



Programmer Guide

RELEASE 1.3

8. September 2006

Frank Lützenkirchen (Essen/Duisburg)

Jens Kupferschmidt (Leipzig)

Detlev Degenhardt (Freiburg)

Heiko Helmbrecht (München)

Thomas Scheffler (Jena)

Anja Schaar (Rostock)

Harald Richter (Essen/Duisburg)

Kathleen Krebs (Hamburg)

Vorwort

Dieser Teil der Dokumentation ist vor allem für Applikationsprogrammierer gedacht. Er beschreibt die Designkriterien und ihre Umsetzung in der vorliegenden Version 1.3. Mit Hilfe dieser Dokumentation sollte es Ihnen möglich sein, Details von MyCoRe zu verstehen und eigene Anwendungen konzipieren und implementieren zu können. Dieses Dokument wird stetig erweitert¹.

¹Die aktuellste Version ist unter www.mycore.de/content/main/documentation zu finden.

Inhaltsverzeichnis

1	Softwareentwicklung.....	1
1.1	Vorabbemerkungen.....	1
1.2	CVS-Zugang.....	1
2	Allgemeines zur Implementierung.....	2
2.1	Struktur der Gesamtanwendung.....	2
2.2	Benutzte externe Bibliotheken.....	3
2.3	Allgemeine Klassen / Exception-Modell / MCRCache.....	6
	Allgemeine Klassen.....	6
	Exception Modell.....	7
2.4	Das Datenmodell-API-Konzept allgemein.....	7
2.5	Die Session-Verwaltung.....	7
2.6	Das EventHandler-Modell.....	9
	Das EventHandler-Modell am Beispiel der Metadaten-Objekte.....	9
	Die Konfiguration des EventHandler-Managers.....	10
2.7	Das Vererbungsmodell.....	11
3	Funktionsprinzipien und Implementierungen von Kernkomponenten.....	13
3.1	Das Query-Modell von MyCoRe.....	13
	Operatoren.....	14
	Pfadangaben.....	14
	Abfragen von Objekt-Metadaten.....	16
	Das Resultat der Query.....	16
	Abfragen von Derivaten.....	16
	Abfragen von Klassifikationen.....	17
3.2	Die Benutzerverwaltung.....	17
	Die Geschäftsprozesse der MyCoRe Benutzerverwaltung.....	17
	Benutzer, Gruppen, Privilegien und Regeln.....	18
3.3	Die Backend-Stores.....	19
	Hibernate oder nativ SQL?.....	19
	Das Search-Backend JDOM-Tree.....	20
	Das Search-Backend für IBM Content Manager 8.x.....	21
	Das Search-Backend für XML:DB.....	21
3.4	Die Frontend Komponenten.....	21
	Erweiterung des Commandline-Tools.....	21
	Das Zusammenspiel der Servlets mit dem MCRCServlet.....	23
	Das Login-Servlet und MCRSession.....	27
	Generieren von Zip-Dateien.....	28
3.5	XML Funktionalität.....	28
	URI Resolver.....	28
3.6	Das MyCoRe Editor Framework.....	30
	Funktionalität.....	30
	Architektur.....	31
	Workarounds.....	32
	Beschreibung der Editor-Formular-Definition.....	33
	Syntax der Formularelemente.....	40
	Eingabevalidierung.....	50
3.7	Klassifikationsbrowser.....	55

	Der Konfigurationsblock.....	56
3.8	Klassifikationseditor	58
	Start des Klassifikationseditors.....	59
4	Module.....	60
4.1	Das SimpleWorkflow-Modul.....	60
	Allgemeines.....	60
	Komponenten und Funktionen.....	61
	Installation.....	63
	Konfiguration.....	64
4.2	Das Webservice-Modul.....	65
	Allgemeines.....	65
	Installation des Webservices.....	66
	Client für den Webservice erzeugen	66
5	Anmerkungen und Hinweise.....	67
5.1	Ergänzung der DocPortal-Beispieldaten.....	67
	Ergänzungen in einer Beispielgruppe.....	67
	Hinzufügen einer neuen Beispielgruppe.....	67
6	Anhang.....	69
6.1	Abbildungsverzeichnis.....	69
6.2	Tabellenverzeichnis.....	71

1 Softwareentwicklung

1.1 Vorabbemerkungen

In diesem Kapitel soll kurz in die Design- und Entwicklungsmechanismen des MyCoRe-Projektes eingeführt werden. Als Open-Source-Projekt legen wir Wert auf einen öffentlichen Zugang zu unseren Code-Quellen sowie auf direkten Kontakt zu interessierten Entwicklern und Nachnutzern.

1.2 CVS-Zugang

Sowohl die aktuellen Quellen des MyCoRe-Kerns wie auch einiger beispielhafter Anwendungen werden auf einem CVS-Server in Essen verwaltet. Alle Quellen stehen für einen anonymen Zugang mittels CVS-Client zur Verfügung². Darüber hinaus ist der Zugang zum CVS-Server des MyCoRe-Projekts für interessierte Entwickler mit Schreibberechtigung möglich. Dieser erfolgt nach Freischaltung eines persönlichen Accounts über `ssh`.

Nach dem Freischalten sind folgende Umgebungsvariablen zu setzen:

```
CVS_RSH=ssh
CVSROOT=:ext:mcr_username@server.mycore.de:/cvs
export CVS_RSH CVSROOT
```

Es empfiehlt sich natürlich zuerst die MyCoRe-Quellen herunterzuladen:

```
cvs -d:ext:mcr_username@server.mycore.de:/cvs checkout mycore
```

Danach können sie mit

```
cvs -d:ext:mcr_username@server.mycore.de:/cvs commit -m "Kommentar"
file
```

Daten einstellen.

Voraussetzung ist ein CVS-Login, welches Sie bei Frank Lützenkirchen³ beantragen können. Soll in ein bestehendes Projekt eine neue Datei integriert werden, so legt man sie zunächst lokal im vorgesehenen (und bereits ausgecheckten) Verzeichnis an. Dann merkt man sie mittels `cvs add <filename>` vor. Um sie dann global zu registrieren, erfolgt ein (verkürzt): `cvs commit <filename>`. Neue Unterverzeichnisse werden auf die gleiche Weise angelegt. Weitere und sehr ausführliche Informationen sind im Internet⁴ zu finden.

²siehe MyCoRe-UserGuide

³luetzenkirchen@bibl.uni-essen.de

⁴<http://www.cvshome.org/docs/>

<http://cvsbook.red-bean.com/translations/german/>

<http://panama.informatik.uni-freiburg.de/~oberdiek/documents/OpenSourceDevWithCVS.pdf>

http://www.selflinux.org/selflinux/pdf/cvs_buch_kapitel_9.pdf

2 Allgemeines zur Implementierung

2.1 Struktur der Gesamtanwendung

MyCoRe ist entgegen anderen Projekten kein monolithisches Objekt. Vielmehr wurde ein modularer Ansatz gewählt, um die Vielfalt der Einsatzgebiete sinnvoll abzudecken. Dabei wird strikt zwischen allgemein gültigen Kernaufgaben und Funktionalitäten der schlussendlichen Anwendung unterschieden. Innerhalb dieser Bereiche gibt es wieder Teile, welche unabdingbar erforderlich sind und solche, die im Bedarfsfall in die Endanwendung integriert werden können. Die Einbeziehung der Module geschieht über zusätzliche Teile, welche in die Zielanwendung zu übernehmen sind. Die nachfolgende Skizze soll das verdeutlichen.

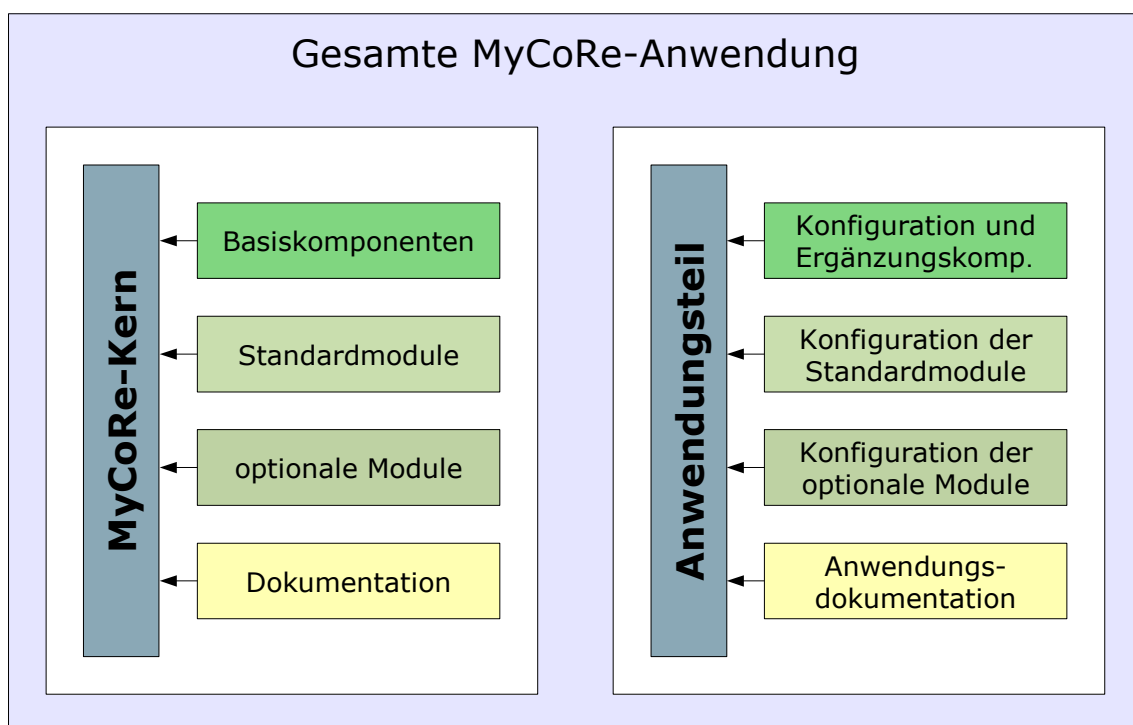


Abbildung 2.1: Grundstruktur des MyCoRe-Projektes

Das Kernsystem des MyCoRe-Projektes (Dateibaum `/mycore`) enthält die allgemein gültigen Dateien, welcher für die Gestaltung der eigenen Anwendung erforderlich sind. Dazu gehören neben den Java-Quellen auch Stylesheets, Schema-Dateien, Konfigurationen und benutzte Java-Archive von Drittanbietern sowie Dateien für den Build-Prozess. Weiterhin sind im Kern eine Reihe von Schriften enthalten, die den Umgang mit einer MyCoRe-Anwendung erklären sollen.

Eine Reihe von Komponenten sind nicht für jede Art von MyCoRe-Anwendung notwendig. Diese wurde in Module zusammengefasst und sind als solche im Verzeichnis `mycore/modules` abgelegt. Standardmodule werden von Hause aus beim Build-Prozess mit integriert. Wenn dies nicht gewünscht wird, so müssen diese Module explizit ausgeschlossen werden. Umgekehrt verhält es sich bei den optionalen Modulen (`UNINSTALLED_modulname`), welche ausdrücklich der Anwendung hinzugefügt werden müssen. Grund sind hier, wie z.B. beim `ImageViewer`,

Lizenzrechte der benutzten zusätzlichen Bibliotheken. Diese müssen vom Endanwender extern hinzugefügt werden und sind nicht im Distributionsumfang von MyCoRe enthalten. Die Integration ist im jeweiligen Modul enthalten. In den jeweiligen Verzeichnissen sind neben dem Quellcode auch Konfigurationsdateien und Dokumentation enthalten.

Die zweite Komponente einer MyCoRe-Anwendung ist die eigentliche Applikation. Sie benutzt die Bibliotheken und anderen Dateien des Kerns, ohne diese noch einmal gesondert anzupassen. Dateien, welche geändert Verwendung finden sollen, werden in der Anwendung gespeichert und überschreiben die Kerndaten. Dies gilt z. B. auch für Java-Klassen. Die Anwendung enthält weiterhin alle Konfigurationen und Ergänzungen, welche für die volle Funktion der Applikation erforderlich sind. Nach erfolgreichem Build-Prozess sind hier auch die Wurzelverzeichnisse der Anwendungsdaten platziert (es sei denn, die Konfiguration wurde geändert).

Zur Integration der Module sind auf der Seite des Anwendungsteils analoge Verzeichnisstrukturen unter `modules` erforderlich. Diese müssen aber nur eine Datei `build.xml` und alle für die spezielle Anwendung notwendigen Ergänzungen enthalten. Es wird auch hier zwischen optionalen und Standardmodulen unterschieden.

Um die Funktionsweise einer vollständigen MyCoRe-Applikation zu verdeutlichen, wird vom Projekt eine Beispielanwendung eines Dokumentenrepositories angeboten. Alle Dateien sind unterhalb des Verzeichnisses `/docportal` enthalten. Die Anwendungen kann auf die speziellen eigenen Erfordernisse angepasst oder für eine völlig neue Applikation adaptiert werden.

2.2 Benutzte externe Bibliotheken

Das MyCoRe-Projekt bemüht sich, bereits etablierte Techniken und Implementationen zu benutzen, um einerseits den eigenen Entwicklungsaufwand so gering wie möglich zu halten, andererseits dem Nachnutzer den Einstieg durch die Verwendung bekannter Komponenten zu erleichtern.

Der Folgende Abschnitt beschreibt alle eingesetzten Bibliotheken. Dabei ist die Lizenzpolitik der Hersteller sehr unterschiedlich. Gemeinsam ist allen jedoch, dass sie den Open Source Gedanken vertreten und für nichtkommerzielle Einsätze frei verfügbar sind. Die Nutzung von MyCoRe unter wirtschaftlichen Aspekten bedarf also eine gesonderten Betrachtung der Lizenzrechte durch den Anwender. Ob hier Rechte verletzt werden, ist im Einzelfall abzuklären.

Files	Beschreibung	Lizenz
lib/activation.jar	?	?
lib/ant-antlr-1.6.3.jar	ANother Tool for Language Recognition.	?
lib/ant-contrib.jar	?	?
lib/antlr-2.7.5H3.jar	ANother Tool for Language Recognition.	
lib/asm.jar	A Java bytecode manipulation framework.	ASM
lib/asm-attrs.jar	A Java bytecode manipulation framework.	ASM
lib/axis.jar	Apache Axis is an implementation of the SOAP for Web services support	AL 2.0
lib/axis-ant.jar	Apache Axis is an implementation of the SOAP for Web services support	AL 2.0
lib/BrowserLauncher2.jar	BrowserLauncher2, a continuation of the BrowserLauncher project, is a library that facilitates opening a browser from a Java application and directing the browser to a supplied url. In most cases the browser opened will be the user's default browser.	LPGL
lib/c3p0-0.9.0.jar	c3p0 is an easy-to-use library for making traditional JDBC drivers "enterprise-ready" by augmenting them with functionality defined by the jdbc3 spec and the optional extensions to jdbc2.	LPGL
lib/cglib-2.1.jar	Byte Code Generation Library is high level API to generate and transform JAVA byte code. It is used by AOP, testing, data access frameworks to generate dynamic proxy objects and intercept field access.	AL 2.0
lib/commons-collections-2.1.1.jar	Apache Jakarta Common - Extends or augments the Java Collections Framework.	AL 2.0
lib/commons-fileupload-1.0.jar	Apache Jakarta Common - File upload capability for your servlets and web applications.	AL 2.0
lib/commons-httpclient-2.0.jar	Apache Jakarta Common - Framework for working with the client-side of the HTTP protocol.	AL 2.0
lib/commons-lang-2.0.jar	Apache Jakarta Common - Provides extra functionality for classes in java.lang.	AL 2.0
lib/commons-logging.jar	Apache Jakarta Common - Wrapper around a variety of logging API implementations.	AL 2.0
lib/commons-net-1.4.1.jar	Apache Jakarta Common - Collection of network utilities and protocol implementations.	AL 2.0
lib/commons-vfs-1.0-dev.jar	Apache Jakarta Common - Virtual File System component for treating files, FTP, SMB, ZIP and such like as a single logical file system.	AL 2.0
lib/dom4j-1.6.jar	dom4j is an easy to use, open source library for working with XML, XPath and XSLT on the Java platform using the Java Collections Framework and with full support for DOM, SAX and JAXP.	W3C
lib/ehcache-1.1.jar	Ehcache is a widely used java distributed cache for general purpose caching, J2EE and light-weight containers.	AL 2.0
lib/fckEditor.zip	This HTML text editor brings to the web many of the powerful functionalities of desktop editors like MS Word. It's lightweight and doesn't require any kind of installation on the client	LPGL

Files	Beschreibung	Lizenz
	computer.	
lib/ftp.jar	Java FTP client library.	LPGL
lib/hibernate3.jar	Hibernate is a powerful, high performance object/relational persistence and query service.	LPGL
lib/hsqldb_1_0_8_0_1.jar	Lightweight 100% Java SQL Database Engine.	HS
lib/jaxen-1.1.jar	The Jaxen Java XPath Engine is an open source cross-API (DOM, JDOM, dom4j, and ElectricXML) XPath library for Java.	
lib/jaxrpc.jar	?	?
lib/jcifs-1.2.0.jar	JCIFS is an Open Source client library that implements the CIFS/SMB networking protocol in 100% Java.	LPGL
lib/jdom-1.0.jar	To provide a complete, Java-based solution for accessing, manipulating, and outputting XML data from Java code.	JDOM
lib/joda-time-1.2.jar	Joda-Time provides a quality replacement for the Java <i>date</i> and <i>time</i> classes.	AL 2.0
lib/jsch-0.1.20.jar	JSch allows you to connect to an sshd server and use port forwarding, X11 forwarding, file transfer, etc., and you can integrate its functionality into your own Java programs.	BSD
lib/jta.jar	JTA specifies standard Java interfaces between a transaction manager and the parties involved in a distributed transaction system: the resource manager, the application server, and the transactional applications.	SUN
lib/jtidy.jar	We have two primary goals. First, to provide a home where all the patches and fixes that folks contribute can be collected and incorporated into the program. Second, a library form of Tidy has been created to make it easier to incorporate Tidy into other software.	W3C
lib/junit-3.8.1.jar	JUnit is a regression testing framework written by Erich Gamma and Kent Beck.	?
lib/log4j-1.2.13.jar	Inserting log statements into your code is a low-tech method for debugging it.	AL 2.0
lib/lucene-analyzers-1.9.1.jar	Apache Lucene is a high-performance, full-featured text search engine library written entirely in Java.	AL 2.0
lib/lucene-core-1.9.1.jar	Apache Lucene is a high-performance, full-featured text search engine library written entirely in Java.	AL 2.0
lib/mail.jar	The JavaMail API provides a platform-independent and protocol-independent framework to build mail and messaging applications.	SUN
lib/poi-2.5-final-20040302.jar	The POI project consists of APIs for manipulating various file formats based upon Microsoft's OLE 2 Compound Document format using pure Java.	AL 2.0
lib/serializer_2_7_0.jar	This is a part of Xalan.	AL 2.0
lib/tm-extractors-0.4.jar	tm-extractors wraps the Word extraction from POI in a nice API.	AL 2.0

Files	Beschreibung	Lizenz
lib/xalan_2_7_0.jar	The Apache Xalan Project is a collaborative software development project dedicated to providing robust, full-featured, commercial-quality, and freely available XSLT support on a wide variety of platforms.	AL 2.0
lib/xercesImpl_2_7_1.jar	The Xerces Java Parser 1.4.4 supports the XML 1.0 recommendation and contains advanced parser functionality, such as support for the W3C's XML Schema recommendation version 1.0 , DOM Level 2 version 1.0 , and SAX Version 2 , in addition to supporting the industry-standard DOM Level 1 and SAX version 1 APIs.	AL 2.0
lib/xml-apis.jar	API-access to XML.	AL 2.0
lib/xmlldb-api-20021118.jar	API and connector to XMLDB's from eXist.	LPGL
lib/xmlldb.jar	API and connector to XMLDB's from eXist.	LPGL

Tabelle 2.1: Übersicht der benutzten externen Bibliotheken

Lizenzen:

- **AL 2.0** - Apache License Version 2.0, January 2004 - <http://www.apache.org/licenses/>
- **ASM** - Copyright (c) 2000-2005 INRIA, France Telecom - <http://asm.objectweb.org/license.html>
- **BSD** - Copyright (c) 2002,2003,2004,2005,2006 Atsuhiko Yamanaka, JCraft,Inc.
- **HS** - Copyright (c) 1995-2000 by the Hypersonic SQL Group. - <http://www.hsqldb.org/web/hsqldbLicense.html>
- **JDOM** - Copyright (C) 2000-2002 Brett McLaughlin & Jason Hunter. - <http://www.jdom.org>
- **LPGL** - GNU Library or Lesser General Public License 2.1
- **SUN** - Sun.com Terms of Use - <http://www.sun.com/termsfuse.html>
- **W3C** - W3C® SOFTWARE NOTICE AND LICENSE Copyright © 1994-2001 World - <http://www.w3.org/Consortium/Legal/copyright-software-19980720>

2.3 Allgemeine Klassen / Exception-Modell / MCRCache

Allgemeine Klassen

Wie in jedem Projekt, so gibt es auch in MyCoRe allgemein verwendet Klassen. Diese sind im Paket *org.mycore.common* und den dort befindlichen Unterpaketen platziert. Hier finden Sie Klassen für

- die Behandlung des MyCoRe-internen EventHandler-Systems,

- die Behandlung und den Umgang mit XML-Daten,
- Ausnahmebehandlungen,
- Caches,
- MyCoRe-Sessions und
- weitere nützliche Funktionen.

In den JavaDocs können Sie sich leicht einen Überblick der implementierten Klassen und Methoden verschaffen.

Exception Modell

MyCoRe nutzt zwei verschiedene Exception-Modelle, die hier kurz erläutert werden sollen. Hinsichtlich der Funktionen und Methoden der Exceptions sollten Sie die JavaDocs des MyCoRe-Projektes konsultieren.

1. MCRException und alle davon abgeleiteten Klassen. Diese Exception-Klasse ist von der Java-Klasse RuntimeException abgeleitet und kann damit Ausnahmestände nach außen tragen.
2. Die MCRCatchException ist hingegen von der Java-Klasse Exception abgeleitet und muss auch jeden Fall abgefangen werden.

2.4 Das Datenmodell-API-Konzept allgemein

2.5 Die Session-Verwaltung

Mehrere verschiedene Benutzer und Benutzerinnen (oder allgemeiner Prinzipale) können gleichzeitig Sitzungen mit dem MyCoRe-Softwaresystem eröffnen. Während einer Sitzung werden in der Regel nicht nur eine, sondern mehrere Anfragen bearbeitet. Es ist daher sinnvoll, kontextspezifische Informationen wie die UserID, die gewünschte Sprache usw. für die Dauer der Sitzung mitzuführen. Da das MyCoRe-System mit mehreren gleichzeitigen Sitzungen konfrontiert werden kann, die zudem über verschiedene Zugangswege etabliert sein können (z.B. Servlets, Kommandozeilenschnittstelle oder Webservices), muss das System einen allgemein verwendbaren Kontextwechsel ermöglichen.

Bei der Bearbeitung einer Anfrage oder Transaktion muss nicht jede einzelne Methode oder Klasse Kenntnis über die Kontextinformationen besitzen. Daher ist es sinnvoll, die Übergabe des Kontextes als Parameter von Methode zu Methode bzw. von Klasse zu Klasse zu vermeiden. Eine Möglichkeit, dies zu bewerkstelligen ist der Einsatz von sog. Thread Singletons oder thread-local Variablen. Die Idee dabei ist, den Thread der den Request bearbeitet als Repräsentation des Request selbst anzusehen. Dazu müssen die Kontextinformationen allerdings an den Thread angebunden werden, was seit Java 1.2 mit Hilfe der Klassen

`java.lang.ThreadLocal` bzw. `java.lang.InheritableThreadLocal` möglich ist. Jeder Thread hat dabei seine eigene unabhängig initialisierte Kopie der Variable. Die `set()` und `get()` Methoden der Klasse `ThreadLocal` setzen bzw. geben die Variable zurück, die zum gerade ausgeführten Thread gehört. Die Klassen der Sessionverwaltung von MyCoRe sind auf Basis dieser Technologie implementiert (siehe Abbildung 2.2).

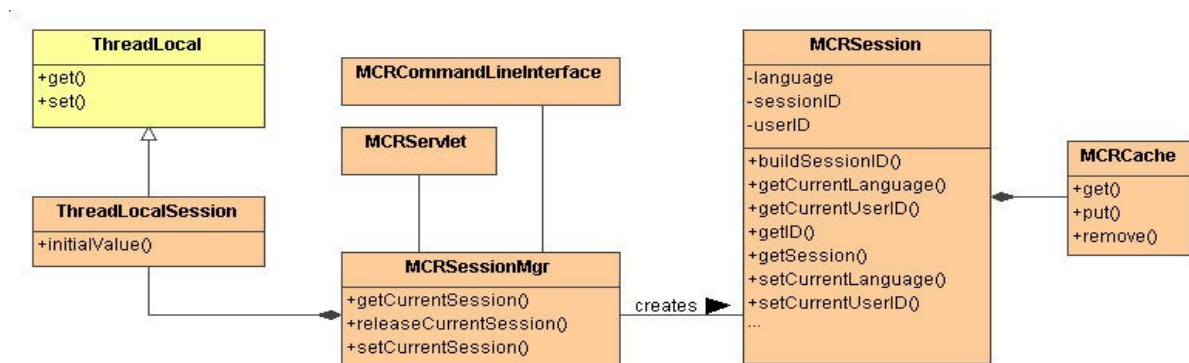


Abbildung 2.2: Die Klassen der Sessionverwaltung

Klienten der Sessionverwaltung sind alle Klassen, die Kontextinformationen lesen oder modifizieren wollen, wie zum Beispiel `MCRCommandLineInterface` und `MCRServlet`. Kontextinformationen werden als Instanzen der Klasse `MCRSession` abgelegt. Diese Klasse bietet Methoden zum Setzen und Lesen der Informationen, wie z. B. der `UserID` der aktuellen Benutzerin, der gewünschten Sprache usw.

