

Museen und Archive im BSZ

**Zweijahresbericht
2006/2007**

**Bibliotheksservice-Zentrum
Baden-Württemberg**

BSZ

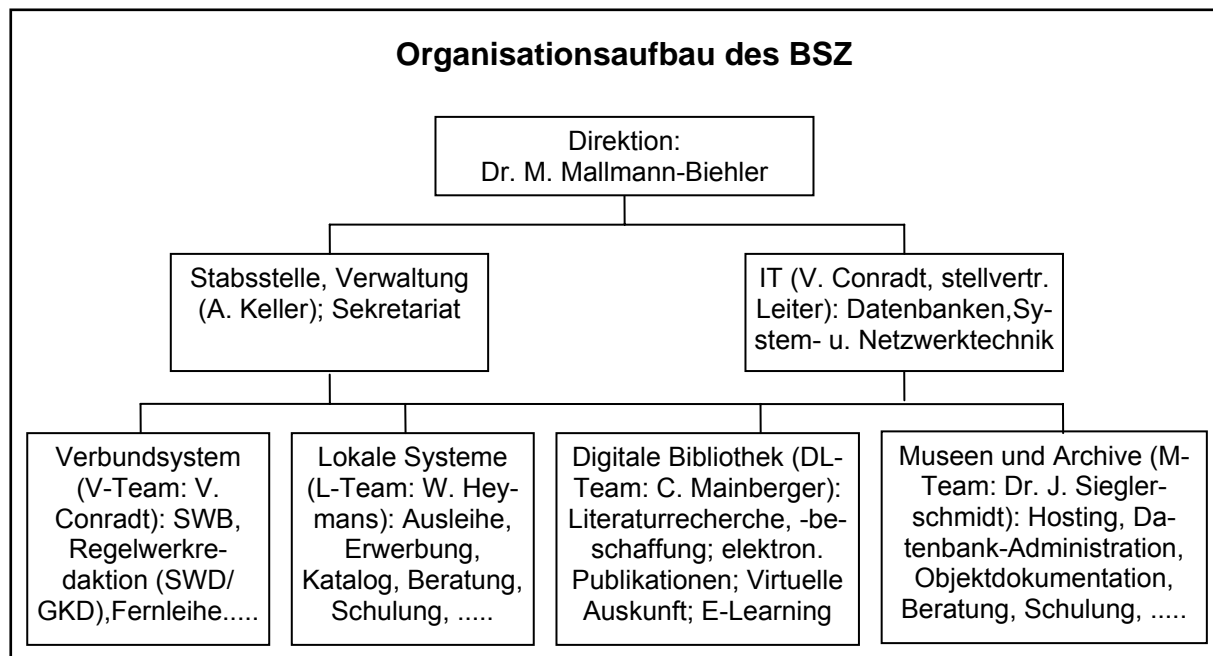
Vorbemerkung

Das *Bibliotheksservice-Zentrum Baden-Württemberg* (BSZ) ist als unselbständige Anstalt des öffentlichen Rechts eine Einrichtung des Landes Baden-Württemberg unter der Dienst- und Fachaufsicht des Ministeriums für Wissenschaft, Forschung und Kunst. Seit dem Jahr 2000 betreut es neben Bibliotheken auch Museen und Archive. In der Satzung sind die Dienstleistungen für Museen inzwischen verankert und damit Bestandteil der Fachaufgaben des BSZ.¹ MusIS ist das Akronym für Museums-Informationssystem.² Das M-Team betreut innerhalb des BSZ die staatlichen Museen in Baden-Württemberg und weitere Museen³. Darüber hinaus sind im Berichtszeitraum die Zuständigkeiten für das BAM-Portal und die Archive hinzugekommen.

Dieser Bericht gibt einen Überblick über die Tätigkeiten des M-Teams in den beiden vergangenen Jahren und versucht, neben dem Sachstand auch Perspektiven aufzuzeigen. Er ist online über die Seiten des BSZ zugänglich.⁴

Dr. Jörn Sieglerschmidt

Konstanz, im März 2008



¹ Satzung in der Fassung vom 6. März 2003, § 2, Ziffer 1: <<http://www2.bsz-bw.de/cms/bsz/satzung/>>.

² Zur Geschichte des Projektes: Jörn Sieglerschmidt: MusIS – Chancen und Probleme großer Dokumentationsprojekte. In: Gerald Maier/ Thomas Fricke (Hrsg.): Kulturgut aus Archiven, Bibliotheken und Museen im Internet. (Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg. Serie A Landesarchivdirektion, Heft 17) Stuttgart: Kohlhammer 2004, 48-63.

³ <<http://www.bsz-bw.de/museen/musis/teilnehmer/>>

⁴ <<http://www2.bsz-bw.de/cms/museen/musis/publ/kurzinfo/musisbericht2006-2007.pdf>>.

Inhaltsverzeichnis

Vorbemerkung	3
Inhaltsverzeichnis	4
1. Einleitung	6
2. Hardware	8
2.1 Museen	8
2.2 Archive	10
2.3 BAM-Portal	10
3. Software	11
3.1 Systemsoftware	11
3.2 Datenbanksoftware	11
3.3 Anwendersoftware	11
3.3.1 IMDAS-Pro	11
3.3.2 Pan-Net Catalogue	12
3.3.4 Andere Software	13
3.3.5 BAM-Portal	13
4. Migrationen	17
4.1 Einmalige Datenimporte nach IMDAS-Pro	17
4.2 Datenimporte in das BAM-Portal	17
4.3 Schnittstellen zu anderen Datenbanken	18
5. Regelwerke	19
5.1 Sachen	20
5.1.1 Objektbezeichnungen	20
5.1.2 Material und Technik	20
5.2 Orte	21
5.3 Personen und Körperschaften	21
5.4 Zeiten	22
5.5 Metadaten	22
6. Kundenbetreuung	23
6.1 Kundenanfragen	23
6.2 Schulungen	23
6.3 Nutzung des BAM-Portals	24
7. Tagungen	27
7.1 Vom M-Team bzw. einzelnen Mitarbeiter/innen organisierte Tagungen	27
7.2 Teilnahme an Tagungen anderer Veranstalter und Messen	27
7.3 Gremien und feste Arbeitsgruppen	28
7.3.1 Deutscher Museumsbund	28
7.3.2 Minerva-/EUBAM-Spiegelgruppe	28
7.3.3 Deutsche Digitale Bibliothek	28
7.3.4 Wissenschaftliche Institutskommission des Instituts für Museumsforschung Berlin	29
8. Projekte	30
8.1 Sacherschließung	30
8.2 Kooperative Erschließung	30
8.3 Datenschutz und IT-Sicherheitskonzept für den Einsatz von Informations- und Kommunikationstechnik in den Kunsteinrichtungen des Landes Baden-Württemberg	30
8.4 Datensicherung, Datennutzung und Langzeitarchivierung für die Staatlichen Museen des Landes Baden-Württemberg	30
8.5 Inventarisierung der Sammlung Kindermedien der Hochschule der Medien in Stuttgart	31

9. Kooperationen	32
9.1 DigiCULT Schleswig-Holstein	32
9.2 Institut für Museumsforschung SPK	32
10. Mitarbeiter/innen und Arbeitsbereiche	33
10.1 Mitarbeiter und Kontaktangaben	33
10.2 Arbeitsbereiche und Ansprechpartner	33
11. Veröffentlichungen zur Museumsdokumentation und zur Museumskunde	35
12. Abkürzungsverzeichnis	36
13. Anhänge	38
Anhang 1: Dienstleistungen des BSZ im Museumsbereich	38
Anhang 2: Convergence of internet services in the cultural heritage sector – the long way to common vocabularies, metadata formats, ontologies	40
Anhang 3: Knowledge organization and multilingual vocabularies	49
Anhang 4: Der Objektbezeichnungsthesaurus des BSZ	57

1. Einleitung

MusIS geht nun in das achte Jahr seines Bestehens als fest institutionalisierter Kooperationsverbund der Staatlichen Museen in Baden-Württemberg. Zuvor als Pilotprojekt des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg (MWK) beim Staatlichen Museum für Naturkunde in Stuttgart angesiedelt, ist die Betreuung der Staatlichen Museen in Baden-Württemberg ab Herbst 2000 zur beständigen Aufgabe des BSZ geworden. Es fungiert damit seit 2000 als Koordinationsstelle für technische und dokumentarische Fragen, die sich in Zusammenhang mit der Museumsdokumentation sowie der Präsentation ihrer Ergebnisse im Internet stellen. Darüber hinaus betreibt das BSZ als sog. application service provider (ASP) für die Mehrheit der staatlichen Museen in Baden-Württemberg und weiterer Museen die Datenbanken, betreut Internetprojekte zur Präsentation von Museumsobjekten und beteiligt sich an nationalen und internationalen Arbeitsgruppen und Projekten zur Weiterentwicklung fachlicher Angebote in diesen Bereichen.

Sowohl die technische Konfiguration als auch die Arbeiten in der Museumsdokumentation sind Teil einer Lösung, deren wichtigste Aufgabe es ist, die Inventarisierung der Museumsobjekte effektiver zu gestalten und damit das wichtigste Eigentum der Museen zugänglich zu machen und zu sichern. Die Museen verfügten und verfügen zumeist weder über die notwendige technische noch die personelle Ausstattung für den Betrieb angemessener Datenbanksysteme und für die Unterstützung der Dokumentation. In dieser Situation lag es angesichts der technischen Entwicklung im Netzbereich nahe, die entsprechenden Dienstleistungen zu zentralisieren.

Zu den vom BSZ betreuten Museen bzw. Sammlungen gehören:

- Badisches Landesmuseum Karlsruhe (BLMK),
- Landesmuseum Württemberg Stuttgart (LMWS)
- Linden-Museum Stuttgart, Staatliches Museum für Völkerkunde (LMS),
- Inventar der Kunsterwerbungen des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg (MWK),
- Reiss-Engelhorn-Museen Mannheim (REMM),
- Staatliche Kunsthalle Karlsruhe (SKK),
- Staatliches Museum für Naturkunde Karlsruhe (SMNK),
- Staatsgalerie Stuttgart (SGS),
- Städtische Museen Freiburg (SMF),
- Stiftung Schloss Friedenstein Gotha (SFG).

Weitere Teilnehmer am Kooperationsverbund sind:

- Archäologisches Landesmuseum Baden-Württemberg (ALM),
- Stiftung Landesmuseum für Technik und Arbeit Mannheim (LTAM),
- Haus der Geschichte Baden-Württemberg Stuttgart (HGS),
- Museum Pfalzgalerie Kaiserslautern,
- Staatliches Museum für Naturkunde Stuttgart (SMNS).

Die einzelnen Dienstleistungen sind im Anhang 1 aufgeführt.

Das Landesarchiv Baden-Württemberg (bis 31.12.2004 Landesarchivdirektion und staatliche Archive) nimmt seit 2000 Dienstleistungen des BSZ in Anspruch, zunächst Webhosting, inzwischen aber darüber hinaus auch *application service* und Datenbankadministration für die Oracle-Datenbanken der staatlichen Archive. Seit 2000 hat das Landesarchiv seine *online*-Angebote erheblich erweitert, u. a. um ein *online*-Bestellsystem, die Auswanderer-Datenbank und die Zollvereinsstatistik. Das württembergische Urkundenbuch wird 2008 *online* gestellt. Das Landesarchiv nimmt mit diesen *online*-Dienstleistungen eine führende, innovative Rolle in Deutschland ein.

2. Hardware

2.1 Museen

Das BSZ stellt für die Museen neben den für die Kommunikation erforderlichen Kommunikations- und Internetdiensten Server bereit, auf denen die Oracle-basierte Dokumentationsanwendung *IMDAS-Pro* der Fa. Joanneum Research, Graz, das Content Management System *Pan-Net Catalogue* der Fa. Pandora Neue Medien Stuttgart, das BAM-Portal und eine *Citrix*-Server-Farm laufen. Außerdem sind die zentral gepflegten Regelwerke wie das *Allgemeine Künstlerlexikon (AKL)* auf diesen Servern abgelegt und für Nutzer/innen zugänglich. Die Software (System, Datenbank, Anwendung) wird zentral gepflegt, d. h. Fehlerbeseitigung, Aufspielen von neuen Versionen bzw. Erweiterungen werden im BSZ vorgenommen, ohne dass der laufende Betrieb dadurch tangiert wird.

Die technische Lösung mit *Citrix* erfordert auf der Seite der Museen nur geringe Bandbreiten und läuft seit einigen Jahren sehr stabil. Die Hardwareanforderungen an die *Citrix*-Clients sind gering. Die *Citrix*-Clients können vor Ort einfach installiert werden, wobei dort auch Rechner mit Betriebssystemen wie Unix oder Mac lauffähig sind. Die Datenübermittlung wird durch eine 128-bit-Verschlüsselung gesichert. Die neueste Version von *Citrix (Presentation Server)* wird über ein Web-Interface einen *secure-gateway*-Zugang zu den Anwendungen ermöglichen (über Port 443). Dieser Zugang wird in Pilotanwendungen ab Januar 2008 getestet. Nach erfolgreich beendeten Tests können Museen diesen Zugang dann bei Bedarf nutzen.

Der Serverbetrieb steht unter den üblichen Anforderungen an Datensicherheit und Datenschutz. Die Daten werden täglich durch einen Bandroboter gesichert. Zugangsbeschränkungen auf allen Ebenen und in allen Arbeitsbereichen sorgen zusätzlich für Datensicherheit. Ein Projekt zur Langzeitarchivierung soll dazu führen, dass auch in diesem Bereich Dienstleistungen für Museen und Archive vom BSZ angeboten werden. Das Datenschutzkonzept wird im Auftrag des *Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg* seit 2007 erarbeitet und noch 2008 vorliegen.

Im Jahr 2005 ist ein neuer *Sun*-Server mit zwei Prozessoren für den Datenbankbetrieb der Museen beschafft und getestet worden. Diese Tests dienten der Überprüfung der Performance unter den neuen Systembedingungen und verliefen erfolgreich. Im Januar/Februar 2006 konnten daher die Museumsdatenbanken auf den neuen Server migriert werden.

Die bereits 2005 begonnene Neukonfiguration des Datenbankservers für die *IMDAS-Pro*-Datenbanken ist 2006 abgeschlossen worden. Während zuvor jede Museumsdatenbank eine eigene Instanz hatte, gibt es jetzt nur noch zwei Instanzen, eine für den Test-, die andere für den Produktionsbetrieb. Dadurch konnten finanzielle und Speicherressourcen geschont werden.

