

Zur Sicherstellung der Datenqualität von Metadaten in digitalen Bibliotheken am Beispiel der Configuration Management Database (CMDB)

Eingereicht von

Gerold Hämmerle

Studienkennzahl: J151

Matrikelnummer: 9252001

Diplomarbeit

am Institut für Informationswirtschaft

WIRTSCHAFTSUNIVERSITÄT WIEN

Studienrichtung: Betriebswirtschaft

Begutachter: Priv. Doz. Dr. Michael Hahsler

Wien, 3. Juli 2007

WIRTSCHAFTSUNIVERSITÄT WIEN

DIPLOMARBEIT

Titel der Diplomarbeit:

Zur Sicherstellung der Datenqualität von Metadaten in digitalen Bibliotheken am Beispiel der Configuration Management Database (CMDB)

Verfasserin/Verfasser: Gerold Hämmerle

Matrikel-Nr.: 9252001

Studienrichtung: J151 Betriebswirtschaft

Beurteilerin/Beurteiler: Priv. Doz. Dr. Michael Hahsler

Ich versichere:

dass ich die Diplomarbeit selbstständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfe bedient habe.

dass ich dieses Diplomarbeitsthema bisher weder im In- noch im Ausland (einer Beurteilerin/ einem Beurteiler zur Begutachtung) in irgendeiner Form als Prüfungsarbeit vorgelegt habe.

dass diese Arbeit mit der vom Begutachter beurteilten Arbeit übereinstimmt.

Datum

Unterschrift

Inhaltsverzeichnis

1	Einführung	2
1.1	Motivation	2
1.1.1	Bedeutung der Qualität von Metadaten für die Qualität einer Digitalen Bibliothek	2
1.1.2	Einfluss auf die Entwurfs- bzw. Auswahlkriterien für digitale Bibliotheken	3
1.2	Aufgabenstellung, Aufbau der Arbeit und Methodik	4
1.3	Zuordnung zum Fachgebiet der Informationswirtschaft	4
2	Theoretischer Hintergrund	5
2.1	Definition von Begriffen im Zusammenhang mit digitale Bibliotheken	5
2.1.1	Digitale Bibliothek	5
2.1.2	Konservierung	7
2.1.3	Daten, Informationen und Wissen	8
2.1.4	Werk, Expression, Manifestation und Exemplar	9
2.1.5	Sammlungen digitaler Werke	10
2.1.6	Bibliotheksobjekte und digitale Objekte	11
2.1.7	Metadaten	12
2.1.8	Katalog	13
2.1.9	Abstract, Schlagwörter und Thesauri	13
2.1.10	Klassifikation	15
2.1.11	Repository	17
2.1.12	Archiv	18
2.2	Metadaten	18
2.2.1	Metadatenstandards	20

2.2.2	Deskriptive Metadaten	23
2.2.3	Strukturelle Metadaten	27
2.2.4	Administrative Metadaten	30
2.2.5	Eindeutiger Schlüssel – <i>Unique Identifier</i>	33
2.2.6	Information Retrieval und Metadaten	36
2.2.7	Löschung von Metadaten	38
2.3	Leistungsmessung und Leistungsindikatoren	38
2.3.1	<i>Precision</i> und <i>Recall</i>	39
2.3.2	Integrität	40
2.3.3	Kriterien aus dem Service Management	41
2.4	Sicherstellung der Qualität von Metadaten	43
2.4.1	Metadaten Mindeststandards	43
2.4.2	Generelle Rahmenbedingungen für den Erfassungsprozess	45
3	Fallbeispiel: Die <i>Configuration Management Database</i> als zentrale digitale Bibliothek für <i>IT Service Management</i> Prozesse	48
3.1	IT Service Management	49
3.2	ITIL und für das IT Service Management relevante Standards	50
3.2.1	IT Infrastructure Library (ITIL)	50
3.2.2	Standards des <i>British Standards Institute</i> (BSI)	52
3.2.3	Control Objectives for Information- and related Technology (CobIT)	53
3.2.4	Microsoft Operations Framework (MOF)	53
3.2.5	Enhanced Telecom Operations Map (eTOM)	54
3.2.6	Der <i>ISO 20000</i> Standard	54
3.2.7	Der Prozess <i>Configuration Management</i> im IT Service Management	56
3.3	Die Configuration Management Database (CMDB)	58
3.3.1	Betrachtung der CMDB als digitale Bibliothek	58
3.3.2	Umfang der CMDB	62
3.4	Benutzergruppen und Prozesse mit berechtigtem Interesse am Zugang zur CMDB	67
3.4.1	Incident Management	68
3.4.2	Problem Management	69
3.4.3	Change Management	70

3.4.4	Release Management	71
3.4.5	Service Level Management	72
3.4.6	Financial Management for IT Services	73
3.4.7	Capacity Management	73
3.4.8	Availability Management	75
3.4.9	IT Service Continuity (ITSCM)	76
3.5	Vorschläge für Attribute – Metadatenelemente	76
3.5.1	Attribute für Configuration Items inklusive deren Beziehungen untereinander	77
3.5.2	Attribute für Services	80
3.5.3	Attribute für Störungen	81
3.5.4	Attribute für Probleme und bekannte Fehler	83
3.5.5	Attribute für RfC und Change	84
3.5.6	Attribute für Dokumentation	86
3.5.7	Attribute für die Beschreibung eines Release	90
3.6	Rahmenbedingungen für die Erfassungsprozess	90
3.6.1	Bei einem neuen Release	92
3.6.2	Bei Veränderungen	93
3.6.3	Bei Störungen	93
3.6.4	Bei der Bearbeitung von Problemen	94
3.6.5	Datenqualität überprüfen, sicherstellen und verbessern	94
3.7	Umsetzung und Verwertung der Ergebnisse	95
4	Zusammenfassung	96
	Literaturverzeichnis	98

Für meine Eltern, Ursula und Philipp.

Zusammenfassung

Das Thema Datenqualität wird für den Gegenstandsbereich der digitalen Bibliotheken zunehmend wichtiger, da fehlerhafte oder nicht ausreichend vorhandene Metadaten die Auffindung von Bibliotheksobjekten erschweren oder gar unmöglich machen können. Insbesondere für nicht im Volltext erschließbare Bibliotheksobjekte wie etwa Bilder, Videoaufzeichnungen, Hardwarekomponenten oder Softwarepakete ist die Verfügbarkeit, Richtigkeit und Vollständigkeit der dazu abgelegten Metadaten wesentlich, damit diese Objekte unter den passenden Stichwörtern wieder aufgefunden werden können.

Im ersten Teil dieser Arbeit wird dazu der umfangreiche Gegenstandsbereich der digitalen Bibliothek anhand wichtiger Begriffe und Konzepte dargestellt, wobei der Begriff und die Organisation der Metadaten besonders ausführlich behandelt werden. Abgerundet wird dieser Teil durch die Darstellung der Möglichkeiten der Leistungsbeurteilung mittels geeigneter Leistungsindikatoren aus dem Information Retrieval und dem IT Service Management.

Der zweite Teil dieser Arbeit widmet sich einer Fallstudie: Die *Configuration Management Database* (CMDB) als zentrales Informationssystem für das IT Service Management übernimmt die Speicherung und die Verwaltung von Metadaten und digitalen Bibliotheksobjekten wie beispielsweise Datenbanken oder Dokumentationen, welche für das IT Service Management von Relevanz sind. Basierend auf Erkenntnissen im Gegenstandsbereich der digitalen Bibliothek werden zusätzliche Vorschläge für Metadatenelemente und deren Strukturierung für die CMDB gemacht.

Zusammenfassend lässt sich festhalten, dass die Betrachtung der *Configuration Management Database* aus der Sicht der digitalen Bibliothek vor allem für Dokumentationen von Komponenten und die Verwaltung der Softwarekomponenten selbst im Rahmen der CMDB viele zusätzliche Metadatenelemente liefern. Als sekundäres Ergebnis soll dokumentiert sein, dass der Gegenstandsbereich der digitalen Bibliothek auch von der Sichtweise aus dem IT Service Management profitieren kann, indem die digitale Bibliothek als Bündel von IT Services betrachtet wird.

Schlagwörter: Digitale Bibliothek, Metadaten, Datenqualität, Konfigurationsmanagement, Configuration Management Database, CMDB, IT Infrastructure Library, ITIL, ISO 20000.

Kapitel 1

Einführung

1.1 Motivation

1.1.1 Bedeutung der Qualität von Metadaten für die Qualität einer Digitalen Bibliothek

Dem Thema Datenqualität wird im Zusammenhang mit Digitalen Bibliotheken vermehrt erhöhte Aufmerksamkeit geschenkt: exemplarisch in Analysen zur Qualitätssicherung bei der Erstellung der Metadaten (Barton et al., 2003), in Studien (Shreeves et al., 2005) oder bei Konferenzen und Workshops wie bei Wilson (2004) oder Ruddy (2006). Begründet wird dieser Trend generell mit einer potenziellen Behinderung des Zugriffs auf Bibliotheksobjekte in digitalen Bibliotheken aufgrund verminderter Datenqualität von Metadaten (Beall, 2005). Dabei kann durch fehlerhafte oder fehlende Metadaten die Suche nach den gewünschten – oft auch digital abgelegten – Inhalten oder anderen Metadaten beeinträchtigt werden.

Fehler im Zusammenhang mit Metadaten wirken sich besonders bei Bilddatenbanken aus, da diese primär über Informationen aus Metadaten erschlossen werden (Beall, 2005, 10). Analog kann daraus geschlossen werden, dass sich Metadatenfehler auch auf Bibliotheks- und Archivdatenbanken für physische und nicht-textliche Gegenstände erheblich auswirken, da im besten Fall digitalisierte Fotos bzw. Videosequenzen der Objekte wie Stand- und Bewegtbildträger, Tonträger, Gemälde, Fahnen, Karteikästen, Hardwarekomponenten, Softwarepakete, Informationen zu IT Services und dergleichen abgelegt werden und häufig keine weiteren Informationen in maschinell durchsuchbarer Textform vorliegen.

Am Beispiel der *Configuration Management Database*¹ (CMDB) als zentrales Informationssystem für IT Service Management Prozesse lässt sich anschaulich zeigen, welchen Einfluss die Qualität von Daten über die erfassten *Configuration Items* (Hardwarekomponenten und

¹Eine detaillierte Beschreibung der Configuration Management Database findet sich im Kapitel „Fallbeispiel“ im Abschnitt „Die Configuration Management Database (CMDB)“.

Software), Dokumente und IT Services hat: IT Services und IT Assets haben in der Regel einen deutlich höheren Wert für den Geschäftsprozess welcher die IT Services nutzt, als die reinen Anschaffungswerte bzw. die Herstellungskosten der zur Serviceerbringung erforderlichen Komponenten (Assirati, 2001b, 125). Korrekte Informationen über diese Configuration Items (CIs) und ihre Dokumentation gelten deshalb als kritischer Erfolgsfaktor (Assirati, 2001b, 79). Störungen können rascher und effektiver behoben werden, wenn Informationen zu exaktem Standort, Softwareversion und die Beziehungen zu anderen beeinflussenden oder beeinflussten Komponenten zur Verfügung stehen. Somit können beispielsweise die Ausfallzeiten eines zentralen Reisebuchungssystems für eine Reisebüro-Kette auf ein Minimum beschränkt werden und durch die erhöhte *Up-time* des Reisebuchungssystems die Produktivität und damit den Tagesumsatz der einzelnen Mitarbeitenden in Beratungs- und Vertriebsfunktionen erhöht werden.

Es ist daher durchaus von Bedeutung, die für die Qualität der Metadaten entscheidenden Entwurfs- bzw. Auswahlkriterien sowie die qualitätsbestimmenden Prozessparameter zu kennen und diese gezielt beeinflussen zu können.

1.1.2 Einfluss auf die Entwurfs- bzw. Auswahlkriterien für digitale Bibliotheken

Im Vorfeld einer Systementscheidung über ein digitales Bibliothekssystem dürfen im Pflichtenheft neben vielen technischen und die Anwendung selbst betreffenden Punkte nicht nur die Informationsbedürfnisse der Benutzenden (welche Datenfelder abgefragt werden können) und der Usability (wie die Benutzerführung konzipiert sein soll) berücksichtigt werden. In Hinblick auf einen qualitativ hochwertigen Datenbestand müssen zusätzlich folgende Themen geklärt werden:

- **Mindestanforderungen an die Attribute** der Metadatenobjekte zur Beschreibung der digitalen Objekte müssen festgelegt werden, denn nur qualitativ hochwertige Metadaten können die funktionalen Anforderungen an ein Bibliothekssystem unterstützen (Marieke et al., 2004).
- **Richtlinien für das Design des Erfassungsprozesses** müssen erarbeitet werden, denn der Prozess der Erstellung der Metadaten ist der Schlüssel für ein erfolgreiches Archiv (Marieke et al., 2004).
- **Richtlinien für Audits des Datenbestands** hinsichtlich der Qualität müssen erarbeitet werden, da die Entwicklung von qualitativ hochwertigen Metadaten eine der wichtigsten Aufgaben für LIS²-Personal darstellt (Robertson, 2005, 295).

Es ist somit evident, dass einerseits der Art und der Struktur der Metadaten, aber andererseits auch den die Metadaten bereitstellenden oder verändernden Prozessen maßgebliche Bedeutung zukommen muss.

²LIS steht für *Library and Information Science*.

1.2 Aufgabenstellung, Aufbau der Arbeit und Methodik

Aufgabe dieser Diplomarbeit ist es,

1. in einem **theoretischen Teil** neben den erforderlichen **begrifflichen Definitionen und Konzepten**
2. verbreitete **Metadatenstandards** und Beispiele für Metadaten in digitalen Bibliotheken betreffen zusammenzustellen und
3. nachfolgend auf **qualitätsbezogene Leistungsindikatoren** einzugehen sowie
4. im **Praxisteil** anhand eines Fallbeispiels die Entwurfs- bzw. Auswahlkriterien zu konkretisieren. Aufgabe ist die Erstellung von **Vorschlägen für Metadatenelemente** für die Auswahl der Attribute und von Hinweisen in der Form von *Do's & Dont's* und für die Gestaltung des Erfassungsprozesses, wobei die Unabhängigkeit von der technischen Durchführung – manuelle oder automatische Erfassung – gegeben sein muss. Als Beispiel wird er Aufbau einer *Configuration Management Database* (CMDB) für die Verwendung in ITIL- bzw. ISO-20000-Prozessen gewählt.

Nicht-Ziele dieser Arbeit sind Betrachtungen der Kostenaspekte, rechtlichen Aspekte oder die Sicherstellung der Vertraulichkeit.

Wesentliche und primär verwendete Methode ist die Literaturrecherche sowie die Erarbeitung einer Case Study.

1.3 Zuordnung zum Fachgebiet der Informationswirtschaft

Die in dieser Arbeit besprochenen Systeme digitaler Bibliotheken sollen die Erfüllung der unterschiedlichen Informationsbedürfnisse von Personen innerhalb und gegebenenfalls auch außerhalb einer Organisation decken. Erklärtes Ziel ist es, Informationen effizient zu verwalten, zu organisieren sowie deren zielgerichteten Verwertung unter Berücksichtigung größtmöglicher Wirtschaftlichkeit zu ermöglichen. Nach (Janko, 1998, 2) und (Fischer, 1999, III) erfüllt die Informationswirtschaft als ein Zweig der Wirtschaftswissenschaften diese Aufgaben.

Kapitel 2

Theoretischer Hintergrund

In diesem Kapitel werden wesentliche Begriffe und Konzepte im Zusammenhang mit einer digitalen Bibliothek definiert um danach als Schwerpunkt auf den Begriff der Metadaten einzugehen. Weiters wird auf den Fehlerbegriff und die Hauptfehlertypen in diesem Zusammenhang eingegangen. Dabei spielen auch die Definition von Leistungsindikatoren und die Leistungsmessung zur Erhebung des jeweils aktuellen Ist-Zustands eine wichtige Rolle. Abschließend werden allgemeine Hinweise für den Entwurf bzw. die Auswahl von Bibliotheks- und Archivsystemen die Metadatenqualität betreffend formuliert.

2.1 Definition von Begriffen im Zusammenhang mit digitale Bibliotheken

Im folgenden Abschnitt werden Schlüsselbegriffe und -konzepte der Domäne „digitale Bibliothek“ erläuternd dargestellt und Abgrenzungen zu ähnlichen Begriffen vorgenommen.

2.1.1 Digitale Bibliothek

Der Begriff *digitale Bibliothek* kann in einer ersten Annäherung als aktiv verwaltete Sammlung von Informationen bezeichnet werden (Arms, 2000). Dabei wird Gewicht auf das Management der Informationssammlung gelegt und somit eine Abgrenzung zu beliebigen Datenströmen oder unstrukturiert abgelegten Informationen formuliert. Auch stellt für Arms der Umfang ein Kriterium dar: eine Datenbank mit Finanzdaten einer Organisation stellt noch keine digitale Bibliothek dar, wobei jedoch eine systematisch organisierte Sammlung von Finanzdaten vieler Organisationen durchaus eine digitale Bibliothek bilden können. Zusammenfassend stellt Arms für digitale Bibliotheken wie folgt fest (Arms, 2000): „The unifying theme is that information is organized on computers and available over a network, with procedures to select the material in the collections, to organize it, to make it available to users, and to archive it“. Hinsichtlich der Nichterfordernis eines

physikalischen Bezugspunktes gleich weit gefasst aber auf Werke mit informierendem Charakter beschränkte Definition vertritt (Wagner-Döbler, 2000, 284), indem er mit Hinblick auf die wissenschaftliche Forschung Objekte mit unterhaltendem Charakter in der Regel ausschließt.

Jedenfalls muss eine digitale Bibliothek nicht nur bibliografische Nachweise von Publikationen in digitaler Form liefern sondern diese Publikationen auch digital im Volltext zur Verfügung stellen. (Umstätter, 2000, 297)

Da in den vorstehend erwähnten Definitionen die Aspekte des spezialisierten Personals, der zielgruppenorientierten Ausrichtung und der dauerhaften Erhaltung der Bibliotheksobjekte nur indirekt über die genannten Prozeduren - und damit zu wenig betont - berücksichtigt werden, wird hier die deutlich detailliertere Definition der *Digital Library Federation* (DLF) bevorzugt (Waters, 1998):

„Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.“

Diese Definition umfasst neben den Bibliotheksobjekten und Dienstleistungen auch die Organisation inklusive intellektueller Ressourcen und IT-spezifischem Know-how sowie vernetzte rechnergestützte Systeme zur Speicherung der Sammlungen und zum Zugriff darauf. Obwohl die Ausrichtung von Bibliotheken auf Dauerhaftigkeit der zur Aufbewahrung übernommenen oder selbst erstellten Objekte eine zentrale Stellung inne hat – und dies für digitale Bibliotheken zumindest der impliziten Wahrnehmung entspricht (Morris, 1998) – argumentiert Waters (1998), dass dieses Ziel bei digitalen Bibliotheken besonders schwer zu erreichen sei und dass bei allen Bemühungen Integrität zu erhalten, Integrität alleine noch keine hinreichende Bedingung für die Dauerhaftigkeit darstellt.

Abgrenzung zum Begriff „Elektronische Bibliothek“

Nach Tennant (1999) werden in der elektronische Bibliothek alle Formen von elektronischen Materialien einschließlich aller analogen Formate (VHS-Video, Audio und dergleichen) berücksichtigt. Der Begriff Elektronische Bibliothek schließt somit auch den der digitalen Bibliothek zur Gänze mit ein.

Abgrenzung zum Begriff „Virtuelle Bibliothek“

Virtuelle Bibliotheken existieren nicht als *eine* physische Bibliothek im klassischen Sinn. Tennant (1999) zufolge benutzen virtuelle Bibliotheken Computer und Netzwerke um Materialien mehrerer unterschiedlicher Bibliotheken zusammenzufassen und über einen zentralen Service der virtuellen Bibliothek abrufbar zu machen. Ein aktuelles Beispiel dafür

stellt die „Europäische Digitale Bibliothek“ dar, welche als kollektives Gedächtnis Europas geschaffen werden soll (Borowski, 2007, 41-42): über ein zentrales EU-Portal sowie nationale Portale sollen einige tausend Kultureinrichtungen wie beispielsweise Bibliotheken, Museen, Mediatheken, Archive oder Theater vernetzt werden und der Zugriff auf Informationen via multilingualer Suche im gesamten Datenbestand vereinheitlicht sowie (Anmerkung: zumindest in quantitativer Hinsicht) umfassender gestaltet und mit ansprechender Präsentation versehen werden.

2.1.2 Konservierung

Um Bibliotheksobjekte dauerhaft zu erhalten sind unabhängig von der Art der Manifestation Aktivitäten zur Konservierung erforderlich. Dabei herrschen zwei sich ergänzende, jedoch technisch und organisatorisch verschiedene Konzepte vor (Wagner-Döbler, 2000, 286-287):

- Das moderne Extrem ist die Positionierung einer Bibliothek als Dienstleistungsbetrieb, wobei die Kommerzialisierung des Informationswesens in Kauf genommen wird und der Informationscharakter von Publikationen in den Vordergrund rückt. Dieser Ansatz forciert die Digitalisierung (scannen) von Publikationen um die Speicherung, das Management und die Weitergabe möglichst effizient zu gestalten.
- Der Gegenpol dazu fordert die museale Pflege der Bibliotheksobjekte, da nicht nur die Information (beispielsweise eines Buches) von Relevanz ist, sondern auch die optischen Qualitäten (Einband aus Leder oder Leinen, Titel in Gold- oder Silberlettern, Papierglanz, usw.) und haptischen Eigenschaften bewahrt werden müssen. Aus diesem Ansatz heraus ist die Entsäuerung der Publikationen die vorzuziehende Erhaltungsmaßnahme und weniger deren Digitalisierung.

Beiden Konzepten gemeinsam ist die Katalogisierung. Dabei werden die Kataloge heute immer in digitaler Form gespeichert und verwaltet, womit jede Bibliothek – auch die der musealen Pflege verbundene – sich mit der Dauerhaftigkeit digitaler Information im Zusammenhang mit Computersystemen beschäftigen muss.

Wesentlicher Aspekt der dauerhaften Erhaltung digital abgelegter Informationen ist die regelmäßige und korrekte Auffrischung der binären Daten. So sieht (Green et al., 1999, 1) als zentrale Erhaltungsmaßnahme das

- *Refreshing*: Dabei werden Dateien von einem Speichermedium auf ein anderes übertragen, ohne das Dateiformat, die interne Struktur der Datensätze in der Datei zu ändern. Diese Vorgangsweise ist solange zur Erhaltung von Datenbeständen angebracht, als es Software gibt, welche diese Dateien lesen und bearbeiten können. Refreshing kann daher nicht als generelle Lösung zur Langzeiterhaltung von Daten dienen.

Morris (1998) macht dabei auf erhebliche Hindernisse aufmerksam, wonach insbesondere der Zugriff auf veraltete und überholte Software sowie Datenformate in digitalen Bibliotheken zu meistern ist. Er sieht für die Erhaltung dieser digitalen Informationen zwei radikale Lösungen: Migration und Emulation. Beide Strategien gehen das Risiko ein, dass Informationen verloren gehen können. Er benennt diese Strategien und Risiken wie folgt (Morris, 1998):

- Wird die *Migrationsstrategie* verfolgt, werden die digitalen Materialien periodisch von einer Hard- und Softwarekonfiguration der einen Generation auf eine andere der nachfolgenden Generationen übertragen. Diese Strategie kann bei spezifischen proprietären Formaten und Systemen nicht möglich sein, denn es kann zwar meist der Inhalt migriert werden, die eingebetteten Funktionalitäten (etwa Makro-Befehle) sowie der Kontext oder die inneren Beziehungen gehen aber verloren. Migration schließt immer auch ein Refreshing mit ein (Green et al., 1999, 1).
- Im Rahmen der *Strategie der Emulation* werden zusätzlich zu den digitalen Materialien Kopien der ursprünglich verwendeten Software (Betriebssysteme, Hilfsprogramme und Anwendungen) und Hardware archiviert. Das Risiko der (teilweisen) Nichtverfügbarkeit dabei besteht, wenn einzelne Komponenten des Systems fehlerhaft werden.

Die Auswahl der jeweiligen Konservierungsmaßnahme ist im Einzelfall sorgfältig abzuwägen.

2.1.3 Daten, Informationen und Wissen

Um die soeben verwendeten Begriffe Daten und Information voneinander abzugrenzen, werden in dieser Arbeit die nachstehenden Definitionen übernommen.

Probst und Knaese beschreiben **Daten** als kontextunabhängig vorliegende Zahlen und Zeichen (Probst und Knaese, 1998, 25), beispielsweise die in binär kodierter Form gespeicherten Werktitel, Autorangaben, Jahresangaben, Abstracts und ähnliches. Nach (Debenham, 1998, 15) müssen Daten (*data things*) zusätzlich fundamental und unteilbar sein. Exemplarisch können der Nachname des ersten Autors oder die Jahreszahl der Veröffentlichung als Daten betrachtet werden.

In weiterer Folge werden Daten zu **Informationen** verdichtet, indem sie in einen sinnvollen Kontext gebracht werden (Probst und Knaese, 1998, 25). Debenham klassifiziert einen Kontext zwischen Daten genau dann als Information, wenn dieser Kontext *nicht durch Anwendung einer Regel* auf Daten ermittelt werden kann und bezeichnet diese als eine *implizite Assoziation* (Debenham, 1998, 15). Ein einfaches Beispiel stellt die Suche nach einem Artikel für eine wissenschaftliche Arbeit dar, wenn folgende Information vorliegt:

Information bestehend aus den Daten:

- Die *Autorinnen* Jane Barton, Sarah Currier und Jessie M. N. Hey

- verfassten einen Artikel mit dem *Titel* „Building Quality Assurance into Metadata Creation: an Analysis based on the Learning Objects and e-Prints Communities of Practice“
- und ist im World Wide Web abrufbar unter URL http://www.siderean.com/dc2003/201_paper60.pdf.

Aufgrund der enthaltenen Information – die URL zum Artikel – kann nun der gesamte Artikel abgerufen und diesen für die weitere Arbeit berücksichtigt werden. Weiters ist die Information enthalten, dass die Autorinnen des Artikels Barton, Currier und Hey sind.

In weiterer Folge stellt sich die Frage nach dem daraus ableitbaren **Wissen** über diesen Artikel. Debenham (Debenham, 1998, 15) klassifiziert einen Kontext zwischen Daten oder Informationen dann als Wissen (*knowledge thing*), wenn dieser Kontext *durch Anwendung einer Regel* auf Daten ermittelt werden kann. Er bezeichnet diese auch als eine *explizite Assoziation*. Es lässt sich somit – unter anderem – folgendes

Wissen aus der Information gewinnen:

- Der Artikel wurde von drei Autorinnen verfasst. Die vereinfachte Regel dazu lautet: „Zähle die Nachnamen“.
- Es handelt sich um eine Dokumentendatei im Format PDF. Die vereinfachte Regel dazu lautet: „Suche in der URL nach dem Dateinamen (hier: *201_paper60.pdf* und extrahiere die Dateiergänzung (hier: *.pdf*)“.

Ergänzend sei der Standpunkt festgehalten, dass Wissen als jene mittels Regeln verknüpften Daten und Informationen gesehen werden kann, die in einer bestimmten Situation sinnvoll verwendet werden kann und damit einen Nutzen bringt (Heinsohn und Socher-Ambrosius, 1999, 1). In der Literatur wird der Begriff oft noch weiter gefasst. Dabei kann Wissen zur effizienten Problemlösung verwendet werden und schließt „neben theoretischen Erkenntnissen auch Erwartungen, Erfahrungen, Wahrnehmungen, praktische Alltagsregeln und Handlungsanweisungen ein“ (Probst und Knaese, 1998, 25).

2.1.4 Werk, Expression, Manifestation und Exemplar

Diese vier Begriffe bezeichnen vier Entitäten zur Beschreibung von Produkten intellektueller und künstlerischer Arbeit und wurden im Rahmen der Empfehlung *Functional Requirements for Bibliographic Records (FRBR)* – eine detailliertere Erläuterung des Modells findet sich weiter unten im Unterabschnitt „Deskriptive Metadaten“ – publiziert. Es handelt sich dabei um ein logisches Denkmodell für bibliografische Beschreibungen in Form eines relationalen Modells mit abstrakten Entitäten (Block et al., 2005, 5). Im folgenden sind die vier zentralen Entitäten kurz umrissen:

- Ein **Werk** (*work*) ist jene abstrakte Entität, von der die Genealogie eines Werkes ausgeht (Block et al., 2005, 8). Arms (2000) führt als Beispiele die Illias, Beethoven’s Fünfte Symphonie, oder das Betriebssystem Unix an.

- Ein Werk ist durch eine oder mehrere **Expressionen** (*expressions*) tatsächlich realisiert. Nach Block et al. (2005) lassen sich im Prinzip nur wenige Informationen aus den aktuellen Katalogen der Entität Expression eindeutig zuordnen. Jedenfalls gehören Angaben zur Veröffentlichung und Sprachangaben dazu. Beispiele für Expressionen sind: die mündliche Überlieferung eines Werks, die Textform oder die musikalische Aufführung. Das Werk einer Anwendungssoftware kann in den unterschiedlichen Expressionen *Sourcecode* oder *Ausführbares Programm* vorliegen (Arms, 2000).
- Eine einzige Expression kann in unterschiedlichen **Manifestationen** (*manifestations*) bereitgestellt werden (Arms, 2000). So können beispielsweise für die deutschsprachige Version eines Anwendungsprogramms die Manifestationen als CD, DVD oder als Download verfügbar sein. Angaben zu einer Manifestation sind hauptsächlich auf die physische Erscheinungsform und die Produktion eines Objektes bezogen, wie die Auflage, Herstellung, Veröffentlichung oder der Vertrieb (Block et al., 2005).
- Ein **Exemplar** (*item*) stellt ein Stück aus einer Auflage einer Manifestation oder eine einzelne Kopie einer Datei beziehungsweise einer Anwendung dar (Arms, 2000). In der Praxis kann dies etwa eine einzelne Benutzerlizenz einer Anwendungssoftware sein.

Abschließend sei festgehalten, dass die vier Begriffe in der tatsächlichen Anwendung oft schwer von einander abzugrenzen sind und sich sowohl für die Erfassung als auch die Übernahme bestehender Kataloge durchaus Problemfelder ergeben. (Block et al., 2005, 8-9)

Weiterführende Informationen zum FRBR-Modell finden sich weiter hinten im Unterabschnitt „Deskriptive Metadaten“.

2.1.5 Sammlungen digitaler Werke

Üblicherweise unterscheiden sich Bibliotheken durch den Schwerpunkt ihrer Sammlungen. Beispiele dafür sind aus dem Bereich der klassischen universitären Bibliotheken die unterschiedlichen Ausrichtungen der Hauptbibliothek der Wirtschaftsuniversität Wien (Sozial- und Wirtschaftswissenschaften) im Vergleich mit den zahlreichen Institutsbibliotheken mit ihren Sammlungen fachspezifischer Bibliotheksobjekte.

Waters (1998) zeigt auf, dass das Profil digitaler Bibliotheken nicht nur die Digitalisierung von Materialien oder gar die digitale Speicherung an sich sein kann, sondern auch ein eigenständiges Profil in thematischer Hinsicht entwickeln müssen. Dabei muss die strategische Frage geklärt werden, wie digitale Formen und klassische Formen integriert werden. Allerdings ist zuerst festzulegen, ob diese Integration erforderlich oder gewünscht ist.

2.1.6 Bibliotheksobjekte und digitale Objekte

Aus der Sicht der Benutzenden sind nach Arms (2000) die einzelnen **Bibliotheksobjekte** (*library objects*) von besonderer Relevanz, da diese von den Benutzenden als eine Einheit betrachtet werden. Dies kann eine Monografie, ein Artikel, ein Interview, eine Anwendungssoftware, eine Datenbank oder eine Urkunde sein.