Die Klasse `MCRSession` besitzt einen statischen Cache, realisiert durch die Klasse `MCRCache`. Bei der Konstruktion einer Instanz von `MCRSession` wird zunächst über die Methode `buildSessionID()` eine eindeutige Id erzeugt und diese als Schlüssel zusammen mit dem Session-Objekt selbst im Cache abgelegt. Auf diese Weise hat man über die statische Methode `getSession()` Zugriff auf die zu einer bestimmten `SessionID` gehörende Instanz.

Damit die Instanzen von `MCRSession` als thread-local Variablen an den aktuellen Thread angebunden werden können, werden sie nicht direkt, sondern über die statische Methode `getCurrentSession()` der Klasse `MCRSessionMgr` erzeugt und später gelesen. Beim ersten Aufruf von `getCurrentSession()` in einem Thread wird über die von `java.lang.ThreadLocal` erbenende, statische innere Klasse `ThreadLocalSession` gewährleistet, dass eine eindeutige Instanz von `MCRSession` erzeugt und als thread-local Variable abgelegt wird. Der Zugriff auf die thread-local Variablen eines Threads kann nur über die Klasse `ThreadLocal` (bzw. `InheritableThreadLocal`) erfolgen. Auf diese Weise ist sichergestellt, dass bei nachfolgenden Aufrufen von `getCurrentSession()` genau die zum aktuellen Thread gehörende Referenz auf die Instanz von `MCRSession` zurückgegeben wird.

Mit der statischen Methode `MCRSessionMgr.setCurrentSession()` ist es möglich, ein bereits vorhandenes Session-Objekt explizit als thread-local Variable an den aktuellen Thread zu binden. Dies ist z.B. In einer Servlet-Umgebung notwendig, wenn die Kontextinformationen in einem Session-Objekt über eine http-Session

mitgeführt werden. (Aktuelle Servlet-Engines verwenden in der Regel zwar Thread-Pools für die Bearbeitung der Requests, aber es ist in keiner Weise sichergestellt, dass aufeinanderfolgende Requests mit dem selben Kontext wieder den selben Thread zugewiesen bekommen. Daher muss der Kontext in einer http-Session gespeichert werden.)

Die Methode `MCRSessionMgr.releaseCurrentSession()` sorgt dafür, dass das thread-local Session-Objekt eines Threads durch ein neues, leeres Objekt ersetzt wird. Dies ist in Thread-Pool-Umgebungen wichtig, weil es sonst möglich bzw. sogar wahrscheinlich ist, dass Kontextinformationen an einem Thread angebunden sind, dieser Thread aber bei seiner Wiederverwendung in einem ganz anderen Kontext arbeitet. Code-Beispiele zur Verwendung der Session-Verwaltung finden sich in `org.mycore.frontend.servlets.MCRServlet.doGetPost()`.

2.6 Das EventHandler-Modell

Mit Version 1.2 wurde in die MyCoRe-Implementierung ein EventHandler-Basispaket integriert. Ziel ist es, eine bessere Trennung der Code-Schichten des Datenmodells und der Backends zu erreichen. Im Datenmodell sollen nur noch Ereignisse ausgelöst werden (z. B. create, delete usw.), welche dann bestimmt durch die Konfiguration in den Property-Dateien verarbeitet werden. Es soll ein allgemein gültiges Template-Modell existieren, welches für die erforderlichen Anwendungsfälle ausgebaut werden kann. Ein singleton-Manager-Prozess nimmt nur ein Ereignis entgegen, wählt die dafür bestimmte Konfiguration aus und startet die Methode `doHandleEvent(MCREvent evt)`. Dies geschieht in der Reihenfolge, welche in der Konfiguration angegeben ist und stellt ein Pipeline-Verfahren dar. Das Event-Objekt wird dabei nacheinander an die Handler durchgereicht. Änderungen an den im Event-Objekt gespeicherten Daten werden also für alle folgenden Handler wirksam. Kommt es bei einem Handler zu einer Ausnahme, so wird diese vom Manager aufgefangen und es wird für alle in der Pipeline davor liegenden Handler die Methode `undoHandleEvent(MCREvent evt)` initiiert. Somit ist ein Rollback möglich. Je nach Anwendung ist es möglich, verschiedene Pipelines für unterschiedliche Abläufe unabhängig voneinander zu implementieren, z. B. eine Pipeline für die Verarbeitung der Metadaten und eine andere für die Volltextindizierung der Dokumente. Die Pipelines und die damit verbundenen Ereignisse unterscheiden sich am Namen der jeweiligen Pipeline.

Das EventHandler-Modell am Beispiel der Metadaten-Objekte

Das EventHandler-Modell wird beispielsweise eingesetzt, um Objekte vom Typ `MCRObject` persistent zu speichern. Das Klassendiagramm (Abbildung 2.3) soll die Zusammenhänge verdeutlichen.

1. Das `MCRObject` ruft in MyCoRe-Version 1.2 zuerst eine Persistence-Layer-Implementierung nach alter Konzeption auf. Hier wurde zur Nutzung des EventHandlers eine Dummy-Klasse `MCRDummySearchStore` geschaffen, welche keine Funktionalität ausführt.

2. Anschließend wird von `MCRObject` ein neues Ereignis erzeugt, welches in diesem Fall die vordefinierte Pipeline `OBJECT_TYPE` und das vordefinierte Ereignis `CREATE_EVENT` nutzt. Es können aber auch beliebige Strings eingetragen werden. Dabei ist aber auf die Konsistenz zu achten.

```
MCREvent evt = new MCREvent(MCREvent.OBJECT_TYPE,
MCREvent.CREATE_EVENT);
```
3. Nun wird dem neuen Ereignis das Datum übergeben, welches an die Handler weitergereicht werden soll. Ein Ereignis kann auch mehrere Daten beinhalten.

```
evt.put("object",this);
```

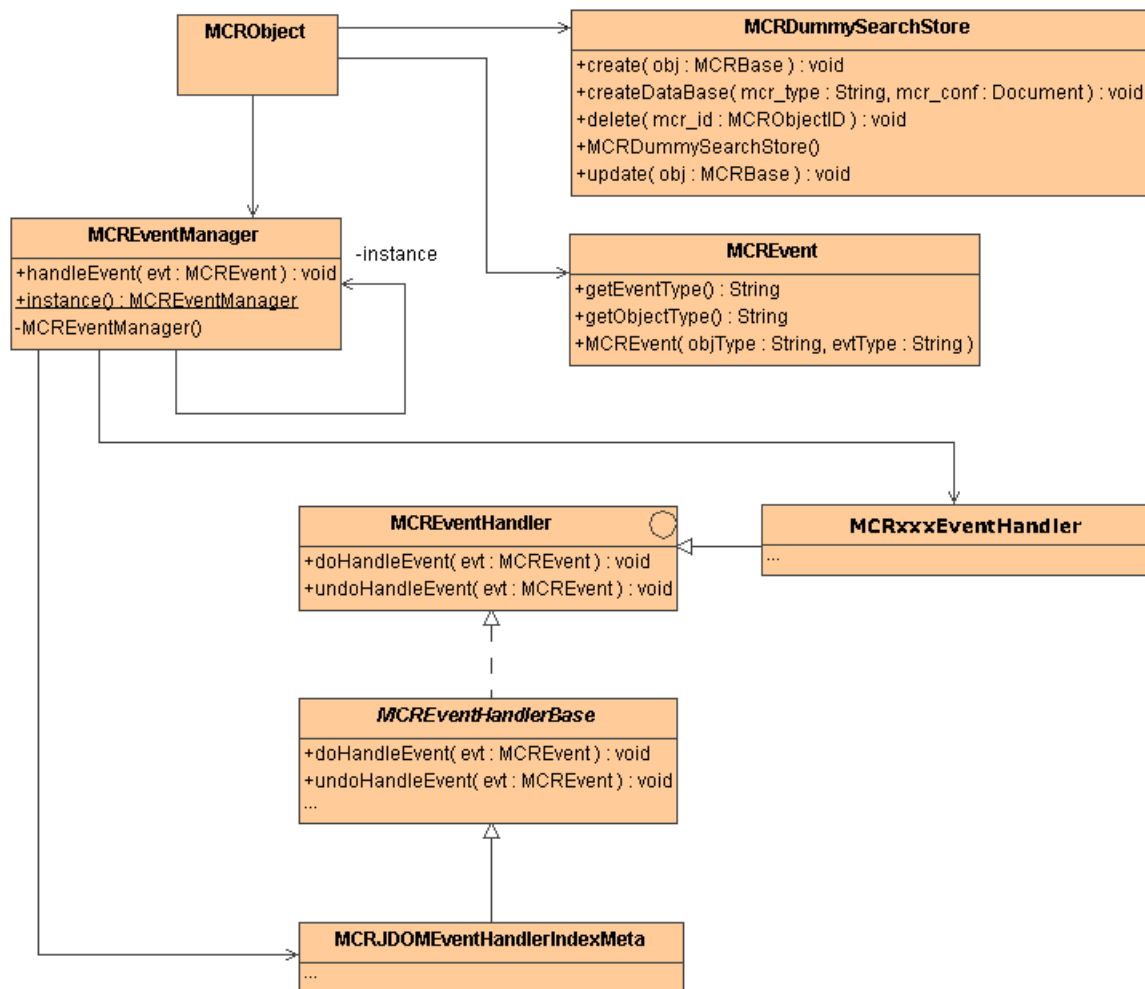


Abbildung 2.3: Klassendiagramm des EventHandler-Modells

4. Die folgende Zeile ruft abschließend den `MCREventManager` auf und stößt die Handler für die Pipeline an.

```
MCREventManager.instance().handleEvent(evt);
```

Die Konfiguration des EventHandler-Managers

Alle Konfigurationen befinden sich im Verzeichnis der Applikation (z. B. DocPortal) in der Datei `mycore.properties.private` (bzw. `mycore.properties.private.template`).

In der Version 1.2 von MyCoRe ist es noch erforderlich, für den jeweiligen SearchStore die dummy-Klasse anzugeben:

```
MCR.persistence_jdom_class_name=org.mycore.common.events.MCRDummySearchStore
```

Nun müssen noch die EventHandler für jede Pipeline (in diesem Fall ist es `MCRObject = MCREvent.OBJECT_TYPE`) in der Reihenfolge ihrer Ausführung angegeben werden. Jeder Handler bekommt dabei eine aufsteigende Nummer.

```
MCR.EventHandler.MCRObject.1.class=org.mycore.backend.jdom.MCRJDOMEventHandlerIndexMeta
```

Wollen Sie eigene EventHandler schreiben und diese einbinden, so ist es ratsam diese direkt von `MCREventHandler` abzuleiten und analog zu den bestehenden Handlern einzubinden. Sie können dafür auch frei neue Pipelines und Ereignisse definieren. Den `MCREventManager` können Sie nun an beliebiger Code-Stelle einbauen und ihm ein von Ihnen definiertes Ereignis übergeben. Diese Komponente ist allgemein verwendbar und nicht auf das MyCoRe-Datenmodell festgelegt.

2.7 Das Vererbungsmodell

In MyCoRe ist innerhalb des Datenmodells für die Metadaten die Möglichkeit einer Vererbung vorgesehen. Diese ist fest in den Kern implementiert und wird ausschließlich durch die steuernden Metadaten des jeweiligen Datensatzes festgelegt. Das heißt, für eine Datenmodell-Definition (z. B. document) können Datensätze mit (Buch mit Kapiteln) und ohne (Dokument) nebeneinander eingestellt werden. Wichtig ist nur, dass die Vererbung nur innerhalb eines Datenmodells oder eines, welches die gleiche Struktur aufweist, jedoch andere Pflichtfelder hat, funktioniert. Vererbung zwischen Datenmodellen mit verschiedenen Metadaten ist ausgeschlossen.

Im Folgenden soll die Vererbung anhand der XML-Syntax zweier Metadaten-Objekte verdeutlicht werden. Beim Laden der Daten wird dann eine Eltern-Kind-Beziehung im System aufgebaut und abgespeichert.

Die beiden Datensätze (Abbildung 2.4 und 2.5) sollen folgendes Szenario widerspiegeln:

- Der Titel soll für die Kind-Daten übernommen werden und durch diese um die Kapitelüberschriften ergänzt werden.
- Die Autorendaten sind an alle Kinder zu vererben.
- Der Umfang des Werkes ist je nach Stufe anzugeben, also für das gesamte Buch die Gesamtzahl der Seiten und für ein Kapitel die Anzahl dessen Seiten.

Entscheidend für die Umsetzung sind folgende Dinge:

```

1 <mycoreobject ... >
2   ...
3   <metadata xml:lang="de">
4     <titles class="MCRMetalangText" heritable="true" notinherit="false"..>
5       <title xml:lang="de">Buchtitel</title>
6     </titles>
7     <authors class="MCRMetalangText" heritable="true" notinherit="false"..>
8       <author xml:lang="de">Erwin der Angler</author>
9     </authors>
10    <sizes class="MCRMetalangText" heritable="false" notinherit="false"..>
11      <size xml:lang="de">100 Seiten</size>
12    </sizes>
13    ...
14  </metadata>
15  ....
16 </mycoreobject>

```

Abbildung 2.4: Auszug aus dem Metadaten-Objektes des Elternsatzes

- Das Attribut `heritable` sagt, ob ein Metadatum vererbbar (`true`) oder nicht (`false`) sein soll.
- Das Attribut `notinherit` sagt, ob das Metadatum von dem Elterndatensatz nicht geerbt werden soll (`true`). Andernfalls wird geerbt (`false`).
- Es muss erst der Elterndatensatz eingespielt werden. Anschließend können die Kinddatensätze der nächsten Vererbungsebene eingespielt werden. Enkeldatensätze folgen darauf, usw. Bitte beachten Sie das unbedingt beim Restore von Sicherungen in das System. MyCoRe ergänzt intern die Daten der Kind-Datensätze für die Suchanfragen um die geerbten Daten.
- Das Attribut `inherited` ist per Default auf 0 gesetzt und beschreibt die Anzahl der geerbten Daten. Das Attribut wird vom System automatisch gesetzt. Kinder erhalten, wenn sie Daten geerbt haben, `inherited=1` usw., je nach Stufe.

```

1 <mycoreobject ... >
2   ...
3   <metadata xml:lang="de">
4     <titles class="MCRMetalangText" heritable="true" notinherit="false"..>
5       <title xml:lang="de">Kapitel 1</title>
6     </titles>
7     <sizes class="MCRMetalangText" heritable="true" notinherit="true"..>
8       <size xml:lang="de">14 Seiten</size>
9     </sizes>
10    ...
11  </metadata>
12  ....
13 </mycoreobject>

```

Abbildung 2.5: Auszug aus dem Metadaten-Objektes des Kindsatzes

Die Ausgabe des Eltern-Datensatzes nach einer Query entspricht dem der Dateneingabe, lediglich im XML-structure-Teil wurde ein Verweis auf die Kinddaten eingetragen⁵.

⁵Siehe XML-Syntax im User Guide.


```

1  <mycoreobject ... >
2      ...
3      <metadata xml:lang="de">
4          <titles class="MCRMetaLangText" heritable="true" notinherit="false"...>
5              <title inherited="1" xml:lang="de">Buchtitel</title>
6              <title inherited="0" xml:lang="de">Kapitel 1</title>
7          </titles>
8          <authors class="MCRMetaLangText" heritable="true" notinherit="false"...>
9              <author inherited="1" xml:lang="de">Erwin der Angler</author>
10         </authors>
11         <sizes class="MCRMetaLangText" heritable="false" notinherit="false"...>
12             <size inherited="0" xml:lang="de">14 Seiten</size>
13         </sizes>
14         ...
15     </metadata>
16     ....
17 </mycoreobject>

```

Abbildung 2.6: XML-Syntax eines Kind-Datensatzes als Query-Resultat

Die Abbildung 2.6 zeigt den Kind-Datensatz, wie er vom System nach einer erfolgreichen Anfrage im Resultats-Container zurückgegeben wird. Dabei ist deutlich die Funktionalität der MyCoRe-Vererbungsmechanismen zu erkennen.

3 Funktionsprinzipien und Implementierungen von Kernkomponenten

3.1 Das Query-Modell von MyCoRe

MyCoRe bemüht sich, das Syntax-Modell der Suchanfragen an die existierenden Standards des W3C für XML Path Language (XPath) Version 1.0 (W3C Recommendation 16. November 1999) anzugleichen, da die derzeit häufigste Abfrage-Syntax im XML-Bereich ist. Dabei muss aber auch Rücksicht auf die Spezifik des MyCoRe-Systems zur Suche von Objekten und Metadaten genommen werden. Daher ist MyCoRe nicht 100% XPath/XQuery-konform, es wird jedoch angestrebt, bestmöglich die Spezifikationen einzuhalten. Grund für diese Abweichungen ist die Praxisorientierung des MyCoRe-Systems, vor allem im Bereich digitaler Bibliotheken. So sollen verschiedene Persistenzsysteme zum Einsatz kommen, welche sehr unterschiedliche Abfrage-Syntaxen, auch außerhalb der XML-Welt, implementieren. Daher bildet die in diesem Abschnitt beschriebene Syntax nur einen kleinen Teil der Möglichkeiten der W3C Spezifikationen ab. Sie genügt aber in der Praxis, um recht komplexe Projekte zu realisieren. Eine Annäherung an XPath2.0 und XQuery 1.0 wird erster folgen, wenn die Standardisierungsphase im Wesentlichen abgeschlossen ist.

MyCoRe muss neben der eigentlichen Query noch erfahren, welche Datenmodell-Typen abzufragen sind. Auch dafür ist die Ursache in der Persistenzunabhängigkeit von MyCoRe zu suchen. Es ist notwendig, die Typen auf die entsprechenden Stores zu mappen, damit die Suche erfolgreich ist. Als Beispiel soll hier der Content Manager 8 ItemType oder die Umsetzung unter eXist stehen. Möglich ist sowohl einzelne Typen wie auch eine Liste davon anzugeben. Die Liste ist ein String mit folgender Syntax:

```
Stringtype_list="type1[,...]"
```

Wichtig ist nur, dass die Elemente, nach denen gefragt wird, in allen Datenmodellen der Typen vorkommen (sonst könnte das Ergebnis der Suche eine leere Resultatliste sein). Normalerweise wird die Type-Liste in der Applikation via Konfiguration festgelegt⁶ und an die entsprechenden Query-Schnittstellen im API (MCRQueryBase) übergeben, z. B. durch das SearchMaskServlet.

Operatoren

MyCoRe verwendet nur die folgenden Operatoren innerhalb eines Test. Dies begründet sich mit der Kompatibilität und Umsetzung gegenüber den einzelnen Persistenzlayern. Wenn weitere Operatoren benötigt werden, so müssen diese in allen Persistenzimplementierungen eingebaut werden.

```
SingleQuery :: OutPath[ Tests ]
              Tests
Tests       :: Test{ AND | OR Test ... }
Test        :: InPath Operator Value
```

- Operator = ist für alle Values zulässig
- Operator != ist für alle Values zulässig
- Operator < ist nur für Datums- und Zahlenangaben zulässig
- Operator <= ist nur für Datums- und Zahlenangaben zulässig
- Operator > ist nur für Datums- und Zahlenangaben zulässig
- Operator >= ist nur für Datums- und Zahlenangaben zulässig
- Operator like versucht im Value den angegebenen String zu finden, als Wildcard ist * zulässig
- Operator contains arbeitet wie like, ist eine TextSearch-Engine verfügbar, so erfolgt die linguistische Suche darin

Die Operatoren like und contains sind eine Ergänzung von MyCoRe, um mit Textsuchmechanismen arbeiten zu können. Sie sind in der von MyCoRe benötigten Syntax-Form so nicht Bestandteil der W3C Spezifikationen, haben sich aber in der Praxis bewährt.

Pfadangaben

Die Pfadangaben für eine Single Query können einfach oder einmal geschachtelt sein, je nachdem, wie die einzelnen Text- bzw. Attributknoten des XML-Datenmodells logisch zusammengehören. Alle hier aufgeführten Möglichkeiten sind relativ zu XML-Dokument-Wurzeln. Andere Pfadangaben wie beispielsweise

⁶mycore.properties: MCR.type_alldocs=document

attribute:: für @ sind aus Gründen der Kompatibilität zu den einzelnen Persistenzlayern nicht erlaubt (und in der Praxis auch nicht erforderlich).

```

OutPath      ::      a
                a/b

InPath       ::      SpecialNodeFunction
                @d
                text()
                c
                c/text()
                c/@d
                c/d/@e
                c/d/text()

SpecialNodeFunction :: ts()
                        text()
                        doctext()

```

Hier noch einige Hinweise:

- * als SpecialNodeFunction darf allein nur in einer Single Query ohne OutPath angegeben werden. Es erfolgt dann die Suche über alle für TextSearch markierte Metadaten.
- text() als SpecialNodeFunction darf allein nur in einer Single Query ohne OutPath angegeben werden. Es wird nach 'ts()' überführt.
- doctext() als SpecialNodeFunction ist für die Abfrage des TextSearch des Dokument-Objektes vorgesehen.
- Weitere SpecialNodeFunctions können in MyCoRe ergänzt werden, wenn in der Applikation Bedarf besteht.

Nun einige gültige Beispiele:

```
text() contains "..."
```

```
*like "..."
```

```
@ID = "..."
```

```
metadata/rights/right = "..."
```

```
metadata/rights/right/text() = "..."
```

```
metadata/masse/mass[text() = "... and @type = "..."]
```

```
doctext() contains "..."
```

Abfragen von Objekt-Metadaten

Unter Objekt-Metadaten sind alle Datenmodell-Typen zu verstehen, welche NICHT `class` oder `derivate` sind⁷. Alle XML-Files der Objekt-Metadaten Typen haben als Master-Tag `/mycoreobject` und genau auf diesen Knoten als Return-Value zielen auch alle Queries unter MyCoRe. Die allgemeine Syntax ist also:

```
Query      ::      OneQuery{ AND | OR OneQuery { AND | OR ... } }
OneQuery   ::      /mycoreobject[SpecialNodeFunction]
```

Dies bedeutet, es können mehrere Abfragen hintereinander mit `AND` oder `OR` verknüpft werden. Für jedes Metadatum des Datenmodells ist eine `OneQuery` zu formulieren. Dass diese Anfragen alle auf denselben Datenraum laufen, dafür sorgt die oben beschriebene Festlegung des Type-Elementes. Somit erhalten Sie ein korrektes Gesamtergebnis.

Das Resultat der Query

Alle Antworten als Resultat der Anfrage werden in einem XML-Container zusammengefasst und der Anwendung zurückgegeben. Der Aufbau des Containers ist dabei persistenzunabhängig. Nachfolgend die XML-Struktur des Resultates einer Query:

```
1 <mcrresults parsedByLayoutServlet="...">
2   <mcrresult host="..." id="..." rank="0" hasPred="..." hasSucc="...">
3     <mycoreobject ...>
4     </mycoreobject>
5   </mcrresult>
6   ...
7 </mcrresults>
```

Abbildung 3.1: XML Syntax des Query-Resultates

- Das Attribut `parsedByLayoutServlet` ist ein Flag, welches eine etwaige Vorverarbeitung des Ergebnisses durch das `LayoutServlet` anzeigt.
- Das Attribut `host` beinhaltet entweder 'local' oder den Hostalias des Servers, von dem das Resultat stammt.
- Das Attribut `id` ist die entsprechende `MCRObjectID` des Resultates.
- Das Attribut `hasPred` gibt mit `true` an, dass dieses Resultat einen Vorgänger hat. Dies ist für die Präsentation der Einzeldaten von Interesse.
- Das Attribut `hasSucc` gibt mit `true` an, dass dieses Resultat einen Nachfolger hat. Dies ist für die Präsentation der Einzeldaten von Interesse.

Abfragen von Derivaten

Für die Derivate gelten dieselben Regeln wie für die Queries. Beachten Sie jedoch, dass das Datenmodell fest vorgeschrieben ist und das Master-Tag anders heißt. Das Resultat auf eine Anfrage wird auch in einen `mcrresults`-Container verpackt.

⁷Die Typen `class` und `derivate` sind reservierte Typen.

```
Query      :: OneQuery{ AND |OR OneQuery { AND | OR ... }}  
OneQuery   :: /mycorederivate[SpecialNodeFunction]
```

Abfragen von Klassifikationen

Klassifikationen werden in MyCoRe-Anwendungen erfahrungsgemäß am häufigsten abgefragt. Die Klassifikationen können nur nach einem sehr festen Schema abgefragt werden, wobei entweder die gesamte Klassifikation oder nur eine einzelne Kategorie zurückgegeben wird. Die XML-Struktur entspricht dabei immer der einer Klassifikation.

```
Query::      /mycoreclass[@ID="..." { and @category like "..."}]
```

3.2 Die Benutzerverwaltung

Dieser Teil der Dokumentation beschreibt Funktionalität, Design, Implementierung und Nutzung des MyCoRe Subsystems für die Benutzerverwaltung.