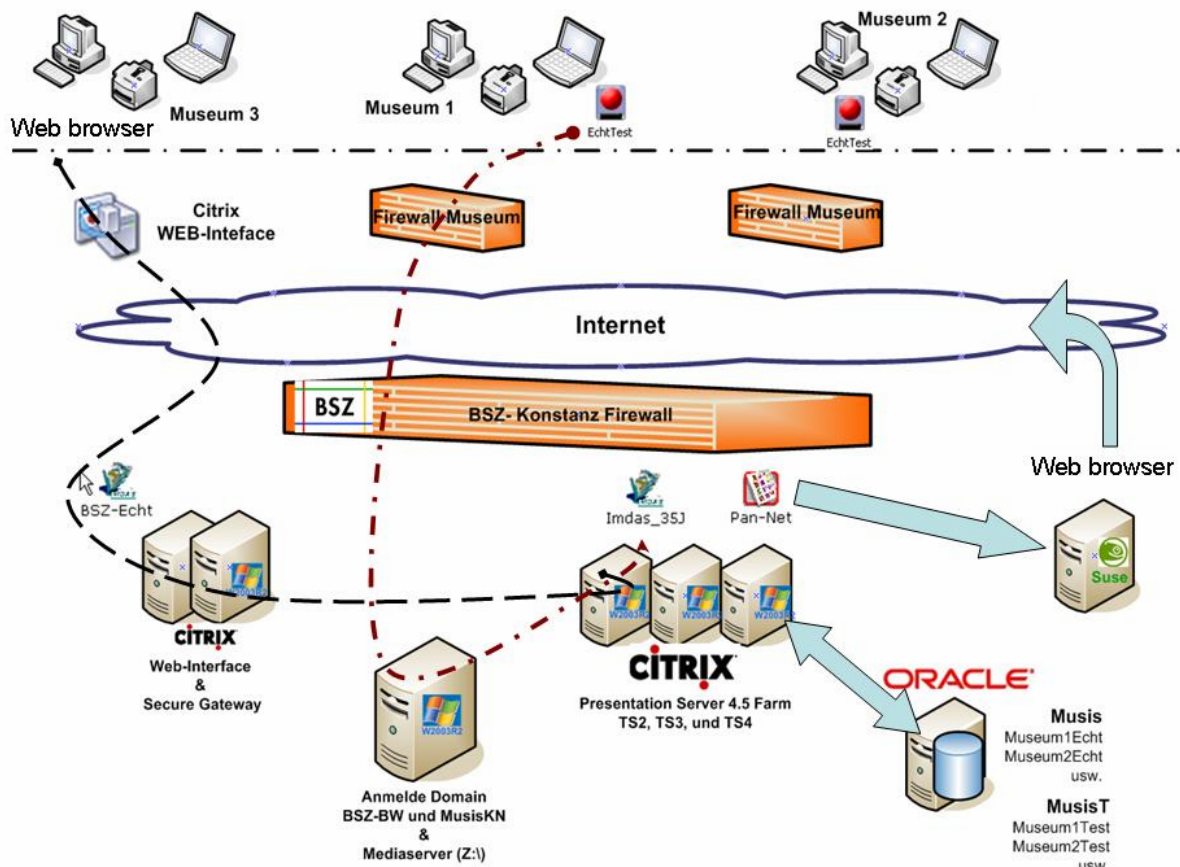


Abbildung 1: Technische Anbindung der Museen

Zugleich mit diesen Änderungen ist ein geregelter Betrieb für die Versionierung eingeführt worden. Grundsätzlich laufen alle Museumsdatenbanken unter einer Version, um den Betreuungsaufwand möglichst gering zu halten.

Auch die Domänenstruktur ist bereinigt, d. h. sowohl die Anmeldestruktur als auch die *active-directory*-Domäne für die Terminalserver sind konsolidiert worden. Die gesamte technische Konfiguration ist skalier-, d. h. je nach Bedarf erweiterbar. Der MusIS-Webserver wird unter *Linux* betrieben und erlaubt die Teststellung sowie die Erprobung neuer Dienste. Die Citrix-Server-Farm ermöglicht eine optimale Lastverteilung sowie Ausfallsicherheit und ist jederzeit bei Bedarf erweiterbar. Der Datenbankserver von *Sun* wird mit *Solaris* betrieben. Das Datenbanksystem ist *Oracle* (Version 9i). Teil der Serverfarm sind schließlich zwei Domaincontroller (Verwaltung der Nutzerberechtigungen) und mehrere Fileserver (Dateiverwaltung).

Im Hintergrund dieser technischen Konfiguration läuft die Bandsicherung für die Serverfarm und das Datenbanksystem. Sie ist Teil eines Sicherheitskonzeptes, das auch die Absicherung der Datenbanken und den Schutz der Datentransfers vor unbefugten Eingriffen gewährleistet.

Die Abb. 1 zeigt die schematische Darstellung der technischen Konfiguration. Die Nutzer kommen über den *Citrix*-Client, die *firewall* des Museums, das Internet und die *firewall* des BSZ zur Anmelde-Domäne des BSZ und schließlich zu den Terminal-Servern, auf denen die Anwendungen laufen (*IMDAS-Pro*, *Pan-Net Catalogue*). Von

dort wird auf den *Oracle*-Datenbankserver zugegriffen. *Pan-Net Catalogue* spielt die Daten über den Webserver ins Internet aus.

2.2 Archive

Im Archivbereich werden Server für die zahlreichen Webdienste des Landesarchivs betrieben. Es handelt sich um unterschiedliche Hardware-Komponenten auf Intel- und AMD-Basis.

Für das Landesarchiv ist in 2007 ein eigenes Netz mit einer eigenen *firewall* aufgebaut worden (vgl. Abb. 2). Auch der eigentliche Webserver, auf dem mit dem *SixCMS* die Informationsangebote des Landesarchivs laufen, wird Anfang 2008 im Netz hinter der eigenen *firewall* sein. Nur noch bei der Sicherung der Daten wird auf den *backup*-Server im Netz des BSZ zugegriffen. Die Datenbankserver der staatlichen Archive stehen jeweils als Einzelrechner mit einem eigenen Netz für den laufenden Betrieb vor Ort zur Verfügung. Lediglich für die Wartung der Datenbanken und das *backup* wird vom BSZ auf die einzelnen Datenbankserver zugegriffen.

2.3 BAM-Portal

Um einen reibungslosen und performanten Betrieb des BAM-Portals zu gewährleisten, wurden zwei neue Server im Berichtszeitraum beschafft. Auf einem Server wird die Website des Portals zusammen mit der Suchmaschine sowie einem FTP-Server, der von den Teilnehmern zur Datenlieferung genutzt wird, gehostet. Der zweite Server dient der kontinuierlichen Indexierung neuer Daten für die Produktivversion des Portals.

3. Software

3.1 Systemsoftware

Der MusIS-Webserver läuft unter *Linux*. Der Datenbankserver der Fa. *Sun* nutzt das *Sun*-eigene Betriebssystem *Solaris*. Alle anderen Server laufen unter Windows-2000/2003-Server. Die Server des Landesarchivs laufen unter den *Linux*-Distributionen *Open Suse* und *Debian Etch 4.0*, die BAM-Server ebenfalls unter *Debian Etch 4.0*.

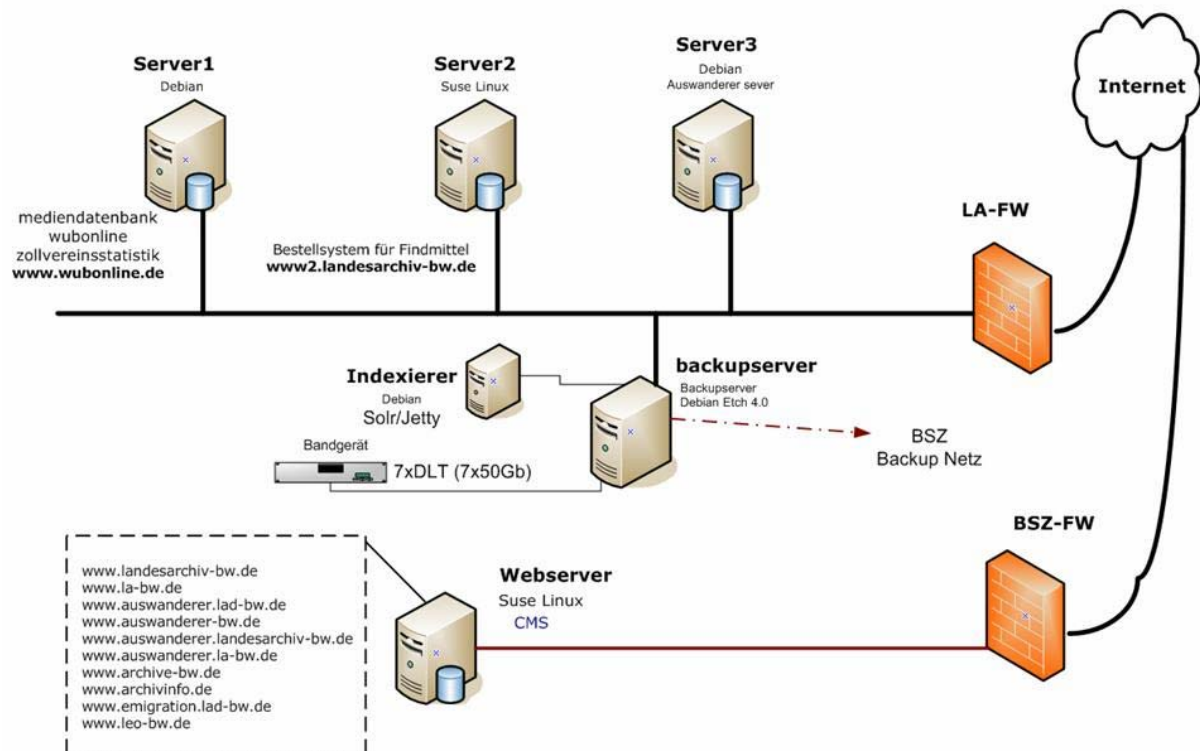


Abbildung 2: Technische Anbindung der Landesarchive

3.2 Datenbanksoftware

Als Datenbanksoftware wird für die MusIS-Datenbanken seit Beginn des Betriebes im BSZ und wie im Soll- und Lösungskonzept von 1995/1996 vorgesehen *Oracle* in der Version 9i genutzt. Die bisherige Konfiguration der einzelnen Datenbanken als eigene Instanz ist mit dem Umzug auf einen neuen Server dahingehend geändert worden, dass es nur noch zwei Instanzen für Echt- und Testdatenbanken gibt. Auch das Landesarchiv betreibt seine Datenbanken auf *Oracle*, aber für die sechs Server in jeweils einer eigenen Instanz.

3.3 Anwendersoftware

3.3.1 IMDAS-Pro

Die Anwendersoftware *IMDAS-Pro* von Joanneum Research (JR) ist seit 1998 in baden-württembergischen Museen im Einsatz und hat sich als vielseitiges Instrument für die Dokumentation von Museumsobjekten und die Sammlungsverwaltung bewährt.

Im Moment ist die Version 3.5 installiert. 2008 wird es einen Versionswechsel nach 4.0 geben. Diese Version bringt unter anderem eine entscheidende Neuerung in der Benutzerverwaltung mit sich, die damit ausschließlich von *IMDAS-Pro* aus und nicht wie bisher über die Oracle-Datenbank gesteuert wird. Seit 2007 bietet JR Erweiterungen und Änderungen der Software als Online-Update an, was das Aufspielen neuer Versionen und Patches erheblich erleichtert, weil dadurch die aufwändige Installation von neuen Versionen entfällt.

Seit 2006 wird das Layoutmodul genutzt, um formularmäßige Ausdrücke zu erzeugen. *IMDAS-Pro* hat als Grundlage hierfür eine Fremdsoftware eingesetzt. Die Komplexität dieser Funktion ist hoch und die Benutzungsfreundlichkeit gering, weswegen die Drucklayouts nur auf Vorlage und in Absprache vom BSZ erstellt werden.

IMDAS-Pro ist in den letzten Jahren in seiner Funktionalität erheblich erweitert worden. War bereits die Umstellung auf benutzerdefinierte Sichten ein erheblicher Schritt in Richtung größerer Handlungsspielräume für benutzerdefinierte Masken und Bereich, so bringt die seit 2007 mögliche Einbeziehung der Felder aus den Geschäftsprozessen zusätzliche Möglichkeiten, mit den v o r h a n d e n e n Strukturen individuelle Sichten für die einzelnen Museen zu gestalten. Allerdings sieht es das BSZ als seine Aufgabe, diese Möglichkeiten so zu nutzen, dass das Datenbanksystem nicht inkonsistent wird. Angesichts der zunehmenden Anforderungen an benutzerdefinierte Sichten ist es nicht auszuschließen, dass gleiche oder identische Informationen auf diese Weise an unterschiedlichen Stellen abgelegt werden. Aus Konsistenzgründen ist es daher notwendig, dass die benutzerdefinierten Masken im BSZ eingerichtet werden.

3.3.2 Pan-Net Catalogue

Pan-Net Catalogue hat sich neben dem BAM-Portal als Präsentationswerkzeug für digitalisierte Museumsobjekte bewährt. Nach der Pilotanwendung im BLMK⁵ sind weitere Präsentationen dazugekommen:

- Die SKK hat anlässlich der gleichzeitigen Ausstellung zu Radierungen Stefano Della Bellas eine Auswahl online präsentiert, die auch jetzt noch zu sehen ist. Außerdem sind die Präsentationen zum besonderen Blatt gesammelt worden.⁶
- Das LMWS zeigt die antiken Gläser aus der Glassammlung Ernesto Wolf.⁷

Es ist inzwischen möglich, die im BAM-Portal gefundenen Treffer aus diesen Sammlungen direkt zum jeweiligen Objekt zu verlinken, so dass auch aus dem BAM-Portal Zugriffe auf diese Seiten erfolgen.

Mit den Anforderungen der verschiedenen Internetausspielungen wurde und wird der Funktionsumfang der Software beständig ausgebaut (z. B. Implementierung einer Bildlaufleiste, Konnektierung neuer Felder).

Die nächsten digitalen Präsentationen sind geplant.

⁵ <http://www.bsz-bw.de/museum/blm/dk/der_digitale_katalog.htm>.

⁶ <http://musis.bsz-bw.de/pan/kunsthalle-karlsruhe/kupferstichkabinett/DasbesondereBlatt/das_besondere_blat.html>;
<http://musis.bsz-bw.de/pan/kunsthalle-karlsruhe/kupferstichkabinett/DellaBella/virtuelle_ausstellung.htm>.