Von Interesse ist die weite Bandbreite von Dokumenttypen von Monographie über einzelne Artikel bis hin zum Film. Alleine für wissenschaftliche Printveröffentlichungen listet Stock (2000):

- Monographie, etwa ein Lehrbuch oder eine umfassende Abhandlung,
- eine Monographie als Herausgeber mit Beiträgen mehrerer Autorinnen und Autoren,
- Artikel in einer referierten Zeitschrift,
- Artikel in der Hauszeitschrift des eigenen Instituts,
- Habilitationsschrift,
- Dissertation,
- Rezension,
- Patent,
- Notiz in einer Zeitschrift („letter to the editor“),
- Artikel in einer Tageszeitung,
- wissenschaftlicher Film.

Bei genauerer Analyse bestehen die Bibliotheksobjekte selbst einerseits **aus einem oder mehreren digitalen Objekten** (*digital objects*) wie digitalisiertem Text, Programme, digitalisiertem Foto, anderem digitalisiertem Material und andererseits aus **strukturellen Metadaten** (*structural metadata*) welche die interne Struktur eines Bibliotheksobjekts wie Datenformate und Relationen zwischen den digitalen Objekten dokumentieren. Bibliotheksobjekte können also durchaus **aus mehreren miteinander in Beziehung stehenden digitalen Objekten** bestehen. Beispielsweise besteht ein per Video aufgezeichnetes Interview aus den digitalen Objekten *Tonspur* und *Bildsignal*. Die Verknüpfung (Synchronisation) erfolgt dabei durch strukturelle Metadaten. Die Manifestationen digitaler Objekte können nach Arms (Arms, 2000) entsprechend ihrer Beziehung zueinander wie folgt kategorisiert werden:

- Als **komplementäre Manifestationen**, wie die beiden sich ergänzenden Objekte Audio und Video aus vorgenanntem Interview.

- Bei **alternativen Manifestationen** wird den Benutzenden die Wahlmöglichkeit zwischen unterschiedlichen Manifestationen gegeben, wie beispielsweise die Auswahl zwischen einem Bild in Bildschirmauflösung (*Low-Resolution*) und demselben Bild in höherer Auflösung für den Ausdruck (*High-Resolution*). Arms (2000) führt als komplexeres Beispiel das Bibliotheksobjekt „Wetter am San Francisco Airport“ an, wobei die Benutzenden zwischen einer laufend aktualisierten Fotografie oder einer textorientierten Darstellung von Temperatur, Luftfeuchtigkeit und Windgeschwindigkeit wählen können.
- Digitale Objekte können als **beschreibende Manifestation** (Surrogate) für andere digitale Objekte dienen. In diesem Fall handelt es sich um Metadaten.

Arms (2000) weist aber auch darauf hin, dass diese Kategorienbildung nicht eindeutig und ausschließlich ist: wie ist beispielsweise ein verkleinertes Bild (*Thumbnail*) zu klassifizieren? Als alternative Darstellung oder als Metadaten über das Originalbild? Darauf kann ohne zusätzliche Informationen keine eindeutige Antwort gegeben werden.

Eine besondere Herausforderung stellen **Bibliotheksobjekte** dar, welche teilweise oder zur Gänze **dynamischen Ursprungs** sind und **dynamisches Verhalten** zeigen. Dies können nach Arms (2000) beispielsweise Programme wie Java Applets, Simulationen oder Videospiele sein und haben gemeinsam, dass die Benutzenden bei jedem Zugriff auf das Bibliotheksobjekt andere Resultate erhalten.

Arms (2000) hebt in diesem Zusammenhang auch **Datenbanken** hervor, da diese aus einer Vielzahl von Datensätzen bestehend bei jedem Zugriff unterschiedliche Abfragemöglichkeiten erlauben und der Inhalt des Bibliotheksobjekts sich für die Benutzenden von Abfrage zu Abfrage verändert.

2.1.7 Metadaten

Metadaten sind Daten, welche andere Daten beschreiben (Surrogate). In digitalen Bibliotheken dienen sie der Beschreibung von Objekten

- auf unterschiedlicher konzeptioneller Ebene ausgehend vom Werk über Expressionen und Manifestationen bis hin zum einzelnen Exemplaren,
- in unterschiedlichen Klassen wie die Entitäten Bibliotheksobjekt, Person oder Organisation,
- in unterschiedlichen Ausprägungen wie elektronisch gespeicherte Objekte und Bücher, Archivmaterial, Kunstgegenstände oder technische Einrichtungen als Objekte der realen Welt sowie
- der Beschreibung der Beziehungen der Objekte untereinander

Arms (2000) weist darauf hin, dass der Begriff Metadaten von der Betrachtungsweise, vom Standpunkt abhängt: So sind beispielsweise *Abstracts* gewöhnlich als Metadaten über die jeweiligen Arbeiten zu sehen. Andererseits sind für eine Abstract-Datenbank die *Abstracts* jedoch die eigentlichen Daten.

Eine detailliertere Betrachtung von Metadaten findet sich im selben Kapitel im Abschnitt „Metadaten“.

2.1.8 Katalog

Ein Katalog (*catalog*) hat im Zusammenhang mit dem Bibliothekswesen eine klare Bedeutung als Sammlung von bibliografischen Informationen in Form von zusammenfassenden Datensätzen, welche nach strengen Regeln erstellt und gewartet werden. Er stellt demnach eine besonders stark strukturierte Sammlung von Metadaten über eine breite Palette an Bibliotheksobjekten in unterschiedlichsten Medien und Formaten dar. (Arms, 2000)

Es werden zwei Arten von Katalogen unterschieden (Universität des Saarlandes (Hrsg.), 2007):

- Ein **alphabetischer Katalog** erschließt Bibliotheksobjekte nach Autor, Titel, Jahr und vergleichbaren Attributen – nicht jedoch nach dem tatsächlichen Inhalt. Die Katalogisierung im deutschsprachigen Raum erfolgt dabei nach den *Regeln zur Alphabetischen Katalogisierung* (RAK).
- Ein **Sachkatalog** erschließt Bibliotheksobjekte inhaltlich nach Themen, die im Wortlaut nicht zwingend im Titel oder anderen beschreibenden Attributen vorkommen, und listet diese geordnet auf.

Nicht alle Informationen in Katalogen entstammen den Bibliotheksobjekten selbst: Informationen über Autoren oder die Herkunft eines Artefakts werden anderweitig ermittelt und erfasst. (Arms, 2000)

2.1.9 Abstract, Schlagwörter und Thesauri

Die Bereitstellung von **Abstracts** bzw. Schlagwörter (*index terms*) stellt insbesondere für Artikel in Periodika ein geeignetes Mittel dar, um eine effiziente Recherche möglich zu machen (Arms, 2000). Dabei werden Metadaten in der Form von stark verkürzte Aufzeichnungen über den Inhalt eines Artikels erstellt und Schlagwörter zur Beschreibung des Bibliotheksobjekts vergeben. Die Herstellung dieser kann auf mehrere Arten erfolgen (Arms, 2000):

- *Manuell durch Experten*: Experten lesen jeden Artikel einer Reihe von Periodika, vergeben Schlagwörter und erstellen Abstracts.

- Beim *Maschinellen Indexieren* werden Indexterme automatisch einem Dokument zugeordnet indem Stoppwörter entfernt werden, die Grundformen der Wörter erzeugt werden, aus Wörter regelbasiert normierte Indexterme generiert werden (Stemming) sowie Phrasen generiert und gewichtet werden. (Lepsky, 2005, 32-34)
- *Vom Verlag bereitgestellte Abstracts*: Der Verlag liefert den vom Autor des Artikels selbst verfassten oder einen vom Verlag verfassten und editierten Abstract zu jedem Artikel.

Für die Verwendung von **Schlagwörtern** werden im Zusammenhang mit Bibliotheken vier Konzepte verwendet (Umlauf, 2007):

- *Stichwörter* werden aus dem Dokument unverändert übernommen und erfasst. Das autorenspezifische Wording kann leicht zu Missverständnissen und Schwierigkeiten bei der Informationsrecherche führen und hat zur Folge, dass verschiedene Schreibweisen und Wortformen (Einzahl/Mehrzahl) getrennt gesucht werden müssen (Köhler, 2005, 21).
- *Freie Schlagwörter* oder *freie Termini* werden im Anlassfall gewählt und erfasst. Die individuelle Wahl der Schlagwörter mündet jedoch ebenfalls in einer uneinheitlichen Verschlagwortung und erschwert ein effektives Information Retrieval.
- *Gebundene Schlagwörter* oder *kontrollierte Termini* werden dem Kontext entnommen und mittels einer Liste vereinbarter und kontrollierter Schlagwörter (*controlled vocabulary*) wie beispielsweise der Schlagwortnormdatei (SWD) im deutschsprachigen Raum normalisiert (Lemmatisierung) (Österreichische Zentralbibliothek für Physik (Hrsg.), 2007). Dort wo verschiedene Begriffe ein und dasselbe Konzept beschreiben und als Synonyme verwendet werden können, wird ein zulässiges „gebundenes Schlagwort“ unter Anwendung von Regeln angesetzt. Dieser aufwändigere und oft IT-unterstützte Erstellungsprozess wird beispielsweise in der *Library of Congress* bei den *subject headings* eingesetzt. (Arms, 2000)
- *Deskriptoren* werden nach einer Inhaltsanalyse des Dokuments einem **Thesaurus** entnommen und präsentieren den Inhalt in retrievalfähiger Form. In Ergänzung zu einem Stichwortkatalog liefert ein Thesaurus zusätzlich auch die gleichrangigen und hierarchischen Relationen zu anderen Deskriptoren. Eine Anwendung für thesaurusbasierte Deskriptoren findet sich in der *National Library of Medicine* wo *MeSH headings* (*Medical Subject Heading*) als Fachthesaurus zur Erstellung von Deskriptoren für medizinische Artikel herangezogen werden und in der Datenbank *MEDLINE* erfasst werden.

Die Struktur von Schlagwortkatalogen ist einerseits weniger formal als die eines alphabetischen Katalogs, andererseits aber deutlich stärker ausgeprägt als im Fall eines Abstracts.

Abstracts sowie Stich- und Schlagwörter stellen den Kern der inhaltlichen Erschließung von Bibliotheksobjekten dar. Die regelkonforme Nutzung der Möglichkeiten der Beschlagwortung entscheidet im Fall einer Suchanfrage über die Anzahl und Qualität der gefundenen Titel.

2.1.10 Klassifikation

Der Begriff der Klassifikation hat im Bibliothekswesen eine dreifache Bedeutung (Manecke, 2004, 127):

1. Eine Klassifikation benennt ein **Klassifikationssystem** an sich. Dies ist das Resultat des im folgenden Punkt genannten Klassenbildungsprozesses.
2. Er beschreibt aber auch den **Prozess der Bildung der Klassen**, der Erarbeitung von Klassifikationen im ersten Sinn.
3. Klassifikation steht aber auch für den Prozess des Klassierens beziehungsweise des Klassifizierens, dem **Zuordnen von Objekten zu den Klassen** eines Klassifikationssystems der ersten Bedeutung. Alle Objekte einer Klasse haben somit ein gemeinsames Merkmal, das sie von den Objekten der anderen Klassen unterscheidet (Buchanan, 1989, 9).

An dieser Stelle soll vornehmlich die Klassifikation im ersten Sinn näher betrachtet werden. Diese kann im Wesentlichen als Gruppierung oder als Einteilung von Wissen nach einheitlichen methodischen Regeln verstanden werden (Manecke (2004)). Dabei umfassen **Universalklassifikationen** das gesamte Wissen aller Wissenschaften und ihrer Disziplinen und **Spezialklassifikationen** das Wissen eines einzelnen Zweigs der Wissenschaften – dafür mit höherem Detaillierungsgrad. Eine weitere Charakterisierung von Klassifikationen ist die Unterteilung in **hierarchische Klassifikationen** die einen top-down Ansatz verfolgen und jedem Bibliotheksobjekt genau einen gültigen Platz¹ durch Auswahl der zutreffendsten Hauptklasse und Unterklasse zuweisen und den **Facettenklassifikationen** (*Colon Classification*) in welchen für ein Bibliotheksobjekt eine vorgegebene Anzahl voneinander unabhängigen Teilklassifikationen (Facetten) zu erstellen und mittels Doppelpunkte (*colons*) aneinander zu reihen sind. (Panny, 2004)

International bekannte und verbreitete hierarchische Universalklassifikationen sind nach Schulz (2006) die

- nach eigenen Angaben am weitesten verbreitete und in 22. Auflage vorliegende (OCLC (Hrsg.), 2006) **Dewey Decimal Classification (DDC)** mit über 110.000 DDC Nummern und den folgenden Hauptklassen (OCLC (Hrsg.), 2003, 3):

¹In klassischen Bibliotheken entspricht dies oft auch dem Aufstellungsort des Bibliotheksobjekts.

000 Computer, Information & Allgemeines
100 Philosophie & Psychologie
200 Religion
300 Sozialwissenschaften
400 Sprache
500 Naturwissenschaften und Mathematik
600 Technologie
700 Künste & Freizeit
800 Literatur
900 Geografie & Geschichte
sowie die

- **Universal Decimal Classification (UDC)** mit über 56.000 Klassen und den gegenüber DDC leicht variierten Hauptklassen

0 Allgemeines
1 Philosophie, Psychologie
2 Religion, Theologie
3 Sozialwissenschaften
5 Naturwissenschaften
6 Technologie
7 Künste
8 Sprache, Linguistik, Literatur
9 Geografie, Biografie, Geschichte

wobei die Hauptklasse mit der Nummer „4“ derzeit nicht vergeben ist. (UDC Consortium (Hrsg.), 2006)

- Weiters unterhält eine der größten Bibliotheken der Welt, die *Library of Congress*, die **LC-Klassifikation (LCC)**, deren ursprüngliche Aufgabe die Aufstellung der eigenen Bestände war. Die LCC ist heute eine weltweit anerkannte Klassifikation. (Panny, 2004) Da für die Einteilung der Hauptklassen Buchstaben statt dezimalen Zahlen verwendet werden, können mehr als 10, nämlich 21 Hauptklassen wie folgt gebildet werden (Library of Congress (Hrsg.), 2007):

A – General Works
B – Philosophy, Psychology, Religion
C – Auxiliary Sciences of History

D – World History and History of Europe, Asia, Africa, Australia, New Zealand, etc.
E, F – History of the Americas
G – Geography, Anthropology, Recreation
H – Social Sciences
J – Political Science
K – Law
L – Education
M – Music and Books on Music
N – Fine Arts
P – Language and Literature
Q – Science
R – Medicine
S – Agriculture
T – Technology
U – Military Science
V – Naval Science
Z – Bibliography, Library Science, Information Resources (general)

Schulz (2006) zufolge verwendet die Hälfte der wissenschaftlichen Bibliotheken in Deutschland jedoch Individualklassifikationen während in englischsprachigen Ländern vorzugsweise Einheitsklassifikationen Verwendung finden. Erst seit kurzem verwendet die Deutsche Bibliothek in der Deutschen Nationalbibliographie die Dewey Decimal Classification. Andererseits konnten sich Facettenklassifikationen nur in einzelnen Disziplinen durchsetzen wodurch die Verbreitung dieser gering geblieben ist (Panny, 2004).

2.1.11 Repository

Nach Arms (2000) ist die primäre Aufgabe von Repositories in digitalen Bibliotheken die, ein Computersystem zur Speicherung der digitalen Objekte und je nach Konzeption auch von Metadaten bereitzustellen. Er vergleicht die Aufgabe mit jener von Bücherregalen in klassischen Bibliotheken.

2.1.12 Archiv

Ein Archiv ist ein auf die Langzeitspeicherung von digitalen Objekten hin ausgerichtetes Repository. Dazu ist nach Arms (2000) auch eine spezialisierte Organisation des Repository angebracht, da Archive oft den Auftrag haben, Materialien auf unbestimmte Zeit zu übernehmen und den Zugriff auf die Bibliotheksobjekte auf Dauer zu gewährleisten. An anderer Stelle hält er fest, dass die digitale Archivierung schwierig ist und dass digitale Archive länger als heute existierende Computersysteme aufrecht erhalten werden müssen (Arms, 2000).

Dies stellt nach heutigem Stand der Technik durchaus eine ernsthafte Herausforderung für unterschiedliche Disziplinen und Managementtechniken dar.

2.2 Metadaten

Nachdem im vorangehenden Abschnitt „Definitionen von Begriffen im Zusammenhang mit digitalen Bibliotheken“ der Begriff Metadaten kurz erläutert wurde, wird in diesem Kapitel eine weitere Detaillierung hinsichtlich Funktionen und Kategorisierungen vorgenommen.

Die ersten Metadaten über Bibliotheksobjekte fallen bereits bei der Erfassung an, beispielsweise beim Scanvorgang. Dabei interessieren vor allem Informationen dazu (Hard- und Softwarekonfiguration) und zum technischen Format der digitalen Objekte. Metadaten fallen auf jeden Fall im Zuge der Erschließung von Bibliotheks- und Archivobjekten an. Im Bibliotheksbereich hat vor allem die bibliografische Erschließung zur Erstellung eines alphabetischen Katalogs eine lange Tradition. Im deutschsprachigen Raum konnten sich diesbezüglich die Regeln für die alphabetische Katalogisierung (RAK) etablieren. Im Gegensatz dazu erfolgt im Archivbereich die Dokumentation und die Erschließung in den so genannten Findbüchern nach ihrem Entstehungszusammenhang und ihrer Herkunft (Provenienzprinzip). (Borowski, 2007, 89-90)

Die Aufgabe der Metadaten in Bibliotheken ist die Unterstützung von **Schlüsselfunktionen in einer (digitalen) Bibliothek**. (Puglia et al., 2004, 6) nennen diese wie folgt:

- Identifizierung
- Erfassung
- Management
- Zugriff
- Benutzung
- Aufbewahrung

Aufgrund der zu unterstützenden Bibliotheksfunktionen sind neben den die Objekte beschreibenden Metadaten auch solche mit anderen Aufgaben notwendig und müssen vielfältige Themenbereiche abdecken (Puglia et al., 2004, 7-21):

- Neben der prominentesten, die Objekte beschreibende Funktion

werden Metadaten auch verwendet, um für das Management und die Erhaltung der Objekte wichtige Informationen zu folgenden Bereichen zu speichern:

- Administration
- Rechtemanagement
- Technische Darstellung inklusive Informationen für das Rendering und die Darstellung am Client Computer
- Interner struktureller Aufbau von Bibliotheksobjekten
- Verhalten von Bibliotheksobjekten und Informationen zur Steuerung des Verhaltens dieser
- Konservierung der (digitalen) Objekte
- Qualitätsbeurteilung
- Datensatzmanagement als Ergänzung zu administrativen Metadaten zum Management, zur Nachverfolgbarkeit und zur Integration von Metadaten
- Tracking
- Batch Level Metadaten zur Beschreibung des Erstellungsprozesses von Bibliotheksobjekten, beispielsweise zur verwendeten Hard- und Softwarekonfiguration beim Scanvorgang für eine Reihe von Objekten
- Meta-Metadaten, beispielsweise zum Entstehungs- und Erfassungsprozess von Metadaten oder deren Migration beziehungsweise Konvertierungen oder Anpassungen an geänderte Datenformats- und Inhaltsstandards

Dabei unterscheiden Puglia et al. (2004) generell zwischen

- permanenten Metadaten und
- temporären Metadaten.

An der vorangehenden Aufstellung lässt sich eine Vielzahl von Aufgaben und Entstehungszeitpunkten von Metadaten erkennen. Zudem werden kaum für alle digitalen Objekte alle Ausprägungen von Metadaten vorhanden oder erforderlich sein. Für eine übersichtliche Einteilung sind daher wenige den Kernaufgaben angepasste Kategorien hilfreich. Das vorherrschende Verständnis fasst die unterschiedlichen Teilaspekte von Metadaten in drei Kategorien zuzügliche einer eindeutigen Identifikationsmöglichkeit zusammen (Arms, 2000; Hurley et al., 1999):

- **Deskriptive** oder **beschreibende Metadaten** beinhalten im Wesentlichen die bibliografischen Informationen und zusätzliche Informationen zum Inhalt wie Abstract, Schlagwörter oder Deskriptoren.
- **Strukturelle Metadaten** liefern Informationen über Formate und interne Strukturen der Bibliotheksobjekte, das sind im wesentlichen Informationen über Beziehungen zwischen den das Bibliotheksobjekt konstituierenden digitalen Objekte.
- **Administrative Metadaten** bestehen aus Informationen zu Rechten und Informationen zur Steuerung des Zugriffs auf die Bibliotheksobjekte.
- **Eindeutige Schlüssel** (*identifier*)

Eine alternative Betrachtung unterteilt Metadaten nur in die zwei Kategorien der deskriptiven und der strukturellen Metadaten, wobei strukturellen Metadaten auch die administrativen Metadaten umfassen (Hurley et al., 1999, 21). Im Anschluss an eine kurze Übersicht über die vorherrschenden Metadatenstandards werden Metadaten den vier Gruppen „Beschreibende Metadaten“, „Strukturelle Metadaten“, „Administrative Metadaten“ und „Eindeutige Schlüssel“ zugeordnet, erläutert und Beispiele für Attribute gegeben.

2.2.1 Metadatenstandards

Nur wenige Metadatenelemente sind universell und können unabhängig vom originalen Format (Text, Bild, Audio, Video, o.ä.) verwendet werden. Dennoch empfehlen (Puglia et al., 2004, 6) ein gesamthaftes und für alle in Betracht kommenden Formate gültiges Datenmodell zu erstellen. Diese Vorgehensweise hat den Vorteil, dass redundante Datenelemente vermieden und die Strukturbeschreibungen konsistenter werden.

Nach Borowski (2007) können Standards für Metadaten unterschieden werden in Datenformatstandards, Inhaltsstandards und Normdatenstandards.

Datenformatstandards

Datenformatstandards unterstützen die Festlegung der Datenstruktur in welchen später bibliografische Nachweise abgelegt werden können bis hin zu den einzelnen Datenelementen

nach welchen die in digitalen Bibliotheken benötigten Datenbanken und Schemata entwickelt werden. Neben der Fähigkeit der bibliografischen und inhaltlichen Beschreibung von Bibliotheksobjekten ist die Austauschbarkeit von Metadaten zwischen Repositories ein weiterer wichtiger Aspekt welcher beim Entwurf oder bei der Auswahl der Metadatenstruktur zu berücksichtigen ist. Gebräuchliche Datenformatstandards sind nach (Borowski, 2007, 93-94):

- Der Datenformatstandard **Dublin Core (DC)** ist ein Schema für beschreibende Metadaten, welches aufgrund der einfach handzuhabenden Metadatenelemente für eine Vielzahl von unterschiedlichen Bibliotheksobjekten verwendet werden kann. Das Schema deckt drei Aspekte ab: der semantische Aspekt definiert wie Elemente zu interpretieren sind (ein Datum oder ein Autor), der syntaktische Aspekt legt fest wie die Metadatenelemente zu formulieren sind und der strukturelle Aspekt bestimmt die Beziehung von Metadatenelementen untereinander. DC hält in der vereinfachten Version *Dublin Core Metadata Element Set* 15 optional und mehrfach verwendbare Datenfelder für beschreibende Metadaten bereit, welche im folgenden Abschnitt „Deskriptive Metadaten“ genannt und beschrieben werden.²
- Der **Encoded Archival Description (EAD)** ist ein Datenformat zum Austausch von Erschließungsdaten aus Archiven.
- Der Datenformatstandard **Metadata Encoding and Transmission Standard (METS)** berücksichtigt administrative Metadaten inklusive Informationen zur Erschließung mit Angaben zum Original sowie strukturelle Metadaten zur internen Struktur und zur Gruppenbildung. Da METS EAD-Findbücher einbinden und auf diese verweisen kann, besteht auch die Möglichkeit zur engeren Vernetzung mit Archiven.
- **Machine Readable Cataloguing (MARC)** stellt eine ganze Familie von Datenformaten zum Austausch und zur Darstellung von Bibliotheksdaten bereit. Die Datenstruktur nach MARC21 ist im Standard **ISO 2709** festgelegt.
- **Maschinelles Austauschformat für Bibliotheken (MAB)** liegt in der Version MAB2 vor und ermöglicht ebenfalls den Austausch und die Präsentation von bibliografischen Metadaten. MAB ist ein in Österreich und Deutschland verbreitet verwendetes Datenformat für Bibliotheken (Block et al., 2005, 1).
- **SPECTRUM** ist ein britischer Standard zum Management von Museumsobjekten.
- **Text Encoding Initiative (TEI)** stellt einen XML-basierten Datenformatstandard zum Austausch von digital vorliegenden Texten dar.

²Weitere Informationen zur Dublin Core Metadata Initiative unter <http://dublincore.org/usage/terms/dc/current-elements/>

- **Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)** stellt Methoden bereit, um basierend auf dem Protokoll *HTTP* und der Markup Language *XML* Metadaten aus digitalen Repositories abzufragen und zusammenzuführen. Das Framework definiert nach Lagoze und van de Sompel (2004) *Data Provider* welche Metadaten verfügbar machen und *Service Provider* welche die Metadaten mit einem Mehrwert versehen. Harvester werden von *Service Providern* betrieben und sammeln Metadaten aus Repositories. Das OAI-PMH Framework stellt klar definierte Anforderungen an Repositories, insbesondere wie welche Anfragetypen vom Repository zu bearbeiten sind.
- Das **Conceptual Reference Model (CIDOC CRM)** wird im Rahmen der Dokumentation des kulturellen Erbes verwendet und stellt ein objektorientiertes Referenzmodell zur Beschreibung von Begrifflichkeiten und Beziehungen sowie von Prozessen der Wissensentstehung dar.
- Das abstrakte relationale Modell **Functional Requirements for Bibliographic Records (FRBR)** wird im Unterabschnitt „Deskriptive Metadaten“ behandelt. Die zentralen Entitäten Werk, Expression, Manifestation und Exemplar wurden bereits im Abschnitt „Definition von Begriffen im Zusammenhang mit digitale Bibliotheken“ kurz erläutert.
- Das **Resource Description Framework (RDF)** wurde für diesen Zweck des Austauschs von Metadaten entwickelt. RDF unterstützt den Datenaustausch insofern, als mittels eines RDF-Schemas die Syntax (wie Metadaten auszudrücken sind) und die Struktur (beispielsweise dass die Objekte *Jahr*, *Tag* und *Monat* die Komponenten des Objekts *Datum* repräsentieren) der Metadaten in XML-Form zwischen Computersystemen kommuniziert werden kann. (Arms, 2000)

Inhaltsstandards

Ein Inhaltsstandard unterstützt nach (Borowski, 2007, 94-95) die Erstellung von beschreibenden Metadaten durch semantischen Beschreibung der Inhalte unter Zuhilfenahme von Regelwerken. Die so nach strengen Regeln erhobenen deskriptiven Metadaten werden einem Datenformatstandard entsprechend – beispielsweise MAB oder MARC – abgelegt beziehungsweise getauscht. Der im deutschsprachigen Raum vorherrschende Inhaltsstandard nennt sich

- **Regeln für die alphabetische Kategorisierung (RAK)** und legt die formale Katalogisierung von Bibliotheksobjekten fest.

Normdatenstandards

Die Aufgabe der Normdatenstandards ist die Regelung betreffend Personennamen, Schlagworten oder die Erstellung von Thesauri und erlangen insbesondere beim Datenaustausch

zwischen digitalen Bibliotheken und zentralen Portalen eine erhöhte Relevanz. Mit Fokus auf den deutschsprachigen Raum stehen die folgenden Standards zur Verfügung (Borowski, 2007, 95):

- Der **Encoded Archival Context (EAC)** regelt die digitalen Speicherung von Personen-, Orts- und Einrichtungennamen im Archivbereich.
- **Personennamendatei (PND)** ist ein deutscher Standard zur Ansetzung von Personennamen im Rahmen der Katalogisierung.
- Die **Schlagwortnormdatei (SND)** und
- die **Gemeinsame Körperschaftsdatei (GKD)** regeln die Verwendung von Schlagworten und Körperschaftsbezeichnungen.

Die Dateien PND, SWD und GKD werden von der Deutschen Bibliothek auf CD-ROM herausgegeben und zur Verfügung gestellt (Borowski, 2007, 95).

2.2.2 Deskriptive Metadaten

Deskriptive bzw. beschreibende Metadaten (*descriptive metadata*) dienen dem Auffinden und der Identifizierung von Ressourcen und liefern nach Puglia et al. (2004) Antworten auf die Fragen **Wer**, **Was**, **Wann** und **Wo** im Bezug auf Ressourcen. Sie beschreiben den Inhalt eines Bibliotheksobjekts (mittels Katalogdatensätzen, Abstrakten oder Schlagwörter), stellen Einstiegspunkte (*access points*; gebundene Schlagwörter) zur Verfügung und etablieren logische oder hierarchische Beziehungen zu anderen Objekten. Neben bibliografischen Informationen werden auch Daten zu Medientypen, Datenformaten und Zustand erfasst. Beschreibende Metadaten sind liegen meist in Textform vor, könnten aber auch als Bilder inklusive Kartenmaterial, Audio oder Computerprogramme verfügbar sein (Arms, 2000).

Die in den Katalogen gespeicherten Metadaten sind hoch strukturiert, folgen einem Datenformatstandard, werden vorzugsweise im Format MAB (Maschinelles Austauschformat für Bibliotheken) oder MARC (Machine Readable Cataloging, in der Version MARC21) gespeichert und häufig online mittels *Online Public Access Catalogue* (OPAC) verfügbar gemacht (Borowski, 2007, 89). Neuere Standardisierungsversuche befinden sich im Stadium der Umsetzung (FRBR) oder werden für weniger strenge Anforderungen wie beispielsweise zur Anreicherung von Webseiten mit Metadaten unter Zuhilfenahme des *Dublin Core* Standards eingesetzt (Hurley et al., 1999, 21).

Deskriptive Metadaten werden gewöhnlich von Experten erstellt und gewartet und repräsentieren einen enormen Aufwand durch ausgebildetes Fachpersonal. Häufig werden Metadaten über lange Zeit – in Einzelfällen über Jahrhunderte – gepflegt und weiter entwickelt (Arms, 2000). Vor diesem Hintergrund ergeben sich zwei Randbedingungen:

- Die getätigten Investitionen müssen geschützt werden.

- Es werden große Anstrengungen unternommen, die Metadatenerstellung zu automatisieren oder die Erstellung durch Personal mittels Werkzeugen der Informationstechnologie zu verbessern und zu beschleunigen.

Neben den bereits genannten und etablierten Metadatenstandards wird in verschiedenen Arbeitsgruppen die Implementation des abstrakten relationalen Modells **FRBR** geprüft. Die Abkürzung FRBR steht für **Functional Requirements for Bibliographic Records** und wurden 1998 als Empfehlung der *International Federation of Library Associations and Institutions* (IFLA) erarbeitet um Bibliothekskataloge vor allem hinsichtlich der Navigation durch diese neu zu strukturieren (Block et al., 2005, 3).³

Es handelt sich dabei um ein logisches Denkmodell für bibliografische Beschreibungen in Form eines relationalen Modells mit abstrakten Entitäten (Block et al., 2005, 5). Die Diskussion darüber befindet sich im deutschsprachigen Raum noch im Anfangsstadium (Block et al., 2005, 1) und ist hinsichtlich der Umsetzung bisweilen massiver Kritik ausgesetzt (Block et al., 2005, 3). Mit ein Grund ist sicherlich, dass nur in etwa 20% der vorhandenen MAB/MARC-Katalogsätze alle vier FRBR-Ebenen beschrieben sind bzw. sich aus vorhandenen Katalogdaten ableiten lassen (Block et al., 2005, 4). Block et al. nennen als gut darstellbar exemplarisch die Bibel, den Koran oder Goethes Faust aber verweisen jedoch nachdrücklich darauf, dass für die große Masse des Materials die Ebenen Work, Expression und Manifestation wenig ausgeprägt sind. In ihrer Schlussbetrachtung (Block et al., 2005, 11) streichen sie die große Chance auf einen potenziellen Mehrwert durch verbesserte Navigation durch Kataloge und die elegantere Aufbereitung von großen Treffermengen heraus, welche durch das relationale Modell realisierbar scheint.