Die Geschäftsprozesse der MyCoRe Benutzerverwaltung

Das Benutzermanagement ist die Komponente von MyCoRe, in der die Verwaltung derjenigen Personen geregelt wird, die mit dem System umgehen (zum Beispiel als Autoren Dokumente einstellen). Zu dieser Verwaltung gehört auch die Organisation von Benutzern in Gruppen. Eine weitere Aufgabe dieser Komponente ist das Ermöglichen einer Anmelde-/Abmeldeprozedur.

Ein Use-Case-Diagramm (siehe Abbildung 3.2) soll eine Reihe typischer Geschäftsprozesse des Systems zeigen (ohne dabei den Anspruch zu haben, alle Akteure zu benennen oder alle Assoziationen der Akteure mit den Geschäftsprozessen zu definieren).

Offensichtlich dürfen nicht alle Akteure des Systems die Berechtigung haben, alle Geschäftsprozesse durchführen zu können. Daher musste ein System von Privilegien und Regeln implementiert werden: Benutzer/innen haben Privilegien (z.B. die Berechtigung, neue Benutzer/innen zu erstellen). Die Vergabe der Privilegien wird durch die Mitgliedschaft der Benutzer/innen in Gruppen geregelt. Darüber hinaus muss das System definierten Regeln gehorchen. So genügt z.B. das Privileg 'add user to group' allein nicht, um genau das Hinzufügen eines Benutzers zu einer Gruppe definieren zu können. Die Regel ist in diesem Fall, jeder andere Benutzer mit diesem Privileg aber maximal Zugehörigkeit zu Gruppen vergeben kann, in denen er oder sie selbst Mitglied ist. Auf diese Weise wird verhindert, dass sich ein Benutzer oder eine Benutzerin selbst höhere Privilegien zuweisen kann. Die Privilegien und Regeln der MyCoRe-Benutzerverwaltung werden weiter unten ausgeführt.

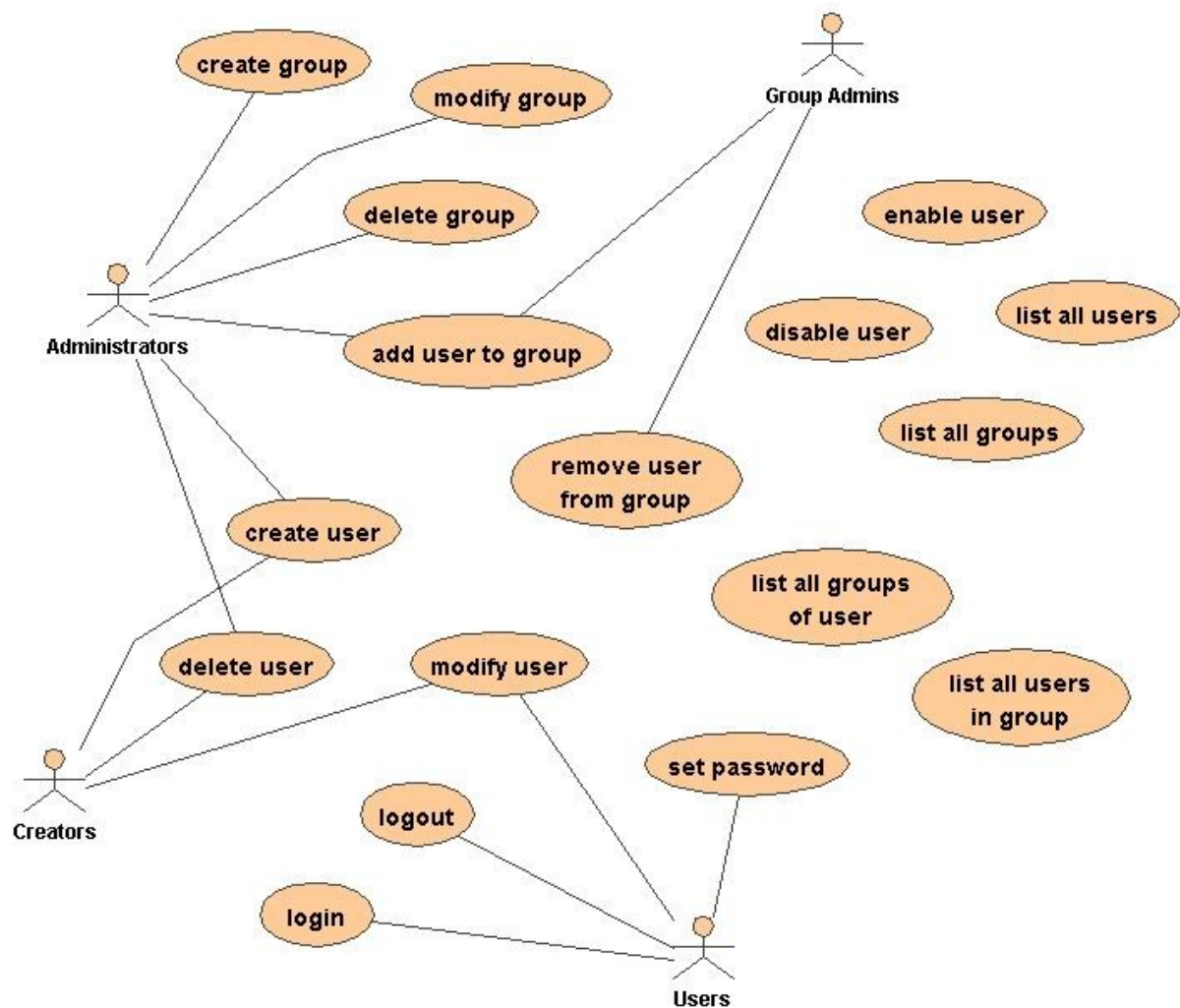


Abbildung 3.2: Geschäftsprozesse der Benutzerverwaltung in MyCoRe

Benutzer, Gruppen, Privilegien und Regeln

Die Attribute von Benutzern/innen des Systems können in drei Bereiche klassifiziert werden, den Account-Informationen wie ID, Passwort, Beschreibung usw., den Address-Informationen wie Name, Anrede, Fakultätszugehörigkeit usw. sowie den Informationen über die Mitgliedschaft zu Gruppen. Die aktuell implementierten Benutzerattribute kann man an folgender beispielhafter XML-Darstellung erkennen:

[ToDo]: to be continued ...

3.3 Die Backend-Stores

Backend	Organisation	Suche	Ablage	Volltexte	Bemerkung
HSQLDB	x				frei verfügbar; für kleine Lösungen
MySQL	x				frei verfügbar; für mittlere Lösungen
IBM DB2	x				kommerzielles Produkt; für große Lösungen
JDOM-Tree		x			in MyCoRe included; für kleine Lösungen
nativ SQL		x			benötigt ein SQL / Hibernate-Backend; je nach Backend
Lucene		x		x	frei verfügbar; für mittlere bis große Lösungen
eXist		x			frei verfügbar; für mittlere Lösungen
IBM CM 8.x		x	x	x	kommerzielles Produkt, für große Lösungen
FileSystem			x		je nach Plattenplatz
Helix			x		kommerzielles Produkt; für Audio-/Video-Lösungen

Tabelle 3.1: Übersicht der MyCoRe-Backends

Im Backend-Bereich muss zwischen den verschiedenen Aufgaben der Stores unterschieden werden. Es gib Stores für die Speicherung organisatorischer Informationen (z. B. User-Daten, XML-Daten usw.), Stores für die Suche und für die Ablage der eigentlichen digitalen Objekte und deren Volltexte. Dabei kann die genutzte Backend-Software ggf. auch für mehrere Stores verwendet werden. Eine Übersicht gibt Tabelle 3.1.

Hibernate oder nativ SQL?

MyCoRe in der Version 1.2 bietet 2 Möglichkeiten, die Stores für die organisatorischen Daten einzubinden.

Zum einen wird über das Package `org.mycore.backend.sql` ein direkter Zugriff auf relationale Datenbanken via JDBC realisiert. Der Vorteil davon sind optimale Zugriffszeiten. Nachteilig kann sich auswirken, dass nur einige der am Markt verfügbaren Datenbanken integriert und getestet haben. Es kann also bei

Verwendung anderer Datenbanken ggf. zu Problemen kommen (besonders beim automatischen Anlegen der Tabellen), da diese nicht getestet wurden.

Der zweite Weg ist die Nutzung unserer Hibernate-Integration. Hier übernimmt das freie Paket Hibernate die Anpassung an die jeweils darunter liegende Datenbank. Es wird also der gesamte Zugriff über ein fest definiertes API geregelt. Der Nachteil ist ein leichter Performance-Verlust, da ja alle Daten durch das API verwaltet werden. Die Klassen zur Arbeit mit Hibernate stehen in `org.mycore.backend.hibernate`.

Welche der beiden Zugriffsarten nun in Ihrem konkreten Projekt genommen wird hängt von den ganz spezifischen Eigenschaften der Anwendung und deren Umgebung sowie den personellen Ressourcen ab. Das MyCoRe-System wurde mit beiden Varianten getestet.

Hinweis:

Während es in der nativ-SQL-Anwendung möglich ist, für jeden XML-Daten-Store eine eigene Tabelle anzugeben, werden die XML-Daten unter Hibernate immer in einer gemeinsamen Tabelle gehalten (Property `MCR.xml_store_sql_table`).

Das Search-Backend JDOM-Tree

Das MyCoRe-Paket bietet eine simple Standard-Lösung für die Suche in kleinen Beispielanwendungen, ohne dass zusätzliche externe Produkte verwendet werden müssen. Die XML-Daten werden im JDOM-Tree zum Startzeitpunkt der Applikation direkt aus der XML-SQL-Tabelle gelesen, teilweise hinsichtlich der Umlaute normalisiert und im weiteren Verlauf der Anwendung im Hauptspeicher verwaltet. Create, Update, Delete auf die XML-SQL-Tabellen wird direkt mit dem im Speicher befindlichen Datenbaum synchronisiert.

Die XPath-Suchanfrage wird in ein XSL-Stylesheet umgewandelt, welches gegen die XML-Daten im Hauptspeicher läuft. Als Ergebnis wird die Liste der Treffer-IDs zurückgegeben. Für die Datumssuche wurde eine zusätzliche XSL-Funktion implementiert. Suchdaten im Operator `contains` werden hinsichtlich der Umlaute normalisiert (z. B. `contains("Eindrücke")` → `contains("eindruecke")`).

Metadaten-Typ	Felder
MCRMetaLangText	<code>tag/subtag/text()</code>
MCRMetaPersonName	<code>tag/subtag/firstname/text()</code> <code>tag/subtag/callname/text()</code> <code>tag/subtag/surname/text()</code> <code>tag/subtag/fullname/text()</code>
MCRMetaInstitutionName	<code>tag/subtag/fullname/text()</code>

Tabelle 3.2: Felder mit Umlautnormalisierung im Search-Store

Hinweis:

Dieser Search-Store ist nur bis zu einigen 100 Datensätzen performant und nicht für Produktionssysteme gedacht.

Das Search-Backend für IBM Content Manager 8.x

Hinweis:

Strings, welche NUR Zahlen enthalten, werden als Zahl interpretiert. Läuft die Anfrage gegen ein als `VarChar` definiertes Attribut, kommt es zum Fehler durch den Content Manager (CM). Ergänzen Sie den Zahlen-String in der Suche z.B. mit einem `*` und benutzen Sie den Like-Operator. Beispiel `like 0004*`.

Das Search-Backend für XML:DB

Auf dem Markt gibt es eine Reihe von XML-DBs. Im Bereich der kommerziellen Systeme wurde für MyCoRe einmal prototypisch Tamino getestet. Aus Kostengründen wurde jedoch auf eine Produktionsumgebung verzichtet. Im Bereich freier Software kommt bei MyCoRe das Produkt eXist ins Spiel. Für Projekte, die mittlere Leistungsanforderungen haben (bis ca. 8000 Datensätze) ist eXist auch im Produktionsbetrieb im Einsatz. Bei größeren Datenmengen kann es derzeit zu Performance-Problemen im Bereich Update kommen.

eXist ist extern zu beziehen und zu installieren. Anschließend wird es über die Konfiguration eingebunden (siehe UserGuide). Analog zum JDOM-Backend wird auch für das XML:DB Backend eine Umlautnormalisierung für die durchsuchbaren Daten und die Queries vorgenommen. Als Ergebnis der Suche wird die Liste der Treffer-IDs zurückgegeben.

Hinweis:

eXist ignoriert derzeit einige Schachtelungskonstrukte einer Query mit '[...]'. Daher kann das Ergebnis der Anfrage mehr Treffer haben, als bei korrekter Ausführung der Query.

3.4 Die Frontend Komponenten

Erweiterung des Commandline-Tools

Dieser Abschnitt beschäftigt sich mit der Struktur des Commandline-Tools und dessen Erweiterung mit eigenen Kommandos. Dem Leser sei vorab empfohlen, den entsprechenden Abschnitt im MyCoRe-UserGuide durchzuarbeiten.

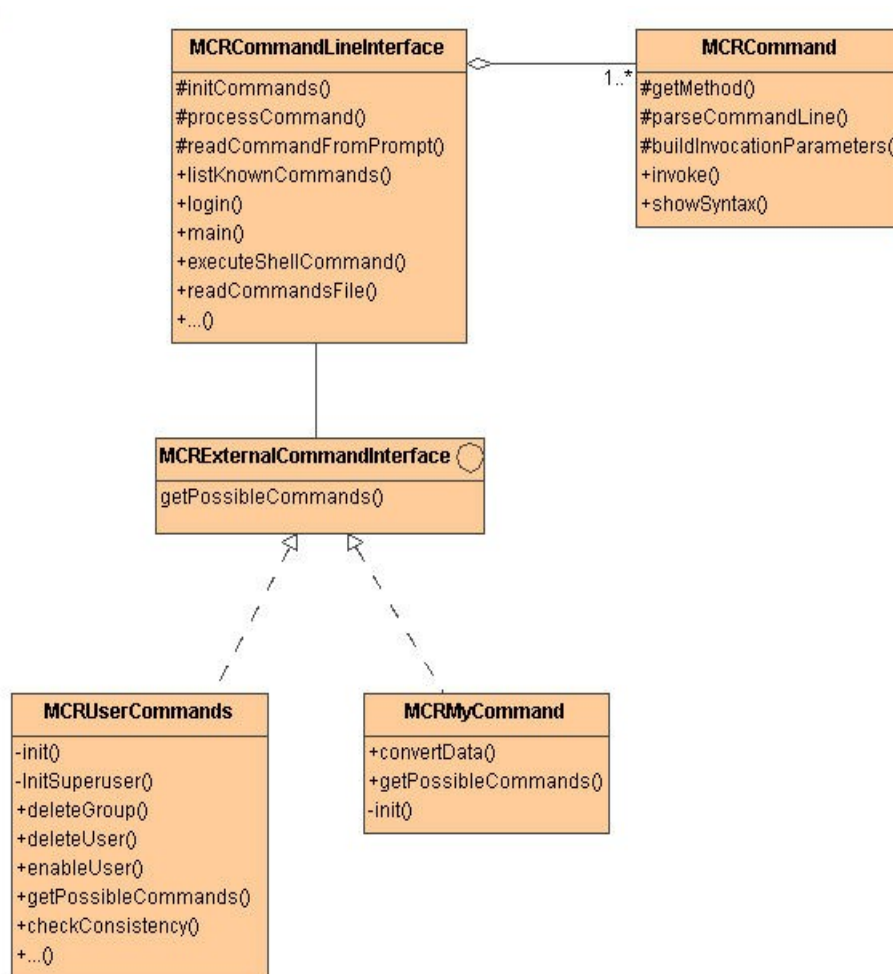


Abbildung 3.3: Zusammenhang der Java-Klassen

Das Commandline-Tool ist die Schnittstelle für eine interaktive Arbeit mit dem MyCoRe-System auf Kommandozeilen-Basis. Sie können dieses System ebenfalls dazu verwenden, mittels Skript-Jobs ganze Arbeitsabläufe zu automatisieren. Dies ist besonders bei der Massendatenverarbeitung sehr hilfreich. In DocPortal werden Ihnen schon in den Verzeichnissen `unixtools` bzw. `dostools` eine ganze Reihe von hilfreichen Skripten für Unix bzw. MS Windows mitgegeben.

All diese Skripte basieren auf dem Shell-Skript `bin/mycore.sh` bzw. `bin/mycore.cmd`, welches im Initialisierungsprozess der Anwendung via `ant` mit gebaut wird (`ant create.unixtools` bzw. `ant create.dostools`). Sollten Sie zu einem späteren Zeitpunkt eventuell einmal `*.jar`-Dateien in den `lib`-Verzeichnissen ausgetauscht haben oder sonstige Änderungen hinsichtlich des Java-CLASSPATH durchgeführt haben, so führen Sie für ein Rebuild des MyCoRe-Kommandos ein `ant scripts` durch.

Die Abbildung 3.3 soll einen Überblick über die Zusammenhänge der einzelnen Java-Klassen im Zusammenhang mit der nutzerseitigen Erweiterung des Commandline-Tools geben.

Es ist relativ einfach, weitere Kommandos hinzuzufügen. In DocPortal sind bereits alle nötigen Muster vorhanden.

1. Im Verzeichnis `~/docportal/sources/org/mycore/frontend/cli` finden Sie eine Java-Klasse `MCRMyCommand.java`. Diese ist ein Muster, kopieren Sie sie in eine Java-Klasse z. B. `MCRTestCommand.java`. Die Klasse kann im Package `org.mycore.frontend.cli` liegen, sie können Sie aber auch in den Bereich tun, zu dem es logisch gehört.
2. Ersetzen Sie alle `MCRMyCommand-String` durch `MCRTestCommand`.
3. Im Konstruktor werden nun alle neuen Kommandos definiert. Hierzu werden der `ArrayList command` jeweils zwei weitere Zeilen hinzugefügt. Die erste enthält den Text-String für das Kommando. Stellen, wo Parameter eingefügt werden sollen, sind mit `{...}` zu markieren, wobei ... eine fortlaufende Nummer beginnend mit 0 ist.⁸ In der zweiten Zeile ist nun der Methodenaufruf anzugeben. Für jeden Parameter ist das Schlüsselwort `String` anzugeben.
4. Nun muss das eigentliche Kommando als Methode dieser Kommando-Klasse implementiert werden. Orientieren Sie sich dabei am mitgelieferten Beispiel.⁹
5. Compilieren Sie nun die neue Klasse mit `cd ~/docportal; ant jar`
6. Als letztes müssen Sie die Klasse in das System einbinden. Die mit dem MyCoRe-Kern mitgelieferten Kommandos sind bis auf die Basis-Kommandos über die Property-Variable `MCR.internal_command_classes` in `~/docportal/mycore.properties` dem System bekannt gemacht. Für externe Kommandos steht hierfür in der Konfigurationsdatei `mycore.properties.application` die Variable `MCR.external_command_classes` zur Verfügung. Hier können Sie eine mit Komma getrennte Liste Ihrer eigenen Kommando-Java-Klassen angeben.
7. Wenn Sie nun `mycore.sh` bzw. `mycore.cmd` starten und DEBUG für den Logger eingeschaltet haben, so sehen Sie Ihre neu integrierten Kommandos.

Das Zusammenspiel der Servlets mit dem MCRServlet

Als übergeordnetes Servlet mit einigen grundlegenden Funktionalitäten dient die Klasse `MCRServlet`. Die Hauptaufgabe von `MCRServlet` ist dabei die Herstellung der Verbindung zur Sessionverwaltung (siehe Die Session-Verwaltung). Das Zusammenspiel der relevanten Klassen ist im Klassendiagramm (Abbildung 3.4) verdeutlicht.

⁸Siehe Beispielcode

⁹Die Methode `convertData` sollten Sie in Ihrer Klasse löschen. Ebenso die Definition in der `commands-ArrayList`.

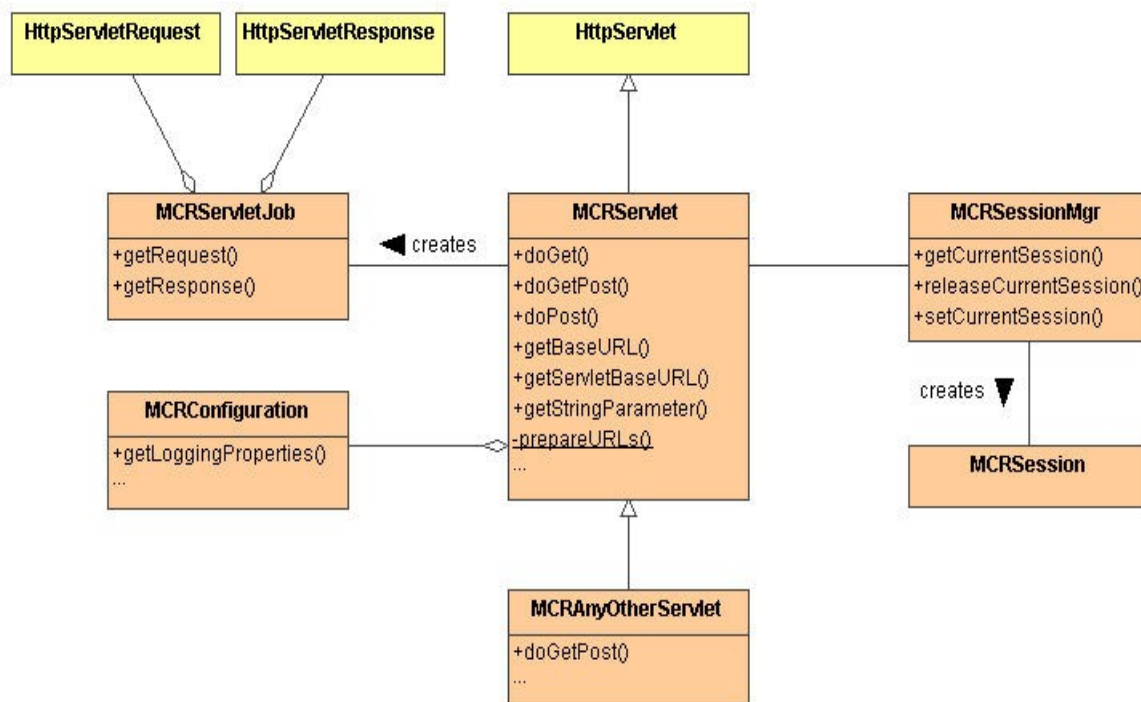


Abbildung 3.4: Klassendiagramm Common Servlets

Wie an anderen Stellen im MyCoRe-System auch, kann auf Konfigurationsparameter wie zum Beispiel den Einstellungen für das Logging über das statische Attribut `MCRConfiguration` zugegriffen werden. Dies wird ausführlich in einem anderen Kapitel beschrieben.

`MCRServlet` selbst ist direkt von `HttpServlet` abgeleitet. Sollen andere Servlets im MyCoRe-Softwaresystem die von `MCRServlet` angebotenen Funktionen automatisch nutzen, so müssen sie von `MCRServlet` abgeleitet werden. Im Klassendiagramm ist das durch die stellvertretende Klasse `MCRAnyOtherServlet` angedeutet. Es wird empfohlen, dass die abgeleiteten Servlets die Methoden `doGet()` und `doPost()` nicht überschreiben, denn dadurch werden bei einem eingehenden Request auf jeden Fall die Methoden von `MCRServlet` ausgeführt.

Der Programmablauf innerhalb von `MCRServlet` ist im folgenden Sequenzdiagramm (siehe Abbildung 3.5) dargestellt. Bei einem eingehenden Request (`doGet()` oder `doPost()`) wird zunächst an `MCRServlet.doGetPost()` delegiert.¹⁰

¹⁰Bei dieser Delegation wird ein Parameter mitgeführt, über den feststellbar ist, ob es sich um einen GET- oder POST-Request gehandelt hat.