⁷ <<http://hybris.bsz-bw.de/pan/WLM/GlassammlungWolf/tp50101.htm>>.

3.3.4 Andere Software

Im Landesarchiv wird als Anwendersoftware für die Erfassung der Archivalien *scopeArchive* eingesetzt.⁸ Das BSZ betreut das Landesarchiv nicht in fachlicher Hinsicht, so dass die Betreuung dieser Software nicht zu seinen Aufgaben gehört.

3.3.5 BAM-Portal

Die Software des BAM-Portals ist eine von der DFG geförderte Eigenentwicklung unter Verwendung der im Folgenden beschriebenen *open-source*-Komponenten. Das System gliedert sich, wie in Abbildung 3 dargestellt in folgende Teile:

Der für den Nutzer sichtbare Teil ist das Suchportal.⁹ Es umfasst die Suchmaschine sowie den Webserver Apache und einen, oder mehrere *Tomcat-Server* (beides OpenSource). Die *Tomcat-Server* dienen der Generierung dynamischer Webinhalte, während der *Apache-Server* statische Seiten ausliefern kann, sowie die HTTP-Anfragen, die über Port 80 an das Portal gerichtet werden, an einen der *Tomcat-Slaves* weiterleitet. Hierbei kann, bei mehr als einem *Tomcat*, eine gleichmäßige Lastverteilung für die Slaveserver erreicht werden. Die Suchmaschine ist eine Javawebapplikation und wird durch eine *Tomcat*-Instanz angesprochen. Sie nimmt Suchanfragen entgegen und beantwortet diese anhand ihres Volltextindexes. Ihre wesentlichen Teile sind dabei *Lucene* sowie darauf aufbauend *Solr*, beide stammen von der *Apache-Softwarefoundation*. *Lucene* ist eine Javabibliothek zur Konstruktion von Volltextsuchmaschinen und *Solr* stellt eine Infrastruktur zur Konfiguration eines Luceneindexes sowie Werkzeuge zur Replikation von Indizes zwischen einem Masterserver, sowie einer Anzahl von Slaveservern dar. Momentan kommt das Portal mit einem Slaveserver aus, dieser ist identisch mit dem Suchportalserver.

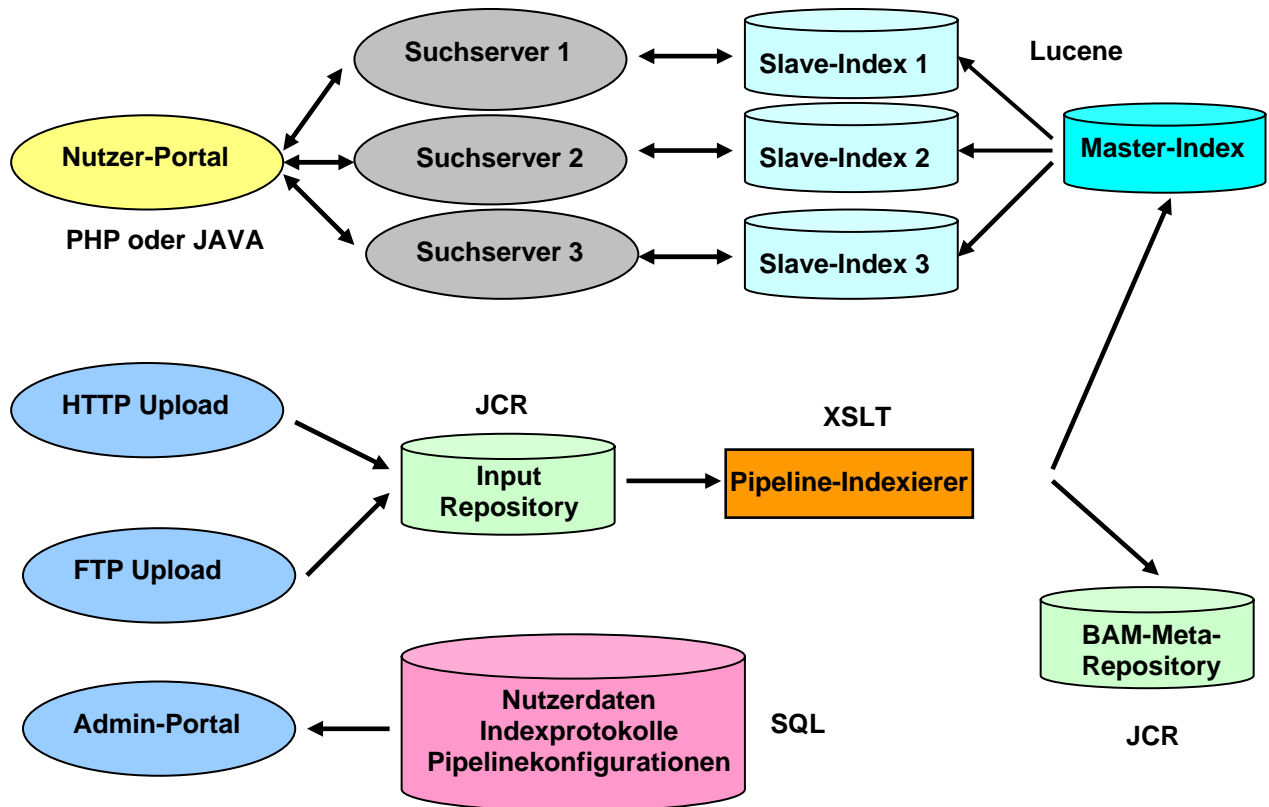
Der Index des Suchportals wird auf einer zweiten Maschine erstellt, auf welcher das Upload-Portal läuft. Auch dieses umfasst wieder eine Apache-, sowie eine *Tomcat*-Instanz. Im Upload-Portal, welches auch einen in Java implementierten FTP-Server enthält, werden die zu indexierenden Daten entweder per FTP oder direkt via HTTP hochgeladen, in ein kanonisches Standardformat transformiert und dann an den Solrserver zur Indexierung weitergereicht. Der Index wird mehrmals täglich mit den Indizes der Suchserver mittels rsync-Protocoll synchronisiert.

Dieses Schema ist gut auf größere Serverfarmen skalierbar, so dass die Architektur von BAM nun auch eine spätere Erweiterung um zusätzliche Suchserver erlaubt. Eine weitere Funktion des zweiten Servers ist der Betrieb einer Testversion des Portals, die sowohl dem Test der Software dient als auch der probeweisen Indexierung neuer Daten. Hierdurch soll die Einhaltung eines hohen Qualitätsstandards gesichert werden.

⁸ <<http://www.scope.ch/>>.

⁹ <<http://www.bam-portal.de>>.

Abb. 3 Architektur BAM-System



Für die angestrebte Visualisierung der multidimensionalen Treffermenge mittels eines Navigationsbaumes war eine Umstrukturierung der Indexstruktur notwendig. Die vorherige Aufteilung in vier spartenspezifische Indices wurde ersetzt durch einen gemeinsamen Index für alle Sparten. Die Abbildung der BAM-Indexfelder auf spartenspezifische Indexfelder wurde von der Suchzeit auf die Indexierungszeit verlagert. Die spartenspezifische Information, welche aus der heterogenen Masse der Metadaten extrahiert wird, wird in einem kompakten, alle Sparten berücksichtigenden XML-Format in komprimierter Form im Index selbst gespeichert (aber nicht indexiert). Somit wird ein effizienter Zugriff auf diese Information während der Darstellung der Suchergebnisse ermöglicht. Die Darstellung geschieht durch sparten- und gegebenenfalls institutionenspezifische XSL-Stylesheets, die die XML-Metadaten in XHTML wandeln.

Zur effizienten Darstellung der Treffermenge in Baumform wurde auf die Möglichkeiten des Solr-Servers zurückgegriffen. Solr erlaubt von Hause aus die Möglichkeit der facetierten Suche und bietet sich somit an zur Realisierung einer hierarchischen Unterteilung der Treffermenge. Hierbei soll jeder Teilbaum auch die Information über die Summe der Treffer seiner Unterbäume tragen. Um eine n-malige Wiederholung der Suche zur Feststellung der jeweiligen Trefferzahlen eines Baumes mit insgesamt n Zweigen und Endpunkten zu vermeiden, welche die Performanz der Suchmaschine stark beeinträchtigt hätte, ist es notwendig, die unfacetierten Treffermengen als Bitvektor darzustellen und mittels Bitmanipulationen in einem Durchgang die gesamte Facetteninformation zu extrahieren. Ein solcher Algorithmus ist in Solr implementiert. Die Laufzeit der Gesamtsuche wird hierdurch nicht signifikant verlängert.

Die zugrundeliegende Theorie der Facettenklassifikation¹⁰ ist noch um einiges allgemeiner. Sie erlaubt multiple Klassifikationen eines Objektes, die in nicht von vornherein festgelegter Weise geordnet sein können. Dies steht im Gegensatz zu starren Taxonomien, die z. B. im Bibliothekswesen gebräuchlich sind. Es ist deshalb geplant, den jetzigen Treffernavigationsbaum, der nach Sparten, Institutionen und Katalogen differenziert, um weitere Aspekte zu verallgemeinern. Als Facetten denkbar wären unter anderem Schlagworte, Material, Technik, Form, Orte oder Personen. Hierdurch wird die multidimensionale Struktur der Treffermenge sichtbar. Der Nutzer kann entlang einzelner Raumrichtungen die Treffermenge durchwandern.

Die Darstellung des Suchergebnisses in XHTML-Form mittels XSL-Technologien sowie die Integration neuer Daten und neuer Datenformate machte es notwendig, das im vorherigen Berichtszeitraum entwickelte XML-Pipelining-Framework zu erweitern. Dieses Framework basiert auf der einfachen Idee des Fließbandes. Am Anfang des Fließbandes werden Daten in beliebiger Form in einen Generator eingefüttert. Dieser Generator hat die Eigenschaft, XML (genauer gesagt: SAX-Events) auszugeben. Diese werden dann in einem oder mehreren Transformatoren weiterverarbeitet. Jeder Transformator nimmt XML entgegen und gibt XML aus (genauer gesagt: SAX-Events). Am Schluss der Pipeline steht ein Serialisierer, welcher XML entgegennimmt und dieses nun in ein Ausgabeformat umwandelt. Dieses kann ein Lucene-Index sein, oder XHTML zur Darstellung in einem Browser, oder auch eine Datenbank.

Auf dem Markt befinden sich viele solcher XML-Pipelining-Frameworks. Das am weitesten entwickelte ist wohl Cocoon (<http://cocoon.apache.org>). Aufgrund des angestrebten Zieles, in kurzer Zeit Ergebnisse erzielen zu wollen, wurde auf seinen Einsatz aber zunächst verzichtet. Die entstandenen Komponenten im BAM-Pipelining-Framework sind jedoch mit minimalem Aufwand nach Cocoon portierbar.

Um zukünftige Anforderungen zu erfüllen (z. B. Generierung von internationalisierbarem XHTML, flexible Darstellung der Detailtreffer und der Präsentation der Teilnehmer in BAM, welche diese durch Bearbeitung von XML-Templates selbst beeinflussen können), soll Cocoon eingesetzt werden. Auch bei der geplanten Zusammenarbeit mit dem EU-Projekt MICHAEL^{plus}, welches ebenfalls Cocoon einsetzt, werden sich Synergien ergeben.

Um den zukünftigen Arbeitsaufwand zur Integration neuer Teilnehmer und Daten zu minimieren, wurde mit der Konzeption einer grafischen Webschnittstelle begonnen. Diese wird dem Betreiber des BAM-Portals die Administration der Suchmaschine und der Indexierungspipelines erleichtern. Neue Teilnehmer sollen in die Lage versetzt werden, ihre Stammdaten selbst anzulegen und später auch zu pflegen. Die Offenlegung der HTML-Präsentation, sowohl der Stammdaten, als auch der Kurz- und Detailtrefferanzeige in Form von XSL-Dateien wird das ermöglichen. Voraussetzung ist natürlich das Vorhandensein entsprechender Ressourcen beim Teilnehmer. Möglich ist aber

¹⁰<http://en.wikipedia.org/wiki/Faceted_classification>;
dt.: <http://de.wikipedia.org/wiki/Facettierte_Klassifikation>

selbstverständlich auch eine Übernahme der Konfigurationsarbeiten durch Mitarbeiter des künftigen BAM-Konsortiums, welche dann auch entgeltpflichtig sein könnte.

4. Migrationen

4.1 Einmalige Datenimporte nach IMDAS-Pro

Für die Migration von Altdaten nach IMDAS-Pro wird die seit 2006 vorliegende, laufend optimierte und erweiterte BSZ-Importroutine benutzt. Mit dieser Importroutine wurden bisher rund 350.000 Datensätze importiert. Die größten Datenbestände kamen aus dem LMS (rund 135.000), SMNS (rund 80.000), SMF (rund 77.000), SMNK (rund 32.300), LMWS (rund 16.100). Eine Übersicht der vom BSZ für die Museen durchgeführten Migrationen bietet Tabelle 1.

Museum	Migrierte Sammlung / Datenbank	Zeitpunkt der Migration	Migrierte Datensätze
SMNK	Entomologie: Mantodea	2006	2.456
SMNK	Entomologie: Ceratopogoniden	2007	456
SMNK	Geologie: Paläontologie	2007	6.406
SMNK	Zoologie: Wirbeltiere	2007	23.060
BLMK	Depotverwaltung: Metallobjekte	2007	1.596
BLMK	Graphikdepot	2007	470
BLMK	Bildarchiv Staufen	2007	5.633
SMF	Datenbank Augustinermuseum	2007	77.000
LMS	Nordamerika/Ethnologie	2006	3.375
LMS	Konvolut/Ethnologie	2006	236
LMS	Ost Asien	2007	14.927
LMS	Orient	2007	12.013
LMS	Ozeanien	2007	28.843
LMS	Südamerika	2007	18.000
LMS	Afrika	2007	58.000
LMWS	Volkskunde/NSG-Objekte	2006	223
LMWS	Archäologie/Frühmittel-alter	2007	15.900
SGS	Italien/Europa/Barock	2007	687
SGS	Verschiedenes	2007	4.566
SKK	Plastik/Kunst	2007	3.379
SFG	Kulturgeschichte/SGA	2007/2008	413
Gesamt			277.416

Tab. 1: Datenmigrationen 2006/2007

Durch die Importroutine wurden nicht nur erhebliche Datenbestände zur Nutzung in IMDAS-Pro zur Verfügung gestellt. Vielmehr tragen die Importe durch sehr intensive inhaltliche Überarbeitung der Altdaten im Rahmen der Migration auch zu einer qualitativen Aufwertung des Gesamtdatenbestandes bei.