Wesentliches Merkmal des FRBR ist die relationale Struktur von in drei Gruppen zusammengefassten Entitäten sowie Relationen. Im folgenden werden diese nach Block et al. (2005) kurz dargestellt:

- Die **erste Gruppe** mit den vier Entitäten
 - Work
 - Expression
 - Manifestation
 - Exemplar

welche bereits im Abschnitt „Definition von Begriffen“ beschrieben wurden.

- Die **zweite Gruppe** umfasst die Entitäten für Informationen zu den Akteuren welche künstlerische Werke, Expressionen oder Manifestationen schaffen oder damit in Zusammenhang stehen. (Block et al., 2005, 3)
 - Personen (*Person*)

³Nach weiteren vertiefenden Informationen zur genannten Empfehlung kann ausgehend von der Webseite <http://www.ifla.org/VII/s13/frbr/frbr.htm> recherchiert werden.

- Körperschaften (*Corporate Body*)
- In den Entitäten der **dritten Gruppe** werden Informationen über den Gegenstand von Werken abgelegt und dienen der sachlichen Erschließung. (Block et al., 2005)
 - Begriffe und Konzepte (*Concept*)
 - Gegenstände (*Object*)
 - Ereignisse (*Event*)
 - Orte (*Place*)

Die im FRBR definierten **Relationen** weisen für jede der drei Gruppen eine unterschiedliche Bedeutung auf (Block et al., 2005, 3-4):

- Für die erste Gruppe von Entitäten – Werk, Expression, Manifestation und Exemplar – sind es *immanente Relationen* wie beispielsweise „Werk A ist realisiert durch Expression X“ und setzen somit die bibliografischen Entitäten untereinander in Beziehung.
- Für die zweite Gruppe der Entitäten – den Personen und Körperschaften – wird mit der Relation die *Rolle* der Personen und Körperschaften im Entstehungszusammenhang definiert, indem die Relationen mit Attributen wie Verfasser, Herausgeber, Verleger, Übersetzer und dergleichen versehen werden.
- Auch für Entitäten der dritten Gruppe stellen die Relationen jeweils die Beschreibungen der *Rolle*, jedoch hier jene des im Werk beschriebenen *Begriffs, Gegenstands, Orts oder Ereignisses* dar. Hier ist die Bedeutung der Relationen auf „ist Gegenstand von“ oder „gehört zu“ beschränkt.

Die Möglichkeiten der Verknüpfung beschränken sich nicht auf FRBR. MAB ermöglicht bereits mittels Verknüpfungsfeldern eine gezielte Suche, beispielsweise nach einer präferierten Sprache oder nach einem bevorzugten Medium. Allerdings können mittels MAB nur Manifestationen miteinander verknüpft werden. (Block et al., 2005, 10)

Im folgenden werden die fünfzehn Metadatenelemente des **Dublin Core** aufgelistet. Dabei sind alle Elemente optional und müssen nicht verwendet werden, können aber auch mehrfach verwendet werden, beispielsweise um *alle* Autoren eines Artikels zu nennen. (Arms, 2000):

- **Ersteller** (*Creator*) als jene Person oder Organisation, welche als Autor, Fotograf, Grafiker usw. für den intellektuellen Inhalt einer Ressource verantwortlich ist.
- **Titel** (*Title*), der dem Bibliotheksobjekt vom Autor oder dem Herausgeber gegeben wurde.

- **Thema** (*Subject*) des Objekts.
- **Beschreibung** (*Description*) des Objekts in Textform inklusive Abstract und Inhaltsbeschreibungen bei nichttextlichen Objekten.
- **Herausgeber** (*Publisher*) als Person oder Organisation, welche für die Veröffentlichung verantwortlich zeichnet.
- **Beiträge von** (*Contributor*) stellt eine Person oder Organisation dar, welche maßgebliche aber dem Ersteller nachrangige Beiträge leistet, wie beispielsweise ein Illustrator oder Übersetzer.
- **Datum** (*date*) der Erstellung oder Herausgabe eines Objekts.
- **Kategorie** (*type*).
- **Format** (*format*) der Daten des Objekts. Die Angabe ist für die Auswahl geeigneter Hard- und Software erforderlich.
- **Schlüssel** (*identifier*) zur eindeutigen Identifikation von Metadatensätzen und Objekten.
- **Quelle** (*source*) als weiteres Objekt, auf welchem das aktuelle Objekt basiert.
- **Sprache** (*language*) des Objektinhalts.
- **Beziehung** (*relation*) als Schlüssel eines anderen Objekts mit einer Bezeichnung der Beziehung.
- **Abdeckung** (*coverage*) im Sinne einer räumlichen und zeitlichen Zuordnung des Objekts.
- **Rechte** (*rights*), welche der Nutzung des Objekts zu Grunde liegen.

Neben dem Vorteil der Flexibilität und der einfachen Anwendbarkeit ohne fachspezifische Einschulung zeigt sich hier auch die Schwäche des DC in der fehlenden Präzision in der Anwendung der Metadatenelemente.

Die **Beziehung zwischen Metadaten und Bibliotheksobjekt** muss sichergestellt werden, auch wenn Metadaten in Form eines Katalogs oder Index in digitalen Bibliotheken getrennt von den Bibliotheksobjekten gespeichert werden. Beiden Varianten – Speicherung der Metadaten getrennt von den digitalen Objekten oder gemeinsam mit diesen im selben Repository wie diese – weisen sowohl Vorteile als auch Nachteile auf (Arms, 2000):

- Speicherung der Metadaten getrennt von den digitalen Objekten: Ermöglicht die Verwendung unterschiedlicher Technologien und Prozesse für die Metadatenverwaltung und die Verwaltung der digitalen Objekte, erfordert aber eine stabile und sichere Verlinkung der Metadaten mit den digitalen Objekten.

- Speicherung der Metadaten gemeinsam mit den digitalen Objekten: Ermöglicht die Einbettung (*embedding*) der Metadaten in die digitalen Objekte selbst und erleichtert dadurch einen Austausch oder eine einfache und effiziente Weitergabe des digitalen Objekts gemeinsam mit den dazugehörigen Metadaten.

Im Einzelfall ist abzuwägen, welcher dieser beiden Varianten insbesondere im Hinblick auf die effiziente Wartbarkeit der Daten und die Integrität der Beziehungen zwischen Metadaten und digitalen Objekten der Vorzug gegeben wird.

2.2.3 Strukturelle Metadaten

Strukturelle Metadaten (*structural metadata*), auch als Strukturdaten bezeichnet, sind erforderlich, um ein Bibliotheksobjekt den Benutzenden präsentieren zu können. (Hurley et al., 1999, 21) bezeichnen sie als Daten, die die Nutzungsmöglichkeiten und die Navigation durch ein Bibliotheksobjekt beschreiben.

Während deskriptive Metadaten in den Formaten MAB oder MARC gespeichert werden, werden strukturelle Metadaten unabhängig von diesen in relationalen Tabellen oder in den SGML⁴-Dateien selbst gespeichert. (Hurley et al., 1999, 21)

Strukturelle Metadaten beschreiben ein Bibliotheksobjekt auf zweierlei Ebenen: Auf der Ebene des Bibliotheksobjekts selbst und auf der Ebene der einzelnen untergeordneten digitalen Objekte.⁵ Diese beiden Ansatzpunkte werden nachstehend beschrieben. Eine dritte Ebene stellen die Schema-Ebene dar: Schemadefinitionen stellen eine Form sehr grundlegender struktureller Metadaten dar, da sie Informationen über die Struktur der eingesetzten Metadaten bereitstellen. (Hurley et al., 1999, 32).

Strukturelle Metadatenelemente für die Ebene des Bibliotheksobjekts

Auf der Ebene des Bibliotheksobjekts (*object level*) werden als verpflichtend die folgenden Metadatenelemente angeführt (Hurley et al., 1999, 26):

- Eine *unique identifier reference* zur Identifikation des digitalen Objekts.

⁴SGML steht für *Standard Generalized Markup Language*, liegt als ISO Standard ISO 8879 vor und stellt einen geräte- und systemunabhängigen Dokumentenbeschreibungsstandard dar. (Connolly, 2004)

⁵Der Zusammenhang zwischen MARC und strukturellen Metadaten am Beispiel eines eingescannten Buchs findet sich als übersichtliche Grafik unter der URL <http://www.loc.gov:8081/ndlint/repository/rep-examples/bib-hdl-obj.gif>. Ein Beispiel für die Verwendung von strukturellen Metadatenelementen anhand des angeführten Exemplars findet sich zweigeteilt: einerseits die Metadatenelemente für das digitale Bibliotheksobjekt des gescannten Buchs insgesamt unter URL <http://www.loc.gov:8081/ndlint/repository/rep-examples/att-use-ex-toc.html> und andererseits die Metadatenelemente für ein einzelnes digitales Objekt – welches eine einzelne gescannte Seite umfasst – unter URL <http://www.loc.gov:8081/ndlint/repository/rep-examples/dance031-meta-inst2-05.html>. Die Dokumentation dazu findet sich unter <http://www.loc.gov:8081/ndlint/ndlpindex.html>. (Alle URLs mit Stand 2007-04-20)

- Der *content type* beschreibt das Format des verfügbaren Objekts, welches nicht notwendigerweise dem Format bei der Erfassung entsprechen muss. Dazu dienen einerseits Deskriptoren für Seitenbild, Text, Audio und andererseits die Nennung der verfügbaren Dateiformate, beispielsweise HTML, PDF, TIFF oder GIF.
- Das *extent*-Element beschreibt detailliert den Aufbau des Bibliotheksobjekts unter Angabe der untergeordneten Objekte (z.B. 100 Bilder von Buchseiten, 10 Abschnittsüberschriften und 5 Farbbilder). Obwohl die Benutzenden ein Exemplar eines Bibliotheksobjekt nach Arms (2000) als eine Einheit betrachten und dessen innere Struktur sie nicht interessieren, ist diese interne Struktur zur Darstellung und zur Navigation durch das Bibliotheksobjekt von Bedeutung. Dem kann in technischer Hinsicht dadurch Rechnung getragen werden, indem für verschiedene Arten von Objekten unterschiedliche Objektmodelle wie beispielsweise Webobjekte, Textdokumente mit SGML oder Computerprogramme bereitgestellt werden. Für jede Art von Objekt wird dabei ein Satz Regeln aufgestellt, welcher die Beziehungen unter den einzelnen digitalen Objekten festlegt. Arms (2000) führt illustrierend an, dass ein und derselbe Ausstellungskatalog als PDF oder in SGML unterschiedliche Struktur aufweisen und über unterschiedliche Objektmodelle verfügen. Er weist aber auch darauf hin, dass die Betreiber von digitalen Bibliotheken ihre Datenmodelle laufend an die während des Betriebs der digitalen Bibliothek gemachten Erfahrungen anpassen und dadurch die Interoperabilität zwischen unabhängigen Repositories erschwert wird.

Strukturelle Metadatenelemente für die Ebene der einzelnen digitalen Objekte

Die Struktur auf der Ebene der einzelnen digitalen Objekte (*sub-object level*) beschreibt die einzelnen Dateien der digitalen Objekte mit den folgenden Elementen (Hurley et al., 1999, 26):

- *Sub-object reference* beinhaltet verpflichtend die Informationen, welche zum Zugriff auf das Objekt erforderlich sind. Idealerweise ist dies ein URN. Ansonsten sollte es sich zumindest um einen *identifier* handeln, welcher auf der URN des übergeordneten Objekts basiert. (Borowski, 2007, 86) weist darauf hin, dass die Dateien nach einheitlichen Konventionen eindeutig in durchnummerierten Dateinamen benannt sein müssen. Darüber hinaus sollte der Dateiname ein Bestandteil eines persistenten Pfads sein und innerhalb des verwendeten Systems eindeutig sein. Ordner, die die einzelnen digitalen Objekte des Digitalisats aufnehmen, sollten Borowski (2007) zufolge mit selbsterklärenden Namen versehen sein. Zusätzlich empfehlen Borowski (2007), dass die Metadaten den Dateinamen inklusive dem gesamten Pfad umfassen. bei der Namensgebung sind Sonderzeichen – darunter auch Leerzeichen – im Hinblick auf eine Konvertierung oder Migration der digitalen Objekte zu vermeiden und die Namen so kurz als möglich zu halten.
- *Sub-object Format* beinhaltet verpflichtend die Informationen über den Dateityp eines digitalen Objekts: beispielsweise GIF-Bild, PDF-Datei oder TIFF-G4-Bild.

Nach Arms (2000) bieten sich dazu die so genannten *MIME*⁶ *types* an. Dabei wird mittels *Typ* (z.B. „Text“), *Untertyp* (z.B. „HTML“) und einem oder mehreren *Parametern* (z.B. „charset=windows-1252“) eine Datei von der erforderlichen Anwendung bis zum Zeichensatz beschrieben. Der MIME Typ des Objekts aus den angeführten Beispielen lautet dann: „text/html; charset=windows-1252“. Neben den offiziell registrierten MIME Typen können eigene MIME (Unter-) Typen, beispielsweise „audio/x-pn-realaudio“, Verwendung finden. Ein Nachteil von MIME ist, dass komplexe Objekte nicht im erforderlichen Ausmaß mit den einfachen MIME Typen beschreibbar sind. Arms (2000) nennt beispielhaft Textdokumente mit eingebetteten Bildern.

- ***Sub-object sequence*** beinhaltet verpflichtend die in numerischer Form vorliegende Informationen über die Reihenfolge des Objekts: beispielsweise handelt es sich um die vierte Seite eines bereits auf *object level* spezifizierten Buchs.
- ***Sub-object dimensions*** beinhaltet nicht verpflichtend die Informationen zur Form des Objekts und sind zur Navigation und Manipulation hilfreich: abgelegte Informationen zu einem Bild können die angebotenen Auflösungen (nicht die, mit welcher erfasst wurde) und Höhe und Breite in Pixeln sein.
- ***Sub-object size*** ist nicht verpflichtend und liefert eine der beiden Informationen: „reference“ oder „full“.
- ***Data file size*** enthält – wenn verwendet – die Dateigröße des digitalen Objekts.
- ***Sub-object type*** enthält gegebenenfalls die Information über die Nutzungsmöglichkeit des Objekts: z.B. Inhaltsverzeichnis.
- ***Sub-object value*** nimmt für Seitenobjekte die tatsächliche Seitenzahl der gedruckten Vorlage auf.
- ***Structural division(s)*** werden dann verwendet, wenn das Objekt mehrere Abschnitte aufweist.
- ***Sub-object relationship(s)*** zu anderen digitalen Objekten, die hierarchisch höher, tiefer oder gleichwertig stehen (*parent*, *children* oder *siblings*).

Metadaten zum **ausführbares Verhalten** von digitalen Objekte werden bei Puglia et al. (2004) als eigenständige Kategorie geführt, in dieser Arbeit aber Hurley et al. (1999) und Arms (2000) folgend den strukturellen Metadaten zugerechnet, da die das Verhalten beschreibenden Metadaten Auskunft über die innere Struktur und die Nutzungsmöglichkeiten eines digitalen Objekts geben. Diese können sein (Puglia et al., 2004, 11):

- Den verschiedenen Genres von Objekten (z.B. Bücher, Bilder, Präsentationen) zugeordnete Verhaltensweisen.

⁶MIME steht als Abkürzung für „Multipurpose Internet Mail Extensions“ und stellt ein Standard für die Beschreibung der Dateitypen von per E-Mail ausgetauschten oder im World Wide Web bereitgehaltenen Dateien dar.

- Den verschiedenen Dateiformaten zuordenbare Verhaltensweisen.
- Verhaltensweisen, welche durch Zuordnung von ausführbarem Code zu digitalen Objekten möglich werden.

So können digitale Bibliotheksobjekte aus einzelnen digitalen Objekten aus Repositories gezielt abgerufen und erstellt werden sowie Navigations- und Manipulationsmöglichkeiten den Benutzenden präsentiert werden.

2.2.4 Administrative Metadaten

Administrative Metadaten beinhalten Informationen, die dem Repository die Verwaltung der Sammlung digitaler Objekte möglich machen. Dabei lässt sich die Grenze zwischen administrativen und strukturellen Metadaten nicht immer exakt ziehen, denn manche Metadatenelemente sind einerseits von administrativem Nutzen, dienen aber andererseits auch der Strukturierung eines digitalen Objekts (Hurley et al., 1999, 23). Dabei lässt sich auch kein genereller Standard anwenden, denn sie sind durch projektspezifische Anforderungen wie Vereinbarungen für die Aufnahme von Bibliotheksobjekten in das Repository oder die Verrechnung und Workflows individuell festzulegen (Puglia et al., 2004, 8).

Administrativen Metadaten kommen vielfältige Aufgaben zu: neben den Informationsbedürfnissen für technische Aufgaben müssen Konservierungsaufgaben⁷ oder verrechnungstechnische Aufgaben abgedeckt werden. In dieser Arbeit wird eine Unterteilung der Metadatenelemente in *Technische administrative Metadaten* und *Nichttechnische administrative Metadaten* verwendet.

Technische administrative Metadaten

Bei der Erstellung der digitalen Objekte – beispielsweise beim Scannen – fallen automatisch technische Metadaten an (Borowski, 2007, 92). Daneben müssen aber auch Informationen bereitgestellt werden, die zur Betrachtung und Benutzung erforderlich sind (Hurley et al., 1999, 23).

Beispiele für technische Metadaten sind die folgenden, Borowski (2007) und Hurley et al. (1999) entnommenen Metadatenelemente, wobei diese primär den Anforderungen der Erstellung von digitalen Bibliotheksobjekten durch scannen folgen:

- Auflösung.
- Farbtiefe.

⁷Dabei fallen Metadaten zu Aktivitäten der Konservierung wie Refreshing, Migration oder Emulationskonfigurationen und deren Ergebnis über die gesamte Aufbewahrungsdauer hinweg an und sind für Entscheidungsprozesse über den Langzeitwert von Erhaltungsmaßnahmen relevant (Puglia et al., 2004, 11).

- Abmessungen des Digitalisats in Pixel.
- Lichtquelle mit Angaben zur Lichtart und Lichttemperatur.
- Farbmanagement, falls angewendet.
- Prozesse manueller oder computergestützter Verbesserung bzw. Veränderung des Objekts. Dabei kann es sich auch um Anreicherung zusätzlich hinzugefügte Artefakte handeln.
- Hardware: der Typ des eingesetzten Scanners.
- Software, welche beim Scannen eingesetzt wurde.
- Komprimierungsgrad inklusive Angabe der zur Komprimierung verwendeten Software.
- Dateiformat. Dabei soll das Dateiformat der im Repository gespeicherte Datei und gegebenenfalls das Dateiformat des übertragenen Objekts angegeben werden
- Farbraum (CMYK, RGB usw.)
- Dateinamen und persistenter Pfad zur Verknüpfung der Metadaten mit den Objektdateien (Bilddateien).

Weiters sollten vergrößerte Bildauszüge dem Gesamtbild auch hinsichtlich der Lage oder Reihenfolge zuordenbar sein (diese Informationen stellen auch strukturelle Metadaten dar), oder die Version der digitalen Instanz des Objekts erkennbar gemacht werden können. (Hurley et al., 1999, 30)

Nichttechnische administrative Metadaten

Nichttechnische administrative Metadaten fallen teilweise in der Vorbereitung der Erstellung der digitalen Objekte – beispielsweise durch Scannen – an, teilweise werden diese erst danach im Zuge einer Zugänglichmachung der Objekte erfasst.

Neben zuverlässigen **Angaben über die Scanvorlage** beziehungsweise **das Original** – es sollen die durch Scannen entstehenden digitalen Objekte eingeschlossen sein – müssen Angaben über die **rechtlichen Rahmenbedingungen**, **Verrechnung** oder zur **Vernetzung über Sammlungsgrenzen hinweg** abgelegt werden.

Die Eigentümerschaft und Autorenschaft die digitalen Objekte betreffend sowie Rechte inklusive Reproduktionsrechte müssen für jedes digitale Objekt dokumentiert werden damit das Urheberrecht, Copyright und Bildrecht gewahrt werden kann ((Hurley et al., 1999, 23) und (Borowski, 2007, 93)). (Puglia et al., 2004, 8) weisen im Zusammenhang mit dem Rechtemanagement (*rights management*) auf ungelöste Fragen insbesondere bei der Konservierung hin, da mit konservierenden Maßnahmen wie Migration aber auch Emulation eine

Änderung der Formate, Strukturen und Eigenschaften einhergehen kann. Metadaten zum Rechtemanagement sind von doppelter Bedeutung: einerseits dienen sie dem *LIS-Personal* zur Bestimmung der die Rechte haltenden Partei, andererseits werden diese Informationen auch zur Zugriffs- und Nutzungskontrolle durch das Repository oder das Bibliothekssystem genutzt (Puglia et al., 2004, 8). Darüber hinaus sollen bei Objekten, die unter die Haager Konvention⁸ oder die EU-Ausfuhrbestimmungen fallen, die geforderten Angaben für den Zoll bzw. die Polizei in administrativen Metadatenelementen abgebildet werden können (Borowski, 2007, 93).

Die Vernetzung über die Grenzen der jeweiligen Sammlung hinweg stellt Anforderungen an eventuell zusätzliche Metadatenelemente, wenn beispielsweise die Einbindung in zentrale Verzeichnisse oder Portale in Angriff genommen werden soll. Dazu sind zusätzlich Metadatenelemente, welche eine gesamte digitale Sammlung beschreiben⁹, erforderlich. (Borowski, 2007, 93)

Die im folgenden angeführten Beispiele für Metadatenelemente sind wiederum Borowski (2007), aber insbesondere Hurley et al. (1999) entnommen:

- Identifikation des Originals durch eine Quellenangabe.
- Art des Originals wie beispielsweise Foto, gedruckte Seite usw.
- Charakteristik des Originals, die geeignet ist das Ergebnis der Erfassung (Scannen) zu beeinflussen.
- Originalgröße.
- Datum der Aufnahme.
- Inhaber des Copyright des digitalen Objekts als Name bzw. Schlüsselfeld.
- Jahr des Copyright.
- *credit line* als Text, welcher bei jeder Anzeige eines Objekts mit angezeigt werden muss. Beispielsweise „Copyright Berkely Art Museum, 1987. All rights reserved.“.
- Einschränkungen für die Anzeige und die Übertragung (*display/transmission restrictions*) wie etwa „This file may be displayed or transmitted across a network only by person(s) who have signed a license agreement with ...“.

⁸Die als „Haager Konvention“ bekannte Übereinkunft zum Schutz von Kulturgütern im Fall von kriegsrischen Auseinandersetzungen ist am 14. Mai 1954 mit dem Hintergrund der umfangreichen Zerstörungen während des Zweiten Weltkriegs vereinbart worden und ist mittlerweile von mehr als 100 Staaten ratifiziert worden. Heute umfasst die Haager Konvention auch Schutzmaßnahmen während Friedenszeiten: Die Registrierung von unbeweglichen Kulturgütern von besonderem Wert, die Markierung durch Tafeln mit dem Emblem der Konvention, die Erlangung eines weitreichenden Schutzes für diese Kulturgüter bis hin zur Ahndung von Verletzungen der Konvention. (UNESCO (Hrsg.), 2007)

⁹Das sind dann um eigens zu definierende Metadaten für eine Sammlung oder anders formuliert um Meta-Metadaten über die Menge der verwalteten Metadaten.

- Einschränkungen für den Nachdruck und die Verbreitung (*copy/distribution restrictions*) wie beispielsweise „Copy and distribution of this file is prohibited without the express written consent of ...“.
- Lizenzzeitraum als Datum des Beginns der Gültigkeit der Lizenzvereinbarung und des Ablaufs der Lizenzvereinbarung für ein digitales Objekt.

Die weiteren erforderlichen Metadatenelemente gestalten sich je nach Anwendungsfall unterschiedlich und werden daher an dieser Stelle nicht weiter aufgelistet. Folgende Themen müssen aber durch Metadatenelemente Berücksichtigung finden:

- Verrechnung
- auf das Repository maßgeschneiderte Arbeitsabläufe
- Haager Konvention und Ausführbestimmungen
- Zugriffs- und Nutzungskontrolle
- Vernetzung über Sammlungsgrenzen hinaus

Darüber hinaus gehenden Anforderungen können hier ebenfalls noch mit weiteren Metadatenelementen Rechnung getragen werden.

2.2.5 Eindeutiger Schlüssel – *Unique Identifier*

Ein *Unique Identifier* ist eine unverwechselbare Kennzeichnung eines Elements in einer Gruppe von Elementen und zur Identifikation, Lokalisierung und zum Management der digitalen Objekte erforderlich (Arms (2000) und Puglia et al. (2004)). Damit die Eindeutigkeit weltweit gewährleistet werden kann, wurden und werden eine Vielzahl an Standards erarbeitet. Beispiele dafür sind nach Arms (2000):

- *Serial Item and Contribution Identifier Standard* (SICI)
- *International Standard Book Number* (ISBN) – die ISBN wurde nach Anforderungen des Buchhandels entworfen und identifiziert verkaufsfähige Produkte in ihren unterschiedlichen Formaten wie beispielsweise Hardcover und Paperback. Im Bibliothekswesen ist die Unterscheidung jedoch primär nach dem Inhalt und nicht nach der Form von Bedeutung. Der Unterschied zwischen ISBN und einer Kennzeichnung auf Exemplarebene wird besonders deutlich, wenn seltene Werke in einer Sammlung als einzelne Exemplare eindeutig identifizierbar sein müssen.

- *Domain names* und *Uniform Resource Locators* (URLs) – sind die am weitesten verbreiteten Identifier und haben sich als außerordentlich erfolgreich bewährt. Die mangelnde Dauerhaftigkeit von URLs wird im Zusammenhang mit digitalen Bibliotheken jedoch langfristig zu einer Problem.
- *Persistent Names* und *Uniform Resource Names* (URN) – stellt global eindeutige Namen zur dauerhaften Benennung von Ressourcen unabhängig von IT Systemen zur Verfügung.
- *Digital Object Identifier* (DOI) – ein numerischer Identifier initiiert von der *Association of American Publishers*. Dieser besteht aus zwei Teilen: einem Identifier für einen Verlag als *naming authority* und einem Identifier zur Bezeichnung einer Manifestation. Arms (2000) nennt als Beispiel ein von *Academic Press* veröffentlichtes Buch mit dem DOI „doi:10.1006/0121585328“.

In weiterer Folge sollen URI, URL und URN detaillierter betrachtet werden.

Uniform Resource Identifier (URI)

Ein *Uniform Resource Identifier* ist eine Zeichenkette, welche eine physikalische oder abstrakte Ressource bezeichnet und ist durch die folgende Begriffe charakterisiert (Berners-Lee et al., 2005, 1-5):

- Uniform - die Einheitlichkeit soll unabhängig vom Zugriffsmechanismus und Kontext gewährleistet sein.
- Resource - Ressourcen sind nach Berners-Lee et al. (2005) im weitesten Sinn als konzeptuelles Mapping auf eine Entität oder Gruppe von Entitäten zu verstehen: beispielsweise Dokumente, Bilder, Services im Internet oder ganzen Sammlungen von Ressourcen aber auch nicht per Netzwerk übertragbare Ressourcen wie etwa Organisationen oder gebundene Bücher.
- Identifier - Ein Identifier wird als eigenständiges Objekt verstanden, welches als Referenz auf eine Entität Verwendung finden kann.

Ein URI Schema definiert den Namensraum des URI und beinhaltet möglichst viele allgemeine und dadurch in einer Vielzahl von URI-Schemen erforderliche Elemente (Berners-Lee et al., 2005, 7).

Nach aktueller Sichtweise kommt der *Uniform Resource Identifier* in den beiden Ausprägungen „Locator“ und „Name“ vor. Typischer Vertreter von URI als Locator ist der *Uniform Resource Locator* (URL), jener als Name der *Uniform Resource Name* (URN) ((Berners-Lee et al., 1998, 4) und URI Planning Interest Group (2001)). Beide sind in den beiden nachfolgenden Unterabschnitten kurz umrissen.

Die Überarbeitung des betreffenden *Request for Comments* (RFC¹⁰) durch Berners-Lee et al. (2005) führt die Generalisierung im Rahmen des URI-Schemas weiter, indem er die Möglichkeit offen lässt, dem *Uniform Resource Locator* und dem *Uniform Resource Name* weitere gleichrangige Ausprägungen von URIs hinzuzufügen.

Zwei Beispiele für URIs mit der Angabe der jeweiligen syntaktischen Komponenten sind nach (Berners-Lee et al., 2005, 16):

```

foo://example.com:8042/over/there?name=ferret#nose
  \_/  \-----/\-----/ \-----/ \_/
  |      |           |           |      |
scheme authority path query fragment
  |
  / \ /
urn:example:animal:ferret:nose

```

Der *Uniform Resource Identifier* URI stellt demnach ein abstraktes Konzept zur Generalisierung von URL und URN dar.

Uniform Resource Locator (URL)

Der Begriff *Uniform Resource Locator* (URL) steht für ein Subset von URIs und hat die Funktion der Angabe eines Speicherorts von Ressourcen (*location*) und die Deklaration eines primären Zugriffsmechanismus (Berners-Lee et al., 2005, 7) wie beispielsweise „http“ oder „ftp“.

Uniform Resource Name (URN)

Der Begriff *Uniform Resource Name* (URN) stellt ebenfalls ein Subset von URIs dar, bezeichnet aber einen Namen für eine Ressource (*name*) auf eindeutige Weise und ist vom Speicherort (*location*) unabhängig. Im Gegensatz zum URL ist die primäre Aufgabe eines URN die dauerhafte Benennung einer Ressource mittels *identifier*, sogar dann, wenn die mit einem Namen bezeichnete Ressource nicht mehr verfügbar ist. Neben der Dauerhaftigkeit ist eine weitere Anforderung an URNs die, dass URNs weltweit eindeutig sind. (Daigle et al. (1999) und Berners-Lee et al. (2005))

Damit ein URN gültig (*valid*) ist, müssen zwei Bedingungen erfüllt sein (Daigle et al., 1999, 1-2):

- Die Zuweisung einer URN ist kontrolliert und muss unter Anwendung der für einen Namensraum geltenden Regeln geschehen

¹⁰Die Verwendung der Abkürzung RFC weicht im Zusammenhang mit der Dokumentation von Internetprotokollen und -Standards durch *The Internet Engineering Task Force* (IETF) von der im IT Service Management üblichen Belegung mit „Request for Change“ ab.

- Der URN Namensraum (*URN namespace*) selbst ist kontrolliert

Diese Bedingungen haben zur Folge, dass nicht alle syntaktisch korrekten URNs auch valide sind (Daigle et al., 1999, 2).

2.2.6 Information Retrieval und Metadaten

Das Fachgebiet **Information Retrieval** (IR) entstand aus der Notwendigkeit aus wachsenden Beständen an wissenschaftlichen Katalogen die für eine wissenschaftliche Fragestellung relevante Literatur zu finden (Ferber, 2003, 4).

Ferber (2003) nennt als Methoden um Literatur zu finden:

- Befragen von Expertinnen und Experten
- Bücher und Tagungsbände
- Bibliotheksrecherchen
- Suche im World Wide Web

In dieser Arbeit werden die Bibliotheksrecherchen unterstützenden Methoden des IR behandelt.

