und bereitstellen. Allerdings lassen sich lediglich 4 Spei-

cherorte für Archive in ELOprofessional verwalten. Zusätzlich kann ein Pfad für eine Kopie angegeben werden. Eine direkte Zuordnung von Archiven oder Ordnern zu bestimmten Speicherpfaden ist nicht möglich. Somit lässt sich keine physische Trennung der gespeicherten Daten vornehmen. Ebenso ist es nicht möglich die Informationen nach Wichtigkeit, Aufbewahrungsschutz oder anderen Kriterien auf unterschiedliche Medien gezielt zu verteilen. Dadurch ergibt sich für den Bereich der Aussonderung eine aufwendige Migration, wenn gefordert wird, dass bestimmte Daten auch physikalisch vernichtet werden müssen, da die Daten z. B. gemischt auf WORM-Medien vorliegen.

Deliver

Im Bereich Deliver besteht die Möglichkeit gespeicherte Dateien in PDF oder TIFF zu konvertieren. Eine Kompression der Daten, wie einem ZIP-Dateiformat wird nicht unterstützt. Die Sicherheitstechnologien umfassen den Einsatz von digitalen Signaturen zur Sicherung der Authentizität von Dokumenten.

3.5. DocuWare 5.1 ENTERPRISE

DocuWare ist seit 1988 Anbieter von DMS-Software und befindet sich wie eingangs des Kapitels dargestellt auf Platz 2 bei den eingesetzten DMS in Deutschland. Die Software setzt ihre Schwerpunkte auf die Bereiche DM und elektronische Archivierung. Die aktuelle Version 5.1 wird seit dem Jahr 2006 angeboten und setzt sich selbst zum Ziel integriertes DM anzubieten und zukünftig den gesamten ECM-Bereich abzudecken [Doc05].

Aufbau und Architektur

Die Architektur von DocuWare 5 folgt einem mehrschichtigen Systemansatz (s. Abbildung 3.12). Die DocuWare-Server-Komponenten basieren auf dem Microsoft .NET-Framework und stellen die Dienste für die Funktionen, auf einem oder mehreren Servern verteilt, bereit wodurch erhöhte Ausfallsicherheit und eine Lastverteilung auf mehrere Server vorgenommen werden kann. Für jeden Bereich wird ein eigener Serverdienst (*Content-Server*, *Authentication-Server*, *Workflow-Server* und *Web-Server*) eingesetzt [Doc05]. DocuWare setzt bei der Speicherung der Informationen auf eine dateibasierte Speicherung der Binärdaten und auf eine relationale SQL-Datenbank (MS SQL-Server, MySQL oder Oracle) für die Speicherung der Meta- und Indexdaten. Dabei stehen für den Zugriff, neben einer eigenen Benutzerverwaltung, Anbindungen an LDAP-Systeme zur Verfügung. Die Clients umfassen einen Web-Client mit reduziertem Funktionsumfang und einen .NET-basierten Rich-Client, der alle Funktionen beinhaltet.

tet. Durch die Realisierung der DocuWare-Server auf Basis der .NET-Architektur von Microsoft können diese nur auf der Windows-Plattform von Microsoft eingesetzt werden [Doc05]. Die weiter benötigten Infrastruktur-Komponenten Datenbank und Dateisystem können jedoch auch über Nicht-Windows-Plattformen bereitgestellt werden.

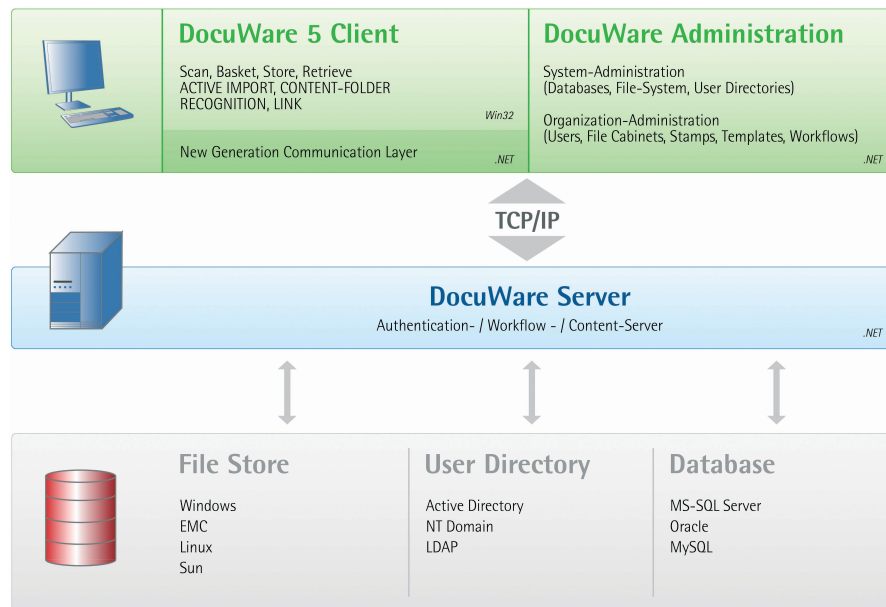


Abbildung 3.12.: DocuWare-Architektur [Doc05]

Capture

DocuWare 5 unterstützt die Erfassung von NCI- und CI-Dokumenten. Für die Erfassung von Papierdokumenten besitzt der DocuWare Rich-Client integrierte Scan- und OCR-Funktionen. Es können ein- oder mehrseitige Dokumente gescannt werden. Für die nachträgliche Bearbeitung der Dokumente stehen Funktionen zum Zusammenfassen bzw. Trennen von Dokumenten, Löschen von Leerseiten und Drehen von Dokumenten zur Verfügung. Die Speicherung der gescannten Dokumente erfolgt im TIFF- oder JPEG-Format. Die OCR-Erkennung unterstützt die vollständige Indizierung von Dokumenten und die Erkennung von einzelnen markierten Bereichen, um Indexdaten aus bestimmten Teilen des Dokuments zu extrahieren.

Weiterhin ist eine Barcodeerkennung zur Zuordnung von Dokumenten vorhanden. Zur Formularverarbeitung können Vorlagen mit Musterdokumenten erstellt werden, mit denen diese in Bereiche unterteilt werden, um die Indexdaten automatisch aus den Dokumenten extrahieren zu können. In Abbildung 3.13 ist die Konfiguration für die Erstellung der Erkennungsbereiche zu sehen. Zusätzlich kann die Erkennungsrate er-

höht werden, indem Wertebereiche in den jeweiligen Feldern eingeschränkt werden. Die erkannten Daten können gegen Informationen aus externen Datenquellen geprüft und abgeglichen werden. Somit lassen sich kleinere Erkennungsfehler korrigieren, etwa bei der Anschrift eines Kunden. Damit stehen für die Erfassung von Papierbelegen die Funktionen der manuellen Erfassung und die Erfassung über Vorlagen zur Verfügung. Eine automatische Erkennung mit Klassifizierung ist nicht vorhanden.

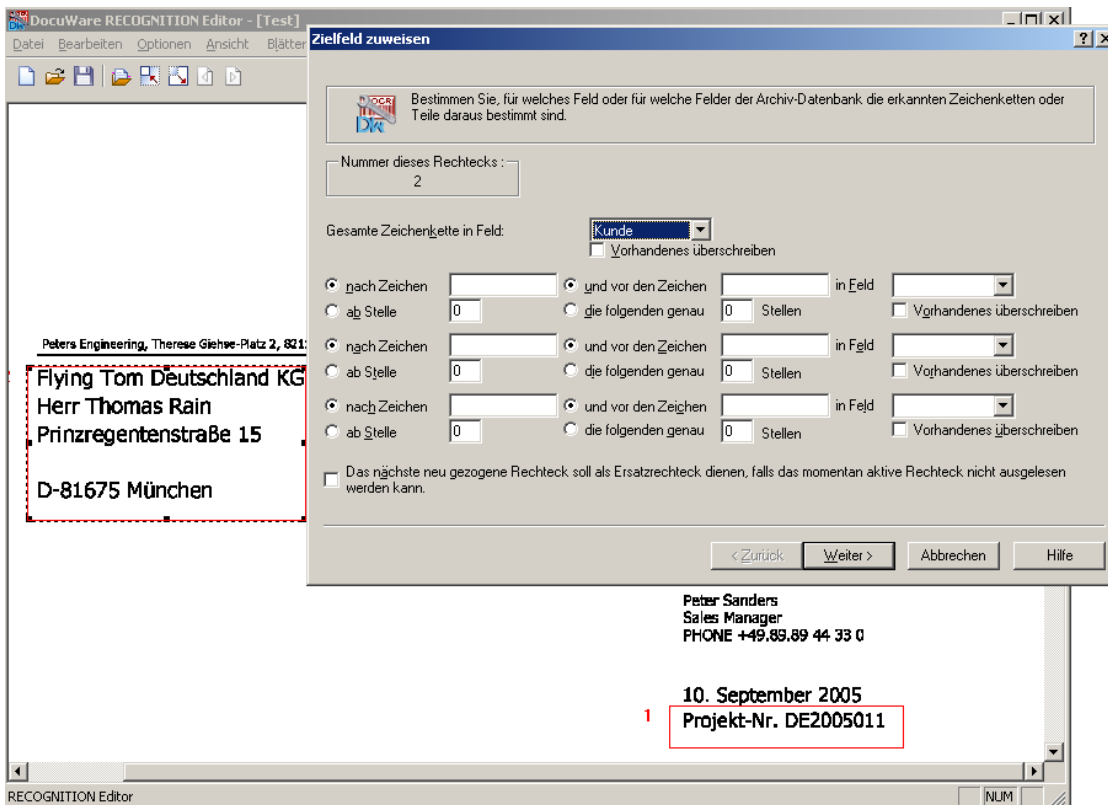


Abbildung 3.13.: Zuweisung von Indexfeldern

Die Erfassung von COLD-Daten wird über ein entsprechendes Modul bereitgestellt, mit dem die erfassten Daten entweder im TIFF-Format oder in einem proprietären DocuWare-Format (CL5) verarbeitet werden, das Layoutinformationen und Inhalte getrennt voneinander abspeichert. Die Zuordnung von Layout und Daten wird durch DocuWare verwaltet. Darüber hinaus werden spezielle Schnittstellen zur Archivierung von Daten aus ERP-Systemen angeboten. Der DocuWare Rich-Client umfasst zusätzlich einen virtuellen Archivdrucker, mit dem Druckausgaben aus beliebigen Anwendungen im TIFF-Format in DocuWare abgespeichert werden können.

Für die Erfassung aus Anwendungen lassen sich mit dem Modul *Link* Informationen aus jeder Applikation, die auf dem Client läuft, über den angezeigten Fensterinhalt

auslesen. Diese gelesenen Informationen können für Suchanfragen oder zum Ausfüllen von Indexwerten bei der Erfassung verwendet werden. Je nach Programm können die Daten direkt als CI-Informationen ausgelesen werden oder über einen Screenshot mit anschließender OCR-Erkennung extrahiert werden. Darüber hinaus steht ein Plugin für die MS Office-Suite zur Verfügung, mit dem Dokumente direkt in DocuWare gespeichert bzw. geöffnet werden können. Attribute und Metadaten können damit bei der Speicherung direkt an DocuWare übergeben werden.

Dokumenten-Management

Der Bereich DM ist einer der Hauptfunktionsbereiche von DocuWare 5. Der Rich-Client bietet den vollen Funktionsumfang. Der Web-Client umfasst lediglich Basis-Funktionen zum Dokumentenzugriff und zur Bearbeitung. Der Zugriff auf die Dokumente über eine Dateisystemintegration ist nicht vorhanden. Es existieren allerdings Erweiterungen von Drittanbietern, die diesen Bereich vergleichbar zu ELOprofessional, MOSS oder Alfresco abdecken.

Die DM-Funktionen erfassen alle wichtigen Funktionen wie Check-In/Check-Out, Versionierung, Verwaltung von Zugriffsberechtigungen, Indexdatenverwaltung und eine strukturierte Ablage der Dokumente. Die Ablage ist getrennt in den Archivbereich, in dem Versionierung und die Langzeitarchivierung unterstützt werden, und in Briefkörbe, die für die temporäre Bearbeitung lebender oder neu erfasster Dokumente dienen. Die Briefkörbe werden für jeden Benutzer getrennt bereitgestellt und werden direkt im Dateisystem des Clients oder einer Netzwerkfreigabe außerhalb des DocuWare Content-Servers gespeichert. Ein Zugriff durch andere Benutzer ist dadurch nicht möglich. Zur Anzeige der gespeicherten Dokumente steht ein integrierter Viewer zur Verfügung, der alle gängigen Dateiformate anzeigen kann. Außerdem lassen sich über den Viewer zusätzliche Informationen zu den Dokumenten, wie Notizen, Stempel oder Markierungen speichern.

Die Archive werden über die DocuWare-Administrationsoberfläche angelegt und auf dem Server zentral gespeichert. Der Zugriff auf Dokumente erfolgt über eine Suchanfrage im jeweiligen Archiv. Grundlegende Voraussetzung für die Einrichtung eines Archivs sind die Auswahl eines Content Servers, die Auswahl eines Speicherortes sowie die Auswahl einer Datenbankverbindung. Für die gespeicherten Dokumente kann definiert werden, ob eine Versionierung aktiviert werden soll, die bei jedem Check-In eine neue Version erstellt. Eine direkte Konvertierung von Dokumenten in TIFF oder PDF steht zur Verfügung. Allen Dokumenten kann eine Sprachaufzeichnung für weitere Informationen hinzugefügt werden. Für die mobile Nutzung können Inhalte aus dem DocuWare-Archiv auf einen Client-PC synchronisiert werden. Die vorgenommenen Änderungen werden bei der nächsten Verbindung wieder mit dem DocuWare-Server repliziert. Dokumente können im Client über die Suchfunktion gefunden werden. Da-

bei kann bei der Suche eine vordefinierte Suchmaske aus dem Administrationsbereich ausgewählt werden. Eine Suche kann auch über mehrere Archive hinweg erfolgen.

Records-Management

DocuWare 5 deckt den Bereich RM durch Funktionen aus den Bereichen DM und durch die Archivierungsfunktionen ab. So können Aktenpläne für die strukturierte, hierarchische Ablage von Records angelegt werden. Records können im Aktenplan mit Aufbewahrungsrichtlinien versehen werden. Diese werden vom System überwacht. Durch Berichte werden die zur Löschung anstehenden Records ermittelt, die anschließend aus den Archiven entfernt werden können. Durch die individuelle Zuordnung von Archiven zu Speicher- oder Archivspeicherorten werden die Daten auch gezielt auf den für die Records definierten Aufbewahrungsmedien gespeichert. Dadurch wird auch ermöglicht, dass Datenträger nur bestimmte Records enthalten, die dann gezielt beim Erreichen von Löschfristen entnommen und vernichtet werden. Papierakten und Lagerorte können durch passende Konfiguration der Indexdaten durch die DocuWare-Datenbank mitverwaltet werden. Durch die Rechteverwaltung kann der Zugriff auf die Records über Benutzer, Gruppen und Rollen definiert werden. Protokollierungen über die Änderungen an den Records werden aktiviert, um Nachweise über die Art der Änderung und die Zuordnung zu Benutzern zu erbringen. Die Authentizität der Records wird über die Anbringung von Stempeln oder handschriftlichen Signaturen auf den Dokumenten sichergestellt. Darüber hinaus werden digitale Signaturen unterstützt.

Web-Content-Management

Der Bereich Web-Content-Management beschränkt sich im Wesentlichen auf Funktionen des Web-Clients aus dem DM-Bereich. Um verschiedene Bereiche für den Webzugriff bereitstellen zu können, müssen die Webseiten manuell durch Änderungen an den CSS- und HTML-Dateien vorgenommen werden. Für eine Portalintegration steht eine Integration in MOSS von Microsoft zur Verfügung. Hierüber können dann echte WCM-Funktionen umgesetzt werden. Dabei können zugleich die Einschränkungen des MOSS im Hinblick auf Archivierung abgedeckt werden, da die Integration ebenso wie bei ELOprofessional einen transparenten Zugriff auf die in MOSS gespeicherten Informationen ermöglicht. Die Archivierung kann im Gegensatz zu ELOprofessional in bestimmte Archive gesteuert werden. Somit lassen sich auch unterschiedliche Aufbewahrungsregeln anhand der in MOSS gespeicherten Informationen abbilden.

Collaboration

Der Bereich Collaboration wird von DocuWare nicht abgedeckt. Es existieren keine gemeinsam nutzbaren Arbeitsbereiche oder Koordinationsfunktionen wie Terminplanungen. Durch die Integrationsmöglichkeiten in MOSS können die Funktionen von MOSS für Collaboration genutzt werden.

Workflow

DocuWare 5 steuert sowohl systeminterne Vorgänge, wie die Konvertierung von Dateiformaten, Imports aus Fremdsystemen, Löschen von Daten, als auch einfache Freigabeprozesse über Workflows.

Beide Arten von Prozessen werden über den *Workflow-Server* gesteuert. Die Ausführung von Geschäftsprozesse wird in DocuWare 5 über Stempel realisiert. Dazu wird für jede Aktivität ein Stempel definiert, der benutzerbezogen im DocuWare Viewer zur Verfügung gestellt wird. Ein Stempel enthält die Benutzerzuordnung, einen frei definierbaren Text und optional eine digitale Signatur. Bei Einrichtung des Stempels wird die zugehörige Aktivität definiert.

Das Anbringen eines Stempels auf einem Dokument wird durch einen Eintrag in der DocuWare-Datenbank vermerkt. Diese Einträge werden vom *Workflow-Server* auf Änderungen zyklisch überprüft. Das Dokument wird dann anhand des angebrachten Stempels in den damit verknüpften Dokumentenpool verschoben. Die Stempel dienen neben dem Dokumentenrouting auch der Nachverfolgung eines Prozessverlaufs, indem alle angebrachten Stempel auf dem Dokument angezeigt werden. Informationen und Daten aus Drittanwendungen können nicht in Workflows eingebunden werden. Mit den bereitgestellten Funktionen über die DocuWare-Stempel lassen sich damit nur sehr einfache und kurze Workflows realisieren. Eine übersichtliche Anzeige über den genauen Prozessverlauf existiert nicht. Der Benutzer kann lediglich anhand der an beliebigen Stellen auf dem Dokument angebrachten Stempel den Prozessverlauf nachvollziehen.

Store

Der Bereich Store wird vom Content-Server abgedeckt. Die Indexdaten werden in der relationalen Datenbank abgelegt, die auch die Ergebnisse von Suchanfragen beantwortet. Optional kann ein externer Volltextindex angelegt werden. Die Nutzdaten werden über den Content-Server gespeichert und wieder abgerufen bzw. werden im Fall der Postkörbe außerhalb des DocuWare-Servers im Dateisystem gespeichert. Die Speicherung kann direkt im Dateisystem des Servers erfolgen oder auf angebundenen NAS-

oder SAN-Systemen oder NFS¹⁸-Freigaben. Außerdem werden Archiv-Server von EMC und IBM zur Ablage unterstützt.

Preserve

Die Archivierungskomponenten werden vom DocuWare Content-Server verwaltet. Er stellt die Schnittstellen und Zugriffe auf Archivsysteme wie Jukeboxen mit optischen Medien, Bandlaufwerken oder CAS-Systemen bereit. Die Speicherorte lassen sich hier durch die Konfiguration der Archive definieren. Die Informationen zu Aufbewahrungs- und Löschrufen werden dabei aus dem Bereich RM verwendet.

Deliver

Der Bereich der Transformationstechnologien wird von DocuWare neben den integrierten Konvertierungsfunktionen des Rich-Clients über das Einrichten von *Konvertierungsworkflows* unterstützt. Die Personalisierung beschränkt sich beim Rich-Client auf das Auswählen der angezeigten Archive. Der Web-Client kann zusätzlich manuell im Layout angepasst werden. Daten können in DocuWare sowohl im Originalformat abgespeichert werden, als auch in komprimierter Form durch ZIP-Dateien. Die Daten werden vom Server automatisch bei der Speicherung komprimiert und beim Zugriff entsprechend entpackt. Bei den Sicherheitstechnologien können digitale Signaturen zur Sicherstellung der Authentizität verwendet werden. Bei der Verteilung der Dokumente stehen die Funktionen aus dem Bereich Workflow zur Verfügung, um Inhalte an andere Benutzer oder Systeme weiterzureichen.

3.6. EASY ENTERPRISE.x 3.0

Die EASY Software AG ist seit 1990 Anbieter von DMS-Software und Marktführer im Bereich DMS in Deutschland. Schwerpunkte der EEX¹⁹-Software stellen wie bei ELO-professional und DocuWare 5 die elektronische Archivierung und das Dokumenten-Management dar.

¹⁸Network File System

¹⁹EASY ENTERPRISE.x

Aufbau und Architektur

EEX basiert auf einer mehrschichtigen Softwarearchitektur wie in Abbildung 3.14 dargestellt. Die drei Schichten (ENTERPRISE.x Kernel, ENTERPRISE.x Beans und ENTERPRISE.x Web-Server) beinhalten die wesentlichen Komponenten der Software. Die Speicherung der Daten erfolgt für die Indexdaten in einer relationalen Datenbank (MySQL, MS SQL-Server, Oracle, PostgreSQL oder IBM DB2). Für die Volltextindexinformationen kann zwischen einer Apache Lucene oder Intrafind Datenbank gewählt werden. Die Ablage der Nutzdaten erfolgt im Dateisystem bzw. über die Anbindung anderer Speichersysteme. Die Anbindung der Speichersysteme erfolgt über den EEX-Kernel, der die Komponenten zur Speicherung und Archivierung der Dokumente und Daten bereitstellt.

Die EEX-Beans verbinden den Kernel über Java RMI mit dem Web-Server und stellen die Java-basierten Schnittstellen für den Zugriff der EEX Rich-Clients, sowie weiterer Module oder eigener Anwendungen mit Java-Schnittstelle bereit. Die EEX-Beans setzen einen J2EE-Applikationsserver voraus, der aus einem *JBoss*, *IBM Websphere* oder *BEA Weblogic* Server bestehen kann [JBo09, IBM09a, Ora09]. Der EEX-Web-Server (*Apache Tomcat*) stellt die Schnittstelle für die Web-Clients und Portale gegenüber dem Benutzer dar. Der XML-Server stellt eine Schnittstelle für ältere Module und Anwendungen von Drittherstellern bereit.

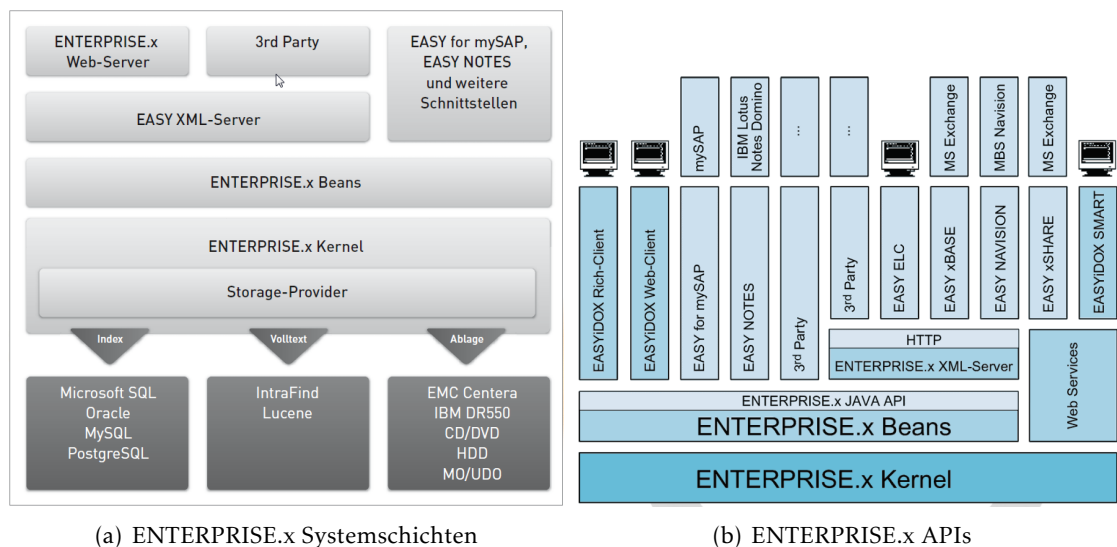


Abbildung 3.14.: EASY ENTERPRISE.x-Architektur [EAS06b]

Alle abgebildeten Systemkomponenten können sowohl komplett auf einem Server zusammengefasst oder auf mehrere Server(-Cluster) verteilt werden. Neben einer inte-

grierten Benutzerverwaltung können verschiedene LDAP-Verzeichnisdienste für die Authentifizierung und Benutzerverwaltung integriert werden.

Im Bereich der Clients stehen zwei .NET-basierte Clients (*EASYiDOX SMART* und *EASYiDOX JUMP!*) zur Verfügung. Ersterer integriert sich über Plugins direkt in die MS Office-Produkte Word, Excel und Outlook und bietet Recherchefunktionen in EEX-Archiven. Der zweite integriert sich außerdem in die .NET-Komponenten von Windows, um sich z. B. in die Desktopsuche von Windows XP oder Vista zu integrieren. Daneben existieren zwei Eclipse-basierte Clients [Ecl09] (*EASYiDOX* und *EASY Experience*), die alle DM-Funktionen inkl. Scannen abdecken. Letzterer besitzt zusätzlich eine Unterstützung für Adhoc-Workflows. Zuletzt existiert noch der *EASYiDOX Web-Client*, der den Webzugriff auf die Daten ermöglicht.

Capture

EEX unterstützt die Erfassung von NCI- und CI-Dokumenten. Für die Erfassung von Papierdokumenten stehen in den Rich-Clients Scanfunktionen zum Erfassen ein- oder mehrseitiger Dokumente zur Verfügung. Für die nachträgliche Bearbeitung bieten die Rich-Clients Funktionen zum Zusammenfassen bzw. Trennen von Dokumenten, Löschen von Leerseiten und Drehen von Dokumenten. Die Speicherung der gescannten Dokumente erfolgt im TIFF-, PNG- oder JPEG-Format. Die integrierte OCR-Erkennung ermöglicht die vollständige Indizierung von Dokumenten oder markierten Teilbereichen wie in Abbildung 3.15 dargestellt.

Neben den integrierten Funktionen zur Erfassung von Papierdokumenten ermöglicht das in Abbildung 3.15 dargestellte Modul *EASY Capture* die Stapelerfassung von größeren Mengen an Papierdokumenten. Dazu wird ein eigener Scan-Client verwendet, der die Aufgaben des Scannens und der Indexierung, sowie Sortieren, Zusammenfassen bzw. Trennen und Bildkorrekturen vornehmen kann. In diesen Client sind diverse Erkennungsmodule von Drittanbietern für die Bereiche OCR, Barcode, ICR oder HCR integriert. Um eine höhere Erkennungsrate zu erreichen, werden die Dokumente beim Scannen von mehreren Erkennungs-Engines analysiert. Durch ein Voting-Verfahren wird das wahrscheinlichste Muster anschließend ausgewählt [EAS06a]. Erkannte Barcodes können mit dem Datenbestand aus anderen Quellen abgeglichen werden, um z. B. Dokumente einem Auftrag aus dem ERP-System zuordnen zu können. Zusätzlich können Vorlagen für Dokumententypen erstellt werden, in denen die Bereiche für die Extraktion der Indexdaten definiert werden können (s. Abbildung 3.16). Nicht automatisch zugeordnete Dokumente können dabei zur manuellen Prüfung zurückgestellt werden.

COLD-Daten können über ein COLD-Modul erfasst werden. Es separiert die enthaltenen Dokumente und extrahiert die Indexfelder und den Volltext. Zusätzlich existiert auch ein virtueller Drucker, der aus beliebigen Anwendungen die Druckdaten

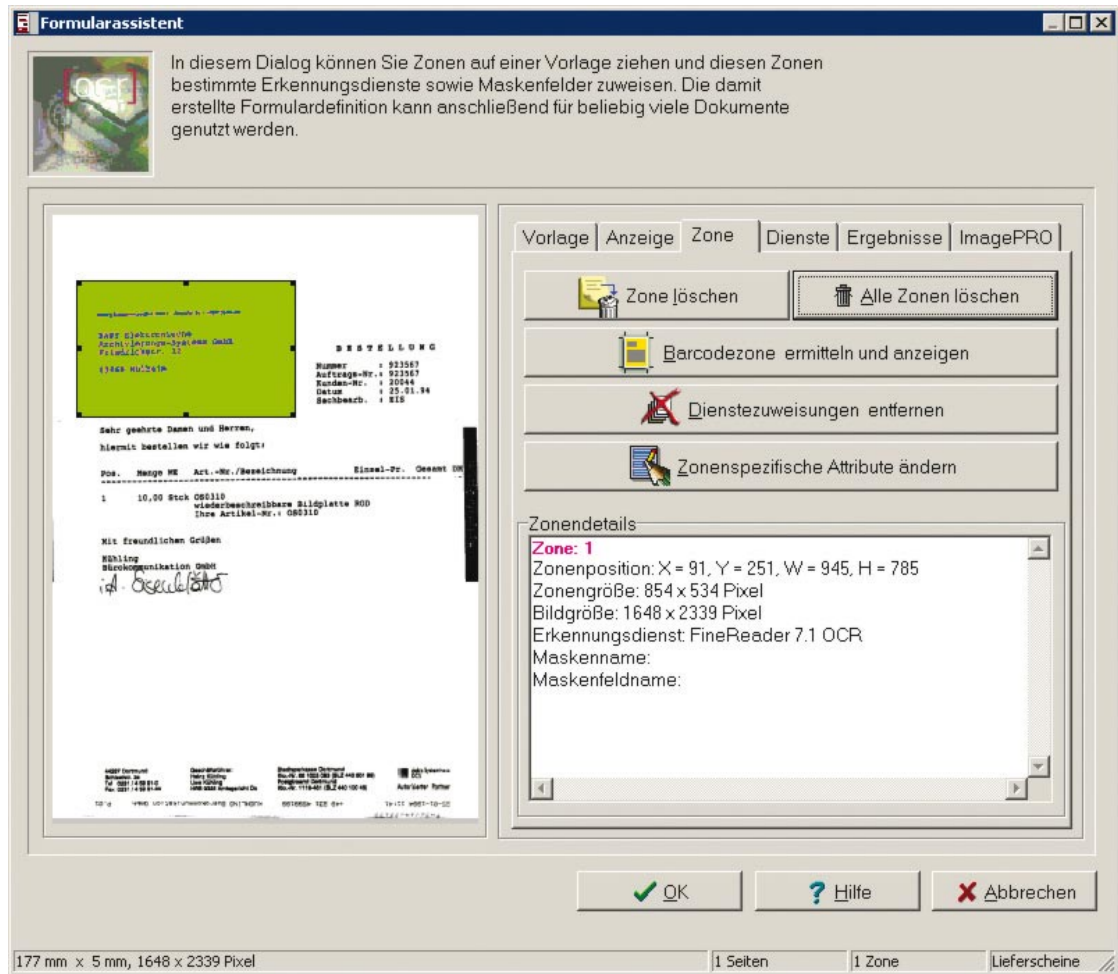


Abbildung 3.15.: EASY-Formularassistent

entgegennimmt und im TIFF- oder PDF-Format in EEX ablegen kann. Für diverse ERP-Systeme stehen Schnittstellen zur Archivierung der ERP-Daten, Integration von Suchfunktionen aus dem ERP-Client und zusätzliche Funktionen, z. B. automatische Buchungsvorgänge anhand erfasster Papierbelege, zur Verfügung.

Die Erfassung aus MS Office-Anwendungen werden über Plugins realisiert, mit denen Office-Dokumente direkt in EEX abgespeichert und geöffnet werden können. Weiterhin können Dateien eines Dateiservers transparent in EEX abgelegt werden. Dabei werden die Daten vom Dateiserver in EEX gespeichert und durch Verweise ersetzt, die durch einen Treiber auf dem Client wieder aufgelöst werden und den Datenzugriff ermöglichen.

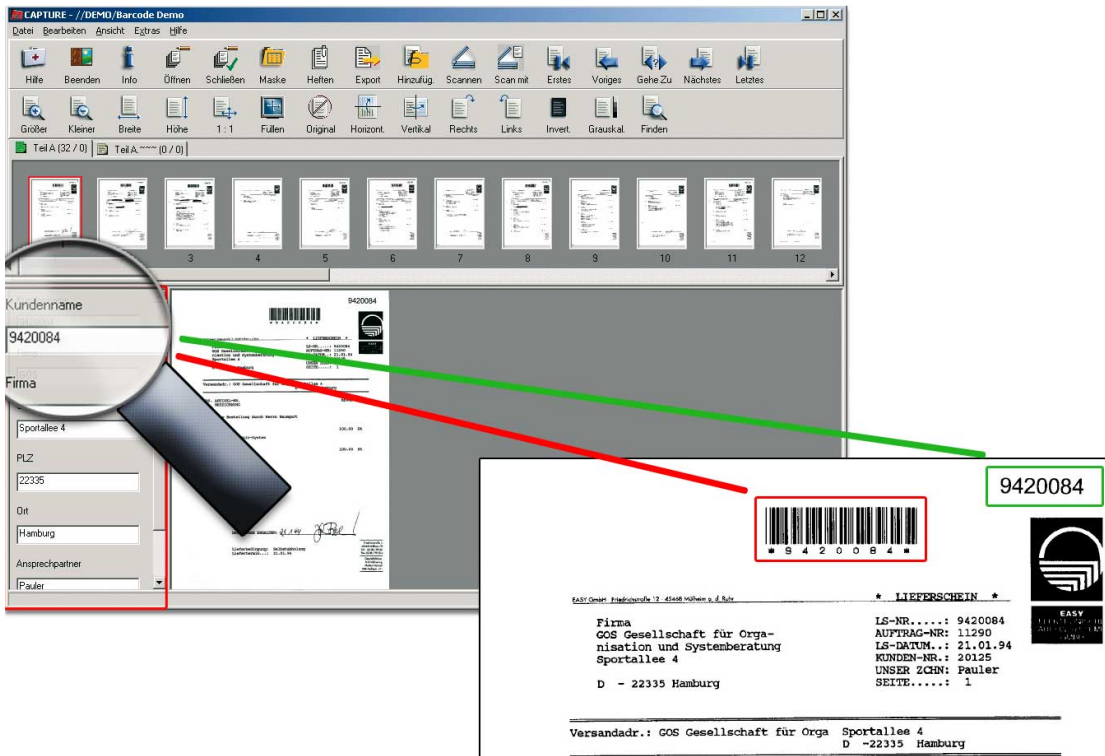


Abbildung 3.16.: Barcode-Erkennung mit EASY Capture [EAS06a]

Dokumenten-Management

Für den Bereich DM stehen die in Abschnitt 3.6 beschriebenen Clients zur Verfügung. Es werden alle wichtigen Funktionen, wie Check-In/Check-Out, Versionierung, Verwaltung von Zugriffsberechtigungen, Indexdatenverwaltung und eine strukturierte Ablage abgedeckt.

Für die Verwaltung der Dokumente stehen im Client private Ordner, gemeinsam genutzte Ordner und der Zugriff auf die EEX-Archive zur Verfügung. In den privaten Ordnern können beliebige Dokumente erstellt und bearbeitet werden. Zur Anzeige der Dokumente steht ein integrierter Viewer zur Verfügung, der alle gängigen Dateiformate unterstützt. Die Anzeigefunktion kann entweder vom Client übernommen werden, der die Berechnung der Anzeige vornimmt, oder über den Server erfolgen, der die Aufbereitung übernimmt und die generierten Bilder an den Client zur Anzeige überträgt. Im Viewer stehen Funktionen wie Stempel und das Anbringen von Notizen oder Textmarker-Funktionen zur Verfügung. Die Archive werden zentral am EEX-Server in der Administrationskonsole verwaltet. Der Aufbau und die Verwaltung von Archivordnern zusammen mit Dokumentenschemata entspricht dem Aufbau von Records und wird daher im Abschnitt Records-Management behandelt.

Zur Suche können über die Administrationskonsole oder den Client selbst individuelle Suchanfragen generiert und zur Wiederverwendung gespeichert werden. Die Suchformulare können beliebige Archive und Felder aus den gespeicherten Dokumenten enthalten. Für die mobile Nutzung können die privaten Ordner offline verwendet werden. Ein Zugriff auf Archive ist offline nicht möglich. Hierzu müssen die gewünschten Dokumente zuvor aus den Archiven in den privaten Bereich ausgecheckt werden oder aus dem Archiv exportiert werden, um Konflikte bei Änderungen zu vermeiden.

Records-Management

Für den Bereich RM stehen umfassende Funktionen zur Verfügung. In EEX besteht ein Record aus mehreren Komponenten, die immer aus einem oder mehreren Nutzdanelementen, den eigentlichen Dokumenten, bestehen. Jeder Record gehört zu einem Schema, in dem Aufbau und Struktur der Records definiert werden. Records werden in einem Pool abgelegt für den definiert wird, welche Schemata er zur Speicherung zulässt. Änderungen an Schemata sind nur zulässig, wenn sie die Konsistenz von bereits mit einem Schema gespeicherten Record nicht verletzen.

Die Attribute von Records werden in der SQL-Datenbank abgelegt. Dateien und weitere Metadatenfelder können getrennt auf unterschiedliche Speicherorte und Datenbanken verteilt gespeichert werden. Dokumente können entweder in einem Schema eingebettet sein oder durch Links verknüpft sein. Eingebettete Dokumente werden fest mit dem Mutterdokument verbunden und zusammen gespeichert. Durch Links verknüpfte Dokumente sind eigenständig und können auch bei bestehenden Verweisen gelöscht werden. Zusätzlich ist einem Schema eine Verhaltensklasse zugeordnet, die das Verhalten der Systemattribute steuert, z. B. ob sie bei einer Suchanfrage nur dem Besitzer des Dokuments angezeigt werden oder allen Benutzern mit Leseberechtigungen. Dies schafft eine sehr feingranulare Definition von Zugriffsberechtigung und die Voraussetzungen für die Erfüllung von Compliance-Kriterien. Berechtigungen können bis auf einzelne Datenelemente innerhalb von Records und Dokumenten getrennt und auf Funktionen, wie Drucken oder E-Mail-Versand, vergeben werden.

Aktenpläne können über hierarchische Beziehungen zwischen den Archiven realisiert werden. Weitere inhaltliche Berechtigungen, die unabhängig vom Speicherort vergeben werden, stehen ebenfalls zur Verfügung. Beispielsweise darf Benutzer A Rechnungen sehen, Benutzer B darf keine Rechnungen sehen. Ein Record kann mit den RM-Richtlinien verwaltet werden, indem der Dokumentenklasse des zugrunde liegenden Schemas ein Aktenplan zugeordnet wird. Für die Einhaltung und Definition eines Thesaurus stehen Auswahllisten zur Verfügung, deren Inhalte anhand des Aktenplans und der Dokumentenklasse für Dokumentenattribute verwendet werden.

Auch die Audit-Trail-Funktionen sind bei EEX sehr umfangreich und bieten die Möglichkeit alle Aktionen inkl. lesendem Zugriff an Dokumenten mit Benutzer und Zeit-

stempel zu protokollieren. Dabei können die Protokolle auf verschiedene Protokolldateien aufgeteilt werden.

Jede Record-Klasse kann mit individuellen Aufbewahrungsrichtlinien verknüpft werden. Nach Ablauf der Aufbewahrungsfrist kann festgelegt werden, ob die Records dann noch durch Benutzer zugreifbar sein sollen oder für den Zugriff gesperrt werden. Aussondernde Records werden über *Retention-Cleanup-Jobs* identifiziert. Sie können damit direkt automatisch gelöscht werden oder alternativ nur als gelöscht markiert werden. Damit sind die Records dann nur noch für Benutzer mit den entsprechenden Berechtigungen sichtbar. Die Records müssen dann manuell gelöscht werden.

Web-Content-Management

Der Bereich WCM wird über das ELC²⁰ abgedeckt. ELC stellt ein Webportal bereit, das die Funktionen WCM, DM und einen integrierten Webshop beinhaltet. Fremdanwendungen lassen sich über eine Java-API in das Portal integrieren. Für die Erstellung von Webinhalten ist ein Redaktionssystem mit integrierten Editoren für HTML, Layout und einfache Bearbeitungsfunktionen für Bilder vorhanden. Webinhalte lassen sich so ohne Programmierkenntnisse erstellen. Durch die Integration von Dokumenten aus EEX kann dadurch ein weiterer Zugang für den Bereich DM und die Bereitstellung von Inhalten über das Webportal genutzt werden. Neue Inhalte können durch einfaches Drag'n'Drop in den Webbrowser eingestellt werden. Die Authentifizierung erfolgt über die vorhandene EEX-Benutzerverwaltung. Abbildung 3.17 zeigt die ELC-Architektur.

Für die Integration in andere Portale stehen Schnittstellen zu MOSS und zum *SAP ENTERPRISE-Portal* zur Verfügung. Im Fall von MOSS werden hierüber wie bei ELO-professional und DocuWare5 die Funktionen von MOSS um die Archivierungsfunktionen durch EEX erweitert. Der Zugriff auf die in MOSS gespeicherten Daten erfolgt transparent durch den Austausch der Daten gegen Verweise in EEX-Archive. Zusätzlich wird die Suchfunktion von MOSS erweitert, damit Inhalte in EEX-Archiven recherchiert und zugegriffen werden können, die nicht über MOSS gespeichert wurden.