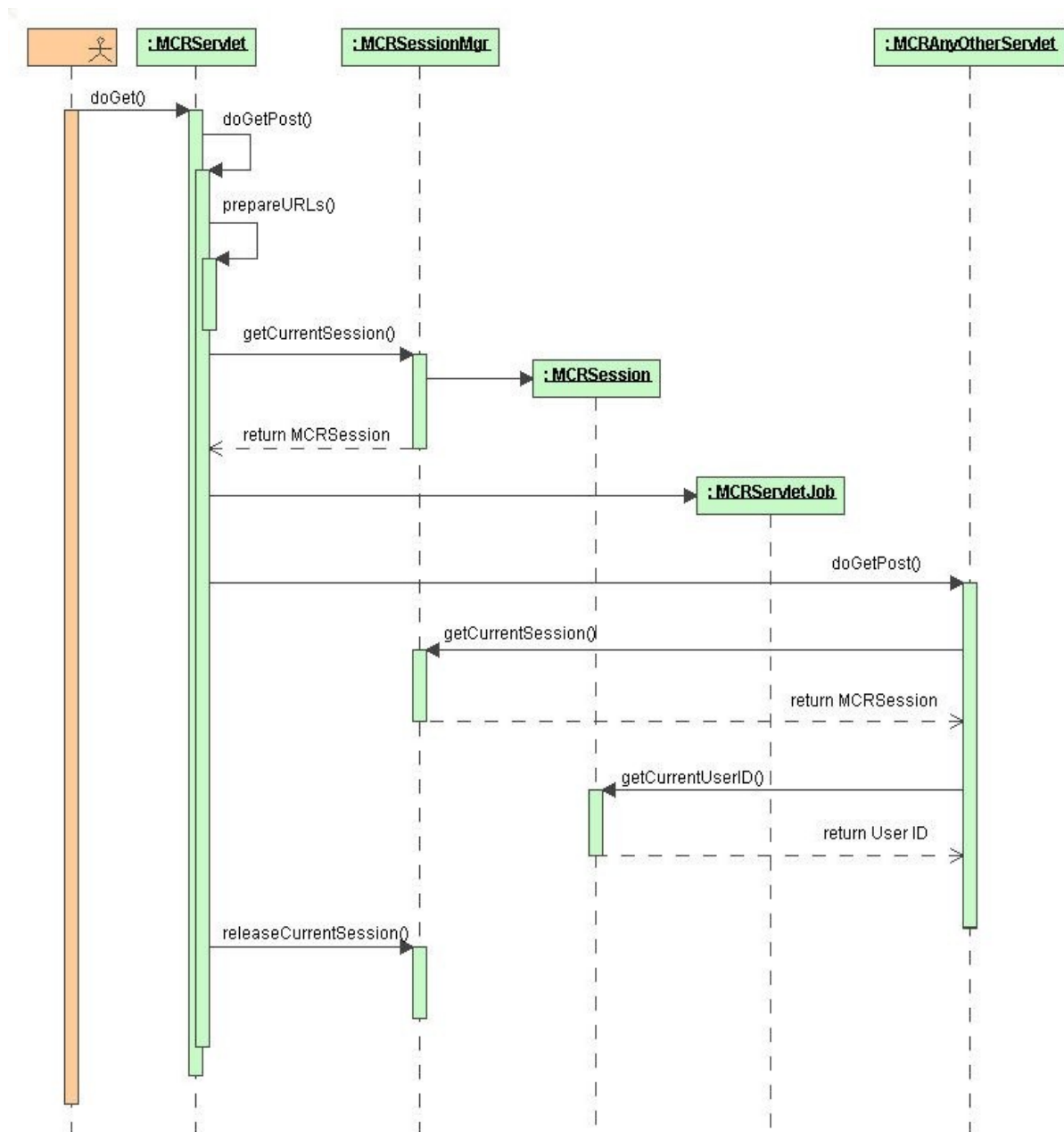


Abbildung 3.5: Sequenzdiagramm Common Servlets

Falls nicht schon aus vorhergehenden Anfragen an das `MCRServlet` bekannt, werden in `doGetPost()` die Base-URL und die Servlet-URL des Systems bestimmt. Dabei besteht die Servlet-URL aus der Base-URL und dem angehängten String 'servlets/'. Darauf folgend wird die für diese Session zugehörige Instanz von `MCRSession` bestimmt. Das Verfahren dazu ist im Ablaufdiagramm (Abbildung 3.6) dargestellt.

Die Session kann bereits durch vorhergehende Anfragen existieren. Falls dies der Fall ist, kann das zugehörige Session-Objekt entweder über eine im `HttpServletRequest` mitgeführte `SessionID` identifiziert oder direkt der `HttpSession` entnommen werden. Existiert noch keine Session, so wird ein neues Session-Objekt über den Aufruf von `MCRSessionMgr.getSession()` erzeugt. Nachfolgend wird das Session-Objekt an den aktuellen Thread gebunden und zusätzlich in der `HttpSession` abgelegt.

Im Sequenzdiagramm gehen wir davon aus, dass die Sitzung neu ist und deswegen ein Session-Objekt über `MCRSessionMgr.getCurrentSession()` erzeugt werden muss. Schließlich wird eine Instanz von `MCRServletJob` erzeugt. Diese Klasse ist nichts weiter als ein Container für die aktuellen `HttpServletRequest` und `HttpServletResponse` Objekte und hat keine weitere Funktionalität (siehe Klassendiagramm, Abbildung 3.4).¹¹

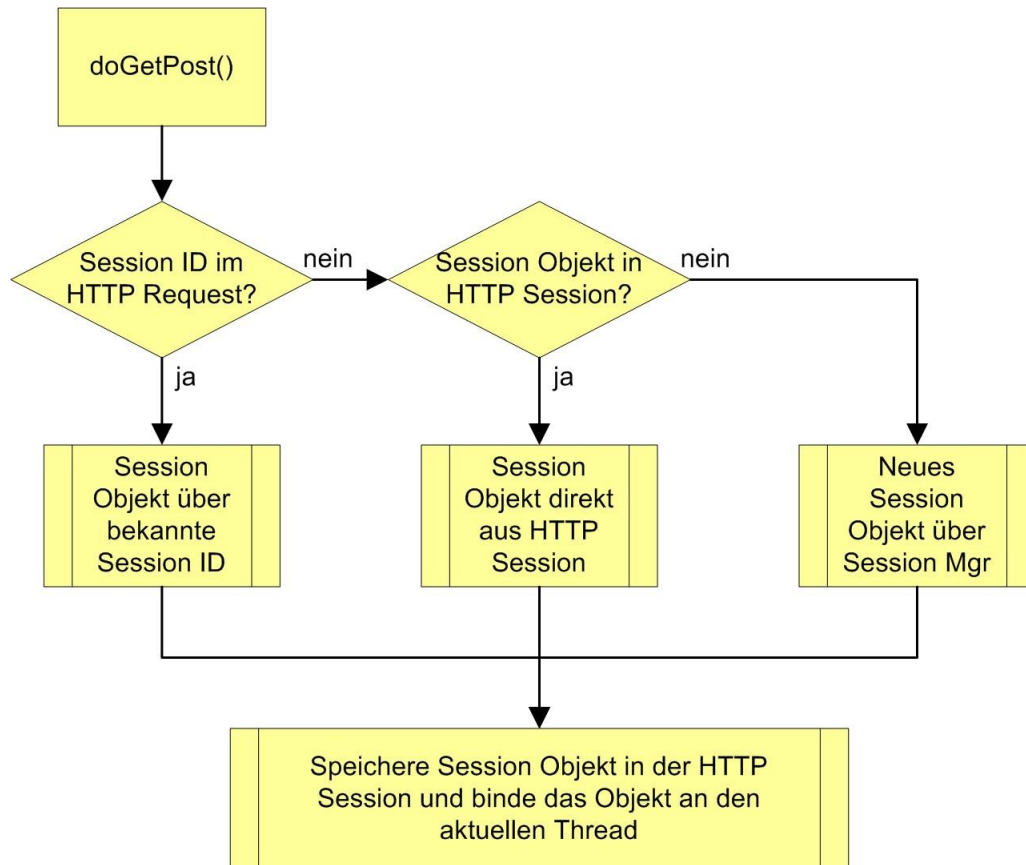


Abbildung 3.6: Ablaufdiagramm für `MCRServlet.doGetPost()`

An dieser Stelle wird der Programmfluss an das abgeleitete Servlet (in diesem Beispiel `MCRAnyOtherServlet`) delegiert. Dazu muss das Servlet eine Methode mit der Signatur

```
public void doGetPost(MCRServletJob job) {}
```

implementieren. Wie das Sequenzdiagramm beispielhaft zeigt, kann `MCRAnyOtherServlet` danach gegebenenfalls auf das Session-Objekt und damit auf die Kontextinformationen zugreifen. Der Aufruf an den SessionManager dazu wäre:

```
MCRSession mcrSession=MCRSessionMgr.getCurrentSession();
```

Es sei bemerkt, dass dies nicht notwendigerweise genau so durchgeführt werden muss. Da wegen der geschilderten Probleme mit `threadlocal` Variablen in Servlet-Umgebungen das Session-Objekt auch in der `HttpSession` abgelegt sein muss,

¹¹Das Speichern des Session-Objekts in der `HttpSession` ist notwendig, weil in einer typischen Servlet-Engine mit Thread-Pool Umgebung nicht davon ausgegangen werden darf, dass bei aufeinander folgenden Anfragen aus demselben Kontext auch derselbe Thread zugewiesen wird.

könnte man die Kontextinformationen auch aus der übergebenen Instanz von `MCRServletJob` gewinnen.

Das Login-Servlet und MCRSession

Das `LoginServlet`, implementiert durch die Klasse `MCRLoginServlet`, dient zum Anmelden von Benutzern und Benutzerinnen über ein Web-Formular. Die Funktionsweise ist wie folgt: Wie in Abschnitt 3.7.2 empfohlen, überschreibt `MCRLoginServlet` nicht die von `MCRServlet` geerbten Standard-Methoden `doGet()` und `doPost()`. Meldet sich ein Benutzer oder eine Benutzerin über das `MCRLoginServlet` an, so wird daher zunächst die Funktionalität von `MCRServlet` ausgenutzt und die in Abschnitt 3.7.2 beschriebene Verbindung zur Sessionverwaltung hergestellt. Wie dort ebenfalls beschrieben, wird der Programmfluss an das Login-Servlet über die Methode `MCRLoginServlet.doGetPost()` delegiert.

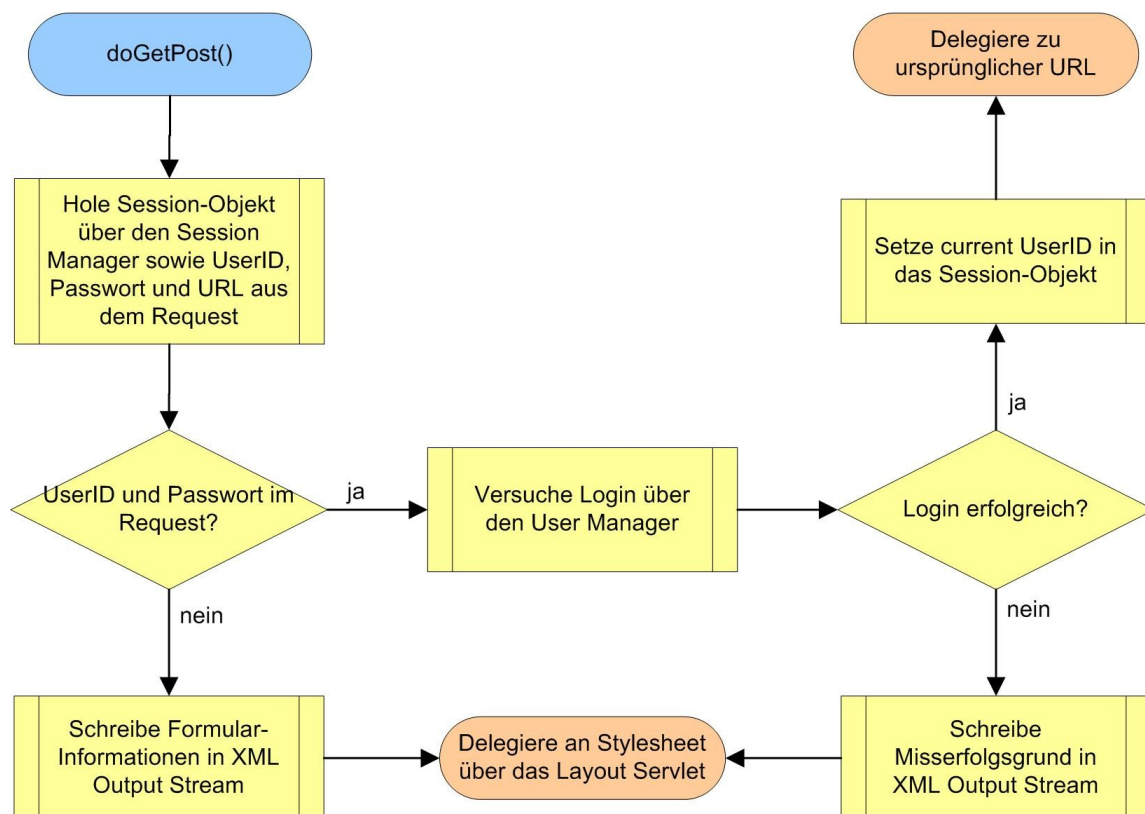


Abbildung 3.7: Ablaufdiagramm für `MCRLoginServlet.doGetPost()`

Der Ablauf in `doGetPost()` wird im Diagramm auf Abbildung 3.7 dargestellt und ist selbsterklärend.

Der resultierende XML Output-Stream muss vom zugehörigen Stylesheet verarbeitet werden und hat die in Abbildung 3.8 gezeigte Syntax.

Bei einer missglückten Anmeldung wird der Grund dafür in Form eines Attributes auf `true` oder `false` gesetzt. Das Stylesheet kann dann die entsprechende Meldung ausgeben. Die Gast-User-ID und das Gast-Passwort werden aus einer

```

1 <mcruser unknowuser="true||false" userdisabled="true||false" invalidpassword="true||false">
2   <guestid>...</guestid>
3   <guestpwd>...</guestpwd>
4   <url>...</url>
5 </mcruser>

```

Abbildung 3.8: XML Output des LoginServlets

Konfigurationsdatei gelesen. Die URL schließlich wird dem Http-Request entnommen und sollte dort von der aufrufenden Seite bzw. vom aufrufenden Servlet gesetzt sein. Ist sie nicht gesetzt, so wird die Base-URL des MyCoRe-Systems verwendet.

Generieren von Zip-Dateien

Das Zip-Servlet, implementiert durch die Klasse `MCRZipServlet`, dient dem Ausliefern der Derivate und der Objektmetadaten als gepackte Zip-Datei. Aus der Konfigurationsdatei `mycore.properties.zipper` holt sich das Servlet über die Variable `MCR.zip.metadata.transformer` den Namen des Stylesheets, welches die Metadaten transformation in das gewünschte Auslieferungsformat vornimmt. In DocPortal verwenden wir hierfür Qualified Dublin Core.

Aufrufmöglichkeiten des Servlets:

```
$ServletsBaseURL/MCRZipServlet?id=MCRID
```

```
$ServletsBaseURL/MCRZipServlet?id=MCRID/foldername
```

MCRID ist die ID eines Objekts vom Typ `<mycoreobject>` oder `<mycorederivate>`. Im Fall von `<mycoreobject>` werden die Dateien aller dem Objekt zugeordneten Derivate und ein XML-File mit den Metadaten des Objekts zusammengepackt. Im Fall von `<mycorederivate>` werden alle Dateien des angegebenen Derivats zusammengepackt. Die Option `MCRID/foldername` ist nur zulässig, wenn MCRID ein Objekt vom Typ `<mycorederivate>` bezeichnet. Dann wird nur der mit `foldername` angegebene Ordner des betreffenden Derivats gezippt.

Wer geschützte Inhalte anbietet, sollte das Zip-Servlet erst dann in seine Anwendung integrieren, wenn die Zugriffskontrolle in MyCoRe gewährleistet werden kann. Dies ist momentan (04.2005) noch nicht der Fall, das Zip-Servlet lässt sich mit jeder MCRID aufrufen.

3.5 XML Funktionalität

URI Resolver

Die Klasse `org.mycore.common.xml.MCRURIResolver` implementiert einen Resolver, mit dem an verschiedenen Stellen im MyCoRe-System XML-Daten über URI's gelesen werden können. Der Resolver wird zur Zeit an folgenden Stellen eingesetzt:

- Bei der Verarbeitung von Stylesheets im LayoutServlet, wenn XML-Daten über die XSL-Funktion `document()` in ein Stylesheet nachgeladen werden oder wenn ein untergeordnetes Stylesheet mittels `xsl:include` nachgeladen wird.

- Beim Import von Editor-Definitionsteilen mittels des `include`-Elementes des Editor-Frameworks.

Der Resolver unterstützt die folgenden Schemata bzw. Protokolle:

```
file://[Pfad]
```

liest eine statische XML-Datei vom Dateisystem des Servers

Beispiel: `file:///usr/local/tomcat/conf/server.xml`
liest die Datei `/usr/local/tomcat/conf/server.xml`

```
webapp:[Pfad]
```

liest eine statische XML-Datei vom Dateisystem der Web-Applikation. Im Gegensatz zur `file()`-Methode kann der Pfad der zu lesenden Datei relativ zum Wurzelverzeichnis der Web-Applikation angegeben werden. Der Zugriff erfolgt direkt, d.h. ohne HTTP Request oder Anwendung eines Stylesheets.

Beispiel: `webapp:config/labels.xml`

```
http://[URL]  
https://[URL]
```

liest eine XML-Datei von einem lokalen oder entfernten Webserver

```
request:[Pfad]
```

liest eine XML-Datei durch einen HTTP Request an ein Servlet oder Stylesheet innerhalb der aktuellen MyCoRe-Anwendung. Im Gegensatz zur `http/https` Methode ist der Pfad relativ zur Basis-URL der Web-Applikation anzugeben, die `MCRSessionID` wird automatisch als HTTP GET Parameter ergänzt.

Beispiel: `request:servlets/MCRQueryServlet?XSL.Style=classif-to-items&type=class&hosts=local&lang=de&query=/mycoreclass[@ID='DocPortal_class_00000002']`

```
resource:[Pfad]
```

liest eine XML-Datei aus dem CLASSPATH der Web-Applikation, d.h. die Datei wird zunächst im Verzeichnis `WEB-INF/classes/` und als nächstes in einer der `jar`-Dateien im Verzeichnis `WEB-INF/lib/` der Web-Applikation gesucht. Diese Methode bietet sich an, um statische XML-Dateien zu lesen, die in einer `jar`-Datei abgelegt sind.

Beispiel: `resource:ContentStoreSelectionRules.xml`

```
session:[Key]
```

liest ein XML-Element, das als JDOM-Element in der aktuellen `MCRSession` abgelegt ist. Mittels der `put()` Methode der Klasse `MCRSession` kann analog zu einer Java-Hashtable unter einem Schlüssel ein Objekt abgelegt werden. Ein Servlet kann so z. B. ein JDOM-Element in der `MCRSession` ablegen, den Schlüssel einem Stylesheet

über einen XSL-Parameter mitteilen. Der MyCoRe Editor kann dieses JDOM-Element dann mittels der `get()` Methode aus der Session lesen.

Beispiel: `session:mylist`
liest das JDOM XML-Element, das als Ergebnis von
`MCRSessionMgr.getCurrentSession().get("mylist");`
zurückgegeben wird.

```
mrobject:[MCRObjectID]
```

liest die XML-Darstellung der Metadaten eines MCRObject.

Beispiel: `mrobject:DocPortal_document_07910401`

```
query:[Query]
```

liest ein XML-Element, das als Wurzelement eines Ergebnisses der Suchanfrage zurückgeliefert wird. `[Query]` ist dabei ähnlich dem Query-Parametern in URLs aufgebaut, d.h. dass Parameter durch ein "&" getrennt werden (maskiert als `&`) und ein Name-Wert-Paar darstellen. Der Name ist vom Wert durch ein "="-Zeichen getrennt. Namen und Werte müssen URL-encodiert vorliegen, der Zeichensatz ergibt sich aus der Property `MCR.request_charencoding`. Folgende Parameter werden ausgewertet.

- `host`: Eine kommaseparierte Liste von Host-Alias-Namen, an die die Anfrage gestellt wird. Dieser Parameter ist optional.
- `type`: Der Dokumenttyp, über den gesucht wird.
- `query`: eine gültige MyCoRe-Query (siehe Kapitel 3.1)

Bei der Verarbeitung von `include`-Anweisungen in Editor-Definitionen dürfen die folgenden URI-Schemata verwendet werden:

```
file http https request resource session webapp mrobject
```

Beim Aufruf der XSL-Funktion `document()` innerhalb eines Stylesheets können die folgenden URI-Schemata verwendet werden:

```
file http https resource session query webapp mrobject
```

3.6 Das MyCoRe Editor Framework

Funktionalität

Das Metadatenmodell einer MyCoRe Anwendung ist frei konfigurierbar. Dementsprechend benötigt ein MyCoRe System auch einen Online-Editor für diese Metadaten, der frei konfigurierbar ist. Aus dieser Anforderung heraus entstand das MyCoRe Editor Framework, das aus einem XSL Stylesheet und einer Menge von Java-Klassen besteht.

Verschiedene MyCoRe Anwendungen können über XML-Definitionsdateien nahezu beliebige Online-Eingabemasken für Metadaten gestalten. Das Framework verarbeitet diese Editor-Definitionsdateien und generiert daraus den HTML-Code der Webseite, die das Online-Formular enthält. Nach Abschicken des Formulars generiert das Framework aus den Eingaben ein dem Metadatenmodell entsprechendes XML-Dokument, das an ein beliebiges Servlet zur endgültigen Verarbeitung (z. B. zur Speicherung) weitergereicht wird. Ebenso können existierende XML-Dokumente als Eingabe in die Formularfelder des Editors dienen, so dass sich vorhandene Metadaten bearbeiten lassen. Das Framework regelt dabei die Abbildung zwischen den XML-Elementen und -Attributen und den Eingabefeldern der resultierenden HTML-Formularseite, indem es die in der Editor-Definitionsdatei hinterlegten Abbildungsregeln verarbeitet.

Inzwischen ist das Editor Framework auch in der Lage, einzelne Dateien zusammen mit den Formulareingaben in das System hochzuladen und zur Weiterverarbeitung an ein Servlet durchzureichen. Die Validierung der Eingabefelder ist strukturell vorbereitet, aber derzeit noch nicht implementiert. Prinzipiell erlaubt das Editor Framework, beliebige XML-Dokumente in HTML-Formularen zu erzeugen oder zu bearbeiten.

Architektur

Die folgende Abbildung zeigt die Architektur des MyCoRe Editor Frameworks:

[ToDo]: Bild folgt

Ein HTML Formular, das man mit Hilfe des Frameworks realisieren möchte, wird im folgenden *Editor* genannt. Jeder Editor besitzt eine eindeutige ID (z. B. `document`¹²) und eine Definitionsdatei im XML-Format (`editor-document.xml`). In dieser Definitionsdatei ist festgelegt, aus welchen Eingabefeldern in welcher optischen Anordnung der Editor besteht und wie die Abbildung zwischen den Eingabefeldern und der zugrunde liegenden XML-Darstellung der Daten aussieht.

Ein Editor ist üblicherweise in eine umgebende Webseite eingebunden, die das Formular enthält. Die umgebenden Webseiten referenzieren über die Editor ID den Editor, der an einer bestimmten Stelle der sichtbaren Webseite eingebunden werden soll. Der HTML-Code der Webseite selbst wird in einem MyCoRe System ja aus einem beliebigen XML-Dokument (z.B. `anypage.xml`¹³) mittels eines dazu passenden XSL Stylesheets (`anypage.xsl`¹⁴) generiert. Um in eine solche Webseite einen Editor integrieren zu können, muss `anypage.xml` ein XML-Element enthalten, das auf den einzubindenden Editor verweist, zusätzlich muss `anypage.xsl` das Stylesheet `editor.xsl` einbinden. Das Stylesheet verarbeitet die Referenz auf den Editor und integriert den HTML-Code des Formulars in die umgebende Webseite. So ist es möglich, Editor-Formulare in beliebige Webseiten einzubinden, unabhängig von ihrem Layout oder ihrer Struktur.

¹²In den MyCoRe-Applikationen meist eine Mischung aus dem MCRObjektID-Typ und der Bezeichnung eines Verarbeitungsschrittes (`editor-commit-document.xml`).

¹³Z. B. `editor_form_...xml` in DocPortal

¹⁴Z. B. `MyCoReWebPage.xsl` in DocPortal

```
1 <editor id="document">
2   <include uri="webapp:editor/editor-document.xml"/>
3 </editor>
```

Abbildung 3.9: Einbindung des Editors in eine Webseite

Das Stylesheet `editor.xsl` aus dem Framework verarbeitet die Definition in `editor-document.xml` und erzeugt daraus eine HTML-Tabelle mit den entsprechenden Beschriftungen und Formularfeldern für die Eingabe der Metadaten. Es lassen sich jedoch nicht nur neue Daten eingeben, auch existierende Metadatenobjekte können bearbeitet werden. Das Stylesheet kann dazu von einer beliebigen URL ein XML-Dokument einlesen, verarbeitet dieses dann entsprechend den Abbildungsregeln aus `editor-person.xsl` und füllt die Eingabefelder des Formulars mit den Inhalten der XML-Elemente und -Attribute.

Wenn der Benutzer im Browser die Eingaben in die Formularfelder getätigt hat und das Formular abschickt, werden die Eingaben durch das Servlet `MCREditorServlet` aus dem Framework verarbeitet. Das Servlet generiert aus den Eingaben ein XML-Dokument, entsprechend den Regeln aus `editor-document.xml`. Dieses XML-Dokument wird anschließend an ein beliebiges anderes Servlet weitergereicht, welches in der Formulardefinition angegeben ist und welches dann die Daten endgültig verarbeiten kann und z. B. das XML-Dokument als MyCoRe Metadaten-Objekt speichert.

`MCREditorServlet` übergibt dabei dem Ziel-Servlet (z. B. `AnyTargetServlet`) ein Objekt vom Typ `MCREditorSubmission`, das nicht nur das XML-Dokument, sondern auch eventuell gemeinsam mit den Formulardaten übertragene Dateien als `InputStream` enthält. Damit ist es möglich, auch einfache Datei-Upload-Formulare über das Editor Framework zu realisieren.

