

# DocPortal

## Projekt

Umsetzung eines  
Dokument-Servers

als MyCoRe-Anwendung

von

Armin Angelus (Uni Halle)  
Uwe Franke (Uni Leipzig)  
Kathleen Krebs (Uni Hamburg)  
Ulrike Krönert (Uni Jena)  
Jens Kupferschmidt (Uni Leipzig)

**Leipzig, 11. Sep. 2006**  
**Dokumentationsversion 1.3**

## **Vorwort**

In der folgenden Dokumentation sollen alle Design-Kriterien für einen Dokumenten-Server auf Basis des MyCoRe-Projektes festgehalten werden. Das Datenmodell soll dabei sowohl für Hochschulschriften, als auch für Lehr- und Lernmaterialien und allgemeine Multimedia-Objekte verwendbar sein. Gleichzeitig sollen Schnittstellen zu Einrichtungen wie der Deutschen Nationalbibliothek und anderer Einrichtungen Beachtung finden. Das DocPortal-Projekt baut auf dem MyCoRe-Kern in der derzeitigen stabilen Version 1.3 inklusive der aktuellen Bug-Fixes auf.<sup>1</sup>

---

1 <http://www.mycore.de>

## Inhaltsverzeichnis

1	Ziele des Projektes.....	5
1.1	Ziel aus Sicht der Bibliothek.....	5
1.2	Ziele der Fakultäten, Institute und Einrichtungen.....	5
1.3	Die Ziele aus technischer Sicht.....	5
1.4	Zusammenarbeit mehrerer DocPortal-Anwendungen.....	6
2	Die Gestaltung eines allgemeinen Dokumenten-Servers mit DocPortal.....	7
2.1	Allgemeines.....	7
2.2	Das Datenmodell.....	8
2.2.1	Allgemeines zum MyCoRe-Datenmodell.....	8
2.2.2	Erweiterung der MyCoRe Datentypen durch DocPortal.....	11
2.2.3	Das Datenmodell von DocPortal.....	11
2.2.3.1	Eine Übersicht.....	11
2.2.3.2	Die Document-Daten.....	11
2.2.3.3	Das Personen- und Einrichtungen-Datenmodell.....	16
2.2.4	Klassifikationen.....	19
2.2.4.1	Subjekte.....	19
2.2.4.2	Herkunft.....	20
2.2.4.3	Typ.....	21
2.2.4.4	Format.....	22
2.2.4.5	Sprache.....	22
2.2.4.6	Nationalität.....	22
2.2.4.7	Übersicht der DocPortal-Klassifikationen.....	22
2.3	Das Datensicherungskonzept von DocPortal.....	23
2.3.1	Allgemeines.....	23
2.3.2	Die Dateneingabe und erforderliche Rechte .....	23
2.4	Arbeiten mit dem System.....	24
2.5	Features von MyCoReSample auf Basis von DocPortal.....	24
3	Die Erweiterung des Dokumenten-Servers zu einem Dissertations-Server.....	26
3.1	Allgemeines.....	26
3.2	Änderung der Pflichtfelder.....	27
4	Workflow für den Dissertations-Server der UB Leipzig.....	29
4.1	Eingaben des Doktoranten / Habilitanten.....	29
4.1.1	Eingabe der Daten zur Person.....	29
4.1.2	Eingabe der Daten zur Dissertation / Habilitation.....	29
4.2	Prüfung durch die Bibliothek.....	30
4.2.1	Prüfung und Ergänzungen in den Personendaten.....	30
4.2.2	Prüfen und Ergänzungen der Dokumentdaten.....	30
5	Links, Anmerkungen und Sonstiges.....	31
5.1	Anmerkung.....	31
5.2	Links.....	31
6	Anbindung und Integration anderer Dokumenten-Systeme.....	32
6.1	Mapping auf Datenmodelle der Deutschen Bücherei.....	32
6.1.1	Mögliches Mapping der Daten des xMetaDiss Formates der Deutschen Bibliothek und des Projektes 'Dissertation Online'.....	32
6.2	Einbindung in SWB.....	35
6.2.1	Vorbemerkung.....	35
6.2.2	Datenmodell.....	36
6.3	Einbindung in aus dem PICA-Katalog.....	37



# 1 Ziele des Projektes

## 1.1 Ziel aus Sicht der Bibliothek

Die Erfassung, Organisation und Publikation elektronischer Volltexte in Fach- und Hochschulen, so an der Universität Leipzig, unter Wahrung der Urheberrechte ist Ziel eines Dokumenten- und Publikations-Servers. Der Server muss bezüglich der Recherche eine Volltextsuche gestatten und über Schnittstellen auch weitere Recherchen, z. B. mit OAI oder WebServices ermöglichen. Der Server erlaubt das Einbringen von Dissertationen, Habilitationen und anderen Hochschulschriften zur Veröffentlichung in unterschiedlich festgelegten Datenformaten und Nutzungsregeln. Seine Nutzung im Rahmen von Verbänden muss möglich sein. Die Verantwortung für die Aufnahme von Dissertationen und Habilitationen auf dem Dokumenten-Server obliegt der Universitätsbibliothek und den Autoren. Der Server dient zugleich der Erleichterung der Aufnahme und Weitergabe der Daten durch die Bibliothek über die Realisierung von Schnittstellen.

Derzeit werden jährlich rund 600 Dissertationen und Habilitationen an der Universität Leipzig verfasst. Die Fakultäten sollen schrittweise durch Umstellung der Promotionsordnung in das Projekt einbezogen werden. Erster Nutzer ist die Fakultät für Veterinärmedizin. Hochschulschriften wie Diplomarbeiten, Vorlesungsverzeichnisse liegen in der Zuständigkeit den Fakultäten und Instituten.

## 1.2 Ziele der Fakultäten, Institute und Einrichtungen

Auch außerhalb des Bibliotheksumfeldes entstehen an einer Hochschule große Mengen an Lehr- und Lernmaterialien sowie Dokumentationen. Diese können die Form einfacher Dokumente, aber auch komplexer digitaler Objekte, wie Film oder in HTML eingebundene Animationen haben. Die Forderung dieser Einrichtungen geht also hin zu einem oder mehreren im Verbund arbeitenden einfach zu handhabenden Server(n) für multimediale Objekte, welcher von alle Bereichen der Lehre, Forschung und Verwaltung gemeinsam genutzt werden kann. Dabei soll ein einfacher Datenaustausch und -export zu anderen Systemen möglich sein.

## 1.3 Die Ziele aus technischer Sicht

Ziel des Projektes ist es, ein möglichst vielfältig anwendbares Datenmodell zu erstellen, welches zum einen für den digitalen Dissertations- und Habilitation-Server gelten soll, zum anderen aber auch im Bereich allgemeiner Multimedia- und Dokumenten-Server einsetzbar ist. Dabei muss beachtet werden, dass sich die Datenmodelle der bestehenden MILESS-Implementationen (Essen, Jena, Münster, Freiburg, ...) darin abbilden lassen. Eine Nachnutzung im MyCoRe Umfeld wird angestrebt.

Besonderes Augenmerk gilt auch der Kompatibilität zu den Datenmodellen der Deutschen Nationalbibliothek (xMetaDiss) wie auch dem Südwest-Verbund (SWB) oder PICA. Hier soll eine Anbindung über eine Datentransformation mittels XSLT erfolgen. Über diese Schnittstellen kann auch ein Verbund mit Opus 3.0 realisiert werden. Entwickelte Schnittstellen sollen dann wieder in den MyCoRe-Kern zurückfließen. Weiterhin ist ein Export der Metadaten nach Dublin Core und die Bereitstellung für OAI vorgesehen.

Dokumente und Objekte, welche nachhaltig dargeboten werden sollen, werden eine eindeutige URL erhalten.

Das gesamte Projekt soll eine Anwendung des MyCoRe-Kerns sein. Somit ist auch für die Zukunft ein Migrationspfad auf Basis dieser Open-Source-Lösung garantiert. Neue Backends können durch die Weiterentwicklung von MyCoRe problemlos integriert werden. Gleichzeitig sorgt die Nutzung der MyCoRe-Komponenten in anderen Anwendungsbereichen wie universitäre Sammlungen für eine tiefgründiges Anwenderwissen bei geringem Eigenentwicklungsaufwand.

## 1.4 Zusammenarbeit mehrerer DocPortal-Anwendungen

MyCoRe ist so konzipiert, dass mehrere separat laufende Anwendungen von MyCoRe mit dem selben Datenmodell (z. B. DocPortal) zu einer virtuellen Anwendung zusammengeschaltet werden können. Die Abfrage/Antwortmechanismen zwischen den Anwendungs-Instanzen sind dabei über einen Webservice realisiert.

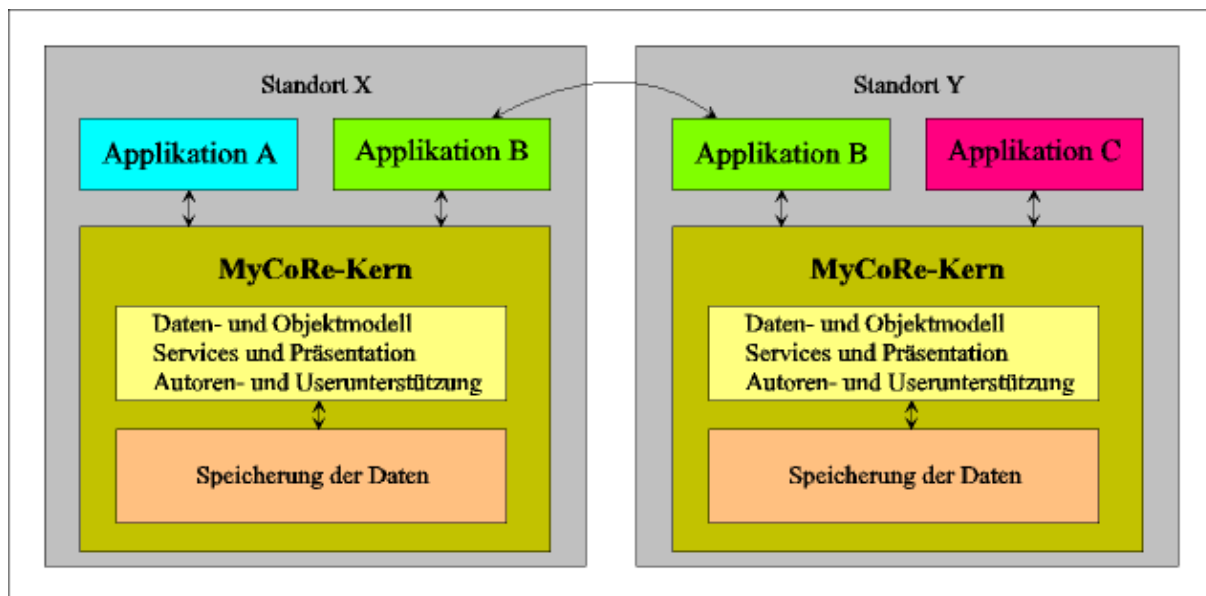


Abbildung 1: Verbindung mehrerer MyCoRe Anwendungen

## 2 Die Gestaltung eines allgemeinen Dokumenten-Servers mit DocPortal

### 2.1 Allgemeines

In diesem Kapitel soll die Gestaltung eines allgemeinen Dokumenten-Servers auf Basis des MyCoRe-Kerns, der DocPortal-Basisschicht sowie ggf. der speziellen Applikationsschicht beschrieben werden. Diese Einteilung ist sinnvoll, um jeweils Kernkomponenten von anwendungsspezifischen Dingen und diese wiederum von Spezialausprägungen wie Farbe und Logo zu trennen. Eine Übersicht der einzelnen Layer liefert die folgende Grafik.