Collaboration

ELC stellt auch die Collaboration-Funktionen bereit, mit denen eine personalisierte Anpassung der Portalansichten durch vordefinierte Layouts ermöglicht wird. ELC stellt über das Webportal Collaboration-Funktionen wie Diskussionsforen, Terminkalender und einen E-Mail-Client bereit. Die Funktionen decken damit allerdings nur einen kleinen Teil für Collaboration ab. Durch die Integration von Terminkalendern und dem

²⁰EASY Logistics Center

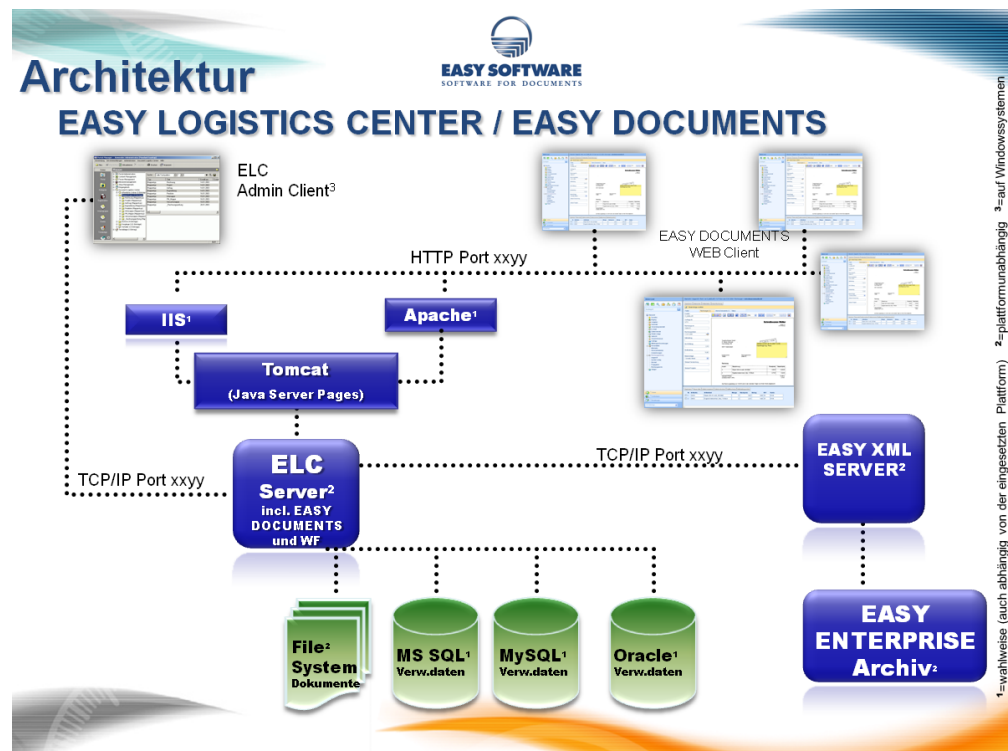


Abbildung 3.17.: ELC-Architektur [EAS07]

E-Mail-Client wird der Bereich der Koordination unterstützt. Weitergehende Funktionen wie gemeinsame Arbeitsbereiche oder Instant-Messaging sind jedoch nicht vorhanden. Damit bietet die ELC lediglich Basisfunktionen in diesem Bereich.

Workflow

Der Bereich WF wird in EEX durch die Adhoc-Workflow-Komponente in EEX und die Workflow-Funktionen aus dem ELC abgedeckt. Die in Abbildung 3.18 dargestellten Adhoc-Workflows stehen ausschließlich über den *EASY Experience Client* zur Verfügung. Sie bieten die Möglichkeit, Workflows zu erstellen, um Dokumente zur Freigabe oder Kenntnissnahme zu versenden.

Bei der Erstellung des Workflow werden alle Dokumente ausgewählt, die versendet werden sollen. Dabei stehen die Optionen parallel, sequentiell, erster gewinnt bei paralleler Versendung, Archivieren von Dokumenten oder Versenden einer E-Mail, zur Verfügung. Stellvertreterregelungen steuern die Weiterleitungen an andere Benutzer zur Bearbeitung bei Abwesenheit. Bei Workflows, die zur Bearbeitung von Dokumenten verwendet werden, wird immer das ursprüngliche Dokument genutzt. Dadurch

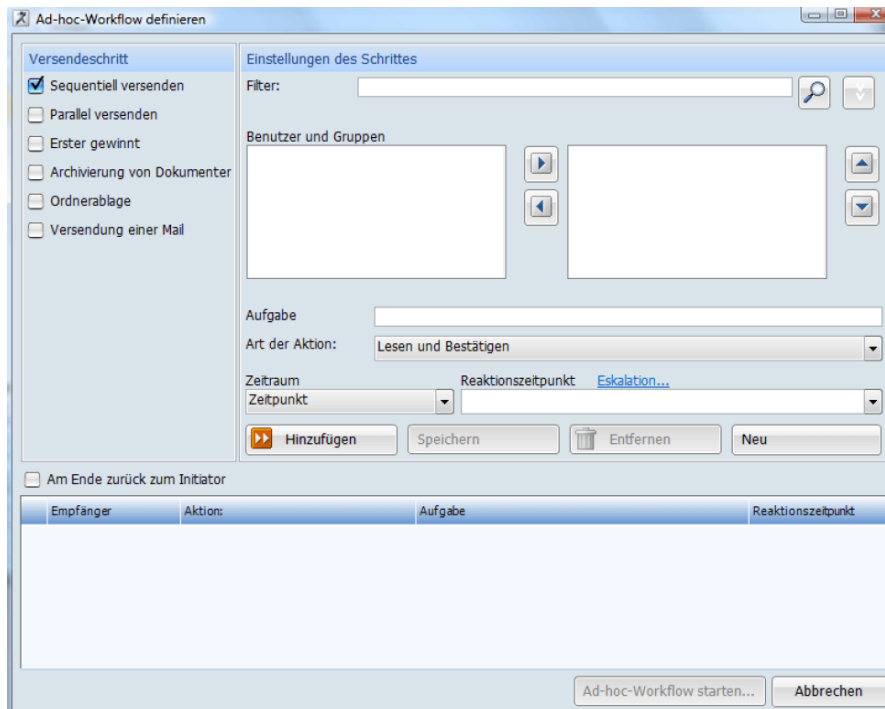


Abbildung 3.18.: EEX-Adhoc-Workflow definieren

muss die Anzahl der Bearbeitungen auf eine pro Prozess beschränkt werden, da ansonsten Inkonsistenzen bei der Bearbeitung entstehen. Bei Überschreiten von Bearbeitungsfristen können über eine Eskalationsfunktion alternative Prozessschritte definiert werden. Der Verlauf eines Prozesses kann über die Monitoring-Funktionen von allen beteiligten Teilnehmern nachvollzogen werden. Zusätzlich stehen bei der Ausführung eines Prozessschrittes die Funktionen Wiedervorlage für den Benutzer, Rückfrage, letzte Aufgabe wiederholen und Workflow abbrechen zur Verfügung.

Die *Standard-Workflows* werden über das ELC abgewickelt. Hierbei kommt auch eine andere Workflow-Engine, als bei den *Adhoc-Workflows* zum Einsatz. Dadurch können die *Standard-Workflows* auch nur über den ELC-Web-Client und nicht mit den normalen Rich-Clients erfolgen. Die Workflows in ELC basieren im Wesentlichen auf dem Prinzip der elektronischen Umlaufmappe.

Die Definition eines Workflows wird mit MS Visio realisiert, das durch Erweiterungen um spezielle *Workflow-Shapes* für die Modellierung ergänzt wird. Diese werden dann in ELC für die Ausführung importiert. Ein mit den speziellen Visio-Shapes erstellter Beispielprozess ist in Abbildung 3.19 dargestellt.

Beginn und Ende dieser Workflows werden durch spezielle Start- bzw. Endknoten festgelegt. Dazwischen können sequenzielle oder parallel verschaltete *Aktionsknoten* und

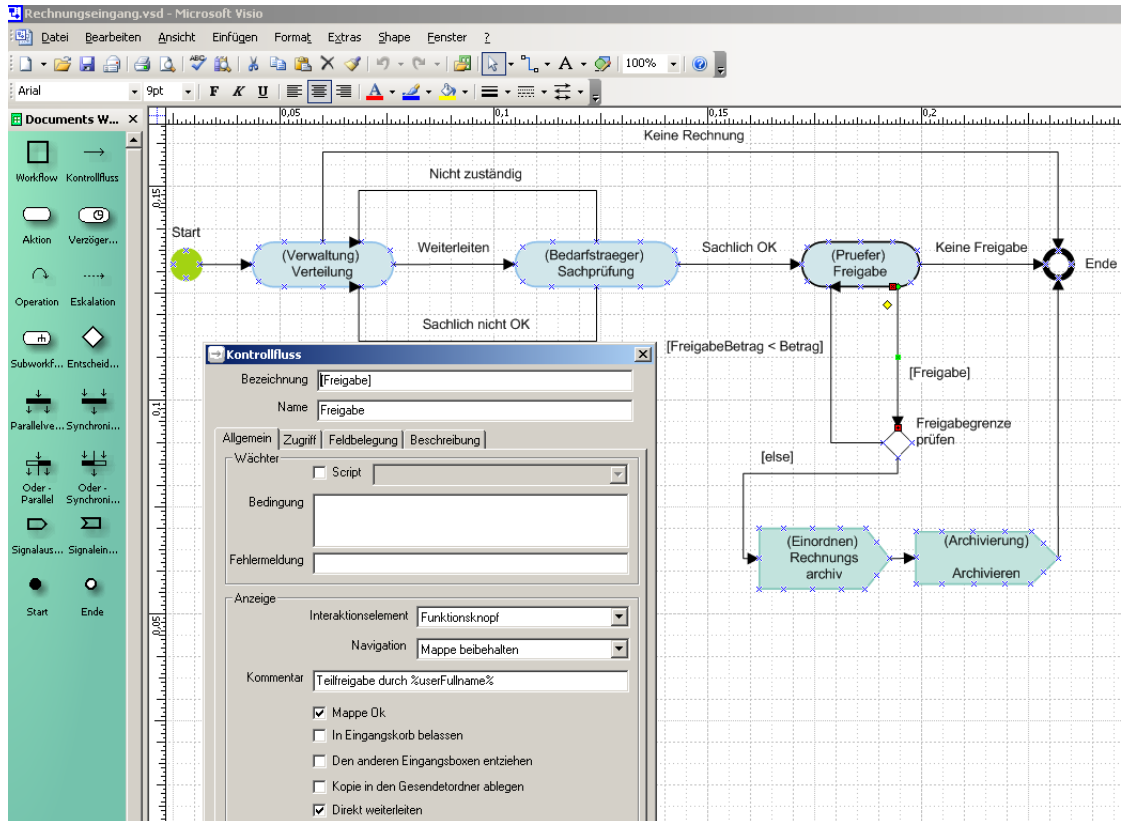


Abbildung 3.19.: ELC-Prozessmodellierung mit Visio

Entscheidungsknoten verwendet werden. Aktionsknoten können verschiedene Bearbeitungsschritte, wie Aufgaben, Eskalationen oder Skripte für das Portal beinhalten. Bei der Ausführung können Daten in einem XML-Format an einem Speicherort in EEX abgelegt werden oder Funktionen mit Javascript aufgerufen werden.

Store

Der Bereich Store wird in EEX über *Storageprovider* verwaltet. Die *Storageprovider* stellen eine Schnittstelle zwischen EEX-Kernel und dem jeweiligen Speichersystem dar. Unterschieden wird zwischen relationalen Datenbanken, Volltextdatenbanken und Datenspeichern (Dateisystemspeicher, optischen Speichergeräten, Bandlaufwerken oder CAS-Systemen). Die *Storageprovider* können jeweils auch mehrere Speichersysteme der gleichen Art verwalten. Sie stellen damit physische Speicherbereiche zur Verfügung, die über die Administrationskonsole verwaltet werden.

Die Speicherung von Inhalten erfolgt über *Pools*, die als logischer Speicherort dienen. Um bestimmte Dokumente in einem Pool ablegen zu können, werden hierzu die Dokumenten- und Recordschemata mit den Pools verknüpft. Dadurch wird sichergestellt, dass nur Dokumente mit einem freigegebenen Schema in einem Pool gespeichert werden dürfen. Pools werden durch einen *Storageprovider* mit den physischen Speicherorten verknüpft. Dabei können für die einzelnen Attribute der Dokumentenschemata verschiedene *Storageprovider* verwendet werden. Somit ist es möglich, eine physikalische Trennung bei der Speicherung der Dokumente zu realisieren. Anforderungen an Speicherpreise, Zugriffszeiten o. ä. lassen sich dadurch besser erfüllen. *Storageprovider* können für Blob-Daten, neue Versionen, Notizen und Anmerkungen, Volltextindizierung und die Dokumenten-Attribute getrennt angegeben werden. Zusätzlich existiert ein spezieller *SingleInstanceStorageprovider*, der sicherstellt, dass Dokumente nur einmal physikalisch gespeichert werden.

Für den Zugriff auf die Daten stellen die Store-Komponenten eine eigene Anfragesprache EQL²¹ bereit. EQL basiert im Wesentlichen auf einer eingeschränkten SQL-Syntax mit Erweiterungen für den Zugriff auf die Volltextdatenbanken. Damit lassen sich vordefinierte Anfragen für die Recherche oder spezielle Suchanfragen auf die Inhalte formulieren.

Preserve

Der Bereich Archivierung wird über dieselben Komponenten wie im Bereich Store realisiert. Hier kommt ein weiterer *Storageprovider* hinzu, der *ArchiveStorageprovider*. Dieser unterscheidet sich von den eben beschriebenen dadurch, dass dieser alle Daten eines Dokuments zusammen abspeichert, um ein Dokument mit all seinen Daten zu archivieren. Ausgenommen sind hier lediglich die Indexdaten, die nur von den Suchkomponenten verwendet werden.

Deliver

Im Bereich Deliver steht für die Transformation von Archivdaten *EASY Report* und *Automation* zur Verfügung. *EASY Report* ist ein Reportgenerator mit dem Layouts für die Archivdaten definiert werden und für den Druck oder den PDF-Export aufbereitet werden können. Damit kann auch COLD für das Output-Management bereitgestellt werden. Mit *EASY Automation* können die Serverprozesse gesteuert und automatisiert werden, um z. B. Daten zu im- oder exportieren oder in andere Formate zu transformieren und die Daten mit Hilfe von regulären Ausdrücken auf Korrektheit zu überprüfen.

²¹EASY Query Language

Außerdem lassen sich mit EQL Anfragen formulieren, die mit Suchmasken verknüpft werden können, um z. B. die Zahl angezeigter Dokumente zu reduzieren. Automatische Kompressionen bzw. Datenreduktion kann über den *SingleInstanceStorageprovider* realisiert werden. Dokumente können in einem ZIP-Archiv gespeichert oder importiert werden. Das Packen und Entpacken erfolgt transparent auf dem Server. Die Syndication-Funktionen werden über die die Funktionen im Store-Bereich, sowie der beschriebenen Transformationen abgedeckt. So ist es im Store-Bereich auch möglich, Daten physikalisch mit mehrfachen Kopien vorzuhalten. Transformationen lassen sich beispielsweise auf die Kopien der Daten anwenden, um sie für andere Anwendungen optimiert anzubieten.

Im Bereich der Sicherheitstechnologien lässt sich die komplette Kommunikation zwischen den Server- und Client-Komponenten in eine PKI integrieren. Außerdem können Dokumente mit digitalen Signaturen versehen werden. Weitere interne Verschlüsselungen für die Daten sind ebenfalls möglich. Je nach Sicherheitsstufe geht dies allerdings zu Lasten der Suchfunktionen, da keine verschlüsselten Daten in die Suchindexbereiche übernommen werden.

3.7. Funktionaler Vergleich

Nach den Darstellungen der CMS und DMS werden die Systeme nun abschließend anhand der aus Kapitel 2 identifizierten Funktionen und der Evaluierungskriterien am Anfang dieses Kapitels funktional verglichen. Dabei zeigen die Ergebnisse aus Tabelle 3.2, dass kein System den vollen Umfang von ECM abdecken kann. Weiterhin wird deutlich, dass sich die Systeme grob in die zwei Kategorien CMS und DMS unterteilen lassen. Alfresco und MOSS erfüllen vor allem Anforderungen an die Bereiche WCM, Collaboration und DM und lassen sich daher in die Kategorie CMS einordnen. ELO-professional, DocuWare ENTERPRISE und EEX decken die Bereiche DM, Preserve und Capture ab und lassen sich daher in die Kategorie DMS einordnen.

Der Bereich Capture wird von DMS-Anbietern wie ELO, DocuWare oder EASY gut abgedeckt. Bei diesen Systemen sind sowohl Funktionen für die Erfassung und Verarbeitung von NCI-, als auch CI-Dokumenten vorhanden. Außerdem besitzen sie umfangreiche Funktionen für die Formularverarbeitung und die Erfassung von COLD-Daten. Bei den CMS hingegen werden in diesem Bereich, bis auf die Integration von Office-Anwendungen, keine Funktionen angeboten. Teilweise lassen sie sich von Produkten von Drittanbietern nachrüsten.

Der Bereich Dokumenten-Management wird hingegen von allen Anbietern für den Bereich lebende Dokumente im Wesentlichen abgedeckt. Hierbei ist festzustellen, dass die DMS die Möglichkeiten der Nutzung über mehrere Clients anbieten, in die auch die Funktionen der anderen Manage-Komponenten integriert sind. Die CM-Systeme

bieten hierbei durch ihren webbasierten Ansatz sämtliche Funktionen über den Browser an.

Der Bereich Records-Management hat vor allem in den letzten Jahren durch eine Vielzahl neuer gesetzlicher Anforderungen an Bedeutung gewonnen. Dieser Bereich wird bei den DMS vor allem durch die Kombination von Funktionen aus den Bereichen DM und Preserve realisiert. Die umfassendsten Funktionen bietet hierbei EEX. Hier sind die Funktionen aus dem RM-Bereich direkt mit der Archivierung verknüpft und erlauben somit eine konsistente, bis auf die Ebene der physischen Speicherung, durchgängige Umsetzung. Ebenso sind umfassende Einstellungen bzgl. der Speicherung, Berechtigungen auf Inhalte und der expliziten Trennung von logischen und physischen Speicherorten bis auf einzelne Elemente von Records und Dokumenten möglich. DocuWare bietet dabei ebenso eine durchgängige Umsetzung der RM-Funktionen bis auf die physische Speicherebene, besitzt jedoch in der Umsetzung Einschränkungen. Beispielsweise fehlt eine getrennte Speicherung von Dokumenten und Records nach einzelnen Inhaltskomponenten. Bei ELOprofessional sind die Funktionen in diesem Bereich weniger konsequent integriert. So lassen sich z. B. Anforderungen an die Speicherung von Records auf unterschiedliche Speichersysteme und -medien nur sehr begrenzt realisieren.

Die beiden untersuchten CMS können in diesem Bereich zwar grundlegende Funktionen erfüllen. Durch die begrenzten bzw. fehlenden Archivierungskomponenten gestaltet sich eine durchgängige Umsetzung von RM-Konzepten jedoch schwierig und inkonsistent. Da die Archivierungskomponenten von externen Systemen bereitgestellt werden müssen, ergeben sich dadurch aufwendigere Umsetzungen, da Einstellungen in verschiedenen Systemen getrennt umgesetzt und konsistent gehalten werden müssen.

Beim WCM stellt sich die Situation umgekehrt dar. Hier können die CMS von Alfresco und Microsoft ausgereifte Funktionen vorweisen, während die DMS-Anbieter ELO und DocuWare hier bislang nur sehr rudimentäre Funktionen bzw. keine Funktionen bereitstellen. Bei EEX sind hier zwar umfangreichere Funktionen vorhanden, jedoch ist dieser Bereich wenig in die anderen Komponenten integriert. Außerdem kommen stark veränderte Bedienkonzepte zum Einsatz.

Da der Collaboration-Bereich typischerweise auf Web-Technologien aufsetzt und mit Hilfe von Portalen Funktionen zur Zusammenarbeit, gemeinsame Arbeitsbereiche oder Terminkalender, usw. bereitgestellt werden, setzt sich die Unterstützung für diesen Bereich wie bei den WCM-Komponenten fort. Bisher sind bei den klassischen DMS-Anbietern kaum Ansätze erkennbar, um diesen Bereich besser zu unterstützen. Die Umsetzung bei MOSS kann hier als umfangreich betrachtet werden. Es werden bereits viele Funktionen direkt abgedeckt und mitgeliefert. Darüber hinaus existieren eine Vielzahl von Anwendungen, die sich direkt in das MOSS-Portal integrieren lassen.

Der Bereich Workflow/BPM ist bei allen Systemen auf einfache und kurze dokumentorientierte Workflows innerhalb der Systeme beschränkt. Die zugrunde liegenden Konzepte und Funktionen unterscheiden sich hierbei sehr stark. Die Funktionen sind, teilweise deutlich erkennbar, nicht ausgereift, um sie unternehmensweit einsetzen zu können. Als Beispiel ist hier DocuWare zu nennen, bei dem Zusammenhänge einzelner Prozessschritte, durch den Ansatz Workflows über Stempel zu realisieren, nicht ersichtlich werden. Bei EEX ist die inkonsistente Bereitstellung von Workflow-Funktionen zu nennen. So können Adhoc-Workflows nur über einen der angebotenen Clients genutzt werden. Production-Workflows stehen nur über den Browser und das ELC zur Verfügung. Außerdem kommen hierfür verschiedene Workflow-Engines zum Einsatz. ELO stellt hierzu einen ausgereiften graphischen Prozessmodellierungseeditor bereit, mit dem auch etwas größere und komplexere Workflows realisiert werden können. Bei Alfresco und MOSS sind Workflows außerhalb der wenigen mitgelieferten Vorlagen i. d. R. mit Programmieraufwand verbunden, was meist zu aufwendigen und fehlerträchtigen Prozessen führt, die nur schwer handhab- und wartbar sind.

Den Bereich BPM hingegen kann keines der untersuchten Systeme zufriedenstellend umsetzen, da bei keinem der Systeme umfangreiche Prozesse ohne weiteres realisiert werden können. Die Integration von Anwendungen zur Ausführung von Aktivitäten ist nicht vorgesehen.

Der Bereich Store ist bei den DMS sehr eng in die DM- und Archivierungskomponenten eingebunden. Bei den CMS sind sie eher als Dienste vorhanden, die die Manage-Komponenten für DM und WCM nutzen. Bei den DMS lassen sich darüber hinaus weitere Speichersysteme wie Bandlaufwerke integrieren. Bei Alfresco steht hierbei nur die relationale Datenbank und das Dateisystem zur Speicherung zur Verfügung. MOSS beschränkt sich komplett auf die Speicherung in einer relationalen Datenbank für alle Inhalte.

Der Bereich Preserve wird wie weiter oben erwähnt von Alfresco und MOSS bisher kaum oder überhaupt nicht adressiert. Beide Anbieter setzen in diesem Bereich auf Drittanbieter, die sich in deren Repositories einklinken und die Daten in Archive extrahieren. Bei ELO gestaltet sich die Archivierung derart, dass neben den verwalteten Dokumenten noch eine Reihe von ERP-Systemen angebunden werden können. Eine detaillierte Unterscheidung nach Inhalten, verschiedenen Speichermedien oder HSM-Funktionen sind hier nicht oder nur sehr rudimentär vorhanden. DocuWare und EASY gehen hier einen Schritt weiter und erlauben detaillierte Zuordnungen zu verschiedenen Archivmedien. Außerdem wird eine physische Trennung der Inhalte bei der Archivierung ermöglicht. Bei EASY sind die Funktionen in diesem Bereich am umfangreichsten, da die Archivierung über mehrere Abstraktionsebenen in logische und verschiedene physische Speichermedien erfolgen kann. Weitergehend werden automatisierte Migrationen auf andere Speichermedien und Archivgeräte im Hintergrund unterstützt.

Der Bereich Deliver ist bei keinem der Systeme explizit durch eigene Komponenten realisiert. Die Funktionen werden je nach System und Architektur hierbei von unterschiedlichen Komponenten bereitgestellt. Bei den CMS sind die Funktionen zur Transformation und der Bereitstellung von Inhalten vor allem auf die webbasierte Nutzung ausgerichtet. Die DMS bieten hierbei zusätzlich umfangreichere Export- und Aufbereitungsfunktionen, um Daten an andere Systeme zur Verarbeitung weiterzureichen.

Die Ergebnisse der Evaluation sind in Tabelle 3.2 aufgeführt. Dabei sind die Bereiche wie in den eingangs beschriebenen Evaluierungskriterien in die verschiedenen Funktionsbereiche gegliedert. Die Erfüllung bzw. Nichterfüllung der jeweiligen Funktionen werden in der Tabelle durch die folgenden Symbole dargestellt. Mit einem vollen Punkt ● wird eine gute Umsetzung der Funktion, eine nicht zufriedenstellende oder unvollständige Umsetzung wird mit ◐ dargestellt. Nicht vorhandene oder nur durch Drittanbieter abgedeckte Funktionen bzw. unzureichend abgedeckte Funktionen sind mit einem Kreis ○ gekennzeichnet.

3.8. Zusammenfassung

In diesem Kapitel wurde eine Auswahl an CM- und DM-Systemen aus einer großen Anzahl am Markt befindlicher Systeme getroffen. Die Systeme wurden vorgestellt und hinsichtlich ihres Funktionsumfangs auf die Erfüllung der im ECM-Konzept gestellten Anforderungen untersucht. Bei der Evaluierung wurde deutlich, dass kein System den vollen Umfang von ECM abdecken kann. Weiterhin hat sich gezeigt, dass die Systeme erwartungsgemäß in den Funktionsbereichen Stärken aufweisen, aus denen sie ursprünglich entstanden sind. So setzen die CMS-Vertreter auf die Verwaltung von elektronischen Inhalten und versuchen über Web-Technologien neben der Erstellung von Webinhalten, Portalfunktionen, DM-Funktionen und Collaboration-Funktionen bereitzustellen. Sie verzichten dabei sowohl auf die Integration von Papierdokumenten, als auch auf die elektronische Archivierung. Diese Bereiche überlassen sie den Integrationen von Drittanbietern. Die DMS hingegen besitzen alle neben der Dokumentenverwaltung, Funktionen zur Erfassung von Papierdokumenten, COLD-Daten und zur elektronischen Langzeitarchivierung. Es sind Ansätze zur Erweiterung um Web-Technologien und Collaboration-Funktionen vorhanden. Diese sind jedoch bislang wenig ausgereift oder noch nicht implementiert. Große Unterschiede zeigten sich bei der Umsetzung des Workflow-Bereichs. Hier werden sehr unterschiedliche Ansätze gewählt, die von einer sehr einfachen Umsetzung von Dokumentenweiterleitungen mit dem Anbringen von elektronischen Stempeln auf Dokumenten (DocuWare) über graphische Editoren zur Prozessmodellierung (ELO, EASY und Alfresco) bis hin zu selbst zu erstellendem Programmcode, mit dem Prozesse erstellt werden (MOSS und Alfresco), reichen. Ebenso wurde deutlich, dass die Systeme keine klare Trennung der Funktionsbereiche aufweisen und häufig einzelne Funktionsbereiche durch verschiedene Komponenten in Kombination abgedeckt werden.

Tabelle 3.2.: Funktionaler Vergleich der evaluierten Systeme

	Funktion	Alfresco Labs 3.0	MOSS 2007	ELOprofessional 6	DocuWare 5	Easy Enterprise.x 3.0
Capture	NCI-Erfassung	○	○	●	●	●
	CI-Erfassung	⊙	⊙	●	●	●
	Formulärerfassung	○	○	⊙	⊙	●
	COLD / ERM	○	○	●	●	●
	Erkennungstechnologien	○	○	●	⊙	●
	Indexierung	⊙	⊙	●	●	●
Dokumenten- Management	Dokumentenverwaltung	⊙	⊙	●	●	●
	Check-In/Check-Out, Versionierung	●	●	●	●	●
	Web-Client	●	●	⊙	⊙	⊙
	Desktop-Client	○	○	●	●	●
	Suche / Recherche	⊙	●	⊙	●	●
	Offlinenutzung	⊙	⊙	●	●	⊙
Records-Management	Erstellung von vorgegebenen Ablagestrukturen	⊙	⊙	⊙	●	●
	Dokumenten- und Recordschemata	⊙	⊙	⊙	●	●
	Aufbewahrungsrichtlinien	⊙	○	⊙	●	●
	Audit-Trails	●	⊙	⊙	●	●
	Berechtigungssystem	⊙	⊙	⊙	⊙	●
	Thesaurus	●	●	●	●	●
	Richtlinien Bereichsübergreifend nutzbar	○	⊙	○	⊙	⊙
	Erfüllung rechtlicher Anforderungen	⊙	○	⊙	●	●
WCM	Webinhalte erstellen	●	●	○	○	⊙
	Webinhalte verwalten und veröffentlichen	●	●	○	○	●
	Bereitstellung eines Portals	⊙	●	○	○	⊙
	Integration in Portale	●	○	●	●	●
	Nutzung von Inhalten aus DM-Bereich	●	●	○	⊙	●
Collab- oration	Zusammenarbeit	⊙	●	○	○	○
	Informationssammlungen (Knowledgebase, Skill-Datenbanken, Foren)	●	●	○	○	⊙
	Groupware-Funktionen / Integration in Groupware-Anwendungen	⊙	●	○	○	○
	Kommunikationsfunktionen (Instant Messaging, Videokonferenzen)	⊙	●	⊙	○	○
Workflow und BPM	Adhoc-Workflows	⊙	⊙	●	○	⊙
	Production-Workflows	●	●	●	⊙	⊙
	BPM	○	○	○	○	○
	Integration in WCM und DM	●	●	⊙	⊙	⊙
	Graphische Modellierung von Workflows	⊙	⊙	●	○	●
Store	Library Services	●	●	⊙	●	●
	Repositories	⊙	○	⊙	⊙	●
	Speichertechnologien	○	○	⊙	●	●
Preserve	Archivierung	○	○	⊙	●	●
	HSM	○	○	○	⊙	●
	Migration	○	○	○	●	●
	Umsetzung von RM-Richtlinien	○	○	⊙	●	●
Deliver	Transformationstechnologien	●	⊙	●	●	●
	Sicherheitstechnologien	●	●	⊙	⊙	⊙
	Distribution	⊙	⊙	○	⊙	⊙
	Output-Management	○	○	⊙	⊙	●

4. Anforderungen an die Integration von DMS, CMS und PMS

Wie aus der Evaluation in Kapitel 3 hervorgegangen ist, kann keines der untersuchten Systeme alle im ECM-Konzept beschriebenen Bereiche vollständig abdecken. Durch eine Integration von Systemen aus den ermittelten Kategorien der CMS und DMS lassen sich jedoch Lücken in den jeweiligen Systemen gegenseitig schließen. Beispielsweise können fehlende Funktionen in MOSS durch eine Integration mit ELOprofessional, DocuWare oder EEX ausgeglichen werden. Im Gegenzug erhalten die drei DMS-Systeme eine Aufwertung in den Bereichen WCM und Collaboration durch die MOSS-Erweiterung. Da eine Integration mit MOSS bereits für alle untersuchten DMS existiert, wird auf eine Betrachtung der Aspekte der Integration dieser Systeme verzichtet.

Im Bereich Workflow besitzen allerdings alle untersuchten Systeme Schwächen. So bieten sie Funktionen, um einfache, vom Benutzer initiierte Adhoc-Workflows durchzuführen, mit denen das Versenden von Dokumenten an weitere Benutzer zur Überarbeitung oder Durchsicht realisiert werden kann. Production-Workflows mit vordefinierten Prozessen lassen sich ebenfalls noch mit den Systemen umsetzen, sofern die Prozesse nur Inhalte einbeziehen, die in den Systemen selbst gespeichert sind. Die Integration von Anwendungen in die Prozessausführung ist nur sehr eingeschränkt möglich. Auch die Erstellung komplexer Workflows gestaltet sich mit den vorgestellten Systemen teilweise sehr unübersichtlich, sodass der Gesamtüberblick über einen Prozess und damit die Prüfung auf eine korrekte Umsetzung kaum noch gewährleistet werden kann. BPM-Funktionen werden von keinem der untersuchten Systeme zufriedenstellend bereitgestellt.

Aus diesen Gründen werden in diesem Kapitel Anforderungen an eine Integration eines DMS mit einem PMS vorgestellt. Es wird zunächst eine Einführung in PMS gegeben, mit der die Funktionsweise dieser Systeme näher vorgestellt wird. Darauf aufbauend wird diskutiert, welchen Stellenwert Dokumente in Geschäftsprozessen besitzen und wie sich die Synergieeffekte der unterschiedlichen Sichtweisen von PMS und DMS nutzen lassen, um Geschäftsprozesse im PMS abzubilden und auszuführen. Um alle Bereiche des ECM-Konzepts einzubeziehen, werden die Funktionsbereiche anhand des ECM-Modells abgegrenzt. Damit wird definiert welche Funktionen DMS, CMS bzw. PMS bereitstellen.

Daran schließt sich eine Diskussion an, auf welcher Basis die Systeme integriert werden können. Hierzu werden Auszüge aus der Records-Management-Spezifikation MoReq2 vorgestellt, um zu erfüllende Anforderungen im Umgang und in der Aufbewahrung von geschäftsrelevanten Informationen einzubeziehen. Die Anforderungen werden daraufhin um spezifische Anforderungen bzgl. der Integration der Systeme ergänzt.

4.1. Prozess-Management-Systeme

Prozess-Management-Systeme unterstützen die Abwicklung von Geschäftsprozessen gemäß definierter Arbeitsabläufe. Sie ermöglichen, Geschäftsprozesse abzubilden, auszuführen und den Prozessablauf zu dokumentieren [DAH05, AH04].

Geschäftsprozesse sind ständigen Änderungen unterworfen, etwa durch Reaktionen auf Marktveränderungen oder durch die Erfüllung neuer rechtlicher Vorschriften. Die Umsetzung dieser Änderungen setzt voraus, dass die Erfassung und Analyse von Prozessen von den Systemen unterstützt wird. Außerdem sollen Geschäftsprozesse, die zur Erstellung von Leistungen entlang der Wertschöpfungskette dienen, kontinuierlich optimiert werden [SJW05].

PMS versuchen daher die Abläufe von Geschäftsprozessen ganzheitlich zu unterstützen. Das PMS übernimmt dazu die Durchführung der Abläufe und koordiniert diese. Zur Bearbeitung einzelner Aufgaben während der Durchführung von Prozessen integrieren sie dazu Anwendungskomponenten in die Prozessausführung. Dadurch sollen auch die Benutzer prozessorientiert unterstützt und der Ablauf von Prozessen überwacht werden [JBS97].

PMS versuchen dies zu erreichen, indem sie eine Trennung von Anwendungs- und Prozesslogik vornehmen (s. Abbildung 4.1). Abläufe werden dazu anhand der konkreten Anforderungen aus den Geschäftsprozessen heraus mit Hilfe der (graphischen) Modellierungswerkzeuge des PMS abgebildet. Prozesse müssen daher nicht mehr an den Funktionen in verschiedenen Anwendungen zur Bearbeitung einzelner Prozessschritte ausgerichtet werden. Dazu werden mit dem PMS die verschiedenen Aspekte eines Geschäftsprozesses, wie Kontroll- und Datenflüsse oder Bearbeiterzuordnungen unterstützt. Weiterhin integriert ein PMS die zur Bearbeitung erforderlichen Anwendungen als Programmbausteine und ermöglicht die Verknüpfung dieser Bausteine mit den Aktivitäten innerhalb des Prozesses. Sie werden bei ihrer Ausführung mit Parametern aufgerufen, um einzelne Aufgaben zu bearbeiten. Nach Beendigung der Bearbeitung liefert die Anwendung Daten an die an den Aktivitäten hinterlegten Parameter zurück an das PMS.

Die zur Durchführung benötigten Informationen in Form von Daten, Dokumenten, Formularen usw. werden durch die Prozess-Engine in einem modellierten Datenfluss

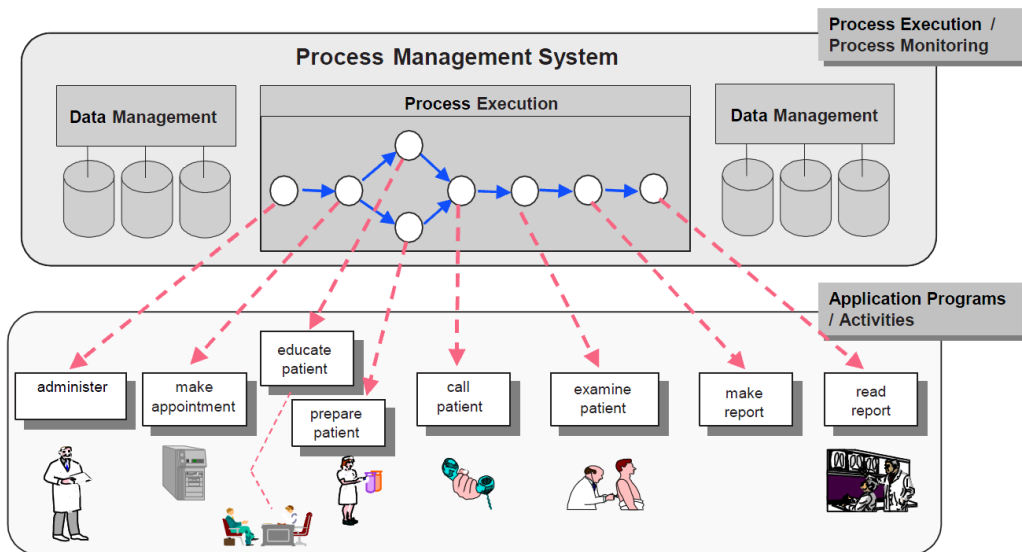


Abbildung 4.1.: Trennung von Prozesslogik und Anwendungscode [Rin04]

transportiert und den Benutzern bzw. Anwendungen bei der Ausführung bereitgestellt, wodurch die Benutzer aktiv bei der Bewältigung ihrer Aufgaben unterstützt werden können [BDS⁺93, Mül05].

Damit bleiben die Zusammenhänge zwischen Teilaufgaben und der Gesamtaufgabe durch den Prozess erhalten. Die Ausführung der Prozesse wird durch eine Prozess-Engine gesteuert. Sie stellt den korrekten Ablauf und die Reihenfolge der Bearbeitungsschritte sicher. Dabei lassen sich verschiedene Abläufe, wie z. B. die parallele Bearbeitung von Aufgaben oder Entscheidungen in Prozessen darstellen. Außerdem verteilt das PMS die zur Bearbeitung anstehenden Aufgaben an die an den Prozessschritten hinterlegten Benutzer. Die Aufgaben werden den Benutzern dabei über eine Arbeitsliste vom PMS bereitgestellt. Um die Auswahl der Bearbeiter nach unterschiedlichen Kriterien vorzunehmen, können typischerweise zusätzliche Angaben mit den modellierten Prozessschritten verknüpft werden. Dadurch kann die Bearbeitung einer Aufgabe durch einen Benutzer, der die entsprechenden Fähigkeiten oder Berechtigungen besitzt, sichergestellt werden. Zur Laufzeit kann das PMS weitere Aspekte einbeziehen, wie z. B. die Anzahl der in der Arbeitsliste vorhandenen Einträge, um die Aufgaben gleichmäßig auf alle Bearbeiter zu verteilen.

4.2. Dokumente in Geschäftsprozessen

Bis zu 90% aller gespeicherten Informationen liegen in Unternehmen in Form von Dokumenten oder Content vor [Jen05]. Wichtige Unterlagen wie Verträge, Rechnun-

gen, Urkunden, technische Beschreibungen, E-Mails, Korrespondenz und damit ein beträchtlicher Teil des Wissens eines Unternehmens sind in Form verschiedener Dokumente vorhanden, die von Mitarbeitern erstellt, bearbeitet oder eingesehen werden. Neben der reinen Informationsspeicherung und -versorgung besitzen Dokumente aber auch gleichzeitig Steuerungsfunktionen bei der Durchführung von Prozessen. Die in Dokumenten enthaltenen Informationen bestimmen häufig direkt den Kontrollfluss an die entsprechenden Bearbeiter [Sei06]. Oft ist der Eingang eines Dokuments auch direkt mit der Initiierung eines Prozesses verbunden. Aktivitäten können jedoch nur ausgeführt werden, wenn die in Dokumenten gespeicherten Informationen dem Bearbeiter vorliegen. Zum Beispiel wird bei einer Eingangsrechnung durch den Inhalt des Rechnungsdokuments festgelegt, welcher Abteilung oder Kostenstelle die Rechnung zuzuordnen ist und welche Mitarbeiter für die inhaltliche Kontrolle der aufgeführten Positionen zuständig sind. Darüber hinaus legen die Positionen und die Höhe der Beträge zusätzlich fest, welche Mitarbeiter die erforderlichen Befugnisse zur Genehmigung von Zahlungen besitzen.

Anhand von Studien wird ebenfalls deutlich, dass Dokumente vielfach zentrale Bestandteile in Prozessen darstellen. So wurde in einer XEROX-Studie ermittelt, dass 92% aller Kanten eines 263 Aktivitäten umfassenden Kerngeschäftsprozesses durch Dokumente determiniert wurden [Rau95, MS91]. Auch aus der in Kapitel 2 erwähnten Studie zur Verbreitung elektronischer Rechnungen geht hervor, dass nach wie vor viele dieser Dokumente auch noch in Papierform vorliegen [May09]. Damit fixieren Dokumente einen bestimmten Informationsstand zu einem bestimmten Zeitpunkt für Personen und organisatorische Stellen, und sie dienen der Kommunikation zwischen in- und externen organisatorischen Einheiten im Rahmen von Geschäftsprozessen [Krä95].

4.3. Unterschiedliche Sichtweisen von DMS und PMS

Die primäre Aufgabe von DM-Systemen besteht darin, alle Aspekte des Dokumentenlebenszyklus abzudecken. Zusätzlich werden Anwendungen integriert, um Dokumente im DMS zu speichern oder auf Dokumente zugreifen zu können. Hierzu bietet ein DMS mit seiner daten- bzw. dokumentenorientierten Sichtweise die Möglichkeit, Daten und Dokumente dem Benutzer anhand der in Dokumenten gespeicherten inhaltlichen Kriterien zu präsentieren. Durch die Archivierungsfunktionen und die Integration mit weiteren Systemen, die ihre Daten im DMS archivieren, wird ein Großteil der Daten eines Unternehmens im DMS-Repository abgespeichert. Durch die bereitgestellten Funktionen zur Verwaltung und Recherche ermöglicht ein DMS eine umfassende Übersicht über im Unternehmen gespeicherte Informationen auch über mehrere Anwendungen hinweg. Historien der Änderungen von Daten und Dokumenten lassen sich über die Versionierungsfunktionen nachvollziehen. Diese sind durch die integrierten Archivierungsfunktionen vollständig über den gesamten Lebenszyklus der Dokumente darstellbar.