Das Ziel-Servlet erhält zusätzlich auch die ursprünglichen HTTP Request Parameter aus dem abgeschickten Formular, so dass es bei Bedarf auch direkt auf die Werte bestimmter Eingabefelder zugreifen kann und nicht zwingend unbedingt das resultierende XML-Dokument beachten muss. Dadurch ist es möglich, das Editor Framework auch mit bereits existierenden Servlets zu nutzen, die bisher keine XML-Dokumente als Eingabe verarbeiten. Das Framework ist daher auch in der Lage, die Eingaben an eine beliebige URL statt an ein Servlet zu senden, so dass das Ziel der Eingaben an sich auch ein externes CGI-Skript oder ein Servlet in einem entfernten System sein könnte. Das schafft ein hohes Maß an Flexibilität, durch das das Editor Framework auch genutzt werden kann, um z. B. Suchmasken oder Login-Formulare zu erzeugen. Die Nutzung ist nicht allein auf die Realisierung von Metadaten-Editoren beschränkt.

Workarounds

Einige Funktionen sind derzeit noch nicht implementiert, es existieren aber Workarounds dafür:

- Eingabevalidierung:
Soll: Die Editor-Definitionsdateien werden dafür Regeln enthalten, die nach

Abschicken des Formulars durch `MCREditorServlet` überprüft werden.

Workaround: Die Validierung kann im Ziel-Servlet erfolgen, entweder auf Basis der ursprünglichen Formularfelder oder auf Basis des daraus erzeugten XML-Dokumentes.

- Namespaces:

Soll: Namespaces können in der Editor Definition deklariert werden, zusammen mit der Abbildung zwischen den Eingabefeldern und der XML-Darstellung der Daten.

Workaround: Namespaces werden im Ziel-Servlet hinzugefügt, indem das generierte JDOM-Dokument entsprechend nachbearbeitet wird. Dies könnte auch ein Stylesheet tun.

Beschreibung der Editor-Formular-Definition

Die Grundstruktur

Im folgenden soll die Gestaltung eines Editor-Formulars näher beschrieben werden. Als Beispiel ist ein Formular für ein Dokument gedacht. Die Daten werden z. B. in der Datei `editor-document.xml` gespeichert.

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <!DOCTYPE editor>
3
4 <editor id="document">
5     ...
6 </editor>

```

Abbildung 3.10: Rahmen der Formular-Definition

Als nächstes definieren wir die Eingabefelder des Editors. Ein Editor besteht aus verschiedenen Komponenten, die innerhalb des Elements `components` deklariert werden. Dies können einfache Komponenten z. B. für Auswahllisten und Texteingabefelder sein, oder aber zusammengesetzte Panels, mit deren Hilfe man einfache Komponenten in einer Tabellengitterstruktur anordnen kann. Panels können wiederum andere Panels beinhalten, so dass sich komplexere Layouts erzeugen lassen. Jede Komponente besitzt eine innerhalb des Editors eindeutige ID, über die sie referenziert werden kann. Jeder Editor besitzt ein Root-Panel, die äußerste Struktur, mit der die Anordnung der Komponenten begonnen wird. Die ID dieses Root-Panels wird als Attribut im Element `components` angegeben.

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <!DOCTYPE editor>
3
4 <editor id="document">
5     <components root="document">
6         <panel id="document" lines="on">
7             ...
8         </panel>
9     </components>
10 </editor>

```

Abbildung 3.11: Definition des Root-Panels

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <!DOCTYPE editor>
3
4 <editor id="document">
5   <components root="document">
6     <panel id="document" lines="on">
7       <cell row="1" col="1">
8         <text><label xml:lang="de">Titel:</label></text>
9       </cell>
10      <cell row="1" col="2">
11        <textfield width="80" />
12      </cell>
13      <cell row="2" col="1">
14        <text><label xml:lang="de">Autor:</label></text>
15      </cell>
16      <cell row="2" col="2">
17        <textfield width="80" />
18      </cell>
19    </panel>
20  </components>
21 </editor>

```

Abbildung 3.12: Definition eines einfachen Formulars

Innerhalb eines Panels lassen sich andere Komponenten oder Panels anordnen. Solche Komponenten werden im einfachsten Fall einfach in die Gitterzellen eines Panels eingebettet. Jede Zelle entspricht einem `cell`-Element innerhalb des Panels. Jede Zelle besitzt eine `row`- und `col`-Koordinate, mit deren Hilfe die Zelleninhalte von links nach rechts in Spalten (`col`) und von oben nach unten in Zeilen (`row`) angeordnet werden. Dabei ist für Textfelder noch die Angabe der Sprache möglich.

Innerhalb der Gitterzellen füllen die Komponenten in der Regel nicht den gesamten Platz aus, daher können sie in ihren Zellen mittels des `anchor`-Attributs anhand der Himmelsrichtungen (`NORTH`, `NORTHEAST`, `EAST`, ... bis `CENTER`) angeordnet werden. Wir ordnen die Beschriftungen rechtsbündig, die Texteingabefelder linksbündig an. Außerdem wird nun in der untersten rechten Zelle rechtsbündig einen "Speichern" Button zum Absenden des Formulars ergänzt.

Im nächsten Schritt soll nun die Definition aus dem Root-Panel ausgegliedert werden. Sie verwenden dazu ein eigenes Panel. Diese Technik wird verwendet, um Panels mehrfach zu nutzen und den Quellcode des Formulars zu strukturieren. Dabei wird gezeigt, wie sich Panels untereinander aufrufen (Abbildung 3.13).

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <!DOCTYPE editor>
3
4 <editor id="document">
5   <components root="document">
6     <panel id="document" lines="on">
7       <cell row="1" col="1" anchor="EAST" ref="form">
8         <cell row="2" col="1" anchor="EAST">
9           <submitButton width="100px" >
10             <label xml:lang="de">[Speichern]</label>
11           </submitButton>
12         </cell>
13       </panel>
14
15       <panel id="form">
16         <cell row="1" col="1" anchor="EAST">
17           <text><label xml:lang="de">Titel:</label></text>
18         </cell>
19         <cell row="1" col="2" anchor="WEST">
20           <textfield width="80" />
21         </cell>
22         <cell row="2" col="1" anchor="EAST">
23           <text><label xml:lang="de">Autor:</label></text>
24         </cell>
25         <cell row="2" col="2" anchor="WEST">
26           <textfield width="80" />
27         </cell>
28       </panel>
29     </components>
30 </editor>

```

Abbildung 3.13: Auslagern von Definitionen aus dem Root-Panel

Soll das Panel `form` nun mehrfach genutzt werden, z. B. in mehreren Formularen, so ist es sinnvoll das Panel in eine gemeinsame Definitionsdatei `imports-common.xml` auszulagern. Im eigentlichen Formular wird dann nur noch auf das `form`-Panel verwiesen. Die Einbindung der ausgelagerten Panels erfolgt mit der `include`-Anweisung (Abbildung 3.14). Die einzufügende Datei (Abbildung 3.15) wird nun im System mittels `EntityResolver` gesucht und eingebunden. Auf diese Weise lassen sich elegante und wartungsarme Formulare gestalten.

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <!DOCTYPE editor>
3
4 <editor id="document">
5   <components root="document">
6     <include uri="webapp:editor/imports-common.xml" />
7     <panel id="document" lines="on">
8       <cell row="1" col="1" anchor="EAST" ref="form">
9       <cell row="2" col="1" anchor="EAST">
10         <submitButton width="100px" >
11           <label xml:lang="de">[Speichern]</label>
12         </submitButton>
13       </cell>
14     </panel>
15   </components>
16 </editor>

```

Abbildung 3.14: Auslagern von Definitionen aus dem dem Editor-Formular

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <!DOCTYPE imports>
3
4 <imports>
5     <panel id="form">
6         <cell row="1" col="1" anchor="EAST">
7             <text><label xml:lang="de">Titel:</label></text>
8         </cell>
9         <cell row="1" col="2" anchor="WEST">
10            <textfield width="80" />
11        </cell>
12        <cell row="2" col="1" anchor="EAST">
13            <text><label xml:lang="de">Autor:</label></text>
14        </cell>
15        <cell row="2" col="2" anchor="WEST">
16            <textfield width="80" />
17        </cell>
18    </panel>
19 </imports>

```

Abbildung 3.15: Die imports-Formular-Definition

Abbildung zwischen Eingabefeldern und XML-Strukturen

Als nächstes definieren wir, wie die Eingabefelder auf XML-Elemente abgebildet werden. Dies geschieht über das Attribut `var` verschiedener Elemente der Editor-Definition. Zunächst wird im `var`-Attribut des `components`-Elements der Name des Root-Elements der XML-Darstellung (`document`) festgelegt. Die Abbildung zwischen Eingabefeldern oder Panels und ihrer Zuordnung zu XML-Elementen geschieht über das `var`-Attribut in den `cell`-Elementen, welche die Eingabekomponente enthalten (Abbildung 3.16).

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <!DOCTYPE editor>
3
4 <editor id="document">
5     <components root="document" var="/document">
6         <include uri="webapp:editor/imports-common.xml" />
7         <panel id="document" lines="on">
8             <cell row="1" col="1" anchor="EAST" ref="form">
9             <cell row="2" col="1" anchor="EAST">
10                <submitButton width="100px" >
11                    <label xml:lang="de">[Speichern]</label>
12                </submitButton>
13            </cell>
14        </panel>
15    </components>
16 </editor>

```

Abbildung 3.16: Einfügen des XML-Main-Tag-Namen

Es gibt für Tags und Attribute innerhalb der Root-Komponente verschiedene Möglichkeiten der Gestaltung des Ausgabe-XML-Baumes. Die einfachste Art ist die direkte Zuordnung. Hier werden die Werte des Eingangs-XML-Files dem Feld in der Editormaske zugeordnet. Weiterhin kann dem Feld noch ein Standardwert durch das Attribut `default` mitgegeben werden. Weiterhin können Felder, welche von der Dateneingabe ausgenommen werden sollen, deren Inhalte aber an die Ausgabeseite durchgereicht werden sollen, als `hidden`-Felder mitgeführt werden. Auch hier

können Standardwerte angegeben werden, welche eingesetzt werden, wenn keine Daten für das Tag/Attribut vorhanden sind.

```

1      <cell row="1" col="2" anchor="WEST">
2          <textfield width="80" var="dc/title" default="Mein Buch" />
3      </cell>
4      <hidden var="dc/title/@lang" default="de" />

```

Abbildung 3.17: Integration von einfachen XML-Tags

Sollten Sie mehrere gleiche XML-Strukturen im Editor zur Bearbeitung anbieten wollen, so wird dies über eine Integer-Zahl in Klammer organisiert. Achtung, es werden immer nur soviel XML-Elemente übernommen und an die Ausgabe weitergereicht, wie angegeben sind. Um eine beliebige Anzahl zu präsentieren und zu bearbeiten nutzen sie die weiter unten angegebenen Möglichkeit des Repeaters. Für hidden-Felder gibt es noch die Variante, mittels des `descendants="true"` - Attributs alle Kindelemente, Attribute und Wiederholungen des Elementes an den Ausgang durchzureichen.

```

1      <cell row="1" col="2" anchor="WEST">
2          <textfield width="80" var="dc/title" default="Mein Buch" />
3      </cell>
4      <hidden var="dc/title/@lang" default="de" />
5      <cell row="2" col="2" anchor="WEST">
6          <textfield width="80" var="dc/title[2]" default="" />
7      </cell>
8      <hidden var="dc/title[2]/@lang" default="de" />
9      <hidden var="dc/source" descendants="true" />

```

Abbildung 3.18: Integration von mehrfachen XML-Tags

Integration und Aufruf des Editor Framework

Bereits zum Anfang des Kapitels wurde eine einfache Syntax zum Aufruf des Editor Framework vorgestellt. Dieser Abschnitt soll nun umfassend zu diesem Thema informieren.

```

1  <editor id="document">
2      <include uri="webapp:editor/editor-document.xml"/>
3  </editor>

```

Abbildung 3.19: Einbindung des Editors in eine Webseite

Das Editor Framework ist so konzipiert, dass es als XML-Sequenz in einer durch das `MCRLLayoutServlet` darzustellenden Webseite integriert werden kann. D.h. Sie erstellen eine XML Seite welche durch das Servlet und die zugehörigen XSLT-Stylesheets nach HTML konvertiert werden. Durch den Import (`include`) der Editor-Framework-XSL-Dateien in das Layout zum Erzeugen der HTML-Seiten ermöglichen Sie die Einbindung ihrer Editor-Formulare. In der oben gezeigten XML-Sequenz wird jedoch nur die Definitionsdatei des Formulars beschrieben: „nimm aus der Web-Applikation die Datei `editor-document.xml` aus dem Verzeichnis `editor`“. Dabei ist noch nichts über die Datenquelle bekannt. Hierfür gibt es noch einige Parameter, welche durch das Framework ausgewertet werden. Diese können sowohl direkt im Browser in der Aufrufsequenz oder über Parametereinstellungen in einem Servlet mitgegeben werden.

```

http://.../document.xml?XSL.editor.source.new=true&type=...

1 String base = getBaseURL() + myfile;
2 Properties params = new Properties();
3 params.put( "XSL.editor.source.new", "true" );
4 params.put( "XSL.editor.cancel.url", getBaseURL()+cancelpage);
5 params.put( "type",mytype );
6 params.put( "step",mystep );
7 job.getResponse().sendRedirect(job.getResponse().encodeRedirectURL(
8   (buildRedirectURL( base, params ))));

```

Abbildung 3.20: Codesequenz zum Aufruf des Editor Framework

Oder in einem Servlet als Codestück, wie Abbildung 3.20 zeigt.

Start-Parameter für das Lesen der zu bearbeitenden XML-Quelle:

Über XSL Parameter kann beim Aufruf des Editors gesteuert werden, ob bzw. aus welcher Quelle die zu bearbeitenden XML-Daten gelesen werden. Die zu bearbeitenden XML-Daten werden von einer URL gelesen. Falls die URL mit `http://` oder `https://` beginnt, wird sie als absolute URL direkt verwendet, andernfalls wird die URL als relativ zur `WebApplicationBaseURL` (context root), dem Startpunkt der `WebApplication`, interpretiert. Es gibt die folgenden vier Optionen:

1. `XSL.editor.source.new=true`
Es wird das Formular mit leeren Datenfeldern gestartet.
2. Statische URL eines XML-Dokumentes (z. B. eines Metadaten-Templates), die in der Editor-Definition als Top-Level-Element angegeben ist (d.h. auf gleicher Ebene wie die target-Deklaration), z. B.

```
<source url="templates/document.xml" />
```
3. URL, die durch ein Servlet oder Stylesheet zur Laufzeit gebildet wird, und die beim Aufruf des Editors über einen XSL Parameter übergeben wird:

```
XSL.editor.source.url=templates/document.xml
```
4. URL, die aus einer Kombination einer statischen URL und eines ID=Tokens gebildet wird, das beim Aufruf des Editors über einen XSL Parameter übergeben wird. Im folgenden Beispiel wird zur Laufzeit das Token XXX durch den Parameterwert 4711 ersetzt:

```
XSL.editor.source.id=4711
<source url="servlets/SomeServlet?id=XXX" token="XXX" />
```

Start-Parameter für die Definition des Abbrechen-Buttons:

`XSL.editor.cancel.url` / `XSL.editor.cancel.id`

Diese Parameter werden im Abschnitt zur Verwendung eines Cancel-Buttons (Abbrechen-Knopf) erläutert, sie steuern die URL, zu der bei Klick auf einen „Abbrechen“ Button zurückgekehrt wird.

Start-Parameter für die Weitergabe an das Zielservlet:

In der Regel werden die Eingaben nach der Bearbeitung mit dem Editor an ein Zielservlet gesendet (`target type="servlet"`). Man kann diesem Zielservlet auch HTTP Parameter mitschicken, die der Editor bei Abschicken des Formulars mitsendet.

So können Informationen an das Zielservlet weitergegeben werden, die nicht Teil des zu bearbeitenden XML sind, z. B. die Information, ob es sich um einen Update- oder Create-Vorgang handelt, die Objekt-ID etc. Der Editor reicht automatisch alle HTTP Request Parameter an das Zielservlet durch, deren Name nicht mit XSL beginnt.

Beispiel: `http://.../editor-document.xml?XSL.editor.source.id=4711&action=update&mode=bingoBongo`

Ausgabeziele

Das Editor-Framework ist auch hinsichtlich der Verarbeitung der entstandenen Daten flexibel konfigurierbar. Wie bereits erwähnt, wird nach dem Betätigen des Submit-Buttons das `MCREditorServlet` für eine erste Verarbeitung angestoßen. Diese reicht die Daten dann an ein weiteres konfigurierbares Ziel weiter. Welches das ist, wird mit der `target`-Zeile in der Formulardefinition festgelegt. Das Attribut `method` definiert, ob dazu HTTP GET oder POST verwendet wird, zu empfehlen ist in der Regel immer POST, was auch der Default-Wert ist und zwingend bei File Uploads gesetzt wird. Möglich sind:

- die Ausgabe als HTML-Seite zum Debugging des Ergebnisses:
`<target type="debug" method="post" format="xml" />`
- die Weiterleitung der Daten in ein nachgeschaltetes Servlet im XML-Format:
`<target type="servlet" method="post" name="MCRCheckDataServlet" format="xml" />`
- die Weiterleitung der Daten an ein nachgeschaltetes Servlet in der Form `name=value`, d. h. ohne Generierung eines XML-Dokumentes:
`<target type="servlet" method="post" name="MCRCheckDataServlet" format="name=value" />`
- die Weiterleitung an `LayoutServlet` zur Anzeige des generierten XML-Dokumentes als Test:
`<target type="display" method="post" format="xml" />`

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <!DOCTYPE editor>
3
4 <editor id="document">
5     <!--
6     <target type="debug" />
7     -->
8     <target type="servlet" name="MCRCheckDataServlet" method="post" format="xml" />
9     ...
10 </editor>

```

Abbildung 3.21: Rahmen der Formular-Definition

- die Weiterleitung an eine beliebige URL als HTTP GET oder POST Request:
`<target type="url" method="post" format="name=value" url="cgi-bin/ProcessMe.cgi" />`

- die Weiterleitung an einen anderen Editor, für den dieser Editor als „Subdialog“ fungiert. Dies wird in einem separaten Abschnitt erläutert.
`<target type="subselect" method="post" format="xml" />`

Mit der nachfolgenden Java-Code-Sequenz kann nun auf die vom MCRServlet durchgereichten XML- und -File-Daten zugegriffen werden.

```

1  /**
2   * This method overrides doGetPost of MCRServlet.<br />
3   */
4  public void doGetPost(MCRServletJob job) throws Exception
5  {
6      // read the XML data
7      MCREditorSubmission sub = (MCREditorSubmission)
8          (job.getRequest().getAttribute("MCREditorSubmission"));
9      org.jdom.Document indoc = sub.getXML();
10     List files = sub.GetFiles();
11
12     ...
13 }
```

Abbildung 3.22: Java-Code-Sequenz für den Zugriff

Neben der Übernahme der Ausgabewerte des MCRServlets als XML-Baum können diese auch in der Form 'name=value' durchgereicht werden. Hierzu müssen u. a. die var-Attribute entsprechend in eine einfache Form gebracht werden. Der Ausschnitt der Editor-Definition soll das verdeutlichen.

```

1  <?xml version="1.0" encoding="ISO-8859-1"?>
2  <!DOCTYPE editor>
3
4  <editor id="document">
5      <target type="servlet" name="MCRStartEditorServlet" method="get" format="name=value" />
6      <components root="formular">
7          <panel id="document" lines="on">
8              <cell row="1" col="1" anchor="EAST">
9                  <text><label xml:lang="de">Label:</label></text>
10             </cell>
11             <cell row="1" col="2" anchor="WEST" var="label">
12                 <textfield width="80" />
13             </cell>
14             ...
15         </panel>
16         ...
17     </components>
18 </editor>
```

Abbildung 3.23: Rahmen der Formular-Definition mit name=value

Syntax der Formularelemente

Aufbau einer Zelle

Das cell-Element stellt ein elementares Gebilde innerhalb eines Formulargitters dar. Seine Position wird durch die Attribute col, row und anchor beschrieben. col gibt dabei die relative Spaltenzahl, row die relative Zeilenzahl an. Mit anchor kann die Ausrichtung gesteuert werden. Mögliche Werte sind hier NORTHWEST, NORTH, NORTHEAST, WEST, CENTER, EAST, SOUTHWEST, SOUTH und SOUTHEAST. Mit dem Attribut

`ref` kann auf ein mehrfach verwendbares Element via ID referenziert werden. Über das Attribut `sortnr` kann die Reihenfolge der zu generierenden XML-Daten für die Ausgabe geregelt werden. Dieses Attribut bezieht sich dabei auf das gesamte Formular.

```

1 <cell var="..." col="..." row="..." anchor="..." ref="..." sortnr="..." />
2
3 oder
4
5 <cell var="..." col="..." row="..." anchor="..." sortnr="...">
6     [Formularfelddefinition]
7 </cell>

```

Abbildung 3.24: Syntax des cell-Elements

Texteingabefelder

Ein einzeliges Texteingabefeld wird mittels eines Elementes `textfield` erzeugt, ein mehrzeiliges mittels eines Elementes `textarea`. Das Attribut `width` gibt die Breite des Texteingabefeldes in Anzahl Zeichen an. Das Attribut `height` gibt für mehrzeilige Texteingabefelder die Anzahl Zeilen an. Das Attribut `maxlength` gibt bei `textfield`-Elementen die maximale Länge der Eingabe an.

Texteingabefelder können einen Vorgabewert enthalten, der wahlweise über ein Attribut oder ein Kindelement mit dem Namen `default` angegeben wird. Die Verwendung eines `default`-Elementes statt eines `default`-Attributes ist sinnvoll, wenn mehrzeilige `default`-Werte angegeben werden sollen.

```

1 <textfield id="tf.name" width="80" default="Standardtext" maxlength="200" />
2 <textarea id="tf.anschrift" width="80" height="5">
3     <default>
4         Musterstrasse 69
5         4711 Musterstadt
6     </default>
7 </textarea>

```

Abbildung 3.25: Syntax von `textfield` und `textarea`

Alternativ kann auch ein `AutoFill`-Wert über das Attribut bzw. Element `autofill` angegeben werden. Während `default`-Werte immer Teil der Eingabe werden, werden `AutoFill`-Werte nur dann als Eingabe betrachtet, wenn der Nutzer sie ergänzt oder ändert. Unveränderte `AutoFill`-Werte werden also ignoriert. Für alle Mitarbeiter der Universität sind z. B. Teile der Email-Adresse oder der Anschrift in den meisten Fällen identisch. Falls der Nutzer diese Angaben im Eingabefeld ändert oder ergänzt, wird dies als Eingabe betrachtet. Ansonsten wird der `AutoFill`-Wert ignoriert. `Default`-Werte dagegen werden immer als Eingabe betrachtet. Wichtig ist, dass `textfield`- und `textarea`-Elemente ein innerhalb des Editors eindeutiges ID-Attribut besitzen.

```

1 <textfield id="tf.email" width="80" autofill="@uni-duisburg-essen.de" />
2 <textarea id="ta.anschrift" width="80" height="5">
3     <autofill>
4         Universitätsstr. 9-11
5         45141 Essen
6     </autofill>
7 </textarea>

```

Abbildung 3.26: AutoFill-Werte in textfield und textarea

Passwort-Eingabefelder

Passwort-Eingabefelder werden über das Element `password` erzeugt. Sie können keine default- oder autofill-Werte besitzen. Bei der Eingabe werden Sternchen statt der eingegebenen Zeichen angezeigt. Das Attribut `width` gibt die Breite des Eingabefeldes in Anzahl Zeichen an. Leider ist es bei allen Browsern so, dass dieses Feld keine vorhandenen Inhalte bearbeiten kann, d. h. Passworte sind aus Sicherheitsgründen immer neu einzugeben.

```
1 <password width="10" />
```

Abbildung 3.27: Syntax des password-Elements

Einfache Checkboxes

Das Element `checkbox` generiert eine alleinstehende Checkbox, bei der über einen „Haken“ ein Wert ein- oder ausgeschaltet werden kann. Das Attribut `value` gibt den Wert an, der gesetzt werden soll, wenn die Checkbox markiert ist. Das Attribut `checked` gibt mit den möglichen Werten `true` und `false` an, ob als default die Checkbox markiert sein soll oder nicht. Die Beschriftung der Checkbox erfolgt über ein Attribut `label` bzw. über untergeordnete mehrsprachige label-Elemente, vgl. dazu den Abschnitt zu Beschriftungen.

```

1 <cell ... var="@raucher">
2     <checkbox value="true" checked="false" label="Ja, ich bin Raucher" />
3 </cell>

```

Abbildung 3.28: Syntax für Checkboxes

Auswahllisten

Auswahllisten werden über das Element `list` erzeugt. Die in der Liste darzustellenden Werte werden über geschachtelte `item`-Elemente angegeben. Das Attribut `type` gibt an, in welchem Format die Listenwerte dargestellt werden. Der Typ `dropdown` stellt die Listenwerte als einzeilige Dropdown-Auswahlliste dar. Das Attribut `width` gibt dabei die Breite des Auswahlfeldes in CSS-Syntax an, d.h. es sind Angaben in Anzahl Zeichen oder Pixeln etc. möglich. Der Typ `multirow` stellt die Listenwerte als mehrzeiliges Auswahlfeld dar. Das Attribut `rows` gibt dabei die Anzahl Zeilen an, die maximal zu sehen sind. Enthält die Liste mehr Werte als Zeilen, werden Scrollbalken dargestellt. Das Attribut `multiple` mit den möglichen Werten `true` und `false` gibt an, ob eine Mehrfachauswahl möglich ist. Das Attribut `default` gibt ggf. eine Vorauswahl vor.