4.2 Datenimporte in das BAM-Portal

Das BAM-Portal ist im Berichtszeitraum erheblich gewachsen. Im BAM-Portal werden momentan gut 38.5 Millionen Datenobjekte nachgewiesen, davon 35.74 Millionen aus dem Südwestdeutschem Bibliotheksverbund (SWB) und dem Gemeinsamen Bibliotheksverbund (GBV),¹¹ 1.88 Millionen Archivobjekte vornehmlich aus dem

¹¹ Gemeinsamer Verbundkatalog (GVK) des GBV <<http://www.gbv.de/vgm/>>.

Landesarchiv Baden-Württemberg, dem Bundesarchiv und den hessischen Staatsarchiven, 0.81 Millionen aus anderen Institutionen, vornehmlich aus der Zentralkartei der Autographen und Nachlässe (ZKA) mit ihrer Online-Datenbank *Kalliope* sowie aus dem Zentralverzeichnis digitalisierter Drucke (ZVDD),¹² und schließlich die Museen mit 0.11 Millionen Objektnachweisen, wobei die baden-württembergischen Museen besser vertreten sein könnten. Die Museumsobjekte können nach Herkunft aufgeschlüsselt im BAM-Portal recherchiert werden.

Die Integration neuer Daten in das BAM-Portal stellt noch immer ein Problem dar, da insbesondere aus den Museen häufig Daten mit proprietären Formaten geliefert werden. Die Situation hat sich durch das *harvesting format museumdat* bereits geändert und wird in Zukunft auch automatische Importe möglich machen bzw. das automatische Einsammeln von Objektdaten von anderen Seiten. So sind bereits jetzt die Daten des ZVDD über eine OAI-Schnittstelle eingesammelt worden.

4.3 Schnittstellen zu anderen Datenbanken

Die Schnittstelle zwischen dem vom BSZ betriebenen SWB und *IMDAS-Pro* (Z39.50) ist nach der Migration des Verbundes auf die *PICA*-Anwendersoftware nochmals geprüft worden. Angesichts der laufenden Umstellungs- und Anpassungsarbeiten kann diese Schnittstelle erst zu einem späteren Zeitpunkt realisiert werden.

Die Schnittstellenprogrammierung zum AKL des Saur Verlages läuft seit 2005 ohne wesentliche Probleme. Änderungen in der Datenlieferung machen Anpassungen in der Schnittstellenprogrammierung notwendig.

Es haben Vorgespräche mit *DigiCult Schleswig-Holstein* (DSH) und dem Zuse-Zentrum für Informationstechnik Berlin (ZIB) stattgefunden hinsichtlich der Programmierung einer Schnittstelle zum gemeinsamen Server für die Verwaltung von Thesauri. Eine *online*-Plattform ist bereits in Betrieb. Softwarehersteller, u. a. auch und erfolgreich JR, haben eine Schnittstelle für das Suchen und Herunterladen von Begriffen in die lokale Datenbank programmiert. Was fehlt, ist eine sog. aktive Komponente, mit der neue Begriffe aus der lokalen Datenbank wiederum in das zentrale Verzeichnis eingestellt und dort redaktionell weiterverarbeitet werden können. DSH hat eine in Betrieb befindliche Programmkomponente, die angepasst und für den gemeinsamen Betrieb verwandt werden könnte. Ein wesentliches Problem muss im Vorfeld gelöst werden, nämlich die eindeutige Kennzeichnung eines jeden Begriffes z. B. über ein *global(ly) unique identifier* (GUID). Nur so können Begriffe in unterschiedlichen technischen Umgebungen bearbeitet und neue Versionen bzw. neue Begriffe eines Regelwerkes an die lokalen Systeme zurückgespielt werden.

¹² <<http://www.kalliope-portal.de>>; <<http://www.zvdd.de>>.

5. Regelwerke

Regelwerke für die Sacherschließung bilden den Grundpfeiler einer qualitativollen Dokumentation von Museumsobjekten. Für diejenigen, die die Situation in den deutschen Museen nicht kennen, sei zum einen ausdrücklich betont, dass im Gegensatz zu den Bibliotheken, die seit den sechziger Jahren des letzten Jahrhunderts viel Arbeit in die gemeinsame Entwicklung von Regelwerken für die Formal- wie die Sacherschließung gesteckt haben, die deutschen Museen bisher weder gemeinsame Regelwerke noch Institutionen haben, die diese Arbeit kontinuierlich koordinieren würden. Zum anderen ist - ebenfalls im Vergleich zu den Bibliotheken - keine strikte Trennung zwischen Formal- und Sacherschließung möglich. Auch die Orientierung daran, was am Objekt direkt ablesbar ist und was erschlossen werden muss, ergibt eine solche nicht. Daher ist bei der Beschreibung von Objekten die Nutzung von normierten Werkzeugen der Sacherschließung wichtig.

Diese Situation hat sich in den letzten beiden Jahren durch die gemeinsamen Anstrengungen der *Fachgruppe Dokumentation im Deutschen Museumsbund* entscheidend zum Positiven gewandelt. Die Einrichtung der *online*-Plattform für Museumsvokabular ist ein erster Anfang, um zu gemeinsam gepflegten Vokabularen zu kommen.¹³ Wesentliche Probleme wie die Organisation der dann notwendig werdenden redaktionellen Arbeit, sind freilich noch zu lösen.

Das BSZ verfolgt hinsichtlich der Museumsdokumentation eine klare Strategie: Eingaben in die Datenbanken möglichst unter Einsatz von Regelwerken, deren Begriffe über Identnummern eindeutig referenzierbar und – soweit möglich – in einen weiteren semantischen Raum eingeordnet sind. Das gilt für Orte, Personen/Körperschaften, Zeiten und Sachen. Fortschritte bei diesen Bemühungen sind nicht einfach zu erreichen, da zum einen die Ressourcen fehlen, um zügig verwendbare Regelwerke für die betreuten Museen zu entwickeln, zum anderen in den Museen, teilweise auch bei einzelnen Arbeitsbereichen derselben, vollkommen unterschiedliche Bedürfnisse vorhanden sind. Museumssammlungen weisen eben nicht dieselbe Universalität und Homogenität auf wie Bibliothekssammlungen.

Das BSZ macht es sich angesichts der mangelnden Kapazitäten in den Museen über den engeren Aufgabenbereich hinaus deswegen auch zur Aufgabe, Thesauri zu entwickeln und laufend zu betreuen. Gleichwohl ist das BSZ überzeugt, dass der einzige vielversprechende Weg zu einem gemeinsamen Vokabular derjenige der Kooperation mit den Bibliotheken ist, konkret mit der Deutschen Nationalbibliothek (DNB) in Frankfurt, wo die Zentralredaktion der Schlagwortnormdatei (SWD) angesiedelt ist. Bei allen möglichen Einwänden bleibt klar, dass für die meisten Museumssammlungen die SWD ein ausreichendes Vokabular bietet. Die bestehende Kooperation mit der DNB verspricht in den nächsten Jahren eine Erweiterung des Vokabulars, die den Museen zugute kommt.

¹³ <<http://www.museumsvokabular.de/>>

Die Thesaurusarbeit für die MusIS-Teilnehmer ist in einem wichtigen Punkt verbessert worden: Die meisten Anwendersysteme konnten zwar Thesauri verwalten, aber nur monohierarchische. Seit 2006 kann *IMDAS-Pro* auch polyhierarchische Thesauri verwalten, wodurch Inkonsistenzen in der Begriffseinordnung vermieden werden.

Zum einen dienten zahlreiche Tagungsteilnahmen der bundesweiten Koordination der Regelwerksarbeiten (vgl. 7.) Zum anderen sind diese Fragen Gegenstand der seit 2004 wieder eingerichteten MusIS-Arbeitsgruppe Regelwerke, die allerdings im Berichtszeitraum nur einmal, am 26. Juni 2006, getagt hat.

5.1 Sachen

5.1.1 Objektbezeichnungen

Die seit 2005 mit Sondermitteln des BSZ vorangetriebene Vereinheitlichung der bisher von den MusIS-Museen angelegten Objektbezeichnungsthesauri konnte Ende 2006 abgeschlossen werden. In diesem Rahmen wurde der Thesaurus auch in eine polyhierarchische Struktur gebracht. Seitdem erfolgt eine kontinuierliche Pflege und Ausbau dieses allgemeinen Objektbezeichnungsthesaurus am BSZ. Der OBZ-Thesaurus deckt den allgemeinen Sammlungsbereich ab und soll durch vertiefende Spezialthesauri ergänzt werden.

Für den OBZ-Thesaurus gilt:¹⁴ Grundsätzlich sollen die Objektbezeichnungen aus der SWD kommen oder – so sie dort nicht enthalten sind – als Begriffsansetzungen vorgeschlagen werden. Inwieweit den rigiden Regeln für den Schlagwortkatalog (RSWK) dabei immer entsprochen werden muss, wird in Absprache mit der DNB erörtert, die ja auch für die öffentlichen Bibliotheken Ansetzungen von Schlagwortketten erlaubt, die nicht mit den RSWK konform gehen.

Seit einigen Jahren nimmt das BSZ im bibliothekarischen Bereich die Aufgaben als Zentralredaktion der SWD für den SWB wahr. Seit März 2007 erfasst auch das M-Team im Rahmen der Zentralredaktion kontinuierlich neue Begriffe für die MusIS-Museen in der SWD-Datenbank. Damit ist das BSZ bisher die einzige Institution im deutschen Museumsbereich, die am Redaktionsverfahren der DNB laufend teilnimmt und den Begriffsbestand der SWD aktiv erweitert. Es steht zu hoffen, dass durch diese Arbeit die SWD auch für die Museen zu einem akzeptierten Werkzeug bei der Sacherschließung wird.

5.1.2 Material und Technik

Der Materialthesaurus ist im Grundbestand abgeschlossen und wird bereits eingesetzt. Allerdings fehlt die Meldung von neuen Begriffen an die SWD-Redaktion. Soweit spezielle Erfordernisse für die Beschreibung bestimmter Materialgruppen vorhanden sind, werden solche Begriffe ergänzt.

Der Technikthesaurus wird, das ist bereits jetzt absehbar, kaum mit der SWD kompatibel sein, also unabhängig von dieser entwickelt werden. An ihm wird ein Problem deutlich, das auch auf den Materialthesaurus zutrifft: Die Museen haben sammlungsabhän-

¹⁴ Nähere Informationen zu Aufbau, Struktur und künftiger Entwicklung des OBZ-Thesaurus siehe Anlage 4

gig sehr unterschiedliche Anforderungen an einen Thesaurus und seine Detailgenauigkeit. Daher sollte den Museen zunächst ein Grundthesaurus angeboten werden, der nur die oberen Hierarchieebenen enthält. Soweit einzelne Teile des Thesaurus sehr fein ausgearbeitet sind, können diese nach Bedarf ergänzt werden, so dass für das einzelne Museum keine überflüssigen Teile in der Datenbank vorgehalten werden. Grundsätzlich ist zu beachten, dass bei der Thesaurusarbeit auch pragmatische Gründe maßgebend sein dürfen, um den Aufbau des Thesaurus nicht zu komplex oder unhandlich werden zu lassen.

Es ist bereits klar, dass im Technikthesaurus die Materialgruppen die oberste Hierarchieebene bilden. Grundsätzliche Fragen, wie z. B. die Unterscheidung von Urformen, Verarbeiten und Bearbeiten eines Werkstoffes, müssen mit bereits vorhandenem Vokabular verglichen und möglichst entsprechend normierter Übung verwandt werden.

5.2 Orte

Obwohl der *Thesaurus of Geographic Names* (TGN)¹⁵ des *J. Paul Getty Trust* vom M-Team 2001 als weniger geeignet für den Einsatz im Verbund eingestuft worden ist, hat sich die Situation insofern verändert, als die *Staatlichen Museen zu Berlin* den TGN einsetzen, dieser auch seither stark erweitert worden ist und in nächster Zeit nochmals auf etwa 5,5 Mio. Ortsnamen weltweit wachsen soll. Der TGN vereinigt alle Anforderungen an ein geographisches Vokabular: politische bzw. Verwaltungseinteilung, geographische Namen wie Berge, Seen usw., geographische Koordinaten, mehrsprachige Ansetzung der Deskriptoren, (historische) Informationen zu den Orten. 2007 ist daher eine Lizenzvereinbarung mit dem *J. Paul Getty Trust* vorbereitet worden, die im März 2008 unterschrieben worden ist. Sie ermöglicht die Nutzung des TGN in den Datenbanken der vom BSZ betreuten Museen und eröffnet die Beteiligung am Redaktionsverfahren des TGN. Geplant ist auf der Basis des *Geographic Names Server*,¹⁶ deutsche Ortsnamen zu ergänzen und möglichst gemeinsam mit anderen dadurch den TGN für deutsche Verhältnisse uneingeschränkt nutzbar zu machen.

5.3 Personen und Körperschaften

Für die Ansetzung von Personennamen kommen drei Quellen infrage: das AKL (s. dazu unter 4.2), die SWD und die Personennamendatei (PND). Für Körperschaften stehen die Gemeinsame Körperschaftsdatei (GKD) oder die SWD bereit, soweit – wie bei den Personennamen und wie vom BSZ empfohlen – ein kontrolliertes Vokabular genutzt wird. Für die üblichen Personennamen- und Körperschaftsansetzungen ist die SWD – bei Künstlernamen das AKL – bisher ausreichend, da sie auch einen für die Museen nutzbaren Namensbestand zu natürlichen und juristischen Personen der Geschichte umfasst. Die Nutzung der GKD wird momentan nicht in Erwägung gezogen.

¹⁵ <http://www.getty.edu/research/conducting_research/vocabularies/>.

¹⁶ <<http://gnswww.nga.mil/geonames/GNS/index.jsp>>.

5.4 Zeiten

Im Herbst 2005 ist für die Vor- und Frühgeschichte sowie die Antike ein Datierungsthe-
saurus erarbeitet worden, der jetzt auf der *online*-Plattform *museumsvokabular.de*
angeboten wird. Es steht zu hoffen, dass dieser Thesaurus in gemeinsamer Arbeit der
interessierten Museen weiterentwickelt wird. Ergänzend wird ein Datierungsthesaurus
mit der Angabe von Zeiträumen nach der Konvention des Deutschen Historischen
Museums Berlin und ein Stilepochen-Thesaurus angeboten.