Die Zusammenhänge von gespeicherten Informationen und ihrer Entstehung oder Zuordnung zu Prozessen, aus denen sie ggf. entstanden sind, ist durch die datenorientierte Sichtweise eines DMS jedoch nur schwer möglich. Durch diese Sichtweise wird auch nicht ersichtlich, welche Prozesse überhaupt vorhanden sind oder welche Aufgaben als nächstes anstehen. Darüber hinaus besitzen Dokumente häufig einen gewissen Prozesscharakter, indem sie sowohl als Informationsträger in Prozessen dienen, als das Auslösen von Prozessen verursachen können. Dieser Prozesscharakter wird jedoch durch die datenorientierte Sichtweise von DMS alleine nur unzureichend unterstützt [Sei00].

PMS dienen hingegen dazu, den Benutzern und Anwendungen alle für die Ausführung eines Prozessschritts relevanten Informationen und Daten zum Zeitpunkt der Bearbeitung zur Verfügung zu stellen. Diese Sichtweise von PMS geht damit vom Geschäftsprozess aus und zerlegt alle auszuführenden Aufgaben in kleine Teile zur Bearbeitung. Den Benutzern werden zur Bearbeitung einzelner Aufgaben nur die dafür notwendigen Informationen und Daten bereitgestellt. Dadurch wird jedoch eine umfassende Sicht auf vorhandene Informationen verwehrt.

Da Anwendungsprogramme nur für bestimmte Einzelaufgaben an den entsprechenden Stellen im Prozess aufgerufen werden, wird damit auch nicht mehr der volle Funktionsumfang, den eine Anwendung bereitstellt, dem Benutzer präsentiert, sondern nur noch die Teile, die für die Bearbeitung der konkreten Aufgabe relevant sind. Die dabei erzeugten und bereitzustellenden Daten werden durch das PMS gesteuert und nicht mehr von den beteiligten Anwendungen selbst weitergereicht. Die Speicherung und der Zugriff auf die Prozessdaten erfolgt zumeist ebenfalls prozessbezogen. Dies beschränkt jedoch die Möglichkeiten für Rechercheaufgaben ein, wenn nach bestimmten Daten oder Dokumenten gesucht werden soll, die in einem Prozess verwendet wurden. Beispielsweise kann dies eine Anfrage nach allen Rechnungen, die in einem bestimmten Zeitraum an einen Kunden versandt wurden sein.

Zusammenfassend kann festgehalten werden, dass sich ein PMS auf die Darstellung und Durchführung von Prozessen fokussiert, die Sicht auf gespeicherte Daten und die darin enthaltenen Informationen jedoch nur eingeschränkt oder nicht ermöglicht. Durch die Integration eines DMS mit einem PMS können die Anforderungen an Prozessunterstützung, die Verwaltung von unstrukturierten Inhalten und die korrekte Aufbewahrung erfüllt werden.

4.4. Abgrenzung der Funktionsbereiche

Um durch eine Integration der Systeme (DMS, CMS, PMS) den gesamten Umfang von ECM abzudecken, werden zunächst die Funktionsbereiche der Systeme abgegrenzt, die teilweise redundant in mehreren Systemen gleichermaßen vorhanden sind. Anhand des Dokumentenlebenszyklus können die einzelnen Stufen, die ein Dokument von der

Erstellung bis zur Vernichtung durchläuft, sowie eine Zuordnung zu den Aufgabenbereichen durchgeführt werden. Die Funktionen aus den einzelnen Bereichen beziehen sich dabei auf die untersuchten Systeme aus Kapitel 3.

Tabelle 4.1 fasst die Abgrenzung und Aufteilung der Funktionsbereiche auf die drei zu integrierenden Systeme zusammen.

Tabelle 4.1.: Abgrenzung von DMS, CMS und PMS

	DMS	CMS	PMS
Capture	●		
Dokumenten-Management	●	● ¹	
Records-Management	●		
Web-Content-Management		●	
Collaboration		●	
Adhoc-Workflows	●		
Production-Workflows und BPM			●
Store	●	● ³	● ³
Preserve	●		
Deliver	● ²	● ²	

¹ im Rahmen von WCM und Collaboration

² Bereitstellung von Diensten je nach Anwendungsbereich

³ temporäre Speicherung

Aufgrund der vorhandenen Funktionen eines DMS können die Bereiche Capture, DM, RM, Adhoc-Workflows, Store, Preserve und teilweise Deliver abgedeckt werden. Da die Bereiche Capture und DM zu den Kerngebieten eines DMS gehören und daher auch die Funktionen entsprechend umfangreich ausfallen, werden sie durch das DMS abgedeckt. Durch den Archivierungshintergrund, der den meisten DM-Systemen zugrunde liegt, sind die Funktionen aus den Bereichen RM und Preserve hier ebenfalls am umfangreichsten ausgeprägt. Für den Bereich Store, in dem lebende und noch zu bearbeitende Informationen abgespeichert werden, bietet sich daher ebenfalls das DMS für die Speicherung des Großteils der Inhalte an, da die erwähnten Funktionen direkten Zugriff auf das DMS-Repository besitzen. Der Bereich Workflow/BPM wird hierbei geteilt, da Adhoc-Workflows zum Versenden einzelner Dokumente über die DMS-Komponenten abgewickelt werden können. Adhoc-Workflows sollen keine wiederkehrenden Prozesse oder Geschäftsprozesse abbilden. Sie dienen lediglich der Benachrichtigung oder der spontanen Verteilung von Dokumenten. Außerdem können sie zum Einholen von Feedback zu Dokumenten an bestimmte Benutzer oder Gruppen gesendet werden. Da diese Workflows von jedem Benutzer erstellt und verwendet werden sollen, wäre hierfür eine Modellierung und Ausführung über das PMS zu aufwendig, da jeder Benutzer hierzu Zugriff und Kenntnisse über die Modellierungswerkzeuge des

PMS benötigen würde. Da der Bereich Deliver vielfältige Querschnittsfunktionen bereitstellt, werden von den DMS typischerweise nicht alle benötigten Funktionen angeboten, wie z. B. Funktionen zur Personalisierung von Portalansichten oder die Konvertierung in weboptimierte Anzeigeformate.

Die Bereiche WCM und Collaboration bieten sich vor allem für ein CMS an. DM-Funktionen werden dabei in die Collaboration- und WCM-Funktionen integriert, um Dokumente gemeinsam zu bearbeiten oder in die Webinhalte zu integrieren. Hier hat sich in Kapitel 3 gezeigt, dass CMS wie MOSS oder Alfresco in diesen Bereichen Stärken ggü. DMS aufweisen. Die Store-Komponenten dienen bei der Integration lediglich der temporären Speicherung von Inhalten oder stellen Inhalte als Kopie bereit. Der Bereich Store dient somit der Speicherung von aktuellen Inhalten in den Bereichen WCM und Collaboration. Für die Langzeitarchivierung werden diese Inhalte ins DMS-Repository übertragen. Durch die bereits vorhandenen Integrationen von CMS und DMS lassen sich die Inhalte nahtlos in das DMS-Repository integrieren.

Das PMS übernimmt schließlich die Bereiche Production-Workflow und BPM. In diesen Bereichen werden Prozesse abgebildet, optimiert und ausgeführt. Die für die Durchführung von Prozessen benötigten Anwendungen werden über das PMS integriert. Durch die Anbindung an das DMS-Repository werden zum einen die Inhalte aus den Bereichen DM, RM und WCM für Prozesse im PMS zugänglich gemacht, zum anderen können hierüber Daten aus dem PMS im DMS-Repository abgelegt werden. Dadurch werden sie den anderen Komponenten der genannten Bereiche zugänglich gemacht. Archivierungsfunktionen stehen somit auch für Inhalte aus Prozessen zur Verfügung. Da während der Ausführung von Prozessen Daten durch direkte Eingabe eines Benutzers oder Daten aus integrierten Anwendungen im Datenfluss weitergeleitet werden müssen, kommen dem PMS auch Teile aus dem Bereich Store zu. Dies bezieht sich auf die temporäre Speicherung von Daten, solange sich diese im Prozess in der Bearbeitung befinden, bis zur Archivierung im DMS.

4.5. Schnittstellen

Um die Systeme zu integrieren, ist es erforderlich, gemeinsame Schnittstellen für die Kommunikation der beteiligten Systeme zu verwenden. Dafür folgt eine kurze Betrachtung einiger Standards für CM bzw. DM-Systeme. Im Bereich der CMS und DMS konnte sich bislang jedoch noch keine einheitliche Schnittstelle durchsetzen, die von allen oder zumindest einem Großteil der Systeme unterstützt wird.

Erste Standardisierungsversuche wurden mit DMA¹ und ODMA² unternommen. Die DMA-Spezifikation sah vor, die Zugriffe auf DMS auf Serverseite mit systemübergrei-

¹Document Management Alliance

²Open Document Management API

fenden Interoperabilitätsschnittstellen auszustatten, um eine Nutzung ohne hersteller-spezifische APIs zu ermöglichen. Mit ODMA wurde versucht, eine einheitliche Schnittstelle für Anwendungen bereitzustellen, um Funktionen wie Recherche, Archivierung oder einen direkten Dokumentenzugriff DMS-unabhängig zu gestalten. Beide Ansätze können jedoch als gescheitert betrachtet werden, da es kaum praktische Umsetzungen dieser Standards in konkreten Produkten durch DMS-Hersteller gibt und die beiden Standards nicht mehr weiter voran getrieben werden. [Ass02, DMw08, GSMK04, ZGB05].

JCR

Mit JCR wird seit dem Jahr 2002 eine Spezifikation für eine Java-Plattform-API für Content-Repositories entwickelt, mit der ein einheitlicher Zugriff auf Contentobjekte beschrieben wird. Die verfolgten Ziele sind, vergleichbar mit dem DMA-Ansatz, einheitliche Schnittstellen für den Zugriff auf Content-Repositories bereitzustellen, um die Integration mit anderen Anwendungen unabhängig vom konkret genutzten DMS zu gestalten. Darüber hinaus geht die Spezifikation auf interne Aspekte, wie den Aufbau von Repositories ein.

Ein JCR-basiertes Repository stellt eine objektorientierte Datenbank bereit, die sowohl die von relationalen Datenbanken bekannten Datenstrukturen, Suchfunktionen, Transaktionen oder referentielle Integrität als auch ein Dateisystem mit hierarchischen Beziehungen, Zugriffskontrolle und Check-In/Check-Out-Funktionen umfasst. Darüber hinaus sind DMS-typische Funktionen zum Speichern von Dokumenten, Volltextindizes oder Versionierung vorhanden [Nue06, Nue09]. Im Gegensatz zu DMA besteht bei JCR von Herstellerseite bislang eine breitere Unterstützung. Konkrete Implementierungen in Produkten befinden sich bislang jedoch noch in der Entwicklungsphase.

CMIS

Neben der JCR-Spezifikation existiert mit CMIS eine plattformneutrale Web-Service-Schnittstellenspezifikation. Die Architektur (s. Abbildung 4.2) setzt dazu auf einen SOA-orientierten Ansatz, um plattformneutrale Schnittstellen zu definieren. CMIS verzichtet auf Vorgaben zum internen Aufbau des Repositories und beschreibt lediglich die Schnittstellen. Die Systemen sollen damit frei in ihren internen Implementierungen bleiben und lediglich eine gemeinsam nutzbare Schnittstelle implementieren, über die Anwendungen auf Inhalte in CMS und DMS zugreifen können. CMIS erfährt bislang eine ähnlich breite Unterstützung der Hersteller wie JCR und schließt auch Unternehmen wie Microsoft ein, die sich nicht an Java-Standards orientieren. Da die Spezifikation erst Ende 2009 in Version 1.0 fertiggestellt wurde, sind bislang lediglich prototypische Umsetzungen vorhanden [EIM08].

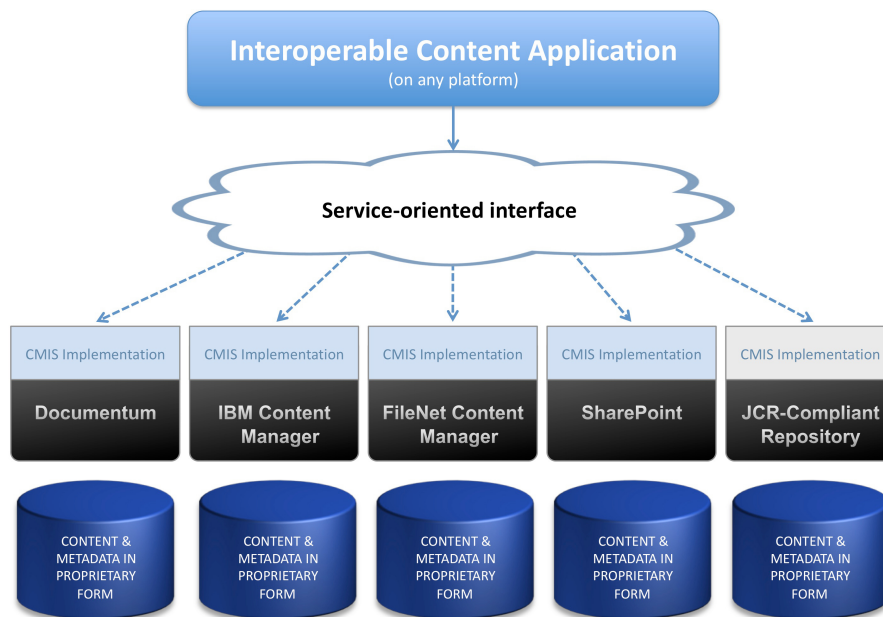


Abbildung 4.2.: CMIS-Architektur [EIM08]

Zum momentanen Zeitpunkt kann durch die in diesem Abschnitt beschriebenen Spezifikationen und deren Umsetzungen in konkreten Systemen keine für die Integration von DMS und PMS standardisierte Schnittstelle für den Zugriff auf das DMS-Repository vorausgesetzt werden. Hier muss auf die jeweiligen herstellereigenen APIs zurückgegriffen werden.

4.6. Records Management als gemeinsame Basis für DMS und PMS

Für die Integration von DMS und PMS wird eine gemeinsame Basis zur Definition und Verwendung der Informationsobjekte benötigt, die in Geschäftsprozesse integriert werden müssen bzw. die als Resultat aus der Durchführung von Geschäftsprozessen entstehen.

Wie in den Kapiteln 2 und 3 bereits angesprochen, ist der Übergang zwischen den Store- und Preserve-Bereichen bei DMS fließend. Ähnlich verhält es sich auch bei den Bereichen Dokumenten- und Records-Management. Auch hier sind die Übergänge zwischen Dokumenten und der Klassifikation als Records in den Systemen fließend. Der physische Speicherort bleibt den Verwaltungskomponenten durch die transparente Integration der Speichersysteme verborgen. Die Unterscheidung zwischen nicht mehr veränderbaren Records, und Dokumenten, die noch bearbeitet werden dürfen, wird

lediglich durch den Status geregelt. Mit dem Übergang zur Archivierung wird ein Dokument zu einem Record.

Bei der Integration soll das DMS-Repository Dokumente und Content speichern, da dieses über Komponenten zur Erfassung, Verwaltung und Archivierung verfügt. Außerdem sollen zu archivierende Daten aus dem PMS ebenfalls dort abgelegt werden, um zum einen den Zugriff auf die Inhalte für die anderen DMS-Komponenten zu ermöglichen, und zum anderen die Archivierungs- und RM-Funktionen zur Aufbewahrung zu nutzen. Die Definition entsprechender Aufbewahrungsrichtlinien erfolgt über die RM-Komponenten, mit denen festgelegt wird, welche Informationen wie lange aufzubewahren sind. Um dies sicherzustellen müssen dazu entsprechende Strukturen in Form von Records für die Daten definiert werden, die angeben, welche Informationen zu einem Geschäftsvorgang aufzubewahren sind.

Um diese Daten in die Prozesse des PMS zu integrieren, muss der strukturelle Aufbau bereits bei der Modellierung der Prozesse festgelegt werden. Die Kenntnis über den strukturellen Aufbau wird benötigt, damit das PMS möglichst viele Informationen über die Inhalte des DMS erhält. Diese können dann im Prozess als Entscheidungsparameter oder zur Auswahl eines zum Dokument passenden Prozesses genutzt werden. Beispielsweise können die zu einer gescannten Rechnung erhobenen Metadaten dazu verwendet werden, um festzustellen, welche Benutzer für die Prüfung zuständig sind, und diesen die Aufgabe in die Arbeitsliste zu stellen.

Umgekehrt sollen die während der Ausführung von Prozessen erhobenen bzw. erstellten Prozessdaten im DMS archiviert werden. Das PMS wird damit um Archivierungsfunktionen erweitert. Hierfür muss die Struktur der zu speichernden Daten bekannt sein, um mit Hilfe von Records-Management-Funktionen die Struktur von Records und damit die aufzubewahrenden Informationen festzulegen. Hierüber wird sichergestellt, dass genau die Informationen aufbewahrt werden, die aufbewahrt werden müssen – sei es durch gesetzliche Vorschriften oder auch unternehmensinterne Vorgaben [SMS08].

Um diese gemeinsamen Strukturen zu nutzen, werden im Folgenden die RM-Konzepte als gemeinsame Basis zur Definition von Schemata verwendet, die den Aufbau beschreiben. Zum Umfang gehört der logische Aufbau des Ordnungssystems, in dem die Informationen abgelegt werden, der strukturelle Aufbau von Klassifikationsschemata, die Dokumente und Records mit den zugehörigen Metadaten, Aufbewahrungsfristen, sowie die Festlegung der Zugriffs- und Bearbeitungsberechtigungen.

4.7. Anforderungen

Mit den aus den Kapiteln 2 und 3 ermittelten Grundlagen und Systemen, sowie den in diesem Kapitel vorgestellten Ansätzen werden nun die Anforderungen an die Konzep-

tion der Integration eines DMS mit einem PMS formuliert. Die Anforderungen lassen sich grob in die drei Bereiche der Integration auf Daten-, Organisations- und Clientebene einordnen. Die Datenintegration bezieht sich auf Aspekte, wie das PMS Daten aus dem Repository des DMS nutzen kann, und wie sich die Prozessdaten zur Archivierung und für den Zugriff über das DMS integrieren lassen. Da RM-Konzepte als gemeinsame Basis für die Integration dienen, wird dazu die MoReq2-Spezifikation näher betrachtet. MoReq2 dient dabei als neutrale Spezifikation, da sich die RM-Funktionen in den DMS-Implementierungen sehr stark unterscheiden und daher kaum allgemeingültige Aussagen getroffen werden können. Dabei werden Anforderungen aus dem Bereich Records- und Dokumentenmanagement ermittelt, die bei der späteren Konzeption zu beachten sind. Die Organisationsintegration dient der einheitlichen und konsistenten Verwaltung der Benutzer und der Rechtevergabe auf die Inhalte, so dass sichergestellt wird, dass die Daten bei der Durchführung von Prozessen durch den Benutzer zugegriffen werden können. Die Clientintegration betrachtet, wie sich die Clientkomponenten für die Benutzer, die die Systeme bedienen, vereinen lassen. Dem Benutzer soll dadurch eine ganzheitliche Sicht auf beide Systeme zur Verfügung gestellt werden, um die Vorteile der prozess- und datenorientierten Sicht zu realisieren.

4.7.1. MoReq2

MoReq2 wurde auf Basis eines Anforderungsdokuments des DLM-Forums³ im Jahre 2006 von der Europäischen Kommission als Projekt zur Erarbeitung einer Records-Management-Spezifikation ausgeschrieben. MoReq2 liefert eine detaillierte funktionale Anforderungsliste an elektronisches Records-Management und ist die einzige branchenunabhängige Spezifikation für den Aufbau und den Betrieb von DM- und RM-Systemen in Europa. MoReq2 bezieht außerdem weitere RM-Standards in die Spezifikation mit ein, wie z. B. den *ISO 15489-Standard* [ISO02]. MoReq2 ist eher ein allgemeiner Anforderungskatalog, an dem sich konkrete, individuelle Anforderungen orientieren können. Hierauf geht die Spezifikation bereits selbst ein und erlaubt bzw. ermutigt, eigene Anforderungen mit aufzunehmen oder darin beschriebene Anforderungen zu streichen, falls diese nicht erforderlich sind. Die vollständige Spezifikation findet sich in [Eur08a] und [Eur08b].

4.7.1.1. Klassifikationsschema

MoReq2 sieht grundsätzlich eine hierarchische Struktur für die Ablage von Records vor. Dies wird damit begründet, dass sich dadurch eine stabile und klare Organisation von Records ergibt und hierarchisch aufgebaute Ablagestrukturen am weitesten verbreitet sind. Für den Aufbau der Ablagestruktur ist zunächst ein Klassifikationsschema zu definieren, mit dem neben dem Aufbau der Hierarchie auch gleichzeitig

³<http://www.dlmforum.eu/>

festgelegt wird, wo die Records abzulegen sind. Ein Klassifikationsschema, wie in Abbildung 4.3 dargestellt, besteht bei MoReq2 aus *Classes*, *Files*, *Sub-Files* und *Volumes*, in denen Records gespeichert werden können.

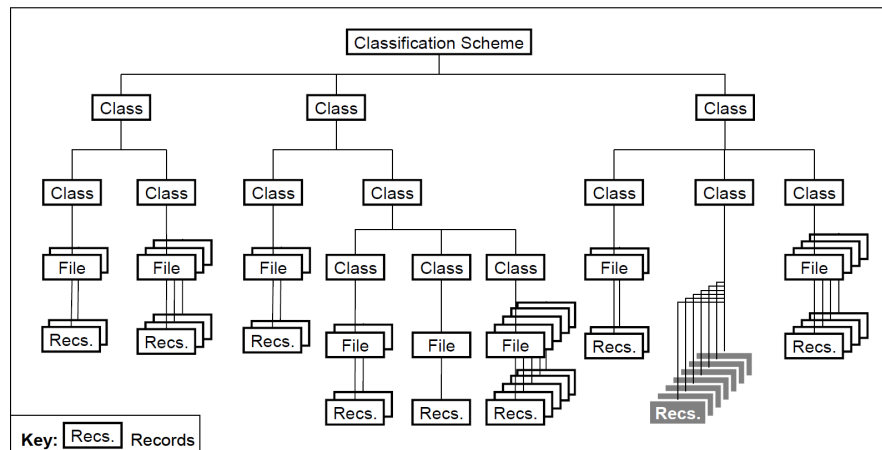
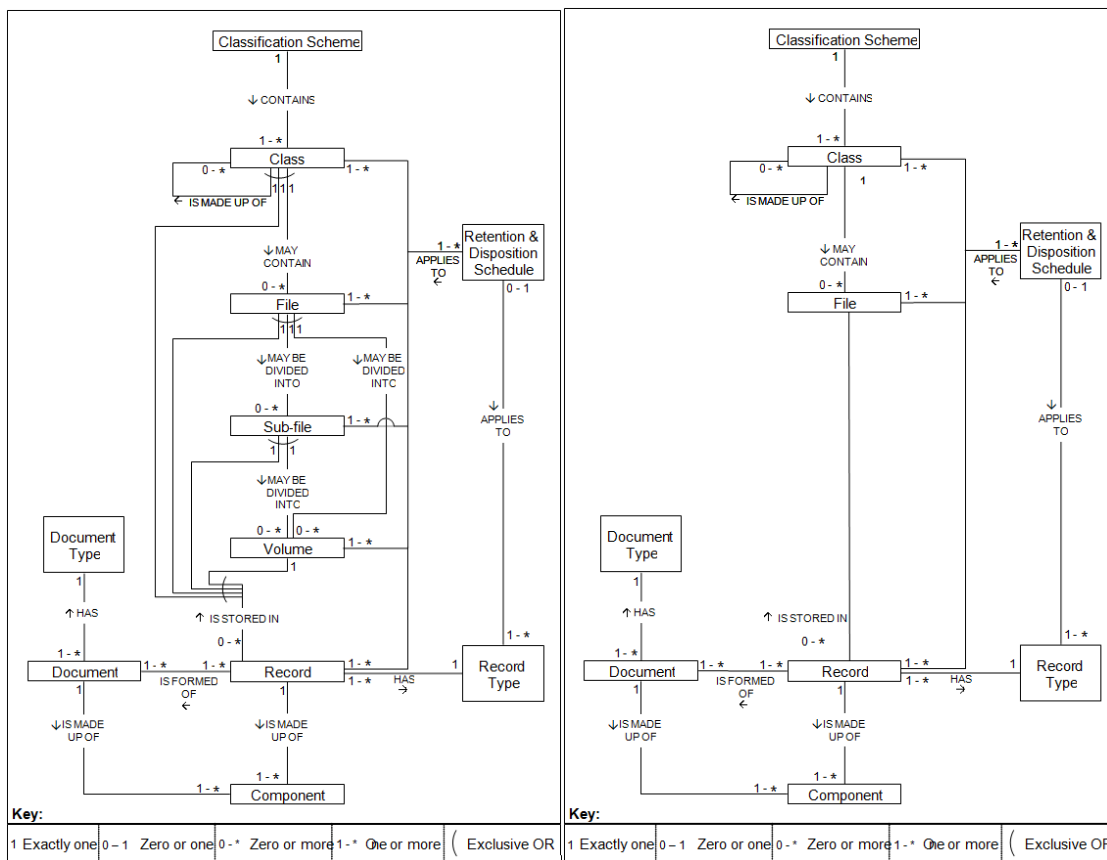


Abbildung 4.3.: MoReq2-Klassifikationsschema [Eur08b]

Für den genaueren Aufbau der Datenstrukturen bietet die Spezifikation das in Abbildung 4.4 gezeigte Entity-Relationship-Modell. In diesem Modell sind *Sub-files* und *Volumes* zur Organisation enthalten, die jedoch lediglich der Verwaltung und Ablage von papierbasierten Records dienen und entsprechen dort den klassischen Papierordnern und Einlagen in den Ordnern. Für die Integration elektronischer Records sind sie nicht relevant. Deshalb wurde das ER-Modell für die Anforderungen in dieser Arbeit abgewandelt und auf *Sub-Files* und *Volumes* verzichtet. Die Vereinfachung des ER-Modells ist nach wie vor konform zur MoReq2-Spezifikation, da eine Deaktivierung der Unterteilung in *Sub-files* und *Volumes* zulässig ist.

Die Entitäten des vereinfachten ER-Modells sind in Tabelle 4.2 aufgeführt. Mit ihnen wird der Aufbau des Klassifikationsschemas definiert. Darin enthalten sind die verschiedenen Klassen der Records und die Verzeichnisstruktur, die mit Records verbunden sind. Records werden in Verzeichnissen gespeichert. Die Definition von Aufbewahrungs- und Vernichtungsfristen erfolgt über entsprechende Schedules, die Klassen, Verzeichnissen, *Record Types* oder Records zugewiesen werden. Die *Entity-Class* besitzt in der MoReq2-Spezifikation eine Beziehung zu sich selbst. Diese Relation ist so zu verstehen, dass eine *Class* aus einer oder mehreren anderen *Classes* zusammengesetzt sein kann.



(a) MoReq2-ER-Modell [Eur08b]

(b) vereinfachtes MoReq2-ER-Modell

Abbildung 4.4.: MoReq2-ER-Modell

4.7.1.2. Kontrolle und Sicherheit

Der Bereich Kontrolle und Sicherheit definiert Anforderungen an Zugriffsbeschränkungen, Kontrollmechanismen und den Schutz der in Records gespeicherten Inhalte. In diesem Bereich bezieht sich MoReq2 vor allem auf die Vorgaben zur Authentizität und Integrität aus dem *ISO 15489-Standard* [ISO02].

Berechtigungen werden nicht direkt Benutzern zugeordnet, sondern sie werden Rollen oder Gruppen zugewiesen. Sie können hierbei Teilen des Klassifikationsschemas, der Verzeichnisse, Records und einzelnen Komponenten der Records zugewiesen werden. Eine Vererbung auf darunter liegende Verzeichnisse ist durch den hierarchischen Aufbau möglich. Durch die Zusammenfassung von Benutzern zu Gruppen bzw. Rollen erhalten diese Berechtigungen für die Zugriffe auf Records. Änderungen an den Berechtigungen dürfen nur durch bestimmte administrative Rollen erfolgen. Sämtliche

Tabelle 4.2.: MoReq2-Entitäten

Entity	Beschreibung
Classification Scheme	beschreibt eine Ablagestruktur für die Organisation
Class	Klassen dienen der Klassifikation von Records
Retention and Disposition Schedule	Aufbewahrungs- und Vernichtungsfristen
File	Files dienen der Aggregation von Records (vgl. Akte)
Record	Records bestehen aus Dokumenten und Components und können nicht mehr verändert werden
Record Type	Ein Record, der aus einem Dokument mit dem zugehörigen Dokumententyp erstellt wurde
Document	gespeicherte Informationseinheit, aber keine Klassifizierung und kein Schutz gegen Veränderung
Document Type	Beschreibt gemeinsame Eigenschaften (Layout, Inhalt, Metadaten), z. B. Formular, Vertrag, Rechnung
Component	Ein oder mehrere Dateien, die eine Einheit bilden z. B.: eine HTML-Datei mit einer JPEG-Bilddatei

Zugriffe dürfen dabei nur nach einer erfolgreichen Authentifizierung des Benutzers stattfinden, z. B. über Benutzername und Passwort.

Die Protokollierung von Zugriffen erfolgt über Audit-Trails, mit denen alle durchgeführten Vorgänge durch Benutzer, administrative Rollen oder automatisierte Vorgänge des Systems aufgezeichnet werden. Das System muss dabei sicherstellen, dass es alle Aktionen, die einen lesenden oder schreibenden Zugriff auf Klassifikationsschemata, Records und darin enthaltene Dokumente oder Komponenten durchführen, automatisch protokolliert. In einem Protokolleintrag sind Benutzer und Zeitpunkt, sowie Art des Zugriffs festzuhalten. Kein Benutzer oder Administrator darf Audit-Trails verändern können. Wird eine Historie der Dokumentenentwicklung benötigt, muss eine Versionierung aktiviert werden. Dadurch soll sichergestellt werden, dass Richtlinien zum Zugriff überprüft und unautorisierte Vorgänge identifiziert und verfolgt werden können.

4.7.1.3. Aufbewahrung und Vernichtung

Im vereinfachten ER-Modell können bei MoReq2 *Classes*, *Files* und *Record Types* mit Aufbewahrungsplänen versehen werden. Jeder Record muss zwingend einen zugeordneten Aufbewahrungsplan besitzen. Wie aus dem ER-Modell ersichtlich wird, können durch die Hierarchie der *Classes* und *Files* mehrere Pläne einem Record zugeordnet sein, die ggf. widersprüchliche Anweisungen enthalten. Hierzu sieht die Spezifikati-

on vor, dass immer nur ein Plan zur Anwendung kommen darf. Dies wird erreicht, indem die Definitionen zur Aufbewahrung in der Hierarchie absteigend überschrieben werden, sodass nur ein Aufbewahrungsplan einem Record zugeordnet wird. Auf *Record Types* angewendete Pläne dienen der Definition von Ausnahmeregelungen für bestimmte Recordtypen.

Darüber hinaus müssen die RM-Komponenten eine Vernichtungssperre bereitstellen, um einzelne Records vor der planmäßigen Vernichtung zu schützen. Ein typisches Beispiel ist eine gerichtliche Auseinandersetzung, die erfordert, dass die Daten mindestens bis zum Abschluss des Verfahrens weiter aufbewahrt werden.

Die Aufbewahrungsfrist beginnt zu dem Zeitpunkt, ab dem die im DMS gespeicherten Inhalte als Record klassifiziert werden und damit nur noch der Aufbewahrung und lesendem Zugriff dienen.

4.7.1.4. Dokumentenmanagement

MoReq2 definiert in einem optionalen Modul Anforderungen an integrierte Systeme, die sowohl DM-, als auch RM-Funktionen beinhalten. In letztere Kategorie fallen die drei DMS ELOprofessional, DocuWare ENTERPRISE und EASY ENTERPRISE.x aus Kapitel 3.

Systeme, die sowohl DM, als auch RM-Funktionen beinhalten, sollten in der Lage sein, sowohl Dokumente, als auch Records mit den gleichen Klassifikationsschemata und Zugriffsberechtigungen zu verwalten. Es muss jedoch direkt erkennbar sein, welche Inhalte als Record und welche als Dokument deklariert sind. Durch die gemeinsame Verwaltung in den gleichen Klassifikationsschemata muss das System es ermöglichen, dass alle so gespeicherten Dokumente als Records deklariert werden können. Weiterhin muss das System über die Erfassungskomponenten Funktionen bereitstellen, mit denen Daten direkt als Record oder als Dokument erfasst werden. Beispielsweise könnte eine Eingangsrechnung direkt als Record erfasst werden, um sicherzustellen, dass die Rechnung nicht mehr verändert werden kann und die Aufbewahrungsfristen sofort angewendet werden können.

Bei der Bearbeitung von Dokumenten muss sichergestellt werden, dass der Bearbeiter das Dokument durch einen Check-Out-Mechanismus vor gleichzeitigen Bearbeitungen durch andere Benutzer schützen kann. Weiterhin müssen administrative Rollen einzurichten sein, die den Check-Out-Vorgang abrechnen können, um die Sperre zu entfernen. Dadurch darf es dem Benutzer, der das Dokument ausgecheckt hat, nicht ermöglicht werden, das Dokument direkt einzuchecken. Beim Check-In muss das DMS automatisch eine neue Version des Dokuments erzeugen.

4.7.2. Integration von Dokumenten in Prozesse

Dokumente können Auslöser von Prozessen sein, etwa der Eingang einer Rechnung mit anschließendem Prüfungs- und Freigabeprozess, als Informationsträger in Prozessen auftreten oder als Ergebnis der Ausführung von Aktivitäten entstehen. In allen Fällen muss die Möglichkeit bestehen, Dokumente in den Datenfluss des Prozesses zu integrieren, um sie den Bearbeitern an der entsprechenden Stelle im Prozess zur Verfügung zu stellen. Außerdem können Dokumente im Verlauf eines Prozesses bearbeitet oder erstellt werden. Die hieraus resultierenden Dokumentenflüsse müssen bei der Modellierung von Prozessen berücksichtigt werden. Außerdem müssen die in Prozessen verwendeten, bearbeiteten oder erstellten Dokumente denselben strukturellen Aufbau besitzen, wie Dokumente, die im DMS abgelegt sind. Dies dient weiterhin dazu, dass Dokumente nicht als „Daten-Blackbox“ durch die Prozesse geschleust werden, sondern die in den Metadaten enthaltenen Informationen von Dokumenten genutzt werden können. Die zur Verfügung stehenden Metadaten müssen dazu bei der Modellierung eines Prozesses im PMS bekannt sein, um die dadurch bereitgestellten Zusatzinformationen für Entscheidungen im Prozess nutzen zu können. Dadurch kann die obligate oder optionale Versorgung von Metadatenfeldern sowohl im DMS, als auch im PMS gewährleistet werden. Wenn die Metadaten bekannt sind, kann das PMS darauf aufbauende Formulare generieren, um Benutzer bei der Eingabe von Daten zu unterstützen [Mic09a].

4.7.3. Archivierung von Prozessdaten

Die in Prozessen erzeugten Daten, die der Durchführung von Geschäftsprozessen dienen, müssen zur langfristigen Aufbewahrung archiviert werden. Die Archivierung erfolgt über die RM-Komponenten des DMS. Durch die Integrationen von Anwendungen bei der Ausführung von Aktivitäten durch das PMS werden hierbei Daten aus den Anwendungen in das PMS übernommen, Dokumente erstellt und im Datenfluss des Prozesses transportiert. Sofern diese in Prozessen genutzten Daten für die spätere Nutzung über Recherchefunktionen oder durch gesetzliche Vorschriften langfristig aufzubewahren sind, müssen diese Daten im Verlauf des Prozesses, spätestens jedoch bei Beendigung des Prozesses im DMS als Records abgelegt werden. Da Records eine definierte Struktur durch ein Recordschema für die Aufbewahrung besitzen, muss sichergestellt sein, dass die Daten auch der Struktur des Records entsprechen. Ansonsten wird eine Speicherung im DMS zurückgewiesen. Deshalb müssen die Daten im Datenfluss des PMS mit den Recordschemata abgeglichen werden.

Um eine prozessorientierte Ablage und den Bezug zum Prozess bei der Archivierung zu erhalten, muss in den Records die Prozessinstanz enthalten sein. Die so abgelegten Daten sind ab dem Zeitpunkt der Speicherung im DMS über die DMS-Clients für

berechtigte Benutzer verfügbar. Sie können so in die Recherche oder der weiteren Verarbeitung, z. B. im Rahmen des Output-Management genutzt werden.

Ein weiterer Aspekt ist, dass die Datenbanken des PMS klein und performant bleiben, da nur die Daten der aktiv laufenden Prozessinstanzen vorgehalten werden müssen und nicht über Jahre oder Jahrzehnte hinweg in der PMS-Datenbank verbleiben.

4.7.4. Organisationsintegration

Sowohl das DMS als auch das PMS erfordern, dass die Benutzer und Organisationsstrukturen des Unternehmens für die Nutzung der Software geeignet abgebildet werden. Daher sind neben der Integration auf Datenebene auch die jeweiligen Organisationsmodelle zu integrieren, um einen einheitlichen Zugriff auf die gespeicherten Informationen aus DMS und PMS zu ermöglichen. Da über das PMS bei der Ausführung von Prozessen auf Inhalte aus dem DMS-Repository zugegriffen wird, müssen sich die Benutzer des PMS dafür zuerst dem DMS gegenüber authentisieren. Weitergehende Anforderungen an die Compliance erfordern außerdem, dass auch der Bezug zur Person bei lesenden Zugriffen auf Informationen und bei Änderungen am Datenbestand protokolliert wird. Dies erfordert, dass die Benutzer in beiden Systemen eindeutig identifiziert werden.

Außer der eindeutigen Identifikation und Zuordnung der Benutzerkonten muss darüber hinaus auch sichergestellt werden, dass die zur Ausführung von Aktivitäten zuständigen Benutzer auch über ausreichende Berechtigungen beim Zugriff auf DMS-Inhalte verfügen. Dafür ist es erforderlich, dass bereits zur Entwurfszeit von Prozessen diese Berechtigungen geprüft werden, um den Prozessmodellierer über fehlende Berechtigungen zu informieren. In größeren Organisationen können sehr viele Benutzer in Prozesse involviert sein. Durch die Verwaltung von Benutzern in Gruppen- oder Rollenkonzepten ist nicht immer direkt ersichtlich, welche konkreten Benutzer die Aufgaben ausführen. Daher ist es erforderlich, dass diese Überprüfung automatisiert erfolgt.

Da PMS und DMS die eingangs beschriebenen unterschiedlichen Sichtweisen besitzen, übernehmen sie unterschiedliche Aufgabenbereiche. Da sich die Verwaltung von Benutzerkonten, Gruppen, Rollen, usw. nicht in allen Systemen gleich darstellt, sollten beide Systeme diesen Aspekt der Verwaltung und Zuweisung von Benutzern zu Gruppen oder Rollen auch bei der Integration selbst verwalten. Daher beschränken sich die Anforderungen an die Integration auf Benutzerobjekte.

Um mehrfaches und fehlerträchtiges Anlegen von Benutzern zu verhindern sollten beide Systeme die Benutzer über eine gemeinsame Benutzerdatenbank integrieren. Hier haben sich LDAP-Verzeichnisdienste zur Verwaltung von Benutzern etabliert und sind auch bereits in nahezu allen Unternehmen vorhanden.

4.7.5. Clientintegration

Bei der Integration der Clients von DMS und PMS lassen sich keine allgemeingültigen Anforderungen definieren, da hier die spezifischen Anforderungen im Wesentlichen von den durchzuführenden Aufgaben abhängen. Weiterhin richten sich die benötigten Funktionen auch an den konkreten Prozessen aus. Darüber hinaus unterscheiden sich die Implementierungen und Anpassungsmöglichkeiten der Anbieter hinsichtlich der Clients sehr stark. Rich-Clients werden in Form einer auf dem Arbeitsplatzcomputer des Benutzers installierten Anwendung realisiert, während webbasierte Clients im Browser laufen. Wie die Evaluierung in Kapitel 3 gezeigt hat, trifft dies jedoch bereits bei den fünf untersuchten Systemen nicht auf alle zu. So verfügen Alfresco und MOSS über keinerlei Rich-Clients. Die drei DMS-Systeme besitzen mindestens einen Rich-Client mit vollem Funktionsumfang und einen Web-Client, der jedoch bei allen Systemen nur einen eingeschränkten Funktionsumfang bietet.

Daher sollen in der Konzeption der Integration lediglich mögliche Ansätze zur Integration auf Clientebene aufgezeigt werden. Beispiele hierfür sind das Starten eines Prozesses aus dem DMS-Client, die Integration der Recherchefunktionen des DMS-Clients in den PMS-Client oder die Integration der Viewer-Komponenten des DMS in den PMS-Client, um Dokumente direkt im PMS-Client anzeigen zu können.

4.8. Zusammenfassung

In diesem Kapitel wurden zunächst Aspekte zur Integration von DMS, CMS und PMS betrachtet, um alle Bereiche des ECM-Konzepts durch Integrationen der verschiedenen Systeme abdecken zu können. Dazu wurden die unterschiedlichen Sichtweisen betrachtet, die einem DMS und einem PMS zugrunde liegen und welche Funktionen Dokumente innerhalb von Geschäftsprozessen besitzen. Darauf aufbauend wurden die Funktionsbereiche des ECM-Konzepts in einer Abgrenzung auf die verschiedenen Systeme verteilt, um herauszustellen welches System die entsprechenden Funktionsbereiche übernimmt.