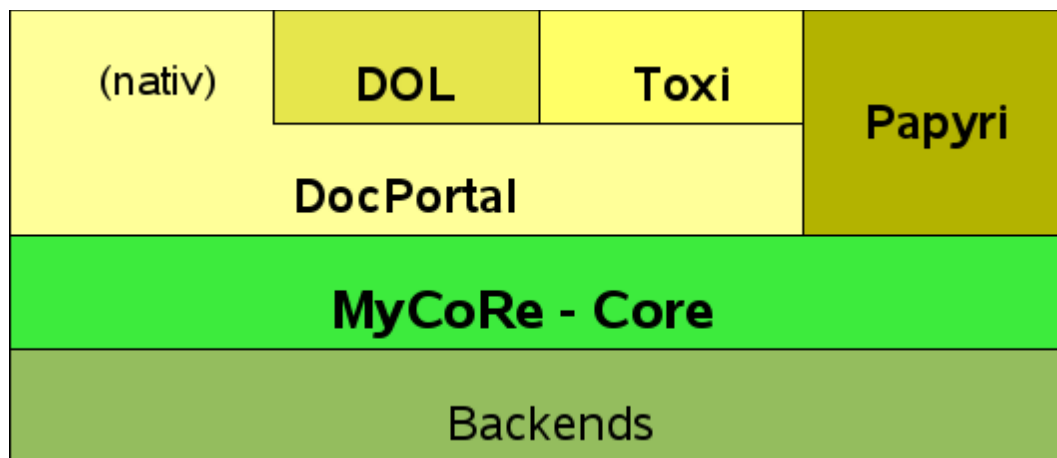


Abbildung 2: Allgemeine Komponentenübersicht

Die DocPortal-Anwendung besteht aus vier Komponentenschichten, welche dafür sorgen, dass mehrfach verwendbare Code-Teile nur einmal im Gesamtsystem gehalten werden müssen. Die einzelnen Schichten können wie folgend charakterisiert werden:

#### Backends

Dieser Teil bezeichnet Komponenten von Drittanbietern, welche zur Speicherung der Daten einer MyCoRe-Anwendung genutzt werden. MyCoRe unterstützt zur Zeit als kommerzielles Produkt den IBM DB2 Content Manager. Als weiteres Paket steht eine Lösung aus den freien Produkten MySQL oder HSQLDB mit Apache Lucene zur Verfügung. Am Ausbau für weitere Backends wird derzeit gearbeitet. Die Einbindung der Backends ist im MyCore-Kern implementiert.

#### MyCoRe-Kern

Der MyCoRe-Kern enthält alle allgemeinen Klassen, Konfigurationsdateien und Stylesheets, welche zur Gestaltung einer beliebigen MyCoRe-Anwendung erforderlich sind. Alle Code-Dateien des MyCoRe-Kerns sind in dem JAR-File *mycore.jar* enthalten. Diese Datei sowie die Stylesheets und Modul-Daten werden komplett so in die Anwendungsebene übernommen.

#### DocPortal

DocPortal ist eine Anwendungsbasis. Hier sind alle gemeinsamen Konfigurationsdateien für die eigentlichen Applikationen eines Dokumenten-Servers gespeichert. Gemeinsame Dateien sind u. a. die Konfiguration des Datenmodells, Konverter oder auch

---

diese Dokumentation. Beim Zusammenbau der eigentlichen Anwendung werden alle Daten aus DocPortal in den *webapps*-Baum kopiert und dann durch die Daten der eigentlichen speziellen Anwendung ergänzt. Für die Uni Leipzig wurde eine angepasste DocPortal-Version erstellt, die als Grundgerüst für zu erstellende Dokumenten-Server in Leipzig dient. Welche Version genutzt wird, hängt vom Source-CVS System

### **MyCoReSample**

Das MyCoReSample ist eine MyCoRe-Endanwendung, welche einem nativ DocPortal entspricht und die Funktionalität von MyCoRe demonstrieren soll. Sie basiert auf dem Datenmodell und Konzept von DocPortal und soll einen nicht modifizierten Dokumenten-Server bereitstellen. Die Original-MyCoReSample-Applikation läuft als Demo unter <http://mycoresamplelinux.dl.uni-leipzig.de/>. MyCoReSample entspricht der offiziellen DocPortal-Version unter <http://www.mycore.de/>.

### **TOXI und SEPT**

Sind zwei Dokumenten-Server, welche vom Leipziger DocPortal abgeleitet wurden und das selbe Look & Feel haben.

### **DOL**

Eine Erweiterung des Leipziger DocPortal und stellt den Dokumenten-Server DOL dar. Hier wurde die spezielle Applikationsschicht eingebaut, um die Verarbeitung von Dissertationen/Habilitationen erweitert. Gleichzeitig wurde der Geschäftsprozess der UBL für diese Arbeiten für die Arbeit mit 'Dissertationen Online' umgesetzt.

### **Papyri**

Das Papyrus-Projekt ist eine vollkommen andere MyCoRe-Anwendung, die aber den gemeinsamen MyCoRe-Kern benutzt. Somit müssen nur anwendungsspezifische Teile wie das Datenmodell oder die Präsentation neu entwickelt werden. Die Anwendung ist unter <http://papyri.uni-leipzig.de> zu finden.

Die gesamten Endanwendungen wie MyCoReSample, TOXI, SEPT, DOL oder Papyri laufen ihrerseits in einer Jetty / Apache oder IBM WebSphere Umgebung. Diese stellt die erforderliche Servlet-Engine zur Verfügung. Virtuelle Hostnamen werden unter Zuhilfenahme eines Apache-Web-Servers realisiert.

## **2.2 Das Datenmodell**

### **2.2.1 Allgemeines zum MyCoRe-Datenmodell**

In diesem Abschnitt sollen kurz die verwendeten Typen des MyCoRe-Datenmodells beschrieben werden. Eine vollständige Beschreibung finden Sie in der MyCoRe-Dokumentation.<sup>2</sup> Im MyCoRe-Datenmodell sind für die wichtigsten Typen von Daten XML-Abschnitte modelliert worden, welche wieder verwendbar eingesetzt werden können. Darüber hinaus erlaubt das System natürlich eine Entwicklung eigener Datentypen. Alle in diesem Projekt verwendeten Typen sind nachfolgend notiert.

Unter Klassifikationen wird in MyCoRe eine Sammlung feststehender Begriffe verstanden, welche hierarchisch angeordnet werden können und über eine Klassifikations- und Kategoriennummer eindeutig identifizierbar sind. Die angegebenen Datentypen können in beliebig

---

2 <http://www.mycore.de/>



vielen Kardinalitäten angegeben werden, so es das Datenmodell zulässt. Zur Vereinfachung wurden die Attribute für die Steuerung von Suche und Vererbung an dieser Stelle weggelassen. Für alle Typen erfolgt eine Plausibilitätskontrolle während der Verarbeitung. Dabei kommt sowohl Java-Code wie auch eine XML Schema Prüfung zum Einsatz.

<i>MyCoRe-Typ</i>	<i>Funktion</i>	<i>Daten</i>	
MCRMetaAddress	Enthält eine Adresse	@type @xml:lang country/text() state/text() zipcode/text() city/text() street/text() number/text()	Ein frei wählbarer Type-String. Sprache, die für dieses Datum gilt. Landesname als Text. Staat als Text. Postleitzahl als Text. Stadt als Text. Straße als Text. Hausnummer als Text.
MCRMetaBoolean	Enthält einen logischen Wert	@type text()	Ein frei wählbarer Type-String. Ein boolescher Wert.
MCRMetaClassification	Verweis auf eine Kategorie einer Klassifikation	@type @classid @catid	Ein frei wählbarer Type-String. ID der Klassifikation. ID der Kategorie.
MCRMetaInstitutionName	Enthält den Namen einer Einrichtung	@type @xml:lang name/text() nickname/text() property/text()	Ein frei wählbarer Type-String. Sprache, die für dieses Datum gilt. Name der Einrichtung. als Text Kürzel der Einrichtung. als Text Rechtlicher Stand als Text.
MCRMetaISBN	Enthält eine ISBN	@type @xml:lang text()	Ein frei wählbarer Type-String. Sprache, die für dieses Datum gilt. Die ISBN.
MCRMetaISO8601	Eine Datumsangabe	@type text()	Ein frei wählbarer Type-String. Der Datumswert.
MCRMetaNBN	Enthält eine NBN	@type @xml:lang text()	Ein frei wählbarer Type-String. Sprache, die für dieses Datum gilt. Die NBN.
MCRMetaLink	Enthält den Verweis auf eine URI	@type @xlink:type @xlink:href @xlink:title @xlink:label	Ein frei wählbarer Type-String. Der Link-Typ (immer locator). Die URI-Adresse. Der Titel des Links. Zusätzliche Textinformationen zum Link.
MCRMetaLinkID	Enthält den Verweis auf eine MCRObjektID	@type @xlink:type @xlink:href @xlink:title @xlink:label	Ein frei wählbarer Type-String. Der Link-Typ (immer locator). Die MCRObjektID. Der Titel des Links. Zusätzliche Textinformationen zum Link.
MCRMetaLangText	Enthält Text	@type @xml:lang text()	Ein frei wählbarer Type-String. Sprache, die für dieses Datum gilt. Freitext
MCRMetaPersonName	Enthält den Namen eine Person	@type @xml:lang firstnam/text() surname/text() fullname/text() lastname/text()	Ein frei wählbarer Type-String. Sprache, die für dieses Datum gilt. Der Vorname als Text. Der Nachname als Text. Der vollständige /ausländische Namenstext. Der vorherige Name.