Information Retrieval Systeme werden eingesetzt, um einen mittels Anfrage formulierten Informationsbedarf mit Datensätzen zu relevanter Literatur abzudecken (Ferber, 2003, 7). Dazu werden nach Ferber (2003) die folgenden Techniken und Verfahren eingesetzt:

- Boolesches Retrieval – basiert im wesentlichen darauf, dass das Auftreten von Fakten (eine oder mehrere Zeichenketten) erfüllt ist oder nicht.
- Zeichenketten, Wörter und Konzepte – dabei wird
 - entweder mittels *computerlinguistischem Ansatz* versucht Terme in Texten als bestimmte Wortformen zu erkennen und auf ihre grammatikalische Grundform (Reduktion auf den Nominativ Singular bei Substantiven und auf den Infinitiv bei Verben) oder ihre Stammform (Lemmatisierung oder Stemming) zurückgeführt,
 - oder mittels der *Lexikografische Grundformenreduktion nach Kuhlen* Wörter nach einem Regelwerk auf ihre formale beziehungsweise lexikografische Grund- oder Stammform zurückgeführt,
 - oder mittels *lexikonbasierter Morphologie-Analyse* den Wortstamm durch Nachschlagen oder Anwendung von Ersetzungsregeln ermittelt

- und anschließend durch *Auflösen von Mehrdeutigkeiten* für mehrdeutige Wörter unter Rückgriff auf den Kontext den „richtigen“ Stamm mit der richtigen Bedeutung ermittelt.
- Klassifikationen und
- Thesauri wie in dieser Arbeit im Abschnitt „Definition von Begriffen im Zusammenhang mit digitale Bibliotheken“ beschrieben.
- Semantische Netze – mittels Ontologien werden Wissensgebiete strukturiert und Objekte aus den Gebieten zugänglich gemacht.
- Vektorraummodell – dabei werden Terme in einem Dokument gewichtet und erst dann für die Beschreibung eines Dokumentinhalts verwendet. Die passendsten Objekte werden dann unter Anwendung eines Ähnlichkeitsmaßes ausgewählt.

Der bereits weiter oben umrissene Datenformatstandard **Open Archives Initiative Protocol for Metadata Harvesting** (OAI-PMH) kann zur Realisierung von **automatisiertem Information Retrieval** herangezogen werden.

Metadaten werden im Rahmen des OAI-PMH Frameworks in Form von *Items*¹¹ mit eigenem *Unique Identifier* ausgetauscht. Dieser *Unique Identifier* ist – seinem Namen entsprechend – eindeutig und unverwechselbar und darf in diesem Zusammenhang nicht einem eindeutigen Ressourcenbezeichner (*Ressource Identifier*) verwechselt werden. Weiters hat er im OAI-PMH Framework die Aufgabe, zusammengehörende Metadatensätze für verschiedenen Formate zu einem *Item* zu markieren. (Lagoze und van de Sompel, 2004)

Unique Identifier spielen eine doppelte Rolle im OAI-PMH Framework (Lagoze und van de Sompel, 2004):

1. Identifier werden zusammen mit jeder **Antwort** (*response*) auf alle Anfragen mit übertragen.
2. Der Identifier wird bei einer **Anfrage** (*request*) benötigt, um Metadatensätze (*records*) anzufordern. Ein OAI-PMH Request wird technisch als HTTP Request realisiert, wobei auf Seite des Repository die Methoden *get* und *post* unterstützt werden müssen.

In Verbindung mit vorgenannten Techniken und Verfahren werden iterative Recherchen in Katalogen möglich.

¹¹Ein Item stellt im OAI-PMH Framework ein Container dar, der Metadaten über verschiedene Formate einer Ressource als Antwort auf eine Anfrage an ein Repository in Form einer statisch gespeicherten oder dynamisch generierten valide well-formed XML-Datei in UTF-8-Kodierung bereithält. Nach Lagoze und van de Sompel (2004) sind Items die konstituierenden Elemente eines Repositories.

2.2.7 Löschung von Metadaten

Um Schutz vor Desinformation, beispielsweise vor Publikationen welche widerlegte Theorien oder widerlegte Informationen veröffentlichen, bieten zu können, muss der informationelle Charakter der Objekte regelmäßig überwacht werden. Somit ändert sich auch die Funktion der einzelnen Bibliotheksobjekte kontinuierlich (Wagner-Döbler, 2000, 287).

Damit diesem Umstand Rechnung getragen werden kann, ist neben der Ergänzung der Metadaten über ein Bibliotheksobjekt auch die Aufgabe („Löschung“) veralteter Metadaten erforderlich. Wenn ein solcher Datensatz nicht mehr länger verfügbar ist, wird er als gelöscht bezeichnet und als solchen markiert. (Lagoze und van de Sompel, 2004)

Im OAI-PMH Framework müssen Repositories deklarieren, wie diese gelöschte Metadatensätze behandeln. Dies muss in jeder Antwort deklariert werden und kann auf dreierlei Arten erfolgen (Lagoze und van de Sompel, 2004):

1. Das Repository liefert *keine* Informationen zu gelöschten Metadatensätzen.
2. Das Repository *kann* Informationen zu gelöschten Metadatensätzen liefern, gewährleistet dies aber nicht auf Dauer oder bei jeder Anfrage.
3. Das Repository liefert *immer* Informationen zu gelöschten Metadatensätzen und wartet den Status „Gelöscht“ auch auf Dauer.

Somit ist der eine Abfrage absetzende Service in der Lage, zu erkennen ob und welche Historie zu einem digitalen Objekt gehört.

2.3 Leistungsmessung und Leistungsindikatoren

Auch digitale Bibliotheken können sich dem Wirtschaftlichkeitsgedanken nicht entziehen. Dies bedeutet, dass mit möglichst wenig Ressourceneinsatz ein größtmöglicher Output anzustreben ist. Um eine effektive Kontrolle und Steuerung zu ermöglichen, ist eine Messung der Leistung der (digitalen) Bibliothek erforderlich.

Im Rahmen dieser Arbeit wird versucht, durch Optimierung der Struktur der Metadaten ein verbesserter Output zur erreichen. Die Inputfaktoren werden dabei konstant gehalten. Deshalb werden an dieser Stelle keine Betrachtungen etwa zu Themen wie Kapitalbedarf (zur Deckung der laufenden Kosten und für Investitionen), Technologieeinsatz, Personal und Ausbildung des Personals oder der betrieblichen Organisation der digitalen Bibliothek oder des Archivs.

An dieser Stelle wird auf die Bestimmung des Outputs eingegangen: dies erfolgt durch Messung von relevanten Größen wie beispielsweise dem finanziellen Ergebnis (Überschuss oder Defizit), Besucherzahlen, Anzahl von Entlehnungen, Anzahl von Veranstaltungen, Gesamtbestand an Bibliotheksobjekten und dergleichen. Dabei wird der Output in zwei

Dimensionen zu messen sein: quantitativ erfassbare Ergebnisse und qualitativ zu erfassende Ergebnisse. Immer interessiert jedoch die (relative) Größe, Position und das Vorhandensein oder Nichtvorhandensein einer relevanten Eigenschaft (Atteslander, 1995, 261).

Für den **quantitativ erfassbaren Output** sind direkt durch abzählen oder vergleichen zahlenmäßig bestimmbar. Beispiele für Einheiten sind Anzahl Bibliotheksobjekte in Stück, Verhältnis von digitalen Bibliotheksobjekten zur Gesamtzahl der Bibliotheksobjekte in Prozent, Umsatz aus Online-Abonnements in Euro, belegte Speicherkapazität eines Repository in Terabyte, Suchanfragen können von den Benutzenden in ihrem Profil abgespeichert werden (ja/nein) usw.

Der **qualitativ erfassbare Output** wie beispielsweise die Servicequalität wird hingegen meist nicht direkt messbar sein. Daher werden Indikatoren (z.B. die Anzahl der negativen Äußerungen über einen Service) verwendet, welche Rückschlüsse auf die outputbezogene Dimensionen (z.B. Zufriedenheit der Benutzenden mit einem Service) zulassen.

Die Dimension der **Servicequalität** für digitale Bibliotheken fasst Waters (1998) wie folgt zusammen: „A [...] measure of service quality takes account of how willingly and how responsively a digital library makes information available to its patron communities.“

2.3.1 *Precision und Recall*

Aus dem Fachbereich des *Information Retrieval* werden die beiden Kriterien *precision* und *recall* übernommen, wobei die Ergebnistreffer einer Einzelabfrage hinsichtlich zweier Aspekte ausgewertet werden (Arms, 2000):

- Erstens besteht das Ergebnis im Idealfall ausschließlich aus relevanten Treffern. **Precision** ist das Verhältnis von relevanten Treffern zu der Gesamtzahl der Ergebnistreffer in Prozent.
- Zweitens werden im Idealfall *alle* relevanten Exemplare als Treffer ausgegeben, die das Suchkriterium erfüllen. **Recall** ist das Verhältnis von den gefundenen relevanten Treffern zur maximal möglichen Gesamtzahl an relevanten Treffern im Rahmen einer bestimmten Abfrage.

Eine zentrale Voraussetzung für die Beurteilung von Retrieval-Ergebnissen besteht darin, dass die „richtige Antwort“ auf eine zu definierende Anfrage bekannt sein muss. Dazu muss im Vorhinein bekannt sein, welche in der Datenbank vorhandenen Objekte zur „richtigen Antwort“ auf die Anfrage „gehören“. Um die Bedeutung hinter „zur richtigen Antwort gehören“ festzulegen, wird eine Relation zwischen einer konkreten Suchanfrage und Objekten eingeführt: die *Relevanz*. Ferber (2003) definiert zwar Relevanz, verweist aber darauf, dass in der Praxis nicht diese Definition benutzt wird, um die Relevanz von Objekten zu bestimmen, sondern dass im Rahmen einer Befragung Experten gebeten werden die jeweilige

Relevanz einzuschätzen (Ferber, 2003). Um den individuellen Messvorgang zu beschleunigen, sind Testkollektionen verfügbar. Diese weisen jedoch nach Ferber (2003) folgende Nachteile auf:

- Die Sammlungen sind im Allgemeinen eher klein
- Dabei ist häufig unklar, wie diese zustande gekommen sind
- Die verfügbaren Sammlungen sind aufgrund des Alters heute nicht mehr repräsentativ
- Allmählich ergibt sich eine Optimierung der Systeme auf die Sammlungen hin
- Die Sammlungen verhalten sich unterschiedlich
- Die Auswahl einer Sammlung für eine Evaluation ist willkürlich, was zur Folge haben kann, dass Sammlungen verwendet werden, bei der – je nach versteckter Intention der Evaluation – die besten oder schlechtesten Ergebnisse zu erwarten sind.

Eine weitere Hürde ist, dass die Methode der Bestimmung von *precision* und *recall* ihren Ursprung in der Zeit hat, als Einzelabfragen über große Datenmengen den primären Anwendungsfall darstellten. Heute sind Suchmethoden interaktiv und ermöglichen eine iterative Suche mit folgendem Ablauf: initiale Suche, auswerten der Ergebnisse, wiederholte verfeinerte Suche mittels modifizierter oder einschränkender Abfrage. Nach Arms (2000) sei heute für die Messung von *precision* und *recall* der gesamte Vorgang dieser schrittweisen Suche – auch als *session* bezeichnet – heranzuziehen. Dieser Auffassung kann der Autor dieser Arbeit nicht folgen, da anstelle der Qualität des bereitgestellten Services jene des Gesamtsystems inklusive dem menschlichen Einfluss („with a human in the loop“) gemessen wird. So würden bei gleich bleibender Servicequalität unterschiedliche Werte für die Leistungsindikatoren *precision* und *recall* in Abhängigkeit vom Vorwissen der bedienenden Person hinsichtlich Themengebiet und Suchstrategien (z.B. Kenntnis der effektiven Verwendung von Thesauri) ermittelt.

2.3.2 Integrität

Die Integrität (*integrity*) der digitalen Objekte ist eine wesentliche Qualitätsanforderung und wird nach Waters (1998) und Lawrence et al. (2000) durch folgende Dimensionen bestimmt:

- Inhalt (*content*)
- Beständigkeit (*fixity*) der Bitfolge der Dateien oder des Bitstroms
- Bezug (*reference*) auf Dateien in dem Sinn, dass Pfadangaben gültig sind und insbesondere die Dateinamenerweiterungen bestehen (bleiben)

- Herkunft (*provenance*) und Historie eines digitalen Objekts hinsichtlich Zulieferung, Manipulationen oder Migration
- Kontext (*context*) in der Form von Zusammenhängen und Interaktionen mit anderen Dateien oder Objekten, Hardware- und Softwareabhängigkeiten oder Verknüpfungen mit Metadaten und Scripts

Die Hauptschwierigkeit besteht in der Messung dieser Kriterien. Für die Praxis gilt es Leistungsindikatoren zu finden, welche vorstehenden Dimensionen einerseits gut beschreiben indem sie für sie charakteristisch sind und andererseits auch mit vertretbarem Aufwand messbar sind. Als geeignete Methode bietet sich hierzu am ehesten eine Serie von wiederkehrenden Audits an.

2.3.3 Kriterien aus dem Service Management

Wie bereits festgestellt, ist die Qualität des Services „Recherche“ ein wesentlicher Aspekt und das Qualitätsniveau das Resultat des Zusammenspiels von Technologie, Prozessen und den sich in Form von Rollen einbringenden Personen. Durch diese umfassendere Sichtweise des Service Managements ergeben sich noch weitere Anforderungen: Neben der inhaltlichen Qualität und der Qualität von Treffern als Antwort auf eine Anfrage ist die Qualität des Services aus der Sicht der Benutzenden noch um die in diesem Unterabschnitt kurz zusammengefassten Aspekte zu ergänzen.

Aus dem Service Management sind noch weitere Kriterien für eine Service bekannt. Assirati (2001a) bezeichnet diese wie folgt:

- **Servicezeit:** Während dieser vereinbarten Betriebszeit des Services (z.B. 24h je Tag, 7 Tage in der Woche) werden Störungen registriert (gezählt, die Auswirkungen erfasst und in der Dauer gemessen). Falls die garantierte Servicezeit nicht rund um die Uhr sondern beschränkt ist – beispielsweise auf die Zeit von 8 Uhr bis 18 Uhr – können Servicezeiterweiterungen möglich sein. Anhand der Zählung dieser Anforderungen um Servicezeiterweiterungen und Auswertung hinsichtlich Benutzer bzw. Benutzergruppen können Rückschlüsse auf Fehlanpassungen von Angebot und Nachfrage gezogen werden.
- **Verfügbarkeit:** Dabei ist festzustellen, in wie vielen Prozent der vereinbarten Servicezeit der Service tatsächlich verfügbar ist. Aus praktischen Gründen geschieht dies in der Regel über den Umweg der Messung der Zeiträume der Nichtverfügbarkeit des Services. Andererseits ist es auch die Nichtverfügbarkeit, welche Einbußen im Geschäftsprozess dessen Teilprozess die Informationsbeschaffung darstellt bewirken: Forscher verlieren unnötig Zeit mit wiederholten Suchanfragen. Letztendlich wäre diese Beeinträchtigung ermittelbar und könnte mit Kostensätzen hoch gerechnet werden. Für die Praxis ist für jeden einzelnen Service zu definieren, wann Nichtverfügbarkeit vorliegt.

- **Zuverlässigkeit:** Sie ist ein Maß für die Häufigkeit der Nichtverfügbarkeiten unabhängig von der Dauer derselben und kann relativ einfach durch abzählen der Unterbrechungen ermittelt werden oder unter Verwendung von statistischen Größen wie *Mean Time Between Failures* (MTBF) oder *Mean Time Between System Incidents* (MTBSI) angegeben werden.
- **Support für die Benutzenden:** Die Supportzeiten können von den Servicezeiten abweichen: Per World Wide Web können beispielsweise über die gesamten 24 Stunden am Tag kostenpflichtige Downloads von Artikeln durchgeführt werden. Bei Störungen im Zuge dieser Aktivität kann das Serviceteam (z.B. am *Service Desk*) jedoch nur während der erweiterten Bürozeiten von 8 Uhr bis 19 Uhr erreicht und um Hilfeleistung gebeten werden. Messbare Größen stellen hier beispielsweise die Anzahl der Störungen und Serviceanforderungen, die Antwortzeiten, die Lösungs- und Wiederherstellungszeiten die Störungen eines Service betreffend oder die Erstlösungsrate – wie viele Störungen während des Erstkontakts mit der Supportorganisation behoben werden können im Verhältnis zur Gesamtzahl der Störungen – dar.
- **Durchsatz:** Damit werden Dimensionen zusammengefasst, welche das insgesamt schwierig zu definierende und zu messende Thema der *Performance* umfassen: Bandbreiten für die Datenübertragung, Datenvolumen je Zeiteinheit (z.B. Monat), Anzahl Suchanfragen je Stunde, Anzahl gleichzeitige Benutzende, usw.
- **Antwortzeit:** Die Zeit zwischen dem Absetzen der Anfrage und der Übergabe der Antwort an definiertem Ort wird meist in aggregiertem Zustand als durchschnittliche und maximale Antwortzeit auf Suchanfragen und abgesetzten Thesauri- und Katalogrecherchen festgehalten.
- **Bearbeitung von Veränderungen (*Changes*):** Ein definiertes Qualitätsniveau hinsichtlich dem Veränderungsprozess ist primär für die bibliotheksinternen Kunden relevant – es sei denn, dass externe Kunden (Bibliotheksnutzer oder Verlage) ebenfalls die Möglichkeit hätten, Veränderungsanträge (*Requests for Change*, RFCs) zu stellen. Diese Gruppen von beispielsweise internen Experten zur Katalogisierung sind unter Umständen berechtigt Veränderungsanträge zu stellen um die Administrationsmasken für Katalogeinträge neuen Anforderungen anpassen zu lassen. Für diese Gruppe ist nun durchaus von Interesse, wie lange das Veränderungsverfahren durchschnittlich bzw. maximal dauert, wie häufig Changes nicht dem RFC entsprechend implementiert wurden, wie häufig Changes abgewiesen wurden, mit welchen Prioritäten Changes versehen wurden oder insbesondere wie viele Emergency-Changes (es ist bereits ein Schaden entstanden) im Verhältnis zur Gesamtzahl der Changes bearbeitet wurden.
- **Service Continuity:** Dabei geht es darum, über Pläne zur Wiederherstellung von ausgewählten Services inklusive Repositories im Katastrophenfall – gegebenenfalls unter angepassten Servicezielen – zu verfügen und festgelegt zu haben, wie diese

ausgelöst werden können. Leistungsindikatoren können sein: Vorhandensein und Reifegrad des *IT Service Continuity Management Prozesses*, Anzahl der abgedeckten Services, Anzahl der geschulten Mitarbeitenden je Jahr, Anzahl Tests der Wiederherstellungspläne, Anzahl der festgestellten Mängel in der Planung, usw.

- **Security:** Im Rahmen von Sicherheitsüberlegungen müssen die drei Phänomene Vertraulichkeit, Integrität (siehe im vorhergehenden Unterabschnitt) und Verfügbarkeit der auf der Ebene der Objekte und Daten erfasst werden. Dabei müssen Leistungsindikatoren basierend auf präventiven Maßnahmen sowie aus der Störungsregistrierung – vornehmlich Daten zur Häufigkeit und Art der erkannten *Security Incidents* – ausgewertet werden.
- **Verrechnung:** Häufig sind Services digitaler Bibliotheken, etwa der Zugang zu den Katalog- oder Abstractdatenbanken kostenpflichtig. Auch können einzelne Artikel-downloads oder Zusatzleistungen eine Verrechnung zur Folge haben. Dabei sind für die Servicequalität die Benutzerfreundlichkeit der dazu erforderlichen Registrierung, die Anzahl und die Sicherheitsstufen verschiedener Zahlungsmöglichkeiten und Aspekte der Transparenz der Verrechnung – für welche Services muss wann wie viel bezahlt werden – von Interesse. Leistungsindikatoren können beispielsweise Verhältnis von korrekterweise beanstandeten Rechnungen zur Gesamtzahl der Rechnungen sowie die Zahl verschiedener Möglichkeiten der Verrechnung (*prepaid*, monatlich, usw.) und Zahlung (Kreditkarte, offene Rechnung, Online-Zahlungsform, usw.) sein.

Wenn auch diese Aspekte bis auf den der *Security* und *Service Continuity* keinen unmittelbaren Einfluss auf Inhalte der Services einer digitalen Bibliothek haben, beeinflussen sie stark das wahrgenommene Gesamtbild der Services und sind somit für die Qualitätsbeurteilung des Outputs einer digitalen Bibliothek von Relevanz.

In weiterer Folge soll auf den Teilaspekt der Sicherstellung der Qualität von Metadaten näher eingegangen werden.

2.4 Sicherstellung der Qualität von Metadaten

In diesem Abschnitt werden im Hinblick auf qualitativ hochwertige Metadatenbestände einerseits Mindeststandards für Metadaten in Form von Metadatenelementen und andererseits Rahmenbedingungen für den Erfassungsprozess zusammengestellt.

2.4.1 Metadaten Mindeststandards

Die hier gelisteten Elemente können in einen Anforderungskatalog für den Entwurf bzw. die Auswahl von Systemen für digitale Bibliotheken übernommen werden. Mindeststandards sollen auch die weitere Nutzbarkeit von *master images* sicherstellen. Zusätzlich soll ein Set

an Metadaten mit dem Ziel der Identifikation des Objekts zeitgleich mit der Erstellung oder Übernahme der digitalen Objekte erstellt werden. (Puglia et al., 2004)

Puglia et al. (2004) beschäftigten sich eingehend mit der Erzeugung von digitalen Objekten durch Einscannen von in physikalischer Form vorliegenden Bibliotheksobjekten und legten den Minimalstandard für deskriptive, strukturelle und administrative Metadaten für die Ebene der Datei – gegebenenfalls in logischer Hinsicht für das digitale Objekt – wie folgt fest:

- *Primary Identifier* als *unique identifier*
- *Secondary Identifier* stellen andere Schlüssel zur Referenzierung auf das Original dar
- *Record Group Identifier* zur Zusammenfassung von Datensatzgruppen
- *Record Group Descriptor* als Titel der Datensatzgruppe
- *Title* des Objekts
- *Series* als Titel der Serie
- *Location* und *Box* des Originals
- *Structural view* oder *page (sequence)* als Beschreibung des sichtbaren Inhalts, Seitennummer und/oder Dateinummer
- *Publisher* vermerkt den Inhaber des Copyrights des generierten digitalen Objekts
- *Text (Source)* mit Angaben zum ursprünglichen Medium und der ursprünglichen Herstellungsart
- *Film (Source)*, der für die Aufnahme verwendet wurde mit Angaben zum ursprünglichen Medium, Format, Farbmodus und dem ursprünglichen Herstellungszeitpunkt
- *Photo Print (Source)* mit Angaben zum ursprünglichen Medium und Farbmodus
- *Digital Photo (Source)* mit Angaben zum ursprünglichen Format, Farbmodus, Einstellungen, Kompressionsgrad
- Bei doppelseitigen Originalen die Angabe „Vorderseite“ bzw. „Rückseite“
- Bei mehrseitigen Originalen die Angaben „Deckblatt“, „Seite i“, „Seite ii“, „Seite 1“, „Seite 2“, „Doppelseite“, usw.
- *Component part* wird ausschließlich für eine durchlaufende Seitennummerierung verwendet, welche jedoch nicht den gedruckten Seitennummern entsprechen müssen

Des Weiteren werden mindestens die folgenden technische Metadaten mit deutlichem Fokus auf die Digitalisierung von physischen Bibliotheksobjekten durch Scannen aufgezeichnet werden (Puglia et al., 2004, 20):

- *Copy* als Rolle der Abbildung: z.B. Master, optimiert für den Ausdruck oder für den Versand geeignet
- *File format type/Version*, beispielsweise „TIFF“ oder „JPEG“
- *Location* als Pfadangabe innerhalb des Repository
- *Image creation date*
- *Photographer/Operator* als Ersteller/in oder Bediener/in des Scanners
- *Compression Type/Level*
- *Color Mode*, z.B. „RGB“ oder „256 Graustufen“
- *Gamma Correction*
- *Color Calibration*, z.B. „AdobeRGB 1998“ für RGB
- *Pixel Array* als Breite mal Höhe in Pixeln
- *Spatial Resolution* in dpi (z.B. 300)
- *Image quality* als Angaben, welche in weiterer Folge die Abschätzung des Werts der Aufnahme bzw. des Scans ermöglichen
- *File Name*
- *Source Information* mit der Angabe ob vom Original eingescannt wurde oder ob eine Zwischenstufe (z.B. Mikrofilm) verwendet wurde

2.4.2 Generelle Rahmenbedingungen für den Erfassungsprozess

Im Zuge der Vorbereitungen zur Errichtung oder Erweiterung einer digitalen Bibliothek müssen Antworten auf die folgende Fragen gefunden werden (Puglia et al., 2004, 16-18):

- Wer wird Metadaten erfassen?
- Wie werden Metadaten erfasst?
- Wann werden die Metadaten erfasst?
- Wo werden die Metadaten gespeichert?

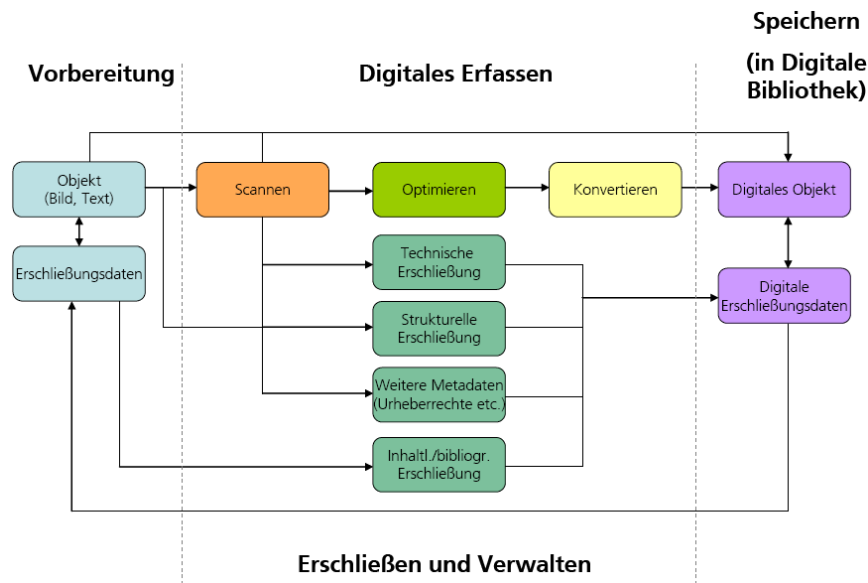


Abbildung 2.1: Erschließung von digitalem Kulturgut (Puglia et al., 2004, 89)

- Wie werden die Metadaten gespeichert?
- Werden die Metadaten mit anderen Systemen ausgetauscht? Mit welchen Systemen basierend auf welchen Standards?
- Gibt es bereits bestehende Übereinkünfte zu Konventionen und Terminologie? Borowski (2007) empfehlen vor allem im Zuge der Erschließung in Museen besonders auf die einheitliche Verwendung von Regelvokabularen und Standards zu achten.

Ein wesentlicher Einfluss auf die Qualität des digitalen Objekts hat die Vorlage. Wenn nicht ein digitales Objekt übernommen werden kann oder vom Original digitalisiert werden soll, muss in der Regel von einer Zwischenstufe – beispielsweise von Mikrofilmvorlagen – digitalisiert werden. Dabei muss diese eine entsprechende Qualität aufweisen (Borowski, 2007, 21). Weiters ist dabei zu beachten, dass dieser Umstand in den Metadaten entsprechenden Niederschlag findet.

Während der Arbeitsablauf für die Aufnahme von bereits in digitaler Form vorliegenden Objekten in eine digitale Bibliothek häufig bereits individuell festgelegt ist, ist der Workflow für die Aufnahme von erst zu digitalisierenden Objekten komplexer. Abbildung 2.1 zeigt deutlich die Abhängigkeit der Metadaten des digitalen Objekts von den Metadaten über das originale, physische Objekt. Insbesondere muss zu jeder Zeit sichergestellt sein, dass die Relation zwischen dem originalen Objekt und dem digitalen Objekt erhalten bleibt.

Zum Workflow bei Digitalisierungsprojekten empfiehlt Borowski (2007) – neben anderen Themen welche eher auf Länder- oder nationaler Ebene zu beantworten wären und daher hier unberücksichtigt bleiben – die folgenden relevanten Themen zu beantworten:

- Digitalisierungsstrategie formulieren.

- Im Austausch mit anderen digitalen Bibliotheken Experten kontaktieren, externe Dienstleister evaluieren und Erfahrungen austauschen um Know-how aufzubauen.
- Inhaltliche, technische und finanzielle Digitalisierungskonzepte erstellen inklusive der Planung der Langzeitarchivierung.
- Workflow planen indem einzelne Arbeitsschritte geplant und aufeinander abgestimmt werden. Bei besonders wertvollen Originalen soll der Digitalisierungsvorgang mit einem Konservator abgesprochen und möglichst vorlagenschonend durchgeführt werden.
- Weiters sind insbesondere im Archiv- und Museumsbereich im Vorfeld von Digitalisierungen logistische Aspekte zu klären. Dazu gehören Objekttransport, Bereitstellung von Räumlichkeiten bis hin zur Objektversicherung.
- Qualität unter Einsatz eines Vier-Augen-Prinzips laufend kontrollieren.
- Die Auswahl von geeigneten Speichersystemen und Techniken zur Datensicherung sollen größtmöglichen Schutz der digitalen Objekte und der Metadaten sicherstellen. Bei großen Projekten empfiehlt Borowski (2007) die Zusammenarbeit mit Rechenzentren an.
- Archiv mit stark eingeschränkten Zugriffsberechtigungen für ein *Fallback*-Szenario aufbauen.

Für Massendigitalisierungen sollte zusätzlich der Einsatz von Scanrobotern evaluiert werden und die Auslastung dieser mit anderen Kultureinrichtungen koordiniert werden, um eine akzeptable Auslastung sicherstellen zu können. (Borowski, 2007, 21)

Borowski (2007) empfiehlt für Archive die Findbücher digital aufzubereiten und online verfügbar zu machen. Eine Digitalisierung des Archivguts sei nachrangig, da zuerst die Strukturierung der Datenmenge geklärt werden müsse.

Kapitel 3

Fallbeispiel: Die *Configuration Management Database* als zentrale digitale Bibliothek für *IT Service Management* Prozesse

Wie später im Abschnitt „Configuration Management Database (CMDB)“ gezeigt wird, kann die CMDB als digitale Bibliothek betrachtet werden. Diese besteht im wesentlichen aus einer Metadatenansammlung welche meist mittels Datenbanken verwaltet wird und aus einem oder mehreren Repositories, welche Dokumentationen, Software und weitere Datenbanken (z.B. über Störungen oder Changes) für den Abruf bereit halten.

Die Erfassung von Metadaten und die Ablage von Bibliotheksobjekten in die Repositories erfolgt – ähnlich den Vorgängen in Bibliotheken – regelbasiert und in mehr oder weniger stark strukturierter Weise. Im Rahmen der Verwendung im *IT Service Management* (ITSM), dem hauptsächlichen Anwendungsgebiet einer CMDB, werden diese Strukturen und Regeln in Form von Prozessen definiert.

Es wird daher hilfreich sein, zuerst einen Überblick über ITSM Frameworks und die darin definierten Prozesse zu erlangen. Um einen Einblick in dieses durchaus umfangreiche Themengebiet zu gewinnen, wird in dieser Arbeit zunächst

- das Schlagwort *IT Service Management* kurz umrissen (Abschnitt 3.1 ab Seite 49), dann
- der Begriff *IT Infrastructure Library* (ITIL) als die Grundlage mancher ITSM Frameworks erläutert und die durch ITIL-Prozesse abgedeckten Gegenstandsbereiche genannt (Abschnitt 3.2.1 ab Seite 50), dann
- ITSM Frameworks und die Entwicklung von ITIL bis hin zum neuen Standard ISO 20000 kurz vorgestellt (Abschnitt 3.2.2 ab Seite 52 bis inklusive Abschnitt 3.2.6) und die abgedeckten Gegenstandsbereiche spezifiziert sowie

- der für viele ITSM-Frameworks zentrale Prozess des *Configuration Management* (Abschnitt 3.2.7 ab Seite 56) beschrieben.

In weiterer Folge wird als nächster Schritt der Verfeinerung

- auf die Betrachtung der CMDB als digitale Bibliothek eingegangen (Abschnitt 3.3.1 ab Seite 58),
- der konzeptionelle Umfang der CMDB abgesteckt (Abschnitt 3.3.2 ab Seite 62),
- die konkreten Informationsbedürfnisse der einzelnen Gegenstandsbereiche an die CMDB zusammengestellt (Abschnitt 3.4 ab Seite 67) um anschließend
- diese Informationsbedürfnisse in Metdatenelemente der CMDB umgesetzt (Abschnitt 3.5 ab Seite 76) und um abschließend
- die Rahmenbedingungen für den Erfassungs- beziehungsweise Aktualisierungsprozess zu konkretisieren (Abschnitt 3.6 ab Seite 90).