5.5 Metadaten

In den letzten Jahren ist vom Institut für Museumsforschung SPK (IfM) in
Zusammenarbeit mit dem ZIB und dem *J. Paul Getty Trust* ein XML-Schema entwickelt
worden, das unter dem Namen *museumdat* firmiert und eine Kernstruktur für die
Beschreibung von Museumsobjekten aus allen Sammlungstypen bereitstellt.¹⁷ Für die
Integration von Daten in das BAM-Portal wird *museumdat* bereits eingesetzt. Die
verschiedenen Metadatenformate der einzelnen Institutionentypen werden derzeit
allerdings auf ein internes, d. h. proprietäres Format im BAM-Portal abgebildet.

¹⁷ <<http://museum.zib.de/museumdat/museumdat-v1.0.pdf>>: museumdat - XML Schema zur
Bereitstellung von Kerndaten in museumsübergreifenden Beständen.

6. Kundenbetreuung

6.1 Kundenanfragen

Die tägliche Arbeit im M-Team ist durch zahlreiche Anfragen zur Softwarebedienung, Benutzerverwaltung, Datenein- und -ausgabe, Bildverwaltung, dem Einsatz von Thesauri und Stammdaten sowie neuen Programmversionen geprägt: Des weiteren gehören zu den laufenden Aufgaben:

- Beratung bei der Organisation von Arbeitsabläufen mit *IMDAS-Pro*,
- die Erarbeitung von Schreibanweisungen und Dokumentationshilfen,
- Erstellung von Druckvorlagen und -formularen,
- Erstellung von benutzerdefinierten Erfassungsmasken,
- die Vorbereitung und Durchführung von Schulungen,
- die Erarbeitung des Schulungsmaterials,
- Aufnahme und Bearbeitung von Verbesserungsvorschläge und Fehlermeldungen in Kooperation mit Joanneum Research.

6.2 Schulungen

IMDAS-Pro-Schulungen					
Nr.	Schulungsort	Teilnehmer	Kurse	Tage	Datum, geschulte Institution
1	Konstanz, BSZ	1	2	2	06./07.02.2006 WLMS
2	Konstanz, BSZ	6	1	2	25./26.09.2006 Administratoren SMF, LMS, SFG, BLMK, MWK, LMWS
3	Karlsruhe, SMNK	6	1	2	16./17.10.2006 SMNK
4	Konstanz, BSZ	7	1	2	08./09.11.2006 Administratoren LMS, LMWS, SKK, SMNK, SGS
5	Stuttgart, BSZ	5	1	2	06./07.12.2006 LMS
6	Stuttgart, BSZ	3	1	2	12./13.12.2006 SFG
7	Karlsruhe, SMNK	6	1	2	16./17.10.2006 SMNK
8	Karlsruhe, SMNK	2	1	2	10./11.01.2007 SMNK
9	Stuttgart, LMW	25	4	4	16.-18.4., 20.4.2007 LMWS
10	Gotha, SFG	7	1	2	12./13.06.2007 SFG/ SM
11	Freiburg, Sozialamt	12	1	2	26./27.06.2007 SMF
12	Mannheim, REMM	19	2	2	24./25.07.2007 REMM
13	Mannheim, REMM	8	1	1	29.08.2007 REMM
14	Mannheim, REMM	7	1	1	30.08.2007 Admin-Schulung REMM
15	Karlsruhe, SMNK	3	1	2	28./29.08.2007 SFG/ MdN
16	Gotha, SFG	4	1	1	11.10.2007 SFG/ SM+MRV
17	Mannheim, REMM	8	1	1	27.11.2007 REMM
18	Stuttgart, LMW	2	1	1	06.12.2007 LMW
19	Karlsruhe, SMNK	4	1	2	09./10.01.2008 SMNK
20	Gotha, SFG	11	2	3	05.-07.02.2008 SFG
Gesamt		146	26	38	

Tab. 2 IMDAS-Pro- Schulungen 2006/2007

Pan-Net Catalogue-Schulungen					
Nr.	Schulungsort	Teilnehmer	Kurse	Tage	Datum, geschulte Institution
1	Stuttgart, LMW	5	1	1	12.12.2007 SGS, LMW, SKK

Tab. 3 Pan-Net Catalogue-Schulungen 2006/2007

6.3 Nutzung des BAM-Portals

Erst mit der verstärkten Bewerbung des Portals und mit der Online-Stellung der neuen Portalversion im Januar 2007 war eine Untersuchung des Benutzerverhaltens sinnvoll.

Monat	Unterschiedliche Besucher	Anzahl der Besuche	Seiten	Zugriffe
Januar	511	1275	7936	22561
Februar	1123	3595	17513	75155
März	2006	5153	31040	138527
April	1472	4608	25530	89219
Mai	1650	4718	19633	90847
Juni	1629	4601	18645	85749
Juli	1544	4896	22902	100476
August	2483	5955	25160	140331
September	6417	11220	41714	356686
Oktober	11269	18538	45375	618551
November 07	15522	22458	54423	848827
Dezember	12589	16590	38743	644116
Total	58215	103607	348614	3211045

Tab. 4 Monatliche Zahlen der Besuche Januar 2007 - Dezember 2007

Seit Anfang 2007 werden die Logdateien ausgewertet, außerdem wurden die Benutzer aufgefordert, einen kurzen Fragebogen zur Bewertung der Funktionalität des Portals auszufüllen. Leider haben bisher zu wenige Besucher der Seite den Fragebogen ausgefüllt. Eine statistische Auswertung ist daher noch nicht sinnvoll gewesen.

Die Übersicht über die monatlichen Nutzerzahlen (Abb. 4) zeigt, dass mit Ausnahme des Dezembers die Nutzungszahlen im Verlauf des Jahres 2007 stetig gestiegen sind. Seit Januar 2008 liegen die entsprechenden Zahlen bei etwa tausend Zugriffen je Tag. Die Zahlen zu den Verweildauern sind noch nicht so aussagekräftig, dass daraus weitergehende Schlüsse gezogen werden können. Die Herkunft der Anfragen zeigt, dass ein großer Teil der Zugriffe über Suchmaschinen erfolgt. Es steht zu hoffen, dass mit der Einbindung von BAM-Referenzen in Wikipedia die Zahl der Zugriffe sich, wie in anderen Portalen, deutlich erhöhen wird.

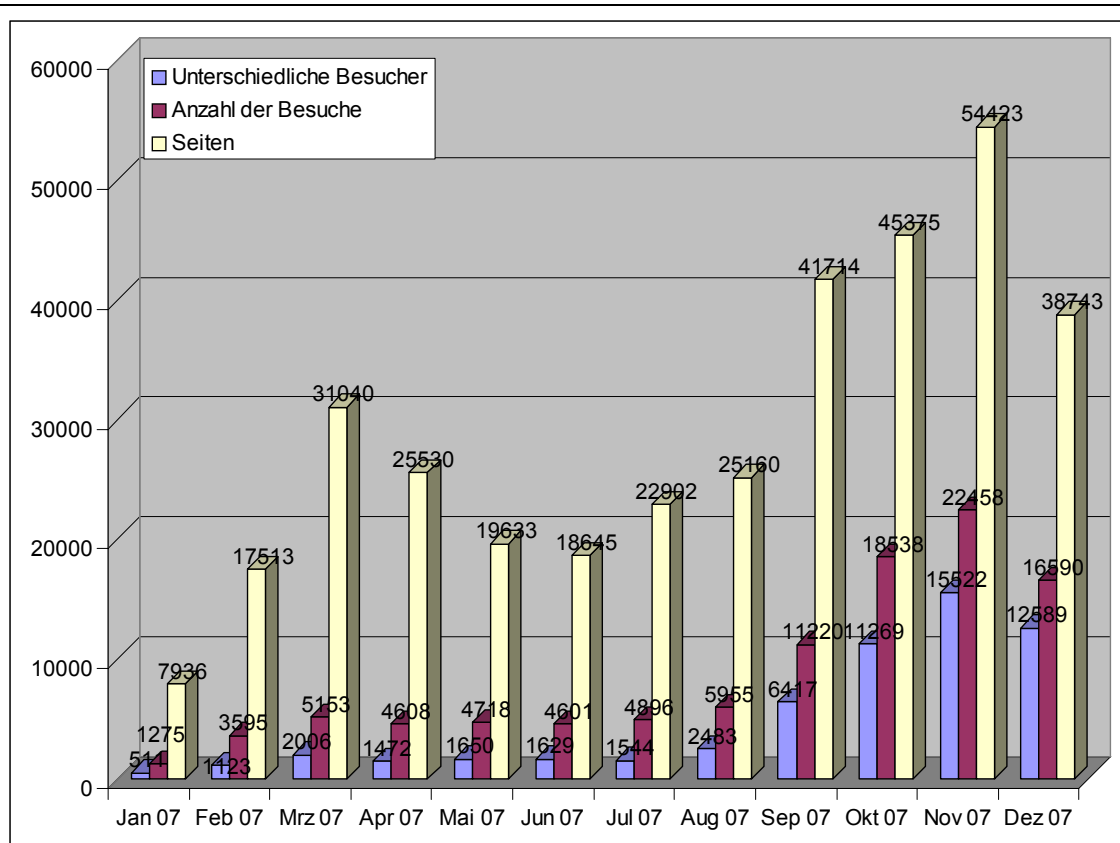


Abb. 4 Monatliche Zahl der Besucher Januar-Dezember 2007

Im August 2007 wurde die Vernetzung des BAM-Portals mit der deutschen Wikipedia entwickelt. Hierzu wurde eine sog. Wikipedia-Vorlage definiert. Ziel einer Wikipedia-Vorlage ist die vereinfachte Definition, Formatierung und Gestaltung von Dokumenten und Dokumentenfunktionen. Im vorliegenden Fall wird durch die Verwendung der Eingabe *BAM|Vincent|Van|Gogh* der Weblink „Informationen zu Vincent, Van, Gogh im BAM-Portal“ erzeugt. Dieser führt zu einer direkten Suche und dem Aufruf des BAM-Suchergebnisses. Dieses Werkzeug wird zur Zeit in ca. 500 Artikeln verwendet (Stand: Februar 2008).

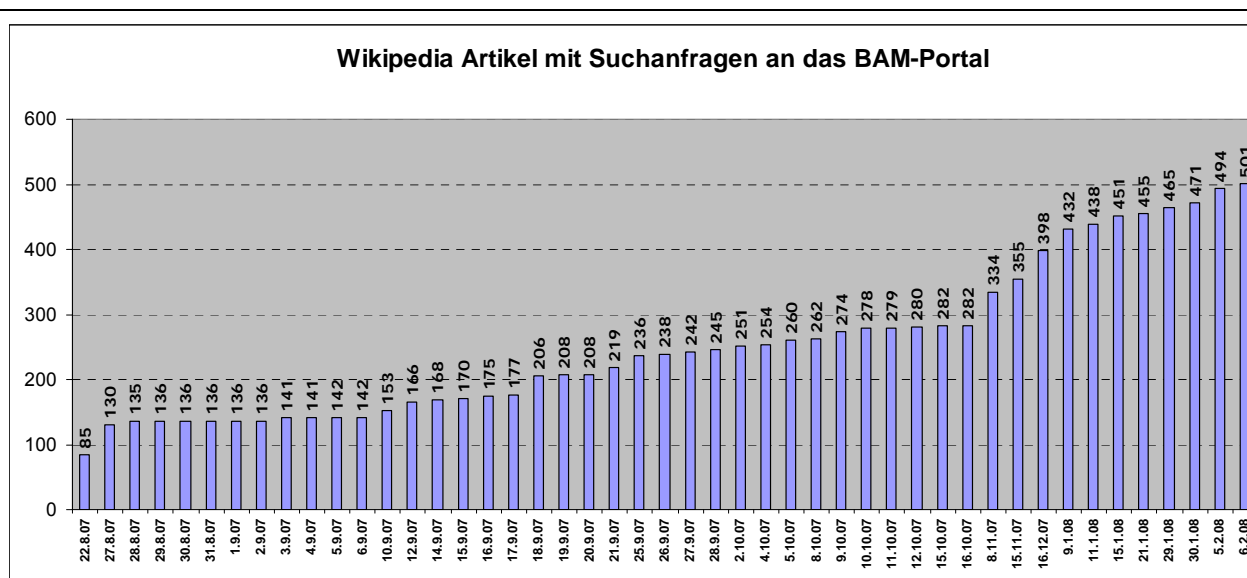
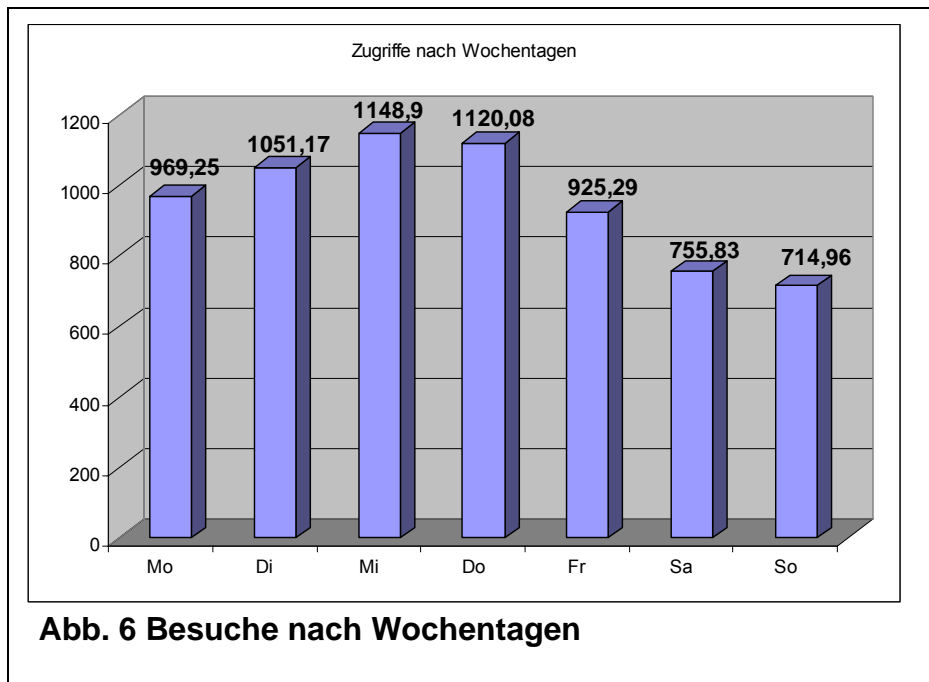


Abb.5 Wikipedia-Artikel mit Suchanfragen August 2007 - Januar 2008



Die nach Wochentagen aufgeschlüsselte Frequenz der Zugriffe (Abb. 6) macht deutlich, dass das BAM-Portal vermutlich vornehmlich von professionell Interessierten besucht wird. Diese Nutzergruppe (Schüler/innen, Studierende, Lehrende, Wissenschaftler/innen) wird vermutlich – wie erwartet und beabsichtigt – auf absehbare Zeit die Hauptgruppe bleiben. Erst wenn das Portal mit zusätzlichen Funktionalitäten auch für andere Gruppen – z. B. touristisch Interessierte – attraktiv wird, kann sich das ändern. Dafür müssten – neben der organisatorischen Abstimmung mit den Tourismusverbänden - aber Kapazitäten für die Entwicklung verfügbar sein.