<i>MyCoRe-Typ</i>	<i>Funktion</i>	<i>Daten</i>	
		callname/text()	Der Rufname oder eine Abkürzung als Text.
		academic/text()	Der akademische Titel als Text.
		prefix/text()	Der Namensprefix als Text.
		peerage/text()	Der Adelstitel als Text.

Tabelle 1: Datentypen des MyCoRe-Projektes

## 2.2.2 Erweiterung der MyCoRe Datentypen durch DocPortal

DocPortal verwendet nur Datentypen, welche bereits im MyCoRe-Kern vorhanden und definiert sind. Alle zugehörigen Beschreibungen sind in der MyCoRe-Dokumentation zu finden.

## 2.2.3 Das Datenmodell von DocPortal

### 2.2.3.1 Eine Übersicht

Die folgende Skizze soll eine Übersicht über die Zusammenhänge der einzelnen Daten geben. In den anschließenden Abschnitten werden diese dann näher erläutert. Die Datenmodelle der Typen **document** und **disshab** sowie **author** und **person** unterscheiden sich lediglich in der Frage der Pflichtfelder. Die Typen disshab und person sind für einen Dissertations-Server (DOL) gedacht und werden in späteren Abschnitten noch einmal besprochen.

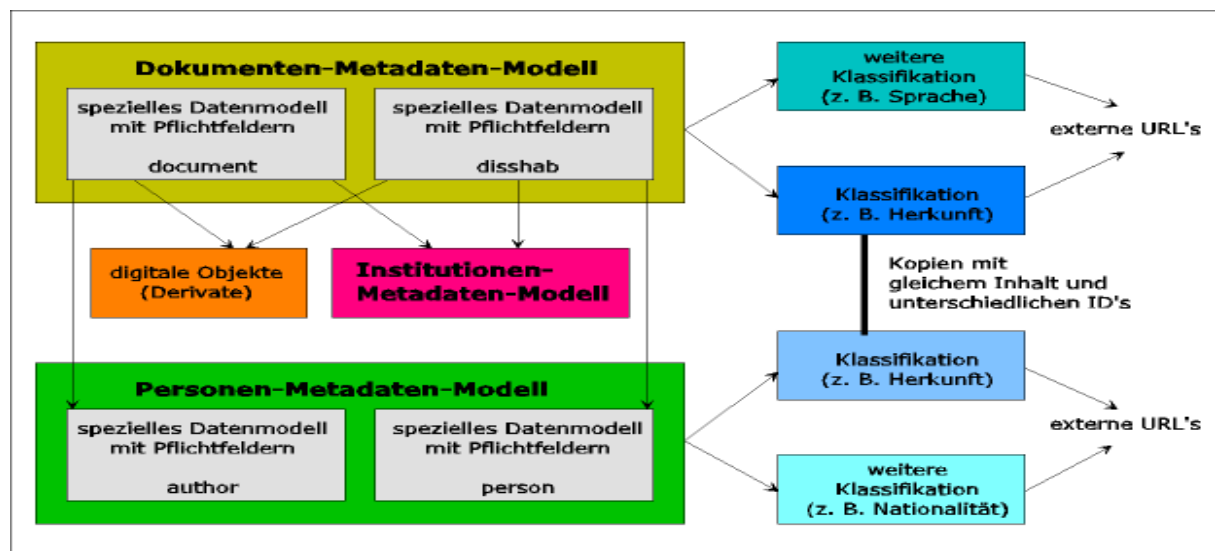


Abbildung 3: Zusammenhänge der einzelnen Metadaten

### 2.2.3.2 Die Document-Daten

Das Datenmodell der Dokumente ist ein Kompromiss zwischen den einzelnen abzubildenden Datenmodellen bisher bestehender Projekte. Gleichzeitig sollen auch Anforderungen an die Zukunft, wie das **xMetaDiss** Datenmodell oder OAI berücksichtigt werden. Entscheidend für die Gestaltung der Daten mit MyCoRe 1.3 sind vor allem die MyCoRe-Datentypen. Die

Festlegungen zur Wiederholbarkeit der Angaben bezieht sich immer auf **eine** Sprache. Die meisten Felder sind optional und können bei Bedarf verwendet werden.

Einzelne Applikationen werden nur einen Teil der angegebenen Felder ausfüllen. Es ist daher sinnvoll sich auf eine allgemeine Mindestmenge von Pflichtfeldern zu einigen, um eine korrekte Suche über mehrere Instanzen und Projekte zu gestatten. Bei einigen Feldern ist dies aber von den Spezifika des jeweiligen Projektes abhängig.

Die Suchmöglichkeiten (parametrisch/Freitext) beziehen sich auf die Metadaten. Hinzu kommt die Volltextsuche im Dokument. Die Felder des **Dublin Core Datenmodells** sollten immer implementiert und wenn möglich mit Daten gefüllt sein. Auf sie beziehen sich auch die Suchen externer Datenanbieter bzw. Teilnehmer am MyCoRe- oder OAI Datenverbund.

Um späteren Problemen mit nicht-Latin1-Sprachen aus dem Weg zu gehen sollen alle Metadatensätze und die Internet-Anwendung UTF-8 als Codierung verwenden.

Nr.	Bezeichner	Bemerkung	Pflicht	wiederhol.	Suche	MCR-Type
1	Titel (DC)	Haupttitel und ggf. weitere Titel	ja	ja	param. / Freitext	Titel MCRMetalangText
2	Creator (DC)	Name eines Autors ohne Verweis	ja	ja	Param. / Freitext	Name MCRMetalangText
3	CreatorLink	Daten des Autor oder Schöpfer des Objektes	nein	ja	param.	Creator MCRMetalinkID
4	Subject (DC)	Ordnungskriterien in Klassifikationen	ggf.	ja	param.	Subject MCRMetalangText
5	Origin	Zugehörigkeit zu einer Einrichtung als Klassifikation	ja	ja	param.	Origin MCRMetalangText
6	Description (DC)	Kurzbeschreibung des Objektes	nein	ja	param. / Freitext	Description MCRMetalangText
7	DescriptURL	Link zu Kurzbeschreibungen	nein	ja	param.	DescriptURL MCRMetalink
8	Publisher (DC)	Name des Veröfentlichters des Objektes	nein	ja	param. / Freitext	Publisher MCRMetalangText
9	Publisher-Link	Daten des Veröfentlichters des Objektes	nein	ja	param.	Publisher MCRMetalinkID
10	Contributor (DC)	Name des Beteiligten an der Erstellung des Objektes	nein	ja	param. / Freitext	Contributor MCRMetalangText
11	ContribLink	Daten des Beteiligten an der Erstellung des Objektes	nein	ja	param.	Contributor MCRMetalinkID
12	Date (DC)	Datumsangaben zum Objekt	nein	ja	param.	Date MCRMetalangText
13	Type (DC)	Typ des Objektes als Klassifikation	ja	ja	param.	Type MCRMetalangText
14	Format (DC)	Format des Objektes als Klassifikation	ja	ja	param.	Format MCRMetalangText
15	Identifizier (DC)	Angaben zur Identifikation des Objektes	nein	ja	param. / Freitext	Identifizier MCRMetalangText
16	Source (DC)	Angaben zu den Quellen des Objektes	nein	ja	param. / Freitext	Source MCRMetalangText
17	SourceLink	Link zu Angaben zu den Quellen des Objektes	nein	ja	param.	SourceLink MCRMetalink
18	Languages (DC)	Sprache als Klassifikation	nein	ja	param.	Languages MCRMetalangText
19	Keywords	Schlüsselworte als verbaler Text	nein	ja	param. / Freitext	Keywords MCRMetalangText
20	Coverage (DC)	Angaben zu der Erstreckung des Objektes	nein	ja	param. / Freitext	Coverage MCRMetalangText
21	CoverageLink	Link zu Angaben zu der Erstreckung des Objektes	nein	ja	param.	CoverageLink MCRMetalink
22	CoverageDate	Datumsangaben zur Erstreckung des Objektes	nein	ja	param.	Date MCRMetalangText
23	Relation (DC)	Textlicher Verweis auf externe Referenzen	nein	ja	param. / Freitext	Relation MCRMetalangText
24	RelationLink	Verweise auf externe Referenzen	nein	ja	param.	RelationLink MCRMetalink
25	RelationISBN	ISBN als Relation.	nein	nein	param.	ISBN MCRMetalangText
26	Rights (DC)	Angaben zu den Rechten des Objektes	nein	ja	param. / Freitext	Rights MCRMetalangText
27	RightsLink	Link zu Angaben zu den Rechten des Objektes	nein	ja	param.	RightsLink MCRMetalink
28	Size	Angaben zu Seitenanzahl, Bilder, Tabellen usw.	nein	ja	param.	Size MCRMetalangText
29	Place	Erscheinungsort	nein	nein	param./ Freitext	ISBN MCRMetalangText
30	ISBN	eindeutige ISBN	nein	nein	param.	ISBN MCRMetalangText
31	NBN	eindeutige NBN	nein	nein	param.	NBN MCRMetalangText
32	URN	eindeutige URN	nein	ja	param.	URN MCRMetalink
33	DDBContact	Eindeutiger Identifizierer der DB	nein	nein	nein	DDBContact MCRMetalangText
34	Notes	Anmerkungen zum Objekt	nein	ja	Freitext	Notes MCRMetalangText
35	Citation	Zitierweise	nein	ja	nein	Citation MCRMetalangText

Tabelle 2: Das neue Metadaten-Modell der Dokumente

Ausfüllhinweise:

Die Zahlen in Klammern geben die maximale Zeichenlänge pro Kardinalität an (wichtig für Systeme mit einem Mapping auf relationale Datenbanken – IBM Content Manager).