Insgesamt soll dieses Kapitel dazu beitragen, die zentrale Bedeutung der CMDB für das ITSM herauszustreichen und einen Zusammenhang zwischen der Welt der CMDB und der der digitalen Bibliothek herzustellen sowie den zusätzlichen Nutzen aus der „bibliothekarischen“ Sichtweise für die CMDB zu konkretisieren.

3.1 IT Service Management

IT Service Management setzt beim zu liefernden IT Service an und betrachtet diesen aus der Sicht des Kunden über den gesamten Lebenszyklus des Services. Im Gegensatz zur traditionellen Sichtweise von IT Management – wonach es primär um das effektive und effiziente Management von Hard- und Softwarekomponenten geht – werden im Rahmen des IT Service Managements **Ziele auf der Ebene der IT Services** definiert und deren Erreichung bewirkt. Es geht somit *nicht nur um die reine Erstellung* beispielsweise einer Kundendatenbank für den Vertrieb sondern auch um die *laufende und optimale Anpassung an Kundenwünsche* bis zur Abschaltung des Services. Wichtig ist dabei der Kontakt mit dem Kunden und den Benutzenden, die bedarfsgerechte Dimensionierung der Verfügbarkeit und der Kapazitätsanforderungen, die (nachhaltige) Behebung von Störungen und die Reaktion im Katastrophenfall, die prozessgeleiteten Veränderungen am Service und der dazu erforderlichen Komponenten und immer auch die Einhaltung vereinbarter Kostenrahmen für diese Leistungen während der gesamten Lebensdauer des Services.

Vorteile durch den Einsatz von prozessgesteuertem IT Service Management sind im wesentlichen (Assirati, 2001b, 17):

- verbesserte Qualität des IT Services,

- klarerer Überblick über die aktuelle Leistungsfähigkeit der IT Organisation,
- besserer Informationsstand über aktuelle Services und Informationen darüber, wo Veränderungen am effektivsten sind,
- verbesserte Kundenzufriedenheit, da die Serviceprovider die Erwartungen an die Services besser kennen und erfüllen können sowie
- verbesserte Sicherheit, Genauigkeit, Geschwindigkeit oder Verfügbarkeit des Services, entsprechend dem vereinbarten Servicelevel.

Diese Vorteile sind in ihrer Ausprägung stark von der jeweiligen Organisation und den dort gesetzten Schwerpunkten bzw. Anforderungen an die IT abhängig.

3.2 ITIL und für das IT Service Management relevante Standards

In diesem Abschnitt werden überblicksweise das Best Practice Modell nach *IT Infrastructure Library* (ITIL) und weitere für das IT Service Management relevante Standards vorgestellt. Die Aufstellung nennt die nach Walter und Krcmar (2006) wichtigsten Frameworks und erhebt keinen Anspruch auf Vollständigkeit.

3.2.1 IT Infrastructure Library (ITIL)

ITIL steht seit den 1980-er Jahren¹ für *IT Infrastructure Library* und stellt Empfehlungen in Form von Best Practices im Rahmen eines Prozessmodells zum Management von IT zur Verfügung. In diesem *public domain framework* finden sich weniger bahnbrechende neue Konzepte als vielmehr bewährte und in einem größeren internationalen Kontext² als führende Praktiken anerkannte Vorgehensweisen wieder. ITIL unterstützt durch definierte Prozesse beispielsweise *total quality frameworks* wie die European Foundation for Quality Management (EFQM) oder Qualitätssysteme nach *ISO 9000* (Assirati, 2001b, 2).

ITIL umfasst die folgenden fünf Hauptelemente (Assirati, 2001b, 4-5):

- **Deliver IT Services**, oft auch als mit „Service Delivery“ bezeichnete Gruppe von Prozessen, mit den Prozessen *Capacity Management*, *Financial Management for IT Services*, *Availability Management*, *Service Level Management*, *IT Service Continuity*

¹Ausgehend von Großbritannien, wo es ursprünglich als Leitfaden für Regierungsorganisationen erstellt wurde, ist ITIL heute weltweit bekannt und implementiert (Assirati, 2001b, 1).

²Der Inhalt der Best Practices wird beispielsweise für den Bereich des Service Managements im Rahmen des IT Service Management Forums (ITSMF) festgehalten und laufend aktualisiert. Weitere Informationen dazu finden Interessierte beispielsweise unter www.itsmf.de.

Management und *Customer Relationship Management*. Das zuletzt genannte Thema findet sich jedoch nicht explizit in den Referenzwerken „Service Support“ bzw. „Service Delivery“ oder in den Prüfungsinhalten zu „ITIL Foundation“ Zertifizierungsprüfungen wieder.

- **Support IT Services**, oft auch als mit „Service Support“ bezeichnete Gruppe von Prozessen, mit der Funktion *Service Desk* und den Prozessen *Incident Management*, *Problem Management*, *Configuration Management*, *Change Management* und *Release Management*.
- **Managing Applications** als Prozessmodell mit dem Schwerpunkt auf der Entwicklung von Anwendungen mit einer Schnittstelle zum diesbezüglichen Projektmanagement.
- **Manage the Infrastructure** mit den Themen *Network Service Management*, *Operations Management*, *Management of Local Processors*, *Computer Installation and Acceptance* und *Systems Management*.
- **The Business Perspective** mit den Themen *Business Continuity Management*, *Partnership and Outsourcing*, *Surviving Change* und *Transformation of business practice through radical Change*.

Thematisch werden die beiden Hauptbereiche „Deliver IT Services“ und „Support IT Services“ mit dem Begriff **Service Management** zusammengefasst. Im Rahmen dieser Arbeit ist der *Configuration Management Prozess* aus dem Hauptelement „Support IT Services“ von Relevanz.

Die *Best Practices* nach ITIL in der aktuellen Version „V2“ aus dem Jahr 2000 erfahren derzeit eine Überarbeitung und resultieren in der „V3“ auch als *ITIL Refresh* bezeichnet, welche voraussichtlich im Herbst 2007 veröffentlicht werden soll. Die Motivation hinter diesem Entwicklungsschritt sind die Anpassung an sich ändernde Anforderungen der ITIL-Anwender im internationalen Umfeld und die leichtere Umsetzbarkeit in kleineren Organisationen (The APM Group Limited (APMG) (Hrsg.), 2005). Neben inhaltliche Ergänzungen werden neue Prozesse eingeführt – der Aspekt des Designs von Services war bisher kaum berücksichtigt – und als wesentlichster Aspekt die Entwicklung weg vom Framework hin zu einem weiterhin auf Prozessen basierenden Lifecyclemodell vorgenommen (Hainschink, 2007). Ein wesentlicher Faktor stellt dar, dass ITIL V3 weiterhin eng mit dem Standard ISO 20000 abgestimmt bleibt (Glenfis AG (Hrsg.), 2007). Spekulation bleibt jedoch, ob sich dadurch auch Nachteile von ITIL V2 – z.B. fehlende übergeordnete Prozesse, nur in Ansätzen berücksichtigtes In-/Outsourcing und mangelhafte Schnittstellenbeschreibungen – ausgemerzt werden können.

3.2.2 Standards des *British Standards Institute* (BSI)

Der 1995 veröffentlichte **PD 0005** *Code of Practice for IT Service Management* des *British Standards Institute* (BSI) basiert auf den ITIL Prinzipien, stellt aber noch kein Prozessmodell sondern vorerst nur eine kohärente, bildhafte Beschreibung der relevanten Themen dar (Assirati, 2001b) und gliedert sich in die folgenden Abschnitte:

- **Service Design & Management Processes** mit den Themen *Security Management, Capacity Management, Financial Management, Availability & Contingency Management, Service Reporting* und *Service Level Management*.
- **Release Processes** mit dem Thema *Release Management*.
- **Supplier Processes** mit den Themen *Customer Relationship Management* und *Supplier Management*.
- **Managing Applications**
- **Resolution Processes** mit den Themen *Incident Management* und *Problem Management*.

ITIL, wie es heute in der Version aus dem Jahr 2000 vorliegt, kann als stetige Erweiterung dieses BSI-Standards betrachtet werden.

Das Arbeitshandbuch zur Selbstbewertung **PD 0015 IT Service Management Self-Assessment Workbook** des BSI liefert einen Fragenkatalog, mit Hilfe dessen eine Einschätzung einer Organisation anhand der Anforderungen der BS 15000 durchgeführt werden kann. (HiSolutions AG (Hrsg.), 2005, 8)

Der 2000 veröffentlichte Standard **BS 15000 Specification for IT Service Management** des BSI baut auf den Dokumenten PD 0005 und PD 0015 auf, definiert eine Reihe zusammenhängender Prozesse und dokumentiert auf diese Weise die *Best Practices* für das IT Service Management für interne IT Organisationen, für IT Outsourcing und für externe IT Dienstleister. Teil 1 (*Specification for Service-Management*) besteht aus den folgenden Abschnitten (BS15000 Associates Group (Hrsg.), 2005):

- Scope
- Terms and Definitions
- Requirements for a Management System
- Planning and Implementing Service Management
- Planning and Implementing New or Changed Services
- The Service Delivery Process

- Relationship Processes
- Resolution Processes
- Release Process
- Control Processes

Teil 2 (*Code of Practice for Service-Management*) dokumentiert die zu Teil 1 gehörenden Empfehlungen zur Etablierung von Service Management. (HiSolutions AG (Hrsg.), 2005, 7)

Weiters gehören die bereits erwähnten Dokumente PD 0005 und PD 0015 als Ergänzungsdokumente zum BS 15000 Standard dazu. (HiSolutions AG (Hrsg.), 2005, 7-8)

3.2.3 Control Objectives for Information- and related Technology (CobIT)

Das Kennzahlensystem wurde von der *Information Systems Audit and Control Association* (ISACA) entwickelt und liefert Vorgaben zur Prüfung von IT Management durch die Etablierung von 34 Kontrollzielen auf Managementebene, wovon 13 direkt auf ITIL Prozessen Bezug nehmen (Turbitt, 2006, 3). In CobIT werden zwar Configuration Management und Assetmanagement die Infrastruktur betreffend abgedeckt, der Bereich Softwareassetmanagement hingegen nicht. (Walter und Krömer, 2006, 4)

3.2.4 Microsoft Operations Framework (MOF)

Das MOF basiert auf ITIL und wird primär im Umfeld von Lösungen der Firma Microsoft angewendet. Keine Abdeckung erfahren im MOF die folgenden Gegenstandsbereiche (Walter und Krömer, 2006, 5):

- Business Perspective
- Planning to Implement Service Management
- Softwareassetmanagement – auch bei MOF wird jedoch Configuration Management und Infrastrukturassetmanagement geregelt

Zusätzlich werden im MOF auch Wissensgebiete und Fähigkeiten aufgelistet, welche erforderlich sind, um eine Team-Rolle ausfüllen zu können und geht somit im Bereich der Rollenverteilung und der Qualifikation der Mitarbeitenden in der IT Organisation deutlich tiefer als ITIL. (Walter und Krömer, 2006, 7)

3.2.5 Enhanced Telecom Operations Map (eTOM)

Das Framework eTOM wurde vom *Tele Management Forum* als branchenspezifisches Referenzmodell für die Telekommunikationsindustrie entwickelt und soll auch zur Spezifikation von Informationssystemen herangezogen werden können. (Walter und Krcmar, 2006, 4)

Gegenüber den auf IT Service Management hin entwickelten Modellen bietet eTOM keine Abdeckung der nachstehenden Gegenstandsbereiche (Walter und Krcmar, 2006, 5):

- Service Desk (Funktion)
- Business Perspective
- Planning to Implement Service Management
- Softwareassetmanagement

3.2.6 Der *ISO 20000* Standard

Basierend auf dem britischen BS 15000-2 Standard³ stellt die ISO 20000 den *ersten weltweit gültigen* (Lanham, 2005) Standard dar, der Anforderungen an interne oder externe IT Dienstleister im Hinblick auf ein professionelles, prozessorientiertes IT Service Management festlegt.

Darüber hinaus ist eine Zertifizierung möglich, die die Konformität mit den Anforderungen der ISO/IEC 20000 bestätigt⁴. (Dugmore, 2005) Das erste in Österreich nach ISO 20000 zertifizierte Unternehmen „Logic4BIZ“ wurde im März 2007 zertifiziert und ist aktuell mit 10 Mitarbeitenden ein eher kleines KMU (Presstext Austria (Hrsg.), 2007). Turbitt (2006) listen die Gründe für eine ISO 20000 Zertifizierung wie folgt auf:

- Branchen, in welchen die Qualität von IT Services entscheidend ist
- Organisationen, welche *Managed Services* oder IT Outsourcing Leistungen anbieten
- Zukünftig kann es durchaus möglich sein, dass Wirtschaftsprüfer nicht nur etablierte IT Management Prozesse sondern auch das Vorhandensein von diesbezüglichen Zertifikaten ihrer Klienten überprüfen werden.

³Die Abweichungen sind geringfügig (ISO, 2007a) und werden in der die Veröffentlichung des Standards anzeigende Presseaussendung mit einer besseren Anwendung im internationalen Einsatz begründet. (Lanham, 2005)

⁴Die ISO 20000 Zertifizierung muss durch eine autorisierte Organisation (Registered Certification Bodies (RCBs)) erfolgen.

Charakteristisch für die ISO 20000 ist eine im Vergleich zu ITIL stärkere Ausrichtung am eigentliche Geschäft der Organisation (Turbitt, 2006, 7) sowie die stärkere Betonung des „geschlossenen Managementzyklus“ (Bundesministerium des Inneren (Hrsg., 2006, 20) basierend auf dem Qualitätskreislauf von W. Edwards Deming mit den zyklisch wiederkehrenden Schritten Plan, Do, Check und Act (Vorbach, 2006, 3).

Inhaltlich baut die ISO 20000 auf dem Prozessmodell für IT Service Management nach ITIL auf und unterstützt dieses vollauf (Lanham, 2005). Abweichend von ITIL fehlt in der ISO die Beschreibung der Funktion „Service Desk“, der ITIL-Prozess „Security Management“ findet sich im ISO-Prozess „Information Security Management“ wieder, die beiden ITIL Prozesse „Availability Management“ und „IT Service Continuity Management“ wurden in einen einzigen ISO-Prozess „Availability und Service Continuity Management“ zusammengeführt. Nach Ergänzung durch zusätzliche Anforderungen und Prozesse ergibt sich die inhaltliche Struktur des Standards ISO 20000 wie folgt (nach BS15000 Associates Group (Hrsg.) (2005), ergänzt durch Bundesministerium des Inneren (Hrsg. (2006)):

- **Gegenstand** (*Scope*)
- **Begriffe und Definitionen** im IT Service Management
- **Anforderungen an ein Management System** mit den Themen Managementverantwortung, Dokumentationsanforderungen, Kompetenz, Bewusstsein und Training
- **Planung und Implementierung von Service Management** mit den Themen Service Management planen, Service Management implementieren, überwachen, messen und begutachten sowie Kontinuierliche Verbesserung
- **Planung und Implementierung von neuen oder geänderten Services**
- Die **Service Delivery Prozesse** Service Level Management, Service Reporting, Availability und Service Continuity Management, Budgeting und Accounting für IT-Services, Capacity Management und Information Security Management
- **Relationship Prozesse** mit den Themen Business Relationship Management und Supplier Management
- Die **Resolution Prozesse** Incident Management und Problem Management
- Die **Control Prozesse** Configuration Management und Change Management
- Der **Release Prozess** Release Management

Der ISO 20000 Standard wurde in zwei Dokumenten veröffentlicht (Bundesministerium des Inneren (Hrsg., 2006, 19):

- Das **Dokument ISO 20000-1:2005** fördert die Verwendung eines integrierten prozessorientierten Ansatzes zur Serviceerbringung (Turbitt, 2006, 5) und beinhaltet die

eigentliche Spezifikation des Standards (ISO, 2007a) in Form von „zu erfüllenden Mindestanforderungen“ (Bundesministerium des Inneren (Hrsg., 2006, 19).

- Das **Dokument ISO 20000-2:2005** liefert Umsetzungshinweise und zusätzliche Anforderungen (ISO (2007b) und (Bundesministerium des Inneren (Hrsg., 2006, 22)), beschreibt die *Best Practices* für Service Management im Rahmen der ISO 20000 Part 1 (Lanham, 2005) und erweitert somit die Mindestanforderungen aus ISO 20000-1:2005 um Anforderungen, „an denen sich die IT Organisation orientieren sollte“ (Bundesministerium des Inneren (Hrsg., 2006, 19).

3.2.7 Der Prozess *Configuration Management* im IT Service Management

Ein wichtiges **Ziel** – und für die Betrachtungen im Rahmen dieser Arbeit das relevante Ziel – des *Configuration Management* Prozesses ist die Bereitstellung von akkuraten Informationen über Konfigurationen und ihre Dokumentation, um alle anderen Service Management Prozesse zu unterstützen (Assirati, 2001b, 121).

Dass nicht nur die Modelle zum IT Service Management sondern auch Modelle zur Softwareentwicklung wie beispielsweise *V-Modell XT*⁵, *Rational Unified Process* (RUP)⁶ oder *Capability Maturity Model Integration* (CMMI)⁷ den Gegenstandsbereich des Configuration Managements kennen und abdecken (Walter und Krcmar, 2006, 5), betont dessen zentrale Bedeutung und rechtfertigt die tiefer gehende Analyse von Anforderungen an eine CMDB.

Damit wird dem hierarchisch höheren Ziel der effektiven sowie effizienten Kontrolle und Steuerung der IT Infrastruktur Rechnung getragen. Eine solide Informationsbasis ist erforderlich, um die Service Management Prozesse effektiv unterstützen zu können, allen voran die Service Support Prozesse *Incident Management*, *Problem Management*, *Change Management* und *Release Management*. Zusätzlich unterstützt ein effektives *Configuration Management* das an Bedeutung gewinnende Management von Softwarelizenzen. Das *Configuration Management* spielt so eine zentrale Rolle für alle Service Management Prozesse (Assirati, 2001b, 11). Die Übereinstimmung der Informationen aus dem Configuration Management mit der realen IT Infrastruktur werden durch regelmäßige Audits und die ständige Verbesserung des Configuration Management Prozesses (*Kontinuierlicher Verbesserungsprozess* – KVP) überprüft und im Fall von Abweichungen prozessgeleitet korrigiert.

⁵Wurde in Deutschland für die Entwicklung von Hard- und Softwaresystemen im Auftrag von Behörden entwickelt und enthält Rollen, Aktivitäten und Vorlagen für die Dokumentation der Ergebnisse. (Walter und Krcmar, 2006, 5, 7)

⁶RUP ist ein Referenzmodell zur Entwicklung von Softwaresystemen bestehend aus Rollen, Aufgaben und Arbeitsprodukten und wurde von Rational Software und IBM entwickelt. (Walter und Krcmar, 2006, 5, 7)

⁷CMMI wurde vom *Software Engineering Institute* entwickelt und beschreibt Best Practices für die Entwicklung und den Betrieb von Produkten und Services mit dem Ziel einen bestimmten Reifegrad zu erreichen (Walter und Krcmar, 2006, 5)

Die **Abgrenzung zum Asset Management** wird in der Literatur wie folgt gesehen: Das Configuration Management geht über das Asset Management hinaus, indem dabei nicht nur die einzelnen Anlagegüter sondern auch die Relationen zwischen den Anlagegütern verwaltet werden, was im Rahmen des Asset Managements kaum geschieht (Assirati, 2001b, 121).

Für unsere weiteren Betrachtungen ist eine kurze Darstellung der **Aktivitäten** im Prozess *Configuration Management* hilfreich. Diese wird nach Assirati (2001b) vorgenommen:

- **Planung** des Zwecks, des Umfangs (*scope*), der Ziele, Policies und Prozeduren sowie des technischen und organisationalen Kontexts.
- **Identifikation**: Identifizieren und auswählen der Konfigurationsstrukturen (*configuration breakdown*) für alle Configuration Items (CI) inklusive ihrem *Besitzer (Owner)* und Beziehungen zu anderen CIs, eindeutige Bezeichnung und Erfassung in der CMDB.
- **Steuerung und Kontrolle (*control*)** soll sicherstellen, dass ausschließlich autorisierte⁸ und identifizierbare CIs akzeptiert, erfasst und über den gesamten Lebenszyklus hinweg vom Wareneingang bis hin zur Entsorgung dokumentiert werden.
- **Statusüberwachung** aller erfassten CIs während des gesamten Lebenszyklus und Erstellung einer Historie. Dadurch werden Veränderungen (*Changes*) nachvollziehbar und Statusänderungen dokumentiert: beispielsweise wird der Status von „Bestellt“ auf „In Produktion“ geändert.
- **Verifikation und Audits** überprüfen die tatsächliche Infrastruktur gegenüber der in der CMDB dokumentierten Infrastruktur auf Abweichungen.

Um dem Configuration Management Prozess wirtschaftlich betreiben zu können, ist die Einrichtung einer **Configuration Management Database (CMDB)** dringend empfohlen. Dabei gilt es aber die Gratwanderung zwischen zu viel abgelegten Informationen und zu wenig Informationen zu meistern. Weitere Überlegungen dazu finden sich im Abschnitt „Umfang der CMDB“.

Eine CMDB kann auf mehreren konzeptionellen Ebenen Verwendung finden. Während sie auf der Ebene der Softwareentwicklung die Steuerung und Verwaltung der Softwarekomponenten aus denen ein Softwaresystem besteht und deren Beziehungen zueinander dokumentiert ist die CMDB im IT Service Management bereits in Ansatz umfassender: Von der Ebene der die Geschäftsprozesse unterstützenden IT Services ausgehend werden die zur Erbringung dieser Services erforderlichen Hilfsservices und Komponenten inklusive der Netzwerktechnik bis auf die Ebene der einzelnen Hard- und Softwarekomponenten herunter gebrochen. Die detaillierte Betrachtung der CMDB findet sich im folgenden Abschnitt.

⁸Dies kann etwa durch die Dokumentation eines autorisierten Changes gegeben sein.

3.3 Die Configuration Management Database (CMDB)

Die CMDB enthält Metadaten zu Objekten der unterschiedlichsten Klassen wie Hardwarekomponenten, Softwarepakete oder -module, IT Services, Dokumentation in digitalisierter und auch gedruckter Form. Abhängig von der jeweiligen Implementation werden weitere Informationen beispielsweise zu Störungen in Form von Störungsdatensätzen oder Änderungen an der IT Infrastruktur in Form von (Request for) Changes abgelegt.

Um die Ziele eines *Configuration Management* Prozesses und damit die Bereitstellung von akkuraten Informationen sicher zu stellen, werden heute primär Datenbanken und datenbankähnliche Systeme verwendet, welche große Datenmengen effizient verwalten können.

Der Zugang zur CMDB ist eingeschränkt und kontrolliert, Änderungen an Struktur und Inhalt der CMDB sind in der Regel streng prozessgeleitet und unterliegen den für die jeweilige Organisation zu definierenden Richtlinien.

In der kommenden Version ITIL V3 wird eine so genannte „Wissensdatenbank für Servicemanagement“ eingeführt, welche in inhaltlicher Hinsicht – Details liegen dem Autor dieser Arbeit jedoch noch nicht vor – im Kern der CMDB entsprechen wird.

3.3.1 Betrachtung der CMDB als digitale Bibliothek

Eine CMDB enthält zu einem wesentlichen Teil Daten über real existierende Configuration Items (CIs), Relationen und Dokumentationen. Dazu werden beschreibende Informationen (Kategorie, Modell usw.) aber auch Informationen administrativer Natur (z.B. *owner responsible*) oder strukturelle Informationen („Service *a* benutzt Serverkomponente *s*“) in Form von Datenelementen und Relationen in der CMDB erfasst und gespeichert. Diese Art von Daten fallen unter den Begriff der Metadaten: beschreibende Metadaten und administrative Metadaten machen einen großen Anteil aus. Strukturelle Metadaten sind ein wesentliches Merkmal einer CMDB, da dadurch beispielsweise der Aufbau von Services oder die Struktur einer umfangreicheren Dokumentation erfasst werden können.

In Tabelle 3.1 werden zum Zweck der Überprüfung auf Vergleichbarkeit der Konzepte für digitale Bibliotheken mit jenen für die Umsetzung einer CMDB im Rahmen eines Configuration Managements diese Konzepte einander gegenüber gestellt, um Hinweise darauf zu finden, ob und inwieweit eine CMDB eine digitale Bibliothek darstellt.

TABELLE 3.1:

VERGLEICH DER DEFINITION DER *digitalen Bibliothek* MIT DEM KONZEPT DER *CMDB*

Definition der <i>digitalen Bibliothek</i>	Entsprechung in der <i>CMDB</i>
(Waters, 1998): „Digital libraries are organizations	Der Prozess „Configuration Management“ mit seinen Aktivitäten, Funktionen und messbaren Zielen einerseits und die funktionale Organisation der IT-Abteilung andererseits
that provide the resources,	CMDB und digitale Repositories basierend auf Hardware, Systemsoftware, Datenbankanwendungen und spezifischen Anwendungen inklusive der Einbindung von physischen Repositories für Bücher, CDs, DVDs usw.)
including the specialized staff,	Configuration Manager und Mitarbeitende im Configuration Management sowie die Person(en) welche die Rolle <i>Librarian</i> übernommen hat bzw. haben
to select,	nur im voraus spezifizierte Informationen und (digitale) Objekte werden in die CMDB aufgenommen
structure,	Metadaten in der CMDB sind in Datenbanken organisiert, die <i>Definitive Software Library</i> (DSL) verfügt über eine definierte Struktur
offer intellectual access to,	Abfragemöglichkeiten, Klassifikation, Kategorisierungen, Beschlagwortung aber auch Verknüpfungen von Daten und Informationen sowie deren statistische Auswertung
interpret,	weniger Interpretation, mehr Berichtswesen
distribute,	Weitergabe und Verteilung an andere ITSM-Prozesse, welche die Informationen als Input benötigen
preserve the integrity of,	Änderungen an Daten und Struktur sind strikt geregelt: durch prozessgeleitete Verfahren und unter dem Einfluss des Security Managements
and ensure the	ITSCM-Pläne, gespiegelte Datenbestände, Ersatzrechen-

persistence over time	zentrum, Backup; andererseits fehlen jedoch Konzepte für Migration und Emulation der Repositories in den vorherrschenden ITSM Frameworks
of collections of digital works	zielgerichtete Sammlung von digitalen Objekten – z.B. alle Software, welche in der Organisation verwendet wird (DSL) bzw. wurde (Archiv), Störungsdatensätze, Change- oder Releasedokumentationen – und Metadaten über digitale Objekte in der DSL oder in physischen Repositories für Datenträger oder herkömmlichen Bücherregalen
so that they are readily and economically available	Alle Metadaten und digitalen Objekte werden in strukturierter Form abgelegt und können mit entsprechend integrierten Tools auch ad hoc abgerufen und aufbereitet werden
for use by a defined community or set of communities.“	die Benutzenden sind aufgrund ihrer Rollen mit den erforderlichen Zugriffsberechtigungen ausgestattet

In Ergänzung dazu sollen nach Umstätter (2000) nicht nur bibliografische Nachweise in digitaler Form

teilweise handelt es sich um Metadaten zu physischen Objekten, wobei die Metadaten auf jeden Fall digital abgelegt sind

sondern auch Objekte digital abgelegt werden wobei Dokumente im Volltext zur Verfügung gestellt werden sollen

der Großteil der Inhalte der CMDB bestehen jedoch aus digital abgelegten Metadaten und digitalen Objekten wie bspw. Datenbanken und Dokumentationen

Die Gegenüberstellung in Tabelle 3.1 liefert deutliche Hinweise darauf, dass es sich bei der CMDB um eine digitale Bibliothek handelt. Jedoch kann nicht die gesamte CMDB sondern nur ein Teil – wenn auch ein erheblicher – als eine digitale Bibliothek bezeichnet werden. Die davon ausgenommenen Aspekte sind:

- Die physische Ablage eines Teils der Definitive Software Library (DSL),
- Die in gedruckter Form vorliegenden Dokumentationen.

Nach den ITSM Frameworks sind jedoch zu beiden Punkten verpflichtend Metadatensätze anzulegen und zu warten. Assirati (2001b) folgend sind die physikalischen Bibliotheken unbedingt erforderlich, um die Originaldatenträger und Originaldokumentation ablegen zu können.

Als Fehlstellen bei Configuration Management Datenbanken im Vergleich zu digitalen Bibliotheken finden sich noch zwei weitere geringfügigere Punkte:

- Erstens erfolgen Interpretationen im Zusammenhang mit Informationen in der CMDB nicht durch Mitarbeitende im Configuration Management – beispielsweise als Librarian – sondern durch die die CMDB nutzenden Personen mit Verantwortlichkeiten in anderen Prozessen (z.B. im Incident Management).
- Zweitens wird dem Thema der Konservierung der digitalen Objekte nur geringe Aufmerksamkeit geschenkt.

Zusätzlich wird die CMDB auch Informationen über Backup-Kopien von CIs aufnehmen. In ITIL wird nur auf die Erfordernis einer historischen Dokumentation und Erhaltung von CIs hingewiesen, die tatsächlichen Maßnahmen dazu jedoch nicht angeführt (Assirati, 2001b, 149). Dies resultiert im Nachholbedarf von Konzepten zur Auffrischung, Emulation und Migration. Zur Behebung dieses Mankos bieten sich zwei Möglichkeiten an: es wird ein eigener Service vereinbart oder eine Erweiterung des Configuration Managements in Richtung Konservierung implementiert.

Die Frage, ob es sich bei CMDBs um digitale Bibliotheken handelt, kann somit für einen großen Teil der CMDB bejaht werden.

Darüber hinaus lässt sich auch die FRBR Struktur in den beschreibenden Metadaten in einer CMDB wieder finden:

- Angaben auf der Ebene des **Werks** sind beispielsweise der Hauptsachtitel „E-Mail Service“ für einen Service, „Microsoft Windows Server 2003 R 2“ für ein Configuration Item oder „Thinking in Java“ und dem Inhalt „Eckel, Bruce“ im Datenfeld der 1. Person für ein Sachbuch.
- Angaben auf der Ebene der **Expression** sind beispielsweise „Englisch“ für das Configuration Item und das Buch sowie zusätzlich „Installationsdateien“ für das CI.
- Angaben auf der Ebene der **Manifestation** sind beispielsweise die Angaben „CD ROM“ für das Configuration Item oder die bibliografischen Angaben „Monografische Druckschrift, New Jersey, 3rd ed., 2003“, für das Buch.

- Angaben auf der Ebene des **Exemplars** sind beispielsweise die Seriennummer des CI sowie die Signatur des Buchs.
- Angaben zur Entität **Personen** sind beispielsweise die für den Service zuständigen Mitarbeitenden.
- Angaben zur Entität **Körperschaften** sind beispielsweise die für den Betrieb des Services erforderlichen Hersteller, Lieferanten und externe Dienstleister.
- Angaben zur Entität **Begriffe** sind beispielsweise Leistungsindikatoren, Schwellenwerte oder bei Störungen die beschreibenden Stichwörter dieser Störungen als Kategorisierung.
- Angaben zur Entität **Gegenstände** sind beispielsweise die von Störungen betroffen CIs.
- Angaben zur Entität **Ereignisse** sind beispielsweise Requests for Changes, Nichtverfügbarkeit oder der Eintritt des Katastrophenfalls.
- Angaben zur Entität **Orte** sind beispielsweise der Ort der Leistungsübergabe in einem SLA, nicht aber Standorte, da diese aus der Sicht einer Bibliothek außerhalb der FRBR in den lokalen Datensätzen abzulegen wäre.

3.3.2 Umfang der CMDB

In der betriebliche Praxis wird die CMDB selten aus einer einzelnen Datenbank bestehen. Vielmehr repräsentiert die CMDB eine Sammlung von Datenbanken und anderen Ressourcen wie gedruckte Handbücher, Textdokumente oder Dateien mit grafischem Inhalt. Alle Ressourcen haben den Zweck, eine IT Infrastruktur und die bereitgestellten Services zu dokumentieren.