7. Tagungen

7.1 Vom M-Team bzw. einzelnen Mitarbeiter/innen organisierte Tagungen

Datum, Ort	Titel
26. April 2006, SMNS	MusIS-Nutzertreffen
26. Juni 2006, BLMK	MusIS-AG Regelwerke
17. April 2007, ALMS Konstanz	Tagung der Museen und Schlösser der Euregio Bodensee: <i>Dokumentation in Museen</i>
3. Mai 2007, SMNS	MusIS-Nutzertreffen
27.-29. Juni 2007, LTAM	BAM-Kolloquium: <i>Quo vadis Digitalisierung?</i>
20. November 2007, LMWS	MusIS-Informationsveranstaltung: <i>Einsatzmöglichkeiten von Pan-Net Catalogue und InfoDesk in Museen und Museumsbibliotheken</i>

7.2 Teilnahme an Tagungen anderer Veranstalter und Messen

Datum, Ort	Titel der Tagung/Messe
9. bis 15. März 2006, Hannover	CeBIT, Präsentation am Stand des Landes Baden-Württemberg
5. bis 7. Juli 2006, Wien	Jahrestagung der ISKO (Vortrag Sieglerschmidt)
10. bis 14. September 2006, Göteborg	Jahrestagung des Comité international pour la documentation (CIDOC) du Conseil international des musées (ICOM) (Vortrag Sieglerschmidt/ von Hagel)
19. bis 22. September 2006, Konstanz	46. Deutschen Historikertag (Stand BSZ und Landesarchiv Baden-Württemberg)
20. September 2006, Bregenz	29. Österreichischer Bibliothekartag: <i>Wa(h)re Information</i> (Vortrag Sieglerschmidt)
31. Januar 2007, Regensburg	Universitätsbibliothek: Blockseminar <i>Archive und kulturelle Überlieferung</i> in der Studieneinheit <i>Informationskompetenz</i> (Vortrag Sieglerschmidt)
7. Februar 2007, Göttingen	6. Medienforum des IWF: <i>Audiovisuelle Wissensmedien online - Repositories, Archive und Möglichkeiten der Vernetzung</i> , (Vortrag Sieglerschmidt)
21. bis 22. Februar 2007, Berlin	<i>Eudico-Konferenz Europas kulturelles und wissenschaftliches Erbe in einer digitalen Welt.</i> (Vortrag Mallmann-Biehler)
7./8. März 2007, Freiburg	Jahreskonferenz der Deutschen Gesellschaft für Klassifikation: <i>Zutrauen in die digitale Zukunft</i> (Vortrag Sieglerschmidt)
19. bis 22. März 2007, Leipzig	3. Leipziger Kongress für Information und Bibliothek: <i>Information und Ethik</i> (Vortrag Sieglerschmidt)
17. April 2007, Konstanz	Tagung der Museen und Schlösser der Euregio Bodensee: <i>Dokumentation in Museen</i> (Vortrag Lill, Kirchhoff)
6. Juni 2007, Frankfurt	Tagung der Fachgruppe Dokumentation im Deutschen Museumsbund im Museum Giersch (Vortrag Lill)
20. bis 22. August 2007, Wien	Jahrestagung des CIDOC (Vortrag Sieglerschmidt)
2. September 2007, Dresden	Nationaler Aktionstag Restaurierung und Digitalisierung (Vortrag Mallmann-Biehler: <i>BAM – Bibliotheken, Archive, Museen</i>)
4. September 2007, Weimar	7. fachübergreifende Tagung der Bauhaus-Universität <i>Mensch & Computer 2007: Blended museum</i> (Vortrag Schweibenz)
17. bis 21. September 2007, Graz	Odok: Informationskonzepte für die Zukunft, im Workshop Kulturdatenmanagement: <i>Datenverbund und virtueller Katalog</i> (Bibliotheken, Archive, Museen) (Vortrag Kirchhoff/ Sieglerschmidt)

Datum, Ort	Titel der Tagung/Messe
25. bis 28. September 2007, Mannheim	Deutscher Archivtag (Vortrag Sieglerschmidt)
8. Oktober 2007, Berlin	Herbsttreffen der Fachgruppe Dokumentation im Rahmen der AG Multimedia im Zuse-Institut (Vortrag Mönicke, Schweibenz)
30. Oktober bis 3. November 2007, Köln	Exponatec: Stand des BSZ

7.3 Gremien und feste Arbeitsgruppen

7.3.1 Deutscher Museumsbund

Die Mitarbeit in den Gremien des DMB, insbesondere in der Fachgruppe Dokumentation und ihren Arbeitsgruppen, ist für das M-Team wichtig, nicht nur als Forum für eigene Aktivitäten, sondern auch als Möglichkeit, Informationen aus anderen, vergleichbaren Kooperationsverbänden zu bekommen und Kooperationspartner für Projekte im Dokumentationsbereich zu finden. Solche Informationen werden an die am Kooperationsverbund beteiligten Museen weitergegeben, um diese über aktuelle Entwicklungen im Bereich Dokumentation/Internetpräsentation zu unterrichten. Neben dem Frühjahrs- und Herbsttreffen der Fachgruppe Dokumentation wird versucht, in folgenden Arbeitsgruppen kontinuierlich mitzuarbeiten: AG Regelwerke, AG Datenaustausch, AG Softwarevergleich, AG Sammlungsmanagement, AG Multimedia.

7.3.2 Minerva-/EUBAM-Spiegelgruppe

Die Arbeitsgruppe EUBAM ist vom damaligen Bundesbeauftragten für Kultur und Medien (BKM) initiiert worden, damit die Digitalisierungsarbeiten in Deutschland koordiniert und vor allem gegenüber der EU besser vertreten werden. Eine ihrer Aufgaben ist daher der fach- und bundesländerübergreifende Austausch über nationale Digitalisierungsprojekte, die Vermittlung zu EU-relevanten Arbeitsstellen und die Unterstützung der deutschen Vertreter für Bund und Länder in der Einbringung deutscher Konzepte und Interessen auf europäischer Ebene in der *National Representatives Group* (NRG). Das ist eine europäische Arbeitsgruppe mit offiziell benannten Vertretern aus allen EU-Mitgliedstaaten zu Fragen des Umgangs mit digitalen Daten im Kulturbereich, die zweimal jährlich tagt und von der EU über das Projekt Minerva Plus unterstützt wird. Gleichzeitig ist es Aufgabe der EUBAM Gruppe, die deutschen Kultureinrichtungen über europäische Projekte zu informieren. Die sog. Spiegelgruppe hat die Aufgabe, die EUBAM- und MINERVA-Arbeitsgruppen durch Bereitstellung von Informationen über die deutschen Aktivitäten im Bereich Digitalisierung (im weiteren Verständnis des Begriffes) aktiv zu unterstützen. Mittlerweile ist eine durch die DFG finanzierte neue Internetplattform für den Nachweis deutscher Digitalisierungsprojekte aufgebaut worden.¹⁸

7.3.3 Deutsche Digitale Bibliothek

Seit Herbst 2007 beteiligt sich das BSZ an der Vorbereitung des Aufbaus der Deutschen Digitalen Bibliothek. Sie ist neben den anderen nationalstaatlichen Anstrengungen zur Präsentation des europäischen Kulturerbes ein Baustein der

¹⁸ <http://www.dl-forum.de/deutsch/foren/25_1333_DEU_HTML.htm>.

*Europeana*¹⁹ (zuvor: European Digital Library). Eine vom Kulturausschuss der Kultusministerkonferenz eingesetzte Bund-Länder-Fachgruppe hat ein fachliches Konzept ausgearbeitet, das die fachlichen Anforderungen an ein deutsches Internet-Portal formuliert. Mit ersten Ergebnissen ist bis Ende 2008 zu rechnen.

7.3.4 Wissenschaftliche Institutskommission des Instituts für Museumsforschung Berlin

Seit Dezember 2005 bis 2009 ist Dr. Jörn Sieglerschmidt gewähltes Mitglied – im Gegensatz zu den kraft Amtes vertretenen - im o. g. Gremium, das Aufgaben als wissenschaftlicher Beirat des IfM erfüllt. Die erste Sitzung fand am 5. Dezember 2005 statt.

¹⁹ <<http://www.europeana.eu/>>

8. Projekte

8.1 Sacherschließung

Nachdem OBZ- und Materialthesaurus sowie verschiedene Datierungsthesauri vorliegen, muss jetzt noch der Technikthesaurus aufgebaut werden. Ergebnisse dieser Arbeit können im Herbst 2008 erwartet werden.

8.2 Kooperative Erschließung

Seit längerer Zeit ist ein Projekt geplant, bei dem es um die kooperative Erschließung von Museumsobjekten geht. Die Mehrheit der Museumsobjekte gilt als Unikat, das eine gemeinsame Erschließung nicht erlaubt. Es gibt aber Museumsbestände, bei denen Objekte in Masse oder zumindest in größerer Zahl produziert wurden, wie z. B. Druckgraphiken oder Fundmünzen, die eine gemeinsame Erfassung durch die Museen erlauben, wie sie auch im Bibliotheksbereich üblich ist. Zusammen mit zwei staatlichen Museen in Baden-Württemberg und der Stiftung Preußischer Kulturbesitz (SPK) sollen zum einen analoges Material digitalisiert und zum anderen *online*-Werkzeuge entwickelt werden, die eine regelwerksgestützte Erfassung der digitalisierten Objekte erlauben.

8.3 Datenschutz und IT-Sicherheitskonzept für den Einsatz von Informations- und Kommunikationstechnik in den Kunsteinrichtungen des Landes Baden-Württemberg

Seit Dezember 2006 läuft u. a. in Kooperation mit dem LMS ein Pilotprojekt zur Einführung und Umsetzung eines IT-Sicherheitskonzeptes nach dem vom Land Baden-Württemberg empfohlenen Standard des *Bundesamtes für Sicherheit in der Informationstechnik* (BSI). Das Ergebnis dieser Arbeit wird u. a. ein Handbuch mit Empfehlungen für Maßnahmen zur Datensicherheit und zum Datenschutz (sogenannter IT-Grundschutz) sein.

Laut der Rahmenempfehlung des MWK, ist der sichere Einsatz der Informationstechnik (IT) nachzuweisen. Empfohlen wird dabei ein Vorgehen nach den Standards des BSI. Die Umsetzung dieser Anforderungen betrifft sowohl die einzelnen Einrichtungen, bei denen schützenswerten Daten anfallen, als auch das BSZ, das Datenverarbeitung im Auftrag dieser Institutionen ausführt. Das BSZ wird nicht nur die Anforderungen des IT-Grundschutzes für sich selbst umsetzen, sondern auch die betreuten Institutionen darin unterstützen. Diese Unterstützung stellt für die Landesmuseen eine wichtige Hilfe dar, weil in den Institutionen in der Regel keine Erfahrung mit der Umsetzung solch komplexer Anforderungen besteht.

8.4 Datensicherung, Datennutzung und Langzeitarchivierung für die Staatlichen Museen des Landes Baden-Württemberg

Das BSZ erarbeitet derzeit ein Konzept für die Sicherung und Langzeitarchivierung von Bilddateien, aber auch anderen Medienobjekten, aus den Staatlichen Museen. Der erste Schritt ist die externe Sicherung dieser Museumsdateien durch das BSZ, damit außerhalb der Institutionen eine Sicherungskopie dieser teilweise unersetzlichen Daten

vorliegt. In einem weiteren Schritt soll eine weitergehende Nutzung der Medienobjekte ermöglicht werden, beispielsweise in IMDAS-Pro und im BAM-Portal sowie in Digital Media Asset Management Systemen für den Vertrieb von Medien über das Internet. Parallel dazu sollen die Medienobjekte für die Langzeitarchivierung vorbereitet werden. Dazu müssen sie mit den erforderlichen Prüfsummen und Metadaten versehen werden, die für eine sichere Archivierung unabdingbar sind. Ist dies erfolgt, können sie entsprechend archiviert werden.

In Kooperation mit der SGS hat das BSZ beim MWK einen Antrag auf Förderung eines Projekts zur Vorbereitung der Datensicherung und Langzeitarchivierung von Bilddaten gestellt. Die Realisierung dieses Projektes ist für das Jahr 2008 vorgesehen. Als erste Vorarbeiten wurden bereits im Herbst 2007 Testübertragungen von Bilddaten mit Prüfsummen durch die SGS vorgenommen. Im Rahmen des Projekts sollen nun sowohl die erforderlichen Programme für die Datenprüfung (beispielsweise auf Standardkonformität der Formate, Abgleich der Prüfsummen, Umkopieren der Daten in andere Formate, etc.) als auch die organisatorischen Abläufe erprobt werden. Die Ergebnisse des Projektes werden allen Staatlichen Museen zur Verfügung stehen.

8.5 Inventarisierung der Sammlung Kindermedien der Hochschule der Medien in Stuttgart

An der Hochschule der Medien (HdM) in Stuttgart (Prof. Dr. Richard Stang, Prof. Dr. Manfred Nagl) wird ein Ausstellungsprojekt durchgeführt, das hinsichtlich der Dokumentation und der Ausstellungsplanung vom BSZ unterstützt wird. Das Institut für angewandte Kindermedienforschung (IfaK) der HdM verfügt über eine Sammlung von Kindermedien, die bis in das 19. Jh. zurückreicht. Studierende des Studiengangs Bibliotheks- und Informationsmanagement werten die Sammlung aus und entwickeln Konzepte für eine mobile Ausstellung, die bis Anfang 2009 realisiert werden soll.