<i>Nr.</i>	<i>Bezeichner</i>	<i>Ausfüllhinweis</i>
1	Titel (1024)	<ul style="list-style-type: none"> <li>• Pro Titeltyp können mehrere Sprachen benutzt werden.</li> <li>• Haupttitel werden mit <b>type="main"</b> markiert.</li> <li>• Alternative Titel werden mit <b>type="alt"</b> markiert.</li> <li>• Untertitel werden mit <b>type="subtitle"</b> markiert.</li> </ul>
2	Creator (128)	<ul style="list-style-type: none"> <li>• Für Habilitationen und Dissertationen Verfasser in Vorlageform.</li> <li>• Sonst Autorname als Text.</li> </ul>
3	CreatorLink	Verweis auf einen Personen-Datensatz.
4	Subject	<ul style="list-style-type: none"> <li>• Verweise auf Kategorien von Klassifikationen, in die das Dokument eingeordnet ist.</li> <li>• Für Habilitationen und Dissertationen in Leipzig --&gt; Regensburger Systematik (ID der RS)</li> <li>• Weitere bibliotheksinterne Sachgruppen</li> </ul>
5	Origin	Auswahlliste, zu welcher Einrichtung das Objekt gehört.
6	Description (4096)	<p>Beschreibende Informationen.</p> <ul style="list-style-type: none"> <li>• Für den Beschreibungstext <b>type="description"</b>.</li> <li>• Für Abstract <b>type="abstract"</b>.</li> <li>• Für die Inhaltsangabe <b>type="content"</b>.</li> </ul>
7	DescriptURL	<p>Verweis auf externe Beschreibungstexte.</p> <ul style="list-style-type: none"> <li>• Für den Beschreibungstext <b>type="description"</b>.</li> <li>• Für Abstract <b>type="abstract"</b>.</li> <li>• Für die Inhaltsangabe <b>type="content"</b>.</li> </ul>
8	Publisher (128)	Textliche Bezeichnung des/derHerausgeber(s).
9	PublisherLink	Verweis auf einen Personen- oder Institutionen-Datensatz.
10	Contributor (128)	<p>Textliche Bezeichnung des/der Beteiligten.</p> <ul style="list-style-type: none"> <li>• Für den Typ können Informationen wie „<b>advisor</b>“, „<b>mentor</b>“, „<b>drawer</b>“ usw. angegeben werden.</li> </ul>
11	ContributorLink	<p>Verweis auf einen Personen- oder Institutionen-Datensatz.</p> <ul style="list-style-type: none"> <li>• Für den Typ können Informationen wie „<b>advisor</b>“, „<b>mentor</b>“, „<b>drawer</b>“ usw. angegeben werden.</li> </ul>
12	Date	<ul style="list-style-type: none"> <li>• Für allgemeine Dokumente Datum der Einstellung <b>type="create"</b>.</li> <li>• Für das Datum der Einreichung zur Dis./Habil.der Arbeit <b>type="submit"</b>.</li> <li>• Für das Datum der Verteidigung zur Dis./Habil.der Arbeit <b>type="accept"</b>.</li> <li>• Für das Datum der Beschlussfassung zur Dis./Habil. der Arbeit <b>type="decide"</b>.</li> </ul>
13	Type	Auswahlliste inklusive „Dissertation“ und „Habilitation“
14	Format	Auswahlliste
15	Identifier (128)	ggf. Bibliothekssignatur
16	Source (1024)	
17	SourceLink	
18	Language	Siehe Anmerkungen zur Sprachnotation.
19	Keywords (128)	Verbal anzugebende Schlüsselworte / Stichworte.
20	Coverage (1024)	

<b>Nr.</b>	<b>Bezeichner</b>	<b>Ausfüllhinweis</b>
21	CoverageLink	
22	CoverageDate	
23	Relation (1024)	<ul style="list-style-type: none"> <li>• Angabe zum Erscheinen des Werkes.</li> <li>• Verbaler Verweis auf vorangegangene Versionen.</li> </ul>
24	RelationLink	Link auf vorangegangene Versionen.
25	RelationISBN	Verweis auf eine ISBN Nummer.
26	Right (1024)	Verbale Beschreibung der Urheberrechte.
27	RightsLink	Verweis auf eine URL mit den Lizenz- und/oder Urheberrechten.
28	Size (1024)	Verbale Aufzählung von Seite, Abbildungen, usw. lt. Vorgabe der Bibliothek. Für Leipzig z. B. : <b>xxx S. : Ill. , graph. Darst.</b>
29	Place (1024)	Erscheinungsort
30	ISBN (32)	
31	NBN (256)	
32	URN (256)	
33	DDBContact (1024)	
34	Notes (4096)	<ul style="list-style-type: none"> <li>• Für SWB Fußnoten <b>type="feet"</b>.</li> <li>• Für SWB Kommentare <b>type="coment"</b>.</li> </ul>
35	Citation 4096)	Verbale Angabe der Zitierweise.

Tabelle 3: Ausfüllhinweise zum neuen Datenmodell der Dokumente

### 2.2.3.3 Das Personen- und Einrichtungen-Datenmodell

In der bisherigen MILESS-Anwendungen wurde zur Darstellung der Personen und Institutionen das Modell der **LegalEntity** verwendet. Eine Erweiterung und Aufspaltung in natürliche Personen und Institutionen erscheint sinnvoll, da doch ein ganzer Teil der Datenstruktur unterschiedlich ist und so auch Personendaten (z. B. zur Identifizierung des Authors) weiterverarbeitet werden können. Hier sind für natürliche Personen erheblich mehr Datenfelder vorzusehen. Eine mögliche Alternative könnte die vCard<sup>3</sup> darstellen, diese ist aber noch offen.

Für **Institutionen** ergibt sich folgendes Datenmodell:

3 <http://www.w3.org/TR/2001/NOTE-vcard-rdf-20010222/>



Nr.	Bezeichner	Bemerkung	Pflicht	wiederhol.	Suche	MCR-Type
1	Name	Angaben zum Namen einer Institution	ja	nein	param. /Freitext	Corporation MCRMetaCorporationName
2	Address	Angaben zur Adresse	nein	ja	nein	Address MCRMetaAddress
3	Phone	Telefonnummern / Fax	nein	ja	nein	Phone MCRMetaLangText
4	URL	URL's	nein	ja	nein	URL MCRMetaLangText
5	eMail	eMails's	nein	ja	nein	eMail MCRMetaLangText
6	Note	Bemerkungen	nein	ja	nein	Note MCRMetaLangText

Tabelle 4: Das neue Metadaten-Modell für Institutionen

## Ausfüllhinweise:

Nr.	Bezeichner	Ausfüllhinweis
1	Name	
2	Address	
3	Phone	<ul style="list-style-type: none"> <li>• für Telefonnummern <b>type=phone</b></li> <li>• für Faxe <b>type="fax"</b></li> </ul>
4	URL	
5	eMail	
6	Note	
7	ServiceDates	
8	ServiceFlags	

Tabelle 5: Ausfüllhinweise für das neue Metadaten-Modell für Institutionen

Für **Personen** hingegen ist ein umfangreicheres Datenmodell erforderlich:

Nr.	Bezeichner	Bemerkung	Pflicht	wiederhol.	Suche	MCR-Type
1	Name	Angaben zum Namen einer Person	ja	nein	param. Freitext	Person MCRMetaPersonName
2	Female	Angaben zum Geschlecht der Person	ja	nein	param.	Person MCRMetaBoolean
3	Institution	Verweis auf die Institution, zu der die Person gehört	ja	ja	param.	Institution MCRMetaClassification
4	Address	Angaben zur Adresse	nein	ja	nein	Address MCRMetaAddress
5	Phone	Telefonnummern	nein	ja	nein	Phone MCRMetaLangText
6	Date	Datumsangaben wie Geburtsdatum, usw.	nein	ja	param.	Date MCRMetaISO8601
7	Profession	Berufsbezeichnung / Amt	nein	ja	nein	Profession MCRMetaLangText
8	ProfClass	Berufsbezeichnung / Amt als Klassifikationseintrag	nein	ja	param.	Profession MCRMetaClassification
9	National	Nationalität	nein	ja	param.	National MCRMetaClassification
10	URL	URL's	nein	ja	nein	URL MCRMetaLink
11	eMail	eMails's	nein	ja	nein	eMail MCRMetaLangText
12	Reference	Externe Referenzen	nein	ja	nein	Reference MCRMetaLink
13	Note	Anmerkungen	nein	ja	nein	Note MCRMetaLangText
14	Publications	Publikationen	nein	ja	nein	Publication MCRMetaLangText

Tabelle 6: Das neue Metadaten-Modell für Personen

## Ausfüllhinweise:

Nr.	Bezeichner	Ausfüllhinweise
1	Name	
2	Female	• weiblich ist <b>true</b> ; männlich ist <b>false</b>
3	Institution	
4	Address	• Für das Büro ist <b>type="office"</b> anzugeben. • Für die private Adresse ist <b>type="private"</b> anzugeben.
5	Phone	
6	Date	• Für das Geburtsdatum ist <b>type="birth"</b> anzugeben.
7	Profession	• Für den Beruf ist <b>type="profession"</b> anzugeben. • Für die Tätigkeit ist <b>type="job"</b> anzugeben.
8	ProfClass	
9	National	
10	URL	
11	eMail	
12	Reference	
13	Note	
14	Publications	

Tabelle 7: Ausfüllhinweise für das neue Metadaten-Modell für Personen

## 2.2.4 Klassifikationen

Klassifikationen sollen eine Suche / Präsentation von einheitlichen Begriffen gewährleisten. Da letztendlich jede Einrichtung selbst für die Auswahl der verwendeten Klassifikation verantwortlich ist, können hier nur Empfehlungen ausgesprochen werden, um eine gemeinsame Instanzen-übergreifende Suche zu ermöglichen. Achtung, in einem Verbund sollte man sich darüber im Klaren sein, dass Streichungen von Kategorien einer Klassifikation zu Fehlern führen können! Ergänzungen hingegen sind unkritisch.

Werden Änderungen an im Dokumenten-Server benutzten Klassifikationen durchgeführt, werden auch alle Umarbeitungen (an Büchern, etc) durchgeführt. Da sich ständig etwas ändert, muss auch die Dokumenten-Server-Administration in den Newsletter/Update-Workflow der Bibliothek einbezogen werden.

### 2.2.4.1 Subjekte

Die Klassifikationen der Subjekte sind wohl das größte Problem bei der Suche nach einem gemeinsamen Nenner. Andererseits kann hier auch toleriert werden, dass eine große Individualität der einzelnen Anwendungen herrscht, wenn man bereit ist, Diskrepanzen bei der Suche hinzunehmen bzw. dieses Feld von einer **gemeinsamen übergreifenden** Suche auszuschließen. Einige Klassifikationen sind schon für MyCoRe realisiert, andere bedürfen noch einer Umsetzung. Eine Übersicht gibt die folgende Tabelle.