Da die Vielfalt der berücksichtigenswerten Informationen insbesondere zu Beginn einer Implementation einer CMDB nur sehr schwer strukturierbar ist, bietet sich die Erstellung eines **Configuration Management Plans** an. Dieses Konzept ist ein Metadatenmodell welches Objekte und ihre Beziehungen untereinander dokumentiert als auch die Person(en) nennt welche die Datenverantwortung besitzen (Vogt, 2002, 66). Somit wird auch eine (anfänglich) dezentrale Verwaltung der CMDB möglich, ohne auf den Qualitätsaspekt der Herkunftsangabe (Provinienz) verzichten zu müssen.

Basierend auf Assirati (2001b), Assirati (2001a), Vogt (2002) und Lindinger (2005) kann diese Sammlung die folgenden **Informationen bzw. Datenbestände** umfassen⁹:

1. Informationen zu Configuration Items – so werden jene Komponenten bezeichnet, aus welchen die IT Infrastruktur besteht – mit den diese beschreibenden Attributen

⁹Für alle angeführten Begriffe gilt im Unklarheit oder Mehrdeutigkeit die in ITIL V2 definierte Bedeutung.

in technischer, administrativer (z.B. den Status: im Test, freigegeben, in Produktion, archiviert), logistischer (Standort), wirtschaftlicher (z.B. Anschaffungskosten, Nutzungsdauer), rechtlicher (z.B. Lizenzen, Wartungsverträge) und datensicherungs-technischer Sicht. Nach Assirati (2001b) und (Lindinger, 2005) sollten die folgenden Komponenten als CIs erfasst werden:

- Hardware, das sind Endgeräte, Server, Storage und den relevanten aktiven und passiven Netzwerkkomponenten inklusive Telekommunikationseinrichtungen sowie nach (Rachlin und Kneiling, 2007, 10) auch Raum (auf Serverräume bezogen) zumindest als Attribute, Stromversorgung und Klimatisierung
 - alle Hardwarekomponenten, die sich auf Lager im *Definitive Hardware Store* (DHS) befinden und auf den Einsatz als Ersatzteile warten oder Hardwarekonfigurationen die auf den Einsatz in der Produktionsumgebung vorbereitet wurden
 - Systemsoftware inklusive Betriebssysteme
 - Businessanwendungen inklusive kundenspezifischen Entwicklungen und Eigenentwicklungen
 - *Commercial off-the-shelf* (COTS) Softwarepakete und andere Standardprodukte inklusive Datenbanklösungen
 - Physische Datenbanken
 - Datenfeeds zwischen Datenbanken und Anwendungen inklusive EDI-Links
2. Beziehungen zwischen Configuration Items in logischer und physischer Weise. Beispiele für Relationen sind „ist installiert auf“ oder „ist eingebaut in“.
 3. Historie der Configuration Items.
 4. Zusätzlich wird die CMDB auch Informationen über Backup-Kopien von CIs aufnehmen. Das Ausmaß der historischen Dokumentation und Erhaltung hängt von der Relevanz für die Organisation ab und ist regelmäßig einer Überprüfung zu unterziehen. (Assirati, 2001b, 149)
 5. Links zu Störungs- und Problemdatensätzen: Im Störungsdatensatz *Incident Record* sollten Benutzende *User* und das die Störung verursachende CI *Failing Component* auf jeden Fall erfasst werden.
 6. Releaseeinheiten und Beziehungen zu den betreffenden Configuration Items.
 7. Detaillierte Informationen zu den bereitgestellten Services (Servicekatalog).
 8. Informationen zu den für die Serviceerbringung erforderlichen Hilfsservices.

9. Beziehungen zwischen Services und Configuration Items bzw. zwischen Services und Geschäftsprozessen oder Kunden. Beispiele für diese Art von Relation sind „gehört zu Service x “ oder „wird genutzt von Kunde y “. Die Komponenten müssen Services zugeordnet werden können, um *Service Level Agreements* (SLAs) erstellen und überwachen zu können.
10. Die Vereinbarung zwischen IT Organisation und Kunde über den zu erbringenden Service (das *Service Level Agreement* – SLA) sowie die beiden Absicherungsverträge: einerseits die organisationsinternen Vereinbarungen *Operational Level Agreements* (OLAs) und andererseits die Vereinbarungen mit externen Dienstleistern *Underpinning Contracts* (UCs). (Vogt, 2002, 308)
11. *Configuration Baselines* als Abbildungen einer funktionierenden Konfiguration einer gesamten IT Infrastruktur zu einem bestimmten Zeitpunkt oder Teilen davon.
12. IT-relevante technische Dokumentationen zu den eingesetzten oder archivierten Systemen und Schnittstellen inklusive aller zusätzlichen Unterlagen, beispielsweise Schulungsunterlagen welche
 - in Form von gedruckten Unterlagen oder
 - elektronisch in unterschiedlichen Dateiformaten (z.B. HTML, PDF, DOC) und
 - auf unterschiedlichsten physikalischen Datenträgern (z.B. CD, DVD, Festplatte, Magnetbändern) oder
 - virtuell über das World Wide Web abrufbar sind (z.B. Verweise auf Tutorials, Herstellerinformationen).
13. Organisationsinterne IT-bezogene Dokumentationen zu Organisation und Prozessen, Prozeduren bis hin zu Anweisungen in Form von Ablaufdiagrammen, Beschreibungen oder Organigrammen.
14. Organisationsinterne Informationen welche teilweise besondere Schutzbedürfnissen unterliegen wie solche zu Usern, Personal, IT-Personal, Unternehmensorganisation, Standorte und Lieferanten gegebenenfalls in Form von Verknüpfungen zu den spezifischen Datenbeständen.
15. Die gesamte Bibliothek aller freigegebenen Software (*Definitive Software Library, DSL*) in welcher die Masterkopien aller freigegebenen *Software Configuration Items* gespeichert und geschützt sind. Die Speichermedien innerhalb einer logischen DSL setzen sich üblicherweise aus einem oder mehreren Festplatten- oder Storage-Systemen mit kontrollierten Zugriff – Bedingungen für den Zugang, Entnahme und Abgang von Software CIs – und Datenträgern wie anderen physischen Datenträgern wie CDs oder DVDs in einem sicheren (feuerfesten und versperrenbaren) Schrank zusammen. Diese Bibliothek ist Assirati (2001b) zufolge getrennt von Entwicklungs-, Test- und Produktionsbibliotheken (Fileservern) zu bewirtschaften und unterliegen der Kontrolle durch

das Change und Release Management. Die Sicherungen der DSL sind ebenfalls außerhalb der Produktionsumgebung aufzubewahren, unterliegen aber der Erfassung in der CMDB (Vogt, 2002, 154). Weiters soll die DSL nach Vogt (2002) zur Aufnahme von ausführbarem Code und Quellcode inklusive Dokumentation, Linkprogramme und Compiler in der jeweils relevanten Version vorbereitet sein und unterliegt der strikten Kontrolle des Change Managements und des Release Managements.

16. Störungen (*Incidents*), auf CIs bezogen.
17. Unbekannte Ursachen für eine oder mehrere tatsächliche oder potenzielle Störungen (*Problems*) und bekannte Fehler (*Known Errors*) bilden, jeweils auf CIs bezogen, die so genannte *Knowledge Database*. Diese kann in die CMDB eingegliedert sein.
18. Veränderungsanträge (*Requests for Change*) und geplante Veränderungen (*Changes*) mit deren Status (z.B. eingereicht, autorisiert, implementiert oder getestet), jeweils auf die betroffenen CIs bezogen. Zusätzlich zu den Details von Änderungen gilt die Übersicht über kommende Veränderungen in Form eines Änderungszeitplans (*Forward Schedule of Change*) als besonders hilfreich für das Incident Management und als Zusatzinformation für die Mitarbeitenden am Service Desk.
19. Release-Grundsätze (*release policies*), Planungsunterlagen für geplante neue Releases und Release-Dokumentation für jeden neue Release aus dem Releasemanagement mit den Beziehungen zu den betroffenen Hard- und Software-CIs.
20. Kapazitätspläne inklusive Schwellwertdaten und Kapazitätsdatenbank (*Capacity Database – CDB*) aus dem Kapazitätsmanagement liefern nach Assirati (2001a) Daten auf den drei nachstehenden Ebenen:
 - *Business Capacity* (Trends im Geschäftsverlauf der Organisation, Absatzplanung und zukünftige Geschäftsanforderungen) als digitale Objekte in der CMDB oder als Verweise auf die relevanten Dokumente,
 - *Service Capacity* (Dimensionierung der Servicekapazität um die Geschäftsanforderungen aus der Sicht der IT abdecken zu können, Überwachungs- und Analysedaten, Berichte zur Performance der Services, Nutzungsprofile der Services inklusive Daten zur Beeinflussung der Nachfrage) und
 - *Ressource Capacity* (Dimensionierung der Komponenten, Auslastung der Komponenten, Minimalanforderungen an Kapazität und Performance)
21. Verfügbarkeitspläne mit Daten zur Verfügbarkeit sowie Zuverlässigkeit einzelner CIs und Services sowie CI-spezifische Angaben zu Wartbarkeit¹⁰ und Servicefähigkeit¹¹.

¹⁰Wartbarkeit bedeutet nach Assirati (2001a) die Möglichkeit, eine IT Komponente nach einem Störfall wieder in einen betriebsfähigen Zustand bringen zu können. Dazu müssen sieben Schritte durchführbar sein: Vorhersehbarkeit von Fehlern, Fehlererkennung, Fehlerdiagnose, Fehlerlokalisierung, Fehlerbehebung, Zurücksicherung von Daten und IT Services sowie die Möglichkeit zu Präventivmaßnahmen um Ausfälle zu verhindern.

¹¹Servicefähigkeit bedeutet nach Assirati (2001a), dass für eine IT Komponente externer Support ver-

Weiters sollen Daten zur Generierung von Analysen und Vorausberechnungen (*forecasts*) abgelegt werden können.

22. IT Finanzpläne inklusive Budgetziele. Nach Vogt (2002) sind noch Prozesskosten und die Kostenstruktur der Services relevant.
23. IT Service Continuity: Ein Plan für den Fall, dass der im Vorhinein definierte Ausnahmezustand einer Katastrophe eintritt. Dazu werden einer Notsituation entsprechende Entscheidungsbefugnisse und Wiederherstellungsprozeduren frühzeitig – jedenfalls bevor eine derartige Notsituation eintritt – im Rahmen eines Planungsprozesses festgelegt und dokumentiert. Sinnvollerweise wird zur diesbezüglichen Dokumentation nicht alleine auf die CMDB zurückgegriffen sondern es werden auch in alternativen Medien außerhalb der CMDB (etwa gedruckte Pläne bei den in Krisensituationen Verantwortlichen) abgelegt sein. Immer jedoch gilt es, diese einzelnen Ausgaben der *IT Service Continuity Pläne* (Exemplare) zu administrieren und diesbezügliche Metadaten zu verwalten.

Was im Einzelfall tatsächlich dem Configuration Management Prozess unterworfen und in die CMDB aufgenommen wird, muss vor Aufbau der CMDB festgelegt werden, kann aber in weiterer Folge durchaus erweitert werden. Dieser Umstand wirft folgende Fragen nach dem Umfang (*scope*) auf (Vogt, 2002):

- **Worüber** müssen Informationen in der CMDB abgelegt werden?
- **Was muss mindestens** in der CMDB abgelegt werden, um die Ziele der damit arbeitenden Prozesse nicht zu gefährden? Diese Frage nach dem „was wir darüber wissen müssen“ zielt auf die Attribute der Metadaten hin. Bis zu welcher Detailinformation (Detaillierungsgrad) soll in der CMDB erfasst werden? Diese Frage soll Antworten liefern, ob beispielsweise für einen Client PC die Festplatte als eigenes CI erfasst wird oder ob die gesamte Zentraleinheit des Client PC als eine Einheit betrachtet wird und ein Tausch der Festplatte „nur“ ein *Change* des Client-PC-CIs und somit eine Änderung eines Attributs darstellt. Anders wird sich dieser Umstand beim Einsatz einer mobilen externen Festplatte gestalten.

In der vorherrschenden Literatur finden sich verteilt durchaus Informationen zu den beiden vorstehenden Fragen, jedoch kaum konkrete Anforderungen an die Metadatenelemente. Bedeutend weniger Empfehlungen finden sich zur Beantwortung der folgenden Frage:

Welche konkreten Metadatenelemente müssen in der CMDB abgelegt werden, um aus der Sicht der *alphabetischen* und *inhaltlichen Erschließung*, der Struktur sowie aus administrativer Sicht ausreichend Informationen zur Verfügung zu haben.

traglich mittels *Underpinning Contract* vereinbart werden kann, wodurch die Verfügbarkeit, Zuverlässigkeit und Wartbarkeit der Komponente abgesichert werden kann.

Von besonderer Bedeutung ist der Ausgleich zwischen *zu viel* und *zu wenig* abgelegten Informationen. Dabei wird die Grenze im Einzelfall und projektabhängig zu ziehen sein. An dieser Stelle soll die folgende Überlegung angestellt werden:

Aus dem Gebot zur Wirtschaftlichkeit ergibt sich ein **Nachteil bei zu vielen abgelegten Informationen zu einem CI oder anderen Objekt**, da diese mit oft erheblichem Aufwand erfasst werden müssen, deren Kenntnis aber keinen wesentlichen Vorteil darstellt. Wenn beispielsweise die Seriennummer aller Netzwerkkarten von Client-PCs einer IT Infrastruktur erfasst werden würden, wäre dies ein erheblicher Aufwand: entweder in Form einer manuellen Erfassung oder der Beschaffung und des Betrieb von entsprechenden Scannern. Der Nutzen jedoch im Falle einer Störung der Netzwerkkarte bleibt fraglich, denn die Seriennummer der Netzwerkkarte wird in den seltensten Fällen Hinweise für die Störungsbehebung oder die Ursachenforschung geben. Hilfreicher sind tendenziell Informationen etwa zu Hersteller (aus dem Release Management) oder Einbauzeitpunkt (aus der Dokumentation des Changes) welche ohnehin bereits erfasst sind und im gegebenen Fall zumindest zukünftige Störungen rascher einzugrenzen helfen oder durch die Erleichterung von pro-aktive Maßnahmen vermeiden helfen. Ein konkretes Beispiel für eine pro-aktive Maßnahme wäre der Austausch aller Netzwerkkarten des Typs x des Herstellers y durch Netzwerkkarten vom Typ z , nachdem erkannt wurde, dass die Netzwerkkarten des Typs x des Herstellers y in bestimmten Betriebszuständen Störungen verursachen. Für den Fall, dass für eine Garantieleistung eines Lieferanten die Seriennummer erforderlich ist, kann diese kostengünstiger im Anlassfall erhoben werden. Vor allem entfällt bei dieser Vorgangsweise für Netzwerkkarten in Client-PCs der Aufwand der ständigen Aktualisierung nach allen das CI betreffenden Veränderungen.

Zu wenige Informationen zwingen bei einer Störungsbehebung oder Ursachenforschung diese nachträglich zu erheben. Beispielsweise wird nach Arbeitsende eine Störung an einem Router gemeldet. Um den Standort der fehlerhaften Komponente (z.B. Wien oder Linz) ausfindig zu machen muss dieser Ort für den fehlerhaften Router telefonisch erfragt werden. Dazu ist allerdings der Arbeitsbeginn am nächsten Tag abzuwarten. Eine Erhebung von wichtigen Informationen kann so mit Zusatzaufwand verbunden sein und einen erheblichen Zeitverlust bedeuten, was die Ausfallszeit erhöht und somit die Produktivzeit für die Benutzenden reduziert.

3.4 Benutzergruppen und Prozesse mit berechtigtem Interesse am Zugang zur CMDB

Nach Arms (2000) müssen im Rahmen der Konzeption von digitalen Bibliotheken neben vielen technischen und organisatorischen Aspekten auch die **Benutzergruppen** berücksichtigt werden, welche ein Interesse am Zugang zur digitalen Bibliothek haben und deren Services nutzen werden.

Die weiter oben genannten *IT Service Management Frameworks* sind in der Regel prozess-

orientiert, wobei jeweils der gesamte vom Framework abgedeckte **Gegenstandsbereich** in einzelne Prozesse mit jeweils klar definierten Zielen gegliedert ist. Zur Erreichung der Prozessziele werden Informationen – primär über die IT Infrastruktur aber auch über Services, Dokumentationen, Vereinbarungen (SLA) und dergleichen – benötigt.

Da bei Arms (2000) explizit von „Benutzergruppen“ gesprochen wird, aber im ITSM sich die handelnden Personen immer nur durch Rollen (z.B. ITIL) oder über die Definition von Wissensgebiete und Fähigkeiten der Team-Rollen (z.B. MOF) in die entsprechenden Prozesse einbringen, werden die einzelnen Anforderungen an die CMDB in abstrahierter Form auf der Ebene der einzelnen **Gegenstandsbereiche mit eigenem Ziel** beschrieben. Die Analyse auf der Ebene der einzelnen Gegenstandsbereiche ermöglicht zudem die Unabhängigkeit von der Prozessorientiertheit so dass jegliches ITSM Framework - egal ob prozessorientiert oder nicht – ihre Berücksichtigung finden kann.

Weiters sind folgende Faktoren zu berücksichtigen:

1. Die aufgrund der unterschiedlichsten Motivationen gewünschten Informationen sind nicht immer im selben Ausmaß zur Zielerreichung innerhalb eines Gegenstandsbereichs relevant.
2. Der Aufwand eine Information zu beschaffen und aktuell zu halten kann höher sein, als der Aufwand die selbe aktuelle Information im Bedarfsfall zu erheben. Dies trifft insbesondere bei sehr selten benötigten Informationen zu.
3. Manche Informationen unterliegen von Gesetzes wegen besonderen Schutzbedürfnissen wodurch der Zugang dazu nur eingeschränkt erfolgen darf.
4. Ein Teil der Informationen kann aus organisationsinternen vertraulichen Informationen in unterschiedlichen Vertraulichkeitsstufen bestehen, zu welchen der Zugang ebenfalls nur eingeschränkt gewährt werden kann.

Neben dem generellen Interesse an einer Information ist demnach auch der Umstand des **berechtigten Interesses** an der Information zu berücksichtigen. Dazu lassen sich keine generellen Annahmen treffen: Es ist auf Projektbasis zu entscheiden, welche Informationsbedürfnisse berechtigt sind und welche keine Rechtfertigung finden. In weiterer Folge sollen daher nur Informationen erfasst werden, an welchen zumindest ein Gegenstandsbereich ein berechtigtes Interesse hat.

Die kommenden Unterabschnitte benennen nun jene Gegenstandsbereiche, welche gleichartige Interessen am Zugang zu Informationen und zu Services des Configuration Managements aufweisen.

3.4.1 Incident Management

Im Rahmen des Incident Managements ist der Umstand hilfreich, dass auf die folgenden Informationen zugegriffen werden kann (Vogt, 2002, 308):

- Details zu allen bisherigen Störungen – diese gelten als für das Incident Management wesentliche Inputs (Assirati, 2001b)
- Als weiterer wesentlicher Input gelten nach Assirati (2001b) die Details zur IT-Infrastruktur (Services, Hardware, Software, Netzwerk, Dokumentationen und Beziehungen untereinander)
- Probleme, bekannte Fehler, (Umgehungs-) Lösungen aus internen Quellen oder von Lieferanten bzw. Dienstleistern
- Requests for Change, geplante und durchgeführte Veränderungen, Forward Schedule of Changes
- Release Notes und Roll-out Pläne
- Vereinbarte Services, Vereinbarungen zu Supportlevel aus SLAs, OLAs und UCs, Auswirkungen und Konsequenzen bei Nichterfüllung der Vereinbarungen, aber auch Auswirkungen auf die Geschäftsprozesse des Kunden
- Prozesskosten, Aufwand für die Störungsbearbeitung
- Schwellenwerte aus dem Kapazitätsmanagement
- Notfallplan aus dem IT Service Continuity Management
- Anfragen und Beschwerden von Kundenseite bzw. von den Benutzenden
- Prozessziele als Vorgaben durch das Management der Organisation

Durch Deckung dieses Informationsbedarfs sind die Voraussetzungen für die raschestmögliche Behebung von Störungen, effiziente Bearbeitung von Anfragen (*Service Requests*) und ein effizienter Übergabepunkt für Beschwerden der Benutzenden gegeben.

3.4.2 Problem Management

Vogt (2002) dokumentiert die zur Ursachenforschung und zur Behebung von strukturellen Schwachstellen benötigten Informationen wie folgt:

- Störungsdetails – vor allem zu den am häufigsten auftretenden Störungen im vergangenen Beobachtungszeitraum
- Detailinformationen zur IT Infrastruktur sind nach Assirati (2001b) als wesentlicher Input zu betrachten
- Umgehungslösungen welche im Zuge von Störungsbehebungen entstanden – diese stammen in der Regel aus dem Incident Management

- Die Ergebnisse von *Post Implementation Reviews* (PIR), das sind Überprüfungen nach jeder Veränderung, ob die Veränderung korrekt implementiert wurde, ob sie funktionierte und ob die Veränderung den bekannten Fehler tatsächlich beheben konnte
- Release Notes sind für die Entwicklung von Lösungen mittels bereits freigegebener Komponenten erforderlich
- Vereinbarungen zu Services, Supportlevel und Auswirkungen auf die Geschäftsprozesse des Kunden
- Prozesskosten
- Performance- und Kapazitätsengpässe sowie Trends hinsichtlich der erforderlichen (für das pro-aktive Problem Management) und tatsächlichen (im Störfall für das reaktive Problem Management) Performance und Kapazität
- Auf die Verfügbarkeit bezogene Trends
- Notfallplan
- Bekannte Fehler, (Umgehungs-) Lösungen, Know-how und Erfahrungen von externen Lieferanten und Dienstleistern
- Zielvorgaben für das Problem Management vom Management der Organisation

Mit diesen Informationen sollte das Problem Management in der Lage sein, die IT-Infrastruktur nachhaltig zu verbessern: fehlertoleranter zu gestalten und *Single Points of Failure* (SPOFs) zu reduzieren, im Idealfall zu eliminieren.

3.4.3 Change Management

Nach Vogt (2002) erfordert die effektive und effizient Bearbeitung von Veränderungsanträgen die folgenden Informationen:

- Informationen über die aktuelle Infrastruktur
- Requests for Change (RfC) und Anfragen – welche in RfCs resultieren könnten – vom Incident-, Problem-, Servicelevel-, Capacity- und Availability Management
- Releaserichtlinien, Testergebnisse der Tests vor den Freigaben von Komponenten, Roll-out Planungsunterlagen, Ergebnisse der Fortschrittskontrolle sowie Reports des Release Managements
- Input für das *Change Advisory Board*¹² (CAB) vom Kunden

¹²Das *Change Advisory Board* berät über die vorgelegten Veränderungsanträge.

- Anforderungen an die Performance – diese sollten nach Assirati (2001a) als Attribute zu den CIs gewartet werden
- Wartungspläne mit der Angabe von Wartungsfenstern – bereitgestellt vom Availability Management – als Zeiträume für die Durchführung von Veränderungen welche die Verfügbarkeit beeinträchtigen wie beispielsweise Wartungsarbeiten an Services
- IT Continuity Plan
- Prozesskosten
- Prozessziele

Den umfangreichsten Input dazu liefern neben dem Releasemanagement primär die so genannten planerischen Gegenstandsbereiche, allen voran das Configuration- und Availability Management oder das *IT Service Continuity Management* (ITSCM).

3.4.4 Release Management

Nach Vogt (2002) werden die folgenden Informationen zur Erstellung von neuen Releaseeinheiten benötigt:

- Informationen über die aktuelle Infrastruktur inklusive Hardwarespezifikationen, Installationsanleitungen und Netzwerkkonfigurationen (Assirati, 2001b)
- Autorisierte *Request for Change* (RfCs) als Arbeitsgrundlage
- Auswirkungen auf die Geschäftsprozesse aus dem Service Level Management
- Ergebnisse der Capacity Audits von Roll-outs
- Tests durch Benutzende der Kundenorganisation(en)
- Informationen zu neuen oder geänderten Hard- und Softwarekomponenten von den externen Herstellern oder Dienstleistern
- Prozesskosten
- Prozessziele
- Detaillierte Informationen über den Inhalt der Definitve Software Library (DSL) – Assirati (2001b) verweist zusätzlich darauf, dass zur Erreichung der Ziele des Release Managements die gesicherte Aufbewahrung von Masterkopien der eingesetzten Software im Rahmen der in die CMDB eingegliederte DSL erforderlich ist.

Hier wird besonderer Wert auf Informationen zu den potenziellen Auswirkungen von Veränderungen auf IT Services und die darauf aufbauenden Geschäftsprozesse gelegt, um die Notwendigkeit und das Risiko abschätzen, beziehungsweise Testszenarien erstellen zu können.

3.4.5 Service Level Management

Nachdem eine wesentliche Aufgabe des Service Level Managements das Halten des Kontakts zum Kunden ist, muss dafür auf eine zuverlässige Informationsbasis zurückgegriffen werden können. Vogt (2002) umreißt die benötigten Informationen wie folgt:

- Informationen über die aktuelle Infrastruktur – insbesondere über die Komponenten aus welchen die Services bestehen und der Bezug von Ereignissen wie Störungen oder Veränderungen auf Services hergestellt werden kann (Assirati, 2001a)
- Servicekatalog¹³ als Übersicht über alle aktuell lieferbaren Services dienen
- *Service Level Agreements* (SLAs), *Operational Level Agreements* (OLAs) und *Underpinning Contracts* (UCs), wozu nach Assirati (2001a) auch die Festlegung von Eskalationswegen dazugehören
- Berichte aus dem Incident-, Problem-, Change- und Availability Management – nach Assirati (2001a) sind auch aktuelle und vergangene Daten zu Performance, Kapazitäten und Auslastung heranzuziehen, um erreichbare Zielvorgaben in SLAs definieren zu können
- Roll-out Planung aus dem Release Management
- Kosten der Leistungserstellung mit Servicekostenprofil inklusive Planvorgaben aus dem Financial Management for IT Services
- Prozesskosten und Budget aus dem Financial Management for IT Services
- Notfallplan aus dem ITSCM
- Kundenanforderungen und Zufriedenheit aus Kundenangaben
- Security Policy – in Zusammenarbeit mit dem Kunden und dem Security Management
- Zielvorgaben für das SLM vom Management der Organisation

Im Gegenstandsbereich des Service Level Managements werden alle den Kunden und den Service betreffenden Informationen benötigt um die mit dem Kunden im SLA vereinbarte Servicequalität in der IT Organisation durchsetzen zu können und damit die Kunden zufrieden zu stellen, beziehungsweise Abweichungen vom vereinbarten Serviceniveau argumentieren zu können.

¹³Nach Assirati (2001a) kann der Servicekatalog in einfachster Form aus einer Matrix bestehen, welche angibt, welche Kunden welche Services nutzen. In ausführlicherer Form kann er als hierarchische Darstellung der Services (z.B. E-Mail) mit deren beteiligten Services (z.B. Virenschutz, dieser kann auch als eigenständiger Service zusätzlich zur Fileserverbereitstellung vereinbart sein) und Hilfsservices (z.B. Netzwerk) sowie den damit verknüpften Servicelevel und Kunden aufgebaut sein. Weiters dient er als Grundlage für die Überwachung der Services.

3.4.6 Financial Management for IT Services

Das Financial Management for IT Services benötigt für die IT-bezogene Budgeterstellung, die Kostenermittlung und Leistungsverrechnung die nachstehenden Informationen (Vogt, 2002):

- Aufbau der IT Infrastruktur (CIs) und Services
- Voraussichtlicher Arbeitsaufwand und Materialeinsatz für die Aufwandsabschätzung von Veränderungsanträgen
- Roll-out Planung, Releaseplan, Releaseeinheiten
- Vereinbarte Services mit Kosten aus den Absicherungsverträgen OLAs und UCs
- Anschaffungs-, Leasing- oder Mietkosten für Komponenten (Hardware, Software) und zugehörige Leistungen mit den Rechnungen von Lieferanten
- Kapazitätsplan, Ressourcen- und Risikoprofile
- Verrechnungspolitik
- Zahlungen des bzw. der Kunden, falls vereinbart
- Organisationsziele, Plan- und Budgetvorgaben
- Erfolgsrechnung

Zusätzlich werden weitere Informationen benötigt, welche jedoch nicht in der CMDB abgelegt werden, sondern in anderen Informationssystemen verwahrt sind. Darunter fallen Informationen kostenrechnerischer oder finanzbuchhalterischer Art. Um einen Zugriff auf diese externen Daten zu erleichtern, sollten zu diesen Quellen – meist Datenbanken oder Berichte aus IT-fremden Teilen der Organisation – zumindest Metadaten erfasst werden.

3.4.7 Capacity Management

Das Capacity Management benötigt viele Einzeldaten zu technischen Belangen der Configuration Items (CIs), zu den Services und deren jeweiligen Nutzung durch die Geschäftsprozesse, finanzielle Daten und quantitative Informationen zu den Geschäftsprozessen selbst. Diese Daten werden meist als Attribute zu CIs gewartet (Assirati, 2001a). Dem Autor sind keine Empfehlungen bekannt, wonach Geschäftsprozesse als eigene Klasse in der CMDB berücksichtigt werden sollen. Aus der Sicht der Gesamtorganisation macht dies durchaus Sinn, denn das Management und die dazugehörige Dokumentation der Geschäftsprozesse erfolgt üblicherweise außerhalb der IT.

Vogt (2002) liefert die nachstehende, gegliederte und nach Assirati (2001a) ergänzte Aufstellung:

- Von zentraler Bedeutung für das Kapazitätsmanagement ist Assirati (2001a) folgend die *Capacity Database* (CDB) als wichtiger Teil der CMDB, indem sie kapazitätsbezogene Daten zu Geschäftsprozessen (*Business Capacity*), Services (*Service Capacity*) und einzelnen Komponenten (*Ressource Capacity*) bereithält
- Informationen über die aktuelle Infrastruktur, welche nicht bereits durch die CMDB abgedeckt sind
- Kapazitätsbezogene Ereignisse (Alarmer, Störungen) aus dem Incident Management
- Kapazitätsprobleme aus dem Problem Management
- Kapazitätsauswirkungen von auf die Verfügbarkeit zielende Maßnahmen
- Kapazitätsauswirkungen von auf die Verfügbarkeit zielende Planungen hinsichtlich IT Service Continuity
- Angaben der jeweiligen Hersteller zur Performance von Komponenten
- Forward Schedule of Changes und Details zu Changes (Assirati, 2001a), damit die Performanceanforderungen von geplanten Veränderungen abgeschätzt werden können
- Roll-out Planung, Releaseplan, Releaseeinheiten
- Entwicklung des Geschäftsgangs des bzw. der Kunden, Anforderungen der Geschäftsprozesse
- Security Policy
- Prozesskosten
- Prozessziele

An dieser Stelle lassen sich nur wenige konkrete Metadatenelemente ausmachen, da deren Strukturierung sehr von der jeweiligen Intensität des Einsatzes und vom Fokus des Capacity Managements abhängt.

3.4.8 Availability Management

Nach Vogt (2002) erleichtern die folgenden Informationen die Planung und Durchsetzung von Anforderungen an die Verfügbarkeit:

- Informationen über die aktuelle Infrastruktur – diese sind nach Assirati (2001a) ein wichtiger Input für das Availability Management und müssen ausreichend sein, um eine Risikoanalyse und eine *Failure Impact Analysis* durchführen zu können
- Verfügbarkeitsbezogene Ereignisse (Alarme, Störungen)
- Probleme im Zusammenhang mit der Verfügbarkeit von Services oder Komponenten
- Forward Schedule of Changes
- Roll-out Planung, Releaseplan, Releaseeinheiten
- Auswirkungen auf Geschäftsprozesse, Anforderungen an die Verfügbarkeit von Services aus Kundensicht für unterschiedliche Servicelevel
- Kapazitätsplan, Ressourcen- und Risikoprofile
- Verbindliche Aussagen von Herstellern bzw. Lieferanten zu Zuverlässigkeit, Wartbarkeit und Servicefähigkeit
- Prozesskosten
- Prozessziele

Die im Availability Management verwendeten Tools und Techniken – beispielsweise die *Component Failure Impact Analysis* (CFIA) oder *Fault Tree Analysis* (FTA) – sind oft identisch mit jenen, welche im Capacity Management oder im IT Service Continuity Management verwendet werden (Assirati, 2001a), da beide Gegenstandsbereiche in technischer Hinsicht ähnliche Ziele verfolgen. Auch wird eine abgestimmte Planung effizient sein und nur *ein* System für das Monitoring von Komponenten und Services eingesetzt werden.