Die in einer Liste dargestellten Werte werden über `item`-Elemente angegeben. Diese können auch mehrstufig geschachtelt sein, um eine hierarchische Struktur darzustellen (d.h. item-Element können wiederum item-Element enthalten). Eine

```

1 <list type="dropdown" width="100px" default="2" />
2   <item value="1"><label xml:lang="de">eins</label></item>
3   <item value="2"><label xml:lang="de">zwei</label></item>
4   <item value="3"><label xml:lang="de">drei</label></item>
5 </list>
6 <list type="multirow" width="100px" rows="3" multiple="true">
7   <item value="a"><label xml:lang="de">Fred</label></item>
8   <item value="b"><label xml:lang="de">Willi</label></item>
9   <item value="c"><label xml:lang="de">Otto</label></item>
10 </list>

```

Abbildung 3.29: Syntax einer Auswahlliste

Schachtelung ist nur für die Listentypen `dropdown` und `multirow` erlaubt und sinnvoll. Die Hierarchie wird dabei automatisch durch Einrückungen dargestellt. Die Beschriftung der `item`-Element wird über geschachtelte `label`-Attribute oder Elemente definiert und kann mehrsprachig angegeben werden, vgl. dazu den folgenden Abschnitt zu Beschriftungen. Die Verwendung von `<list type="multirow" multiple="true">`, d. h. mehrzeiliger Auswahllisten mit Mehrfachauswahl ist nur sinnvoll, wenn das aktuelle `var`-Attribut der Zelle auf ein Element verweist, denn Attribute sind ja grundsätzlich in XML nicht wiederholbar.

```

1 <cell ... var="programmiersprache">
2   <list type="multirow" multiple="true" rows="3">
3     <item label="Java" value="java" />
4     <item label="C++" value="cpp" />
5     <item label="Pascal" value="pas" />
6   </list>
7 </cell>

```

Abbildung 3.30: Beispiel für Mehrfachauswahl

Das Beispiel erzeugt eine dreizeilige Mehrfachauswahlliste, die ggf. in mehrfach generierte XML-Elemente „programmiersprache“ resultiert.

Radio-Buttons und Checkbox-Listen

Der Typ `radio` stellt die Listenwerte als eine Menge von Radiobuttons dar, es kann also nur ein Wert ausgewählt werden. Der Typ `checkbox` stellt die Listenwerte als eine Menge von Checkboxes dar, in diesem Fall können mehrere Werte ausgewählt sein. Bei diesen beiden Typen werden alle Werte der Liste nacheinander in einer Tabellenstruktur mit einer gewissen Anzahl Zeilen und Spalten ausgegeben. Es kann entweder die Anzahl Zeilen (Attribut `rows`) oder die Anzahl Spalten (Attribut `cols`) dieser Tabelle angegeben werden, der jeweils nicht angegebene Wert folgt automatisch aus der Anzahl der Werte in der Liste. Der Spezialfall `rows="1"` entspricht also einer Anordnung aller Listenwerte von links nach rechts in einer Zeile, der Fall `cols="1"` entspricht einer Anordnung von oben nach unten in nur einer Spalte.

Analog zu `multirow/multiple` Listen ist die Verwendung von Checkbox-Listen nur sinnvoll für die Abbildung auf XML-Elemente, da XML-Attribute nicht wiederholbar sind (vgl. vorangehenden Abschnitt).

```

1 <list type="radio" rows="2" default="new">
2   <item value="1"><label xml:lang="de">eins</label></item>
3   <item value="2"><label xml:lang="de">zwei</label></item>
4   <item value="3"><label xml:lang="de">drei</label></item>
5 </list>
6 <list type="checkbox" cols="3">
7   <item value="a"><label xml:lang="de">Fred</label></item>
8   <item value="b"><label xml:lang="de">Willi</label></item>
9   <item value="c"><label xml:lang="de">Otto</label></item>
10 </list>

```

Abbildung 3.31: Syntax von Radio-Button und Checkbox-Listen

Beschriftungen

Beschriftungen von Eingabefeldern lassen sich mit den Elementen `text` und `label` definieren. Das Element `text` erzeugt eine frei platzierbare Beschriftung etwa für eine Texteingabefeld. Beschriftungen dürfen auch XHTML Fragmente enthalten, etwa um eine fette, kursive oder mehrzeilige Darstellung zu erreichen. Mehrsprachige Beschriftungen von Elementen können über mehrere `label`-Elemente realisiert werden, wobei das Attribut `xml:lang` die Sprache angibt.

Die aktive Sprache wird wie bei anderen MyCoRe-Stylesheets über den `XSL.Lang` Parameter definiert. Wenn ein `label` existiert, bei dem `xml:lang` der aktiven Sprache entspricht, wird dieses ausgegeben. Ansonsten, wird das `label` der Default-Sprache (Parameter `XSL.DefaultLang`) ausgegeben, falls es existiert. Ansonsten wird das `Label` ausgegeben, das im `label`-Attribut oder im ersten `label`-Element enthalten ist, das kein `xml:lang`-Attribut enthält. Ansonsten wird das erste `label`-Element verwendet, das überhaupt existiert.

```

1 <text>
2   <label xml:lang="de"><b>Eins</b></label>
3   <label xml:lang="en"><b>One</b></label>
4 </text>

```

Abbildung 3.32: Syntax eines Textfeldes

Internationalisierung

Das Editor-Framework verfügt über mehrere Möglichkeiten der Mehrsprachigkeit einzelner Auswahl- oder Textfelder. Bis MyCoRe-Version 1.2 war hier nur der Einsatz von `<label xml:lang="...">`-Tags (wie oben beschrieben) möglich. Mit Version 1.3 ist die direkte Nutzung von I18N¹⁵ möglich. Um eine Abwärts-Kompatibilität zu erreichen sind die `<label>`-Tags auch weiterhin im Framework gültig. Es ist aber zu empfehlen, langfristig ältere MyCoRe-Anwendungen entsprechend umzustellen, damit wird eine strikte Trennung von Layout und sprachabhängigem Text erreicht.

Mit MyCoRe v1.3 wurde das `i18n`-Attribut in alle Editor-Framework-Stellen eingebaut, in den bisher `<label>`-Tags vorgesehen waren. Entsprechend der I18N-Spezifikation sind alle Texte nun über eine Property-Variable bekannt zu machen und in sprachabhängigen Dateien anzulegen. Diese tragen die Namen `messages_{lang}.properties` und sind bei MyCoRe im `config`-Verzeichnis untergebracht. Sollte das System für die angegebene Sprache keine solche Datei finden, so wird als Fallback die Datei `messages.properties` gesucht. Sind beide

¹⁵<http://www.debian.org/doc/manuals/intro-i18n/>

Angaben vorhanden, hat das `i18n`-Attribut Vorrang. Die nachfolgende Abbildung verdeutlicht die neue Funktion.

Abbildung 3.33: Beschriftung mit I18N

Abstandshalter

Ein Abstandshalter erzeugt leeren Raum als Platzhalter zwischen Elementen. Die Attribute `width` und `height` geben Breite und Höhe des Abstandes in CSS-Syntax, etwa in Anzahl Pixeln, an.

```
1 <space width="0px" height="50px" />
```

Abbildung 3.34: Syntax eines Abstandhalters

Externes Popup-Fenster

Ein externes Popup-Fenster kann z. B. ein Hilfetext zu einem Eingabefeld im HTML-Format enthalten. Es können aber auch beliebige andere Informationen und Funktionalitäten wie Eingabehilfen dargestellt werden. Der Startpunkt ist ein Button im Formular, der mit einer frei wählbaren Beschriftung versehen werden kann. Standardmäßig wird er als `[?]` Button im Formular dargestellt. Bei Anklicken des Buttons erscheint das codierte HTML in einem Popup-Fenster. Die Attribute `width` und `height` steuern die Breite und Höhe dieses Fensters. Das Attribut `css` gestattet die Verwendung eines externen CSS-Files. Wichtig ist, dass jedes `helpPopup` Element eine eindeutige ID besitzt. Optional kann der Popup-Fenster auch von einer externen URL geladen werden. Falls die URL mit „`http://`“ oder „`https://`“ beginnt, wird sie als absolute URL direkt verwendet, andernfalls wird die URL als relativ zur `WebApplicationBaseURL` (`context root`), dem Startpunkt der WebApplication, interpretiert.

Innerhalb des `helpPopup`-Tags gibt es mehrere Tags, welche das Aussehen und den Inhalt des Aufruf-Buttons und des Popup-Fensters bestimmen. Jedes Tag verfügt über das `xml:lang`-Attribut. Somit kann jedes Tag für die im Projekt benötigten Sprachen implementiert werden. Zusätzlich gibt es das Konstrukt `xml:lang="all"`, das anzeigt, dass das entsprechende Tag immer unabhängig von der Spracheinstellung des Systems präsentiert werden soll.

Im Einzelnen sind die Tags wie folgend definiert:

- `button` gibt die Zeichenkette an, welche der Start-Button haben soll,
- `title` gibt den Titel des Popup-Fensters an,
- `close` gibt den Text zum Schließen des Popup-Fensters an,
- `label` beinhaltet den eigentlichen darzustellenden HTML-Code oder Text.

```

1  <helpPopup id="url.help" width="400px" height="200px" css="/mycss.css">
2      <button xml:lang="all">!</button>
3      <title xml:lang="de">Hilfetext</title>
4      <close xml:lang="de">Beenden</close>
5      <label xml:lang="de">
6          In diesem Feld geben Sie bitte die <b>WWW Adresse (URL)</b>
7          der Webseite ein. Bitte achten Sie darauf, dass die Adresse mit
8          <i>http://</i> beginnt, da sonst der Link unter Umständen nicht
9          funktioniert.
10     </label>
11 </helpPopup>
12
13 <helpPopup id="url.help" width="400px" height="200px" url="hilfe/url.html" />

```

Abbildung 3.35: Syntax eines Popup-Fensters

Buttons

Über das Element `button` wird ein Knopf erzeugt, der bei Anklicken zu einer externen URL wechselt. Das Attribut `width` legt die Breite des Knopfes fest, das Attribut `label` oder ggf. mehrsprachige enthaltene `label`-Elemente legen die Beschriftung des Knopfes fest. Falls die URL mit „http://“ oder „https://“ beginnt, wird sie als absolute URL direkt verwendet, andernfalls wird die URL als relativ zur `WebApplicationBaseURL` (`context root`), dem Startpunkt der `WebApplication`, interpretiert.

```

1  <button width="100px" url="index.html">
2      <label xml:lang="de">Start</label>
3  </button>

```

Abbildung 3.36: Syntax eines einfachen Buttons

SubmitButton

Über das Element `submitButton` wird ein Knopf erzeugt, der bei Anklicken die Eingaben an das in der Konfiguration angegebene Ziel sendet. Das Attribut `width` legt die Breite des Knopfes fest, das Attribut `label` oder ggf. mehrsprachige enthaltene `label`-Elemente legen die Beschriftung des Knopfes fest.

```

1  <submitButton width="100px">
2      <label xml:lang="de">Start</label>
3  </submitButton>

```

Abbildung 3.37: Syntax des Submit-Buttons

CancelButton

Über das Element `cancelButton` wird ein Knopf erzeugt, der bei Anklicken die Bearbeitung des Formulars abbricht. Das Attribut `width` legt die Breite des Knopfes fest, das Attribut `label` oder ggf. mehrsprachige enthaltene `label`-Elemente legen die Beschriftung des Knopfes fest.

```

1  <cancelButton width="100px">
2      <label xml:lang="de">Start</label>
3  </cancelButton>

```

Abbildung 3.38: Syntax des Cancel-Buttons

Die Ziel-URL des Buttons kann auf drei verschiedene Weisen gebildet werden. Falls die URL mit „http://“ oder „https://“ beginnt, wird sie als absolute URL direkt verwendet, andernfalls wird die URL als relativ zur `WebApplicationBaseURL` (`context root`), dem Startpunkt der WebApplication, interpretiert.

Statische URL, die in der Editor-Definition als Top-Level-Element angegeben ist (d.h. auf gleicher Ebene wie die `target`-Deklaration), z. B.

```
<cancel url="pages/goodbye.html" />
```

1. URL, die durch ein Servlet oder Stylesheet zur Laufzeit gebildet wird, und die beim Aufruf des Editors über einen XSL-Parameter übergeben wird:

```
XSL.editor.cancel.url=pages/goodbye.html
```

2. URL, die aus einer Kombination einer statischen URL und eines ID=Tokens gebildet wird, das beim Aufruf des Editors über einen XSL Parameter übergeben wird. Im folgenden Beispiel wird zur Laufzeit das Token XXX durch den Parameterwert 4711 ersetzt:

```
XSL.editor.cancel.id=4711
```

```
<cancel url="servlets/SomeServlet?id=XXX" token="XXX" />
```

Ausgabe von Werten aus dem Quelldokument

Das Element `output` kann bei der Bearbeitung einer existierenden XML-Quelle verwendet werden, um den Inhalt von Attributen oder Elementen im Formular als nicht bearbeitbaren Text auszugeben. Dabei ist zu beachten, dass der ausgegebene Wert bei Abschieken des Formulars nicht weitergegeben wird. Hierfür muss ggf. ein `hidden`-Element verwendet werden. Das Attribut `default` gibt den Wert an, der ggf. ausgegeben wird, wenn das Dokument keinen Wert enthält.

Beispiel:

```
<text><label>Dokument-ID:</label></text>
```

```
<output var="@id" default="(neu)" />
```

gibt den Wert des Attributes `id` aus, sofern dieses im XML-Quelldokument existiert, ansonsten wird „(neu)“ ausgegeben.

Wiederholbare Elemente mit Repeatern erstellen

Häufig sind einzelne Eingabefelder oder ganz Panels wiederholbar. Das Element `repeater` schachtelt auf einfache Weise ein Eingabeelement oder ein ganzes Panel und macht dieses wiederholbar. Das `var`-Attribut der umgebenden Zelle muss auf ein Element verweisen, da Attribute nicht wiederholbar sind. Das Attribut `min` gibt an, wie oft minimal das wiederholte Element im Formular dargestellt wird, das Attribut `max` gibt die Maximalzahl an Wiederholungen an. In jedem Fall wird das Element minimal so oft dargestellt, wie es in der Eingabe auftritt.

Das Beispiel wiederholt die Komponente mit der ID `pcreator`. Wiederholbare Elemente werden mit +/- Buttons dargestellt, mit denen neue Eingabefelder hinzugefügt bzw. dargestellte gelöscht werden können. Bei mehr als einer aktuell dargestellten Wiederholung werden Pfeile angezeigt, mit deren Hilfe die Reihenfolge der Elemente vertauscht werden kann.

```

1 <cell ... var="creators/creator">
2   <repeater min="1" max="10">
3     <textfield id="tf.creator" width="80" />
4   </repeater>
5 </cell>
6 wiederholt ein einzelnes Eingabefeld oder eine beliebige andere Komponente.
7
8 <cell ... var="creators/creator">
9   <repeater min="3" max="6">
10    <panel ...>
11      </panel>
12    </repeater>
13 </cell>
14 wiederholt ein ganzes Panel, dass wie hier im Beispiel auch direkt eingebettet definiert
    werden kann.
15
16 <cell ... var="creators/creator">
17   <repeater min="3" max="6" ref="pcreator" />
18 </cell>

```

Abbildung 3.39: Beispiel Repeater

Der Framework-interne FileUpload

Das Element `file` erlaubt es, eine einzelne Datei zusammen mit den Formulareingaben hochzuladen oder eine auf dem Server vorhandene Datei zu löschen oder zu aktualisieren. Voraussetzung ist hierfür die vollständige Integration des Upload-Services in den jeweiligen Java-Klassen. Das Attribut `maxlength` gibt optional die maximale Grösse der hochzuladenen Datei an. Der FileUpload über ein HTML-Formular ist nur für relativ kleine Dateien mit wenigen MB zu empfehlen, andernfalls ist der in MyCoRe implementierte externe FileUpload zu nutzen. Das Attribut `accept` gibt optional den akzeptierten MIME Typ an. Ob diese Angaben beachtet werden, ist jedoch vom Browser abhängig und daher nicht verlässlich. Die Darstellung eines Datei-Upload Feldes hängt ebenfalls sehr stark vom Browser ab und kann nicht über CSS gestaltet werden. In der Regel bestellt ein solches Feld aus einem Texteingabefeld (dessen Breite in Anzahl Zeichen das `width`-Attribut angibt) und einem Button „Durchsuchen“. Beschriftung und Layout dieser Elemente sind nicht steuerbar und fest vom Browser vorgegeben, CSS Angaben funktionieren leider nicht zuverlässig. Über den Durchsuchen-Button kann eine Datei von der lokalen Platte gewählt werden, alternativ kann der Dateipfad im Texteingabefeld eingegeben werden. Mit Abschicken des Formulars wird der Dateinhalt und der Dateiname an den Server übertragen, was unter Umständen lange dauert.

```

1 <cell row="1" col="1" anchor="WEST" var="pathes/path">
2   <file width="60" maxlength="5000000" />
3 </cell>

```

Abbildung 3.40: Syntax des FileUpload

Zu beachten ist noch, dass für den FileUpload einige Property-Werte des Systems von Bedeutung sind.

- Dateien bis zu dieser Größe werden im Hauptspeicher des Servers gehalten:
MCR.Editor.FileUpload.MemoryThreshold=1000000
- Dateien bis zu dieser Größe werden für HTTP Uploads akzeptiert:
MCR.Editor.FileUpload.MaxSize=5000000

- Temporärer Speicherort für hochgeladene Dateien:
MCR.Editor.FileUpload.TempStoragePath=/tmp/upload

Das Auslesen der Dateien zur weiteren Verarbeitung kann z. B. wie folgt durchgeführt werden.

```

1  import org.apache.commons.fileupload.*;
2  ...
3  public void doGetPost(MCRServletJob job) throws Exception
4  {
5      MCREditorSubmission sub = (MCREditorSubmission)
6          (job.getRequest().getAttribute("MCREditorSubmission"));
7      List files = sub.GetFiles();
8      for( int i=0; i<files.size(); i++)
9      {
10         FileItem item = (FileItem)( files.get(i) );
11         String fname = item.getName().trim();
12     ...
13         File fout = new File(dirname,fname);
14         FileOutputStream fouts = new FileOutputStream(fout);
15         MCRUtils.copyStream( item.getInputStream(), fouts );
16         fouts.close();
17     }
18     ...
19 }
```

Abbildung 3.41: Java-Code zum Lesen des FileUpload

Weitere Hinweise zum Auslesen der hochgeladenen Dateien finden Sie in der JavaDoc-Dokumentation der Klassen `org.mycore.frontend.editor2.MCREditorSubmission` und `MCREditorVariable` sowie in der Online-Dokumentation des Apache File Upload Paketes (<http://jakarta.apache.org/commons/fileupload/>).

Die Methode `MCREditorSubmission.listFiles()` liefert eine `java.util.List` von hochgeladenen `FileItem`-Objekten. Die Apache-Klasse `FileItem` besitzt Methoden `getName()` und `getInputStream()`, die den Dateinamen und den hochgeladenen Dateiinhalt liefern. Die Methode `getFieldName()` liefert den Pfad im XML-Dokument, zu dem die hochgeladene Datei gehört.

Alternativ kann auch über diesen Pfad auf eine bestimmte hochgeladene Datei direkt zugegriffen werden: Die Eingaben aus dem Editor werden als JDOM-Dokument an das Zielservlet übergeben und stehen über `MCREditorSubmission.getXML()` zur Verfügung. Mit JDOM-Operationen kann man nun zu dem JDOM-Element oder -Attribut navigieren, für das ggf. eine Datei im Formular hochgeladen wurde. Die Methode `MCREditorSubmission.getFile()` erwartet als Argument dieses JDOM-Objekt und liefert das dazu gehörende `FileItem`.

`FileItem`-Objekte sollten unmittelbar verarbeitet werden, da die Dateiinhalte nur temporär gespeichert sind. Der hochgeladene Inhalt kann mittels `FileItem.write()` in ein persistentes `java.io.File` kopiert werden oder in einem `MCRFile` gespeichert werden.

Zu beachten ist, dass FileUploads nicht in Editoren verwendet werden sollten, die Repeater enthalten. Will man mehr als eine Datei hochladen, sollte das file-Element „manuell“ mehrfach im Editor definiert werden, es darf jedoch kein Repeater verwendet werden.

Integration externer Datenquellen

Das Editor-Framework gestattet die Einbindung externer Datenquellen in das Formular. Dies ermöglicht eine flexible Gestaltung von Eingabewerten zur Laufzeit. Dabei werden Teile der Formulardefinition „on the fly“ erzeugt und im Moment der Darstellung erst integriert.

Einfügen eines Request aus einem Servlet

Diese Beispiel zeigt den Zugriff auf ein Servlet zur Laufzeit. Dabei werden die Ausgabedaten der Servlet-Antwort vor der Integration in das Framework mittels XSLT in eine passende Form gebracht. In Beispiel wird eine MyCoRe-Klassifikation in

```

1 <list id="COrigin" width="300" type="dropdown">
2   <item value=""><label xml:lang="de">(bitte wählen)</label></item>
3   <include uri="request:servlets/MCRQueryServlet?XSL.Style=
4     classif-to-items&am p;type=class&hosts=local&lang=
5     de&query=/mycoreclass%5b@ID='DocPortal_class_00000002'%5d" />
6 </list>

```

Abbildung 3.42: Include Servlet-Request

eine Auswahlliste eingefügt.

Einfügen eines Definitionsabschnittes aus der Session

Es besteht auch die Möglichkeit, Daten dynamisch in der `MCRSession` zu hinterlegen und von dort in die Gestaltung des Formulars zu integrieren. Im ersten Schritt wird in einer Java-Klasse ein `JDOM-Element` erzeugt, welches dann im zweiten Schritt in das Formular eingebaut wird. Das Attribut `cacheable` spezifiziert dabei, ob die Daten permanent vorgehalten werden sollen. Für ständig wechselnde Daten ist hier `false` anzugeben.

```

1 org.jdom.Element items = new org.jdom.Element("items");
2 org.jdom.Element item = new org.jdom.Element("item");
3 item.setAttribute(...)
4 items.addContent(item);
5 ...
6 MCRSessionMgr.getCurrentSession().put("mylist",items);

```

Abbildung 3.43: Java-Code

Eingabevalidierung

Einführung

Die Eingabefelder von Editor-Formularen können durch Integration von Validierungsregeln direkt durch das Editor-Servlet geprüft werden, noch bevor die Eingaben an das weiterverarbeitende Zielservlet weitergegeben werden. Wenn eine Eingabe fehlerhaft ist, wird das Formular noch einmal angezeigt. Es erscheint eine Fehlermeldung. Eingabefelder mit fehlerhaften Eingaben werden im Formular optisch markiert. Zunächst ein einfaches Beispiel:

Das Element `validationMessage` enthält die Meldung, die am Kopf des Formulars erscheint, wenn ein oder mehrere Eingabefelder fehlerhaft ausgefüllt wurden. Dieses

```

1 <editor id="validation.demo">
2   <source ... />
3   <target ... />
4
5   <validationMessage>
6     <label xml:lang="de">
7       Eingabefehler: Bitte korrigieren Sie die markierten Felder.
8     </label>
9   </validationMessage>
10
11  <components root="root" var="/document">
12    ...
13    <panel id="root" lines="on">
14      <cell row="1" col="1" anchor="EAST">
15        <text label="Titel:" />
16      </cell>
17      <cell row="1" col="2" anchor="WEST" var="title">
18        <textfield id="tf.title" width="80">
19          <condition id="cond.title" required="true">
20            <label xml:lang="de">Bitte geben Sie einen Titel ein!</label>
21          </condition>
22        </textfield>
23      </cell>
24    </panel>
25    ...
26  </components>
27 </editor>

```

Abbildung 3.44: Einfaches Beispiel für Eingabevalidierung

Element ist optional, es kann ein sprachunabhängiges label-Attribut oder ein oder mehrere sprachabhängige label-Elemente enthalten.

Jedes zu validierende Eingabefeld muss eine eindeutige ID besitzen! Neben Texteingabefeldern können auch alle anderen Feldtypen wie z. B. Auswahllisten validiert werden. Nur hidden-Felder können momentan nicht validiert werden.

Das zu validierende Feld muss ein oder mehrere condition-Elemente enthalten, die wiederum eine eindeutige ID besitzen müssen. Die Attribute dieser condition-Elemente enthalten die zu prüfenden Validierungsregeln. Ein condition-Element kann ein sprachunabhängiges label-Attribut oder ein oder mehrere sprachabhängige label-Elemente enthalten, die eine Meldung für den Benutzer enthalten. Bei einer fehlerhaften Eingabe wird diese Meldung angezeigt, wenn der Benutzer mit der Maus über das Fehlersymbol fährt, das neben dem Eingabefeld angezeigt wird.