9. Kooperationen

9.1 DigiCult Schleswig-Holstein

In Hinsicht auf die Betreuung der Museen verfolgt DSH.²⁰ denselben Ansatz wie das BSZ. So verfolgt auch DSH eine klare Linie hinsichtlich der Thesaurusarbeit und hat sich für die SWD als Standardvokabular entschieden. DSH hat seit einigen Jahren Software implementiert, die eine kooperative Entwicklung gemeinsamer Vokabularien erlaubt.

9.2 Institut für Museumsforschung SPK

Die Zusammenarbeit mit dem IfM in allen Fragen der Dokumentation, insbesondere Regelwerksfragen, besteht seit Beginn des MusIS-Kooperationsverbundes. Da das IfM die einzige länderübergreifende Beratungsinstitution für deutsche Museen ist, hat es zugleich die Aufgabe, die länderübergreifende Kooperation der Museen zu fördern. Das geschieht hinsichtlich der Dokumentationsfragen in der 1994 geschaffenen entsprechenden Fachgruppe des Deutschen Museumsbundes (DMB). Da das BSZ regional gleichartige Aufgaben wahrnimmt, ergibt sich ein ständiger Abstimmungsbedarf.

²⁰.<<http://www.digicult-sh.de/>> und < <http://www.museen-sh.de/ml/digicult.php> >.

10. Mitarbeiter/innen und Arbeitsbereiche

10.1 Mitarbeiter und Kontaktangaben

Name	Telefon	E-Mail
Thomas Kirchhoff; Dipl.-Physiker	07531/88-4120	thomas.kirchhoff@bsz-bw.de
Jens Lill; Dipl.-Bibliothekar (FH)	07531/88-4932	jens.lill@bsz-bw.de
Kathleen Mönicke, Dipl.-Museologin (FH)	07531/88-4846	kathleen.moenicke@bsz-bw.de
Francisco Peña Diaz, Dipl.-Ing. (FH)	07531/88-4931	francisco.pena@bsz-bw.de
Giovanna Ratini, Datenbankadministratorin	07531/88-4932	giovanna.ratini@bsz-bw.de
Werner Schweibenz, M.A.	07531/88-4279	werner.schweibenz@bsz-bw.de
Dr. Jörn Sieglerschmidt, Oberkonservator	07531/88-3049	joern.sieglerschmidt@bsz-bw.de
Sophie-Luise Wollrab, Dipl.-Museologin (FH)	07531/88-4846	sophie.wollrab@bsz-bw.de

Die E-Mail-Adresse für alle Anfragen an das M-Team lautet m-team@bsz-bw.de

10.2 Arbeitsbereiche und Ansprechpartner

Arbeitsbereiche	Ansprechpartner/innen
Allgemeines Künstlerlexikon	Ratini, Schweibenz
Arbeitsgruppe Multimedia im DMB	Schweibenz, Mönicke
Arbeitsgruppe Regelwerke im DMB	Lill
Arbeitsgruppe Sammlungsmanagement im DMB	Mönicke, Wollrab
Archäologisches Landesmuseum Baden-Württemberg	Sieglerschmidt
Archive	Sieglerschmidt
Badisches Landesmuseum Karlsruhe	Wollrab
BAM-Portal	Kirchhoff, Sieglerschmidt
Betriebssysteme, Netzwerk- und Hardwarebetreuung	Peña Diaz, Ratini
BSZ-Importroutine	Ratini, Wollrab
Datierungsthesaurus	Mönicke
IMDAS-Pro-Beratung: GIS-Modul	Sieglerschmidt
IMDAS-Pro-Beratung: Layoutmodul	Sieglerschmidt, Schweibenz
IMDAS-Pro-Beratung: Media-Modul, Fotoverwaltung	Schweibenz

Arbeitsbereiche	Ansprechpartner/innen
IMDAS-Pro-Beratung: Konfiguration benutzerdefinierter Masken	Wollrab
IMDAS-Pro-Beratung: Thesaurusfragen	Lill
IMDAS-Pro-Stammdaten: Pflege und Beratung	Wollrab
Landesmuseum Württemberg Stuttgart	Mönicke
Langzeitarchivierung	Schweibenz
Linden-Museum Stuttgart	Schweibenz
Material- und Technikthesaurus	Sieglerschmidt
Ministerium für Wissenschaft, Forschung und Kunst	Wollrab
Objektbezeichnungsthesaurus; SWD-Erfassung	Lill
Ortsthesaurus	Sieglerschmidt
Pan-Net Catalogue	Mönicke
Personenstammdaten	Sieglerschmidt
Reiss-Engelhorn-Museen Mannheim	Mönicke
Softwarebetreuung	Peña Diaz
Staatliche Kunsthalle Karlsruhe	Lill
Staatliches Museum für Naturkunde Karlsruhe	Wollrab
Staatsgalerie Stuttgart	Mönicke
Städtische Museen Freiburg	Schweibenz
Stammdaten: Pflege und Beratung	Wollrab
Stiftung Schloss Friedenstein Gotha	Lill

11. Veröffentlichungen zur Museumsdokumentation und zur Museumskunde

Jens Lill:

Der Objektbezeichnungsthesaurus des BSZ. Informationsveranstaltung zu Fragen der Dokumentation in Museen. In: <http://www2.bsz-bw.de/cms/museen/musis/veranstaltungen/Lill_textfassung-OBZ.pdf> (s. Anhang 4).

Werner Schweibenz:

Besucherinteressen im traditionellen und virtuellen Museum. Konferenz Mensch & Computer Bauhaus-Universität Weimar 2007. In: <http://www2.bsz-bw.de/cms/museen/musis/publ/Schweibenz_KMD-Track_Weimar_Besucherinteressen.pdf>. [im Druck]

Jörn Sieglerschmidt:

(Hg. zus. mit H. Peter Ohly, Christian Swertz): Wissensorganisation und Verantwortung: Gesellschaftliche, ökonomische und technische Aspekte. (Fortschritte in der Wissensorganisation 9) Würzburg: Ergon 2006.

Aufbau und Nutzbarkeit von Portalen – theoretische und praktische Überlegungen zu öffentlich finanzierten Kulturportalen. In: ebda. 2006, 330-337.

Convergence of internet services in the cultural heritage sector - the long way to common vocabularies, metadata formats, ontologies. ISKO-Konferenz 2006 Wien. In: Ohly, H. Peter/ Netscher, Sebastian/ Mitgutsch, Konstantin (Hgg.): Kompatibilität, Medien und Ethik in der Wissensorganisation (Compatibility, media, and ethics in knowledge organization). (Fortschritte der Wissensorganisation/ Advances in knowledge organization 10) Würzburg: Ergon 2008, 154-160.

(zus. mit Frank von Hagel) Convergence of internet services in the cultural heritage sector - the long way to common vocabularies, metadata formats, ontologies. CIDOC-Jahrestagung 2006. In: <<http://www2.bsz-bw.de/cms/museen/musis/publ/goete12vhjs.pdf>> (s. Anhang 2).

Herbarium. In: JÄGER, Friedrich (Hg.): Enzyklopädie der Neuzeit. Bd. 5. Stuttgart: Metzler 2007, 378-381.

Knowledge organization and multilingual vocabularies. CIDOC-Jahrestagung 2007. In: <<http://www2.bsz-bw.de/cms/museen/musis/publ/wien-cidoc2007.pdf>> (s. Anhang 3).

12. Abkürzungsverzeichnis

- AKL
Allgemeines Künstlerlexikon 8
- ALMS
Archäologisches Landesmuseum
Stuttgart/ Konstanz 7
- ASP
application service provider 6
- BAM
Gemeinsames Portal zu Bibliotheken,
Archiven und Museen 3
- BKM
Bundesbeauftragter für Kultur und
Medien 28
- BLMK
Badisches Landesmuseum Karlsruhe
6
- BSZ
Bibliotheksservice-Zentrum Baden-
Württemberg 3
- CIDOC
Comité international pour la
documentation 27
- DMB
Deutscher Museumsbund 19
- DNB
Deutschen Nationalbibliothek 19
- DSH
DigiCult Schleswig-Holstein 18
- GBV
Gemeinsamer Bibliotheksverbund 17
- GUID
global(ly) unique identifier 18
- GVK
Gemeinsamer Verbundkatalog 17
- HdM
Hochschule der Medien 31
- HGS
Haus der Geschichte Baden-
Württemberg Stuttgart 7
- ICOM
International Council of Museums 27
- IFaK
Institut für angewandte
Kindermedienforschung 31
- IfM
Institut für Museumsforschung 22
- ISKO
International Society for Knowledge
Organization 27
- JR
Joanneum Research 12
- LABW
Landesarchiv Baden-Württemberg 7
- LMS
Lindenmuseum Stuttgart 6
- LMWS
Landesmuseum Württemberg
Stuttgart 6
- LTAM
Stiftung Landesmuseum für Technik
und Arbeit in Mannheim 7
- MPK
Museum Pfalzgalerie Kaiserslautern
7
- MusIS
Museums-Informations-System 3
- MWK
Ministerium für Wissenschaft,
Forschung und Kunst Baden-
Württemberg 6, 8
- OBZ
Objektbezeichnungsthesaurus 20
- PND
Personennamendatei 21
- REMM
Reiss-Engelhorn-Museen Mannheim
6
- RSWK
Regeln für den Schlagwortkatalog 20
- SFG
Stiftung Schloß Friedenstein Gotha 7
- SKK
Staatliche Kunsthalle Karlsruhe 6
- SMF
Städtische Museen Freiburg 6
- SMNK
Staatliches Museum für Naturkunde
Karlsruhe 6
- SMNS
Staatliches Museum für Naturkunde
Stuttgart 7
- SWB
Südwestdeutscher Bibliotheksverbund
17
- SWD
Schlagwortnormdatei der Deutschen
Nationalbibliothek 19
- TGN
Thesaurus of Geographic Names 21
- ZIB

Zuse-Zentrum für Informationstechnik
Berlin 18
ZKA
Zentralkartei der Autographen und
Nachlässe 18

ZVDD
Zentralverzeichnis digitalisierter
Drucke 18

13. Anhänge

Anhang 1: Dienstleistungen des BSZ im Museumsbereich

lfd. Nr.	Leistungsbereiche	Einzelleistungen
Technische Betreuung		
1	Sicherheit	Sicherheitsmaßnahmen (u. a. Firewall, Virenschutz)
2	Sicherheit	Datensicherung (Backup)
3	Beratung	Beratung bei der Definition von Schnittstellen
4	Beratung	Beratung in technischen Fragen, insbesondere System- und Netzwerkfragen
5	Koordination	BelWue-Anschluss
6	Koordination	Informationsbeschaffung und Weitergabe der Informationen an die Museen (z. B. technische Neuentwicklungen, Langzeitarchivierung usw.)
7	Koordination	Verhandlungen mit Dritten, z. B. Anbieterfirmen
8	Dienstleistung	Web-Hosting von Daten für Projekte etc.
9	Querschnittskosten	Server- und PC-Ausstattung, Datenübertragungskosten, Bandsicherungsroboter, etc.
IMDAS-Pro		
1	technisch	Lizenzgebühr für Datenbanksoftware (Oracle)
2	technisch	Installation und Betreuung der Objektdatenbank
3	technisch	Installation und Betreuung des Client-Server-Systems
4	technisch	Unterstützung bei der Installation der Client-Software
5	technisch	Benutzer- und Lizenzverwaltung
6	fachlich	Konfiguration von IMDAS-Pro, z. B. Einrichtung der Nutzerberechtigungen, Import von Thesauri usw.
7	fachlich	Software-Schulungen (Anwender, Administration)
8	fachlich	Beratung der Nutzer/innen (Hotline)
9	fachlich	Entwicklung von Datenfelddefinitionen und Online-Hilfen
10	technisch	technische Unterstützung bei der Datenmigration und Beratung bei der vorherigen bzw. Nachbearbeitung der Daten
11	technisch	Import und Export von Daten, z. B. bei Massendatenkorrekturen
12	technisch	Anpassungen des Drucklayouts
13	technisch	technische Integration von Normdaten: Allgemeines Künstlerlexikon
14	technisch	technische Integration von Normdaten: ICONCLASS, ggf. TGN
15	technisch	technische Integration von Webservices, z. B. für Thesauri
16	Beratung	Herstellung fachlicher Kontakte und Weitervermittlung an die Museen
17	Koordination	Ansprechpartner für Hersteller und Anwender/innen
18	technisch	Systemtests und Abnahme bei Software-Versionswechseln
19	technisch	Verhandlungsführung mit Herstellerfirmen über technische Weiterentwicklung der Software und Fehlerbehebung
20	technisch	Beauftragung von individuellen Programmiererweiterungen und -anpassungen mit Herstellerfirmen
Fachliche Betreuung der Dokumentation (nur in Verbindung mit IMDAS-Pro und Datenbankbetreuung)		

lfd. Nr.	Leistungsbereiche	Einzelleistungen
1	Koordination	Organisation und Koordinierung der kooperativen Sacherschließung im MusIS-Verbund
2	Koordination	Einrichtung von Diskussionslisten und -foren
3	Koordination	Organisation und Koordinierung der Arbeitsgruppen im MusIS-Verbund (z.B. Regelwerke)
4	Beratung	Veranstaltung von Fortbildungen und Tagungen
5	Beratung	Beratung in Fragen der Museumsbibliothek
7	Beratung	Beratung bei der Projektentwicklung
8	Koordination	Einbeziehung in überregionale Projekte und Portale, z. B. BAM-Portal
Fachliche Betreuung der Dokumentation (ohne Datenbankbetreuung durch MusIS)		
1	fachlich	Dokumentationsschulung
2	fachlich	Beratung der Nutzer/innen (Hotline)
3	fachlich	Beratung in Fragen der Sacherschließung (Thesauri, Klassifikationen usw.)
4	fachlich	Organisation und Koordinierung von Arbeitsgruppen
5	fachlich	Veranstaltungen von Fortbildungen und Tagungen
6	technisch	Verhandlungsführung mit Herstellerfirmen über technische Weiterentwicklung der Software und Fehlerbehebung
7	technisch	Entwicklung von Datenfelddefinitionen und Online-Hilfen
8	Koordination	Einbeziehung in überregionale Projekte und Portale, z. B. BAM-Portal
Pan-Net Catalogue		
1	technisch	Bereitstellung von Speicherplatz, Datensicherung
2	technisch	Lizenzgebühr für das CMS (Basistarif, 1 Arbeitsplatz)
3	technisch	Installation und Betreuung des Web-CMS
4	technisch	Gebühr für das Layout des CMS
5	technisch	Betreuung der Anwendung
6	technisch und fachlich	Verhandlungsführung mit Herstellerfirmen über technische Weiterentwicklung der Software und Fehlerbehebung
7	fachlich	Software-Schulungen
8	fachlich	Beratung der Nutzer/innen (Hotline)

Anhang 2: Convergence of internet services in the cultural heritage sector – the long way to common vocabularies, metadata formats, ontologies

Jörn Sieglerschmidt, Bibliothekservice-Zentrum Baden-Württemberg, Konstanz
Frank von Hagel, Institut für Museumskunde, Staatliche Museen zu Berlin, Stiftung Preußischer Kulturbesitz²¹

Summary

Since several years it has been observed that information offered by different knowledge producing institutions on the internet is more and more interlinked. This tendency will increase, because the fragmented information offers on the internet make the retrieval of information difficult or even impossible. At the same time the quantity of information offered on the internet grows exponentially in Europe – and elsewhere - due to many digitization projects. Insofar as funding institutions base the acceptance of projects on the observation of certain documentation standards the knowledge created will be retrievable and will remain so for a long time. Otherwise the retrieval of information will become a matter of chance due to the limits of fragmented, knowledge producing social groups.