<i>Klassifikation</i>	<i>in DocPortal</i>	<i>MCRObjectID</i>	<i>im Sample</i>
Basisklassifikation (GBV)	nein		
Sachgruppen der Deutschen Nationalbibliographie (DNB)	ja	DocPortal_class_00000007	ja
Dewey Decimal Classification (DDC)	ja	DocPortal_class_00000009	
Regensburger Verbundklassifikation (RKV)	nein		
Mathematics Subject Classification (msc1991)	nein		
Physics and Astronomy Classification Scheme (pacs2003)	ja	DocPortal_class_00000010	
ACM Computing Classification System (ccs1998)	nein		
Zentralblatt für Didaktik der Mathematik (zdm)	nein		
Library of Congress Classification Scheme (LCC)	nein		
Universam Decimal Classification Scheme (UDC)	nein		
National Library of Medicine Classification Scheme (NLM)	nein		
Mathematics Subject Classification (msc2000)	nein		
Schlagwortnormdatei (SWD)	nein		
Library of Congress Subject Headings Scheme (LCSH)	nein		
Medical Subject Headings Scheme (MeSH)	nein		
Unified Medical Language System Scheme (UMLS)	nein		

Tabelle 8: Potentielle Klassifikationen für DocPortal

### 2.2.4.2 Herkunft

Diese Klassifikation beinhaltet eine Liste von Grundeinträgen für die beteiligten Einrichtungen. Diese Grundeinträge sind durch den jeweiligen DocPortal-Anwender für sich um eine weitere Untergliederung näher zu spezifizieren. Ziel dieser Grundeinrichtung ist eine grobe Suchbarkeit der Daten auch von Anwendungen anderer Portal-Teilnehmer aus (Beispiel: Suche von Objekten aus Leipzig in Jena). Es erscheint sinnvoll, für die Herkunft die Möglichkeit einer Anbindung von URL's an die jeweilige Kategorie, wie sie MyCoRe bietet, auszunutzen, somit kann direkt auf die Web-Seite eines Institutes oder einer Einrichtung referenziert werden. Die im DocPortal gelieferte Grundstruktur ist:

- ◆ Unis
  - Unis.Dresden
  - Unis.Essen (steht für Duisburg-Essen)
  - Unis.Freiburg
  - Unis.Greifswald
  - Unis.Halle
  - Unis.Jena
  - Unis.Kiel

- Unis.Leipzig
- Unis.Muenster
- Unis.Rostock
- ◆ Firma
  - Firma.IBM
  - Firma.DUR
- ◆ Andere

In den abgeleiteten, auf DocPortal basierenden, Leipziger Applikationen DOL, SEPT und TOXI wurde der Eintrag der Uni Leipzig um Unterkategorien der Fakultäten und Institute erweitert.

### 2.2.4.3 Typ

Da sich das xMetaDiss Konzept nur für Dissertation und Habilitationen verantwortlich fühlt, ist eine Integration in MILESS/MyCoRe relativ einfach möglich. DocPortal sieht daher alle in MILESS bisher verwendeten Typen vor. Ergänzungen durch die Anwender von DocPortal können problemlos durchgeführt werden.

#### **In MILLES/DocPortal werden verwendet:**

- ◆ Abstrakt
- ◆ Anleitung
- ◆ Artikel
- ◆ Bild
- ◆ Buch
- ◆ Bibliographien
- ◆ Datensatz, Programm
- ◆ Diplomarbeit, Magisterarbeit
- ◆ Dissertation
- ◆ Examensarbeiten
- ◆ Forschungsarbeit
- ◆ Habilitationen
- ◆ Journal
- ◆ Lehrmaterial
- ◆ Musik (Audio)
- ◆ Musiknoten
- ◆ Organisationsinformationen
- ◆ Personen-Informationen
- ◆ Preprint
- ◆ Proceedings
- ◆ Rede, Vortrag
- ◆ Report
- ◆ Service
- ◆ Sonstiges
- ◆ Statistische Daten
- ◆ Studienapparate
- ◆ Universitäre Veröffentlichungen
- ◆ Video

#### 2.2.4.4 Format

Die Festlegung der Format-Klassifikation ist sehr schwierig. METADISS sieht hier zum Beispiel die MIME-Types vor. In MILESS/MyCoRe werden hier verbale Einstufungen benutzt. Dabei sollten wir auch bleiben, da wir ja ggf. eine Menge von Objekten mit einem Metadaten-satz versehen (Derivate). Aus der Speichertabelle dieser Derivate können die erforderlichen Mime-Types gewonnen werden.

**Im DocPortal wird verwendet** (dies entspricht auch dem SWB):

- ◆ collection
- ◆ dataset
- ◆ event
- ◆ image
- ◆ interactive\_ressource
- ◆ service
- ◆ sound
- ◆ text
- ◆ unbekannt

#### 2.2.4.5 Sprache

Die Language Klassifikation ist eine Abbildung der Sprachen nach ISO-639-1. Ggf. werden diese Sprachkürzel um Länderkürzel nach ISO Norm 3166 erweitert werden (z. B. eng-US). Um Konform zur XML-Notation zu sein, wird gemäß Spezifikation die Form **...[-CC]** gewählt, wobei ... der 3-stellige Sprachcode ist. Diesem kann sich das Länderkürzel mit Minuszeichen anschließen.

#### 2.2.4.6 Nationalität

Die Klassifikation der Nationalitäten umfasst eine einfache Liste der gängigsten Nationalitäten. Eine Ergänzung ist jederzeit möglich.

#### 2.2.4.7 Übersicht der DocPortal-Klassifikationen

Die nachfolgende Tabelle gibt eine Übersicht der mitgelieferten Klassifikationsmuster, wie sie Verwendung finden. Diese Muster müssen für eine Nachnutzung entsprechend angepasst werden. Gleichzeitig erfolgt mit dem weiteren Projektausbau auch die Erweiterung dieser Klassifikationen.

<i>MCRObjectID</i>	<i>Inhalt der Klassifikation</i>
DocPortal_class_00000001	Eine Liste der möglichen Nationalitäten.
DocPortal_class_00000002	Eine Grundliste der beteiligten Universitäten und Firmen. Diese Liste ist nicht fein Strukturiert, dies Müssen die Anwender selbst vornehmen. Im <b>description</b> Attribut der Kategorien können MCR-ObjectID's der Institutionen abgelegt werden.
DocPortal_class_00000003	Analog zu DocPortal_class_00000002..
DocPortal_class_00000004	Eine Liste der möglichen Sprachen.
DocPortal_class_00000005	Eine Liste der Typen, siehe Einbindung in SWB
DocPortal_class_00000006	Eine Liste der Formate, siehe Einbindung in SWB
DocPortal_class_00000007	Eine Liste der Sachgruppen DNB
DocPortal_class_00000008	Eine Liste der Berufe , siehe Einbindung in SWB
DocPortal_class_00000009	Die DDC-Klassifikation
DocPortal_class_00000010	Die PACS-Klassifikation

Tabelle 9: DocPortal-Klassifikationen

## 2.3 Das Datensicherungskonzept von DocPortal

### 2.3.1 Allgemeines

Wie im Geasnten MyCoRe-Projekt unter Version 1.3 so wird auch in DocPortal zur Sicherung der Daten das ACL-Konzept verwendet. Es löst das bisherige System der Privilegien, wie es bis 1.2 üblich war, ab. Im neuen ACL-System sind nun für ein Objekt alle berechtigten Nutzer und Gruppen pro Aktion (z. B. Ändern) eingetragen. Das betrifft sowohl Rechte auf Einstellen, Ändern und Löschen wie auch spezielle Leserechte.

### 2.3.2 Die Dateneingabe und erforderliche Rechte

Im DocPortal sind die Benutzergruppen und damit auch die Rechte auf Daten über ACL's miteinander verkettet. Eine Übersicht der Gruppen und User im Sample gibt die Tabelle unten.

<i>User</i>	<i>Password</i>	<i>Group</i>
administrator	alleswirdgut	admingroup
reader1A	reader1A	readergroup1
reader2A	reader2A	readergroup2
author1A	author1A	authrogroup1
author1B	author1B	authrogroup1
author2A	author2A	authrogroup2
author2B	author2B	authrogroup2
editor1A	editor1A	editrogroup1
editor1B	editor1B	editrogroup1
editor2A	editor2A	editrogroup2
editor2B	editor2B	editrogroup2

Tabelle 10: Benutzer des MyCoReSample

## 2.4 Arbeiten mit dem System

Für die Arbeit mit dem auf DocPortal basierenden MyCoReSample sowie der darauf basierenden Ableger wie TOXI oder SEPT wurden Kurzanleitungen geschrieben, welche die Arbeiten aus Sicht des Autors und aus Sicht des Bearbeiters verdeutlichen sollen. Bitte Konsultieren Sie diese Dokumente für ein besseres Verständnis der Abläufe. Sie enthalten auch zahlreiche Screen-Shots, die das genaue Vorgehen verdeutlichen. An dieser Stelle soll eine Grafik nur noch einmal den Workflow-Prozess verdeutlichen.

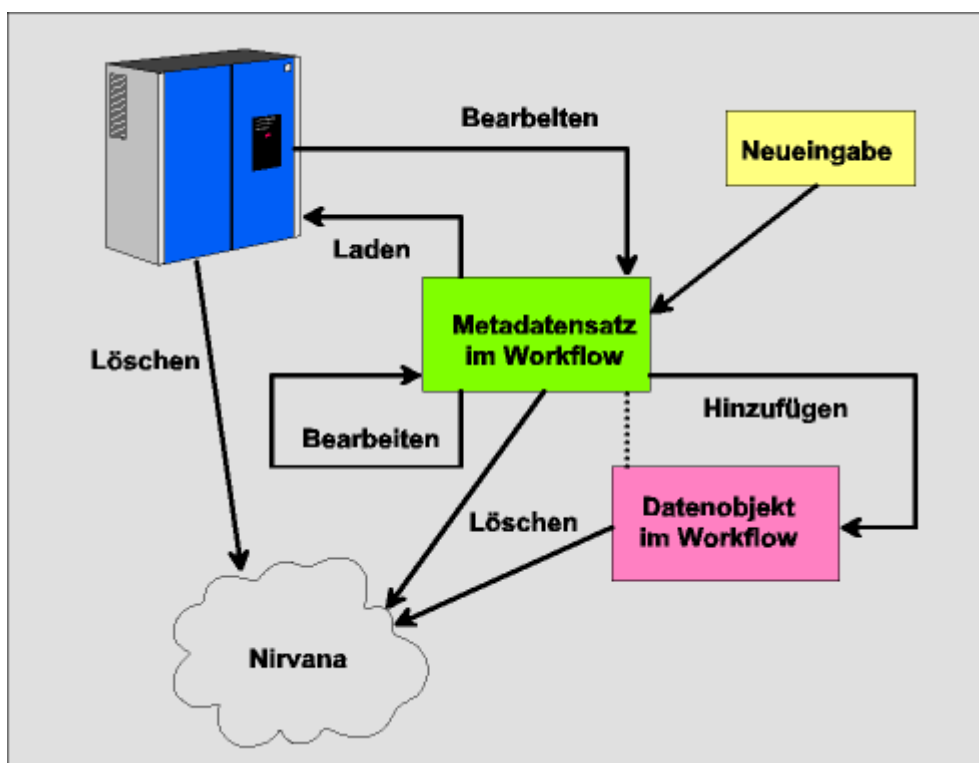


Abbildung 4: Workflow-Prozesse für einen allgemeinen Dokumenten-Server

## 2.5 Features von MyCoReSample auf Basis von DocPortal

MyCoReSample auf Basis von DocPortal hat eine ganze Reihe von Eigenschaften, welche einen vollständigen Dokumenten-Server abdecken sollen. Diese sind in Kurzfassung:

- Einteilung des Datenmodells in Dokumente sowie weitere zusätzliche beschreibende Komponenten wie Institutionen und Autoren/Personen (siehe Grafik).
- Bei der Suche wird über alle eingestellten Dokumente gesucht.
- Eine Volltextsuche ist je nach verwendetem System (IBM oder Linux) über eine Reihe von Dateitypen der Dokument-Objekte möglich (PDF, DOC usw.).
- Es können mehrere Dokumenten-Server auf DocPortal-Basis zu einem virtuellen Server zusammengeschaltet werden (Remote-Abfragen).
- Auf einem System können mehrere spezielle Installationen von DocPortal gleichzeitig und unabhängig voneinander betrieben werden. Voraussetzung ist die Trennung von SQL-Tabellen, Search-Stores und Servlet-Engines.