Abbildung 3.45: Fehlgeschlagene Eingabevalidierung

Ein Eingabefeld kann mehr als ein condition-Element enthalten. Dadurch können komplexe Validierungsregeln schrittweise geprüft werden, und der Nutzer bekommt eine exaktere Rückmeldung, welche dieser Regeln verletzt wurde. Wenn die Validierungsregeln auf mehrere condition-Elemente aufgeteilt wird, sollte die „required“ Regel ggf. die erste sein.

```
1 <condition id="cond.title.required" required="true">
2   <label>Bitte geben Sie einen Titel ein!</label>
3 </condition>
4 <condition id="cond.title.length" maxLength="250">
5   <label>Bitte geben Sie maximal 250 Zeichen ein!</label>
6 </condition>
```

Abbildung 3.46: Beispiel für eine Reihenfolge von Regeln

Oder

```
1 <condition id="cond.title" required="true" maxLength="250">
2   <label>Bitte geben Sie einen Titel ein (maximal 250 Zeichen)!</label>
3 </condition>
```

Abbildung 3.47: Beispiel für Regeln in einer condition

Validierungsregeln für einzelne Felder

Die Validierungsregeln für Eingabefelder werden über Attribute oder Attributkombinationen des condition-Elements definiert. Die folgenden Regeln können geprüft werden:

required="true"	Es ist eine Eingabe erforderlich. Auch Eingaben, die nur aus Leerzeichen bestehen, werden abgewiesen. Ist diese Regel verletzt, werden ggf. weitere vorhandene Regeln nicht mehr geprüft.
minLength="10"	Die Eingabe muss aus minimal 10 Zeichen bestehen.
maxLength="250"	Die Eingabe darf aus maximal 250 Zeichen bestehen.
type="integer" min="0" max="100"	Die Eingabe muss eine ganze Zahl sein. Optional kann über die Attribute <code>min</code> und/oder <code>max</code> der Wertebereich eingeschränkt werden.
type="decimal" format="de" min="0" max="3,5"	Die Eingabe muss eine Zahl (ggf. mit Nachkommastellen) sein. Optional kann über die Attribute <code>min</code> und/oder <code>max</code> der Wertebereich eingeschränkt werden. Das Attribut <code>format</code> muss einen ISO 639 Sprachcode enthalten (z.B. „de“ oder „en“), der bestimmt, wie Kommazahlen formatiert werden. Die <code>min</code> und <code>max</code> Werte müssen entsprechend diesem Format angegeben werden.
type="datetime" format="dd.MM.yyyy" min="01.01.1970" max="31.12.2030"	Die Eingabe muss ein Datums- oder Zeitwert sein. Optional kann über die Attribute <code>min</code> und/oder <code>max</code> der Wertebereich eingeschränkt werden. Das Attribut <code>format</code> muss ein <code>java.text.SimpleDateFormat</code> Pattern enthalten (z.B. „dd.MM.yyyy“), das bestimmt, wie der Wert formatiert sein muss. Die <code>min</code>

und max Werte müssen entsprechend diesem Format angegeben werden.
<code>type="string" min="a" max="zzz"</code>
Eine Eingabe kann ein beliebiger Text sein. Optional kann über die Attribute <code>min</code> und/oder <code>max</code> der Wertebereich eingeschränkt werden.
<code>regexp=".+@.+\\.+."</code>
Die Eingabe muss den angegebenen regulären Ausdruck erfüllen. Die Syntax des regulären Ausdrucks ist durch die Klasse <code>java.util.regex.Pattern</code> definiert.
<code>xsl="contains(., '@')"</code>
Die Eingabe wird gegen eine XSL Bedingung geprüft, wie sie in XSL <code>if</code> oder <code>when</code> Elementen verwendet werden kann. Der Eingabewert wird in der Bedingung durch einen Punkt referenziert.

Validierungsregeln für Feldkombinationen

Es ist ebenfalls möglich, die Eingaben zweier Felder gegeneinander zu prüfen. Dazu wird ein `condition`-Element in das die beiden Felder enthaltende Panel eingefügt. Die Werte der beiden Felder können dann über Vergleichsoperatoren gegeneinander geprüft werden. Beispiele:

- Ein Eingabeformular zum Ändern eines Passwortes enthält zwei Passwortfelder, so dass das Passwort wiederholt eingegeben werden muss. Die Eingaben in diesen beiden Feldern müssen identisch sein.
- Ein Eingabeformular enthält zwei Eingabefelder für einen Datumsbereich. Das Datum im Feld `validFrom` muss kleiner oder gleich dem Datum im Feld `validTo` sein.

Zunächst werden die Werte beider Eingabefelder `validFrom` und `validTo` nach den gleichen Regeln auf die Eingabe eines korrekten Datums geprüft. Wenn mindestens eines der Felder eine Eingabe enthält (was ggf. über ein `required`-Attribut erzwungen werden könnte), werden die Werte der Felder miteinander verglichen.

```

1 <textfield id="tf.date" width="10">
2   <condition id="cond.date" type="datetime" format="dd.MM.yyyy">
3     <label xml:lang="de">
4       Bitte geben Sie das Datum im Format TT.MM.JJJJ ein,
5       z. B. 01.09.2005
6     </label>
7   </condition>
8 </textfield>
9
10 <panel lines="off">
11
12   <cell row="2" col="1" anchor="EAST">
13     <text label="Gültig von:" />
14   </cell>
15   <cell row="2" col="3" anchor="EAST">
16     <text label="Gültig bis:" />
17   </cell>
18
19   <cell row="2" col="2" anchor="WEST" var="validFrom" ref="tf.date" />
20   <cell row="2" col="4" anchor="WEST" var="validTo" ref="tf.date" />
21
22   <condition id="cond.valid" field1="validFrom" field2="validTo"
23     operator="<=" type="datetime" format="dd.MM.yyyy">
24     <label>Datum "Gültig von" muss <= Datum "Gültig bis" sein!</label>
25   </condition>
26
27 </panel>

```

Abbildung 3.48: Codebeispiel

```
field1="validFrom"
```

Erstes zu vergleichende Feld. Syntax und Funktionsweise entsprechen dem Verhalten des Attributes `var` für Zellen.

```
field2="validTo"
```

Zweites zu vergleichende Feld. Syntax und Funktionsweise entsprechen dem Verhalten des Attributes `var` für Zellen.

```
operator="<="
```

Vergleichsoperator (hier `>=`). Es können die Operatoren `=`, `>`, `<`, `>=`, `<=`, `!=` (für ungleich) verwendet werden.

```
type="datetime"
format="dd.MM.yyyy"
```

Datentyp und Format der Eingabe, analog zu den gleichnamigen Attributen, die im vorangehenden Abschnitt bereits erläutert wurden.

Geplante Validierungsmöglichkeiten

Die folgenden Validierungsmöglichkeiten sind noch nicht implementiert, sollen aber kurzfristig realisiert werden:

- Validierung von `hidden` Feldern. Dies ist insbesondere nötig in Verbindung mit der `subselect` Funktion z. B. bei einer externen Auswahl einer Klassifikationskategorie um sicherzustellen, dass mindestens eine Kategorie ausgewählt wird.
- Validierung durch externen Java-Code. Dies ist z. B. nötig, um gültige ISBNs, NBN Prüfsummen, existierende UserIDs etc. über eine externe Java-Methode prüfen zu können.

3.7 Klassifikationsbrowser

Mit dem Klassifikationsbrowser bietet MyCoRe die Möglichkeit über Klassifikationen zu navigieren und so zu den zugehörigen Dokumenten zu gelangen.

Für jede Klassifikation die über einen Link im Navigationsmenü erreicht werden soll, muss ein Beschreibungsblock für das Browsen angelegt werden.

Beispiel für Browsen in Dissertationen:

Der Klassifikationsbrowser für die Klassifikation der Dissertationen wird (z.B. bei einer lokalen Beispiel-Installation) durch folgenden Request aufgerufen:

```
http://localhost:8080/docportal/browse/dissertationen
```

Dabei werden alle Requests von `http://localhost:8080/docportal/browse/*` vom Servlet `MCRClassificationBrowser` entgegengenommen. Hier wird der Browserpfad, der nach `...browse/` stehen muss, ausgewertet und die dem Browserpfad entsprechende Klassifikation der Navigation zugrunde gelegt.

In der Properties-Datei muss zum zugehörigen Browserpfad `dissertationen` ein Konfigurationsblock angelegt sein. Ist ein Parameter im Konfigurationsblock nicht enthalten, wird versucht in einem Default-Konfigurationsblock den entsprechenden Default-Parameter zu lesen. Ist dieser nicht vorhanden erhält man einen `MCRConfigurationException` für die Pflichtfelder.

Über Klassifikationen browsen

14 Dokumente die dem Klassifikationstypen zugeordnet sind.

Liste und Struktur der Fakultäten und Institute der Universität Rostock

- ☞ [__8 Dok.] Universität Rostock
- ☞ [__3 Dok.] Fakultäten der Universität Rostock
- ☞ [__5 Dok.] Zentrale Einrichtungen der Universität
 - ☑ [__5 Dok.] Universitätsbibliothek
 - ☑ [__0 Dok.] Rechenzentrum
 - ☑ [__0 Dok.] Medienzentrum
 - ☞ [__0 Dok.] Kliniken der Uni Rostock



Abbildung 3.49: Klassifikationsbrowser in einer MyCoRe-Anwendung

Der Konfigurationsblock

MCR.ClassificationBrowser.dissertationen.Classification

Dieser Parameter ordnet dem Browserpfad eine Klassifikation (MyCoRe ID der Klassifikation) zu.

Beispiel:

```
MCR.ClassificationBrowser.dissertationen.Classification  
    = Docportal_class_00000009
```

Hinter der Klassifikation mit der ID Docportal_class_00000009 verbirgt sich in diesem Fall die DDC-Klassifikation, nach der typischerweise Dissertationen klassifiziert werden.

MCR.ClassificationBrowser.dissertationen.EmbeddingPage

Legt fest, in welche Datei die Klassifikation eingebettet werden soll. Diese Datei muss das Tag `<classificationbrowser/>` enthalten. An der Position dieses Tags wird der Inhalt (der generierte XML-Klassifikationsbaum) zur Laufzeit eingefügt. Durch die Variabilität ist es möglich, Klassifikationen nach Belieben in beliebige XML-Dokumente einzubetten.

Beispiel:

```
MCR.ClassificationBrowser.dissertationen.EmbeddingPage  
    = mcr_doc_browse.xml
```

Die Datei `mcr_doc_browse.xml` ist im DocPortal enthalten.

MCR.ClassificationBrowser.dissertationen.Style

Legt das Stylesheet fest, mit dem das erzeugte XML-Dokument transformiert wird. Dabei setzt sich der Name des auszuwählenden Stylesheets aus der EmbeddingPage und dem Style zusammen.

Beispiel:

```
MCR.ClassificationBrowser.dissertationen.Style = search
```

Es wird die Datei `mcr_doc_browse-search.xsl` zur Transformation verwendet.

MCR.ClassificationBrowser.dissertationen.EmptyLeafs

Klassifikationen besitzen im Allgemeinen eine baumartige Struktur. Zu jeder im Baum dargestellten Kategorie wird die Anzahl der damit klassifizierten Objekte ermittelt und (je nach Stylesheet) dargestellt. Kategorien die keine Objekte referenzieren (Anzahl der Dokumente=0) können durch diese Option ausgeblendet werden. Mögliche Werte sind `yes|no`.

Beispiel:

```
MCR.ClassificationBrowser.dissertationen.EmptyLeafs = yes
```

Zeigt auch die leeren Kategorien der Klassifikation an.

MCR.ClassificationBrowser.dissertationen.View

Bei der Darstellung einer Klassifikation kann gewählt werden, ob der gesamte Baum oder nur die jeweils aktuelle Ebene angezeigt werden soll.

Durch Zu- und Aufklappen wird die nächste Ebene dann ein- bzw. ausgeblendet. Mit `View=flat` wird dann in die aktuelle Ebene gewechselt und die vorherige ist nicht mehr sichtbar. Mit `View=tree` wird immer der gesamte Baum dargestellt und die gewählten Ebenen werden auf- und zugeklappt, sodass es möglich ist, gleichzeitig mehrere Kategorien verschiedener Ebenen darzustellen.

Mögliche Werte sind `flat|tree`

Beispiel:

```
MCR.ClassificationBrowser.dissertationen.View = flat
```

MCR.ClassificationBrowser.dissertationen.Doctype

Über den Parameter `Doctype` wird die Auswahl der Dokumenttypen gesteuert, die bei der Abfrage der Dokumente berücksichtigt werden sollen.

Der angegebene Typ muss in der Datei `mycore.properties` über den Parameter `MCR.type_[Doctype]=[Dokumententypliste]` definiert sein. Als möglicher Wert, kann jede Angabe aus `MCR.type_*` verwendet werden.

Beispiel:

```
MCR.ClassificationBrowser.dissertationen.Doctype = mydisshabs
```

Es wird `MCR.type_mydisshabs` in der Datei `mycore.properties` herausgesucht, der z.B. mit `MCR.type_mydisshabs=disshab` belegt ist. Somit werden bei dieser Klassifikation alle Dokumente vom Typ `disshab` berücksichtigt.

Für `Doctype=alldocs` wird z.B. der Wert `MCR.type_alldocs=document, article, codice, disshab, professorum, portrait` gefunden und alle Dokumente, die einem Typ aus der Typliste entsprechen, bei der Suche berücksichtigt.

MCR.ClassificationBrowser.dissertationen.Restriction

Es ist möglich für die referenzierten Objekte eine Restriktion zu setzen um die Auswahl durch eine weitere Kategorie einer beliebigen Klassifikation weiter einzuschränken. Es werden dann nur die Objekte berücksichtigt, die beiden Kriterien genügen.

Beispiel:

```
MCR.ClassificationBrowser.dissertationen.Restriction
= Docportal_class_00000005##TYPE0007
```

Mit dieser Restriktion werden nur die Dissertationen berücksichtigt, die auch der Klassifikation `Docportal_class_00000005` mit der Kategorie `TYPE0007` genügen. Das sind im Beispiel als Ergebnis alle Dissertationen im PDF-Format.

MCR.ClassificationBrowser.dissertationen.searchField

Hier wird das bei der Ausführung der Query zu durchsuchende Feld festgelegt.

Beispiel:

```
MCR.ClassificationBrowser.dissertationen.searchField = ddc1
```

Es wird das DDC-Tag in den Metadaten der Dokumente als Suchparameter genutzt.

MCR.ClassificationBrowser.dissertationen.Comments

Kommentare, die optional beim Erstellen der Klassifikation für Kategorien angegeben sein können, werden im Baum mit dargestellt. Mögliche Werte sind `true|false`

Beispiel:

```
MCR.ClassificationBrowser.dissertationen.Comments = true
```

Der Default-Konfigurationsblock

Die Defaultbelegung für die notwendigen Parameter enthält folgende Werte:

```
MCR.ClassificationBrowser.default.EmbeddingPage = mcr_doc_browse.xml
MCR.ClassificationBrowser.default.Style         = search
MCR.ClassificationBrowser.default.EmptyLeafs    = yes
MCR.ClassificationBrowser.default.View          = tree
MCR.ClassificationBrowser.default.Comments      = true
MCR.ClassificationBrowser.default.Doctype       = alldocs
MCR.ClassificationBrowser.default.searchField   = cltype
```

Insbesondere die Belegung des `searchField`-Parameters kann natürlich zu falschen Ergebnissen bei der Dokumentreferenzierung und der damit verbundenen Anzeige der Anzahl der Dokumente, führen!

3.8 Klassifikationseditor

Mit dem Klassifikationseditor als Webinterface, gibt es neben der Kommandozeilen-Schnittstelle eine weitere Möglichkeit, Klassifikationen im System zu erzeugen, zu bearbeiten und zu löschen.

Um Klassifikationen editieren zu können, muss der angemeldete Nutzer je nach Bedarf über folgende Rechte verfügen:

<code>create-classification</code>	zum Erzeugen von Klassifikationen
<code>modify-classification</code>	zum Modifizieren von Kategorieinträgen in einer Klassifikation
<code>delete-classification</code>	zum Löschen einer Klassifikation

(siehe Kapitel 5.3.4 Administrative Privilegien im UserGuide).

Konfiguration

Die Konfiguration des Klassifikationseditors erfolgt über die Angaben in der Datei `mycore.properties.classification`. Wie man Seiten, die im Fehlerfall angezeigt werden, definieren kann, zeigt nachstehender Auszug aus der Datei:

```
MCR.classeditor_page_error_user    = editor_error_user.xml
MCR.classeditor_page_cancel        = classeditor_cancel.xml
MCR.classeditor_page_error_id      = classeditor_error_clid.xml
MCR.classeditor_page_error_move    = classeditor_error_move.xml
MCR.classeditor_page_error_delete  = classeditor_error_delete.xml
```

Analoge Properties zum Klassifikationsbrowser:

```
MCR.classeditor.EmbeddingPage = mcr_doc_browse.xml
MCR.classeditor.Style         = edit
```

Für die Klassifikationen, die über die ID definiert sind, wird analog zum Browsen der Browserpfad definiert. Dies dient dazu, um auch hier die richtige Zuordnung zu den, durch die Klassifikation referenzierten Dokumenten zu besitzen:

```
MCR.classeditor.atlibri_class_00000012 = codices
MCR.classeditor.atlibri_class_00000013 = archiv
MCR.classeditor.atlibri_class_00000009 = ddc
MCR.classeditor.atlibri_class_00000008 = eval
MCR.classeditor.atlibri_class_00000007 = dnb
MCR.classeditor.atlibri_class_00000006 = medien
MCR.classeditor.atlibri_class_00000005 = type
MCR.classeditor.atlibri_class_00000003 = origin
```

Start des Klassifikationseditors

Der Editor wird über den Klassifikationsbrowser mit dem Parameter `mode=edit` gestartet.

Beispiel:

```
http://localhost:8080/docportal/browse?mode=edit
```

Die Handhabung des Klassifikationseditors ist im UserGuide beschrieben.

4 Module

Module sind eine sinnvolle Ergänzung der MyCoRe-Kern-Komponenten. Sie können optional in die Applikationen der Anwender eingebaut werden und sollen diese funktional bereichern.

4.1 Das SimpleWorkflow-Modul

Allgemeines

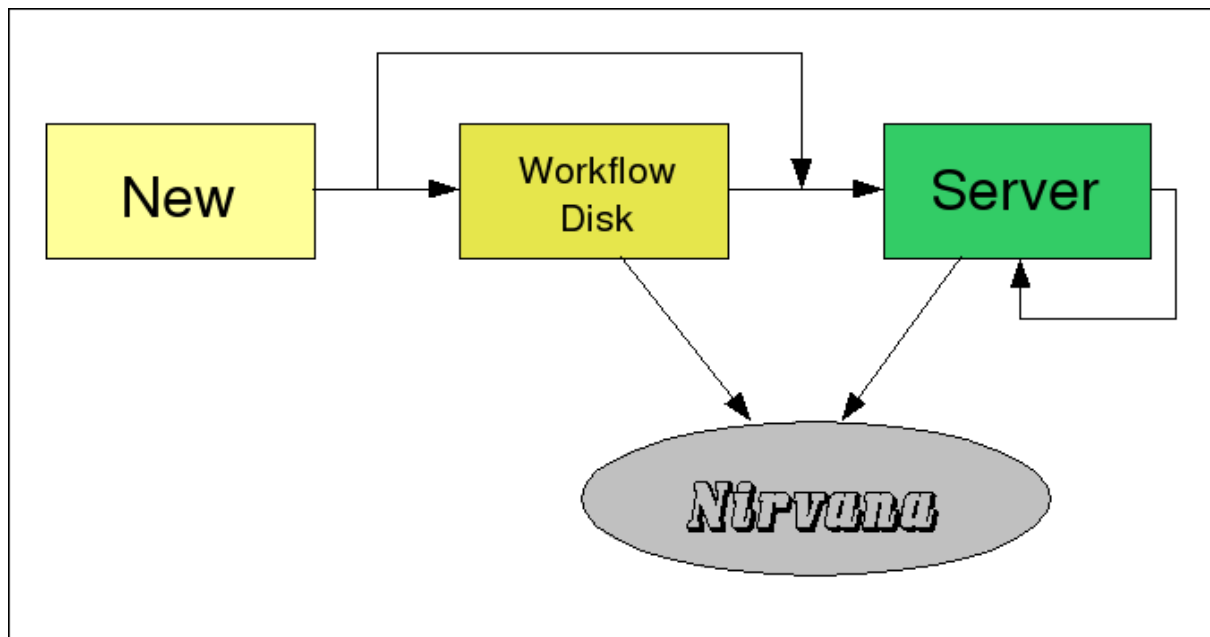


Abbildung 4.1: Grundübersicht des SimpleWorkflow

Für die Erstellung einfacher Anwendungen, welche nur einen relativ primitiven Arbeitsablauf bedingen, war es notwendig ein einfaches Werkzeug zur Gestaltung dieser Abläufe anzubieten. So entstand die Idee des SimpleWorkflow. Eigentlich handelt es sich dabei gar nicht um einen Workflow, sondern eher um eine Menge von kleinen Werkzeugen, die über HTTP-Request zu einem interaktiven Arbeitsablauf zusammengefügt werden können. Der Begriff *Workflow* soll dabei die Bearbeitungsebene zwischen dem ersten Erstellen eines Objektes und dessen Ablage in den Server und die dabei vor sich gehenden Arbeitsschritte beschreiben. Physisch handelt es sich um ein Verzeichnis `workflow`, unter welchem für jeden Objekttypen Unterverzeichnisse angelegt sind und in welchem die Daten zwischengespeichert werden. Konsultieren Sie zum besseren Verständnis auch die Beschreibung im MyCoRe-UserGuide.

Der SimpleWorkflow besteht im wesentlichen aus einer Sammlung von Servlets, welche über HTTP-Request angesprochen, verschiedene Bearbeitungsprozesse initiieren. Dabei wird gleichzeitig eine Berechtigungsprüfung für den Zugriff auf einzelne Aktionen durchgeführt. Einerseits werden dabei die allgemeinen Privilegien des aktuellen Benutzers geprüft, andererseits in den Fällen der Veränderung von Daten werden noch Eignerrechte zum Dokument mit in die Prüfung einbezogen. Die allgemeinen Rechte sind in der folgenden Tabelle zu sehen.

Privileg	Bedeutung
create-<type>	Gestattet das Anlegen neuer Objekte.
modify-<type>	Gestattet das Bearbeiten von Objekten im Workflow (d. h. auf dem Plattenbereich).
delete-<type>	Gestattet das Löschen von Objekten aus dem Workflow (d. h. dem Plattenbereich).
commit-<type>	Gestattet das Einstellen der Objekte aus dem Workflow (d. h. dem Plattenbereich) in den Server oder das Bearbeiten der Daten im Server.
remove-<type>	Gestattet das Löschen von Objekten aus dem Server.
editor	Dies ist ein Spezial-Privileg, um bei nicht entscheidbaren Prüfungen (z.B. durch inkorrekte Einträge) Nutzern die Bearbeitung trotzdem zu gestatten.

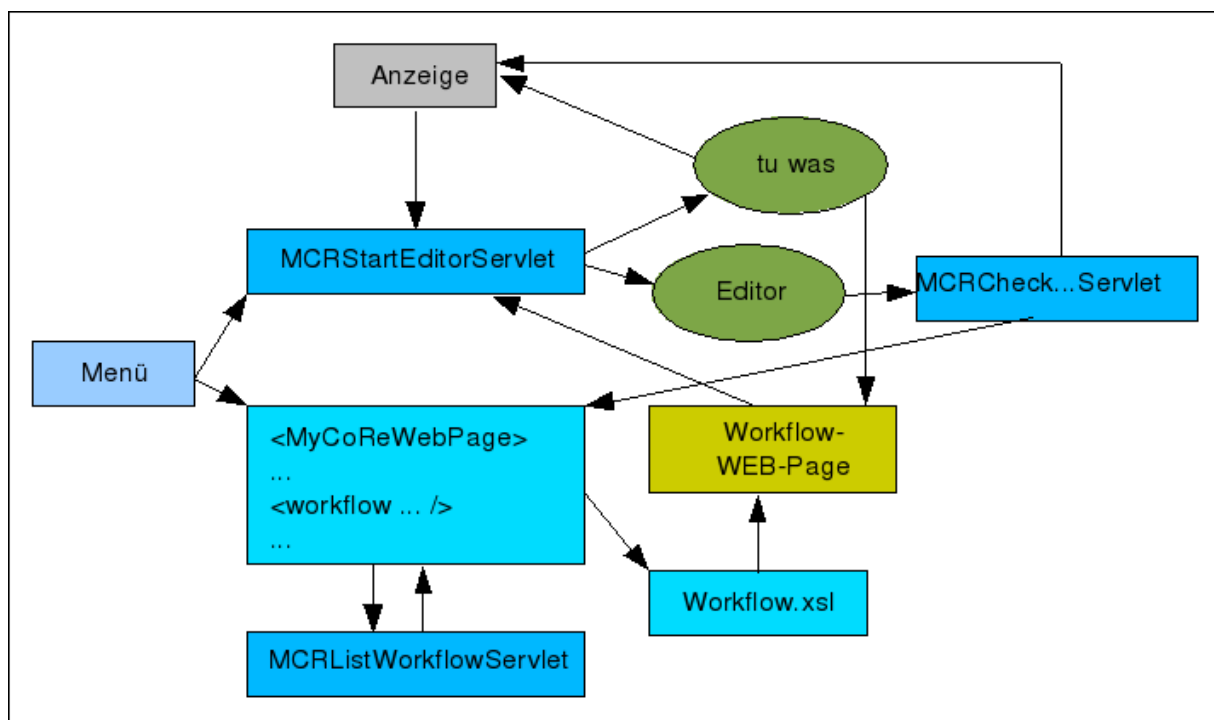
Tabelle 4.1: Privilegienliste für den SimpleWorkflow

Komponenten und Funktionen

Der SimpleWorkflow besteht aus einer Reihe von Servlets. Die folgende Tabelle listet die Servlets auf.