Weiter wurden Anforderungen formuliert, die bei der Integration eines DMS mit einem PMS zu erfüllen sind. Hierzu wurden Aspekte aus der MoReq2-Spezifikation betrachtet, die Anforderungen an den Umgang mit Informationen beschreiben. Es wurde beschrieben, welche Anforderungen an die Integration von Dokumenten in Prozesse und die Archivierung von Daten zu beachten sind, die aus Prozessen heraus erstellt werden. Außerdem wurden weitere Anforderungen an die Integration der Organisationsverwaltungen formuliert, um die Benutzerverwaltungen der Systeme zu integrieren und den Zugriff auf die gespeicherten Informationen zu kontrollieren. Abschließend wurde auf die Aspekte auf Ebene der Clients eingegangen, um dem Benutzer einen integrierten Zugriff auf die Funktionen der verschiedenen Systeme zu ermöglichen.

5. Konzeption der Integration von DMS und PMS

Nachdem in Kapitel 4 Anforderungen für die Integration von Dokumenten- und Prozessmanagement formuliert wurden, wird in diesem Kapitel eine allgemeine Konzeption für die Integration eines DMS und eines PMS vorgestellt.

Zunächst wird die Integration konzeptionell auf Datenebene vorgestellt. Hierzu wird dargestellt, wie sich das DMS als gemeinsam genutztes Repository für Dokumente und die Archivierung darstellt. Dazu werden Erweiterungen für das PMS erläutert, um Dokumente mit ihren Metadaten in Prozesse zu integrieren. Weiter wird vorgestellt, wie diese Metadaten in Prozessen genutzt werden können und wie Daten, die während der Prozessausführung erstellt werden, so im DMS-Repository abgelegt werden können, dass auf diese Daten über die DM-Komponenten zugegriffen werden kann.

Anschließend wird die Integration auf Organisationsebene vorgestellt. Dazu wird gezeigt, wie die Systeme über eine zentrale Benutzerverwaltung über Verzeichnisdienste realisiert werden können und wie sich hierüber Aspekte der Authentifizierung abbilden lassen.

Abschließend werden Aspekte für eine Integration auf Clientebene vorgestellt, damit die Systeme auch bei der Interaktion mit dem Benutzer integriert präsentiert werden können.

5.1. Datenintegration

Dieser Abschnitt stellt die Konzepte zur Integration auf Datenebene eines DMS und eines PMS vor. Dazu wird zunächst beschrieben, welche Erweiterungen PMS-seitig vorgenommen werden müssen, um Inhalte aus dem DMS-Repository zugänglich zu machen. Außerdem wird vorgestellt, wie sich dadurch sowohl Dokumente leicht in Prozesse integrieren, als auch Daten oder Dokumente, die bei Durchführung eines Prozesses erstellt werden, dort ablegen lassen. Anschließend wird beschrieben, wie sich die Integration von Dokumenten mit ihren Metadaten in Prozessen darstellt. Weiterhin wird aufgezeigt, wie in Prozessen erzeugte Dokumente im DMS abgelegt werden

können. Dabei werden die Records-Management-Richtlinien zur Aufbewahrung und Langzeitarchivierung berücksichtigt.

5.1.1. Gemeinsames Repository

Für die Integration wird das DMS-Repository als gemeinsam genutzter Speicher für Dokumente, Records und zur Archivierung von Daten aus Prozessen verwendet, die bei der Durchführung entstehen. Durch den Zugriff auf das gemeinsam verwendete DMS-Repository, erhält das PMS Zugriff auf alle dort gespeicherten Inhalte. Weiterhin wird darüber ermöglicht, dass Daten aus Prozessen im DMS-Repository bereitgestellt werden können. Die im DMS-Repository gespeicherten Informationen werden mit den Management-Komponenten nach den jeweiligen Aspekten, verwaltet.

Um dies zu erreichen, muss das PMS um eine Anbindung an das DMS erweitert werden, damit der grundsätzliche Zugriff auf das DMS-Repository hergestellt werden kann. Wie in Kapitel 4 erwähnt, existieren hierzu bislang keine verbreiteten und standardisierten Schnittstellen, so dass die Integration technisch mit den herstellerspezifischen APIs des DMS erfolgen muss.

In Abbildung 5.1 ist die Integration auf Architekturebene dargestellt. Hierbei werden die physischen Speichersysteme, wie Datenbanken, Speichersysteme für lebende Daten und Archivsysteme transparent durch die Store- und Preserve-Komponenten des DMS bereitgestellt. Auf die Inhalte des DMS-Repository kann dadurch von den RM- und DM-Komponenten, sowie vom PMS zugegriffen werden. Somit können auch die Richtlinien aus den RM-Komponenten auf die gespeicherten Dokumente und Records im Repository angewendet werden.

5.1.2. Schemata

Um auf die im DMS-Repository gespeicherten Inhalte zuzugreifen, und die zugehörigen Metainformationen in Prozessen nutzen zu können, werden für die Integration, wie in Kapitel 4 beschrieben, gemeinsame Schemata für den Zugriff und die Speicherung von Dokumenten und Records benötigt. Die Schemata gewährleisten dabei das Zusammenspiel der Dokumenten-, Records- und Prozessmanagementkomponenten. Durch sie werden der strukturelle Aufbau von Dokumenten, Records und Daten aus Prozessen definiert und übergreifend für alle Komponenten bereitgestellt.

Als Basis dazu dienen die vom DMS bereitgestellten Schemata zur Definition von Dokumenten- und Recordstrukturen, da sie von allen beteiligten Komponenten des DMS bereits übergreifend verwendet werden. Um die Schemata in den Prozessen nutzen zu können, muss das PMS um diese Funktionen mit einer entsprechenden *Schema-komponente* erweitert werden. Diese ermöglicht zum einen, vorhandene Schemata aus

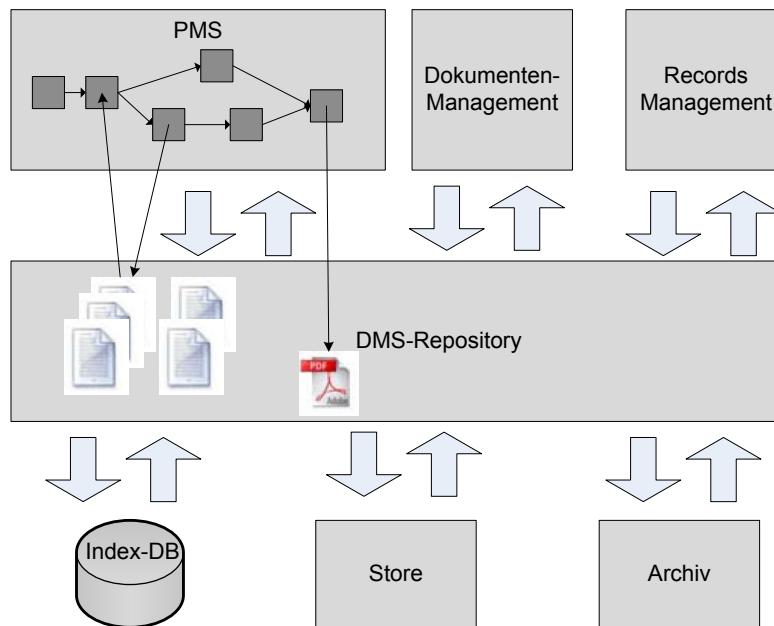


Abbildung 5.1.: Architektur der Integration von DMS und PMS

dem DMS zu übernehmen, und zum anderen, neue Schemata zu erstellen, die im PMS in die Prozessmodellierung eingebunden werden. Auf Basis der Schemata lassen sich damit auch direkt Datenstrukturen im Datenfluss der Prozesse ableiten. Beispielsweise ergeben sich aus den Metainformationen einer eingescannten Papierrechnung im DMS zusätzliche Informationen wie Rechnungsnummer, Rechnungsdatum, Informationen über den Kreditor, Rechnungsbetrag oder Lieferscheinnummer. Diese Metadaten können direkt im Datenfluss des Prozesses repräsentiert werden. Dabei können Informationen aus den Metadaten, wie der Rechnungsbetrag herangezogen werden, um einen Benutzer auszuwählen, der berechtigt ist, entsprechende Rechnungen zu prüfen. Die Kreditoreninformationen aus dem Dokument können im Verlauf des Prozesses bei Zahlungsfreigabe für weitergehende Prozessschritte bei der Zahlung vom PMS genutzt werden.

Um die Schemata übergreifend verfügbar zu machen, werden diese im DMS gespeichert. Die einzelnen Komponenten des DM, RM und das PMS müssen dazu dieselben Schemata verwenden und Zugriff auf diese besitzen. Nur durch eine Verknüpfung mit den Schemata ist es dem PMS möglich die Speicherung von Dokumenten im DMS-Repository durchzuführen. Außerdem werden sie im DMS für die Dokumenten- und Records-Management-Komponenten verfügbar gemacht. Im DMS werden die Schemata als Grundlage für die Speicherung der Dokumente genutzt. Auf ihnen basieren die Dokumentenbeschreibung, Suchmasken und die Struktur der Metadaten. Sie können dabei auch von den Records-Management-Komponenten verwendet werden, da diese die Schemata wie in Kapitel 4 beschrieben, auch für die Verwaltung von Records

verwenden. Hierüber werden auch die zur Archivierung vorgesehenen Daten aus Prozessen im DMS abgelegt und in Form von Records aufbewahrt.

Erstellen von Schemata

Die Erstellung der Schemata erfolgt durch die Zusammensetzung von elementaren Datentypen. Typischerweise werden hier die aus Programmiersprachen bekannten Datentypen wie Integer, Float, String, Boolean oder Binärdaten (BLOB¹s) unterstützt. Zusätzlich können weitere Attribute angegeben werden, die festlegen, ob das Datenfeld obligat oder optional mit einem Inhalt zu versehen ist oder ob Wertebereiche, z. B. über reguläre Ausdrücke, eingeschränkt werden. Weitere Eigenschaften aus dem Records-Management-Bereich sind Aufbewahrungs- und Vernichtungsfristen, die mit einem Schema verknüpft sind oder im Schema hinterlegt werden. Abbildung 5.2 zeigt ein einfaches Dokumentenschema in UML²-Notation [OMG09a, OMG09b], in dem einige Metadatenfelder definiert wurden. Hierbei können die beiden Felder Notizen und Beschreibung optional angegeben werden. Für die Sprache wird über einen regulären Ausdruck festgelegt, dass hier nur das zweibuchstabile Kürzel der Sprache angegeben werden darf, in der das Dokument verfasst wurde.

Dokumentenschema
-Titel : string
-Autor : string
-Notizen (optional) : string
-Beschreibung (optional) : string
-Sprache : string[a-z]{2}
-Version : int
-Datei : Blob
-Aufbewahrungsfrist : Datum

Abbildung 5.2.: Einfaches Beispiel für ein Dokumentenschema

Schemata werden entweder im DMS oder über das PMS erstellt. Die Komponente des PMS muss sich daher an den spezifischen Schemata des DMS orientieren. Um sicherzustellen, dass nur gültige Schemata erstellt werden können, muss dieser Bereich ebenso wie die DMS-spezifische API individuell angepasst werden. Für den Fall, dass Schemata zuerst im DMS erstellt werden, müssen diese durch die Schemakomponenten des PMS importiert werden können. Hierzu liest das PMS die Schemabeschreibung über die Schnittstelle zum DMS aus und erstellt dazu ein entsprechendes Schema im PMS, das auf die entsprechenden Datenstrukturen des PMS abgebildet wird. Um Schemata aus dem PMS heraus zu erstellen, etwa um ein Schema für die Archivierung von Daten aus einem Prozess zu erzeugen, muss hierzu über die Schemakomponente des PMS

¹Binary Large Object

²Unified Modeling Language

das Schema erstellt werden können und anschließend an das DMS übermittelt werden. Um Fehler bei der Übertragung zu vermeiden, sollten die Schemata hierzu automatisch über die Schemakomponente des PMS im DMS angelegt werden können. Hierzu wird ebenfalls die Schnittstelle zum DMS verwendet, um das erstellte Schema im DMS anzulegen. Da beide Systeme dadurch über die gleichen Schemata verfügen, kann sichergestellt werden, dass das PMS Informationen über den Aufbau von Dokumenten aus dem DMS besitzt.

Änderungen an Schemata

Durch Änderungen an Dokumentstrukturen oder geänderte Geschäftsprozesse, kann es erforderlich werden, dass Schemata abgewandelt werden müssen. Die daran vorzunehmenden Änderungen, die im PMS verwendet werden, sind hierbei ebenfalls über die Schemakomponente des PMS durchzuführen. Da die Schemata jedoch im DMS mit Dokumenten und Records verknüpft sind, unterliegen die erlaubten Änderungsoperationen Einschränkungen, um zu verhindern, dass bereits gespeicherte Dokumente ihre Gültigkeit bzgl. ihres Schemas nicht verlieren. Außerdem muss das PMS sicherstellen, dass laufende Prozessinstanzen, die mit einem bestimmten Schema gestartet wurden, auch korrekt terminieren können.

Die Einschränkungen der Änderungsoperationen legen daher fest, dass keine Attribute oder Metadaten entfernt werden dürfen, optionale Felder dürfen nicht obligat werden, und es dürfen keine neuen obligaten Datenfelder hinzugefügt werden. Falls Änderungen vorgenommen werden müssen, die sich nicht im Rahmen der genannten Einschränkungen durchführen lassen, muss ein neues Schema erstellt werden. Durch diese Einschränkungen wird sichergestellt, dass mit einem Schema verknüpfte Dokumente im DMS ihre Gültigkeit behalten. Für laufende Prozessinstanzen, die mit einem bestimmten Schema gestartet wurden, kann damit sichergestellt werden, dass diese bei Terminierung die Daten im DMS ablegen können.

5.1.3. Referenzen

Neben den Schemata, die den strukturellen Aufbau von Dokumenten beschreiben, werden logische Speicherorte innerhalb des DMS-Repository benötigt, um ein Dokument abzuspeichern bzw. auf ein Dokument zuzugreifen. Um die strukturierte Ablage im DMS sicherzustellen, werden daher auch die Schemata mit einem oder mehreren logischen Speicherorten verknüpft. Für die Integration benötigt das PMS daher zusätzlich die Information über die logischen Speicherorte. Diese werden typischerweise über eine (hierarchische) Anordnung von logischen Verzeichnissen realisiert. Das PMS verknüpft hierzu die Schemata mit den Referenzen auf die logischen Speicherorte.

Wenn mehrere Verzeichnisse zur Speicherung der Prozessdaten oder dem Zugriff der Dokumente zur Verfügung stehen, können diese zur Laufzeit durch den Benutzer oder die Prozesslogik bestimmt werden. Die Referenzen auf die Verzeichnisse können hierbei über eine Liste im PMS mit dem Schema zusammen gespeichert werden. Mehrere logische Speicherorte werden benötigt, um einen Dokumententyp an mehreren Stellen im DMS-Repository abzulegen. So könnten beispielsweise von einem Unternehmen erstellte Rechnungen mit einem Dokumentenschema erstellt werden. Durch die Zuordnungen zu mehreren Verzeichnissen lassen sich dann diese Rechnungen den unterschiedlichen Kunden zuordnen. Durch die Verknüpfung von Dokumentenschema und Verzeichnis wird damit sichergestellt, dass Dokumente nur an ihren vorgesehen Speicherorten abgelegt werden können. Beispielsweise wird damit verhindert, dass sensible Informationen, die in Rechnungen enthalten sind, nicht versehentlich in anderen Verzeichnissen abgelegt werden. Dies kann dazu führen, dass Rechnungsdokumente nicht wiedergefunden werden oder dass Benutzer diese einsehen können, obwohl sie dafür keine Berechtigung besitzen.

Wie im Abschnitt über die Schemata beschrieben, müssen die mit einem Schema verknüpften logischen Speicherorte bei der Übernahme von Schemata aus dem DMS vom PMS ebenfalls übernommen werden. Bei der Erstellung von Schemata über das PMS sind hierzu die logischen Speicherorte anzugeben, an denen die Dokumente abgelegt bzw. gelesen werden sollen. Beim automatischen Anlegen eines Schemas über die Schemakomponente, muss diese auch die logischen Speicherorte im DMS mit den Schemata verknüpfen.

5.1.4. Prozessintegration

Um die definierten Schemata in Prozesse integrieren zu können, muss die Schemakomponente die Schemata auf die im PMS vorhandenen Datenelemente abbilden. Damit wird ein Schema im Datenfluss eines Prozesses durch ein oder mehrere Datenelemente repräsentiert, die damit verknüpft sind. Die Eigenschaften der einzelnen Datenfelder aus dem Schema müssen dazu mit den Definitionen im Schema übereinstimmen, um sicherzustellen, dass die Informationen über obligate bzw. optionale Felder sowie die Datentypen erhalten bleiben.

In Abbildung 5.3 ist ein Beispielprozess dargestellt, der den bereits erwähnten Prozess zur Prüfung einer eingehenden Rechnung darstellt. Die hierfür benötigten Dokumente sind im DMS-Repository gespeichert und werden während der Ausführung des Prozesses direkt aus dem DMS geöffnet. Über die angegebenen Metainformationen der Dokumentenschemata kann das PMS bei der Ausführung auch direkt die weiterhin benötigten Dokumente automatisch bestimmen, da in dem gezeigten Beispiel der vorhergehende Beleg in der Kette jeweils durch eine eindeutige ID auf dem Dokument (Rechnung, Lieferschein, Auftragsbestätigung und Bestellung) hinterlegt ist. Zusätzlich können die im Prozess ermittelten Informationen über die Rechnung oder

den Lieferschein direkt als Parameter an die Anwendung übergeben werden, mit der das Reklamationsdokument erstellt wird. Dabei kann das erstellte Dokument durch die Metainformationen der zugehörigen Dokumente verknüpft werden, um den Bezug zwischen den Dokumenten herzustellen. Durch diese Informationen, die sich aus dem Schema ergeben, kann ein Dokument damit vollständig mit allen vorhandenen Metadaten in einen Prozess integriert werden.

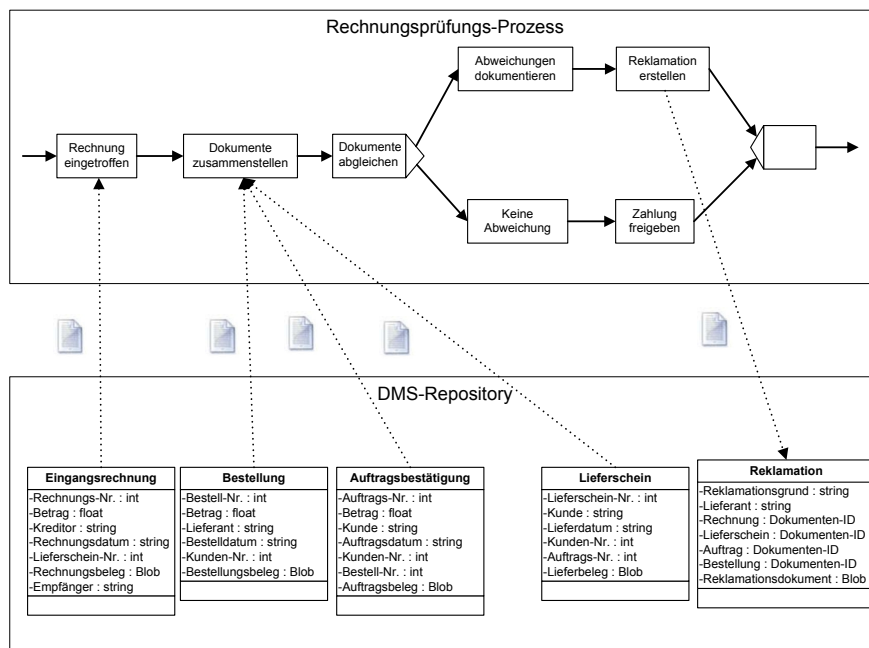


Abbildung 5.3.: Beispielprozess für eine Rechnungsprüfung

Zugriff auf Dokumente

Auf die im DMS gespeicherten Dokumente kann sowohl über das DMS, als auch das PMS zugegriffen werden. Um die Redundanzfreiheit zu erhalten, speichert das PMS immer nur die Referenzen auf das Dokument. Beim Zugriff während der Prozessausführung wird das Dokument dann direkt aus dem DMS geöffnet.

Durch die Anforderungen des Records-Management als gemeinsame Basis ergibt sich, dass alle Änderungen und Zugriffe auf Dokumente und Records im DMS protokolliert werden müssen. Hierzu wird die Versionierung im DMS für Dokumente und Records zwingend aktiviert. Dies bedeutet, dass Änderungen nur über den Check-In/Check-Out-Mechanismus des DMS erfolgen können. Beim Check-In eines Dokuments wird dabei automatisch vom DMS eine neue Version des Dokuments erzeugt.

Für das PMS sind die Zugriffe genau wie über die DM-Komponente zu realisieren. Hierbei können lesende Zugriffe und Zugriffe mit Änderungen unterschieden werden. Wenn ein Dokument oder Teile eines Dokuments im Prozess nur lesend verwendet werden, werden keine Zugriffssperren über einen Check-Out benötigt. Ob ein Zugriff nur lesend stattfindet, kann vom PMS ermittelt werden, indem im Prozess keine Schreibzugriffe auf das Dokument im modellierten Prozess vorhanden sind.

Falls ein Dokument gelesen wird, das während der Ausführung einer Aktivität geändert wird bzw. potentiell geändert werden kann, muss das PMS beim Starten der Aktivität das Dokument aus dem DMS auschecken, um eine gleichzeitige Bearbeitung über das DMS zu verhindern. Um alle Änderungen in die Audit-Trails des DMS einfließen zu lassen, ist nach Beendigung jeder Aktivität, die einen Schreibvorgang vorgenommen hat, ein Check-In mit den Änderungen durchzuführen. Hierbei werden die Änderungen ins DMS eingebracht und eine neue Version erzeugt. Um zu verhindern, dass einmal gelesene und veränderte Daten in einem Dokument während der Ausführung eines Prozesses extern geändert werden, und somit Inkonsistenzen der Daten zwischen zwei Aktivitäten auftreten können, muss das Dokument direkt nach dem Einchecken sofort wieder vom PMS ausgecheckt werden, um die Zugriffssperre für die gesamte Bearbeitungszeit während eines Prozesses zu erhalten. Da jedoch zum Zeitpunkt des Check-Out nicht klar ist, welcher Benutzer den nächsten Schritt im Prozess mit Zugriff auf das Dokument ausführt, muss hierfür die Zugriffssperre von einem gesonderten Benutzeraccount gehalten werden, da wie in Kapitel 4 beschrieben ein Check-In nur von dem Benutzer durchgeführt werden darf, der das Dokument auch ausgecheckt hat. Zwischen dem Einchecken und dem Auschecken eines Dokuments kann jedoch in diesem Moment potentiell ein Check-Out-Vorgang durch das DMS erfolgen. Damit wäre der exklusive Zugriff durch das PMS verloren und es können sich externe Änderungen im Dokument zwischen zwei Aktivitäten ergeben. Um dies zu verhindern, müssen der Check-In und anschließende Check-Out-Vorgang über eine Transaktion im DMS gekapselt werden. Transaktionen stellen sicher, dass entweder alle in der Transaktion angegebenen Vorgänge vollständig oder gar nicht durchgeführt werden. Somit werden durch die Kapselung in einer Transaktion beide Vorgänge nacheinander ausgeführt, ohne dass eine Sperre durch das DMS erfolgen kann. Die in Kapitel 3 untersuchten Systeme stellen bis auf MOSS die dafür benötigten Funktionen über ihre APIs zur Verfügung. MOSS unterstützt bislang keine Transaktionen. Um Transaktionen mit Sharepoint zu realisieren muss dies entweder auf Datenbankebene implementiert werden oder über die Definition eines MOSS-Workflows. Workflows ermöglichen das Setzen von *commit points*, über die ein transaktionales Verhalten nachgebildet werden kann [Mic09b].

Somit können durch die Verwendung von Transaktionen Inkonsistenzen der Daten zwischen zwei Aktivitäten vermieden werden. Durch diese Vorgehensweise wird außerdem auch verhindert, dass zwei Prozessinstanzen auf das gleiche Dokument zugreifen können. Der letzte Check-In-Vorgang und damit die Freigabe eines Dokuments erfolgt nach Beendigung der letzten Aktivität, die einen Schreibzugriff auf das Dokument im

Prozess besitzt. Das PMS kann dies erkennen, indem beim Beenden überprüft wird, ob noch weitere Schreibzugriffe im Prozess modelliert wurden. Durch dieses Verhalten wird erreicht, dass Dokumente bei langlaufenden Prozessen zum frühestmöglichen Zeitpunkt wieder im DMS verfügbar gemacht werden.

Für den Fall, dass ein Dokument bei Ausführung einer Aktivität bereits ausgecheckt ist, kann die Aktivität nicht gestartet werden. Hierbei informiert das PMS den Benutzer über eine entsprechende Fehlermeldung und verhindert das Starten der Aktivität. Dies kann jedoch dazu führen, dass Aktivitäten lange blockiert werden, etwa wenn ein versehentlich ausgechecktes Dokument vorliegt, das vom Benutzer nicht wieder eingechekkt wurde. Um diesem Problem zu begegnen, sollten blockierte Aktivitäten über die Monitoring-Funktionen des PMS an einen administrativen Benutzer gemeldet werden. Dieser kann den Bearbeitungsgrund ermitteln und ggf. dafür sorgen, dass das Dokument wieder freigegeben wird.

Erstellen von Dokumenten

Wie am Beispiel der Reklamation in Abbildung 5.3 dargestellt, können im Verlauf von Prozessen neue Dokumente erstellt werden, die im DMS-Repository abgelegt werden. Dazu ist die Aktivität, die die Erstellung des Dokuments repräsentiert mit dem Schema des Dokuments verknüpft. Durch die definierten Datenfelder mit ihren Datentypen kann das PMS bei Ausführung der Aktivität sicherstellen, dass alle im Dokumentenschema angegebenen Metadaten versorgt werden. Ansonsten darf die Aktivität nicht erfolgreich beendet werden und schlägt fehl.

Wie bei den Schemata beschrieben, muss zusätzlich für die Erstellung des Dokuments ein Speicherort aus der vorgegebenen Liste des Schemas ausgewählt werden. Diese Auswahl kann durch den Prozessmodellierer beim Erstellen des Prozesses erfolgen, durch den Benutzer zur Laufzeit oder implizit (wenn nur ein Speicherort angegeben ist) erfolgen. Beim Beenden der Aktivität wird das Dokument dann vom PMS im DMS-Repository am ausgewählten Speicherort abgelegt und die entsprechenden Datenfelder werden dabei dem Dokument als Metadaten hinzugefügt.

5.2. Organisationsintegration

Neben der Datenintegration müssen auch die Organisationsstrukturen des DMS und des PMS integriert werden, um Zugriffe auf die im DMS gespeicherten Daten für die PMS-Benutzer zu ermöglichen. Außerdem wird durch die Integration ermöglicht, dass das PMS und das DMS die gleichen Benutzerinformationen verwenden, um die Zugriffe in den Audit-Trails des DMS protokollieren zu können.

Hierzu wird zunächst eine Einführung in die Funktionsweise von Verzeichnisdiensten gegeben. Anschließend wird das Konzept vorgestellt, wie sich die Integration der Benutzerverwaltungen über einen Verzeichnisdienst realisieren lässt und wie damit die Authentifizierung beim Zugriff auf das DMS-Repository umgesetzt werden kann.

5.2.1. Verzeichnisdienste

Verzeichnisdienste werden typischerweise dazu verwendet, um Organisationsstrukturen abzubilden und Benutzerdaten zentral zu verwalten und Anwendungen zur Verfügung zu stellen. Die Einsatzgebiete sind weitreichend, etwa von der Ermittlung von E-Mailadressen eines Benutzers bis hin zur unternehmensweiten Verwaltung von Benutzerdaten mit der Abbildung von Organisationsstrukturen des Unternehmens einschliesslich Rechteverwaltung und Authentifizierung.

LDAP

LDAP ist ein nachrichtenorientiertes Kommunikationsprotokoll, das für die Kommunikation zwischen Client und Server dient. Ursprünglich entstand es als Zugriffsprotokoll für X.500-Verzeichnisdienste. LDAP-basierte Verzeichnisdienste gelten in der Industrie inzwischen als de-facto-Standard für die Benutzerverwaltung [KT03]. Kapitel 3 hat außerdem gezeigt, dass alle dort untersuchten CMS und DMS die Integration in LDAP-Verzeichnisdienste unterstützen. Ein PMS bietet normalerweise ebenfalls eine Integration in LDAP-Verzeichnisdienste an, so dass davon ausgegangen werden kann, dass beide Systeme eine LDAP-Schnittstelle bieten. Im Folgenden werden die vier LDAP-Basismodelle vorgestellt, die die Funktionsweise und die darin zu speichernden Daten beschreiben [HSG03, KT03].

Informationsmodell

Die gesamten in einem LDAP-Verzeichnis gespeicherten Informationsobjekte werden als DIB³ bezeichnet. Die DIB besteht aus einzelnen Einträgen, die in einer hierarchischen Beziehung zueinander angeordnet sind. Die Einträge im Verzeichnis (*Entries*) setzen sich aus einer bestimmten Menge an Attributen zusammen, die einen Attributtyp mit einem oder mehreren Werten besitzen. Attribute können obligat (*MUST*) oder optional (*MAY*) sein und werden durch Objektklassen definiert. Ein Eintrag in einem LDAP-Verzeichnis beschreibt damit ein Objekt bzw. eine Instanz einer Objektklasse. Die Objektklassen, die in einem Verzeichnisdienst zur Verfügung stehen, werden in einem LDAP-Schema abgelegt. Es beschreibt, welche Objektklassen vorhanden sind,

³Directory Information Base

welche Attribute darin vorhanden sein müssen bzw. können und welcher Syntax die Attribute folgen. Abbildung 5.4 zeigt die Objektklassen *Person* und *organizationalPerson* und ihre Vererbungshierarchie.

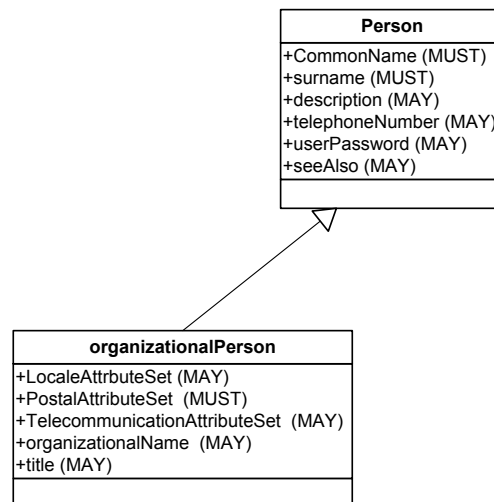


Abbildung 5.4.: Objektklasse Person mit Vererbung

Für den Datenaustausch werden die Daten textbasiert über LDIF⁴ ausgetauscht. Jeder LDAP-Eintrag im LDIF-Format besitzt die Form *Name=Wert* für Objekte oder *Attributname: Wert* für Attribute von Objekten. Tabelle 5.1 zeigt eine Übersicht der wichtigsten Attributtypen.

Tabelle 5.1.: LDAP-Attributtypen

Zeichenkette	Attributtyp
cn	Common Name: Objektname
dn	der komplette Pfad eines Eintrags
ou	Organization Unit: Name der Organisationseinheit
uid	Benutzerkennung
uidnumber	numerische Benutzerkennung

⁴LDAP Interchange Format

Namensmodell

Jeder Verzeichniseintrag kann eindeutig über seinen genauen Pfad in der Verzeichnishierarchie als DN⁵ identifiziert werden. Er wird durch Zusammensetzung von RDN⁶s gebildet (ähnlich wie Pfadangaben im Dateisystem). Die Hierarchie in einem Verzeichnisbaum wird als DIT⁷ bezeichnet und definiert gleichzeitig den Namensraum des Verzeichnisses. In Abbildung 5.5 ist ein einfacher DIT dargestellt.

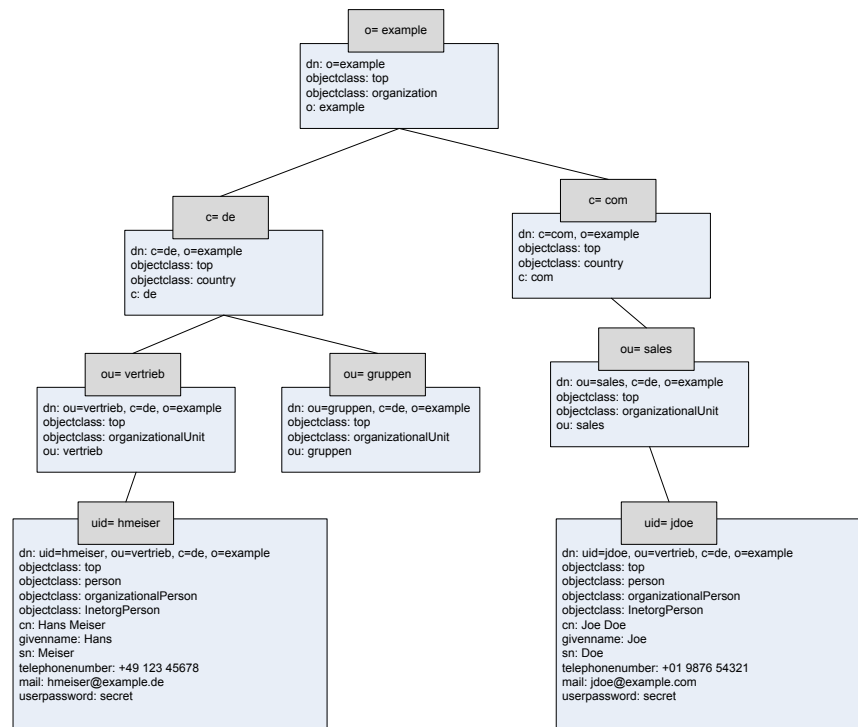


Abbildung 5.5.: Beispiel eines Directory Information Tree nach [HSG03, Kin00]

Funktionales Modell

Das funktionale Modell von LDAP beschreibt die über das LDAP-Protokoll ausführbaren Operationen. Es umfasst neun Basisoperationen, die sich in drei Kategorien einteilen lassen. Die erste enthält die Abfrageoperationen *search* und *compare*, mit denen Anfragen an das Verzeichnis gestellt werden können. Die Änderungsoperationen umfassen *add*, *delete*, *modify* und *modify DN* (Umbenennen). Die Authentifizierungs- und

⁵Distinguished Name

⁶Relative Distinguished Name

⁷Directory Information Tree

Kontrolloperationen umfassen *bind*, *unbind* und *abandon*, mit denen eine Verbindung mit dem Verzeichnis auf- bzw. abgebaut werden kann. *Abandon* teilt dem Verzeichnis mit, dass die Antwort auf die Anfrage verworfen werden kann. Zusätzlich stehen einige zusätzliche Optionen zur Verfügung, die z. B. Einschränkungen des Suchbereichs über Suchfilter im DIT oder Einschränkungen auf bestimmte Attributtypen zulassen. Jedes Objekt im Verzeichnis besitzt eine weltweit eindeutige OID⁸. OIDs werden durch eine Folge von Nummern eindeutig gekennzeichnet, die jeweils durch einen Punkt getrennt werden und sich in den hierarchischen Namensraum eingliedern.

Sicherheitsmodell

Das LDAP-Sicherheitsmodell beschreibt ein Framework, mit dem die im Verzeichnis gespeicherten Informationen vor unautorisierten Zugriffen geschützt werden können. Die einfachste Variante der Authentifizierung besteht darin, dass der Client seinen DN und sein Passwort im Klartext an den Server übermittelt. Ansonsten wird eine anonyme Sitzung etabliert. Seit Version 3 bietet LDAP die Möglichkeit der Authentifizierung mittels SASL⁹, das ein Framework für verschiedene Authentifizierungsfunktionen bereitstellt. SASL stellt alle Verbindungsprotokolle als ein Profil dar. Beim Verbindungsaufbau handeln die beiden Kommunikationspartner dann eine Authentifizierungsmethode und ggf. eine Verschlüsselungsmethode (z. B. SSL¹⁰ oder TLS¹¹) für die Übertragung aus. Über SASL kann eine Anwendung auf dem LDAP-Client den SASL-Protokolltreiber auf dem Server über eine API aufrufen. Als weiteren Parameter wird der DN benötigt, der den Verzeichniseintrag identifiziert (z. B. Benutzerkennung). Für den letzten benötigten Parameter sind die Identifizierungsdaten (Credentials) anzugeben, über die der DN identifiziert werden kann. Das Format hängt von der verwendeten Authentifizierungsmethode ab.

Die mit SASL unterstützten Protokolle unterscheiden sich hierbei vom verwendeten Verzeichnisdienst. Active Directory¹² bietet beispielsweise die drei grundsätzlichen Authentifizierungsmethoden *anonymous*, *simple* und *SASL*. Für SASL sind diverse Authentifizierungsverfahren definiert, etwa Anonymous, CRAM-MD5, Digest-MD5, Kerberos V4, Kerberos V5, X.509. Weitere LDAP-Verzeichnisse existieren von Novell mit eDirectory¹³ oder OpenLDAP¹⁴.

⁸Object Identifier

⁹Simple Authentication and Security Layer

¹⁰Secure Sockets Layer

¹¹Transport Layer Security

¹²<http://www.microsoft.com/windowsserver2008/en/us/active-directory.aspx>

¹³<http://www.novell.com/de-de/products/edirectory/>

¹⁴<http://www.openldap.org/>

5.2.2. Integration der Benutzerverwaltung

Dieser Abschnitt beschreibt, wie das DMS und das PMS die Benutzer über den zentralen Verzeichnisdienst importieren und in ihre Benutzerverwaltungen integrieren können.

Für die Integration auf Organisationsebene sind wie in Kapitel 4 bereits erwähnt, primär die Benutzerobjekte aus dem LDAP-Verzeichnis relevant. Das Zusammenfassen der Benutzerobjekte in Gruppen und die Zuweisung von Rollen zur Ausführung von Prozessen bzw. zum Zugriff auf Dokumente und Records verbleiben in den jeweiligen Systemen, da die Organisation in LDAP-Verzeichnissen häufig nach anderen Aspekten erfolgt, beispielsweise der Rechtevergabe auf Netzlaufwerke oder dem Zugriff auf bestimmte Computer und Server. Da auch Unterschiede in den Organisationskonzepten der verschiedenen Systeme existieren, genügt es daher, Benutzerobjekte aus dem Verzeichnisdienst zu importieren und diese in den jeweiligen Systemen in Gruppen und/oder Rollen zu verwalten.

Der hierarchische Aufbau eines Verzeichnisdiensts wird typischerweise dazu verwendet, Benutzerobjekte in Organisationseinheiten (OUs) zusammenzufassen. Über diese OUs sind eine oder mehrere Positionen im Verzeichnisdienst zu ermitteln, in denen die benötigten Benutzerobjekte zu finden sind. Der Pfad zu einer OU, in der nach Benutzerobjekten gesucht werden soll, kann über den DN der OU angegeben werden, z. B.: *ou=benutzer, dc=example, dc=com*. Die darin enthaltenen Benutzerobjekte werden durch ihren DN eindeutig identifiziert. Er setzt sich aus dem Pfad zur OU und dem eigentlichen Benutzerobjekt zusammen.

Die aus dem Verzeichnisdienst ermittelten Benutzerobjekte werden von DMS und PMS importiert und können über die Benutzerverwaltungskomponenten der beiden Systeme organisiert werden. Hierbei sind die eindeutige Benutzerkennung und der Benutzeranmeldename immer zu importieren. Weitere Attribute, wie E-Mailadresse, Vorname oder Nachname können optional vom DMS und vom PMS importiert werden. Durch die Integration der Benutzerobjekte über den Verzeichnisdienst erhalten beide Systeme die gleichen Benutzerobjekte und es wird eine eindeutige Zuordnung ermöglicht.

Da Benutzerobjekte sich im Laufe der Zeit ändern können, etwa durch eine Namensänderung des Benutzers oder das Ausscheiden von Mitarbeitern aus dem Unternehmen, müssen Änderungs- und Löschoptionen, die im Verzeichnisdienst an den Benutzerobjekten vorgenommen werden, auch in die beiden Systeme übernommen werden. Hierzu können die Systeme in regelmäßigen Abständen die Einträge mit dem Verzeichnisdienst abgleichen und dabei neu hinzugekommene Benutzerobjekte importieren. Geänderte Benutzerobjekte können erkannt werden, indem die ID des Benutzers gleich bleibt, die importierten Attribute jedoch geändert wurden. Die geänderten Werte können dann einfach aktualisiert werden. Beim Löschen von Benutzerobjekten müssen die

Systeme überprüfen, ob das Benutzerobjekt gelöscht werden kann, ohne dass die Funktionalität bestimmter Bereiche beeinträchtigt wird. Hierbei muss vor allem das PMS dafür sorgen, dass keine Prozessschritte existieren, die ohne das zu löschende Benutzerobjekt nicht mehr ausgeführt werden können. Um dieses Problem zu lösen kann das PMS z. B. einen Administrator benachrichtigen, der die Aufgaben des Benutzers auf einen anderen Benutzer überträgt. Der zu löschende Benutzer kann danach entfernt werden. Diese notwendigen Prüfungen sind in den Systemen normalerweise auch bereits vorhanden, da dieses Problem auch bei Benutzerobjekten auftritt, die durch das System selbst verwaltet werden.