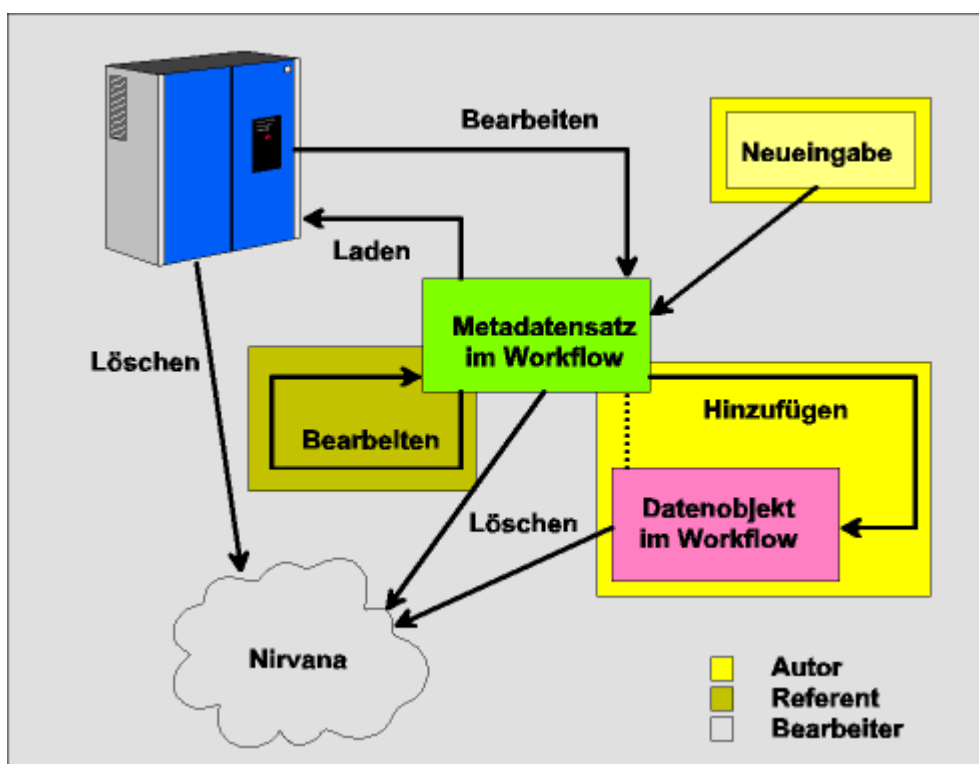


- Es wird ein umfangreiches Autorensystem mit einem integrierten Datei-Upload und kleinem Workflow angeboten.
- Die Daten und Objekte der Autoren gelangen in einen Workflow, welcher durch den Autor und dedizierte Bearbeiter editiert werden können. Die Autoren und Bearbeiter können dann das Objekt freigeben und in den Server laden.
- Es existiert eine Reihe von Klassifikationen, die allgemeingültig sind.
- Es besteht die Möglichkeit des Browsen in Klassifikationen.
- Es existiert ein einfaches, aber wirkungsvolles Benutzer- und Access-Control-System.
- Das System arbeitet mehrsprachig.
- Datenfelder können im Editor mehrfach ausgefüllt werden, z. B. mehrere Titel usw.
- Daten können von berechtigten Autoren auch nach dem Einstellen in den Server bearbeitet, verändert oder gelöscht werden.
- Das System ist für einen Dissertations-Server mit anderen Geschäftsmodellen erweiterbar.

## 3 Die Erweiterung des Dokumenten-Servers zu einem Dissertations-Server

### 3.1 Allgemeines

Dissertations-Server unterscheiden sich von reinen Dokumenten-Servern vor allem in den Geschäftsprozessen und Verantwortlichkeiten. Während beim Dokumenten-Server alle berechtigten Autoren Objekte in den Arbeitsprozess Einbringen, Manipulieren und Laden dürfen ist dies für den Bereich Dissertationen / Habilitationen aus rechtlichen Gründen eingeschränkt. Alleinig verantwortlich hierfür ist die Bibliothek der jeweiligen Universität oder Hochschule. Die Grafik zeigt den geänderten Geschäftsprozess für Dissertationen / Habilitationen.



Die bereits in Kapitel 2 vorgestellte Leipziger DocPortal-Version wurde für die Anwendung als Dissertations- / Habilitations-Server um eine Reihe von Komponenten erweitert. Das System läuft unter dem Namen **DOL**. Die Funktionalitäten, wie sie aus dem MyCoRe-DocPortal bekannt sind, blieben jedoch erhalten. Geändert hat sich:

- Das Datenmodell **document** wurde nach **disshab** kopiert. Lediglich die Pflichtfelder haben sich geändert.
- Das Datenmodell **author** wurde nach **person** kopiert. Auch hier haben sich die Pflichtfelder geändert.
- Die Konfiguration wurde dahingehend geändert, dass andere Tabellen- und Storenamen verwendet werden.

- Das User- und Rechte-System wurde erweitert, es kam die Gruppe der Bibliotheksmitarbeiter hinzu.
- Daraus abgeleitet wurde ein zusätzlicher Arbeitsablauf für Dissertationen / Habilitationen integriert.

## 3.2 Änderung der Pflichtfelder

Abhängig vom Projekttyp können eine unterschiedliche Anzahl von Pflichtfeldern für die Speicherung der Metadaten eines Dokumenten-Servers ausgemacht werden. Um diese zu unterscheiden wird das selbe Datenmodell in verschiedenen MCRObjektID-Typen abgelegt. Unterscheiden tun sich diese Typen nur in der Festlegung der Pflichtfelder. Unterschiedliche Typen können auch genutzt werden um die Pflichtfelder eines Workflow abzubilden. Für einen Dissertations/Habilitations-Server sind dies z. B. **disshab** und **person**. Via Konvertierung/Bearbeitung können die Typen ineinander überführt werden. Auch eine Suche über eine Gruppe von Typen mit gleichem Datenmodell ist in MyCoRe möglich. So bilden die Typen **document** und **disshab** den übergeordneten Typ **alldocs** und **author** und **person** den Typ **allpers**.

<i>Nr.</i>	<i>Bezeichner</i>	<i>disshab</i>	<i>document</i>
1	Titel (DC)	Pflicht	Pflicht
2	Creator (DC)	Pflicht	Pflicht
3	CreatorLink	Eintrag durch Bearbeiter	optional
4	Subject (DC)	Pflicht	Pflicht
5	Origin	Pflicht	Pflicht
6	Description (DC)	Pflicht	eine Angabe Pflicht
7	DescriptURL		
8	Publisher (DC)	optional	optional
9	PublisherLink		
10	Contributor (DC)	optional	optional
11	ContribLink		
12	Date (DC)	Pflicht	optional
13	Type (DC)	Pflicht	Pflicht
14	Format (DC)	Pflicht	Pflicht
15	Identifier (DC)	Eintrag durch Bearbeiter	optional
16	Source (DC)		optional
17	SourceLink		
18	Languages (DC)	Pflicht	
19	Keywords	optional	optional
20	Coverage (DC)	optional	optional
21	CoverageLink		
22	CoverageDate		
23	Relation (DC)	optional	optional
24	RelationLink		
25	RelationISBN		
26	Rights (DC)	Pflicht	Pflicht
27	RightsLink		
28	Size	Pflicht	optional
29	Place	Pflicht	
30	ISBN		
31	NBN		
32	URN		
33	DDBContact	Pflicht	optional
34	Notes	optional	optional
35	Citation	optional	optional

Tabelle 11: Pflichtfelder im Vergleich

## 4 Workflow für den Dissertations-Server der UB Leipzig

Die Arbeitsabläufe in diesem Bereich werden durch die Vorgaben der UBL und durch die entsprechenden Promotionsordnungen festgelegt. Die UBL hat auf diesem Rechner auch Administrationsrechte.

### 4.1 Eingaben des Doktoranten / Habilitanten

#### 4.1.1 Eingabe der Daten zur Person

Zuerst müssen für die Abgabe der Arbeit die Daten zur Person erfasst werden. Diese sind im Formular selbst einzutragen. Hierzu gehören die **Pflichtfelder**

- Name
- Institutionen
- Geburtsdatum
- Berufsbezeichnung
- Nationalität
- eMail-Adresse

#### 4.1.2 Eingabe der Daten zur Dissertation / Habilitation

Nach Eingabe der persönlichen Daten sind nun alle Felder zur abgegebenen Arbeit auszufüllen. Diese sind:

- Titel ggf. mit Alternativ- und/oder Untertitel (**Pflicht**)
- Verweis auf die Personenangaben (**Pflicht**)
- Auswahl der Fakultät bzw. des Institutes (**Pflicht**)
- Eingabe des Abstract
- Je nach Promotionsordnung Angabe des Verteidigungsdatums der Arbeit oder des Verleihungsdatums des Titels oder max. die von der DB in XMetaDiss geforderten Datumsangaben anzugeben. (**Pflicht**)
- Auswahl des Types „Dissertation“ oder „Habilitation“ (**Pflicht**)
- Auswahl des Formates (Medientyp) (Pflicht/ ggf. optional)
- Auswahl der Sprache (**Pflicht**)
- Eingabe von Stich- oder Schlüsselwörtern (sonstige Sachgruppen werden in der UBL nicht vergeben bzw. überprüft - Empfehlung: Klären ob dies für die UB entfallen kann, damit keine Falschen bibliographischen Daten durch das Eingabeformular entstehen; ?DDB/Researchermöglichkeit) (**optional**)
- ggf. Eingabe des Verweises auf etwaige Erscheinungsformen (Zeitschriften usw.) (optional)
- Eingabe zu den Rechten an dieser Arbeit (**automatisch**)
- Eingabe der Anzahl der Seiten, Illustrationen usw. (**Pflicht**)
- Eingabe des Erscheinungsortes (**Pflicht**)
- Eingabe der Zitiervorschrift (**Pflicht**)
- ggf. Eingabe von Bemerkungen (**optional**)

---

## 4.2 Prüfung durch die Bibliothek

Nachdem die Daten durch den Einreicher der Arbeit erfasst wurden, müssen Sie durch Mitarbeiter der Bibliothek geprüft und ergänzt werden.