Trotzdem kann der Wunsch des Availability Managements nach einer eigenen Datenbank in einem eigenen Repository sinnvoll sein. Diese *Availability Management Database* (AMDB) kann dann dem Availability Management eigene Metriken zum Zweck der Dokumentation, statistischen Analyse und vorausschauenden Planung aufnehmen. Dabei ist jedoch zu beachten, dass Informationen nicht redundant – einmal in der AMDB für Analysezwecke des Availability Managements und zusätzlich in der CMDB für den Zugriff auf die Daten für alle anderen Gegenstandsbereiche – abgelegt werden, wodurch ein nicht gerechtfertigter Zusatzaufwand (Synchronisation, Speicherplatz, administrativer Aufwand) entstehen würde. (Assirati, 2001a)

3.4.9 IT Service Continuity (ITSCM)

Um die IT Service Continuity Pläne aktuell – das bedeutet den sich verändernden Umständen angepasst – zu halten, benötigt das ITSCM nach Vogt (2002) in regelmäßigen Abständen die nachstehenden Informationen:

- Informationen zum Aufbau der IT Infrastruktur und Services werden benötigt, um Risikoanalysen und Auswirkungsanalysen (*Component Failure Impact Analysis* oder *Fault Tree Analysis*) durchführen zu können
- Feststellung des Notfalls
- Forward Schedule of Change
- Roll-out Planung, Releaseplan, Releaseeinheiten
- Auswirkungen auf Geschäftsprozesse, Anforderungen an die Verfügbarkeit von Services aus Kundensicht für unterschiedliche Servicelevel
- Kapazitätsplan, Ressourcenprofile
- Risikoprofile und Schwachstellen als Inputs für eine Risikoanalyse
- Business Impact und Business Continuity Planung als Informationen vom Kunden
- Umfang der Standby-Strategie hinsichtlich Komponentenlieferungen und Services von externen Lieferanten
- Prozesskosten
- Prozessziele

Die CMDB und gesichert aufbewahrte Bibliotheken ermöglichen – eine sorgfältige Datensicherung vorausgesetzt – eine Wiederherstellung von Services und der dazu benötigten Infrastruktur durch die Identifikation der erforderlichen CIs und ihrer Konfiguration (Asirati, 2001b).

3.5 Vorschläge für Attribute – Metadatenelemente

Basierend auf der Zusammenstellung der Informationstypen und digitalen Objekte einer CMDB im Abschnitt „Umfang der CMDB“ und der Zusammenstellung der Informationsbedürfnisse der unterschiedlichen Gegenstandsbereiche aus dem Abschnitt „Benutzergruppen und Prozesse mit berechtigtem Interesse am Zugang zur CMDB“ werden als nächsten Schritt der Verfeinerung in den folgenden Unterabschnitten für jeden angeführten Informationstyp die **erforderlichen Metadatenelemente** angeführt und beschrieben.

3.5.1 Attribute für Configuration Items inklusive deren Beziehungen untereinander

Das Standardwerk zum Thema IT Service Management nach ITIL (Assirati, 2001b, 164) sieht die nachstehend aufgelisteten Attribute für Configuration Items, in diesem Fall Hard- und Softwarekomponenten, vor. Um eine bessere Übersicht zu erhalten, wurde für diese Arbeit auch gleichzeitig eine Kategorisierung der Metadatenelemente nach Arms (2000) vorgenommen.

Rachlin und Kneiling (2007) teilen die Attribute weiters ein in **Base Configuration Attributes** zur Beschreibung von allgemeinen Eigenschaften der CIs einerseits und **Extended Attributes** für CI-typische Attribute andererseits, welche meist als Verknüpfungen zu weiteren – mitunter sehr detaillierten – Attributen inklusive der Beschreibung der Zugriffsmöglichkeit auf diese. Diese weiteren Detailinformationen, beispielsweise der Algorithmus für einen Lastverteiler wird nicht in die CMDB kopiert sondern bleibt nur gemeinsam mit der Systemsoftware des Lastvertelers – somit beim „Eigentümer“ der Daten – gespeichert. Durch die erweiterten Attribute ist dieses Attribut zu einem speziellen CI – hier dem Lastverteiler, ein Software-CI – auch spezifiziert und auffindbar.

Identifizier

- **CI Name** als eindeutiger Name des CI.
- **CI Nummer** als fortlaufende und eindeutige Nummer – diese wird in der Literatur (beispielsweise bei Lindinger (2005) oder Vogt (2002)) generell als primäres Unterscheidungsmerkmal der einzelnen CIs geführt Allerdings bedeutet dies nicht, dass damit jede einzelne Instanz einer CI identifizierbar ist: eine CI-Nummer kann für alle PC Mäuse eines Typs stehen.
- **Kopie- bzw. Seriennummer** als Nummer, welche eine Instanz eines CI eindeutig identifiziert. Für Hardware ist dies die Seriennummer, für Software eine eindeutige Kopienummer. Diese unterscheidet bei einem öfters in der Infrastruktur eingesetzten CI die einzelnen Instanzen des CIs wie beispielsweise die Seriennummer bei einem standardmäßig eingesetzten Monitor eines Typs (eine CI-Nummer, mehrere Seriennummern) oder die Kopienummer bei einem verbreitet in der Infrastruktur eingesetzten Softwarepaket (eine CI-Nummer, mehrere Kopiennummern).

Deskriptive Metadaten

- **Kategorie** zur Klassifikation von CIs. Assirati (2001b) führt an dieser Stelle auch die Dokumentation an. Sinnvoll erscheint es dem Autor hier jedoch nur Hard- und Software zu erfassen, denn Dokumente werden im Rahmen dieser Arbeit als eigene Klasse von Bibliotheksobjekten der CMDB betrachtet und nicht den CIs zugerechnet.

Demzufolge wird jegliche Dokumentation zu einem CI nur als Verknüpfung zu diesem erfasst werden. Ein Beispiel für eine Gliederung kann etwa sein:

- Computerhardware,
 - Computerzubehör,
 - Netzwerk,
 - Appliance,
 - Facility.
- **Type oder Unterkategorie** zur Detaillierung der Kategorie: beispielsweise Drucker, Hardware Konfiguration, Softwarepaket, Programmmodul. Ein Beispiel für die Verfeinerung der vorstehenden Kategorie-Gliederung kann etwa sein:
 - Computerhardware: Desktop, Notebook, PDA, Server, Mainframe
 - Computerzubehör: Monitor, Scanner, Drucker, Tastatur, Bandlaufwerke
 - Netzwerk: Unterscheidung in Sprach- und Datennetz
 - Appliance: Firewall
 - Facility: Klimatisierung (Heizung, Lüftung, Kühlung), Stromversorgung, unterbrechungsfreie Stromversorgung, Notstromaggregat, Schränke, Kabeltassen
 - **Modellnummer** Modellbezeichnung des Hardware-CIs, üblicherweise vom Hersteller übernommen (z.B. „Notebook IBM, Modell X66“)
 - **Versionsnummer** des CI
 - **Bemerkung:** Ein Feld für Zusatzinformationen in Textform, beispielsweise über den Unterschied zur Vorversion des CIs

Strukturelle Metadaten

- **Beziehungen zu Parent-CIs** als eindeutige Bezeichner (Name, Kopie, Seriennummer, Modellnummer) des Parent-CIs
- **Beziehungen zu Child-CIs** als eindeutige Bezeichner (Name, Kopie, Seriennummer, Modellnummer) aller Child-CIs
- **Alle anderen Beziehungen** zu CIs ausgenommen Parent-/Child-CIs wie beispielsweise „kann zugreifen auf“ oder „ist installiert auf“

Administrative Metadaten

- **Quelle oder Lieferant (*Source*):** Eigenentwicklung oder von Firma XY gekauft
- **Ende des Zeitraums der Garantie/Gewährleistung:** Angabe darüber, bis zu welchem Datum Garantie- bzw. Gewährleistungsansprüche für ein CI laufen
- **Ort:** Angabe, in welcher Softwarebibliothek oder auf welchem Medium sich ein bestimmtes Software-CI oder in welchem Gebäude und Zimmer sich ein bestimmtes Hardware-CI befindet
- **Verantwortliche Person (*owner*):** Name und/oder die Funktion der für das CI verantwortlichen Person
- **Verantwortlich seit, Datum (*owner responsible*):** gibt an, seit wann die eingetragene Person für das CI verantwortlich ist
- **Sicherheitsstufe** (Vogt, 2002, 71)
- **Lizenznummer** oder ein Verweis auf ein Lizenzabkommen
- **Lieferdatum** – gibt an, wann das betreffende CI an die Organisation geliefert wurde
- **Annahmedatum** – gibt an, wann das betreffende CI von der Organisation getestet und akzeptiert wurde
- **aktueller Status** beinhaltet die Information zum derzeitigen Status, beispielsweise „Geplant“, „Geliefert“, „Im Test“, „In Produktion“ oder „Archiviert“
- **geplanter Status** beinhaltet die Information zum nächsten geplanten Status, beispielsweise für ein CI welches sich derzeit „Im Test“ befindet der nächste geplante Status „In Produktion“. Mit Angabe eines Datums oder Ereignisses, wann diese Änderung eintreten soll
- **RfC-Nummern** für alle ein CI betreffenden Veränderungsanträge
- **Change-Nummern** für alle ein CI betreffende Changes
- **Problemnummern** für alle ein CI betreffende Problemdatensätze (*problem records*)
- **Störungsnummern** für alle ein CI betreffende Störungsdatensätze (*incident records*)

Abhängig von der CI-Kategorie werden oft noch

- (erweiterte) Attribute, beispielsweise die Zugriffsmodalitäten (Datenbanken) und Protokolle (EDI-Links)

zu berücksichtigen sein. Darüber hinaus sollen nach (Vogt, 2002, 71) folgende Informationen zusätzlich zu den genannten Metadaten erfasst werden können:

- Informationen zum auslösenden Prozess und
- Informationen zur Datenverantwortung
- Speicherort des CI-Backup – für Software CIs soll der Speicherort der Backupkopien dokumentiert werden
- Anforderungen an Hard und Software (Betriebssystemversion, eventuell erforderliche Anwendungen und Protokolle)

Letztendlich kann eine

- Verknüpfung mit dem System für die Lagerbewirtschaftung des *Definitive Hardware Stores* (DHS)

für rasche Bestandsabfragen nützlich sein.

3.5.2 Attribute für Services

Identifizier

- Serviceidentifikator – als eine eindeutige Kennzeichnung eines Services
- Name des Services

Deskriptive Metadaten

- Beschreibung des Services – diese wird für die dynamische Erstellung des Servicekatalogs benötigt.
- Kunden – Angabe über die Kunden, welche den Service nutzen – als Verknüpfungen

Strukturelle Metadaten

- Aufbau des Services
- benötigte Unterservices

Administrative Metadaten

- Kunde
- Verantwortlichkeiten für den Service
- Status des Services
- Zeitstempel bei jeder Statusänderung – zur Verfolgung wichtiger Termine wie beispielsweise das Inbetriebnahmedatum

3.5.3 Attribute für Störungen

Um Störungen effizient zu verwalten und die Störungsbehebung effektiv zu unterstützen werden Informationen zu Störungen zusammengefasst und in Form von Störungsdatensätzen in einer Störungsdatenbank gespeichert. Solch ein Datensatz wird auch als Ticket – im Fall von Störungen als **Incident Ticket** – bezeichnet. Die Lebensdauer dieses Tickets beginnt mit der Erfassung der Störung und endet mit dem Abschluss der Störungsbehebung nach der Wiederherstellung. Die eigentliche **Aufgabe eines Tickets** ist

- die Speicherung und
- Kommunikation (Weitergabe) von Störungsdetails für alle an der Wiederherstellung beteiligten Rollen (Personal) sowie
- die Überwachung des Fortschritts und
- die Bereitstellung von Daten in strukturierter Form zur automatische Auswertung für Berichte (darin werden unter anderen Informationen die durchschnittliche und maximale Reaktions-, Reparatur- oder Wiederherstellungszeit weitergegeben).

Zur Speicherung des Incident Tickets werden Metadaten über die Störung erfasst und abgespeichert. Nach Vogt (2002) erhebt und benötigt der Service Desk die folgenden Informationen zur Störungsbehebung:

Identifizier

- Störungsnummer (*incident ticket id*) – eine fortlaufende eindeutige Kennzeichnung einer Störung

Deskriptive Metadaten

- Benutzende Person(en) – neben der Identifizierung der benutzenden Person(en) ist die Anzahl der betroffenen User bei einem Ausfall eine wesentliche Information für das Availability Management
- Betroffener Service – Die Information über den oder die betroffenen Services ist bei einem Ausfall ein weiterer für das Availability Management wichtiger Input
- Auswirkung(en) – sind bei einem Ausfall ebenfalls für das Availability Management wichtig
- Durchgeführte Änderung(en) – als Verknüpfung zu den durchgeführten Veränderungen (Changes)

Strukturelle Metadaten

- Bekannte(r) Fehler – die Ursache(n) für einen Ausfall ist auch für das Availability Management wichtig

Administrative Metadaten

- Kunde – der Kontakt zum Kunden muss gegebenenfalls rasch herstellbar sein
- Meldende Person (*Incident*) oder Komponente (ein automatisch von einem Monitoringtool ausgelöster *Alarm*)
- Vereinbarter Support Level
- Sicherheitsstufe
- Supportstruktur und Verantwortlichkeiten
- administrativer Bearbeiter der Störung und Support Team
- Status
- Umgehungslösung(en)
- Zeitstempel bei jeder Statusänderung – die Zeitpunkte eines Ausfalls und der Wiederherstellung sind für das Availability Management ebenso wichtig wie die Dauer des Ausfalls

3.5.4 Attribute für Probleme und bekannte Fehler

Wenn im Zuge einer schweren Störung die Ursache für diese oder nach einer Häufung gleichartiger Störungen die gemeinsame Ursache der Störungen identifiziert werden muss, wird zur Erfassung der erforderlichen Informationen und zur Verwaltung des Problems ein **Problem Ticket** angelegt. Dieses enthält vergleichbar dem Incident Ticket alle relevanten Informationen. Nach Vogt (2002) werden im Problem Management zur Ursachenforschung und Lösungsentwicklung die folgenden Informationen benötigt:

Identifier

- Problemnummer (*problem ticket id*) – als eine fortlaufende und eindeutige Kennzeichnung eines Problems

Deskriptive Metadaten

- Beschreibung des Problems – dies kann auch durch Verweise auf einzelne Störungen erfolgen
- Service, welcher vom Problem betroffen ist
- CIs, welche vom Problem betroffen sind
- Kategorie – stellt eine Steuergröße in dem Sinn dar, als dass dadurch die Weiterleitung an ein spezialisiertes Expertenteam ausgelöst wird
- Auswirkung – ein Indikator für die Auswirkung stellt die Anzahl der Störungen dar und dient ebenfalls als Steuergröße zur Priorisierung
- Dokumentation des Problems – dies wird durch Verweise auf auf die eigentliche Dokumentation des Expertenteams erfolgen
- Bekannter Fehler – nachdem die (gemeinsame) Ursache der Störung(en) ausgeforscht werden konnte, wird diese als bekannter Fehler registriert und dokumentiert
- Umgehungslösung (*workaround*)
- Lösung (*solution*) für das zum bekannten Fehler gewordene Problem

Strukturelle Metadaten

Im Zusammenhang mit dem Problem Ticket werden strukturelle Informationen zum Problem nicht in einer vorgegebene Menge an Metadatenelementen erfassbar sein, da sich die innere Struktur von Problemen von Fall zu Fall unterscheidet. Die diesbezügliche Beschreibung wird im Rahmen der deskriptiven Metadaten mit der Dokumentation des Problems gemeinsam erfolgen.

Administrative Metadaten

- Dringlichkeit (Steuergröße)
- Priorität (Steuergröße)
- Vereinbarter Service Level
- andere ähnliche Probleme
- Verantwortliche für Support
- Lieferanten
- Offene Changes aus dem *Forward Schedule of Changes*
- Abgeschlossene Changes
- Trends aus dem Capacity Management
- Trends aus dem Availability Management

3.5.5 Attribute für RfC und Change

Zur Bearbeitung von Veränderungsanträgen (RfC) sind in Abhängigkeit von den zu verändernden CIs nach Vogt (2002) die folgenden Informationen erforderlich:

Identifizier

- RfC Nummer oder vergleichbare Kennzeichnung als eindeutige Identifikation

Deskriptive Metadaten

- Verweis auf das zu lösende Problem, falls zutreffend
- Zu verändernde CIs
- Version der zu verändernden CIs
- Betroffene Services
- Kategorie
- Von der Änderung betroffene Kunden
- Antragsteller

- Grund der Änderung – für Veränderungen können definierte Gründe wie Störung, Problem, Kundenwunsch oder Service Request vorgegeben werden
- Auswirkung, falls der Change nicht durchgeführt wird – so kann die Unterlassung der Veränderung zum (erneuten) Auftreten von Störungen führen (die gemeinsame Ursache für mehrere Störungen wurde entdeckt oder wenn der Hersteller einer Komponente eine Aktualisierung empfiehlt), ein Verlust des Wettbewerbsvorteils eines Geschäftsprozesses zur Folge haben (den geänderten Anforderungen des Kunden soll durch einen Change Rechnung getragen werden) oder es droht die Gefahr, dass für veraltete Komponenten keine Wartungs- oder Servicevereinbarung mehr getroffen werden können (ein Software-CI für welches der Support seitens des Herstellers eingestellt wurde)
- Risiken der geplanten Änderung – jede Veränderung birgt das Risiko, dass sich durch die Veränderungen bisher nicht bekannte Störungsquellen ergeben: einerseits durch die Durchführung der Veränderung selbst, andererseits durch die neuen Komponenten oder deren neuartige Kombination, auch wenn diese durch das Release Management bereits freigegeben wurden
- Implementierungsplan
- Backout Verfahren – für den Fall des Scheiterns der Aktualisierung muss klar definiert werden, welche Schritte zum Rückbau auf einen funktionierenden Zustand beziehungsweise für einen Umbau auf einen annehmbaren, funktionierenden Zustand erforderlich sind
- Ergebnis des *Post Implementation Review* (PIR) – zu jeder durchgeführten Veränderung muss einerseits für das Reporting maschinell auswertbar festgehalten werden ob der Change erfolgreich war und ob das Ergebnis der Veränderung auch das im RfC beabsichtigte war sowie andererseits für die zukünftige Fehlervermeidung und Verbesserung in Textform das Resultat im Bezug auf den RfC evaluiert werden
- Auswirkung auf die Servicekontinuität
- Sicherheitsstufe basierend auf den betroffenen CIs und Services

Strukturelle Metadaten

- Beziehungen zwischen den vom Change betroffenen Komponenten

Administrative Metadaten

- Status des Changes (beispielsweise „eröffnet“, „abgelehnt“, „autorisiert“, „durchgeführt“ oder „abgeschlossen“)

- Datum der Beantragung
- Priorität
- Anzahl der zur Veränderung führenden Störungen und deren Auswirkungen
- Risikobewertung und Risikomanagement
- Empfehlung für die Zusammensetzung des *Change Advisory Board*¹⁴ (CAB), welches die Veränderung behandeln und die Entscheidungsgrundlage für die Autorisierung bzw. Nichtautorisierung durch den Change Manager aufbereitet. Die Zusammensetzung erfolgt hinsichtlich des Inhalts und der zu erwartenden Auswirkungen.
- Ressourcenzuweisung
- Autorisierung
- Verantwortliche Person für den Change
- Tatsächliches Implementierungsdatum
- Datum der Erfolgskontrolle (*Post Implementation Review* – PIR)

3.5.6 Attribute für Dokumentation

Unter dem Begriff Dokumentation werden hier alle gedruckten oder digital verfügbaren Schriftstücke unabhängig von deren Umfang zusammengefasst. Darunter fallen insbesondere:

- Installations-, Inbetriebnahme- und Bedienungsanleitungen
- Trainingsunterlagen für die Benutzenden, IT Techniker und Unterrichtende
- Prozessbeschreibungen, Ablaufdiagramme und Organigramme
- *Service Level Agreements*, *Operational Level Agreements* und *Underpinning Contracts* für den Fall, dass sie als Dokumente im Volltext – und nicht als Datensätze in einer eigens dafür bereitgestellten Datenbanktabelle¹⁵ – vorliegen

¹⁴Das *Change Advisory Board* tritt in regelmäßigen Abständen zusammen um die Notwendigkeit, Wirtschaftlichkeit und die Auswirkungen von Veränderungen zu besprechen und somit die Grundlage für die Autorisierung zu legen. In der Praxis werden die Beschlüsse hauptsächlich in Form von Umlaufbeschlüssen gefasst und ermöglichen dadurch umgehende Entscheidungen ohne unnötigen Zeitverlust. (Vogt, 2002, 131)

¹⁵Über die Informationen welche in dieser Tabelle zu Speichern sind können an dieser Stelle keine Empfehlungen abgegeben werden, da die Felder stark von der jeweiligen SLA Struktur und den darin vereinbarten Services abhängen. Der Vorteil der datenbankgestützten Variante von SLA-, OLA- und UC-Vereinbarungen liegt darin, dass auf konkrete Werte, beispielsweise die mittlere und maximale Wiederherstellungszeit, leichter automatisch zugegriffen werden kann.

- Releasegrundsätze (Release Policy)
- Budgets, Finanzpläne und diesbezügliche Zielvorgaben der Organisation an die IT
- Kapazitätspläne, sofern in Dokumentform vorliegend
- Verfügbarkeitspläne, sofern in Dokumentform vorliegend
- IT Service Continuity Pläne, sofern in Dokumentform vorliegend
- Budgets und Finanzpläne

Zur Identifikation, Verwaltung und alphabetischen, inhaltlichen sowie strukturellen Beschreibung sollen als Mindeststandard die 15 grundlegenden Metadatenelemente des Dublin Core – jeweils durch die Angabe „(DC)“ angezeigt – angesetzt werden:

Identifizier

- **Schlüssel** (DC) zur primären und eindeutigen Identifikation
Beispiel: zur Identifikation im lokalen Repository DocId=435
- **URL** als zweiter, allerdings nicht persistenter und durch die CMDB kontrollierter, Schlüssel
Beispiel: die veröffentlichte Version (in Ermangelung einer URN) die Angabe
URL=http://gonzo.uni-weimar.de/~richte10/progsown/Richter_LaTeX_Tipps.pdf

Deskriptive Metadaten

- **Titel** (DC) des Dokuments beziehungsweise der Dokumentation
Beispiel: LATEX - Tipps und Tricks
- **Ersteller** (DC) als jene interne oder externe Person sowie Organisation welche für den intellektuellen Inhalt (Texte, Schema, technische Zeichnungen, Zusammenstellungen, usw.) verantwortlich ist.
Beispiel: Torsten Richter
- **Thema** (DC) des Dokuments – stellt eine Beschreibung des Inhalts in Schlagwörtern dar, wobei diese einem kontrollierten Vokabular oder idealerweise einem Fachthesaurus entnommen werden sollen
Beispiel: Textverarbeitung, Latex, Grundlagenwissen, PDF-Erzeugung, wobei die angegebenen Schlagwörter in diesem Beispiel nicht kontrolliert sind

- **Beschreibung** (DC) in Textform (Abstract)
Beispiel:
Dieser Artikel führt in die Verwendung von Latex zur Dokumentenerstellung ein und gibt Beispiele für die Blatteinteilung, Absatzkontrolle, Verwendung von Boxen, Sonderzeichen, Text hervorhebungen, Gliederungen, Tabellen, Grafiken, Verweise, Erstellung mathematischer Formeln, Silbentrennung sowie die Ausgabe als DVI, PS und PDF.
- **Herausgeber** (DC) welcher für die Veröffentlichung verantwortlich ist
Beispiel: `www.tortools.de`
- **Beiträge von** (DC) Personen oder Organisationen, die nachrangige Beiträge lieferten wie beispielsweise Darstellungen oder Übersetzungen des Dokuments
Beispiel: keine Angabe
- **Datum** (DC) der Erstellung oder Herausgabe des Dokuments
Beispiel: 2004-10-31 als Datum der letzten Aktualisierung
- **Kategorie** (DC) des Dokuments
Beispiel: Bedienungsanleitung oder eventuell Trainingsunterlage
- **Quelle** (DC) auf welcher das aktuelle Dokument basiert
Beispiel: keine Angabe
- **Sprache** (DC) des Inhalts
Beispiel: Deutsch
- **Beziehungen** (DC) zu anderen Objekten (CI, Service, weitere Dokumente, usw.)
Beispiele:
Dokumentation für CI=1529 verknüpft das Dokument mit dem CI 1529, der Softwareinstallationsdatei `TXCSetup_1Beta7_01.exe` in der DSL,
Trainingsdokument für `ServiceName=LatexBereitstellung` verknüpft das Dokument mit einem Service der Latex-Bereitstellung
- **Abdeckung** (DC) im Sinne einer räumlichen und zeitlichen Zuordnung
Beispiel: `Sprachraum=Deutsch`, `VerwendungAb=2007-03-01`, `VerwendungBis=bis auf weiteres`

Strukturelle Metadaten

Strukturelle Metadaten auf der Ebene des digitalen Bibliotheksobjekts sind:

- **Format des digitalen Bibliotheksobjekts** (DC) – als Deskriptoren für Bild (einer Seite), Audio, Video, Software
Beispiel: Textdokument

- **Dateiformat des digitalen Bibliotheksobjekts** – als Hinweis für die zur Präsentation oder Bearbeitung erforderliche Anwendung wie beispielsweise Textdatei, MS Word Dokument, HTML-Dokument, Acrobat PDF Dokument, Bilddateiformat JPEG/TIFF usw., Java Bytecode oder als ausführbare EXE-Datei.
Beispiel: PDF Datei
- **Gliederung des Dokuments** – beschreibt den Aufbau des Bibliotheksobjekts in quantitativer Hinsicht (Anzahl Seiten, Kapitel, Tabellen, grafische Darstellungen, usw.), kann aber auch für nicht digital vorliegenden Bibliotheksobjekte als eingescanntes Inhaltsverzeichnis, das mittels Texterkennung (OCR) bearbeitet und als im Volltext durchsuchbarer Text abgelegt werden
Beispiel: 42 Seiten, Einzeldatei - keine Unterobjekte

Strukturelle Metadaten auf der Ebene der einzelnen digitalen Objekte welche das Bibliotheksobjekts bilden sind:

- **Identifizier des Einzelobjekts** – im Idealfall ein URN
Beispiel: nicht anwendbar
- **Format des Einzelobjekts (Datei)**
Beispiel: nicht anwendbar
- **Reihenfolge des Einzelobjekts** im übergeordneten Bibliotheksobjekt
Beispiel: nicht anwendbar
- **Nutzungsmöglichkeit der Datei** – als ausführbares Verhalten des Objekts, beispielsweise als Inhaltsverzeichnis (mit Hyperlinks) oder als interaktive Darstellung welche aufgrund des Zeigens mit der Maus auf Teilbilder Zusatzinformationen zu diesen anzeigt (was bei einer Flash-Animation möglich sein kann)
Beispiel: mit Inhaltsverzeichnis und Beispielen, Bibliografie
- **Beziehungen** zu anderen digitalen Objekten zur Navigation: über- und untergeordnete Objekte oder gleichrangige Objekte (vor, zurück)
Beispiel: nicht anwendbar

Administrative Metadaten

- **Dateigröße**
Beispiel: 0,8 MB
- **Abmessungen**
Beispiel: A4 hoch

- **Quelle/Original**
Beispiel: die veröffentlichte Version unter
URL=http://gonzo.uni-weimar.de/~richte10/progsown/Richter_LaTeX_Tipps.pdf
- **Art des Originals**
Beispiel: PDF Datei
- **Datum der Erfassung**
Beispiel: Abruf am 2007-03-12
- **Rechte (DC) am Dokument**
Beispiel: Copyright 2001-2004 by Torsten Richter
- **Zugriffseinschränkungen** – welche Benutzergruppen dürfen in welcher Weise auf die Dateien und Informationen zugreifen
Beispiel: keine

3.5.7 Attribute für die Beschreibung eines Release

Zusätzlich zur DSL und den Metadaten in der CMDB werden in dieser auch die *Release Definitions* abgelegt, da in diesen die Testkonfiguration für die Freigabe einer Hard- bzw. Softwarerelease beschrieben sind. Zudem sollen neben den Objekten für *Release Definitions* auch die Relationen zu den betreffenden CIs abgelegt werden, um in späterer Folge – etwa im Rahmen einer Ursachenforschung im Problem Management – rasch und sicher die jeweiligen Testbedingungen eruieren zu können. (Assirati, 2001b)

Diese Dokumente können dann einerseits als Dokumentation und andererseits als Beziehungen zu den CIs erfasst werden, deren Metadatenstruktur bereits unter „Attribute für Dokumentation“ vorgeschlagen wurde.

3.6 Rahmenbedingungen für die Erfassungsprozess

Der Erfolg des Configuration Management wird an der Qualität der CMDB gemessen. Allerdings sind die Nutznießer von qualitativ hochwertigen Daten und Informationen immer nur andere Gegenstandsbereiche. Beispielsweise zeigt Assirati (2001a), dass die Fähigkeit all jene Komponenten zu identifizieren von welchen ein bestimmter Service abhängt, stark von aktuellen und korrekten Informationen in der CMDB abhängt. Dieser Umstand zeigt deutlich, dass die CMDB im wesentlichen für „die anderen“ Gegenstandsbereiche Nutzen stiftet, und die CMDB ohne Abnehmer von Informationen als Selbstzweck nicht erhaltbar ist. Dementsprechend schwierig gestaltet sich eine erfolgreiche Motivation für die Einführung von Configuration Management und die Aufrechterhaltung des definierten Qualitätsniveaus. In der Literatur (Frühauf und Linhofer, 2004) wird das Configuration Management gerne mit Hygienevorschriften in einem Krankenhaus verglichen: sie tragen zwar nicht zur

Gesundung der Patienten bei, sorgen aber dafür, dass diese nicht kranker werden. Trotzdem soll nach Assirati (2001b) wenn immer auch möglich eine zentrale CMDB eingesetzt werden.

Um unter Berücksichtigung dieser Umstände die Vertraulichkeit, Integrität und Verfügbarkeit der in der CMDB und DSL abgelegten Objekte und Informationen sicherzustellen sind die folgenden allgemeinen Maßnahmen angebracht (teilweise nach Assirati (2001b) und Assirati (2001a)):

- Das Configuration Management soll die Integrität der Software-CIs sicherstellen indem es
 - das Speichermedium so wählt, dass Regenerationsfehler und Datenverlust minimal sind,
 - archivierte Software-CIs in Abstimmung mit der Lebensdauer des Speichermediums mittels Refreshing erneuert beziehungsweise auf einen neuen Datenträger umkopiert sowie
 - Duplikate der Datensicherungen an einem anderen Ort als den der CMDB bzw. DSL mit beschränktem Zugriff darauf einlagert damit die Auswirkungen im Katastrophenfall minimiert werden können.
- Die regelmäßige Überprüfung auf archivierte bzw. als gelöscht markierte CIs und Informationen, welche den Anforderungen entsprechend lange genug archiviert wurden, um diese aus der CMDB und der DSL zu entfernen (*housekeeping*).
- Dass insbesondere CIs und ihre Dokumentation durch aktuelle Anti-Virus-Software geschützt ist.
- Der Standort der Hardware des oder der Repository (CMDB und DSL) muss in klimatischer Hinsicht den Spezifikationen der Hersteller entsprechen.
- Werkzeuge und Prozeduren für ein Disaster Recovery müssen bereit gehalten werden.