5.2.3. Zugriffsberechtigungen

Für die Ausführung von Prozessen müssen neben den Benutzerobjekten außerdem die Zugriffsberechtigungen auf die im DMS gespeicherten Inhalte und die Ausführung von Prozessen im PMS erteilt werden. Diese Zuweisungen erfolgen mit den Verwaltungskomponenten des DMS für die DMS-Inhalte und im PMS für die Prozessausführung getrennt. Um den Benutzern die notwendigen Berechtigungen im DMS zu erteilen und um Laufzeitfehler zu vermeiden, muss sichergestellt werden, dass PMS-Benutzer auf DMS-Inhalte lesend bzw. schreibend zugreifen können.

Berechtigungen ermitteln

Bei der Modellierung des Datenflusses werden die zu verwendenden Schemata festgelegt. Durch diese Festlegung werden die im DMS gespeicherten Dokumente identifiziert, die über das PMS erstellt wurden. Die Prozessmodellierung definiert auch, welche Zugriffe lesend oder schreibend erfolgen. Die Aktivitäten, über die auf die Schemata zugegriffen wird, werden bei der Modellierung zusätzlich mit dem Organisationsmodell des PMS verknüpft. Daraus ergeben sich alle Benutzer, die diese Aktivitäten potentiell ausführen können. Die ermittelten Benutzer werden in zwei Gruppen für lesenden und schreibenden Zugriff eingeteilt. Schreibender Zugriff beinhaltet Lesen und Ändern.

Da das DMS keinerlei Informationen über die Prozesse im PMS besitzt, muss die Zuweisung der Benutzerrechte auf die DMS-Inhalte über das PMS erfolgen. Die Zuweisung der Berechtigungen erfolgt individuell für jede Prozessvorlage, um keine Beeinträchtigung oder Konflikte zwischen verschiedenen Prozessvorlagen zu erzeugen. Außerdem wird dadurch in der Rechteverwaltung des DMS erkennbar, welche Einträge durch das PMS vorgenommen wurden. Damit können versehentliche Änderungen oder gar das Entfernen von Berechtigungen über die DMS-Rechteverwaltung reduziert werden.

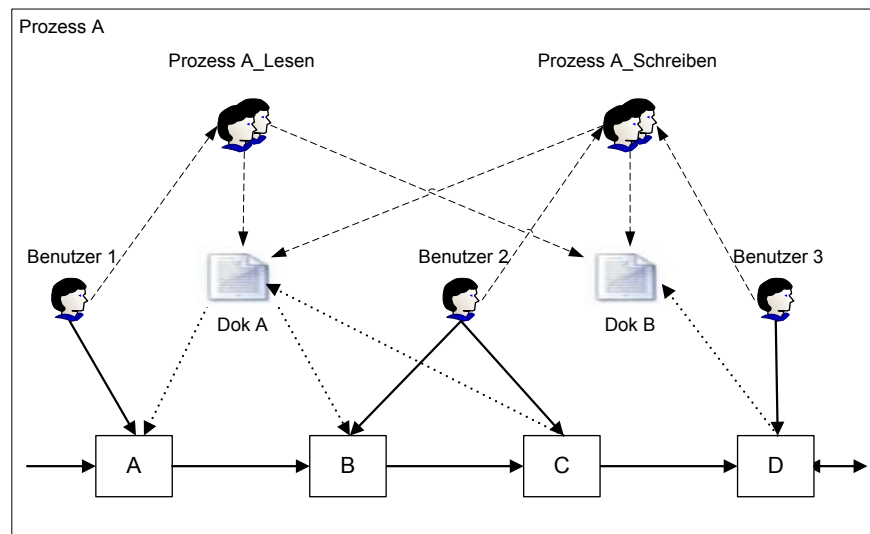


Abbildung 5.6.: Beispiel zur Ermittlung von Berechtigungen auf Dokumente im DMS

In Abbildung 5.6 ist ein Beispiel eines einfachen Prozesses dargestellt, in dem zwei Dokumente und drei Benutzer in einen Prozess eingebunden sind. Durch die lesenden und schreibenden Zugriffe auf die Dokumente wird die Zugriffsart bestimmt. Die mit diesen Aktivitäten verknüpften Benutzer werden hierzu vom PMS ermittelt und einer der beiden Gruppen *Prozess A_Lesen* bzw. *Prozess B_Schreiben* zugeordnet. Die Bezeichnung ist so zu gestalten, dass damit eine Prozessvorlage im PMS eindeutig identifiziert werden kann. Über die Schnittstelle zum DMS werden dann die beiden Gruppen im DMS angelegt, mit den in den Schemata hinterlegten Speicherorten im DMS verknüpft und die Berechtigungen zum Lesen bzw. Schreiben den Gruppen zugewiesen. Durch die gemeinsame Benutzerbasis über den Verzeichnisdienst kann das PMS die Benutzer im DMS den beiden Gruppen zuweisen. Damit wird einerseits erreicht, dass die Berechtigungen auf die DMS-Inhalte eindeutig einem Prozess und den ausführenden Benutzern zugeordnet werden kann, andererseits kann hierüber auch sichergestellt werden, dass die Gruppenbezeichnungen nicht bereits existieren und somit einen Konflikt im DMS verursachen würden. Dazu muss natürlich ausgeschlossen werden, dass im DMS manuell Gruppen angelegt werden, die der Syntax der Prozessidentifizierung entsprechen. Weiterhin muss ausgeschlossen werden, dass Änderungen an den vom PMS erstellten Gruppen und Benutzerzuordnungen über die DMS-Verwaltung durchgeführt werden. Änderungen oder das Entfernen von Benutzerobjekten müssen vom PMS nicht berücksichtigt werden, da diese Änderungen an den Attributen vom DMS übernommen werden und auch beim Löschen die Benutzer automatisch aus allen DMS-Gruppen und -Rollen entfernt werden.

5.2.4. Authentifizierung

Um den Benutzern Zugriff auf das DMS-Repository zu ermöglichen, müssen sich diese wie aus den Anforderungen in Kapitel 4 beschrieben, beim Zugriff auf DMS-Inhalte während der Ausführung von Prozessen mit ihren persönlichen Zugangsdaten dem DMS gegenüber authentisieren, um die Zugriffsberechtigungen auf DMS-Inhalte und die Zuordnungen von Zugriffen und Benutzern in die Audit-Trails des DMS sicherzustellen.

Dazu wird zunächst die Umsetzung ohne einen gemeinsamen Authentifizierungsdienst vorgestellt. Bei dieser Variante werden die Benutzer vom DMS bzw. PMS getrennt authentifiziert. Um mehrfache Anmeldungen, die Verwaltung getrennter Authentifizierungsdienste und die Handhabung der Anmeldeinformationen weiter zu integrieren, wird anschließend die Authentifizierung mittels Single-Credential und SSO erläutert.

5.2.4.1. Getrennte Authentifizierung

Erfolgt die Authentifizierung über das DMS bzw. das PMS, muss das PMS beim Start jeder Aktivität das Passwort des Benutzers abfragen, um dem Benutzer zu ermöglichen sich gegenüber dem DMS mit seinen Identifizierungsdaten zu authentisieren. Dazu muss das PMS dahingehend erweitert werden, dass beim Starten einer Aktivität, die einen Zugriff auf das DMS-Repository erfordert, die Anmeldeinformationen vom Benutzer abgefragt werden. Der Benutzer muss dann seine DMS-Zugangsdaten eingeben, damit die Authentifizierung durch das DMS vorgenommen werden kann. Diese Variante hat den Nachteil, dass in beiden Systemen unterschiedliche Passwörter und ggf. unterschiedliche Benutzernamen vorliegen. Weiterhin müssen Passwortrichtlinien in den Systemen getrennt verwaltet werden.

5.2.4.2. Single Credential

Unter Single Credential wird die Anmeldung an mehreren Systemen mit den gleichen Identifizierungsdaten (Credentials), z. B. Benutzernamen und Passwort verstanden. Der Benutzer muss sich jedoch an jedem System separat anmelden.

Dazu delegieren das DMS und das PMS die Authentifizierung an einen gemeinsamen Authentifizierungsdienst. Hierfür kann der gemeinsam verwendete Verzeichnisdienst genutzt werden, über den auch schon die Benutzer importiert werden. Die Systeme senden hierzu während des Authentifizierungsvorgangs die vom Benutzer erhaltenen Zugangsdaten an den Verzeichnisdienst. Dieser prüft, ob Benutzername und Passwort mit den im Verzeichnisdienst gespeicherten Daten übereinstimmt. Falls diese übereinstimmen, meldet er an das anfragende System die erfolgreiche Authentifizierung und

der Benutzer kann angemeldet werden. Durch die zentrale Verwaltung über den Verzeichnisdienst ergibt sich, dass das Passwort nicht von den Systemen getrennt gespeichert wird und der Benutzer somit an beiden Systemen das gleiche Passwort verwenden kann.

Um die geforderte Authentifizierung des DMS aus Prozessen heraus umzusetzen, muss der Benutzer bei jedem Zugriff authentifiziert werden. Da das PMS keine Kenntnis über das Passwort des Benutzers hat, muss dem Benutzer beim Zugriff auf das DMS-Repository eine Authentifizierungsmöglichkeit beim Start einer Aktivität angeboten werden.

5.2.4.3. Single Sign On

Die von den meisten Benutzern als lästig empfundene mehrfache Eingabe ihres Benutzernamen und Passworts entfällt bei SSO. Die Voraussetzung für SSO bildet ein gemeinsam genutzter Authentifizierungsdienst. Dies kann in diesem Fall durch den Verzeichnisdienst oder aber auch durch spezielle SSO-Systeme realisiert werden. Im Folgenden wird von einem SSO-Dienst ausgegangen, der durch einen Verzeichnisdienst mit dem Kerberos-Protokoll bereitgestellt wird. Kerberos ist ein verteilter Authentifizierungsdienst, mit dem SSO umgesetzt werden kann und der plattform- sowie systemunabhängig ist [BAN89, KN93, Lin96, NT94].

Der SSO-Ansatz mit Kerberos bietet für die Integration Vorteile, da keine Unterbrechung des Arbeitsflusses durch mehrfache Anmeldungen stattfindet und ermöglicht eine Vereinfachung der Administration. Außerdem werden keine Passwörter in mehreren Systemen gespeichert, Passwörter müssen nicht mehrfach über das Netzwerk übertragen werden und fehlgeschlagene Anmeldungen durch den Benutzer entfallen.

Der Nachteil daran ist, dass bei einer Kompromittierung ein Angreifer vollen Zugriff auf alle Systeme und Daten erhält. Außerdem existieren unterschiedliche Kerberos-Implementierungen in verschiedenen Versionen für Unix und Windows, die teilweise nicht kompatibel zueinander sind [Gar03, Hil00].

Um die Funktionsweise von SSO über Kerberos zu erläutern, wird im Folgenden allgemein auf die Authentifizierung mit Kerberos eingegangen. Dazu wird gezeigt, wie der Ablauf einer Authentifizierung eines Benutzers an einem Authentifizierungsserver mit Kerberos abläuft. Darauf aufbauend wird erläutert, wie die Authentifizierung mit weiteren Anwendungen erfolgen kann, ohne dass der Benutzer sich mehrfach authentifizieren muss.

Das Prinzip von SSO über Kerberos besteht darin, dass ein oder mehrere Authentifizierungsserver im Netzwerk vorhanden sind. Die Authentifizierungsserver werden bei

Kerberos KDC¹⁵ genannt. Der KDC ist dafür verantwortlich einen Benutzer zu authentifizieren, indem der Benutzer sich mit seinen Anmeldeinformationen am KDC authentisiert und daraufhin ein Ticket erhält, mit dem er von weiteren Systemen oder Diensten authentifiziert werden kann. In Abbildung 5.7 ist ein Kerberos-Ticket vereinfacht dargestellt. Der obere Teil enthält die unverschlüsselten Informationen, die bei Kerberos aus einem *Realm* und einem *Principal Name* besteht. Der Realm gibt den Gültigkeitsbereich eines KDC an, d. h. alle zu einem Realm gehörenden Benutzer und Systeme werden von ihm verwaltet und vertrauen ihm. Ein Realm kann beispielsweise ein DNS-Domänen-Name sein. Der Principal Name identifiziert den Besitzer des Tickets. Im verschlüsselten Teil liegen:

- einige Ticket-Flags, auf die später eingegangen wird
- der Schlüssel, mit dem Informationen zwischen dem Benutzer und dem Zielsystem verschlüsselt ausgetauscht werden
- verschlüsselte Kopien des Principal Name und des Realm
- die Start- und Endzeiten, die die Gültigkeit des Tickets angeben
- eine oder mehrere IP-Adressen, die den Rechner des Benutzers identifizieren
- die Autorisierungsdaten des Benutzers

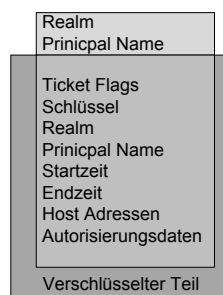


Abbildung 5.7.: Kerberos-Ticket nach [Cha99]

Zur Erläuterung der in den folgenden Abbildungen verwendeten Begriffe und zur Beschreibung der Abläufe der Kerberos-Authentifizierungen wird kurz die Notation der einzelnen Vorgänge beschrieben.

- K_X : Geheimer Schlüssel von X, wobei X der Benutzer (B), der Server (S) oder der KDC (K) sein können
- $\{ \text{Nachricht} \}_{K_X}$: Eine Nachricht, die mit dem geheimen Schlüssel von X verschlüsselt wurde

¹⁵Key Distribution Center

- $\{T\}K_S$: Ein Ticket, das mit dem Schlüssel von S verschlüsselt wurde
- $K_{X,Y}$: Ein Session Key, der zwischen X und Y verwendet wird
- $\{\text{Nachricht}\} K_{X,Y}$: Nachricht, die mit dem Session Key von X und Y verschlüsselt wurde

Erste Anmeldung am KDC: Wenn ein Benutzer sich das erste Mal an einem System anmeldet, z. B. durch Eingabe von Benutzername und Passwort bei der Betriebssystemanmeldung, wird vom Client eine Nachricht an den KDC gesendet, die einen Zeitstempel, das verschlüsselte Passwort des Benutzers und eine Anfrage zum Erhalt eines TGT¹⁶ enthält. Ein TGT ist dabei ein gewöhnliches Ticket wie in Abbildung 5.7. Die Bezeichnung dient lediglich der besseren Unterscheidbarkeit zu anderen Tickets und es dient nur der Anforderung weiterer Tickets. Wenn die Anfrage am KDC eintrifft, überprüft dieser die vom Client gesendeten Informationen und authentifiziert den Benutzer bei erfolgreicher Überprüfung. Anschließend sendet der KDC eine TGT an den Client des Benutzers zurück, das den Benutzernamen und die erwähnten Informationen eines Tickets enthält (s. Abbildung 5.8).

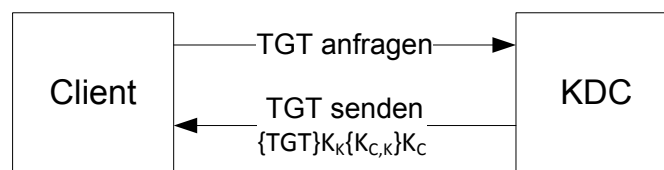


Abbildung 5.8.: Erhalt eines *Ticket Granting Ticket* nach [Cha99]

Authentifizierung an einem entfernten Serverdienst: Nachdem der Benutzer erfolgreich vom KDC authentifiziert wurde und ein TGT erhalten hat, kann er nun die Kommunikation mit einer Serveranwendung initiieren. Dazu startet der Benutzer die Client-Anwendung (z. B. den PMS-Client). Um eine Verbindung zum entfernten Serverdienst herstellen zu können, sendet der Client des Benutzers im Hintergrund eine Anfrage an den KDC, um ein *Service Ticket* für die Kommunikation zwischen Client und Server zu erhalten. In dieser Anfrage sind das TGT des Benutzers, der Name der angeforderten Serveranwendung und ein *Authenticator* enthalten, der als Nachweis dient, dass das TGT zu dem Benutzer gehört. Der Authenticator enthält dazu die aktuelle Zeit, den Benutzernamen und ist mit dem *Session Key* $K_{C,K}$ verschlüsselt, der zwischen Client und KDC verwendet wird. Nach Erhalt der Anfrage entschlüsselt der KDC das gesendete TGT und extrahiert daraus den Session Key $K_{C,K}$. Mit diesem Session Key wird der Authenticator entschlüsselt. Hierüber kann der KDC sicherstellen, dass das

¹⁶Ticket Granting Ticket

Ticket auch wirklich dem Benutzer gehört, da nur der Benutzer den korrekten Session Key $K_{C,K}$ besitzt und zusätzlich der Zeitstempel darin eingeschlossen ist.

Anschließend sendet der KDC das angefragte *Service Ticket* zur Kommunikation zwischen Client und Serveranwendung an den Client. Dazu werden die Informationen bzgl. Benutzername, Realm und Autorisierungsdaten vom TGT in das *Service Ticket* kopiert, die Ticket Flags gesetzt und ein neuer Session Key $K_{C,S}$ generiert. Dieses Ticket wird mit dem geheimen Schlüssel des Servers S verschlüsselt und an den Client mit dem Session Key $K_{C,S}$ gesendet. Der Session Key $K_{C,S}$ wird dabei mit dem Session Key $K_{C,K}$ verschlüsselt.

Nach Erhalt des *Service Tickets* kann der Client nun dieses Ticket an Server S senden. Server S kann das *Service Ticket* mit seinem geheimen Schlüssel entschlüsseln und erhält darüber den Session Key $K_{C,S}$. Dieser Ablauf ist in Abbildung 5.9 veranschaulicht.

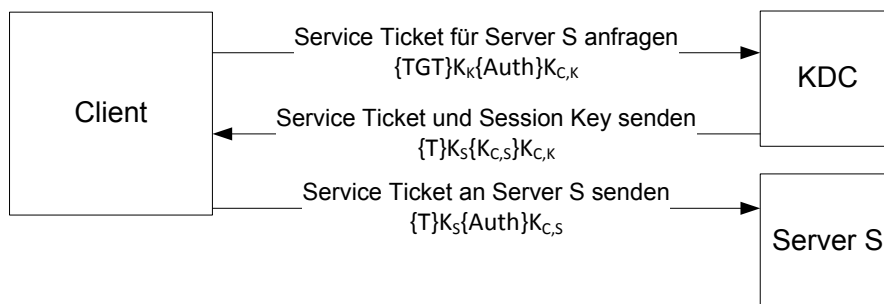


Abbildung 5.9.: Erhalt und Verwendung eines *Service Tickets* nach [Cha99]

Delegation der Authentifizierung: Um einem Server die Möglichkeit zu geben, einen Benutzer an weiteren Anwendungen über Kerberos anzumelden, besteht in Kerberos die Möglichkeit der Delegation. Eine Anwendung auf dem Client des Benutzers kann hierzu das TGT des Benutzers und den zugeordneten Session Key an einen Server S senden, damit dieser weitere Anfragen im Auftrag des Benutzers erstellen kann. Das TGT alleine genügt hierfür nicht, da der Session Key ebenfalls für die Anfrage weiterer Tickets benötigt wird. Dazu wird der Session Key $K_{C,K}$ mit dem Session Key $K_{C,S}$ verschlüsselt übertragen. Damit Server S mit dem TGT des Benutzers weitere Tickets beim KDC anfragen kann, muss dafür das *Forwardable*-Flag im TGT gesetzt sein. Durch das gesetzte *Forwardable*-Flag erkennt der KDC, dass er *Service Tickets* ausstellen kann, obwohl die im TGT enthaltene IP-Adresse nicht mit der von Server S übereinstimmt. Alternativ kann auch Server S beim KDC mit einer Vertrauensstellung zur Delegation ausgestattet werden, um ihm zu ermöglichen, weitere Tickets anzufragen.

Wenn der Client sein TGT mit gesetztem *Forwardable*-Flag an Server S gesendet hat, und Server S auf Server T mit den Anmeldeinformationen des Clients zugreifen möchte, kann Server S das TGT mit einem gültigen Authenticator an den KDC senden, um von diesem ein gültiges Service Ticket für den Client zu erhalten. Dieses neue Ticket enthält dabei die Informationen des Benutzers wie im originalen Ticket, es wird jedoch mit dem geheimen Schlüssel von Server T verschlüsselt. Somit kann Server S dieses Ticket an Server T senden, als ob dieses direkt vom Client gesendet worden wäre. Dieser Ablauf ist in Abbildung 5.10 dargestellt.

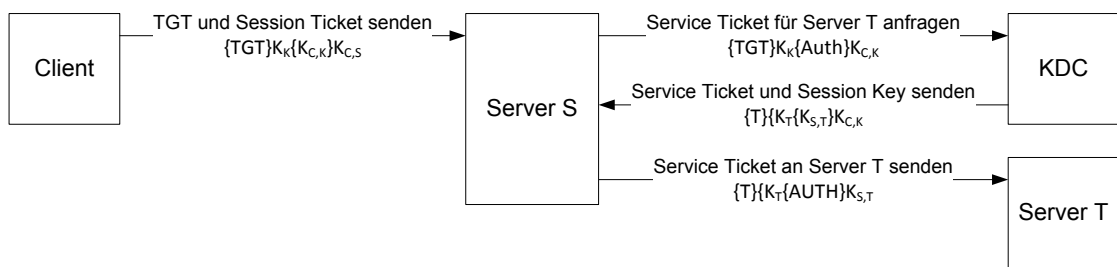


Abbildung 5.10.: Kerberos-Delegation nach [Cha99]

Authentifizierung eines Benutzers am DMS über das PMS: Da der Zugriff auf DMS-Inhalte aus dem PMS heraus erfolgt, kann der Benutzer vom PMS wie bei der Authentifizierung an einem entfernten Serverdienst vom PMS authentifiziert werden. Die Zugriffe auf DMS-Inhalte erfolgen, wie in Kapitel 4 beschrieben, über das PMS. Da bei der Authentifizierung über SSO keine weiteren Anmeldeinformationen vom Benutzer abgefragt werden, muss hierzu das PMS in der Lage sein, die Anmeldeinformationen an das DMS weiterzuleiten. Dies kann wie beschrieben mit Hilfe der Delegation über Kerberos realisiert werden. Hierbei übernimmt das PMS die Position von Server S, und kann über entsprechende Konfiguration des Verzeichnisdienstes bzw. entsprechende Kerberos-Tickets die Authentifizierung am DMS (vgl. Server T) sicherstellen.

Der SSO-Ansatz kann jedoch nur verfolgt werden, wenn sowohl das DMS, als auch das PMS in der Form angepasst werden können, dass beide Systeme den SSO-Server zur Authentifizierung akzeptieren.

5.3. Clientintegration

Dieser Abschnitt geht auf die Aspekte der Integration auf Clientebene ein. Durch die sehr unterschiedliche Gestaltung im Bereich von Clients werden hier nur Vorschläge erörtert, wie Clients eines DMS und eines PMS integriert werden können.

5.3.1. Benutzerpräsentation

Um dem Benutzer einen einheitlichen Zugriff zur Ausführung von Prozessen und den typischen Recherche- bzw. Erfassungsfunktionen des DMS zu ermöglichen, können die Clients der beiden Systeme integriert werden. Dadurch können dem Benutzer alle benötigten Funktionen in einer einzigen Anwendung präsentiert werden. Die konkrete Präsentation der Funktionen dem Benutzer gegenüber ist vom geplanten Einsatzzweck abhängig, etwa ob der Benutzer eher prozessorientierte Aufgaben übernimmt oder eher Tätigkeiten ausübt, die die Funktionen des DMS erfordern. Je nachdem sollte der entsprechende Client bei der Integration in den Vordergrund gerückt werden. Bei der Clientintegration steht vor allem der Abgleich von sehr aktuellen Daten oder die Bearbeitung von Prozessen bzw. der Aufruf von Funktionen aus dem jeweils anderen Client im Vordergrund.

5.3.2. Integration der Rich-Clients

Sowohl das PMS, als auch das DMS bieten meist Rich-Clients an, die auf dem Computer des Benutzers installiert und dort ausgeführt werden. Bei der Client-Integration werden daher die beiden Clients oder Teile davon auf dem Computer des Benutzers gekoppelt. So ist z. B. die Integration der Viewer-Funktionen des DMS-Clients, die meistens eine sehr große Zahl an anzeigbaren Dateiformaten unterstützen, im PMS-Client denkbar. Weitere Integrationsansätze können der Zugriff auf die Recherchefunktionen des DMS-Clients sein oder die Möglichkeit bei der Erfassung von Dokumenten über den DMS-Client direkt einen bestimmten Prozess zu initiieren, so dass die soeben erfassten Dokumente direkt in die eigene Arbeitsliste des PMS-Clients zur Bearbeitung übernommen werden.

Als Beispiel sei hier das Szenario der späten Erfassung von Papierdokumenten am Arbeitsplatz eines Sachbearbeiters genannt, der eingehende Papierdokumente an seinem Arbeitsplatz scannt, die erforderlichen Metadaten und die Dokumentenklassifizierung über den DMS-Client vornimmt und im Anschluss auch selbst über den PMS-Client bearbeiten möchte, ohne dass die Aufgaben anderen Sachbearbeitern automatisch in die Arbeitsliste gestellt werden. Bei diesem Szenario könnte beispielweise die ID des ausgewählten Dokuments mit dem zugehörigen Prozess an den PMS-Client übergeben werden, der dann automatisch einen neuen Prozess startet und den ersten Bearbeitungsschritt sofort dem Benutzer zuweist.

Weiterhin ist eine Integration auf Clientebene in der Form denkbar, dass Teile des einen Clients in die Oberfläche des anderen integriert werden. Hierzu können einzelne Funktionen als Plugin eingebunden werden. Beispiele für die Integration über Plugins sind in Kapitel 3 bzgl. der Integrationen von DM-Funktionen in MS Office bzw. Open Office vorgestellt. Als konkretes Beispiel im Zusammenhang des DMS- und PMS-Client

sei hier die Aufgabenliste des PMS-Clients genannt, die direkt in die Oberfläche des DMS-Clients integriert werden kann, damit der Benutzer die Aufgabenliste immer im Blickfeld behalten kann, auch wenn die Anwendung nicht im Vordergrund erscheint oder gar gestartet ist.

Voraussetzung für die Integration der Clients ist, dass diese ihre Funktionen über eine API bereitstellen können und darüber auch Aufrufe entgegennehmen. Ebenso müssen auch die Architekturen der Clients kompatibel sein, um auf technischer Ebene eine kompatible Schnittstelle verwenden zu können.

5.3.3. Integration der Web-Clients

Neben den beschriebenen Rich-Clients bieten die meisten Systeme heutzutage auch Web-Clients mit teilweise reduziertem Funktionsumfang an. Hier bieten sich Portale an, um die Clients dem Benutzer zu präsentieren. Die hierzu typischerweise verwendeten Technologien zur Einbindung stellen Webparts und Portlets dar. Von den in Kapitel 3 untersuchten Systemen können Alfresco, MOSS und EEX eigene Portale bereitstellen bzw. besitzen bereits vorgefertigte Konfigurationen zur Einbindung in Portale. Die in Abbildung 5.11 dargestellte Integration der Clients über Portlets orientiert sich dabei an der Portlet-Spezifikation [AH03].

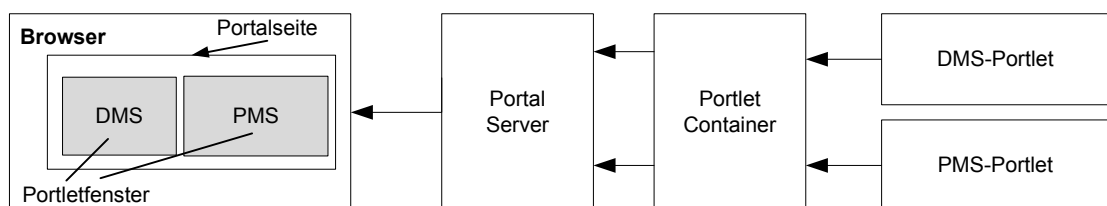


Abbildung 5.11.: Ein Web-Portal mit integriertem DMS- und PMS-Client

Außerdem besteht die Möglichkeit den Web-Client in einen eingebetteten Browser eines Rich-Clients zu integrieren, sofern dies der Rich-Client unterstützt. Im Fall von Alfresco oder MOSS aus Kapitel 3 wäre dies auch die einzige Möglichkeit, da diese Systeme keinen Rich-Client bereitstellen.

5.4. Zusammenfassung

In diesem Kapitel wurde die Konzeption der Integration eines DMS mit einem PMS vorgestellt. Für die Integration auf Datenebene wurde beschrieben, wie das DMS als gemeinsames Repository der beiden Systeme dient. Dazu wurden die dafür notwen-

digen Erweiterungen für das PMS erläutert, damit dieses auf die Inhalte zugreifen kann und um Dokumente im DMS bereitzustellen, die während der Ausführung von Prozessen erzeugt werden. Mit diesen Erweiterungen ist es möglich, Inhalte aus dem DMS-Repository in Prozesse einzubinden und dabei die bereitgestellten Metadaten im PMS zu nutzen. Darüber hinaus wurden die zuvor formulierten Anforderungen des Records-Management umgesetzt, so dass die im DMS-Repository abgelegten Informationen nach diesen Richtlinien verwaltet und am Ende der Aufbewahrungszeit kontrolliert vernichtet werden können.

Durch die Integration der Benutzerverwaltungen über einen zentralen Verzeichnisdienst können die Zuordnungen der Benutzer aus dem PMS bei Zugriffen auf DMS-Inhalte aus Prozessen heraus in die Audit Trails des DMS einfließen. Über die Nutzung der Authentifizierungsfunktionen, die Verzeichnisdienste den Systemen zur Verfügung stellen, können darüber hinaus mehrfache Anmeldungen vermieden werden.

Die vorgestellten Aspekte der Clientintegration schließlich runden die Integration von Dokumenten- und Prozessmanagement ab, indem Möglichkeiten aufgezeigt wurden, wie die Rich-Clients und die Web-Clients, die die Systeme besitzen dem Benutzer gegenüber einheitlich präsentiert werden können. Außerdem wurde anhand von Beispielen erläutert, welche Funktionen in den jeweils anderen Client übertragen werden können.

6. ADEPT2-Metamodell

Da in Kapitel 7 die Integration des Prozess-Management-Systems ADEPT2 mit einem DMS präsentiert wird, werden in diesem Kapitel die Grundlagen des ADEPT2-Metamodells behandelt und in kompakter Form vorgestellt.

ADEPT2 bietet zur Modellierung von Prozessen eine graphische Beschreibungssprache, die auf symmetrischen Blockstrukturen basiert. Für die Modellierung eines Prozesses mit ADEPT2, wird ein Kontrollflussgraph erzeugt, der die Basis für einen ausführbaren Prozess bildet. Hierbei wird bei ADEPT2 zwischen *Prozessvorlagen* und *Prozessinstanzen* unterschieden. Die Prozessvorlage stellt die vom Prozessmodellierer erstellte Beschreibung eines Geschäftsprozesses dar, die zur Laufzeit von ADEPT2 durch Erstellen einer logischen Kopie der Vorlage ausgeführt werden kann. Der modellierte Geschäftsprozess wird damit zu einer Prozessinstanz der Prozessvorlage. ADEPT2 erlaubt als Besonderheit zu den meisten anderen PMS, dass Änderungen an einem Prozessgraphen zur Laufzeit vorgenommen werden können, ohne andere Instanzen des Prozesses zu beeinflussen [Jur06, Rin04]. Bei Änderungen an einem Prozessgraph bleibt die Korrektheit erhalten.

Abbildung 6.1 zeigt drei Prozessinstanzen einer Prozessvorlage. Die angebrachten Markierungen an den Knoten des Graphen repräsentieren den Zustand der in Ausführung befindlichen Prozesse. In Instanz 3 ist zusätzlich durch Einfügen von Knoten G eine instanzspezifische Änderung enthalten. Durch die Ausführungshistorie des Prozesses ist zu jeder Zeit ersichtlich welche Knoten bereits ausgeführt wurden, ausführbar sind oder noch auszuführen sind.

6.1. Kontrollfluss

Prozesse werden in ADEPT2 durch symmetrische Blockstrukturen repräsentiert. Die Blöcke sind immer durch eindeutige Start- und Endknoten eingeschlossen. Blöcke können ineinander verschachtelt oder sequenziell angeordnet werden. Eine Überlappung von Blöcken ist jedoch nicht möglich.

Als Blockkonstrukte werden *Sequenzen*, *Schleifen*, sowie *bedingte* und *parallele Verzweigungen* angeboten. Ein Prozess stellt dabei selbst ebenfalls durch die eindeutigen Start- und Endblöcke einen Block dar.

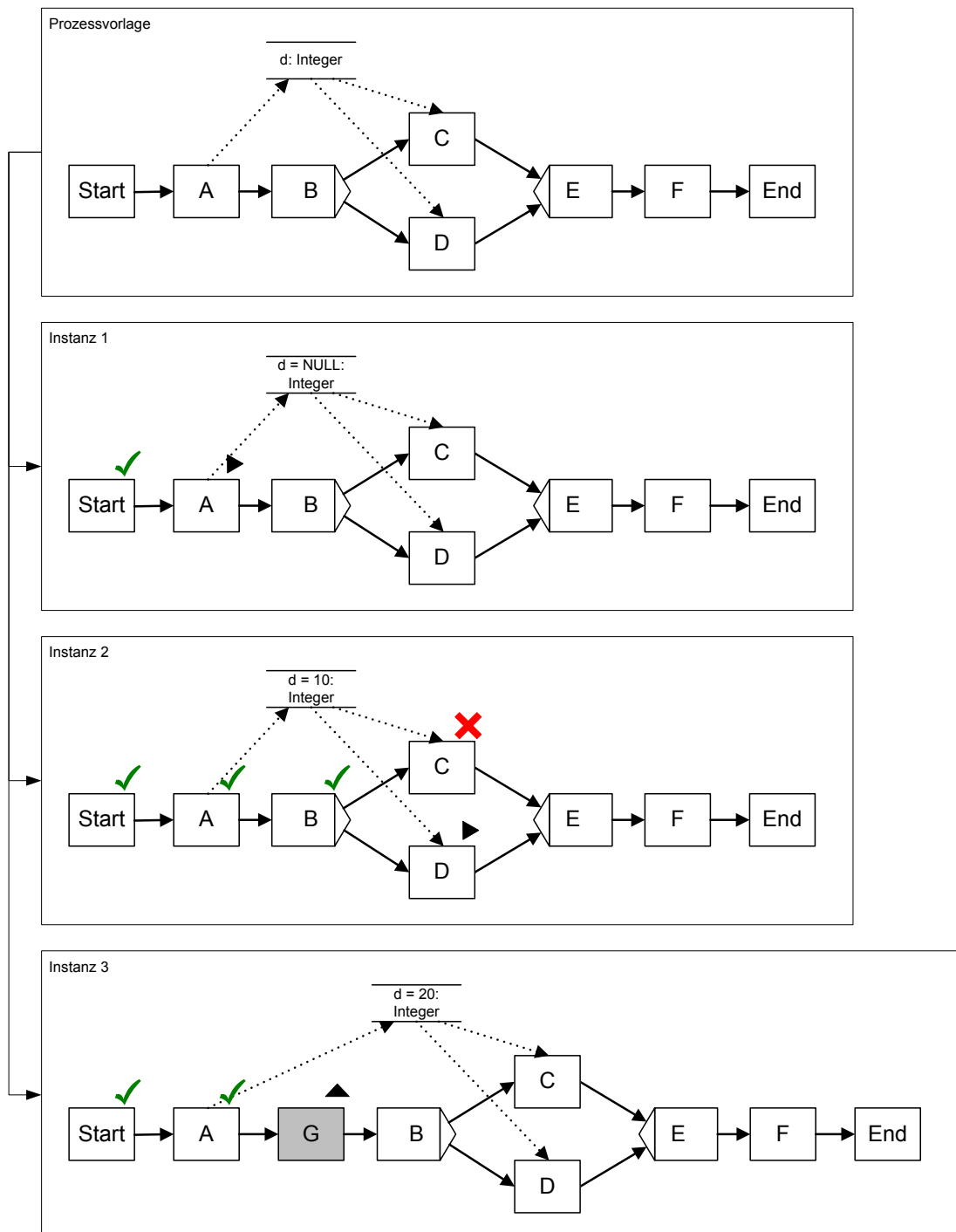


Abbildung 6.1.: Prozessvorlage mit drei Instanzen

Die Verbindung von Knoten untereinander erfolgt über *Kontrollkanten*. Die wichtigsten Blockkonstrukte sind in Abbildung 6.2 dargestellt.

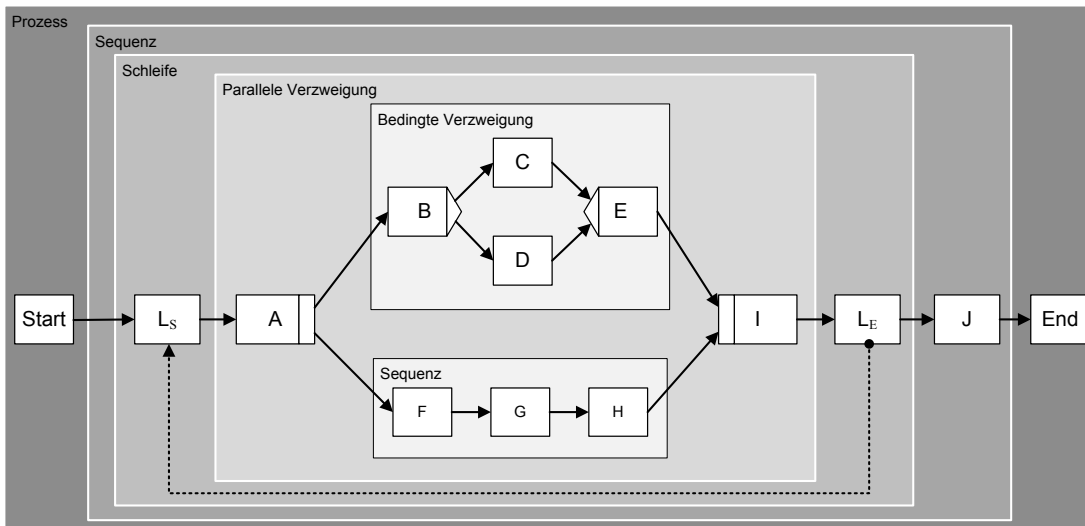


Abbildung 6.2.: Schematische Darstellung der Kontrollflusskonstrukte nach [Rei00]

Diese umfassen Sequenzen, bei denen die Knoten zur Laufzeit nacheinander ausgeführt werden. Die Ausführung eines Knotens ist dabei von der Beendigung des Vorgängerknotens abhängig.

Die *parallele Verzweigung* ermöglicht, nach Beendigung des Verzweigungsknotens mehrere Knoten parallel auszuführen. Erst wenn alle parallel ausgeführten Knoten beendet wurden, kann der Vereinigungsknoten ausgeführt werden.

Bei der *bedingten Verzweigung* wird zur Laufzeit genau ein Nachfolger des Verzweigungsknotens ausgeführt. Der Knoten wird anhand der Verzweigungsbedingung des Verzweigungsknotens bestimmt. Dazu wird jeder ausgehenden Kontrollkante des Verzweigungsknotens ein Wert bzw. Wertebereich zugeordnet, der mit der Auswertung einer Verzweigungsentscheidung übereinstimmen muss, um zur Ausführung zu kommen. ADEPT2 stellt sicher, dass nur ein Zweig ausgewählt wird, indem die Wertebereiche disjunkt gewählt werden müssen. Die nicht gewählten, und damit nicht mehr zur Ausführung kommenden Zweige werden übersprungen und um sicherzustellen, dass sie nicht mehr zur Ausführung kommen durch das *dead path elimination* [LA94, Bau96] Verfahren als nicht mehr ausführbar gekennzeichnet. Sobald der ausgewählte Zweig beendet wurde, kann der Vereinigungsknoten zur Ausführung kommen.

Für mehrfach auszuführende Aufgaben innerhalb eines Prozesses, können Bereiche eines Prozesses in *Schleifen* eingebettet werden. Schleifen besitzen, wie Verzweigungen, ebenfalls eindeutige Start- und Endknoten. Die Anzahl der Ausführungen einer Schlei-

fe werden durch Abbruchbedingungen definiert, die am Ende des Durchlaufs überprüft werden. Eine Schleife wird daher immer mindestens einmal durchlaufen.

Weiterhin existieren noch einige Kontrollflusskonzepte, die sich beispielsweise auf die Reihenfolge parallel ablaufender Prozessschritte, *ODER-Verzweigungen* oder *Subprozesse* beziehen. Diese können in [Rei00] und [Wol08] nachgelesen werden.

6.2. Datenfluss

Der Datenfluss wird in ADEPT2 über *Datenelemente* realisiert, die prozessweit sichtbar sind. Zur eindeutigen Identifikation der Datenelemente besitzen diese jeweils einen eindeutigen Namen innerhalb des Prozesses. Die Verbindung von Datenelementen und Prozessschritten erfolgt über sog. *Aktivitäten*. Aktivitäten repräsentieren die zu einem Prozessschritt gehörenden Informationen, mit denen der Datenaustausch zwischen Prozess und auszuführendem Anwendungsprogramm geregelt wird. Zur Modellierung der Zugriffe werden Datenelemente mit Lese- bzw. Schreibkanten mit den Parametern der Aktivitäten verbunden.

Um zu gewährleisten, dass der zur Modellierzeit festgelegte Datenaustausch zwischen Datenelementen und den *Aktivitätenparametern* zur Laufzeit sichergestellt werden kann, besitzen sowohl die Aktivitätenparameter, als auch die Datenelemente einen Datentyp. Dieser Datentyp muss auf beiden Seiten identisch sein, um eine Verbindung zwischen Aktivität und Datenelement herstellen zu können. Weiterhin muss für Aktivitätenparameter definiert werden, ob sie optional oder obligat sind. Bei optionalen Aktivitätenparametern wird damit bestimmt, ob zur Laufzeit eine Aktivität korrekt ausgeführt werden kann, wenn der Eingabeparameter NULL als Parameter übergeben wird. NULL bedeutet, dass das Datenelement einen undefinierten Wert besitzt, da noch kein Wert geschrieben wurde. Bei Ausgabeparametern wird damit entsprechend festgelegt, dass das Schreiben eines Datenelements nicht erfolgen muss. Bei obligaten Eingabeparametern wird damit definiert, dass das zu lesende Datenelement einen Wert (\neq NULL) enthält. Bei Ausgabeparametern muss die Aktivität einen Wert (\neq NULL) in das Datenelement schreiben. In Abbildung 6.1 sind in den dargestellten Prozessinstanzen jeweils ein schreibender und zwei lesende Datenzugriffe enthalten. Eine graphische Unterscheidung zwischen optionalen und obligaten Zugriffen existiert nicht.