### 4.2.1 Prüfung und Ergänzungen in den Personendaten

Es wird geprüft, ob

- der Satz im SWB vorhanden, wenn ja wird dieser angehängt (Katalogisierung 2.Zeile "LOK 1 (L1 UB)"). Wenn der Satz nicht vorhanden ist, wird neu katalogisiert. Zu Prüfen ist.

Es wird ergänzt:

- der Beruf der Person wird eindeutig einer Kategorie aus der SWB-Berufsklassifikation zugeordnet.

### 4.2.2 Prüfen und Ergänzungen der Dokumentdaten

- Der Autorname für Feld 2 wird den Personendaten entnommen und entsprechend ergänzt.
- Die Einordnung in die Regensburger Systematik (RVK) wird vom Fachreferenten durchgeführt. Als Subjekt wird diese dann von einem Schreibberechtigten eingetragen .
- Weitere Anmerkungen werden ggf. ergänzt.

## 5 Links, Anmerkungen und Sonstiges

### 5.1 Anmerkung

Da der Dokumenten-Server nicht die Aufgabe einer Verbundsoftware übernehmen kann, sind nach dem Eintrag der Autoren der Daten in den Dokumentenserver (Autor/DOL), dem Eintrag der Daten in den Verbund(Lokalsatz) (UB/SWB), nach dem Herunterladen in den Lokalsystem (Liberio). Die Meldung an die DDB muss erfolgen (DDB). Noch fehlende Daten für den Dokumentenserver (DOL) sind aus dem Lokalsatz zu ziehen.

### 5.2 Links

- MyCoRe-Homepage <http://www.mycore.de>
- Dokumenten-Server in Leipzig <http://dol.dl.uni-leipzig.de/>
- Dokumenten-Server der Toxikologie in Leipzig <http://toxi.dl.uni-leipzig.de/>
- Ein anderes MyCoRe-Projekt in Leipzig <http://papyri.dl.uni-leipzig.de/>
- OnlineDB-Regensburg (RVK) [http://www.bibliothek.uni-regensburg.de/rvko\\_neu/](http://www.bibliothek.uni-regensburg.de/rvko_neu/)
- Systematik an der UBL <http://www.bibliothek.uni-regensburg.de/Systematik/systemat.html>

---

## 6 Anbindung und Integration anderer Dokumenten-Systeme

### 6.1 Mapping auf Datenmodelle der Deutschen Bücherei

#### 6.1.1 Mögliches Mapping der Daten des xMetaDiss Formates der Deutschen Bibliothek und des Projektes 'Dissertation Online'

Mit der Beschreibung der Metadaten des Projektes xMetaDiss<sup>4</sup> der Deutschen Bibliothek und des Projektes 'Dissertation Online' in Version 0.9 liegt ein weiterer Orientierungspunkt für die Gestaltung eines attraktiven Datenmodells vor. Es gilt jedoch die Daten auf das in MyCoRe verwendete XML-Metadaten-Gerüst abzubilden. Die nachfolgende Tabelle soll dies beschreiben. Diskrepanzen gibt es leider hinsichtlich der verwendeten Sprach-Codes. Während sich MyCoRe an XML und damit ISO-639-1 und ISO-3166 orientiert, ist bei xMetaDiss ISO-639-2 vorgesehen. Hier muss eine Schnittstelle das Mapping der Sprach-Codes realisieren. Seitens der DocProta-Anwendung wird weiter am XML-Standard mit `xsl:lang` festgehalten. Für alle nicht in ISO-639-1 spezifizierten Sprachen wird der in der Spezifikation erlaubte Sonderweg in der Form `xlm:lang="x-..."` angewandt, wobei ... der dreistellige Sprachcode der ISO-639-2 Norm ist. Hieraus kann der dreistellige Code jederzeit wieder extrahiert werden. Zur Vergabe von Rollen und anderen Attributen kann ggf. das `type` Attribut genutzt werden, welches jedes MCRMeta-Datenobjekt per Default hat.

---

4 <http://deposit.ddb.de/xmetadiss.htm>



<i>xMetaDiss</i>	<i>DocProtal</i>	<i>Anmerkungen</i>
dc:title	Titles	Mit dem Attribut <b>type="main"</b> kann der Titel als Haupttitel markiert werden.dcterms
dcterms:alternative	Titles	Mit dem Attribut <b>type="alt"</b> kann der Titel als alternativer Titel markiert werden.
ddb:isTranslationOf	Titles	Wieder <b>type="main"</b> aber mit <b>xml:lang="..."</b> für eine andere Sprache.
Übersetzungen der alternativen Titel sind in xMetaDiss nicht vorgesehen!!!	Titles	Wieder <b>type="alt"</b> aber mit <b>xml:lang="..."</b> für eine andere Sprache.
dc:creator	CreatorLink	Verweis auf einen Satz des Datenmodells <b>Person</b> .
pc:person	Siehe Datenmodell <b>Person</b>	Alle für die Person folgenden Daten werden in einem extra Datenmodell gespeichert, auf welches dann referenziert wird. Das Geschlecht @gender wird dabei unter Gender gespeichert.
pc:name pc:foreName pc:surName pc:personEnteredUnderGivenName pc:prefix pc:titleOfNobility pc:academicTitle	Siehe Datenmodell <b>Person-Person/Name</b> <b>Person/Name--&gt;firstname</b> <b>Person/Name--&gt;surname</b> <b>Person/Name--&gt;surname</b> <b>Person/Name--&gt;prefix</b> <b>Person/Name--&gt;peerage</b> <b>Person/Name--&gt;academic</b>	Der Name der Person wird im MyCoRe Datentyp MCRPersonName abgelegt. Die Typen können unter <b>type="..."</b> angegeben werden. Möglich sind: <ul style="list-style-type: none"> <li>• pseudonym</li> <li>• earlier</li> <li>• latest</li> <li>• fullerName</li> <li>• realName</li> <li>• variants</li> </ul>
pc:dateOfBirth	Siehe Datenmodell <b>Person-Person/Date</b>	Da das Date-Element mehrere Datumsangaben enthalten kann, sollte dort noch <b>type="birth"</b> spezifiziert werden.
pc:placeOfBrith		<b>Das Geburtsort wird in DocPortal nicht gespeichert!</b>
pc:nativePlace		<b>Der Heimatort wird in DocPortal nicht gespeichert!</b>
pc:profession	Siehe Datenmodell <b>Person-Person/Profession</b>	
pc:workingFields		<b>Das Arbeitsgebiet wird in DocPortal nicht gespeichert!</b>
pc:countryPubl		<b>Das überwiegende Publikationsland wird in DocPortal nicht gespeichert!</b>
pc:fieldOfStudy		<b>Die Studienfächer werden in DocPortal nicht gespeichert!</b>
pc:affiliation	Siehe Datenmodell <b>Person-Person/Institution</b> als Verweis auf einen <b>Institution-Datensatz</b>	Die Daten können als Verweis auf eine Institution gespeichert werden.
cc:universityOrInstitution		siehe pc:affiliation
cc:name		siehe pc:affiliation
cc:place		siehe pc:affiliation
cc:department		siehe pc:affiliation
pc:note	Siehe Datenmodell <b>Person-Person/Note</b>	
pc:publications	Siehe Datenmodell <b>Person-Person/Publications</b>	
pc:homePage	Siehe Datenmodell <b>Person-Person/URL</b>	
pc:email	Siehe Datenmodell <b>Person-Person/eMail</b>	
pc:address	Siehe Datenmodell <b>Person-Person/Address</b>	
dc:subject	Subjects	Die Liste der Klassifikationen finden Sie unter S
dc:description	Description	Freitext mit dem Attribut <b>type="description"</b> .

<i>xMetaDiss</i>	<i>DocProtal</i>	<i>Anmerkungen</i>
dcterms:tableOfContent	Description	Freitext mit dem Attribut <b>type</b> ="content".
dcterms:abstract	Description	Freitext mit dem Attribut <b>type</b> ="abstract".
ddb:note	Notes	Eintrag in Notes mit dem <b>type</b> ="description".
dc:publisher	PublisherLink	Alle Informationen zu einer Institution werden im extra Datenmodell <b>Institution</b> verwaltet, auf welches referenziert wird. Die Rolle kann über das Attribut <b>type</b> festgelegt werden.
dc:contributor	ContributorLink	Alle Informationen zu Beteiligten werden im extra Datenmodell <b>Person</b> verwaltet, auf welches referenziert wird. Die Rolle kann über das Attribut <b>type</b> festgelegt werden. Als Rollen sind vorgesehen: <ul style="list-style-type: none"> <li>• advisor</li> <li>• referee</li> <li>• committeeMember</li> <li>• chair</li> <li>• co-chair</li> </ul>
dc:date	Date	Hier werden alle Datumsangaben gespeichert. Über das type-Attribut kann eine Selektion erfolgen. Folgende Typen sind vorgesehen: <ul style="list-style-type: none"> <li>• accept</li> <li>• submit</li> <li>• created</li> <li>• issued</li> </ul>
dc:type	Type	Klassifikationsausschnitt aus der umfassenden Type-Classification.
dc:identifier	URN NBN	Angaben der URN und NBN
dc:format	Format	F mit verbalen Bezeichnungen.
dcterms:extent	Size	Freitext
dcterms:medium	Siehe Speichermodell der Deirvate	Daten über den MIME-Type werden in MyCoRe in des Speichertabelle für die Dokeument-Objekte (Derivate) separat gehalten.
dcterms:bibliographicCitation	Citation	Freitext
dc:source	Source	Freitext
dc:source	SourceLink	URI
dc:language	Language	Klassifikation, die ein Mapping von ISO-639-2 mit XML-Syntax gestattet.
dc:relation dcterms:isVersionOf dcterms:hasVersion dcterms:isReplacesBy dcterms:replaces dcterms:isRequiredBy dcterms:requires dcterms:isPartOf dcterms:hasPart dcterms:isReferecedBy dcterms:refences dcterms:isFormatOf	Relation RelationLink RelationISBN	Freitext, URI, ISBN, es sind die folgenden Typen im Attribut <b>type</b> möglich: <ul style="list-style-type: none"> <li>• isVersionOf</li> <li>• hasVersion</li> <li>• isReplacedBy</li> <li>• replaces</li> <li>• isRequiredBy</li> <li>• requires</li> <li>• isPartOf</li> <li>• hasPart</li> <li>• isRefercedBy</li> </ul>

<i>xMetaDiss</i>	<i>DocProtal</i>	<i>Anmerkungen</i>
dcterms:hasFormat dcterms:conformsTo		<ul style="list-style-type: none"> <li>• references</li> <li>• isFormatOf</li> <li>• hasFormat</li> <li>• conforemsTo</li> </ul>
dc:coverage dcterms:spatial dcterms:temporal	Coverage CoverageDate	Freitext oder Datum
dc:rights	Rights RightsLink	Freitext oder URI
dcterms:accessRights	ServiceFlags	Die Zugriffsrechte werden als Flags abgelegt. Eine genaue Spezifizierung erfolgt im Abschnitt Sicherheit.
thesis:degree thesis:level thesis:grantor thesis:discipline thesis:name		
dbd:contact	DDBCContact	Textzeile
dbd:fileNumber ddb:fileProperties ddb:identifier ddb:rights ddb.server ddb:description ddb:dateDelivered ddb:useRestrictionInfo dcterms:modified ddb:note	Siehe Speichermodell der Deirvate	Diese Daten werden dem Object Store entnommen und direkt daraus generiert. Einige der Felder werden per Default belegt.