Servlet	Funktion
MCRStartEditorServlet	Das Servlet dient als Startpunkt für alle Arbeiten mit dem SimpleWorkflow.
MCRCheckCommitDataServlet	Wird vom Editor über ein <target>-Tag aufgerufen und schreibt die Metadaten nach Bearbeitung in den Server.
MCRCheckEditDataServlet	Wird vom Editor über ein <target>-Tag aufgerufen und schreibt die Metadaten nach deren Bearbeitung auf die Platte.
MCRCheckNewDataServlet	Wird vom Editor über ein <target>-Tag aufgerufen und schreibt die neuen Metadaten auf die Platte.
MCRCheckCommitFileServlet	Wird vom Editor über ein <target>-Tag aufgerufen und schreibt die Derivate-Daten nach Bearbeitung in den Server.

Servlet	Funktion
MCRCheckNewFileServlet	Wird vom Editor über ein <target>-Tag aufgerufen und schreibt die neuen Derivate-Daten auf die Platte.
MCRFileListWorkflowServlet	Listet die auf der Platte befindlichen Dateien in den Derivaten auf.
MCRFileViewWorkflowServlet	Gestattet den Zugriff auf eine Derivate-Datei auf der Platte.
MCRListDerivateServlet	Listet alle auf der Platte befindlichen Derivate auf.
MCRListWorkflowServlet	Erzeugt einen XML-Baum, welcher zur Darstellung des Workflow (Platteninhaltes) benötigt wird.

Tabelle 4.2: Übersicht der SimpleWorkflow-Servlets**Abbildung 4.2:** Ablaufschema im SimpleWorkflow

Die folgende Abbildung soll noch einmal die Beziehungen der einzelnen Komponenten verdeutlichen. Hier gibt es zwei Komplexe. Der erste arbeitet mit dem Plattenzwischenspeicher und stellt einen simplen *Workbasket* dar. In diesen Korb können Objekte neu eingestellt, bearbeitet, ergänzt, geprüft oder wieder gelöscht werden. Ist dieser Arbeitsschritt fertig, so kann das Objekt in den Server hoch geladen werden. Hier gibt es wieder die Möglichkeit, so der Nutzer die Berechtigung dazu hat, Objekte zu bearbeiten, zu verändern oder zu löschen. Alle diese Schritte arbeiten direkt gegen den Server. Ausgangspunkt aller Aktivitäten ist dabei das `MCRStartEditorServlet`. Hier wird beim Aufruf eine Aktion mit gegeben, welche den weiteren Ablauf bestimmt. Entweder werden jetzt die *ToDo's* direkt ausgeführt

(Löschen) oder es wird z. B. eine Web-Seite mit einer Editor-Maske aufgerufen. Diese wiederum beinhaltet im `<target>`-Tag das zu benutzende Verarbeitungs-Servlet, welches dann je nach Aufgabe wieder zu einer Web-Seite oder der Workflow-Ansicht verzweigt.

Installation

In DocPortal ist ein großer Teil der Funktionen bereits integriert. Bei der Verwendung in anderen Applikationen sind wieder die zwei Funktionskomplexe zu unterscheiden – für die einfache Workflow-Gestaltung auf der Platte müssen sie:

1. das `workflow.xml` Stylesheet in Ihre Web-Applikation kopieren. Da das Stylesheet in der Darstellung der Daten sehr anwendungsspezifisch ist, sollten Sie es an das Design Ihrer Applikation anpassen. Die im Kern mitgelieferte Version kann nur als Vorlage dienen!
2. die folgende Zeile in das XSL-Stylesheet kopieren, welches die Auswertung Ihrer XML-Web-Seiten realisiert (z. B. `MyCoReWebPage.xml`).
`<xsl:include href="workflow.xml" />`
3. Weiterhin finden Sie im Kern das Stylesheet `mycoreobject-document-to-workflow.xml`. Dieses ist eine Transformationsvorlage für die Transformation von den XML-Objekt-Metadaten in eine Workflow-interne XML Struktur. Für jeden Ihrer Metadaten-Typen muss eine solche Konverter-Datei mit Namen `mycoreobject-<type>-to-workflow.xml` in Ihrer Anwendung vorhanden sein.
4. die Icon-Dateien in Ihre Web-Applikation zu kopieren.
5. die Servlets in der Datei `web.xml` Ihrer Web-Applikation aus der mitgelieferten Datei `web.xml` zu ergänzen. Die Java-Klassen sind bereits im `mycore-....jar` Archiv enthalten.
6. Als letztes muss der Workflow in eine XML-Webseite integriert und diese entsprechend über Menüpunkte aufgerufen werden.

Die Integration des SimpleWorkflow in die Präsentationsseiten erfolgt unter Einbeziehung der bereitgestellten Icons und eines dahinter liegenden Linke oder Formulars. Dieses kann an beliebiger Stelle in der Präsentation platziert werden.

```

1  <MyCoReWebPage>
2      <section title="Bearbeiten von Objekten aus dem Workflow" xml:lang="de">
3          <table id="contentArea" width="100%" cellpadding="20" cellspacing="0">
4              <tr>
5                  <td>
6                      <center>
7                          <span class="textboldnormal">
8                              <u>Bearbeiten und Verwalten von Dokumenten<br />
9                                  (nur für berechnigte Editoren)</u>
10                             </span>
11                             <p></p>
12                             <workflow type="document" step="editor" />
13                         </center>
14                     </td>
15                 </tr>
16             </table>
17         </section>
18 </MyCoReWebPage>

```

Abbildung 4.3: Beispiel-Web-Seite für den Aufruf eines Workflow

```

<a href="{ $ServletsBaseUrl }MCRStartEditorServlet{ $HttpSession }?
    tf_mcrId={ $id }&
    se_mcrId={ $id }&
    type={ $type }&
    step=commit&
    todo=seeditobj"
    >
</a>

```

Konfiguration

Die Konfiguration des SimpleWorkflow beschränkt sich auf einige wenige Dinge. Anzugeben sind:

- MCRObjektID-Projektnamen für die einzelnen Metadatentypen,
- die Verzeichnisnamen des Plattenspeichers,
- die Mail-Verteilung und
- die Lokation für das Nachladen von Informationen.

```
#####
# SimpleWorkflow                                                    #
#####
# The project ID
MCR.default_project_id=DocPortal
MCR.default_project_type=document
MCR.document_project_id=DocPortal
...
# Editor path
MCR.editor_document_directory=/home/mcradmin/docportal/workflow/docume
nt
...
# Editor Mail addresses for Messages add1@serv1,add2,@serv2,...
# Editor flags for todo and type
MCR.editor_document_wnewobj_mail=user@domain.xyz
MCR.editor_document_weditobj_mail=
MCR.editor_document_wcommit_mail=user@domain.xyz
MCR.editor_document_wdelobj_mail=
MCR.editor_document_seditobj_mail=user@domain.xyz
...
MCR.editor_mail_sender=user@domain.xyz
MCR.editor_mail_application_id=MyCoReSample
# Generic mail configuration for MCRMailer
MCR.mail.server=myhost.de
MCR.mail.protocol=smtp
MCR.mail.debug=false
# Editor HostAlias for Classification
MCR.editor_baseurl=local
```

Hier noch einige Hinweise:

Kopieren Sie den Abschnitt in Ihre Datei `mycore.properties.private` und ergänzen Sie die Einträge mit den Daten Ihrer Anwendung. Ein Teil der vom `MCRStartEditorServlet` veranlassten Aktionen ist so implementiert, dass sie auf Wunsch eine Mail an eine oder mehrere Adressen schicken können. Wenn Sie für den Konfigurationswert, welcher durch das Paar `type_todo` beschrieben wird, nichts angeben, so wird die Mail unterdrückt. Alle Angaben in diesem Konfigurationsabschnitt sollten selbsterklärend sei.

4.2 Das Webservice-Modul

Allgemeines

Für das Webservice-Module wird das Axis-Framework verwendet (<http://ws.apache.org/axis/>). Über den Webservice können MyCoRe-Objekte geholt und Queries in der neuen Abfragesprache ausgeführt werden. In dem Modul ist auch ein Beispiel eines Clients enthalten, der den installierten Webservice von MyCoRe nutzt.

Installation des Webservices

Hierzu müssen in der `build.xml` des Webservice-Moduls die folgenden Properties angepasst werden:

```
<property name="hostname"      value="localhost" />
<property name="port"          value="8080" />
<property name="servletpath"    value="/services/AdminService" />
```

Docportal wird wie gewohnt gestartet und durch Eingabe der Url

```
http://localhost:8080/servlets/AxisServlet
```

wird geprüft, ob Axis richtig konfiguriert wird. Danach wird das Deployment des Webservice mit `ant deploy` (`mycore/modules/module-webservices/build.xml`) durchgeführt. Ein erneuter Aufruf des AxisServlets zeigt den MyCoRe-Webservice mit dem Namen `MCRWebService` und den Methoden `MCRDoRetrieveObject` und `MCRDoQuery` an. Ein Klick auf die WSDL (Web Service Description Language) von `MCRWebService` zeigt die Parameter und Datentypen der Rückgabewerte an.

Mit `ant undeploy` wird das Undeployment des Webservice `MCRWebService` durchgeführt.

Mit

```
http://localhost:8080/services/MCRWebService?
method=MCRDoRetrieveObject&id=<mcrId>
```

kann überprüft werden, ob der `MCRWebService` Ergebnisse liefert.

Client für den Webservice erzeugen

Hierzu muss das Property der `build.xml` angepasst werden:

```
<property name="webservice"    value="/services/MCRWebService" />
```

Mit '`ant client.cmd`' wird ein Webservice-Client erstellt, der den Webservice `MCRWebService` nutzt. Hierzu werden WSDL-Informationen vom Server geholt und mittels Axis die Stubs generiert und im Verzeichnis `module-webservices/build/src` gespeichert. Anschließend werden alle Daten kompiliert und `module-webservices/build/bin/wsclient.cmd` (Windows) gebaut.

[ToDo]: sh für Linux bauen]

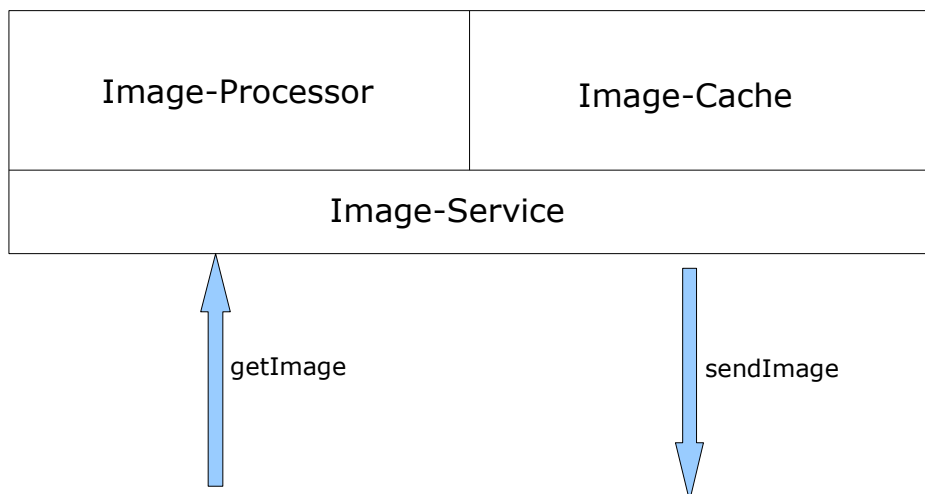
4.3 Bildbetrachter

Allgemeines

DocPortal bietet einen sogenannten „Image-Viewer“ für das komfortable Betrachten von Bilddateien an. Dieser eignet sich sehr gut für Bildarchive, oder jegliche Content-Repositories in denen Bilddaten verwaltet und angezeigt werden sollen.

Der Bildbetrachter basiert, wie MyCoRe, auf Java und XML/XSL. Grundsätzlich wird die Funktionalität in zwei getrennten Modulen, „Module-Imaging“ im MyCoRe-Kern und „Module-IView“ in DocPortal realisiert. Die folgende Abbildung zeigt die allgemeine Systemarchitektur.

MyCoRe - Module-Imaging



DocPortal - Module-IView



Module-Imaging

Das „Module-Imaging“ befindet sich im MyCoRe-Kern und bietet eine API um Bilddaten, in MyCoRe speziell abzuspeichern, performant zu laden und gegebenenfalls zu verändern (Skalierung, Bereichsextraktion).

Das Modul benutzt die Bildbearbeitungsbibliothek „JAI“ von Sun als Grundlage für Bildberechnungen.

Die folgende Abbildung zeigt gibt Überblick, zu den verschiedenen Klassen und deren Zusammenhang.

Bild von Chi.

Bilder lesen

Bilder können über den Image-Service geladen werden. Das Modul organisiert dabei automatisch den für das Bild optimalen Ladevorgang. Entsprechend der Ausgabegröße wird entschieden, ob der Cache benutzt werden kann oder eine Live-Skalierung durchgeführt werden muss.

Hinweis: Beachten sie, dass performante Leseoperation nur garantiert werden können, wenn der Cache eingeschaltet ist. Lesen sie dazu den entsprechenden Abschnitt „Cache“.

Bilder schreiben

Es ist nicht nötig Bilder speziell abzuspeichern. Allein bei eingeschaltetem Cache, werden die Bilder redundant abgelegt. Der gecachten Bilder werden im selben Content-Store gespeichert, wie der restliche Content. Es ist nicht nötig, dafür weitere Konfigurationen vorzunehmen.

Cache

Der Cache speichert grundsätzlich Bilder redundant im Repository. Dabei werden die am häufigsten angeforderten Bildgrößen bereits fertig berechnet vorgehalten. Dadurch müssen bei Leseoperationen keinerlei Berechnungen mehr durchgeführt werden und die Ladezeit wird sehr gering.

Es können grundsätzlich drei verschiedene Caches angelegt werden:

1. Thumbnail-Cache

Dieser Cacheteil speichert eine Thumbnailversion des Bildes ab. Die Größe des Thumbnails wird in der Properties-Datei `$DocPortal/modules/module-iview/mycore.properties.iview` im Bildbetrachter („MCR-IView“) als Pixel angegeben.

```
MCR.Module-iview.thumbnail.size.width=100
```

```
MCR.Module-iview.thumbnail.size.height=75
```

2. Übersichtsbilder

Dieser Cacheteil speichert Bilder in einer gebräuchlichen Bildschirmauflösung um Übersichten („Bildgröße an Seite-, Breite angepasst“) vorzuhalten.

Weiterhin wird er dazu benutzt, Ausgabegrößen, die kleiner sind als diese Cachegröße, beschleunigt zu berechnen.

Zur Konfiguration sollte die am meisten benutzte maximale Monitorauslösung der Nutzer angegeben werden. In der Properties-Datei `$DocPortal/modules/module-iview/mycore.properties.iview` im Bildbetrachter („MCR-IView“) sind die beiden Werte als Pixel anzugeben:

```
MCR.Module-iview.cache.size.width=1280
```



```
MCR.Module-iview.cache.size.height=1024
```

3. Originalgröße

Dieser Cacheeintrag speichert vom Originalbild ein gekacheltes TIFF in der Originalauflösung ab. Das ist nötig, wenn Nutzer innerhalb von Bildern , wie zum Beispiel bei Landkarten, navigieren müssen.

Durch diesen Cacheeintrag müssen beim Lesen von Ausschnitten aus Bildern, nur die Informationen der Ausschnitte geladen werden, und nicht mehr das ganze Bild.

Bilder die bereits im Original im TIFF-Format vorliegen, wird kein eigener Cache erzeugt, sondern das Originalbild gekachelt und überschrieben. Für alle Formate außer TIFF, werden separate Cache-Einträge angelegt.

In der Properties-Datei `$DocPortal/modules/module-iview/mycore.properties.iview` im Bildbetrachter („MCR-IView“) wird dieser Wert als Pixel angegeben.

```
MCR.Module-iview.cacheOrig=false|true
```

Hinweis: Beachten sie, dass dieser Cache-Eintrag das Repository stark vergrößern kann, wenn sie Bilder nicht als TIFF abspeichern. Nur wenn sie wirklich innerhalb der Bilder navigieren möchten, sollte dieser Cache angeschaltet werden.

5 Anmerkungen und Hinweise

5.1 Ergänzung der DocPortal-Beispieldaten

Mit Version 1.1 wurden die Beispieldaten für das DocPortal aus der Distribution des selbigen herausgelöst und in eine separaten CVS-Baum untergebracht. Dies hat den Vorteil, dass

- die Installation des DocPortals nicht mehr von Beispieldaten abhängig ist,
- man nach der Installation ein leeres, betriebsbereites System hat,
- die Distribution des Samples schlanker und der Download damit schneller ist,
- mehr Beispiele in einer extra-CVS-Distribution angeboten werden können und
- die Beispiele gezielt geladen und auch wieder entfernt werden können.

Die Beispieldaten stehen auf dem CVS-Server in Essen¹⁶ in einem extra CVS-Baum mit dem Namen `content` bereit. Dieser Enthält eine Sammlung einzelner Beispieldatengruppen. Nach dem checkout können die Gruppen je nach Wunsch einzeln installiert werden.¹⁷ Dabei spielt das jeweils mitgelieferte `build.xml`-Skript eine wichtige Rolle, hier sind alle Funktionen zur Arbeit mit dem Beispiel definiert.

Um neue Beispieldaten bereitzustellen gibt es zwei Wege: es wird eine Beispielgruppe mit Daten ergänzt oder es wird einen neue Beispielgruppe aufgebaut.

Ergänzungen in einer Beispielgruppe

Folgende Arbeiten sind erforderlich:

- Erzeugen der Metadaten für das Dokument (ggf. mit Daten für den Autor und/oder die Institution).
- Erzeugen des/der Derivate.
- Integration des Ladens und Entfernens im `build.xml`-Skript in den `target`-Abschnitten `load` und `remove`.

Hinzufügen einer neuen Beispielgruppe

Hier sind mehr Schritte erforderlich. Dabei ist immer darauf zu achten, dass die Beispielgruppe in sich vollständig ist, d. h. alle Autoren- und Institutionsdaten mitgeliefert werden. Da zum Laden der Daten das Update-Kommando verwendet wird, ist sichergestellt, dass es keine Doppelung im System gibt. Verwenden Sie möglichst die schon vorhandenen Autoren und Institutionen erneut.

¹⁶server.mycore.de

¹⁷siehe UserGuide

```
1 content
2     +-- mein_sample
3         +-- build.xml
4         +-- readme.txt
5         +-- document
6         +-- derivate-Verzeichnisse (optional)
7         +-- author (optional)
8         +-- institution (optional)
```

Abbildung 5.1: Mindeststruktur einer Beispielgruppe

- Checken Sie die content-Distribution aus.
- Erzeugen Sie ein neues Verzeichnis unter `content`, welches die Wurzel für Ihre Beispielgruppe sein soll.
- Übernehmen Sie die Verzeichnisstruktur von einem bestehenden Beispiel und füllen Sie diese mit Ihren Daten.
- Kopieren Sie das `build.xml`-Skript von einem bestehenden Beispiel und adaptieren Sie es. Es muss mindestens die targets `info`, `load` und `remove` beinhalten. Ggf. sind noch weitere targets zum Kopieren von Stylesheets usw. nötig, hier müssen Sie das `build.xml`-File ergänzen.
- Schreiben Sie ein kurzes ReadMe-File im ASCII-Format mit Installationshinweisen.
- Testen Sie das fertige Beispiel.
- Commiten Sie alles in den CVS-Server.

6 Anhang

6.1 Abbildungsverzeichnis

Abbildung 2.1: Grundstruktur des MyCoRe-Projektes.....	2
Abbildung 2.2: Die Klassen der Sessionverwaltung.....	7
Abbildung 2.3: Klassendiagramm des EventHandler-Modells.....	10
Abbildung 2.4: Auszug aus dem Metadaten-Objektes des Elternsatzes.....	11
Abbildung 2.5: Auszug aus dem Metadaten-Objektes des Kindsatzes.....	12
Abbildung 2.6: XML-Syntax eines Kind-Datensatzes als Query-Resultat.....	12
Abbildung 3.1: XML Syntax des Query-Resultates.....	15
Abbildung 3.2: Geschäftsprozesse der Benutzerverwaltung in MyCoRe.....	17
Abbildung 3.3: Zusammenhang der Java-Klassen.....	20
Abbildung 3.4: Klassendiagramm Common Servlets.....	22
Abbildung 3.5: Sequenzdiagramm Common Servlets.....	23
Abbildung 3.6: Ablaufdiagramm für MCRServlet.doGetPost().....	24
Abbildung 3.7: Ablaufdiagramm für MCRLoginServlet.doGetPost().....	25
Abbildung 3.8: XML Output des LoginServlets.....	25
Abbildung 3.9: Einbindung des Editors in eine Webseite.....	29
Abbildung 3.10: Rahmen der Formular-Definition.....	30
Abbildung 3.11: Definition eines einfachen Formulares.....	31
Abbildung 3.12: Definition des Root-Panels.....	31
Abbildung 3.13: Auslagern von Definitionen aus dem Root-Panel.....	32
Abbildung 3.14: Auslagern von Definitionen aus dem dem Editor-Formular.....	32
Abbildung 3.15: Die imports-Formular-Definition.....	33
Abbildung 3.16: Einfügen des XML-Main-Tag-Namen.....	33
Abbildung 3.17: Integration von einfachen XML-Tags.....	34
Abbildung 3.18: Integration von mehrfachen XML-Tags.....	35
Abbildung 3.19: Einbindung des Editors in eine Webseite.....	35
Abbildung 3.20: Codesequenz zum Aufruf des Editor Framework.....	35
Abbildung 3.21: Rahmen der Formular-Definition.....	37
Abbildung 3.22: Java-Code-Sequenz für den Zugriff.....	37
Abbildung 3.23: Rahmen der Formular-Definition mit name=value.....	38
Abbildung 3.24: Syntax des cell-Elements.....	39
Abbildung 3.25: Syntax von textfield und textarea.....	39

Abbildung 3.26: AutoFill-Werte in textfield und textarea.....	39
Abbildung 3.27: Syntax des password-Elements.....	41
Abbildung 3.28: Syntax für Checkboxes.....	42
Abbildung 3.29: Syntax einer Auswahlliste.....	42
Abbildung 3.30: Beispiel für Mehrfachauswahl.....	42
Abbildung 3.31: Syntax von Radio-Button und Checkbox-Listen.....	43
Abbildung 3.32: Syntax eines Textfeldes.....	44
Abbildung 3.33: Beschriftung mit I18N.....	44
Abbildung 3.34: Syntax eines Abstandhalters.....	45
Abbildung 3.35: Syntax eines Popup-Fensters.....	46
Abbildung 3.36: Syntax eines einfachen Buttons.....	47
Abbildung 3.37: Syntax des Submit-Buttons.....	48
Abbildung 3.38: Syntax des Cancel-Buttons.....	48
Abbildung 3.39: Beispiel Repeater.....	49
Abbildung 3.40: Syntax des FileUpload.....	49
Abbildung 3.41: Java-Code zum Lesen des FileUpload.....	50
Abbildung 3.42: Include Servlet-Request.....	51
Abbildung 3.43: Java-Code.....	52
Abbildung 3.44: Einfaches Beispiel für Eingabevalidierung.....	52
Abbildung 3.45: Fehlgeschlagene Eingabevalidierung.....	53
Abbildung 3.46: Beispiel für eine Reihenfolge von Regeln.....	53
Abbildung 3.47: Beispiel für Regeln in einer condition.....	54
Abbildung 3.48: Codebeispiel.....	55
Abbildung 3.49: Klassifikationsbrowser in einer MyCoRe-Anwendung.....	57
Abbildung 4.1: Grundübersicht des SimpleWorkflow.....	61
Abbildung 4.2: Ablaufschema im SimpleWorkflow.....	63
Abbildung 4.3: Beispiel-Web-Seite für den Aufruf eines Workflow.....	64
Abbildung 5.1: Mindeststruktur einer Beispielgruppe.....	68

6.2 Tabellenverzeichnis

Tabelle 2.1: Übersicht der benutzten externen Bibliotheken.....	5
Tabelle 3.1: Übersicht der MyCoRe-Backends.....	18
Tabelle 3.2: Felder mit Umlautnormalisierung im Search-Store.....	19
Tabelle 4.1: Privilegienliste für den SimpleWorkflow.....	62
Tabelle 4.2: Übersicht der SimpleWorkflow-Servlets.....	63