Die Datenelemente in ADEPT2 erlauben zunächst nur atomare Datentypen abzubilden. Mit den in [For09] eingeführten strukturierten und komplexen Datentypen wurde eine Erweiterung eingeführt, mit der nun zusammengesetzte Datentypen erstellt werden können. Hierzu werden mehrere atomare Datentypen definiert, die als ein zusammengesetzter Datentyp repräsentiert werden. Die Datenfelder innerhalb eines strukturierten Datentyps sind jeweils eindeutig über einen Namen identifiziert. Darüber hinaus lassen sich Einschränkungen auf einfache Datentypen definieren, mit denen

Wertebereiche durch die Angabe von Regeln beschränkt werden können. Sie können als eigenständiger Datentyp und als Datentyp für ein Feld eines zusammengesetzten Datentyps verwendet werden. Ein Beispiel für einen strukturierten Datentyp ist in Abbildung 6.3 dargestellt.

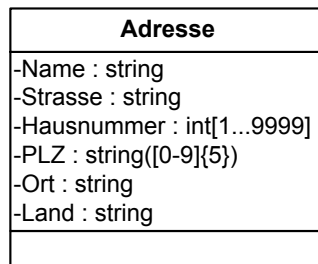


Abbildung 6.3.: Beispiel eines strukturierten Datentyps *Adresse* nach [For09]

Weitergehend können strukturierte Datentypen wieder Felder enthalten, deren Datentyp selbst strukturiert ist. Außerdem ist es neben dieser hierarchischen Beziehung der Datentypen möglich, auf diese eine Vererbung anzuwenden. Hierzu ist in Abbildung 6.4 ein Beispiel dargestellt, bei dem ein Datentyp *Firmenkontakt* zu den eigenen Datenfeldern, auch die des Datentyps *Adresse* besitzt.

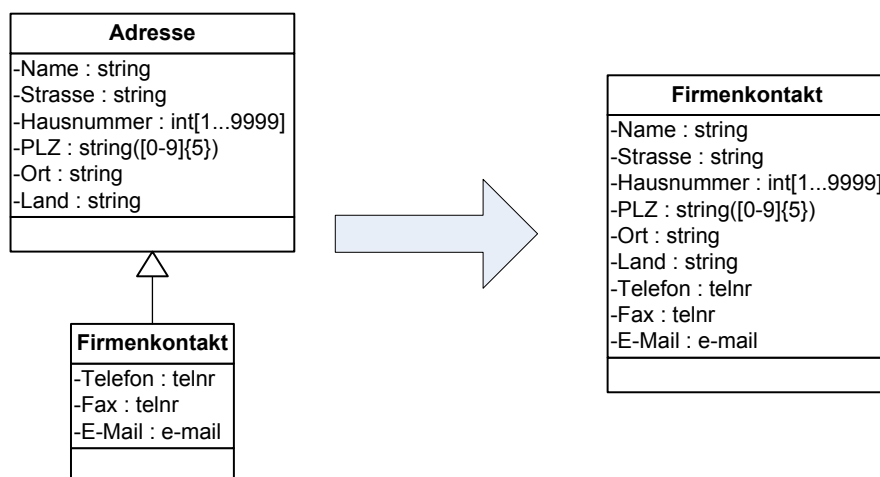


Abbildung 6.4.: Beispiel eines strukturierten Datentyps *Firmenkontakt* einschließlich Vererbung nach [For09]

6.3. Korrektheit

Mit den bisher vorgestellten Modellierungskonstrukten, lassen sich zwar Prozesse modellieren, jedoch stellen diese alleine nicht sicher, dass die damit modellierten Prozesse auch korrekt vom System ausgeführt werden können. So können z. B. Datenelemente obligat gelesen werden, bei denen nicht zuvor durch einen obligaten Schreibzugriff sichergestellt wurde, dass das Datenelement auch wirklich Daten enthält.

Daher muss jeder ADEPT2-Prozess acht Korrektheitskriterien (KF-1 – KF-8) bzgl. des Kontrollflusses erfüllen [Rei00]. Außerdem sind für die Korrektheit des Datenflusses in [Rei00] drei Regeln beschrieben, die sicherstellen, dass obligate Eingabeparameter versorgt werden, parallele Schreibzugriffe nicht aus parallel ausgeführten Zweigen erfolgen dürfen und dass zwischen zwei Schreibzugriffen mindestens ein lesender Zugriff stattfinden muss.

6.4. Organisationsmodell

Das in [Ber05] beschriebene ADEPT2-Organisationsmodell dient der Abbildung von Organisationsstrukturen. Neben der Abbildung der organisatorischen Struktur und der Beziehungen untereinander, dient es auch als Basis für die Definition der Bearbeiterzuordnungen, die für die Ausführung einer Aktivität in Frage kommen, z B. alle Personen, die der Rolle „Sachbearbeiter“ zugewiesen sind. Dazu wird jedem Knoten im Prozess über eine *Bearbeiterformel* eine Menge an möglichen Bearbeitern zugeordnet. Die Bearbeiterformel wird zur Laufzeit ausgewertet, um die Aufgaben den entsprechenden Bearbeitern über eine Arbeitsliste zur Erledigung zuzuweisen.

Zur Abbildung von Organisationsstrukturen werden durch das ausdrucks mächtige Organisationsmodell von ADEPT2 vielfältige Modellierungsoptionen angeboten, mit denen sich Strukturen abbilden lassen. So lassen sich neben den gängigen Modellen, die Benutzer, Gruppen und Rollen berücksichtigen, weitere Aspekte, wie spezielle Projektgruppen, Stellen und Fähigkeiten modellieren. Anhand dieser Modellierungsoptionen lassen sich die Organisationsstrukturen sehr feingranular abbilden und gleichzeitig für die Bestimmung von Benutzern zur Ausführung von Aktivitäten nutzen. Das Organisationsmodell aus [Ber05] ist in Abbildung 6.5 dargestellt.

Um die Komplexität der Organisationsstruktur einfacher gestalten zu können, sind bis auf die Basisentitäten *OrgUnit*, *OrgPosition* und *Agent* nicht benötigte Entitäten ausblendbar, wenn nicht alle Optionen benötigt werden. Außerdem bestehen Anbindungen an LDAP-Verzeichnisdienste, um hierüber bereits vorhandene Strukturen zu übernehmen.

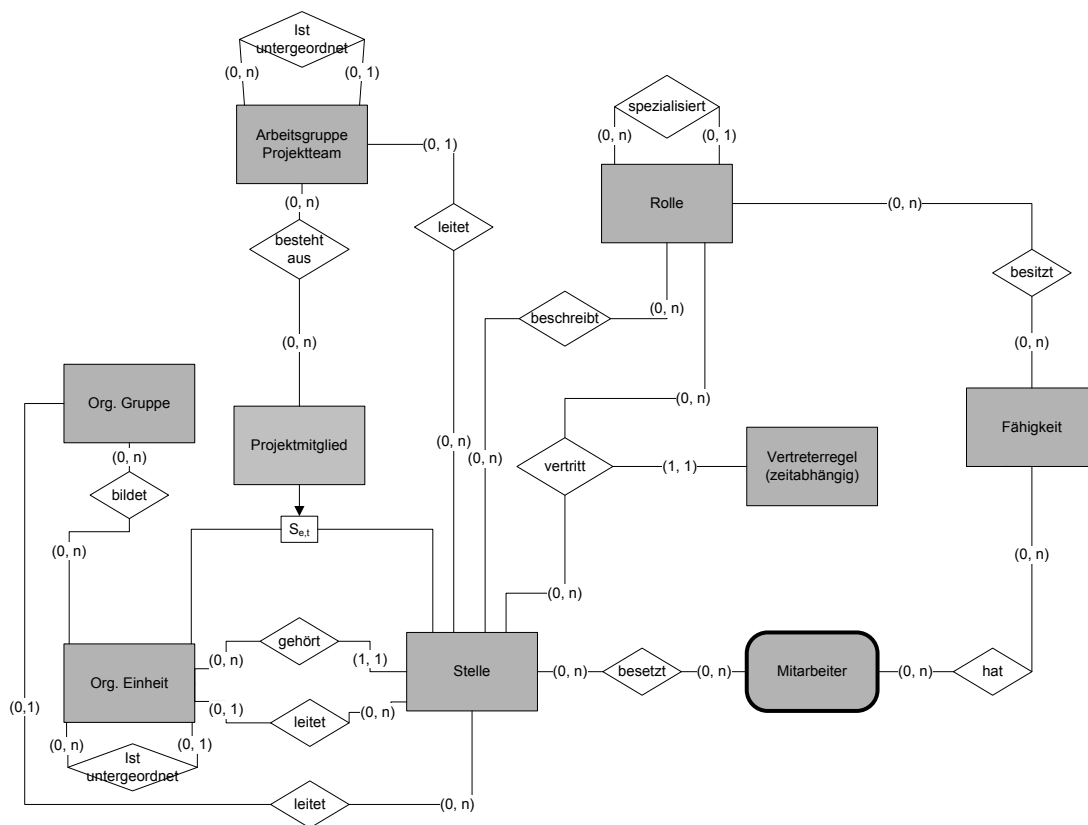


Abbildung 6.5.: ADEPT2-Organisationsmodell [Ber05]

7. Integration von ADEPT2 mit einem DMS

In diesem Kapitel wird die Umsetzung der Konzeption aus Kapitel 5 zur Integration eines DMS mit ADEPT2 beschrieben. Dazu wird zuerst auf die Aspekte der Integration auf Datenebene eingegangen. Darin werden die in der Konzeption beschriebenen Erweiterungen für ADEPT2 anhand der Schemata zur Verwendung einer gemeinsamen Beschreibung der DMS-Inhalte dargestellt. Anschließend wird gezeigt, wie sich diese in ADEPT2-Prozesse eingliedern lassen und wie die notwendigen Zugriffsberechtigungen für die DMS-Inhalte ermittelt werden. Für die Integration auf Organisationsebene werden die aus der Konzeption dargestellten Varianten zur Authentifizierung erläutert. Diese umfassen die getrennte Benutzerverwaltung im DMS und ADEPT2 sowie die Integration über einen Verzeichnisdienst einschliesslich Single Credential und SSO-Authentifizierung. Abschließend werden Integrationsmöglichkeiten auf Client-Ebene des ADEPT2 Rich-Clients und des Web-Clients mit den Clients eines DMS aufgezeigt.

7.1. Datenintegration

Dieser Abschnitt stellt die Umsetzung der in Kapitel 5 beschriebenen Konzepte der Integration auf Datenebene mit dem PMS ADEPT2 vor. Zuerst werden Schemata für ADEPT2 eingeführt, mit denen DMS-Inhalte im Datenfluss von ADEPT2 repräsentiert werden. Sie dienen der Synchronisation des strukturellen Aufbaus von Dokumentenschemata aus dem DMS und dem Datenfluss in ADEPT2. Außerdem erfolgt durch sie der Datenaustausch mit DMS-Inhalten. Dazu wird zunächst der Aufbau von Schemata erläutert und wie diese mit den strukturierten Datentypen von ADEPT2 realisiert werden. Anschließend wird erläutert, wie die Zugriffe auf DMS-Inhalte implementiert werden müssen, um die extern gespeicherten Daten in ein Schema in ADEPT2 einzubringen und wie mit konkurrierenden Zugriffen auf Dokumente im DMS-Repository umzugehen ist.

7.1.1. Schemata

Die Schemata bilden die Basis für die Datenstrukturen der DMS-Inhalte. In ADEPT2 sollen die Schemata daher so gestaltet werden, dass sie sich in die vorhandenen Datenflusskonzepte von ADEPT2 eingliedern lassen, so dass Änderungen an Aktivitäten

und den Datenflusskonzepten möglichst gering ausfallen. Außerdem sollen die Schemata nicht nur für die Integration eines bestimmten DMS ausgelegt sein, sondern so allgemein gehalten werden, dass sie für die Integration unterschiedlicher DMS geeignet sind. Die technische Implementierung der Schnittstellen muss dabei aufgrund der nicht vorhandenen standardisierten Schnittstellen jedoch für jedes DMS individuell erfolgen. Auf diese Aspekte wird daher nicht näher eingegangen.

Die Hauptaufgabe der Schemata stellt die Repräsentation von Datenstrukturen von DMS-Inhalten dar. Sie werden sowohl für die Integration von vorhandenen Inhalten aus dem DMS, als auch zur Ablage von Daten verwendet, die während der Ausführung von Prozessen im DMS abgelegt werden sollen. Hierzu müssen die spezifischen Dokumenten-Schemata je nach verwendetem DMS erstellt und in den ADEPT2-Datenfluss eingebunden werden. Außerdem müssen auch im DMS vorhandene Schemata übernommen werden können, um bereits vorhandene Dokumente in Prozesse einbinden zu können und die Schemata dabei konsistent zu übernehmen.

7.1.2. Aufbau von Schemata

Um die Datenstrukturen von DMS-Inhalten, wie Dokumente in ADEPT2 abzubilden und dabei die zugehörigen Metadaten mit einzubeziehen, wird eine Datenstruktur benötigt, die diese exakt als Einheit abbildet, um alle Metadatenfelder mit den jeweils entsprechenden Datentypen, sowie den enthaltenen Dokumentendateien.

Der Aufbau eines Schemas in ADEPT2 umfasst daher immer drei DMS-spezifische Datenfelder, sowie einen strukturierten Datentyp, der die Struktur von DMS-Inhalten abbildet. Die drei Datenfelder bestehen dabei aus einem Datumsfeld für die Aufbewahrungsfrist, einer Liste mit den Referenzen, die auf die Speicherorte innerhalb des DMS-Repository verweisen und einem ID-Feld, das die eindeutige Identifikation eines Dokuments innerhalb des Speicherortes repräsentiert.

Wie in der Konzeption beschrieben bleiben die DMS-Inhalte, die in Prozesse eingebunden werden, im DMS-Repository gespeichert und werden über Referenzen in den ADEPT2-Datenfluss eingebunden. Typischerweise besitzen DMS eine hierarchisch aufgebaute Ablagestruktur. Diese Struktur kann über die DMS-Schnittstelle von ADEPT2 ausgelesen werden. Die mit einem Schema verknüpften Speicherorte können damit über eine Liste von Verzeichnispfaden repräsentiert werden, die als Referenzen dienen. Die Dokumenten-ID dient der eindeutigen Identifizierung des Dokuments innerhalb des Verzeichnisses. Bei den meisten DMS ist diese ID auch innerhalb des gesamten DMS-Repositories eindeutig. Die Datumsangabe dient vor allem dazu, beim Erstellen von Schemata im DMS eine Aufbewahrungsfrist aufgrund der RM-Anforderungen anzugeben. Zusätzlich kann darüber festgelegt werden, dass z. B. einzelne Dokumente eine andere als die mit dem Schema verknüpfte Aufbewahrungsfrist erhalten. Ein Beispiel hierfür könnten Verträge sein, die abhängig von ihrer Laufzeit individuell aufbe-

wahrt werden sollen. Diese Option ist abhängig vom konkret vorliegenden DMS und dessen Konfiguration, und ob dies für ein bestimmtes Dokumentenschema zulässig ist. Weiterhin kann diese Information dazu dienen, in Prozessen zu erkennen, ob auf ein Dokument aufgrund einer abgelaufenen Aufbewahrungsfrist im DMS nicht mehr zugegriffen werden kann.

7.1.3. Abbildung von DMS-Inhalten auf strukturierte Datentypen

Neben den drei Datenfeldern, die jedes Schema besitzt, sollen die Struktur und die konkreten Inhalte eines Dokuments auf einen strukturierten Datentyp in ADEPT2 abgebildet werden. Ein strukturierter Datentyp ist innerhalb von ADEPT2 eindeutig und besteht aus mindestens einem Feld. Jedes Feld besitzt einen im Kontext des strukturierten Datentyps eindeutigen Namen und einen Datentyp. Der Datentyp kann ein beliebiger in ADEPT2 bekannter Datentyp sein. Jedes Feld kann neben einem elementaren Datentyp auch wieder strukturiert sein. Zusätzlich können Wertebereiche der jeweils verwendeten Datentypen eingeschränkt werden [For09].

Somit kann ein Schema als eine spezielle Form eines strukturierten Datentyps abgebildet werden, der immer die drei Datenfelder für die Referenz mit einem listenwertigen Datentyp, die Aufbewahrungsfrist und die Dokumenten-ID umfasst. Die Datentypen der Referenzliste und der beiden Felder sind dabei vom konkreten DMS abhängig und individuell für jedes DMS bei der Implementierung zu ermitteln.

Da ein strukturierter Datentyp Felder enthalten kann, die selbst wieder strukturiert sind, kann damit der strukturelle Aufbau und die Definition der Datentypen für die Inhaltsfelder eines Dokuments über einen integrierten strukturierten Datentyp realisiert werden. In Abbildung 7.1 ist dies anhand eines Beispiels für ein Schema einer Rechnung verdeutlicht. Dabei werden die drei festgelegten Werte für jedes Schema mit einem strukturierten Datentyp, der den strukturellen Aufbau einer Rechnung repräsentiert, zu einem Rechnungsschema zusammengesetzt.

Durch die Verwendung von strukturierten Datentypen zur Abbildung von Dokumentenschemata des DMS können die Konzepte und die Implementierung der ADEPT2-Datenflusskonzepte weiterhin genutzt werden, ohne die vorhandenen Datenstrukturen von ADEPT2 ändern oder eine neue Datenstruktur einführen zu müssen. Durch die Verwendung eines strukturierten Datentyps kann ein Dokument aus dem DMS auch in ADEPT2 als eine Einheit abgebildet werden, in der alle Attribute und Metadaten eines Dokuments zusammengefasst sind, einschließlich der Datentypen und der Parameter zur obligaten bzw. optionalen Datenversorgung. Dadurch können Überprüfungen bzgl. der Vollständigkeit aller zu einem Schema gehörenden Datenfelder entfallen. Diese wären beispielsweise bei der Verwendung atomarer Datentypen notwendig, bei denen ein Datentyp nur ein einzelnes Datenfeld innerhalb des Schemas repräsentiert. Die Informationen welche Datentypen zu einem Schema im DMS gehören, kann damit nicht

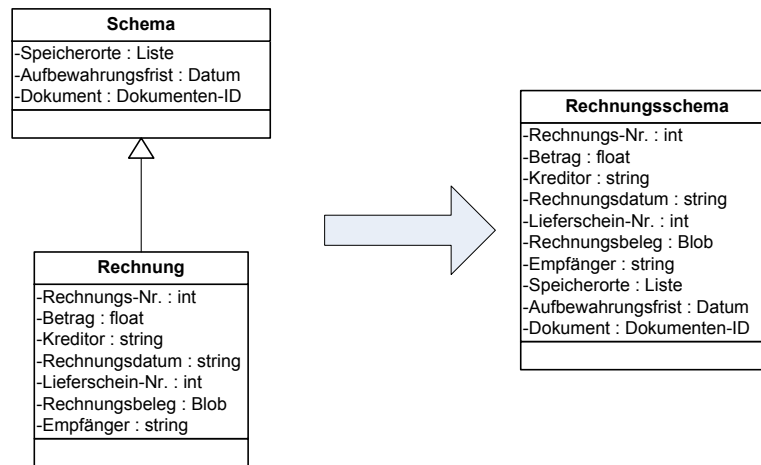


Abbildung 7.1.: Beispiel für die Erstellung eines *Rechnungsschemas* als strukturierten Datentyp

abgebildet werden, wodurch zusätzliche Prüfungen erforderlich wären, um sicherzustellen, dass immer ein vollständiges Dokument im DMS abgelegt wird.

7.1.4. Erweiterungen für Schemata

In ADEPT2 muss jedes Datenelement, auf das lesend zugegriffen wird, zuvor im Prozess durch eine Aktivität schreibend mit Daten versorgt worden sein, da beim Starten einer Prozessinstanz alle Datenelemente zunächst mit NULL-Werten belegt sind. Dies gilt auch für alle in einem strukturierten Datentyp enthaltenen Datenfelder. Hierbei müssen in jedes Datenfeld zuerst Daten geschrieben werden, bevor diese gelesen werden können. Bei der Integration von vorhandenen Dokumenten, die in einem Prozess verwendet werden, ändert sich dieses Verhalten. Hierbei werden die Inhalte der einzelnen Datenfelder nicht über Aktivitäten im Prozess geschrieben, sondern werden durch die Festlegung der Referenz auf den Speicherort und die Dokumenten-ID aus dem DMS-Repository in den strukturierten Datentyp geschrieben. Dies bedeutet, dass die Inhalte des in Abbildung 7.1 dargestellten strukturierten Datentyps (Rechnung) geschrieben werden, sobald im Schema der Speicherort und das Dokument bestimmt sind. Damit muss nicht wie bisher jedes einzelne Feld überprüft werden, ob es zuvor durch eine Aktivität mit Daten versorgt wurde, sondern lediglich, ob zuvor im Prozess das Referenz- und das Dokumenten-ID-Feld eines Schema geschrieben wurden. Durch die im Schema hinterlegten Informationen bzgl. obligater bzw. optionaler Datenfelder können die vorhandenen Konzepte zur Sicherstellung der Datenversorgung beibehalten werden.

7.1.5. Realisierung der Zugriffe auf DMS-Inhalte

Wie in der Konzeption in Kapitel 5 beschrieben, sollen Dokumente möglichst direkt und nahtlos in den Datenfluss von Prozessen integriert werden können. Da die Inhalte eines Dokuments im DMS-Repository gespeichert sind, muss das Auslesen der Informationen des Dokuments und die Übertragung in den strukturierten Datentyp im Hintergrund und transparent für die Aktivitäten erfolgen, um die externe Speicherung zu verdecken. Da Dokumente aus dem DMS-Repository nicht wie Datenelemente nur innerhalb eines Prozesses verfügbar sind, kann der exklusive Zugriff nicht wie bei anderen Datenelementen von ADEPT2 garantiert werden. Um einen exklusiven Zugriff auf DMS-Inhalte zu erhalten, sind hierzu entsprechende Zugriffssperren über den Check-Out/Check-In Mechanismus des DMS von ADEPT2 zu setzen. Hierzu können die drei möglichen Zugriffe *Lesen*, *Schreiben* und *Bearbeiten* von DMS-Inhalten unterschieden werden.

Rein lesende Zugriffe können daran erkannt werden, dass keine Schreibkante zu einem Schema im Prozess vorhanden ist. Rein schreibende Zugriffe liegen vor, wenn ein neues Dokument aus dem Prozess heraus im DMS abgelegt werden soll. Hierbei existiert keine Lesekante auf das Schema. Zugriffe zum Bearbeiten eines Dokuments sind erkennbar, wenn sowohl Lese-, als auch Schreibkanten zu einem Schema im Prozess vorhanden sind. Diese drei Zugriffsarten haben Auswirkungen auf die Check-In/Check-Out-Sperren, die von den Aktivitäten beim Zugriff auf das Dokument anzuwenden sind und im Folgenden beschrieben werden.

Lesender Zugriff

Für den lesenden Zugriff auf ein Dokument aus dem DMS-Repository genügt es, das Dokument und die im Schema zugehörigen Datenfelder über die DMS-Schnittstelle beim Starten einer Aktivität, die eine Lesekante auf das Schema besitzt, zu lesen. Hierzu wählt der Benutzer zunächst einen Speicherort aus der im Schema definierten Liste aus. Nach dieser Auswahl werden alle Dokumente dem Benutzer angezeigt, die mit dem angegebenen Schema an diesem Speicherort abgelegt sind. Sobald das konkrete Dokument ausgewählt wird, können die Inhalte des Dokuments in die Felder des strukturierten Datentyps übernommen werden. Da bei lesenden Zugriffen keine Schreibkante zum Schema im Prozess vorhanden ist, ist auch kein Check-Out des Dokuments erforderlich. Außerdem können auch ausgecheckte Dokumente mit ihrem aktuellen Inhalt gelesen werden. Die einmal gelesenen Informationen bleiben während des Prozesses in dem strukturierten Datentyp vorhanden. Somit können keine Seiteneffekte auftreten, wenn das Dokument im DMS zwischen zwei Lesevorgängen geändert wird. Die Übernahme der Inhalte aus dem DMS sind lediglich beim ersten Lesezugriff nötig, da weitere Lesezugriffe auf dem strukturierten Datentyp in ADEPT2 erfolgen.

Schreibender Zugriff

Ein rein schreibender Zugriff auf das DMS-Repository liegt vor, wenn während der Ausführung von Prozessen ein Dokument erstellt wird, das im Verlauf eines Prozesses im DMS-Repository abgelegt werden soll. Durch den im Schema definierten strukturierten Datentyp wird dabei sichergestellt, dass alle obligaten Datenfelder versorgt werden, die zur Speicherung eines Dokuments mit einem Dokumentenschema des DMS notwendig sind. Außerdem sind im Schema auch alle mit dem Schema verknüpften Speicherorte vorhanden, an denen das Dokument im DMS-Repository abgelegt werden kann. Um zu verhindern, dass Dokumente, die sich noch in der Erstellungsphase eines Prozesses befinden, bereits im DMS-Repository sichtbar werden, wird das Dokument erst beim Abschließen der Aktivität erstellt, die die letzte Schreibkante zum Schema besitzt. Da es sich um ein neu erstelltes Dokument handelt, das nicht im DMS vorhanden ist, werden hierbei keine Sperren über Check-In/Check-Out-Mechanismen benötigt. Durch diese Vorgehensweise wird ein Dokument immer in der neuesten Version im Prozess erzeugt. Die ggf. auftretenden temporären Zwischenversionen während der Prozessausführung werden nicht ins DMS-Repository übernommen.

Bearbeiten

Beim Bearbeiten von Dokumenten während der Ausführung eines Prozesses sind konkurrierende Zugriffe auf das Dokument zu berücksichtigen. Hierbei muss das Dokument zunächst wie bei einem lesenden Zugriff ausgewählt und die Inhalte in den strukturierten Datentyp des Schemas übernommen werden. Dabei muss jedoch, anders als beim rein lesenden Zugriff, beim Starten der Aktivität, die mit der ersten Lesekante auf das Schema verknüpft ist, das Dokument über die Check-Out-Funktion des DMS für eine Bearbeitung durch das DMS gesperrt werden. Falls ein Dokument ausgewählt wird, das bereits durch eine Check-Out-Sperre geschützt ist, schlägt die Aktivität fehl und informiert den Benutzer mit einer entsprechenden Fehlermeldung, dass das Dokument nicht ausgecheckt werden kann.

Bei Beendigung jeder Aktivität, die eine Änderung am Inhalt eines Schemas durchgeführt hat, muss im DMS eine neue Version des Dokuments über einen Check-In und die Übermittlung der aktuellen Version erfolgen. Anschließend wird das Dokument sofort wieder ausgecheckt, um keine externen Änderungen am Dokument zwischen zwei Aktivitäten zuzulassen. Da zwischen dem Check-In und dem Check-Out Vorgang potentiell ein Check-Out über das DMS erfolgen kann, müssen die beiden Vorgänge dabei über eine vom DMS bereitgestellte Transaktion gekapselt werden, damit beide Vorgänge auch sicher durchgeführt werden können. Wenn die Aktivität mit der letzten Schreibkante auf das Schema beendet wird, wird das Dokument über einen Check-In ohne den Check-Out-Vorgang wieder im DMS verfügbar gemacht.

Wie in den Anforderungen in Kapitel 4 beschrieben erfolgt die Bearbeitung mit den Berechtigungen des Benutzers, um die Zugriffe und Änderungen in die Audit-Trails des DMS auch direkt mit dem Benutzer zu verknüpfen. Um die Konsistenz von Dokumenten zu erhalten, darf ein Dokument nur von dem Benutzer eingecheckt werden, der es ausgecheckt hat. Da jedoch zwei Aktivitäten nicht zwingend vom gleichen Benutzer ausgeführt werden, muss die Check-Out-Sperre, die dem Schutz externer Änderungen zwischen zwei Aktivitäten dient, von einem speziellen Benutzeraccount gehalten werden, der explizit hierfür zu verwenden ist. Sobald klar ist, welcher Benutzer die nächste Änderung am Dokument vornimmt, kann der Check-Out-Vorgang abgebrochen werden und mit den Anmeldeinformationen des eigentlichen Benutzers wieder ausgecheckt werden. Dieser Vorgang ist ebenso über eine Transaktion im DMS zu kapseln.

Neben lebenden Dokumenten werden über die Ablagestruktur des DMS auch als Records klassifizierte Dokumente verfügbar gemacht, die nicht verändert werden können. Diese werden vom DMS über ein Attribut realisiert, das dem Dokument zugeordnet ist. Da solche Dokument nicht bearbeitet werden können, ist es sinnvoll, diese bei der Auswahl der vorhandenen Dokumente dem Benutzer nicht anzuzeigen, da eine Bearbeitung eines solchen Dokuments zwangsläufig zu einem Laufzeitfehler führen würde. Dazu überprüft die Aktivität des ersten Lesezugriffs jedes Dokument innerhalb des gewählten Speicherortes auf ein gesetztes „Record-Attribut“.

Durch diese Vorgehensweise kann die Anforderung nach exklusivem Zugriff innerhalb eines Prozesses sichergestellt werden, um die Korrektheit weiterhin zu gewährleisten. Ein vorhandenes Dokument bleibt weiterhin im DMS sichtbar und steht für lesende Zugriffe zur Verfügung. Das Dokument wird bei langlaufenden Prozessen nur solange vor dem Zugriff durch andere Komponenten gesperrt, wie es die Bearbeitung im Prozess erfordert.

7.1.6. Erstellen der Schemata

Für die Erstellung eines Schemas in ADEPT2 wird ein strukturierter Datentyp und die drei beschriebenen Datenfelder für die Referenz, das Dokument und die Aufbewahrungsfrist benötigt.

Jedes Schema benötigt mindestens einen Speicherort im DMS. Die Speicherorte können dabei über einen Verzeichnisbaum aus dem DMS ausgelesen und ausgewählt werden. Die Speicherorte können dann über eine Liste mit den Referenzen repräsentiert werden, bei der jeder Eintrag ein Verzeichnis im DMS-Repository darstellt. Außerdem benötigt ein Schema zusätzlich ein Feld zur eindeutigen Identifikation des konkreten Dokuments. Dieses Feld wird bei lesendem Zugriff mit der eindeutigen Dokumentenbezeichnung im DMS zur Laufzeit versehen. Wird das Schema für die Erstellung eines

neuen Dokuments verwendet, das im DMS abgelegt werden soll, so wird hierzu zur Laufzeit eine Bezeichnung festgelegt.

Damit besteht jedes Schema aus einem strukturierten Datentyp, einer Liste mit Referenzen auf die Speicherorte, einem Feld für die Aufbewahrungsfrist und einem Feld, das die konkrete Identifikation des Dokuments innerhalb des Speicherortes enthält. Alle genannten Datenfelder sind immer obligat.

Der Aufbau des strukturierten Datentyps richtet sich dabei sowohl nach den Möglichkeiten von ADEPT2, als auch nach den unterstützten Strukturen des DMS. Über Konfigurationsoptionen muss hierzu die Mächtigkeit der strukturierten Datentypen, die in einem Schema eingebettet sind, so eingeschränkt werden können, dass nur für beide Systeme gültige Datenstrukturen erstellbar sind. Um sicherzustellen, dass die in ADEPT2 definierten Schemata auch vom DMS genauso abgebildet werden können, müssen die Schemata DMS-spezifisch angepasst werden. Bei der Implementierung der Schemata ist zu ermitteln, welche Datentypen, Datenstrukturen und Einschränkungen von Wertebereichen, z. B. über reguläre Ausdrücke, für einzelne Felder vom DMS abgebildet werden können. Hierzu müssen die ADEPT2-Schemata durch Konfigurationsoptionen anpassbar sein, so dass nur DMS-konforme Schemata erstellt werden können. Dies können z. B. die Entfernung bestimmter nicht unterstützter Datentypen oder die Beschränkung von Wertebereichen für einzelne Felder sein.

Übernahme eines Schemas aus dem DMS

Die zweite Möglichkeit, um ein Schema in ADEPT2 zu erstellen, ist die Übernahme eines vorhandenen Schemas aus dem DMS. Hierzu werden die im DMS vorhandenen Schemata aus dem DMS gelesen und es wird anhand der Struktur ein strukturierter Datentyp erzeugt, der die Metadatenfelder und den Aufbau der Dokumente repräsentiert. Die Informationen über Datentypen, Einschränkungen von Wertebereichen, optionale bzw. obligate Felder sind vollständig in einem Dokumentenschema enthalten und müssen bei der Erzeugung des strukturierten Datentyps aus dem DMS übernommen werden. Da die Aufbewahrungsfristen normalerweise mit dem Dokumentenschema verknüpft sind, wird diese Information ebenfalls aus dem DMS übernommen.

Die bereits mit dem Schema verknüpften Speicherorte lassen sich dabei entweder über das Dokumentenschema, oder über die Verzeichnisstruktur des DMS-Repositories ermitteln.

Lassen sich im DMS über Schemata Datenstrukturen erstellen, die sich nicht mit den strukturierten Datentypen von ADEPT2 direkt abbilden lassen, können diese zunächst nicht von ADEPT2 verwendet werden. Um die Schemata nutzen zu können, besteht die Möglichkeit, dafür spezielle Konverter zu implementieren, welche die Transforma-

tion der Daten vornehmen. Diese können dann z. B. als Vorschaltdienst einer Aktivität [Rei00] mit Zugriff auf das betroffene Schema verwendet werden.

7.1.7. Verwalten von Schemata

Da Schemata in ADEPT2 mit Hilfe von strukturierten Datentypen repräsentiert werden, können diese ebenso über die Datentypverwaltung von ADEPT2 realisiert werden [For09]. Die Verwaltungskomponente stellt hierzu eine Übersicht über alle vorhandenen Schemata und deren Verwendung in Prozessen dar. Diese Informationen werden für das Löschen und Änderungen von Schemata benötigt. Schemata dürfen nur gelöscht werden, wenn sie in keinem Prozess verwendet werden. Falls bereits Instanzen von Prozessen mit einem Schema gestartet wurden, darf dieses nur aus der Anzeige entfernt werden, da ansonsten abgeschlossene Prozessinstanzen ungültig würden. Das zugehörige Dokumentenschema im DMS bleibt davon unberührt, da mit einem Dokumentenschema gespeicherte Dokumente weiterhin im DMS gespeichert bleiben.

Durch geänderte Anforderungen an die Datenstruktur, können sich Änderungen am Schema der strukturierten Datentypen ergeben. Dazu sind die in Kapitel 5 beschriebenen Einschränkungen zu berücksichtigen, um sicherzustellen, dass keine Konflikte mit vorhanden Dokumenten oder Records erzeugt werden, die mit dem Schema im DMS abgelegt sind. Diese Einschränkungen sind:

- Dokumentenattribute dürfen nicht entfernt werden
- neue Felder dürfen nur optional hinzugefügt werden
- obligate Felder dürfen nicht entfernt werden

Zulässige Änderungen an Schemata, die von ADEPT2 genutzt werden, müssen auch über ADEPT2 erfolgen, um sicherzustellen, dass die Änderung in beiden Systemen umgesetzt wird. Ansonsten wären Laufzeitfehler beim Zugriff auf Dokumente oder Records während der Prozessausführung die Folge.

Schemata, die nicht in ADEPT2 verwendet werden, können dabei weiterhin über die DMS-Komponenten verwaltet werden, da sie keinen direkten Einfluss auf die Prozesse haben.

Falls Änderungen erforderlich sind, die nicht innerhalb der genannten Einschränkungen durchführbar sind, muss ein neues Schema erstellt werden, das die Änderungen enthält. Dieses neue Schema kann dabei durchaus wieder mit denselben Speicherorten, wie das alte Schema verknüpft werden, um Dokumente sowohl nach dem alten, als auch nach dem neuen Schema in derselben Ablagestruktur verwalten zu können.

7.1.8. Zugriffsberechtigungen ermitteln und zuweisen

Wenn die Prozessmodellierung und alle Bearbeiterzuweisungen erfolgt sind, müssen für die korrekte Ausführung von Prozessen die Zugriffsberechtigungen auf die Speicherorte im DMS-Repository angegeben werden. Dazu ist für ADEPT2 eine Erweiterung zu entwickeln, die am Ende der Prozessmodellierung alle Benutzer ermittelt und die dafür notwendigen Berechtigungen im DMS für die Benutzer setzt.

Dazu sind alle in Frage kommenden Benutzer zu ermitteln, die eine Aktivität potentiell ausführen können, die über eine Datenkante mit einem Schema verbunden ist. Aus der im Schema gespeicherten Liste der Referenzen zur Ablage im DMS-Repository ergeben sich Speicherorte aus dem verknüpften Schema. Aus der Richtung der Datenkante folgt, ob der Zugriff lesend oder schreibend stattfindet.

Um die konkreten Benutzerobjekte zu ermitteln, die den Aktivitäten zugeordnet sind, werden die in [Ber05] beschriebenen Formeln der Bearbeiterzuordnung aufgelöst, um alle Benutzer zu ermitteln, die die Aktivität ausführen können. Die so ermittelten Benutzer werden dann, wie in der Konzeption dargestellt, in zwei Gruppen für lesenden und schreibenden Zugriff eingeteilt. Ein schreibender Zugriff beinhaltet dabei auch lesenden Zugriff. Wenn beispielsweise Benutzer A potentielle Aktivität 1 und Aktivität 2 ausführen kann, die beide dasselbe Schema und den denselben Speicherort referenzieren, und sich aus Aktivität 1 ein lesender und aus Aktivität 2 ein schreibender Zugriff ergibt, so ist Benutzer A in die Gruppe für den schreibenden Zugriff einzuordnen.

Mit diesen Informationen können anschließend die Berechtigungen im DMS eingerichtet werden. Dazu wird der Prozess einmal durchlaufen und dabei die Informationen der Speicherorte im DMS, sowie der Benutzerobjekte mit lesendem oder schreibenden Zugriff ermittelt. Auf Basis der ermittelten Informationen können die Berechtigungen von ADEPT2 aus, über die Schnittstelle zum DMS eingerichtet werden, indem die Gruppen für lesenden und schreibenden Zugriff im DMS angelegt, die ermittelten Benutzerobjekte den Gruppen zugewiesen und abschließend die Berechtigungen an den Speicherorten im DMS entsprechend gesetzt werden. In der Bezeichnung der Gruppe ist dabei ein Name zu verwenden, der die Prozessvorlage eindeutig identifiziert, um im DMS die Gruppen und die damit verbundenen Zugriffsberechtigungen Prozessen zuordnen zu können.

Da die Bearbeiterzuordnungen bei der Ausführung von Prozessen erst zur Laufzeit aufgelöst werden, kann es vorkommen, dass zwischen dem Anlegen der Gruppen im DMS und der Ausführung einer Aktivität, ein nicht berücksichtigter Benutzer die Aktivität mit einem Zugriff auf ein Schema ausführt. Diese Vorgehensweise ist in ADEPT2 beabsichtigt, um Änderungen im Organisationsmodell zu berücksichtigen. Um dem Problem zu begegnen, müsste in solch einem Fall die Zuordnung des ADEPT2-Benutzers zu den erwähnten DMS-Gruppen zur Laufzeit erfolgen. Wenn es sich bei dem Benutzer beispielsweise um einen neu hinzugekommenen Benutzer handelt, kann es weiterhin

vorkommen, dass der Benutzer im DMS noch nicht bekannt ist. Dies resultiert daraus, dass die Benutzer wie in Kapitel 5 lediglich in bestimmten Abständen aus dem Verzeichnisdienst ins DMS übernommen werden. Damit würde jedoch auch eine automatisierte späte Zuordnung von ADEPT2 zur Laufzeit scheitern. Mögliche Lösungen dieses Problems wären eine manuelle Zuweisung des Benutzers zu den Gruppen im DMS oder die erneute Ausführung der automatischen Zuweisung aus ADEPT2, bei der die aktuellen Benutzer und Bearbeiterzuordnungen erneut ausgewertet werden.

Da die konkret unterstützten Funktionen zur Zuweisung von Berechtigungen und der Aufbau des DMS-Repository bei jedem DMS unterschiedlich sind, müssen diese Informationen bei der Implementierung der Schnittstelle zum DMS individuell implementiert werden.

Falls die beiden Benutzerverwaltungen der Systeme nicht integriert wurden, müssen die Berechtigungen im DMS manuell auf Basis der ermittelten Benutzer eingerichtet werden. Dazu sollte die Komponente zur Ermittlung der Benutzer diese in einer Liste mit den erforderlichen Berechtigungen von ADEPT2 ausgeben. Anhand dieser Liste können dann die Berechtigungen im DMS überprüft werden und ggf. die notwendigen Berechtigungen erstellt werden.

7.2. Organisationsintegration

In diesem Abschnitt werden die drei in der Konzeption vorgestellten Integrationsvarianten der Organisationsaspekte für ADEPT2 behandelt. Zuerst wird die Variante der getrennten Benutzerverwaltungen vorgestellt. Anschließend wird auf die Integration über Verzeichnisdienste eingegangen. Hierzu wird zunächst beschrieben, wie Benutzerobjekte aus einem LDAP-Verzeichnisdienst ermittelt werden können, um darüber eindeutige Benutzerzuordnungen zwischen den LDAP-Benutzerobjekten und den Benutzerobjekten im ADEPT2-Organisationsmodell herzustellen. Darauf aufbauend werden die Authentifizierungsvarianten mittels Single Credential und SSO betrachtet.