Tabelle 12: Ein mögliches Mapping von xMetaDiss

## 6.2 Einbindung in SWB

### 6.2.1 Vorbemerkung

Der SWB führt 2005/2006 eine neue Katalogisierungssoftware ein. Die voraussichtliche Verbundsoftware/Katalogisierungssoftware wird PICA. (Sachsen hat weiterhin als Katalogisierungssoftware Libero) Das aktuelle Datenmodell bezieht sich auf die derzeit im SWB verwendete Katalogisierungssoftware KATWIN.

Zur Klassifikation wird in der UB Leipzig die Regensburger Systematik (RVK) verwendet, diese ist im SWB (noch) nicht vorgesehen. Bei Änderungen im RVK, muss der Administrator die entsprechenden Klassifikationen an den Dokumenten ändern. Eine Erschließung nach anderen Sachgruppen (DDC,...) wird an der UB nicht durchgeführt.

Die Sachgruppenangaben der Deutschen Nationalbibliographie (wird seit 1.1.2004 von der Library of Congress bereitgestellt), werden durch die UB für Dokumente auf dem Dokumentenserver nicht geprüft. Es ist abzuklären, ob Sachgruppen für Habilitationen und Dissertationen notwendig sind, da diese der Autor falsch vergeben werden kann. Somit ist die Suchbarkeit der Diss/Hab. zu klären.

Für einen möglichen Daten-Import in den SWB stehen derzeit folgende Möglichkeiten zur Verfügung:

#### 1.4 Transfer-Medien

Zur Übermittlung der Daten an das BSZ können folgende Medien verwendet werden:

- Attachment per E-Mail
- Disketten
- bei großen Datenmengen per FTP (nach Absprache mit dem BSZ)
- Opus-Schnittstelle, ab 2005 XMetaDiss

Die Größe der Dateien muss mit dem BSZ abgesprochen werden.

BSZ-Kategorien:

<pre>--Schlagw orte-- Schlagw ort / alle Schemata Schlagw ort / SWD Schlagw ort / NLM Schlagw ort / noScheme Schlagw ort / AAT Schlagw ort / Colon Schlagw ort / Cr Schlagw ort / DDC Schlagw ort / ICONCLASS Schlagw ort / JEL Schlagw ort / LCC Schlagw ort / LCSH Schlagw ort / MeSH Schlagw ort / MSC Schlagw ort / PACS Schlagw ort / RCHME Schlagw ort / SHIC2 Schlagw ort / TGMI Schlagw ort / TGN Schlagw ort / UDC Schlagw ort / ULAN</pre>	<pre>--- alle LocalTypes --- Abstract Anleitung (Manual) Artikel  Bild Buch Datensatz, Programm Diplomarbeit, Magisterarbeit Dissertation Habilitation Journal Lernmaterial Musik (Audio) Musiknoten Organisationsinformationen Personen-Informationen Preprint Proceedings Rede, Vortrag Report Research Paper Service Sonstiges Statistische Daten Video</pre>	<pre>text image collection dataset event interactive_resource service softw are sound</pre>
--	--	---

## 6.2.2 Datenmodell

<i>Daten des SWB</i>	<i>DocPortal</i>	<i>Anmerkung</i>
200*900 (Name, Vorname)	Daten aus CreatorLink Verweis Person --> Name --> surname, firstname	Die zur Person gehörigen Daten werden aus dem Personen-Datensatz geholt.
le1 (Geburtsjahr)	Daten aus CreatorLink Verweis Person --> Date	
910 (Verweis)	Daten aus CreatorLink Verweis Person --> Name --> fullname Person -->Name-->lastname	Zusätzlich notierter Verweis-Name
988 (Beruf)	Daten aus CreatorLink Verweis Person --> ProfClass	
999 (Bemerkung / Hinweis)	Daten aus CreatorLink Verweis Person -->Name-->academic Person -->Name-->pereege Person -->Institution	Hier können weitere nicht normierte Hinweise zur Person eingetragen werden.
320 (Titel)	Titel	Titel des Dokuments ( <b>type="main"</b> )
335 (Untertitel)	Titel	Titel des Dokuments ( <b>type="subtitle"</b> )
359 (Verf. in Vorlageform)	Creator	
418 (Ort)	Place	
425 (Jahr)	Date	Das Datum der Bestätigung der Arbeit.
433 (Umfang)	Size	
501 (Fußnote)	Note	Eintrag im Feld Note mit dem <b>type="feet"</b> .
502 (Kommentar)	Note	Eintrag im Feld Note mit dem <b>type="coment"</b> .
503 (Sprache)	Language	Wird aus der Language-Klassifikation abgeleitet und konvertiert.
504 (Sprache)	Language	Nur dt., engl., franz., ital., russ.)
519 (Diss.vermerk)		Wird aus anderen Feldern zusammgebaut. z. B. Leipzig, Univ., Diss., 2004
574 (Abrufzeichen)		immer „hs“
720 (Regensvburger Klass./Systema)	Subject	

Tabelle 13: Ein Mapping für die SWB-Verbunddaten

## 6.3 Einbindung in aus dem PICA-Katalog

<i>PICA3 Katalogisierung</i>	<i>DocPortal</i>	<i>Anmerkung, Pica-Bezeichnung</i>
<b>Bibliographische Ebene</b>		
0100	-	Identifikationsnummer (PPN)
0500	Type	Bibliographische Gattung und Status
1100	Date	Erscheinungsjahr
1108	-	Allgemeine Materialbenennung z.B.: Elektronische Ressource
1110 *	-	Dokumenttyp z.B.: ep (...), ho (...)
1500	Languages	Sprachbezeichnung z.B.: /1de/4en 1 - Sprachbezeichnung des vorliegenden Textes 4 - Sprachbezeichnung des Abstrakts
1700	Coverage	Erscheinungsland
301x 302x	Contributor	Sonstige beteiligte Personen z.B.: vorname@nachname
310x 312x	Publisher	Primärkörperschaft Sekundärkörperschaft z.B.: !PPN der körperschaft!körperschaft
4000 4005 *	Titel	Hauptsachtitel, Zusätze,... Verfasserangabe Unterreihen
4020	-	Ausgabebezeichnung z.B.: [Elektronische Ressource]
4030 *	Publisher	Ort, Verlag z.B.: ort : verlag
4031 *	-	Ort, Datenbankanbieter (Host)
4060	[aus Derivatmodell]	Umfangsangabe, spezifische Materialbenennung, techn. System z.B.: Online-Ressource, Videofile (4,59 Min), ...
4061	Size	Illustrationsangabe, sonstige physische und technische Angaben
4065	-	Besitznachweis für die Verfilmungsvorlage
4083 *	Identifier	Angaben zum Zugriff auf die elektronische Ressource im Fernzugriff
4207	Description	Inhaltliche Zusammenfassung
4229	-	Hinweise auf parallele Ausgaben
4238 *	Format	Technische Angaben zum elektronischen Dokument
530x	Subject	Basisklassifikation
555x *	Subject	Einzel Schlagwort z.B.: Elektronische Publikation, ...
<b>Lokale Ebene</b>		
65xx *	Subject	Lokale Schlagwörter
<b>Exemplarebene</b>		
68xx *	Subject	Lokale Schlagwörter (Exemplar-Ebene) z.B.: eThesis. ePublication, ...
70xx	Date	Datum und Selektionsschlüssel z.B.: 16-04-04 : ...
7133 *	[URL des Objektes]	Lokale Angaben zum Zugriff auf die elektronische Ressource im Fernzugriff
(7901)	-	Abt./Bearb. und Datum der letzten Änderung

<i>PICA3 Katalogisierung</i>	<i>DocPortal</i>	<i>Anmerkung, Pica-Bezeichnung</i>
7800	Identifizier	Exemplar-Produktionsnummer

Tabelle 14: SWB Datenaufbau

Bem.: xx - fortlaufende Ziffern  
\* - Wiederholungen möglich

## Tabellenverzeichnis

Tabelle 1: Datentypen des MyCoRe-Projektes.....	11
Tabelle 2: Das neue Metadaten-Modell der Dokumente.....	13
Tabelle 3: Ausfüllhinweise zum neuen Datenmodell der Dokumente.....	16
Tabelle 4: Das neue Metadaten-Modell für Institutionen.....	17
Tabelle 5: Ausfüllhinweise für das neue Metadaten-Modell für Institutionen.....	17
Tabelle 6: Das neue Metadaten-Modell für Personen.....	18
Tabelle 7: Ausfüllhinweise für das neue Metadaten-Modell für Personen.....	18
Tabelle 8: Potentielle Klassifikationen für DocPortal.....	20
Tabelle 9: DocPortal-Klassifikationen.....	23
Tabelle 10: Benutzer des MyCoReSample.....	23
Tabelle 11: Pflichtfelder im Vergleich.....	28
Tabelle 12: Ein mögliches Mapping von xMetaDiss.....	35
Tabelle 13: Ein Mapping für die SWB-Verbunddaten.....	37
Tabelle 14: SWB Datenaufbau.....	39



## Abbildungsverzeichnis

Abbildung 1: Verbindung mehrerer MyCoRe Anwendungen.....	6
Abbildung 2: Allgemeine Komponentenübersicht.....	7
Abbildung 3: Zusammenhänge der einzelnen Metadaten.....	11
Abbildung 4: Workflow-Prozesse für einen allgemeinen Dokumenten-Server.....	24