Von Fall zu Fall zu konkretisierende Maßnahmen hinsichtlich der Gestaltung von Erfassungs- und Aktualisierungsprozeduren lassen sich wie folgt zusammenzustellen (Assirati, 2001a):

- Es soll versucht werden, möglichst viele Schritte zu automatisieren. Assirati (2001a) zufolge sollen beginnend mit der initialen Befüllung der CMDB möglichst viele Daten automatisationsunterstützt oder gar vollautomatisch übernommen und aktualisiert werden können. Von Seiten des Configuration Management sind entsprechende Tools für die Erfassung, Inventarisierung und für Audits anzustreben. Weiters sind Bemühungen zu unternehmen, die vorhandenen Tools für Netzwerkmanagement oder Assetmanagement mit Schnittstellen zur CMDB auszustatten. Ein möglichst hoher Automtisationsgrad reduziert reduziert Fehler und die Erfassungs- bzw. Aktualisierungskosten.

- Der Zugriff auf die CMDB soll klar geregelt und die Zugriffsrechte durchgesetzt werden.
- Ein wirksamer Schutz vor nicht autorisierten Veränderungen und Zerstörungen muss vorgesehen werden.
- Nur berechtigten Personen darf die Erstellung von kontrollierten Kopien aus der DSL und aus dem Repository für Dokumentationen möglich sein.
- Falls die DSL nicht integraler Bestandteil der CMDB ist, sollen Tools für den automatischen Abgleich zwischen DSL und CMDB sorgen.
- Damit die Verantwortlichkeit für die CMDB und die DSL konkret an einer in Linienverantwortung stehenden Person festgemacht werden kann, empfiehlt Assirati (2001b) die Nominierung einer Person für die Rolle des bzw. der *Configuration Librarian*, welcher für die Masterkopien aller Software und Dokumentationen verantwortlich zeichnet.

In weiterer Folge werden Besonderheiten bei hervorzuhebenden Anwendungsfällen zusammengestellt:

3.6.1 Bei einem neuen Release

Eines der Ziele des Release Managements ist die Sicherung von Masterkopien von – durch das Release Management kontrollierten – Software in der DSL und dass die diesbezüglichen Metadaten in der CMDB aktuell gehalten sind. Dazu wird im Rahmen des ITIL Frameworks die Verantwortung wie folgt umgesetzt: erstens die Verantwortung für die Aufbewahrung von Objekten und zweitens jene für die Aktualisierung der Objekte und dazugehörigen Metadaten. In die Verantwortung der Aufbewahrung fällt, dass alle Software, Parameter, Testdaten, Laufzeitumgebungen (*run-time environments*) und alle weitere für diesen Release erforderliche Software der Kontrolle des Configuration Managements übergeben werden. Andererseits liegt die Verantwortung für Aktualisierungen der CMDB nach Änderungen in der DSL ebenfalls beim Release Management Prozess. (Assirati, 2001b)

Die zu setzenden Rahmenbedingungen müssen sicherstellen, dass zusätzlich zu den *Release Definitions* (Testkonfiguration für die Freigabe) auch die Relationen zu den betreffenden CIs in der CMDB und DSL abgelegt werden, um in späterer Folge – etwa im Rahmen einer Ursachenforschung im Problem Management – rasch und sicher die jeweiligen Testbedingungen eruieren zu können.

Als möglicher Indikator für die Überwachung der diesbezüglichen Performance nennt Assirati (2001b) die zeitnahe Aufzeichnung aller Informationen zu Aktivitäten von Build, Verteilung und Implementierung.

3.6.2 Bei Veränderungen

Nach jeder Veränderung müssen die betroffenen CIs so rasch wie möglich aktualisiert werden. Im Idealfall erfolgt dies automatisch, wenn die Tools für Change Management und die CMDB ausreichend integriert sind.

Vogt (2002) nennt die Ereignisse bei oder vor welchen die CMDB zu aktualisieren ist als die folgenden:

- Neue Komponenten werden übernommen
- Bei jeder Statusänderung einer Komponente
- Bei Änderung der das CI innehabenden Person
- Bei einer Änderung des Standorts des CI
- Bei Änderungen von Abhängigkeiten (Relationen)
- An Ende eines CI-Lebenszyklus und nach der Entsorgung
- Bei Feststellung eines Fehlers in der CMDB

Zu Kontrollzwecken ist regelmäßig zu überprüfen, ob die Daten der autorisierten Veränderungen auch korrekt in der CMDB eingetragen wurden. Vogt (2002) empfiehlt dazu immer kürzere Zeitabstände zwischen den Überprüfungen je mehr fehlerhaft verwaltete CIs aufgedeckt werden.

3.6.3 Bei Störungen

Bei jeder Störung sollen neben dem betroffenen CI auch die Angabe zur benutzenden Person erfasst werden, damit in der CMDB die entsprechenden Relationen zwischen Störung, CI und User hergestellt werden können. Diese Relationen unterstützen die Eingrenzung und Beseitigung von Störungen sowie die Ursachenforschung.

Das für die CMDB verwendete Informationssystem muss nach Assirati (2001a) außerdem in der Lage sein, eine Historie bestehend aus Störungsdatensätzen, Problemdatensätzen, Known Errors sowie Change-Datensätzen ohne Benutzereingriff zu generieren. Dadurch wird die Störungserfassung sowohl in qualitativer Hinsicht als auch im Hinblick auf den Zeitaufwand verbessert und das Reporting erleichtert.

Als weitere Empfehlung nennt Assirati (2001a) die Ablage von Supportvereinbarungen und Eskalationsregelungen in der CMDB. Dies soll so geschehen, dass ein Zugriff darauf sehr leicht möglich ist.

3.6.4 Bei der Bearbeitung von Problemen

In Analogie zur Störungsbehebung sollen bei jedem Problem die betroffenen CIs und Benutzenden erfasst werden, damit eine entsprechenden Beziehung zwischen Problem, CI und User aus Daten der CMDB hergestellt werden kann.

Die CMDB muss auch in der Lage sein, Problem Datensätze mit Störungsdatensätzen in Verbindung bringen zu können (Assirati (2001b) und Assirati (2001a)), wobei nach Assirati (2001b) unerheblich ist, ob sich die referenzierten Datensätze und die Problem Datensätze in einer CMDB befinden oder ob getrennte Problem Datenbanken existieren.

3.6.5 Datenqualität überprüfen, sicherstellen und verbessern

In wiederholten Maßnahmen ist die Qualität der Metadaten und der digitalen Objekte einer CMDB zu überprüfen. Vogt (2002)

Als Zeitpunkte für eine Überprüfung der Datenqualität sind nach Vogt (2002) geeignet:

- in regelmäßigen Abständen
- vor und nach erheblichen und weitreichenden Veränderungen zur Überprüfung auf korrekte und autorisierte CIs
- nach einem Disaster Recovery
- nach der Entdeckung eines CI, welches ohne autorisierten Change implementiert, verändert oder entfernt wurde
- vor Auslieferung eines Software Release oder einer Konfigurationsänderung
- nach der Implementation eines neuen Configuration Management Systems (Prozesse, Werkzeuge)

Dabei kann der Umfang der Überprüfung folgendermaßen gewählt sein:

- im vollen Umfang – wird aus Kapazitätsgründen kaum möglich sein
- auf von einem Change oder einem neuen Release betroffenen CIs beschränkt
- durch zufällige Auswahl von Stichproben

Die Überprüfung kann **manuell** durch eine visuelle (z.B. Druckertyp ablesen) und analytische Kontrolle (z.B. beim selben Drucker zusätzlich die Konfigurationsseite ausdrucken) von Eigenschaften erfolgen, welche mit dem Inhalt von Attributen der CIs in der CMDB verglichen werden. Alternativ dazu bietet sich die Überprüfung **mittels automatisierten**

Werkzeugen, beispielsweise Inventarisierungstools, an (Vogt, 2002, 61). Dabei wird die sich in der IT Infrastruktur befindliche Hard- und Software automatisiert auditiert und die so erfasste Realität mit den Informationen in der CMDB verglichen.

Um gesicherte Informationen über die Ausgangslage für die weitere Entwicklung der IT Infrastruktur zur Verfügung zu haben, werden **Configuration Baselines** erstellt, welche einen funktionierenden Zustand in der Form eines „Schnappschusses“ der CMDB konservieren. Diese Informationen werden im Fall von nicht erfolgreich implementierten Changes für ein Roll-back oder Fall-back benötigt. (Vogt, 2002, 63)

3.7 Umsetzung und Verwertung der Ergebnisse

An dieser Stelle sollen die Möglichkeiten der Verwertung der Ergebnisse der Fallstudie erörtert werden.

Das Ziel des Fallbeispiels im Praxisteil der Arbeit war die Ermittlung und die Dokumentation von Kriterien für den Entwurf beziehungsweise die Auswahl einer *Configuration Management Database* (CMDB) für den Einsatz im IT Service Management. Dabei wurde der Schwerpunkt bewusst auf die Ebene der Metadaten gelegt. Um eine möglichst hohe Datenqualität zu erreichen, wurden Metadatenelemente hinsichtlich ihrer Fähigkeit ausgewählt, die Qualität und Leistungsfähigkeit der CMDB zu verbessern: Diese Metadatenelemente müssen in der Lage sein, eine inhaltliche und alphabetische Katalogisierung zu ermöglichen und die Administration der digitalen Bibliotheksobjekte wie Dokumentationen und Softwarekomponenten nachvollziehbar zu machen und zu unterstützen. Ein nicht zu vernachlässigender Aspekt stellt die Dokumentation von Urheber- und Nutzungsrechten an digitalen Bibliotheksobjekten dar.

Diese im Abschnitt „3.5 Vorschläge für Attribute – Metadatenelemente“ gemachten Vorschläge ab Seite 76 können herangezogen werden, um Mindestanforderungen an eine *Configuration Management Database* festzulegen. Dieser Nutzen ist unabhängig davon, ob die CMDB zugekauft oder *in-house* entwickelt werden soll. Konkret werden im genannten Abschnitt die erforderlichen Metadatenelemente für *Configuration Items* (CIs), Störungen, Probleme und bekannte Fehler, *Request for Changes* (RfCs), Releasedokumentation und Dokumentationen zu Komponenten betreffend zusammengefasst und decken somit den Kern der Informationsbedürfnisse im Rahmen des IT Service Managements ab.

Kapitel 4

Zusammenfassung

Auf den ersten Blick könnte ein Zusammenhang zwischen *digitalen Bibliotheken* und der *Configuration Management Database* (CMDB) als an den Haaren herbeigezogen gelten. Ausgehend vom umfangreichen Konzept der digitalen Bibliothek wird in dieser Arbeit dargestellt, dass eine *Configuration Management Database* in erheblichem Umfang eine digitale Bibliothek darstellt.

So unterschiedlich die Anforderungen auf den ersten Blick auch sind, so vergleichbar sind die Hürden der digitalen Wissensorganisation, der Steuerung des Zugriffs auf die Informationen, der Erhaltung der digitalen Objekte inklusive Datensicherung der Repositories und den Einschränkungen digitaler (Datei-) Formate: Von digitalen Bibliotheken wird die akkurate Auflistung von Nachweisen über digitale Bibliotheksobjekte, die Konservierung dieser und der oft multimedial aufbereitete Zugang zu den Objekten für ein breites Publikum erwartet. Von einer CMDB hingegen werden ad hoc genaue Informationen über die oft weltweit verteilten und relativ kurzlebigen Komponenten von IT Infrastrukturen verlangt.

Dabei zeichnen sich zwei Trends ab, wonach sich die Gegenstandsbereiche der digitalen Bibliothek und jener der *Configuration Management Database* einander mehr und mehr überdecken:

- Für die digitale Bibliothek lässt sich die Entwicklung zu einer ausgeprägten Serviceorientierung hin beobachten.
- Für die CMDB ist es die Entwicklung hin zu einer *Wissensbasis*, welche bereits durch die kommende Version von *ITIL V3 (Refresh)*, vorgezeichnet wird. Nachvollziehbarkeit wird immer wichtiger und die Ansprüche an die Datenqualität steigen. Dadurch entsteht ein Aufholbedarf in der Entwicklung von Metadaten zum Entstehungszusammenhang und administrativen Metadaten, insbesondere zu Refresh-, Migrations- und Emulationsschritten. Weitere Tatsache ist, dass die Archive mit ausgedienten CIs tendenziell umfangreicher werden, wobei die Archivfunktion der CMDB im Zusammenhang mit wirtschaftlich orientierten, beziehungsweise gewinnorientierten Organisationen wohl nie die selbe Bedeutung erlangen wird, wie jene im Zusammenhang mit Kunst- und Kulturgütern.

Es ist festzuhalten, dass die aktuelle Literatur zum Thema der *Configuration Management Database* primär deskriptive Metadaten fordert, in geringerem Ausmaß administrative Metadaten. Abhängig von der Art des Configuration Items oder anderen digitalen Objekts finden sich strukturelle Metadaten häufig in der Form von Relationen zwischen den digitalen Objekten. Nur pauschale oder ungenaue Verweise beziehungsweise gar fehlende Hinweise, werden zu den Themen Urheber- und Nutzungsrechte, Zugriffsteuerung oder der Dauerhaftigkeit von digitalen Objekten (Persistence) eingefordert. Entweder werden diese als Selbstverständlichkeit angesehen und daher nicht spezifiziert oder einfach nur ignoriert. Für eine erfolgreiche Realisierung und um die Qualität der Metadaten zu verbessern ist ein dualer Ansatz wünschenswert, indem

- einerseits die Qualitätsansprüche bereits in der Planungsphase der CMDB in der Form von zusätzlichen Metadatenelementen eingebracht werden und damit später die qualitätsrelevanten Informationen auch abgelegt werden können und
- andererseits durch die qualitätsfördernde Ausgestaltung der Prozessschritte Fehlerquellen ausgeschaltet werden so dass die Informationen auch korrekt erhoben werden und die Erfassung möglichst fehlerfrei erfolgen kann.

Da die Bestimmung des Grads der Outputverbesserung aufgrund von Änderungen der Metadatenstruktur für die gesamte (digitale) Bibliothek sehr komplex ist und in dieser Arbeit nur auf allgemeiner Ebene – und nicht für jeden CI-Typ und CI-Untertyp einzeln – möglich war, wären die nächsten Schritte die Analyse auf der Ebene der einzelnen Repository der CMDB.

Literaturverzeichnis

- Arms, William Y. 2000. *Digital Libraries*. URL <http://www.cs.cornell.edu/wya/DigLib/MS1999/index.html>, Abruf am 2007-03-09.
- Assirati, Bob. 2001a. *Service Delivery*. London: Office of Government Commerce, The Stationary Office, seventh impression 2004 edition.
- Assirati, Bob. 2001b. *Service Support*. London: Office of Government Commerce, The Stationary Office, eighth impression 2004 edition.
- Atteslander, Peter. 1995. *Methoden der empirischen Sozialforschung*. Berlin – New York: de Gruyter, 8., bearb. Aufl.
- Barton, Jane, Sarah Currier und Jessie M. N. Hey. 2003. *Building Quality Assurance into Metadata Creation: an Analysis based on the Learning Objects and e-Prints Communities of Practice*. URL http://www.siderean.com/dc2003/201_paper60.pdf, Abruf am 2007-03-23.
- Beall, Jeffrey. 2005. *Metadata and Data Quality Problems in the Digital Library*. URL <http://jodi.tamu.edu/Articles/v06/i03/Beall/Beall.pdf>, Abruf am 2007-03-05.
- Berners-Lee, Tim, Roy T. Fielding und Larry Masinter. 1998. *Uniform Resource Identifiers (URI): Generic Syntax. Request for Comments No. 2396*. The Internet Society, URL <http://www.ietf.org/rfc/rfc2396.txt>, Abruf am 2007-04-24.
- Berners-Lee, Tim, Roy T. Fielding und Larry Masinter. 2005. *Uniform Resource Identifiers (URI): Generic Syntax. Request for Comments No. 3986*. The Internet Society, URL <http://www.ietf.org/rfc/rfc3986.txt>, Abruf am 2007-04-24.
- Block, Barbara, Christel Hengel, Reinhold Heuvelmann, Cornelia Katz, Beate Rusch et al. 2005. *Maschinelles Austauschformat für Bibliotheken und die Functional Requirements for Bibliographic Records*. Berlin: Konrad-Zuse-Zentrum für Informationstechnik Berlin, URL <http://www.zib.de/Publications/Reports/ZR-05-60.pdf>, Abruf am 2007-03-26.
- Borowski, Marion (Hrsg.). 2007. *Bestandsaufnahme zur Digitalisierung von Kulturgut und Handlungsfelder*. URL <http://www.bundesregierung.de/Content/>

- DE/Artikel/2007/01/anlagen/2007-01-11-fraunhofer-studie-pdf-format, property=publicationFile.pdf, Abruf am 2007-03-28.
- BS15000 Associates Group (Hrsg.). 2005. *The ISO 20000 (BS15000 / BS 15000) ITSM Standard*. Preston (UK): BS15000 Associates Group, URL <http://www.bs15000.org.uk/>, Abruf am 2007-03-28.
- Buchanan, Brian. 1989. *Bibliothekarische Klassifikationstheorie*. München: Saur.
- Bundesministerium des Inneren (Hrsg. 2006. *ITIL und Standards für IT-Prozesse - Prozess-Standards für die Entwicklung der IT-Service-Organisation gemäß ITIL Best Practices*. Berlin: Bundesministerium des Inneren, URL http://www.kbst.bund.de/cln_047/nn_991292/SharedDocs/Anlagen-kbst/itil_und_standards_fuer_it_prozesse,templateId=raw,property=publicationFile.pdf/itil_und_standards_fuer_it_prozesse.pdf, version 1.0.1, Abruf am 2007-03-27.
- Connolly, Dan. 2004. *Overview of SGML Resources*. W3C, URL <http://www.w3.org/MarkUp/SGML/>, Abruf am 2007-04-20.
- Daigle, L., D. van Gulik, R. Iannella und P. Faltstrom. 1999. *URN Namespace Definition Mechanisms*. The Internet Society, URL <http://www.ietf.org/rfc/rfc2611.txt>, Abruf am 2007-04-24.
- Debenham, John. 1998. *Knowledge Engineering. Unifying Knowledge Base and Database Design*. Berlin - Heidelberg: Oldenbourg.
- Dugmore, Jenny. 2005. *ISO/IEC 20000 benchmarks provision of IT service management*. Genf: ISO, URL <http://www.iso.org/iso/en/commcentre/pressreleases/archives/2005/Ref985.html>, Presseartikel der ISO Ref.: 985, Abruf am 2007-03-26.
- Ferber, Reginald. 2003. *Information Retrieval - Suchmodelle und Data-Mining-Verfahren für Textsammlungen und das Web*. Heidelberg: dpunkt.verlag.
- Fischer, Joachim. 1999. *Informationswirtschaft: Anwendungsmanagement*. Wien: Oldenbourg.
- Frühauf, Karol und Gerald Linhofer. 2004. *Konfigurationsmanagement - das Bindeglied der IT-Prozesse*. HMD - Praxis der Wirtschaftsinformatik, 237: 114–127; Zitiert nach: Silvia Knittl (2006): Entwicklung eines Konzepts zur plattformübergreifenden Integration des Release-Managements bei der HVBInfo, S. 69.
- Fuchs-Kittowski, Klaus, Hubert Laitko, Heinrich Parthey und Walther (Hrsg.) Umstätter. 2000. *Wissenschaft und Digitale Bibliothek – Wissenschaftsforschung, Jahrbuch 1998*. Berlin: Gesellschaft für Wissenschaftsforschung, URL www.wissenschaftsforschung.de/Jahrbuch1998.pdf, 1. Auflage 2000, Abruf am 2007-03-20.

- Glenfis AG (Hrsg.). 2007. *ITIL V3 - Service Life Cycle*. Hünenberg, CH: Glenfis AG , URL <http://www.itil.org/de/itilv3-servicelifecycle/index.php>, Abruf am 2007-05-01.
- Green, Ann, JoAnn Dionne und Martin Dennis. 1999. *Preserving the Whole - A Two-Track Approach to Rescuing Social Science Data and Metadata*. Massachusetts: The Digital Library Federation, URL <http://www.diglib.org/pubs/dlf088/dlf088.pdf>, Abruf am 2007-03-13.
- Hainschink, Bettina (Hrsg.). 2007. *Paradigmenwechsel ITIL V3 vs. ITIL V2 (Presseaus-sendung)*. Wien: presstext Nachrichtenagentur GmbH, URL <http://www.presstext.at/pte.mc?pte=070430016>, Abruf am 2007-04-30.
- Heinsohn, Jochen und Rolf Socher-Ambrosius. 1999. *Wissensverarbeitung. Eine Einführung*. Heidelberg - Berlin: Spektrum Akademischer Verlag.
- HiSolutions AG (Hrsg.). 2005. *ITIL und Informationssicherheit - Möglichkeiten und Chancen des Zusammenwirkens von IT-Sicherheit und IT-Service-Management*. Bonn: Bundesamt für Sicherheit in der Informationstechnik, URL www.bsi.de/literat/studien/ITinf/itil.pdf, Abruf am 2007-04-29.
- Hurley, Bernard J., John Price-Wilkin, Merrilee Proffitt und Howard Besser. 1999. *The Making of America II Testbed Project: A Digital Library Service Model*. Massachusetts: The Digital Library Federation, URL <http://www.diglib.org/pubs/dlf089/dlf089.pdf>, Abruf am 2007-03-13.
- ISO. 2007a. *ISO/IEC 20000-1:2005 Information technology – Service management – Part 1: Specification*. ISO, URL <http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=41332>, Abruf am 2007-03-28.
- ISO. 2007b. *ISO/IEC 20000-2:2005 Information technology – Service management – Part 2: Code of practice*. ISO, URL <http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=41333>, Abruf am 2007-03-28.
- IT Service Management Forum Deutschland e.V. 2006. *IT Service Management - Zeitschrift des itSMF Deutschland e.V.* Heidelberg: dpunkt.verlag, 1. Jahrgang, Heft 2.
- Janko, Wolfgang H. 1998. *Informationswirtschaft 1 - Grundlagen der Informatik für die Informationswirtschaft*. Berlin: Springer, 2., überarbeitete Auflage.
- Köhler, Ralph. 2005. *Bibliographienkunde*. URL <http://www.bibliothek.thh-friedensau.de/uploads/media/Bibliographienkunde-1.ppt>, Abruf am 2007-04-16.
- Lagoze, Carl und Herbert van de Sompel. 2004. *The Open Archives Initiative Protocol for Metadata Harvesting*. URL <http://www.openarchives.org/OAI/openarchivesprotocol.html>, Abruf am 2007-03-13.

- Lanham, Paul (Hrsg.). 2005. *ISO/IEC 20000 Unveiled, the new International IT Service Management Standard - Press Release*. London: BSI British Standards, URL <http://www.bsi-global.com/en/About-BSI/News-Room/BSI-News-Content/Sectors/ICT--Telecommunications/ISOIEC-20000-Unveiled-the-new-International-IT-Service-Management-Standard/?recid=3034>, Abruf am 2007-03-28.
- Lawrence, Gregory W., William R. Kehoe, Oya Y. Rieger, William H. Walters und Anne R. Kenney. 2000. *Risk Management of Digital Information: A File Format Investigation*. Washington: Council on Library and Information Resources (CLIR), URL <http://www.clir.org/pubs/reports/pub93/contents.html>, Abruf am 2007-04-28.
- Lepsky, Klaus. 2005. *Methoden und Verfahren des Information Retrieval*. Köln, URL www.iws.fh-koeln.de/institut/personen/lepsy/Skript-2-Methoden-und-Verfahren-des-Information-Retrieval.pdf, Abruf am 2007-04-16.
- Library of Congress (Hrsg.). 2007. *Library of Congress Classification Outline*. Washington: Library of Congress, URL <http://www.loc.gov/catdir/cpso/lcco/>, Abruf am 2007-05-19.
- Lindinger, Markus. 2005. *Templates als Hilfsmittel zur erfolgreichen ITIL Einführung*. Wien: MASTERS Consulting GmbH, URL <http://www.conect.at/files/papers/20050928/Lindinger.pdf>, Abruf am 2007-05-01.
- Manecke, Hans-Jürgen. 2004. *Klassifikation, Klassieren*. In: Grundlagen der praktischen Information und Dokumentation. Hrsg. Rainer Kuhlen et al.- 5. völlig neu gefasste Ausg.- München u.a.: Saur, 2004.
- Marieke, G., A. Powell und M. Day. 2004. *Improving the Quality of Metadata in Eprint Archives*. Ariadne, URL <http://www.ariadne.ac.uk/issue38/guy/>, Abruf am 2007-03-05.
- Morris, James M. 1998. *Cornell Project Will Assess Risks of Migration Strategy*. Washington: Council of Library and Information Resources (CLIR), URL <http://www.clir.org/pubs/issues/issues04.html#cornell>, Abruf am 2007-03-13.
- OCLC (Hrsg.). 2003. *Introduction to Decimal Classification*. Dublin, Ohio, USA: OCLC Online Computer Library Center, Inc., URL <http://www.oclc.org/dewey/versions/ddc22print/intro.pdf>, Abruf am 2007-04-16.
- OCLC (Hrsg.). 2006. *Dewey services*. Dublin, Ohio, USA: OCLC Online Computer Library Center, Inc., URL <http://www.oclc.org/dewey/>, Webseite der OCLC. Abruf am 2007-04-16.

- Panny, Wolfgang. 2004. *Information Retrieval*. Wien: Wirtschaftsuniversität Wien – Abteilung für Informationswirtschaft, Skriptum zur Vorlesung.
- Presstext Austria (Hrsg.). 2007. *Presseaussendung: Erstes Unternehmen mit ISO-20000 Zertifikat ist in KMU - Logic4BIZ: neuer IT-Standard verdoppelt Effizienz im Kundenservice (Presseaussendung)*. Wien: Presstext Austria, URL <http://www.presstext.at/pte.mc?pte=070329007>, Abruf am 2007-03-29.
- Probst, Gilbert J. B. und Birgit Knaese. 1998. *Risikofaktor Wissen - Wie Banken sich vor Wissensverlusten schützen*. Wiesbaden: Gabler.
- Puglia, Steven, Jeffrey Reed und Erin Rhodes. 2004. *Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images*. College Park, MD, USA: U.S. National Archives and Records Administration (NARA), URL <http://www.archives.gov/research/arc/digitizing-archival-materials.pdf>, Abruf am 2007-03-13.
- Rachlin, Seth und John Kneiling. 2007. *Aperture VISTA and the CMDB: An Enterprise Best Practices Approach - Aperture White Paper*. Stamford, CT: TechPar Group, URL <http://www.aperture.com/program/whitepaperweb/vistaandthecmdb1206.pdf>, erfordert Registrierung, Abruf am 2007-05-06.
- Robertson, R.J. 2005. *Metadata Quality: Implications for Library and Information Science Professionals*, volume 54 no. 5. Library Review, zitiert nach: Beall (2005).
- Ruddy, David. 2006. *Course 4: Metadata and Digital Library Development*. URL http://www.ala.org/ala/alcts/alctsconted/alctscevents/workshops/ceig4_schedule.doc, Abruf am 2007-03-23.
- Schulz, Ursula. 2006. *Lerneinheit 2: Klassifizieren, Klassifikationen*. Hamburg, URL <http://www.bui.haw-hamburg.de/pers/ursula.schulz/worg1/le-2.html>, Abruf am 2007-04-13.
- Shreeves, Sarah L., Ellen M. Knutson, Besiki Stvilia, Carole L. Palmer, Michael B. Twidale et al. 2005. *Is Quality Metadata Shareable Metadata? The Implications of Local Metadata Practices for Federated Collections*. URL <http://www.ala.org/ala/acrl/acrlvents/shreeves05.pdf>, Abruf am 2007-03-21.
- Österreichische Zentralbibliothek für Physik (Hrsg.). 2007. *Bibliographie - Leitfaden zur konventionellen Bibliographie mit besonderer Berücksichtigung von CD-ROM- und Web-Ressourcen*. Wien: Österreichische Zentralbibliothek für Physik, URL <http://www.zbp.univie.ac.at/gj/konven/>, Abruf am 2007-04-16.
- Stock, Walter G. 2000. *Was ist eine Publikation?* In: Fuchs-Kittowski et al. (2000), Seiten 239-282.

- Tennant, Roy. 1999. *Digital v. Electronic v. Virtual Libraries*. University of California Regents, URL <http://sunsite.berkeley.edu/mydefinitions.html>, Abruf am 2007-03-13.
- The APM Group Limited (APMG) (Hrsg.). 2005. *ITIL Refresh Statement*. Buckinghamshire: APMG, OGC and TSO, URL <http://www.itil-officialsite.com/AboutITIL/ITILRefreshStatement.asp>, Abruf am 2007-05-01.
- Turbitt, Ken. 2006. *ISO 20000: What's an Organization to Do?* ohne Ortsangabe: BMC Software, URL <http://www.bmc.com/products/documents/52/17/65217/65217.pdf>, Abruf am 2007-03-28, zum Abruf ist eine Registrierung erforderlich.
- UDC Consortium (Hrsg.). 2006. *Outline of the UDC*. PO Box 90407, 2509 LK The Hague, The Netherlands, URL <http://www.udcc.org/outline/outline.htm>, Abruf am 2007-04-16.
- Umlauf, Konrad. 2007. *Einführung in die Regeln für den Schlagwortkatalog RSWK*. Berlin, URL <http://www.ib.hu-berlin.de/~kumlau/handreichungen/h66/>, Abruf am 2007-04-16.
- Umstätter, Walther. 2000. *Die Rolle der Digitalen Bibliothek in der modernen Wissenschaft*. In: Fuchs-Kittowski et al. (2000), Seiten 297ff.
- UNESCO (Hrsg.). 2007. *Convention for the Protection of Cultural Property in the Event of Armed Conflict - 1954*. Paris: United Nations Educational, Scientific and Cultural Organization (UNESCO), URL http://portal.unesco.org/culture/en/ev.php-URL_ID=8450&URL_DO=DO_TOPIC&URL_SECTION=201.html, Abruf am 2007-05-02.
- Universität des Saarlandes (Hrsg.). 2007. *TerminosaurusREX. Die Informationswissenschaft in Begriffen*. D-66041 Saarbrücken: Universität des Saarlandes / Informationswissenschaft, URL <http://server02.is.uni-sb.de/trex/>, Datenbank, Abruf am 2007-04-12.
- URI Planning Interest Group. 2001. *URIs, URLs, and URNs: Clarifications and Recommendations 1.0 - Report from the joint W3C/IETF URI Planning Interest Group. W3C Note 21 September 2001*. W3C, URL <http://www.w3.org/TR/2001/NOTE-uri-clarification-20010921/>, Abruf am 2007-04-24.
- Vogt, Walter. 2002. *fIT for benefit – Services kundenorientiert planen und steuern*. Basel, CH: Perseo Consult AG.
- Vorbach, Stefan. 2006. *Qualitätsmanagement*. Graz: Institut für Innovations- und Umweltmanagement, URL www.uni-graz.at/qm-ws06-07_kapitel_1.pdf, 7., überarbeitete Auflage, Abruf am 2007-04-29.
- Wagner-Döbler, Roland. 2000. *Was ist eine Bibliothek?* In: Fuchs-Kittowski et al. (2000), Seiten 283-295.

- Walter, Sven Markus und Helmut Krcmar. 2006. *Reorganisation der IT-Prozesse auf Basis von Referenzmodellen – eine kritische Analyse*. Heidelberg: dpunkt.verlag, in: IT Service Management Forum Deutschland e.V. (2006), Seiten 3-10.
- Waters, Donald J. 1998. *What Are Digital Libraries?* Washington: Council of Library and Information Resources (CLIR), URL <http://www.clir.org/pubs/issues/issues04.html#dlf>, Abruf am 2007-03-13.
- Wilson, Amanda. 2004. *Optimizing Metadata Generation Practices*. URL <http://www.ala.org/ala/altscontent/catalogingsection/catdisc/catalogingresearch/ResDGan04.doc>, Abruf am 2007-03-23.