7.2.1. Getrennte Benutzerverwaltungen

Für den Fall, dass sich zwischen DMS und ADEPT2 keine gemeinsame Basis für die Benutzer etablieren lässt, muss die Integration über getrennte Benutzerverwaltungen realisiert werden. Hierbei können keine Beziehungen zwischen den Benutzerobjekten der beiden Systeme hergestellt werden.

Bei der Ausführung einer Aktivität, die auf ein Schema zugreift, müssen zuvor Benutzername und Passwort vom Benutzer erfragt werden. Diese Informationen werden dann beim Zugriff auf das DMS zur Authentifizierung von ADEPT2 im Hintergrund

an das DMS gesendet. Schlägt die Authentifizierung fehl, schlägt auch die Ausführung der Aktivität fehl. Nachteilig wirken sich jedoch bei dieser Variante die häufigen Authentifizierungsvorgänge auf die Produktivität des Benutzers aus. Außerdem sind die mehrfachen Eingaben von Benutzername und Passwort auch sehr fehleranfällig. Zudem kann nicht sichergestellt werden, dass der in ADEPT2 angemeldete Benutzer auch wirklich dem Benutzer im DMS entspricht, falls der Benutzer bei der Ausführung andere Anmeldeinformationen für das DMS angibt.

7.2.2. Synchronisation über Verzeichnisdienst

Um eine Zuordnung der Benutzer zwischen DMS und ADEPT2 herstellen zu können, werden, wie in Kapitel 5 beschrieben, die Benutzerobjekte mit einem Verzeichnisdienst synchronisiert. Sowohl ADEPT2, als auch sämtliche in Kapitel 3 untersuchten Systeme unterstützen die Anbindung an mindestens einen LDAP-basierten Verzeichnisdienst.

Für den Import der Benutzerobjekte aus einem Verzeichnisdienst werden dabei die in Tabelle 7.1 angegebenen Informationen benötigt, um den Verzeichnisdienst ansprechen zu können und die Benutzerobjekte zu importieren.

Die Informationen zu Server, Benutzername und Passwort dienen dabei dem Verbindungsaufbau zu einem LDAP-Verzeichnisdienst. Über die Suchbasis werden ein oder mehrere Einträge im hierarchisch aufgebauten DIT bestimmt, ab denen absteigend nach Benutzerobjekten, die importiert werden, gesucht werden soll. Die in LDAP gespeicherten Objekte werden anhand ihrer *objectClass* identifiziert. Hierzu kann z. B. die *objectClass* Person angegeben werden, um nur Objekte zu erhalten, die diese Eigenschaft besitzen. Über die UUID der Benutzerobjekte aus dem Verzeichnisdienst können die beiden Systeme den entsprechenden Benutzer jeweils eindeutig identifizieren. Über das *userField* wird der Anmeldename des Benutzers ermittelt. Die letzten beiden Informationen sind dabei abhängig vom verwendeten Verzeichnisdienst. Die genaue Bezeichnung lässt sich aus der Dokumentation des Verzeichnisdienstes entnehmen.

Durch den Import der Benutzerobjekte aus dem Verzeichnisdienst, können die Benutzer im DMS und in ADEPT2 im jeweiligen Organisationsmodell verwaltet werden. Die spezifischen Modellierungen der Organisation wie Gruppen, Rollen oder Stellen werden dabei innerhalb der Systeme selbst verwaltet, um die systemspezifischen Eigenschaften in vollem Umfang nutzen zu können. Ebenso erfolgt die Zuweisung der Berechtigungen in den Systemen getrennt. Eine Ausnahme stellt die zuvor beschriebene Zuweisung der Berechtigungen für den lesenden bzw. schreibenden Zugriff auf die DMS-Inhalte dar, die bei der Prozessmodellierung ermittelt werden und von ADEPT2 im DMS eingetragen werden.

Tabelle 7.1.: Verzeichnisdienstinformationen

Eigenschaft	Beschreibung
LDAP-Server	Protokoll, IP-Adresse und Port des Verzeichnisdienst-Servers, z. B. ldap://192.168.0.1:389
Benutzername und Passwort	Benutzername und Passwort eines Benutzers im Verzeichnisdienst, der darin eine Suche absetzen darf
UUID Attribute	LDAP-Attribut, das die UUID des Benutzers zur eindeutigen Identifikation enthält, z. B. objectGUID bei Active Directory, entryGUID bei openLDAP
Suchbasis	Basispfad, ab dem im Verzeichnisdienst nach Benutzern gesucht werden soll. Hierfür sind ggf. mehrere Pfade einzutragen, z. B. DC=DOM,DC=DE
objectClassUser	Name einer LDAP-Objektklasse, die LDAP-Benutzer beschreibt, z. B. objectClass = Person, objectClass = user
UUID Attribute Type	Datentyp der LDAP-Attributs, das die UUID enthält, z. B. Byte bei Active Directory, String bei openLDAP
userField	LDAP-Attribut, in dem der Anmeldename enthalten ist, z. B. sAMAccountName bei Active Directory. Dieser ist gleichzeitig der Benutzername des Benutzers

7.2.3. Authentifizierung mit Single Credential und Single Sign On

Um einem Benutzer zu ermöglichen, sich an beiden Systemen mit den gleichen Anmeldeinformationen zu authentisieren, können sowohl das DMS, als auch ADEPT2 die Authentifizierung an den Verzeichnisdienst delegieren. Um die Authentifizierung an den Verzeichnisdienst zu delegieren, muss der Authentifizierungsdienst von ADEPT2 in der Form erweitert werden, dass hierzu das Passwort nicht direkt im Organisationsmodell hinterlegt wird, sondern bei jeder Anmeldung über den Verzeichnisdienst abgefragt wird.

Um sowohl LDAP-Benutzerobjekte, als auch ADEPT2-Benutzer gleichzeitig zu unterstützen, kann der Ablauf des Anmeldevorgangs folgendermaßen vorgenommen werden. Bei der Anmeldung prüft ADEPT2, ob der Benutzername in der Benutzerverwaltung bekannt ist. Wenn ja, wird das eingegebene Passwort in der ADEPT2-Benutzerdatenbank überprüft und der Benutzer wird angemeldet. Wenn nein, prüft das System, ob die Authentifizierung für den Benutzer an einen Verzeichnisdienst delegiert wurde. Ist dies der Fall, werden Benutzername und Passwort an den Authentifizierungsdienst

des Verzeichnisdienstes gesendet und bei erfolgreicher Antwort wird der Benutzer angemeldet.

Wenn sowohl das DMS, als auch ADEPT2 jedoch die Authentifizierung an den Verzeichnisdienst delegieren, steht die Funktionalität des Single Credential zur Verfügung. Da Verzeichnisdienste i. d. R. mehrere Authentifizierungsmethoden unterstützen (s. Kapitel 5), müssen die beiden Systeme nicht die gleiche Authentifizierungsmethode verwenden.

JAAS

Da ADEPT2 auf Java basiert, bietet es sich hierzu an, den in Java integrierten Authentifizierungs- und Autorisierungsdienst JAAS¹ zu verwenden [SUN01b]. JAAS implementiert dazu eine Java-Version des PAM²-Frameworks, um eine standardisierte Schnittstelle zur Authentifizierung und Autorisierung von Benutzern bereitzustellen. Da die Autorisierung aufgrund der unterschiedlichen Organisationsmodelle den jeweiligen Systemen überlassen wird, wird im Folgenden lediglich auf die Authentifizierungsaspekte eingegangen.

JAAS stellt verschiedene Module für die Authentifizierung bereit, um Authentifizierungsdienste z. B. von Unix, Windows, LDAP oder Kerberos zu verwenden. Dadurch bleiben die Anwendungen unabhängig von den zugrunde liegenden Technologien zur Authentifizierung. Bei Änderungen an der zu verwendenden Authentifizierungsmethode muss somit lediglich die Konfiguration angepasst werden. Die Implementierung der Benutzeranmeldung kann unverändert bleiben.

Die Kernklassen von JAAS lassen sich in drei Kategorien aufteilen: allgemein, Authentifizierung und Autorisierung [FSR06]. Da die Autorisierung in dieser Arbeit nicht relevant ist, wird auf diese Klassen im Folgenden nicht weiter eingegangen.

Die allgemeinen Klassen umfassen:

- `Subject (javax.security.auth.subject)`
- `Principal (java.security.Principal)`

Die Authentifizierungsklassen umfassen:

- `Callback (javax.security.auth.callback.Callback)`
- `CallbackHandler (javax.security.auth.callback.CallbackHandler)`

¹Java™ Authentication and Authorization Service

²Pluggable Authentication Module

- `Configuration` (`javax.security.auth.login.configuration`)
- `LoginContext` (`javax.security.auth.login.LoginContext`)
- `LoginModule` (`javax.security.auth.login.LoginModule`)

Das *Subject* steht in JAAS für die Information einer einzigen Entität, z. B. ein Benutzer oder ein Dienst. Das Subject umfasst dabei die *Principals*, die die öffentlichen und privaten Anmeldeinformationen eines Subjects darstellen. Die JAAS-API verwendet dazu das existierende Java-Interface `java.security.Principal`. Während des Authentifizierungsprozesses werden in einem Subject die zugehörigen Principals gespeichert. Ein Subject kann mehrere Principals besitzen, die dazu dienen ein Subject zu identifizieren, z. B. den Namen des Benutzers. Weiterhin kann ein Subject auch sicherheitsrelevante Attribute besitzen, die als *Credentials* im Subject abgelegt werden. Diese sind in öffentliche Credentials (z. B. öffentliche Schlüssel von Zertifikaten oder Kerberos Server Tickets) und private Credentials (z. B. private Schlüssel des Benutzers) unterteilt. Um diese Informationen aus einem Subject abzurufen stehen die Methoden `getPrincipals()`, `getPrivateCredentials()` und `getPublicCredentials()` zur Verfügung [SUN01b].

Authentifizierung eines Subjects

Die Authentifizierung eines Subjects erfordert eine JAAS-Anmeldung, die aus den folgenden Schritten besteht:

1. Instanziierung des `LoginContext` und Übergabe des Namens der Anmeldekonfiguration, sowie einen `CallbackHandler`, um die in der Konfiguration des `LoginModule` erforderlichen `Callback`-Objekte zu füllen.
2. Der `LoginContext` lädt alle in der `Configuration` angegebenen `LoginModule`.
3. Die Methode `LoginContext.login` muss aufgerufen werden.
4. Die Anmeldemethode ruft alle geladenen `LoginModule` auf. Versucht ein `LoginModule` sich beim Subject zu authentifizieren, ruft es die `Handle`-Methode des verknüpften `CallbackHandler` auf. Bei erfolgreicher Durchführung verknüpfen die `LoginModule` die benötigten Principals und Credentials mit dem Subject.
5. Der `LoginContext` liefert den Authentifizierungsstatus zurück. Schlägt die Authentifizierung fehl, wird eine `LoginException` ausgelöst.
6. Bei einer erfolgreichen Authentifizierung kann auf das authentifizierte Subject durch `LoginContext.getSubject` zugegriffen werden.

7. Durch Aufruf von `LoginContext.logout` werden die Verknüpfungen zwischen Subject, Principals und Credentials entfernt.

Durch die Klasse `LoginContext` werden die Methoden für die Authentifizierung der Subjects bereitgestellt. Um nun eine Anmeldung ohne Eingabe der Anmeldeinformationen des Benutzer am Beispiel von Kerberos zu ermöglichen, muss dafür als `LoginModule` die Klasse `Krb5LoginModule` verwendet werden. In der zugehörigen Konfigurationsdatei zu diesem Modul muss dazu die Option `useTicketCache=true` angegeben werden. Durch diese Option ermittelt das `Krb5LoginModule` das Kerberos Ticket Granting Ticket aus dem Standardspeicherort des verwendeten Betriebssystems. Optional kann der Speicherort des Tickets über die Option `ticketCache` in der Konfiguration angegeben werden. Falls kein Ticket ermittelt werden kann, wird der Benutzer zur Eingabe von Benutzername und Passwort aufgefordert. Wenn sichergestellt ist, dass das Kerberos-Ticket immer verfügbar ist, kann der `CallbackHandler` NULL sein, da dieser nur der Interaktion mit dem Benutzer dient [SUN01a]. Der Ablauf einer SSO-Authentifizierung mittels Kerberos und JAAS ist in Abbildung 7.2 graphisch dargestellt.

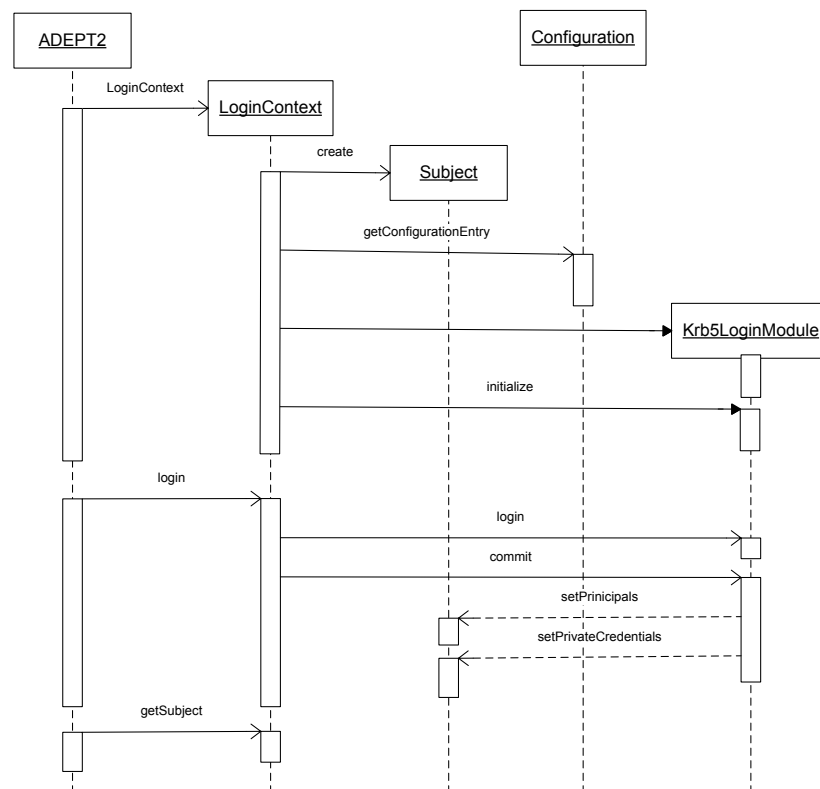


Abbildung 7.2.: Ablauf einer Authentifizierung mit Kerberos nach [SS02]

Um beide Systeme weiter zu integrieren und mehrfache Authentifizierungsvorgänge zu vermeiden, können beide Systeme die Benutzer über einen SSO-Dienst authentifizieren. Der zu verwendende SSO-Dienst muss dabei von beiden Systemen unterstützt werden und kann z. B. von einigen Verzeichnisdienstimplementierungen wie Active Directory bereitgestellt werden.

Wie in Kapitel 5 beschrieben, muss die Authentifizierung des Benutzers bei Verwendung von SSO mit Kerberos durch eine Delegation realisiert werden. Um dies zu realisieren sieht JAAS eine Credentials Delegation vor, mit der der ADEPT2-Server in die Lage versetzt wird, Tickets vom KDC für einen Client zu erhalten, um darüber die Anmeldung am DMS vorzunehmen. Die Verwendung der dazu erforderlichen Tickets mit einem gesetzten *Forwardable*-Flag können über die Klasse `javax.security.auth.kerberos.DelegationPermission` verwendet werden.

7.3. Clientintegration

Um dem Benutzer einen einheitlichen Client für integriertes Dokumenten- und Prozessmanagement zu bieten, können die Clients von ADEPT2 und des DMS-Clients teilweise oder vollständig in einen Client integriert werden. Aufgrund der sehr heterogenen Anforderungen und der sehr unterschiedlichen Implementierungen, sowie der Vielzahl an möglichen Clients, die im Bereich DMS vorzufinden sind, werden an dieser Stelle lediglich allgemeine Integrationsmöglichkeiten aufgezeigt. Diese werden anhand der von ADEPT2 genutzten Client-Techniken erläutert. ADEPT2 besitzt sowohl einen Rich-Client, der auf der Eclipse RCP³ [SDF⁺03] aufsetzt, als auch einen Web-Client, der auf der Eclipse RAP⁴ basiert [Ben08].

7.3.1. Rich-Client

Bei der Rich-Client-Integration werden Teile des DMS- und des ADEPT2-Clients direkt auf dem Rechner des Benutzers gekoppelt. Sie können zum einen der Präsentation beider Programme in einer Anwendung für den Benutzer dienen, ohne dass Kontextwechsel der Programme nötig sind, zum anderen können hierüber einzelne Funktionen des einen Clients in den anderen übernommen werden. Für die Realisierung der Client-Integration existieren mehrere Techniken, die zum einen von der Implementierung der konkret zu integrierenden Clients abhängt, zum anderen von den zu integrierenden Funktionen. Ebenso müssen auch die Architekturen der Clients kompatibel sein, um auf technischer Ebene eine kompatible Schnittstelle verwenden zu können. Weiterhin haben sich auch kommerzielle Anbieter wie z. B. JNBridge [JNB09] etabliert, um spezi-

³Rich Client Platform

⁴Rich Ajax Platform

ell die Interoperabilität von .NET- und Java-Programmen zu ermöglichen. Da ADEPT2 einen Rich-Client auf Eclipse-Basis verwendet, wird im Folgenden von diesem als Basis ausgegangen.

RCP – RCP

Der einfachste Fall einer Clientintegration liegt vor, wenn der DMS-Client ebenfalls auf Eclipse basiert, wie z. B. der Easy Experience-Client aus Kapitel 3. Eigene Anwendungen oder Module werden in den Eclipse RCP-Client über *Plugins* realisiert. Da beide Clients in diesem Fall auf der RCP aufsetzen, können hier die jeweiligen Plugins zusammen in einen Eclipse-Client geladen und angeordnet werden. Die Kommunikation zwischen den Komponenten kann dabei über die herstellereigenspezifische API erfolgen.

RCP – Browser

Manche DMS wie MOSS oder Alfresco verzichten komplett auf einen Rich-Client. Die Bedienung erfolgt ausschließlich über den Browser. Hierbei besteht in Eclipse die Möglichkeit einen integrierten Browser innerhalb des Eclipse-Clients zu öffnen und darüber die Inhalte und Funktionen dem Benutzer zu präsentieren. Um nur bestimmte Komponenten zu integrieren, kann an den Browser die URL⁵ übergeben werden, die die Komponente auf dem Webserver identifiziert. Eine weitere Möglichkeit zur Integration von einzelnen Funktionen besteht darin, eine RSS-View in den ADEPT2-Client zu integrieren, über die der Benutzer über aktuelle Meldungen informiert werden kann. Sowohl MOSS, als auch Alfresco bieten die Möglichkeiten Inhalte über RSS-Feeds bereitzustellen. Diese Integrationsvariante kann genauso für die Web-Clients der anderen DMS angewendet werden.

RCP – native Clients

Für die Integration von anderen Rich-Clients, die nicht auf Java basieren, wie z. B. bei ELOprofessional oder DocuWare, können die Clients durch die Verwendung von *Wrappern* integriert werden. Ein Wrapper kapselt dabei den einzubindenden Client und führt ihn innerhalb der JVM⁶ aus. So kann über die JNI⁷-API von Java beispielsweise ein in C oder C++ erstelltes Programm eingebunden werden [Gor98]. Weitergehend ermöglicht JNI auch die Einbindung einzelner Bibliotheken des DMS-Clients. Um Funktionen aus einer anderen Programmiersprache zu verwenden, muss der Code zu

⁵Uniform Resource Locator

⁶Java Virtual Machine

⁷Java Native Interface

einer dynamisch ladbaren Bibliothek gebunden werden. Um die Funktionen dabei von Java aus aufrufen zu können, müssen dazu die *native*-Methoden implementiert werden. Für diese *native*-Methoden müssen entsprechende Header-Dateien erstellt werden. Die Header-Dateien können über den Aufruf `javah -jni <Klasse>` erzeugt werden. Die erzeugte Header-Datei muss anschließend in den Quellcode der zu integrierenden Bibliothek eingebunden werden. Daraus kann dann die Bibliothek übersetzt werden, um sie in das Java-Programm zu integrieren. Da jedoch nur selten der Quellcode des DMS-Clients zur Verfügung steht, kann diese Variante der Integration in den seltensten Fällen zum Einsatz kommen.

Unter Windows kann für die Integration die zur Verfügung gestellte COM-Schnittstelle zur Kommunikation mit dem DMS-Client verwendet werden. Eclipse bietet die Möglichkeit über das *Package* `org.eclipse.swt.ole.win32` einen *Wrapper* bereitzustellen, mit dem Anwendungen über COM und OLE integriert werden können, die in Views oder Editoren in Eclipse angezeigt werden. Dabei wird die Anwendung über einen COM-Container integriert, über den die Interaktion mit der Anwendung gesteuert wird [SDF⁺03].

RCP – .NET

Eine enge Integration der Clients sind mit den oben vorgestellten Ansätzen nur bedingt möglich, da hierbei typischerweise nur ganze Module oder der komplette Client in die Oberfläche integriert werden können. Die Integration über JNI und die Einbindung externer Bibliotheken gestaltet sich darüber hinaus auch sehr aufwendig. Für die Integration von .NET und Java, sowie die Möglichkeit flexibel einzelne Funktionen in den jeweils anderen Client zu integrieren, existieren kommerzielle Anbieter wie JNBridge. Mit JNBridge können Java-Klassen direkt .NET-Klassen aufrufen, die in C#, C++ oder Visual Basic implementiert sind. Die andere Richtung von .NET zu Java ist ebenso möglich [JNB08]. Die Architektur von JNBridge ist in Abbildung 7.3 dargestellt.

JNBridge realisiert die Zugriffe über *Proxy-Klassen*. Die .NET- bzw. Java-Objekte verbleiben dabei im gleichen Betriebssystemprozess auf dem Rechner und kommunizieren innerhalb des Prozesses über eine *shared-memory*-Kommunikation. Dadurch können die .NET CLR⁸ und die JVM innerhalb eines Prozesses ausgeführt werden. Hierzu werden beim Start der Anwendung beide Laufzeitumgebungen gestartet. Wenn die Anwendung z. B. eine Java-Anwendung darstellt, wird die .NET CLR automatisch innerhalb des Java-Prozesses gestartet. Weitere Kommunikationsmöglichkeiten werden über direkte TCP/IP Kommunikation oder SOAP-Nachrichten unterstützt, die jedoch für die direkte Kommunikation der Clients auf einem Computer nicht relevant sind.

⁸Common Language Runtime

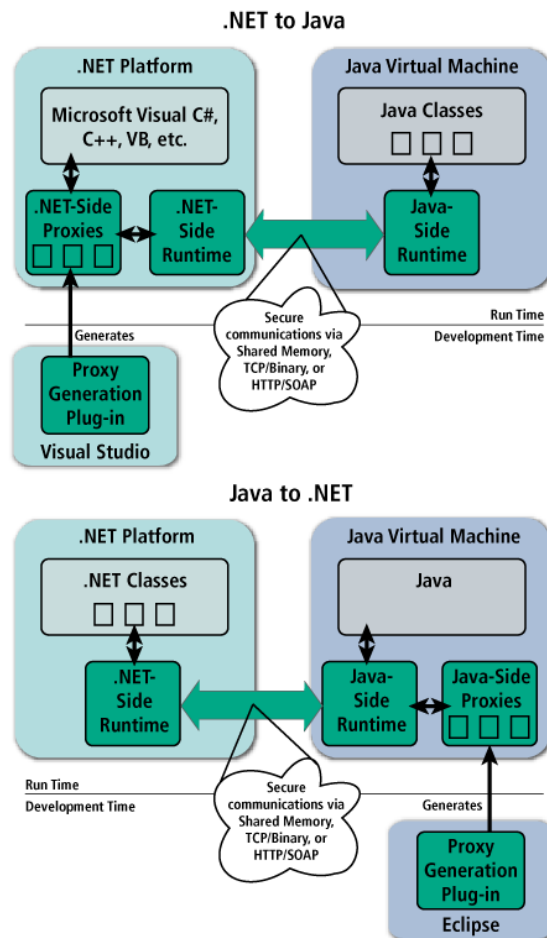


Abbildung 7.3.: JNBridge-Architektur [JNB08]

Um die Proxies zu erstellen, die die Zugriffe von .NET-Klassen aus Java heraus ermöglichen, wird ein Eclipse-Plugin bereitgestellt, mit dem die benötigten .NET-Klassen ausgewählt werden, die integriert werden sollen. Verschachtelte Klassen und Abhängigkeiten werden dabei von JNBridge berücksichtigt und mit eingebunden. Aus diesen ausgewählten Klassen erstellt JNBridge die entsprechenden .NET *assembly* DLL-Dateien bzw. Java-Klassen.

Nachdem die Proxies erstellt und eingebunden wurden, kann der Entwickler direkt die darin enthaltenen Klassen aufrufen. Dabei wird aus dem dahinter liegenden .NET-Code beim Aufruf ein Konstruktor aufgerufen, der ein entsprechendes Java-Objekt erzeugt. Beim Aufruf einer Methode wird durch die Proxy-Klasse die entsprechende .NET-Methode aufgerufen. Vererbung und *Interfaces* können ebenso verwendet werden, indem beispielsweise die Interfaces der Proxy-Klasse implementiert werden.

Bei Methodenaufrufen werden die Datentypen durch die Proxy-Klasse in die jeweils entsprechenden Datentypen konvertiert, z. B. werden native Java-Arrays automatisch in native .NET-Arrays konvertiert. Es können auch Projekte erstellt werden, die sowohl Aufrufe von .NET zu Java, als auch Java zu .NET besitzen.

Für Oberflächen können dabei Java *User Interface*-Elemente in .NET *Windows Forms* verwendet werden und *Windows Forms* in der Java-Oberfläche.

7.3.2. Web-Client

Für die Integration der Web-Clients bietet sich der Einsatz von Portalen an. Dabei kann der ADEPT2-Web-Client in ein vom DMS bereitgestelltes Portal integriert werden. Alternativ können die Web-Clients in einem anderen Portal zusammengefasst werden, falls das DMS kein Portal bereitstellt. Über die gemeinsame Darstellung in einem Portal kann der Benutzer dadurch beide Clients innerhalb eines Browsers bedienen und erhält somit eine ganzheitliche Ansicht der beiden Clients. Idealerweise ist das Portal dabei mit dem DMS und ADEPT2 über einen SSO-Dienst verbunden, so dass der Benutzer nur einmal authentifiziert werden muss.

Zur Integration in Portale existiert durch die *Java Portlet Specification* [AH03] seit dem Jahr 2003 ein Standard für Portlets auf der Grundlage von Java-Technologien. Im Unterschied zu Servlets dienen Portlets dazu, Teile einer Seite und keine komplette Webseite darzustellen, weshalb auch keine Tags verwendet werden dürfen, die sich auf eine ganze Seite beziehen (`base`, `body`, `frame`, `frameset`, `head`, `html`, `title`). Die Spezifikation definiert dazu die Interoperabilität zwischen einem Portlet und einem *Portlet-Container*, in dem die Portlets verwaltet werden. Er steuert den Lebenszyklus und leitet die Requests an das Portlet weiter. Der Lebenszyklus besteht dabei aus der Erzeugung und Initialisierung, der Verarbeitung von eingehenden Requests und der abschließenden Beendigung und Entfernung. Dieser Lebenszyklus wird dabei von der Schnittstelle `Portlet` abgebildet, die im Paket `javax.portlet` enthalten ist. Für das Laden des Portlets verwendet der Portlet-Container einen Classloader und ruft, nachdem es geladen wurde, die Methode `init()` zur Initialisierung auf. Zur Initialisierung werden über Parameter weitere sprachspezifische Informationen aus der *Portlet-Definition* im *Deployment-Deskriptor* des Portlets einbezogen. Nachdem das Portlet initialisiert wurde, kann es Requests empfangen und verarbeiten. Wird das Portlet nicht mehr benötigt, wird es vom Portlet-Container über den Aufruf von `destroy()` beendet [Kus05].

Für die Verarbeitung von Requests werden zwei Arten in der Portlet-API unterschieden. Mit Action-Requests werden Anforderungen zu Zustandsänderungen an das Portlet geschickt, mit Render-Requests wird das Portlet hingegen angewiesen, die Darstellung zu erneuern. Der Ablauf der beiden Requests, die über die Methoden `render()` und `processAction()` realisiert werden, sind in Abbildung 7.4 dargestellt.

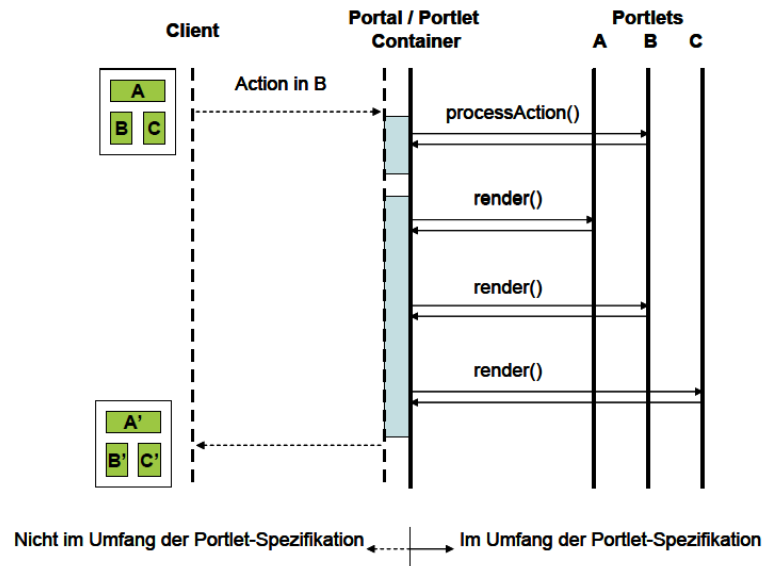


Abbildung 7.4.: Darstellung eines Action-Requests mit `processAction()` und `render()` nach [Kus05]

In dem gezeigten Beispiel sind drei Portlets auf einer Seite angeordnet. Der Client sendet einen Action-Request an Portlet B. Zur Verarbeitung des Action-Requests ruft der Portlet-Container dazu die Methode `processAction()` von Portlet B auf. Nach Beendigung von `processAction()` werden alle `render()`-Methoden der Portlets aufgerufen, um die Seite neu darzustellen. Die Reihenfolge dieser Aufrufe spielen dabei keine Rolle. Sie können auch parallel bearbeitet werden. Die Darstellung durch den Aufruf von `render()` erfolgt entweder durch die Ausgabe auf ein `PrintWriter`-Objekt, durch die Delegation an ein Servlet oder eine Java-Server-Page.

Da der Web-Client von ADEPT2 auf RAP basiert, werden Portlets bislang nicht vollständig unterstützt. Die Unterstützung von RAP für Portlets war für Version 1.2 angekündigt, wurde jedoch auf unbestimmte Zeit verschoben [Lan08]. RAP-Clients können jedoch trotzdem in Portale integriert werden, indem der RAP-Client über einen *iFrame* als Portlet eingebunden wird.

7.4. Zusammenfassung

Im vorliegenden Kapitel wurde gezeigt, wie die beschriebene allgemeine Konzeption aus Kapitel 5 zur Integration von ADEPT2 mit einem DMS umgesetzt werden kann.

Dazu wurde zuerst auf die Aspekte der Datenintegration eingegangen. Mit den Schemata wurden auf Basis der strukturierten Datentypen eine Repräsentation von DMS-Inhalten aufgezeigt, die sich in den Datenfluss von ADEPT2 eingliedert. Somit können die in Dokumenten enthaltenen Informationen in Prozessen genutzt werden. Weiterhin dienen die Schemata auch der Speicherung von Dokumenten im DMS. Um die DMS-Inhalte transparent in den Datenfluss zu integrieren, wurde auf die Realisierung von Lese- und Schreibzugriffen auf das DMS-Repository eingegangen und aufgezeigt wie Bearbeitungen von vorhandenen Inhalten in Prozessen umgesetzt werden können. Dabei wurden konkurrierende Zugriffe berücksichtigt. Außerdem wurde gezeigt, wie DMS-Inhalte exklusiv für die Dauer der Ausführung eines Prozesses in ADEPT2 bereitgestellt werden können. Abschließend wurde auf die Erstellung und die Verwaltung von Schemata in ADEPT2 eingegangen.

Bei der Integration der Organisationsaspekte wurde aufgezeigt, wie sich die Benutzerverwaltungen der beiden Systeme über einen Verzeichnisdienst integrieren lassen. Darauf aufbauend wurde anhand von JAAS erläutert, wie ADEPT2 um Funktionen zur Authentifizierung über einen Verzeichnisdienst erweitert werden kann. Eine Möglichkeit zur Implementierung von SSO-Funktionen in ADEPT2 wurde anhand der Kerberosunterstützung in JAAS dargestellt.

Zuletzt wurden Integrationsaspekte des ADEPT2 Rich- und Web-Clients aufgezeigt. Hierzu wurde vorgestellt, wie sich die Client-Komponenten auf Basis der Eclipse Rich Client Platform integrieren lassen. Weiter wurde darauf eingegangen, wie Client-Komponenten integriert werden können, die nicht in Java vorliegen. Anhand von JNBridge wurde aufgezeigt, wie Funktionen aus Java und .NET gemeinsam durch die Einbindung über die JNBridge Proxy-Klassen realisiert werden kann. Abschließend wurde auf die Integration der Web-Clients in Portale mit Hilfe von Portlets eingegangen und wie der ADEPT2 RAP-Client in ein Portlet integriert werden kann, bis dies offiziell unterstützt wird.

8. Zusammenfassung und Ausblick

Jedes Unternehmen muss sich in der heutigen Zeit mit den Problemen der wachsenden Informationsflut auseinandersetzen. Um die riesigen Informationsmengen zu bewältigen, ist der systematische Umgang mit Dokumenten und Content unerlässlich. In den letzten Jahren sind in diesem Umfeld eine Vielzahl neuer Begriffe und Akronyme entstanden, die suggerieren, dass sich damit die Probleme im Umgang mit den unternehmensweit gespeicherten Informationen lösen lassen. Die meisten dieser Begriffe werden jedoch in erster Linie dazu verwendet, um bekannte Technologien mit einem neuen Schlagwort zu versehen, um die dahinter steckenden Produkte besser am Markt platzieren zu können. Nichtsdestoweniger ist die Notwendigkeit für die Lösung dieser Probleme nach wie vor gegeben.

8.1. Zusammenfassung

In der vorliegenden Arbeit wurde der Umfang des ECM-Konzepts identifiziert. Dazu wurde anhand der Definition der AIIM ein 5-Komponenten-Modell vorgestellt, das die einzelnen Funktionsbereiche beschreibt, die zur umfassenden Verwaltung von unstrukturierten Daten und Informationen dient.

Im Bereich Capture wurden Funktionen und Technologien vorgestellt, mit denen Informationen, die auf sehr unterschiedlichen Medien vorliegen, digitalisiert werden können. Bei bereits digital erzeugten oder vorliegenden Informationen wurde gezeigt, wie diese um zusätzliche Metadaten erweitert werden können, um den Inhalt von Dokumenten für Anwendungen und die Benutzer zugänglich zu machen. In den Bereichen Store und Preserve wurden die verschiedenen Speichertechnologien, die zur Speicherung lebender Informationen und zur Langzeitarchivierung eingesetzt werden können. Ebenso wurden die Abstraktionsschichten vorgestellt, mit denen die physikalischen Speichermedien für Anwendungen und den Benutzer transparent dargestellt werden. Darauf aufbauend wurden die unterschiedlichen Verwaltungsaspekte von ECM erläutert: Der Umgang mit Dokumenten, die Erstellung und Verwaltung von Webinhalten, die sichere und zu rechtlichen Vorschriften konforme Aufbewahrung von Inhalten, die Integration von Koordinations- und Zusammenarbeitsfunktionen und zuletzt die systemgestützte Durchführung von Prozessen umfassen. Abschließend wurde auf die im Rahmen von ECM genannten Funktionen zur Ausgabe und der damit verbundenen

Aufbereitung von Inhalten eingegangen, mit denen unterschiedliche Zielgruppen über verschiedene Distributionskanäle Informationen erhalten können.

Die Untersuchung verschiedener DM- und CM-Systeme hat gezeigt, dass der volle Umfang des ECM-Konzepts mit einem einzelnen System bislang nicht abgebildet werden kann. So umfassen die aus Archivsystemen hervorgegangen DMS hauptsächlich die Bereiche der Erfassung, die Verwaltung, die einfache Verteilung, Speicherung und die Langzeitarchivierung von Dokumenten und Content. Die in den letzten Jahren zunehmende Verbreitung von Web-Technologien und die Verwaltung von Web-Content wird hingegen von diesen Systemen bislang nicht oder nur unzureichend berücksichtigt. Eine Unterstützung zur gemeinsamen Erstellung oder Bearbeitung von Dokumenten bzw. Content werden bislang ebenfalls nicht von den Systemen berücksichtigt. CM-Systeme, die ihren Hintergrund eher im Bereich des WCM haben, adressieren hier eher die Erstellung und Bearbeitung von Content und Dokumenten. Durch ihren weborientierten Aufbau bieten sie auch mehr Funktionen zur Zusammenarbeit und einfachere Zugriffsmöglichkeiten für verteilte Arbeitsgruppen und mobile Benutzer. Jedoch werden bei diesen Systemen die Bereiche der Erfassung durch fehlende Integrationsmöglichkeiten für Papierdokumente und die Langzeitarchivierung, sowie Funktionen zur strukturierten Ablage vernachlässigt und sind nicht oder nur sehr rudimentär vorhanden. Für die systemseitige Unterstützung zur Durchführung von Geschäftsprozessen hat sich gezeigt, dass alle Systeme Schwächen bei der Umsetzung aufweisen. So gestaltet sich die Abbildung von komplexen Geschäftsprozessen mit den meisten Systemen sehr schwierig und aufwendig. Integrationen von Anwendungen, die den Benutzer bei der Bewältigung von Prozessschritten unterstützen, sind mit den Systemen ebenfalls nicht möglich.

Darauf aufbauend wurden Anforderungen an die Integration eines CMS, eines DMS und eines PMS beschrieben, mit denen sich das ECM-Konzept in allen Bereichen umsetzen lässt. Hierbei wurde vor allem auf die Aspekte der Integration des DMS mit einem PMS eingegangen, da zum einen bereits Integrationen von CM- und DM-Systemen existieren, zum anderen weil die Synergieeffekte zwischen DMS und PMS deutlicher ausfallen.

Durch die sich anschließende Konzeption zur Integration von DMS und PMS wurde allgemein vorgestellt, wie sich die im DMS gespeicherten Informationen in Prozesse integrieren lassen. Außerdem wurde gezeigt, wie Informationen aus Prozessen im DMS bereitgestellt werden können, um sie für die längerfristige Aufbewahrung und zur Dokumentation von Prozessen zu nutzen. Weiterhin wurde das PMS dahingehend erweitert, dass Anforderungen durch rechtliche Vorschriften oder unternehmensinterne Vorgaben gemäß dieser Bestimmungen aufbewahrt und am Ende der Nutzung kontrolliert vernichtet werden können.

Abschließend wurde die allgemeine Konzeption auf das PMS ADEPT2 übertragen. Dazu wurde gezeigt, wie sich Dokumente, die in einem DMS gespeichert sind, in Prozesse

integrieren und die darin enthaltenen Informationen für die Durchführung von Prozessen nutzen lassen. Weiter wurde gezeigt, wie sich die Organisationskomponenten von ADEPT2 in Verzeichnisdienste integrieren lassen, um darüber eine gemeinsame Benutzerverwaltung mit anderen Systemen zu realisieren. Dazu wurden verschiedene Authentifizierungsvarianten vorgestellt, die von der getrennten Authentifizierung über Single Credential bis hin zu SSO reichen, bei dem sich ein Benutzer nur einmalig am System anmelden muss. Abschließend wurden Ansätze zur Integration der Clientkomponenten für den ADEPT2 Rich-Client und den Web-Client in Portale aufgezeigt.

8.2. Ausblick

Das ECM-Konzept zielt darauf ab, strukturierte, schwach strukturierte und unstrukturierte Informationen zusammenzuführen. Aus aktueller Sicht hinken existierende Systeme, die versuchen, dieses Konzept vollständig umzusetzen, dem angestrebten Umfang hinterher. Weiterhin bleibt es fraglich, ob in absehbarer Zukunft Systeme entwickelt werden können, die in der Lage sind, alle diese Bereiche zufriedenstellend abzudecken. Besser wäre es, wenn sich Systeme, die sich auf einen oder wenige Bereiche spezialisieren, leicht integrieren lassen könnten.

Hierzu besteht vor allem der Bedarf nach verlässlichen Standards, die zuerst das zentrale Repository betreffen, in dem die Informationen zusammengeführt werden. Bislang existieren hierzu aktuell mit JCR und CMIS zwei Ansätze, um plattformunabhängige Schnittstellen für den Zugriff auf die Informationen zu entwerfen. Hier bleibt abzuwarten, ob sich diese Schnittstellen auf Seiten der Hersteller durchsetzen und diese Schnittstellen Einzug in die Systeme halten. Da JCR jedoch ein Java-Standard ist, ist nicht davon auszugehen, dass er sich über die bereits Java-basierten DMS und CMS hinaus verbreiten wird.

Um ein zentrales Repository, wie im ECM-Konzept beschrieben, zukünftig etablieren zu können, benötigt dieses, neben den standardisierten Schnittstellen, auch gewisse Basisfunktionen, wie die Unterstützung von Transaktionen, um den zu integrierenden Systemen entsprechende Funktionen zum Zugriff auf gemeinsam genutzte Ressourcen zu ermöglichen. Transaktionen stellen eine grundlegende Funktion dar, die ein modernes Repository bereitstellen muss, um vor allem die immer stärker eingesetzten Workflow- und BPM-Funktionen oder gemeinsame Arbeitsbereiche unterstützen zu können. Für beide Bereiche ist es wichtig, dass alle Inhalte, die für eine bestimmte Bearbeitung benötigt werden, vollständig bereitgestellt werden können, ohne dass externe Änderungen Einfluss auf die Konsistenz der Informationen nehmen können. Bislang werden Transaktionen in der CMIS-Spezifikation jedoch nicht erwähnt. Somit ist zumindest in naher Zukunft nicht davon auszugehen, dass diese Funktion genutzt werden kann.

Damit die im ECM-Konzept beschriebenen Verwaltungskomponenten als einzelne spezialisierte Systeme integriert werden können, sind weiterhin auch standardisierte Beschreibungen für Dokumenten und Content notwendig. Mit Hilfe dieser Beschreibungen können die in Dokumenten vorhandenen Informationen, die unstrukturiert gespeichert sind, über ihre Metadaten beschrieben werden, die dann wiederum von den einzelnen Verwaltungskomponenten genutzt werden können. Da der damit beschriebene strukturelle Aufbau mit den zu speichernden Inhalten direkt verknüpft wird, sollten diese ebenfalls im Repository verankert werden können. Somit können auch wirklich unternehmensweit und über Anwendungsgrenzen hinaus die Informationen einheitlich genutzt, bereitgestellt und verwaltet werden. Für die Beschreibung der Metadaten und des strukturellen Aufbaus bietet sich eine XML-basierte Beschreibung an. XML ist ein bereits weit verbreitetes Format und besitzt die Möglichkeiten über XML-Schema beispielsweise den strukturellen Aufbau zu beschreiben und festzulegen.

Da die im Repository gespeicherten Informationen typischerweise mit verschiedenen Anwendungsprogrammen erzeugt werden, stellen diese ebenfalls einen wichtigen Punkt zur gesamtheitlichen Informationsverwaltung dar. Die Integration von Anwendungen wie z. B. ERP- oder CRM-Systeme basiert aktuell typischerweise auf proprietären Anbindungen, die individuell für jedes der Systeme entwickelt werden muss. Da speziell ERP-Systeme sehr viele Daten und Dokumente erzeugen, die Beziehungen zu verschiedenen ECM-Bereichen der Erfassung, der Archivierung, der Einbindung in Geschäftsprozesse und den Bereich Deliver besitzen, zählen diese Systeme zu den wichtigsten Anwendungen, die integriert werden müssen. Obwohl diese Systeme und die weitreichenden Verknüpfungen zu anderen Informationen in Unternehmen bereits seit langem bekannt sind, existieren bis heute nur sehr wenige Standardschnittstellen. Hier bietet beispielsweise der Marktführer SAP mit der Archive-Link-Schnittstelle eine Ausnahme, die eine für SAP-Systeme standardisierte Schnittstelle für die Archivierung mit DMS zur Verfügung stellt [Kor03]. Dadurch wird zum einen die Integration der Archivierung erleichtert, zum anderen wird aber auch aus Sicht des Quellsystems bereits berücksichtigt, dass die erzeugten Informationen an andere Systeme weitergegeben und ggf. von diesen auch verwaltet werden. Insbesondere ist zu berücksichtigen, dass Daten nach einem gewissen Zeitraum der Aufbewahrung gelöscht werden sollen oder auch müssen. Die Anforderungen der Löschung von Daten wird dabei zum einen durch die enorm ansteigende Menge an Informationen erforderlich, zum anderen durch weiter zu erwartende regulatorische Anforderungen, die beispielsweise den Datenschutz betreffen. Die Einbeziehung der mit der Auslagerung von Daten verbundenen Konsequenzen sollten daher bereits in den Konzepten der Anwendungen zukünftig berücksichtigt und implementiert werden, um eine Integration zu vereinfachen bzw. zu ermöglichen.

Eine weitere zukünftige Herausforderung stellt neben der Integration von Daten und Informationen auch die Abbildung der Organisationsstrukturen und damit verbundenen Zugriffsberechtigungen, sowie auf technischer Seite die Authentifizierung zwischen Systemen dar. Aufgrund der Vielzahl existierender und verwendeter Technologien, lassen sich diese Aspekte aktuell nur sehr schwer umsetzen und verwalten. Da

die Autorisierung für den Zugriff und die Veränderung von Daten meist an konkreten Anforderungen und Funktionen innerhalb einer Anwendung ausgerichtet sein müssen, kann hierfür kaum ein allgemeingültiges Modell für alle Anwendungen und Systeme entworfen werden. Darüber hinaus stellen die unterschiedlichen Systeme auch verschiedene Anforderungen, nach welchen Kriterien entschieden wird, ob ein Zugriff zulässig ist oder nicht. Die Aspekte der zentralen und unternehmensweiten Benutzerverwaltung und der Authentifizierung von Benutzern, können jedoch umgesetzt werden. So werden hierfür zentral verwaltete Benutzerdatenbanken und Authentifizierungsdienste benötigt, die von Anwendungen integriert werden. Darüber lassen sich die Zugriffe, Authentifizierungsvorgänge und Anmeldeinformationen für Benutzer vereinheitlichen. Außerdem kann hierüber sichergestellt werden, dass auch immer der gleiche Benutzer über mehrere Systeme hinweg authentifiziert wird und die ihm zugeteilten Berechtigungen erhält. Verzeichnisdienste können bislang zumindest die grundlegenden Funktionen zur Verfügung stellen, auch wenn sie insbesondere in der Abbildung von Organisationsstrukturen nicht immer optimal geeignet sind.

Der letzte Baustein zu einem einheitlichen Gesamtsystem stellt die Kommunikation mit dem Benutzer dar. Hier haben sich zum einen durch das Client-Server-Modell auf dem Rechner des Benutzers installierte Clients, zum anderen webbasierte Clients über einen Browser etabliert. Beide Modelle besitzen ihre spezifischen Vor- und Nachteile und werden daher in Zukunft vermutlich parallel eingesetzt. Auf der webbasierten Seite haben seit geraumer Zeit Portale Einzug gehalten, um dem Benutzer verschiedene Web-Clients und Komponenten integriert anzubieten. Typischerweise kann sich der Benutzer hierbei auch individuelle Ansichten zusammenstellen und einzelne Komponenten kombinieren. Auf Seite der Rich-Clients stehen diese Möglichkeiten bislang kaum zur Verfügung, da jede Anwendung innerhalb eines eigenen Fensters ausgeführt wird. Die Erstellung individueller Anzeigen obliegt dabei den Funktionen des Rich-Clients und beschränkt sich auf die Komponenten des jeweiligen Clients. Um hier ähnlich flexibel wie in Web-Portalen zu agieren, sind zukünftig Client-Frameworks, wie z. B. Eclipse denkbar. Darüber kann ein einzelner standardisierter Client auf alle Rechner der Benutzer ausgerollt werden. Die Anwendungskomponenten könnten dabei als Plugins innerhalb dieses Basis-Clients ausgeführt werden. Weiterhin können hierüber aufwendige Rollouts und Remoteinstallationen auf die Clients reduziert werden, indem z. B. die Plugins über einen Server bereitgestellt werden und beim Start ggf. eine neue Version automatisch vom Server geladen und gestartet wird. Weiterhin könnten die Authentifizierungsaspekte vereinfacht werden, indem der Basis-Client die Authentifizierung des Benutzers übernimmt. Die Plugins könnten dabei automatisch im Kontext des ausführenden Benutzers ausgeführt werden.

A. Abkürzungsverzeichnis

AIIM	Association for Information and Image Management
AO	Abgabenordnung
API	Application Programming Interface
BLOB	Binary Large Object
BGB	Bürgerliches Gesetzbuch
BPM	Business Process Management
CAS	Content Addressed Storage
CI	Coded Information
CIFS	Common Internet File System
CLR	Common Language Runtime
CM	Content Management
CMS	Content Management System
CMIS	Content Management Interoperability Services
COLD	Computer Output on Laserdisc
Collab	Collaboration
CRM	Customer Relationship Management
DAS	Direct Attached Storage
DIB	Directory Information Base
DIT	Directory Information Tree
DM	Dokumenten-Management
DMA	Document Management Alliance
DMS	Dokumenten-Management-System
DN	Distinguished Name
DoD5015.2	Department of Defense 5015.2-STD
DRM	Digital Rights Management
ECM	Enterprise Content Management
EDI	Electronic Data Interchange
EDV	Elektronische Datenverarbeitung
EEX	EASY ENTERPRISE.x
ELC	EASY Logistics Center
EQL	EASY Query Language
ERM	Enterprise Reports Management
ERP	Enterprise Resource Planning
FTP	File Transfer Protocol
GDPdU	Grundsätze zum Datenzugriff und zur Prüfbarkeit digitaler Unterlagen

GoB Grundsätze ordnungsgemäßer Buchführung
GoBS Grundsätze ordnungsgemäßer DV-gestützter Buchführung
HCR Handprint Character Recognition
HGB Handelsgesetzbuch
HSM Hierarchisches Speicher-Management
HTML Hypertext Markup Language
ICR Intelligent Character Recognition
IDR Intelligent Document Recognition
ILM Information Lifecycle Management
J2EE Java Platform Enterprise Edition
JAAS Java™ Authentication and Authorization Service
JBPM JBoss Business Process Management
JCR Content Repository API for Java
JNI Java Native Interface
JVM Java Virtual Machine
KDC Key Distribution Center
LDAP Lightweight Directory Access Protocol
LDIF LDAP Interchange Format
MoReq2 Model Requirements for the Management of Electronic Documents and Records 2
MOSS Microsoft Office Sharepoint Server
NAS Network Attached Storage
NCI Non-Coded Information
NFS Network File System
OCR Optical Character Recognition
ODMA Open Document Management API
OID Object Identifier
OMR Optical Mark Recognition
PAM Pluggable Authentication Module
PMS Prozess-Management-System
PDM Produktdatenmanagement
PKI Public Key Infrastructure
RAID Redundant Array of Inexpensive Disks
RAP Rich Ajax Platform
RCP Rich Client Platform
RDN Relative Distinguished Name
REST Representational State Transfer
RSS Really Simple Syndication oder auch Rich Site Summary
RM Records Management
SAN Storage Area Network
SASL Simple Authentication and Security Layer
SigG Signaturgesetz
SMTP Simple Mail Transfer Protocol
SOA Service-orientierte-Architektur

SOX Sarbanes-Oxley Act
SSL Secure Sockets Layer
SSO Single Sign On
SQL Structured Query Language
TGT Ticket Granting Ticket
TLS Transport Layer Security
UML Unified Modeling Language
URL Uniform Resource Locator
WCM Web Content Management
WebDAV Web-based Distributed Authoring and Versioning
WF Workflow
WORM Write Once Read Many
XML Extensible Markup Language

Abbildungsverzeichnis

2.1. Abgrenzung von Content und Dokument [MV02]	15
2.2. Dokumentenlebenszyklus [GSMK04]	16
2.3. ECM-Modell [KSP07]	20
2.4. ECM-Capture [KSP07]	21
2.5. Scannen und Indexieren [GSSZ02]	25
2.6. ECM-Manage [Kam03b]	26
2.7. Dokumenten-Management [KSP07]	27
2.8. Records-Management [KSP07]	29
2.9. Web-Content-Management [KSP07]	32
2.10. Collaboration [KSP07]	35
2.11. Workflow/BPM [KSP07]	36
2.12. ECM-Store [KSP07]	39
2.13. ECM-Preserve [KSP07]	42
2.14. IBM DR 550 Archiv-System [IBM09b]	44
2.15. Hierarchisches Speicher-Management nach [GSSZ02]	45
2.16. ECM-Deliver [KSP07]	46
3.1. VOI-Studie zur Untersuchung des Marktpotentials 2007 [BHW07]	51
3.2. Alfresco-Architektur [Pot08]	57
3.3. Architektur der Erfassungskomponenten [Sha06]	58
3.4. Portalintegration von Alfresco und Liferay [Pot08]	61
3.5. JBPM Workflow-Editor [Sha06]	63
3.6. MOSS 2007-Architektur [Mic09c]	65
3.7. Workflow Foundation und MOSS [Cha06]	69
3.8. Workflow Designer für Visual Studio 2005 [Cha06]	70
3.9. Architektur von ELOprofessional [ELOb]	73
3.10. ELO-DocXtractor [ELOa]	74
3.11. Workflow-Editor	78
3.12. DocuWare-Architektur [Doc05]	80
3.13. Zuweisung von Indexfeldern	81
3.14. EASY ENTERPRISE.x-Architektur [EAS06b]	86
3.15. EASY-Formularassistent	88
3.16. Barcode-Erkennung mit EASY Capture [EAS06a]	89
3.17. ELC-Architektur [EAS07]	92
3.18. EEX-Adhoc-Workflow definieren	93
3.19. ELC-Prozessmodellierung mit Visio	94

4.1. Trennung von Prozesslogik und Anwendungscode [Rin04]	103
4.2. CMIS-Architektur [EIM08]	109
4.3. MoReq2-Klassifikationsschema [Eur08b]	112
4.4. MoReq2-ER-Modell	113
5.1. Architektur der Integration von DMS und PMS	121
5.2. Einfaches Beispiel für ein Dokumentenschema	122
5.3. Beispielprozess für eine Rechnungsprüfung	125
5.4. Objektklasse Person mit Vererbung	129
5.5. Beispiel eines Directory Information Tree nach [HSG03, Kin00]	130
5.6. Beispiel zur Ermittlung von Berechtigungen auf Dokumente im DMS	134
5.7. Kerberos-Ticket nach [Cha99]	137
5.8. Erhalt eines <i>Ticket Granting Ticket</i> nach [Cha99]	138
5.9. Erhalt und Verwendung eines <i>Service Tickets</i> nach [Cha99]	139
5.10. Kerberos-Delegation nach [Cha99]	140
5.11. Ein Web-Portal mit integriertem DMS- und PMS-Client	142
6.1. Prozessvorlage mit drei Instanzen	146
6.2. Schematische Darstellung der Kontrollflusskonstrukte nach [Rei00]	147
6.3. Beispiel eines strukturierten Datentyps <i>Adresse</i> nach [For09]	149
6.4. Beispiel eines strukturierten Datentyps <i>Firmenkontakt</i> einschließlich Vererbung nach [For09]	149
6.5. ADEPT2-Organisationsmodell [Ber05]	151
7.1. Beispiel für die Erstellung eines <i>Rechnungsschemas</i> als strukturierten Datentyp	156
7.2. Ablauf einer Authentifizierung mit Kerberos nach [SS02]	168
7.3. JNBridge-Architektur [JNB08]	172
7.4. Darstellung eines Action-Requests mit <code>processAction()</code> und <code>render()</code> nach [Kus05]	174

Tabellenverzeichnis

3.1. Unterschiede Sharepoint-Designer und Visual-Studio-Designer	71
3.2. Funktionaler Vergleich der evaluierten Systeme	100
4.1. Abgrenzung von DMS, CMS und PMS	106
4.2. MoReq2-Entitäten	114
5.1. LDAP-Attributtypen	129
7.1. Verzeichnisdienstinformationen	165

Literaturverzeichnis

- [Ada07] ADAM, Azad: *Implementing Electronic Document and Record Management Systems*. Auerbach Publications, 2007
- [AH03] ABDELMUS, Alejandro ; HEPPER, Stefan: *JSR-168 Java Portlet Specification 1.0*. Oktober 2003
- [AH04] ALST, Wil van d. ; HEE, Kees van: *Workflow Management: Models, Methods and Systems*. Mit Press, 2004
- [AII06] ASSOCIATION FOR INFORMATION AND IMAGE MANAGEMENT (AIIM): RECOMMENDED PRACTICE - Analysis, Selection, and Implementation Guidelines Associated with Electronic Document Management Systems (EDMS). 2006. – Forschungsbericht
- [All09] ALLIANCE STORAGE TECHNOLOGIES INC. (Hrsg.): *UDO Media*. Alliance Storage Technologies Inc., 2009. http://www.plasmon.com/archive_solutions/media.html. – Abgefragt 01.09.2009
- [Apa09] APACHE SOFTWARE FOUNDATION (Hrsg.): *Apache Lucene*. Apache Software Foundation, August 2009. <http://lucene.apache.org/>. – Abgefragt 07.08.2009
- [Ass02] ASSOCIATION FOR INFORMATION AND IMAGE MANAGEMENT (AIIM) (Hrsg.): *DMA The Document Management Alliance*. 1100 Wayne Avenue, Suite 1100, Silver Spring, MD 20910, USA: Association for Information and Image Management (AIIM), Januar 2002. <http://dmatech.info/>. – Abgefragt 10.11.2009
- [Ass08] ASSOCIATION FOR INFORMATION AND IMAGE MANAGEMENT (AIIM) (Hrsg.): *What is ECM?* 1100 Wayne Avenue, Suite 1100, Silver Spring, MD 20910, USA: Association for Information and Image Management (AIIM), 2008. <http://www.aiim.org/what-is-ecm-enterprise-content-management.aspx>. – Abgefragt 15.07.2009
- [BAN89] BURROWS, Michael ; ABADI, Martin ; NEEDHAM, Roger: A logic of Authentication / Digital Systems Research Center. 1989 (39). – Forschungsbericht

- [Bau96] BAUMGARTEN, Bernd: *Petri-Netze: Grundlagen und Anwendungen*. 2. Auflage. Spektrum Akademischer Verlag, 1996
- [BDS⁺93] BREITBART, Y. ; DEACON, A. ; SCHEK, H.-J. ; SHETH, A. ; WEIKUM, G.: Merging application-centric and data-centric approaches to support transaction-oriented multi-system workflows. In: *SIGMOD Rec.* 22 (1993), Nr. 3, S. 23–30
- [Ben08] BENTELE, Reinhold: *Konzepte, Technologien und Realisierungsvarianten eines webbasierten, intelligenten Workflow-Klienten*, Universität Ulm, Diplomarbeit, 2008
- [Ber05] BERROTH, Marco: *Konzeption und Entwurf einer Komponente für Organisationsmodelle*, Universität Ulm, Diplomarbeit, 2005
- [BHW07] BERNDT, Oliver ; HALSTENBACH, Volker ; WEISS, Dr. D.: *Dokumentenmanagement in Deutschland 2007 - Untersuchung des Marktpotentials - Ergebnisbericht / VOI (Verband Organisations- und Informationssysteme e.V.)*. 2007. – Forschungsbericht
- [BL94] BERNDT, Oliver ; LEGER, Lothar: *Dokumenten- Management- Systeme. Nutzen, Organisation, Technik*. Luchterhand, 1994
- [Boi04] BOIKO, Bob: *Content Management Bible*. Wiley Publishing, 2004
- [Cha99] CHAPPELL, David: Exploring Kerberos, the Protocol for Distributed Security in Windows 2000. In: *Micrisoft Systems Journal* (1999), August
- [Cha06] CHAPPEL, David ; MICROSOFT CORPORATION (Hrsg.): *Understanding Workflow in Windows SharePoint Services and the 2007 Microsoft Office System*. Microsoft Corporation, September 2006
- [CLR01] CORMEN, Thomas H. ; LEISERSON, Charles E. ; RIVEST, Ronald L.: *Introduction to Algorithms*. MIT Press, 2001
- [DAG⁺06] DADAM, Peter ; ACKER, Hilmar ; GÖSER, Kevin ; JURISCH, Marin ; KREHER, Ulrich. ; LAUER, Markus ; RINDERLE, Stefanie ; REICHERT, Manfred: ADEPT2 - Ein adaptives Prozess-Management-System der nächsten Generation. In: *Tagungsband Stuttgarter Softwaretechnik Forum 2006* (2006), S. 10
- [DAH05] DUMAS, Marlon ; AALST, Wil can d. ; HOFSTEDDE, Arthus H. M.: *Process-Aware Information Systems: Bridging People and Software Through Process Technology*. Wiley-Interscience, 2005
- [DMw08] DMWARE AIIM (Hrsg.): *ODMA Open Document Management API Interoperability Exchange*. 1100 Wayne Avenue, Suite 1100, Silver Spring, MD

- 20910, USA: DMware AIIM, April 2008. <http://odma.info/>. – Abgefragt 10.11.2009
- [Doc05] DocuWARE AG (Hrsg.): *Docuware-Systemarchitektur Version 1.1*. Therese-Giehse-Platz 2, 82110 Germering: DocuWare AG, Dezember 2005
- [DoD07] *DoD 5015.2 -STD ELECTRONIC RECORDS MANAGEMENT SOFTWARE APPLICATIONS DESIGN CRITERIA STANDARD*. U.S. Department of Defense, April 2007
- [DP98] DAVENPORT, Thomas H. ; PRUSAK, Laurence: *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Press, 1998
- [EAS06a] EASY SOFTWARE AG (Hrsg.): *EASY Capture*. Am Hauptbahnhof 4, D-45468 Mülheim an der Ruhr: EASY SOFTWARE AG, Januar 2006
- [EAS06b] EASY SOFTWARE AG (Hrsg.): *EASY ENTERPRISE.x Der neue Dokumentenserver der EASY SOFTWARE AG*. Am Hauptbahnhof 4, D-45468 Mülheim an der Ruhr: EASY SOFTWARE AG, Januar 2006
- [EAS07] EASY SOFTWARE AG (Hrsg.): *EASY LOGISTICS CENTER die Drehscheibe für Information, Dokumente und Vorgänge*. Am Hauptbahnhof 4, D-45468 Mülheim an der Ruhr: EASY SOFTWARE AG, 2007
- [Ecl09] ECLIPSE FOUNDATION (Hrsg.): *eclipse*. Eclipse Foundation, Dezember 2009. <http://www.eclipse.org/>. – Abgefragt: 22.12.2009
- [eCo09] eCOPY, INC. (Hrsg.): *eCopy ShareScan*. eCopy, Inc., September 2009. <http://www.ecopy.de/products-ecopy-sharescan.asp>. – Abgefragt 18.09.2009
- [EIM08] EMC ; IBM ; MICROSOFT: *Content Management Interoperability Services (CMIS) 0.5*. August 2008
- [ELOa] ELO DIGITAL OFFICE GMBH (Hrsg.): *ELO DocXtractor INVOICE*. Heilbronner Straße 150, D-70191 Stuttgart: ELO Digital Office GmbH
- [ELOb] ELO DIGITAL OFFICE GMBH (Hrsg.): *ELOprofessional & ELOenterprise Release 6.0*. Heilbronner Straße 150, D-70191 Stuttgart: ELO Digital Office GmbH
- [EMC08] EMC DEUTSCHLAND GMBH (Hrsg.): *EMC Centera Content-Addressed Storage-System*. Am Kronberger Hang 2 a, D-65824 Schwalbach/Taunus: EMC Deutschland GmbH, 2008
- [EMC09] EMC DEUTSCHLAND GMBH (Hrsg.): *EMC Captiva Family*. EMC Deutschland GmbH, September 2009. <http://germany.emc.com/products/family/captiva-family.htm>. – Abgefragt 12.09.2009

- [Eng07] ENGLISH, Bill: *Microsoft Office SharePoint Server 2007*. Microsoft Press, 2007
- [Eur08a] EUROPEAN COMMISSION (Hrsg.): *Appendix 9 to the MoReq2 Specification - Model requirements for the management of Electronic Records: Metadata Model*. European Commission, 2008
- [Eur08b] EUROPEAN COMMISSION (Hrsg.): *Model Requirements for the Management of Electronic Records - MoReq2 Specification*. European Commission, 2008
- [fil09] FILEMARK CORPORATION (Hrsg.): *SMARTi COLD for Alfresco*. fileMARK Corporation, 2009. <http://filemark.com/SMARTi%20COLD%20for%20Alfresco.pdf>. – Abgefragt 22.08.2009
- [Fis07] FISHER, Layna (Hrsg.): *2007 BPM & Workflow Handbook*. Future Strategies Inc., 2007
- [For09] FORSCHNER, Alexander: *Fortschrittliche Datenflusskonzepte für flexible Prozessmodelle*, Universität Ulm, Diplomarbeit, 2009
- [Fra08] FRAPPAOLO, Carl: *Building an ECM Strategy – Alternatives and Decision Points / AIIM*. 2008. – Forschungsbericht
- [FS04] FELDMAN, Susan ; SHERMAN, Chris: *The High Cost of Not Finding Information / IDC*. 2004. – Forschungsbericht
- [FSR06] FLEURY, Marc ; STARK, Scott ; RICHARD, Norman: *JBOSS 4.0. Das offizielle Handbuch*. Addison-Wesley, 2006
- [Gar03] GARMAN, Jason: *Kerberos: The Definitive Guide*. O'Reilly, 2003. – 272 S.
- [Gau03] GAUS, Wilhelm: *Dokumentations- und Ordnungslehre*. Springer Verlag, 2003
- [Gor98] GORDON, Robert: *Essential JNI: Java Native Interface*. Prentice Hall, 1998
- [GRC⁺07] GANTZ, John F. ; REINSEL, David ; CHUTE, Christopher ; SCHLICHTING, Wolfgang ; McARTHUR, John ; MINTON, Stephen ; XHENETI, Irida ; TONCHEVA, Anna ; MANFREDIZ, Alex: *A Forecast of Worldwide Information Growth Through 2010 / IDC*. 2007. – Forschungsbericht
- [GSA08] *Microsoft SharePoint Server and EDRMS*. Government of South Australia, Juli 2008
- [GSMK04] GÖTZER, Klaus ; SCHNEIDERATH, Udo ; MAIER, Berthold ; KOMKE, Torsten: *Dokumenten-Management*. dpunt.verlag, 2004

- [GSSZ02] GULBINS, Jürgen ; SEYFRIED, Markus ; STRACK-ZIMMERMANN, Hans: *Dokumenten-Management*. Springer Verlag, 2002
- [Hil00] HILL, Paul B.: Kerberos interoperability issues. In: *LISA-NT'00: Proceedings of the 3rd conference on Large Installation System Administration of Windows NT*. Berkeley, CA, USA : USENIX Association, 2000
- [HSG03] HOWES, Timothy ; SMITH, Mark ; GOOD, Gordon: *Understanding and Deploying LDAP Directory Services*. 2nd Edition. Addison-Wesley, 2003. – 936 S.
- [IBM09a] IBM CORPORATION (Hrsg.): *IBM WebSphere Software*. IBM Corporation, 2009. <http://www-01.ibm.com/software/de/websphere/>. – Abgefragt 09.10.2009
- [IBM09b] IBM DEUTSCHLAND GMBH (Hrsg.): *IBM System Storage DR550*. Pascalstraße 100, D-70548 Stuttgart: IBM Deutschland GmbH, 2009
- [ISO02] *Information and documentation - Records management -*. International Organization for Standardization, Dezember 2002
- [Jas09] JASIG CORPORATION (Hrsg.): *Central Authentication Service*. Jasig Corporation, September 2009. <http://www.jasig.org/cas>. – Abgefragt 19.09.2009
- [JBo09] JBOSS COMMUNITY (Hrsg.): *JBoss Portal. The open source and open standards Java Portal*. JBoss Community, September 2009. <http://www.jboss.org/jbossportal/>. – Abgefragt 18.09.2009
- [JBS97] JABLONKSI, S. ; BÖHM m. ; SCHULZE, W.: *Workflow Management: Entwicklung von Anwendungen und Systemen*. Dpunkt Verlag, 1997
- [Jen05] JENKINS, Tom: *Enterprise Content Management*. Open Text Corporation, 2005
- [JNB08] JNBRIDGE, LLC (Hrsg.): *Bridging Java and .NET: A Technical Overview of JNBridgePro*. JNBridge, LLC, September 2008
- [JNB09] JNBRIDGE, LLC (Hrsg.): *JNBridge*. JNBridge, LLC, November 2009. <http://www.jnbridge.com/>. – Abgefragt 15.11.2009
- [Jur06] JURISCH, Martin: *Konzeption eines Rahmenwerkes zur Erstellung und Modifikation von Prozessvorlagen und -instanzen*, Universität Ulm, Diplomarbeit, 2006

- [Kam03a] KAMPPFFMEYER, Ulrich: *Dokumenten-Technologien: Wohin geht die Reise?* Project Consult Unternehmensberatung Dr. Ulrich Kampffmeyer GmbH, 2003
- [Kam03b] KAMPPFFMEYER, Ulrich: *Enterprise Content Management - die unternehmensweite Informationsplattform der Zukunft / Project Consult Unternehmensberatung Dr. Ulrich Kampffmeyer GmbH. 2003. – Forschungsbericht*
- [Kin00] KING, Robert R. ; GORAJ, Suzanne (Hrsg.): *Mastering Active Directory*. Sybex, 2000
- [KJOF01] KOOP, Hans J. ; JÄCKEL, K. K. ; OFFERN, Anja L. ; FEDTKE, Stephen: *Erfolgsfaktor Content Management. Vom Web Content bis zum Knowledge Management*. Vieweg Verlag, 2001
- [KN93] KOHL, J. ; NEUMAN, C.: *RFC 1510: The Kerberos Network Authentication Service (V5)*. <http://www.ietf.org/rfc/rfc1510.txt>. Version: September 1993
- [Kof09] KOFAX PLC (Hrsg.): *Kofax Capture*. Kofax PLC, September 2009. <http://www.kofax.com/document-capture-software/>. – Abgefragt 12.09.2009
- [Kor03] KORSCHEN, Markus: *SAP R/3 Data Archiving*. Vieweg+Teubner, 2003
- [Krä95] KRÄNZLE, Hans-Peter: *Dokumentenmanagement: Technik und Konzepte*. In: *Praxis der Wirtschaftsinformatik* Heft 181 (1995), S. 27
- [KSP07] KAMPPFFMEYER, Ulrich ; SŁOMINSKY, Merten ; PFEIFFER, Stefan: *ECM Enterprise Content Management / Project Consult Unternehmensberatung Dr. Ulrich Kampffmeyer GmbH, IBM Deutschland GmbH. 2007. – Forschungsbericht*
- [KT03] KNECHT-THURMANN, Stephanie: *Active Directory*. Addison-Wesley, 2003
- [Kus05] KUSSMAUL, Timo: *Die Java-Portlet-Spezifikation*. In: *Java Spektrum* (2005), März
- [LA94] LEYMANN, Frank ; ALTENHUBER, Wolfgang: *Managing business processes as an information resource*. In: *IBM Syst. J.* 33 (1994), Nr. 2, S. 326–348
- [Lan08] LANGE, Fabian: *Eclipse Rich Ajax Platform: Bringing Rich Client Into the Web*. Apress, 2008
- [Lif09] LIFERAY INC. (Hrsg.): *Liferay Portal*. Liferay Inc., September 2009. <http://www.liferay.com/web/guest/products/portal>. – Abgefragt 18.09.2009

- [Lin96] LINN, J.: *RFC 1964: The Kerberos Version 5 GSS-API Mechanism*. <http://www.ietf.org/rfc/rfc1964.txt>. Version: Juni 1996
- [May09] MAYER, Thomas: *E-Invocing: crown or catalyst of an efficient billing process?* / Deutsche Bank Research. 2009. – Forschungsbericht
- [MHS08] MOLLIEN, Torsten ; HAUSER, Thomas ; SCHARNAGL, Dieter: *Microsoft SharePoint Server 2007 im Einsatz : die Plattform für Business Intelligence und Zusammenarbeit im Unternehmen*. Addison-Wesley, 2008
- [Mia99] MIANO, John: *Compressed Image File Formats. JPEG, PNG, GIF, XBM, BMP*. Addison-Wesley, 1999
- [Mic07] MICROSOFT CORPORATION (Hrsg.): *Microsoft Office Groove Collaborative Solutions through SharePoint Integration*. One Microsoft Way, Redmond, WA 98052-7329 USA: Microsoft Corporation, März 2007
- [Mic09a] MICHELER, Florian: *Konzeption, Implementierung und Integration einer Komponente für die Erstellung intelligenter Formulare*, Universität Ulm, Diplomarbeit, 2009
- [Mic09b] MICROSOFT CORPORATION (Hrsg.): *How Windows SharePoint Services Processes Workflow Activities*. Microsoft Corporation, 2009. <http://msdn.microsoft.com/en-us/library/ms442249.aspx>. – Abgefragt 11.12.2009
- [Mic09c] MICROSOFT CORPORATION (Hrsg.): *Introduction to SharePoint Products and Technologies for the Professional .NET Developer*. One Microsoft Way, Redmond, WA 98052-7329 USA: Microsoft Corporation, 2009. <http://msdn.microsoft.com/en-us/library/cc537498.aspx>. – Abgerufen: August 2009
- [Mül05] MÜLLER, Joachim: *Workflow-based Integration: Grundlagen, Technologien, Management*. Springer Verlag, 2005
- [MPHS06] MUNKVOLD, Bjørn E. ; PÄIVÄRINTA, Tero ; HODNE, Anne K. ; STANGELAND, Elin: *Contemporary Issues of Enterprise Content Management - The Case of Statoil*. In: *Scandinavian J. Inf. Systems* 18 (2006), Nr. 2
- [MS91] McDONNEL, E.D. ; SOMMERVILLE, G.E.: *Corporate Reengineering that Follows the Design of Document Imaging*. In: *Information Strategy: The Executive's Journal* 3 (1991), S. 5 ff.
- [MV02] MARKL, Harry ; VOGEL, Rüdiger: *DMS Deutschland Sektor Report / CONCORD Equity Research*. 2002. – Forschungsbericht

- [Nak02] NAKANO, Russell: *Web Content Managemern: A Collaborative Approach*. Addison-Wesley, 2002
- [NOS⁺09] NEUROTH, Heike ; OSSWALD, Achim ; SCHEFFEL, Regine ; STRATHMANN, Stefan ; JEHN, Mathias: *nestor Handbuch Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*. Verlag Werner Hülsbusch, 2009
- [NT94] NEUMAN, B.C. ; Ts'O, T.: Kerberos: an authentication service for computer networks. In: *Communications Magazine, IEEE* 32 (1994), Sep, Nr. 9, S. 33–38
- [Nue06] NUESCHELER, David: *JSR-170 Content Repository for Java™ technology API*. <http://www.jcp.org/en/jsr/detail?id=170>. Version: 2006
- [Nue09] NUESCHELER, David: *JSR-283 Content Repository for Java™ Technology API Version 2.0*. <http://jcp.org/en/jsr/detail?id=283>. Version: 2009
- [OHE99] ORFALI, Robert ; HARKEY, Dan ; EDWARDS, Jen: *The Essential Client/Server Survival Guide*. JohnWiley & Sons, 1999
- [OMG09a] OMG: *OMG Unified Modeling Language™ (OMG UML), Infrastructure Version 2.2*. <http://www.omg.org/spec/UML/2.2/Infrastructure>. Version: Februar 2009
- [OMG09b] OMG: *OMG Unified Modeling Language™ (OMG UML), Superstructure Version 2.2*. <http://www.omg.org/spec/UML/2.2/Superstructure>. Version: Februar 2009
- [Ora09] ORACLE CORPORATION (Hrsg.): *Oracle WebLogic Server 11g Standard Edition*. Oracle Corporation, 2009. <http://www.oracle.com/appserver/weblogic/standard-edition.html?origref=http://www.oracle.com/us/products/middleware/application-server/index.htm>. – Abgefragt 09.10.2009
- [Pfa95] PFAFF, Thomas: *Dokumentenmanagement. Das papierlose Büro?: Konzepte, Technologien, Tips*. VDE-Verlag, 1995
- [PGK88] PATTERSON, David A. ; GIBSON, Garth ; KATZ, Randy H.: A case for redundant arrays of inexpensive disks (RAID). In: *SIGMOD '88: Proceedings of the 1988 ACM SIGMOD international conference on Management of data*. New York, NY, USA : ACM, 1988, S. 109–116
- [PM05] PÄIVÄRINTA, Tero ; MUNKVOLD, Bjørn E.: Enterprise Content Management: An Integrated Perspective on Information Management. In: *HICSS*, 2005
- [Pot08] POTTS, Jeff: *Alfresco Developer Guide*. Packt Publishing, 2008

- [Rau95] RAUFER, Heinz: Ein Werkzeug zur Analyse und Modellierung von Geschäftsprozessen als Voraussetzung für effizientes Workflow-Management. In: *Wirtschaftsinformatik* 5 (1995), S. 467 ff.
- [Red09] RED HAT (Hrsg.): *Hibernate Framework*. Red Hat, August 2009. <https://www.hibernate.org/>. – Abgefragt 06.08.2009
- [Rei00] REICHERT, Manfred: *Dynamische Ablaufänderungen in Workflow-Management-Systemen*, Universität Ulm, Diss., 2000
- [Rin04] RINDERLE, Stefanie: *Schema Evolution in Process Management Systems*, Universität Ulm, Diss., 2004
- [Sal06] SALOMON, David: *Data compression: the complete reference*. Springer Verlag, 2006
- [SBC⁺08] SHEGDA, Karen M. ; BELL, Toby ; CHIN, Kenneth ; GILBERT, Mark R. ; MACCOMASGAIGH, Mick: *Magic Quadrant for Enterprise Content Management / Gartner Inc. 2008*. – Forschungsbericht
- [SC05] SHEGDA, Karen M. ; CHIN, Kenneth: *Document Management Choices Improve for SMB Audience / Gartner Inc. 2005*. – Forschungsbericht
- [SDF⁺03] SHAVOR, Sherry ; D'ANJOU, Jim ; FAIRBROTHER, Scott ; KEHN, Dan ; KELLERMAN, John ; MCCARTHY, Pat: *The Java Developer's Guide to Eclipse*. Addison-Wesley, 2003
- [Sei00] SEIDLMEIER, Heinrich: *Prozessmodellierung zur Gestaltung von Dokumentenmanagementsystemen / Hochschule für angewandte Wissenschaften Fachhochschule Rosenheim. 2000*. – Forschungsbericht
- [Sei06] SEIDLMEIER, Heinrich: *Prozessmodellierung mit ARIS. 2. akt. Auflage*. Vieweg Verlag, 2006
- [Sha06] SHARIFF, Munwar ; WALTER, Mike W. (Hrsg.): *Alfresco Enterprise Content Management Implementation*. Packt Publishing, 2006
- [SJW05] SCHEER, August-Wilhelm (Hrsg.) ; JOST, Wolfram (Hrsg.) ; WAGNER, Karl (Hrsg.): *Von Prozessmodellen zu lauffähigen Anwendungen: ARIS in der Praxis*. Springer Verlag, 2005
- [SMS08] STEINBRECHER, Wolf ; MÜLL-SCHNURR, Martina: *Prozessorientierte Ablage : Dokumentenmanagement-Projekte zum Erfolg führen ; praktischer Leitfaden für die Gestaltung einer modernen Ablagestruktur*. Gabler, 2008

- [SS02] SPALL, Andreas ; SOHN, Dirk: JAAS Tutorial: Single Sign-On mit Kerberos. (2002)
- [SUN01a] SUN MICROSYSTEMS, INC. (Hrsg.): *Java™ Authentication and Authorization Service (JAAS) LoginModule Developer's Guide*. SUN Microsystems, Inc., August 2001. <http://java.sun.com/j2se/1.4.2/docs/guide/security/jaas/JAASLMDevGuide.html>. – Abgefragt 22.12.2009
- [SUN01b] SUN MICROSYSTEMS, INC. (Hrsg.): *Java™ Authentication and Authorization Service (JAAS) Reference Guide for the J2SE Development Kit 5.0*. SUN Microsystems, Inc., August 2001. <http://java.sun.com/j2se/1.5.0/docs/guide/security/jaas/JAASRefGuide.html>. – Abgefragt 14.12.2009
- [TH06] THIERAUF, Robert J. ; HOCTOR, James J.: *Optimal Knowledge Management: Wisdom Management Systems Concepts And Applications*. IGI Publishing, 2006
- [Tie02] TIEMEYER, Ernst: *Projekte erfolgreich managen*. Beltz, 2002
- [VMw09] VMWARE CORPORATION (Hrsg.): *SpringSource*. VMware Corporation, August 2009. <http://www.springsource.com/>. – Abgefragt 03.08.2009
- [WLKC01] WEINTRAUB, A. ; LOGAN, D. ; KANDERS, G. ; CALVERT, M.: *Implementing an Integrated Document Management Strategy / GartnerGroup*. 2001. – Forschungsbericht
- [Wol08] WOLZ, Jonas: *Erweiterte Kontroll- und Datenflusskonzepte in ADEPT*, Universität Ulm, Diplomarbeit, 2008
- [ZGB05] ZÖLLER, Bernhard ; GULBINS, Jürgen ; BAUMEISTER, Hans D. ; INFORMATIONSSYSTEME E.V., VOI (Verband O. (Hrsg.): *Dokumenten-Management Vom Archiv zum Enterprise-Content-Management*. Verband Organisations- und Informationssysteme, 2005

Danksagung

An dieser Stelle möchte ich mich bei allen Personen bedanken, die mich während der Erstellung dieser Arbeit gefördert und unterstützt haben:

Allen voran möchte ich mich bei meinen Eltern für ihre Unterstützungen in jeglicher Hinsicht während meines Studiums bedanken. Ohne sie wäre dies alles nicht möglich gewesen.

Prof. Dr. Manfred Reichert danke ich für seine Anregungen, sowie PD Dr. Stefanie Rinderle-Ma für die Bereitschaft, diese Diplomarbeit zu begutachten.

Besonders möchte ich mich bei meinem Betreuer Ulrich Kreher für seine engagierte Unterstützung während der gesamten Zeit bedanken. Dank seiner konstruktiven Kritik an den Zwischenergebnissen und seiner Diskussionsbereitschaft konnte die Arbeit ständig weiterentwickelt werden. Ich wünsche ihm und auch den Mitarbeitern der AristaFlow GmbH alles Gute und eine erfolgreiche Zukunft.

Für die Korrekturen und Diskussionen bedanke ich mich bei Martina Locher, Thomas Dominik, Hannes Rometsch, Alexander Forschner, Oliver Gableske und Johannes Kast.

Matrikelnummer: 494899

Erklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Ulm, den 12. Januar 2010

.....

Marian Lindenlaub