Keywords:

ontologies, metadata formats, authority files, CIDOC-CRM, portals, cultural heritage

On the following pages we shall present some standards and how they are used up to present. Subsequently we shall deal with European Archive-Library-Museum (ALM) projects and, in particular, the German ALM project, its implementation, its claims and its future targets.

Regarding standards we have to distinguish between syntax and semantics. Our use of these terms doesn't or rather only partly correspond to the concepts of information science.²² Data carry the content of a communication. It can therefore be analyzed semantically. The structure of data cannot be completely separated from the content; because application rules influence the content as – reversely – the content or rather the intended communication influences the application rules. Grammar and - in the narrower sense - syntax perform, but don't determine the semantics. Therefore only statements about objects can carry meaning; those objects are the contents (text, image, and sound) offered via the internet.

Texts – and only these, as long as image or sound retrieval devices are not reliable - can be indexed with the help of authority files. Well-known authority files of that kind are for example the thesauri of the Getty Foundation (TGN, AAT, ULAN)²³ or the subject headings of the national libraries (e. g. DE: SWD, FR: RAMEAU, subject headings used by the national libraries UK: LCSH)²⁴. The advantage of using such authority files in linking different internet resources is the use of well-defined ID numbers assigned to

²¹<joern.sieglerschmidt@bsz-bw.de>; <f.v.hagel@smb.spk-berlin.de>.

²²Sieglerschmidt, J., Metadaten: <<http://www2.bsz-bw.de/cms/service/museen/publ/metadaten-js-2002.pdf>>.

²³Getty Vocabularies: <http://www.getty.edu/research/conducting_research/vocabularies/>.

²⁴The Schlagwortnormdatei der Deutschen Bibliothek is not available online; see for basic information <<http://www.ddb.de/eng/standardisierung/normdateien/swd.htm>>; Subject Headings of the Library of Congress: <<http://authorities.loc.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First>>; Répertoire d'autorité-matière encyclopédique et alphabétique unifié der Bibliothèque Nationale de France: <<http://catalogue.bnf.fr/servlet/AccueilConnecte?Autorites=RAMEAU>>.

every term. A museum object like an airplane (German: Flugzeug) could be linked to literature of libraries by using this number (DE: SWD: 4017672-1; FRBNF11931002). Furthermore the user could get information about similar objects and literature by using the semantic net that the terms of a thesaurus offer (synonyms, related, narrower, broader terms etc.). Multilingual tools are indispensable for the future of internet retrieval. The German term *Fahrzeug* (broader term of Flugzeug, SWD-ID 4016320-9) has the same meaning as the French *véhicules* (FRBNF11975775) and the British/American *vehicles* (US/GB: LSCH). The aim of the former MACS²⁵ and the actual CrissCross²⁶ project is to coordinate – not to translate – those terms, in order to allow the retrieval of objects in different foreign language OPACs. Thus it will be possible to use multilingual semantic nets for internet retrieval – similar to the semantic web, where terminological combinations are used for linking customer services²⁷. Back to our example: The reverse path could be followed viz. the reference from literature found in an OPAC to museum objects. The efforts of catalogue enrichment aim at services like that. The example might be sufficient in order to show that here m:n-relations are concerned; a semantic space that has been widened by the use of authority files and that could be expanded by supplementary dimensions. Going beyond such attempts the semantic web offers help designing typical scenarios of social action, i. e. workflows that are usually tied together. Libraries and their online services have realized such workflows already. Thus a book that could not be found in the library nearby will be ordered by interlibrary loan. Or an article can be ordered as pdf-file using the *subito*-portal.²⁸ Archival online platforms have partly realized functions like that, and museums could do that, too.²⁹ Basically the task is to model repetitive workflows (administrative processes) and to extract the relevant information using the model structure and to offer services linked to it.

The syntax of the data offered on the internet is equally important for linking information, because it defines where and which type of data can be found using certain orthographical notation rules. All kinds of databases use such syntactical rules (fields, types of fields, rules for data entry). Usually different data are separated by tags or metadata structures. The Dublin Core Metadata Element Set (DCMES) is well known in that respect, but has lost some importance during the last years.³⁰ While in early ALM projects like the German BAM project the DCMES seemed to be an appropriate metadata format for all three types of documentation structure, current projects like the Finnish MuseoSuomi prefer a combination of ontologies and XML. The metadata formats EAD and TEI are designed for the recording of archival documents and autographs.³¹ The Getty Foundation developed metadata standards for the museums: Cataloguing Cultural Objects (CCO) and Categories for the Description of Works of Art (CDWA, CDWA lite).³² The XML schema ABCD of the Taxonomic Databases Working Group (TDWG) of the Global Biodiversity Information Facility (GBIF) is relevant for natural history collections.³³ The Metadata Encoding and Transmission Standard

²⁵ Multilingual Access to Subjects: <<https://ilmacs.uvt.nl/pub/>>.

²⁶ CrissCross: <<http://www.ddb.de/eng/wir/projekte/crisscross.htm>>.

²⁷ For Information see: <<http://www.w3.org/2001/sw/>>.

²⁸ subito library document delivery service: <<http://www.subito-doc.de/>> engl. Version <<http://www.subito-doc.com/>>.

²⁹ cf. <www.bpk-images.de>.

³⁰ ISO 15836 (draft): <<http://www.niso.org/international/SC4/n515.pdf>>.

³¹ Encoded Archival Description, supported from the Library of Congress: <<http://www.loc.gov/ead/>>; Text Encoding Initiative: <<http://www.tei-c.org/>>.

³² <http://www.getty.edu/research/conducting_research/standards/cdwa/8_printing_options/definitions.pdf>.

³³ <<http://www.gbif.org/>>.

(METS) supported by the Library of Congress likewise is a XML schema aiming at the description and linking of different internet resources.³⁴

All metadata formats have a rigid, rather inflexible structure. The links between metadata elements disappear because the elements are ordered parallelly and not hierarchically. Retrieval strategies must use complex and-combinations in order to get precise and/or complete results. The question e. g., which artists worked in certain places, cannot be answered. Information of a complex structure cannot be modelled appropriately by flat metadata structures.

Ontologies offer a solution for such problems because they are very flexible regarding the adaptation to new requirements. They can model enormously complex structures without producing confusion. The example of the Conceptual Reference Model (CRM) of the Comité International pour la documentation (CIDOC), one of the committees of the International Council of Museums (ICOM), does just this. The CIDOC-CRM is evaluated as draft by the bodies of the International Standards Organization (ISO).³⁵ The CIDOC-CRM corresponds to the requirements of the Resource Description Framework (RDF(S)) and is compatible with object oriented and relational database models (R and OO DBMS) and with XML, especially by its infinitely nested structure. The CRM has been mapped to DCMES, EAD, TEI, and FRBR.³⁶ These mappings show that the CRM is very powerful and more consistent than other standards. The CRM could therefore be recommended as a general reference model for specific metadata and database formats of all kinds. The CRM offers a model for all knowledge producing institutions, because even complex scholarly contents can be handled.

The CRM distinguishes classes (84 *entities* so far) and relations (141 *properties* so far). The CRM has a syntactical structure with (historical) events as main reference points. Classes have content only if a reference is given. Classes can be repeated infinitely. Classes and properties are organised in a hierarchical structure of supraclasses, subclasses, and elementary classes. Supraclasses are apart from the term class itself:

- place,
- temporal entities,
- physical stuff,
- actors,
- conceptual objects.

Primitive values, i. e. classes without relations to others, are:

- number,
- time primitive and
- string.

Detailed information can be downloaded from the internet site of the CIDOC-CRM.

Historical knowledge is incomplete, not finite, and not conclusive. Its elements show a varying statistical stability regarding the revision of knowledge by current discourse, existence in real life, identity and relations to other elements of knowledge. The CRM wants to describe the structure of knowledge and its change. In order to enhance uniformity it structurally and conceptionally prefers transitions instead of fixed states, opinions instead of propositions. It avoids the repetition of relationships and excludes exceptions.

³⁴ <<http://www.loc.gov/standards/mets/>>.

³⁵ CIDOC-CRM in the version 3.4.9 from 2003: <http://cidoc.ics.forth.gr/docs/cidoc_crm_version_3.4.9.pdf>; intended to be ISO 21127 (draft).

³⁶ Functional Requirements for Bibliographic Records: <http://www.ifla.org/VII/s13/frbr/frbr.pdf>> der International Federation of Library Associations and Institutions (IFLA).

The CRM is conceived as a tool of knowledge organization, of knowledge administration and of scientific documentation in a narrower sense. That doesn't mean to exclude a wider, non-academic public, but strives to satisfy the demands of scholarly research through precision and detailed information. Areas of application are

- the integration of information of cultural heritage institutions,
- the improvement of the understanding of this information,
- the representation of this information, e. g. in the form of storytelling.

Basically the issue is knowledge transfer from cultural heritage institutions.

Portals trying to combine such information in one interface were designed since the late nineties, when EU funding mechanisms increasingly demanded the convergence of information from the cultural heritage institutions archive, library, and museum. The requirements of the EU had a visible effect on the development of national initiatives during the last years. A few examples of such initiatives will now form the final part of this contribution. The following portals present digital heritage in the internet:

- Norway: Arkiv, Bibliotek og Museum, ABM utvikling Statens Senter;³⁷
- Denmark: Nordjyllands kulturhistoriske søgebase (NOKS);³⁸
- Netherlands: Cultuurwijzer Nederland;³⁹
- Germany: BAM, Portal für Bibliotheken, Archive, Museen.⁴⁰

Only some of these portals offer foreign language versions of the project description, while the bulk of information is only available in one language. ABM utvikling, e. g. offers brief English and German versions. All portals are currently in the stage of development and are therefore more or less restricted with regard to the range of retrievable objects, although the claim – to be accomplished in the future - is to offer a complete regional/national coverage. The Cultuurwijzer expresses it generally: *De Cultuurwijzer is een platform dat het publiek toegang geeft tot het culturele erfgoed in Nederland. Het biedt in wisselende thema's een schat aan informatie. Deze thema's geven toegang tot de onderliggende informatie.* All portals – with the exception of ABM utvikling – offer content information (text, images, sound) in addition to structured descriptions of specific items (title, author/creator etc.). The projects make an effort implementing national and international standards, e. g. NOKS: *at fremme koordinering og samordning af eksisterende standarder.* Different retrieval strategies are possible: either full-text retrieval or structured retrieval, by – sometimes faceted – classification schemes. Retrieval operations can be restricted to certain institutions or types of institutions, to types of media, or to the provenance of objects. Technical tools employed are not described in detail, so that technical solutions cannot be compared. Such technical description exists only for the Finnish museums portal *Museusuomi*.⁴¹

The main objective of the BAM project is to present the variety of cultural and scientific traditions and to develop a nation wide cultural heritage portal that presents information from archives, libraries and museums on the level of individual items (record, book, object) and of collections, in order to show the relations between and to connect the different strands of cultural knowledge. As a means to this end it is intended

- to cooperate with other relevant initiatives on a national and international level,
- to develop appropriate organizational and administrative structures for the sustainable maintenance of the service after public funding has run out,

³⁷ <<http://www.abm-utvikling.no/om/english.html>>.

³⁸ <<http://www.noks.dk/>>.

³⁹ <<http://www.cultuurwijzer.nl/asp/page.asp?alias=cultuurwijzer.nl>>.

⁴⁰ <<http://www.bam-portal.de/>>.

⁴¹ Hyvönen, Eero e.a.: A cultural community portal for publishing museum collections on the semantic web. In: <<http://www.seco.tkk.fi/publications/2004/hyvonen-saarela-et-al-a-cultural-community-2004.pdf>>; cf. also <<http://www.archimuse.com/mw2004/papers/hyvonen/hyvonen.html>>.

- to define conditions of participation, and to further develop technical functions,
- to coordinate and unify the relevant metadata,
- to support the use of metadata standards and authority files.

The BAM portal is formed by a consortium:

- Stiftung Preußischer Kulturbesitz,
- Bundesarchiv,
- Landesarchiv Baden-Württemberg,
- Landesmuseum für Technik und Arbeit Mannheim,
- Bibliotheksservice-Zentrum Baden-Württemberg.

The collections presented are (May 2006):

- the South West German union catalogue (Südwestdeutscher Bibliotheksverbund) with about 12 million titles,
- the Northern German union catalogue (Gemeinsamer Bibliotheksverbund) with about 22 million titles,
- the Prussian Cultural Heritage Foundation with the library catalogues of about 10 million titles and the online content of the image archive with about 15.000 items,
- the national union catalogue for post-medieval handwritten materials, Kalliope, with about 1 million items,
- the regional archives of Baden-Württemberg with 2500 online inventories,
- the national archives with 500 inventories,
- the Historical Museum of the city of Leipzig (Stadtgeschichtliches Museum) and four South Western State Museums (Landesmuseum für Technik und Arbeit Mannheim, Badisches Landesmuseum Karlsruhe, Staatliche Kunsthalle Karlsruhe, Württembergisches Landesmuseum Stuttgart) with a few thousand objects.

The project is financed by the German Research Council (DFG: Deutsche Forschungsgemeinschaft) in its current third project phase that will last until the middle of 2007. The main objectives of this third project phase are: to improve the technical performance, the functionalities, and the steady growth of content of all three participating types of institutions. The idea of the portal is to produce results from all three types of institutions, whatever topic you are searching for. Therefore the amount of content – and hopefully a balanced one – is crucial for satisfying this expectation of users.

The BAM portal offers to the participating institutions a common, cross-institutional access to their digital catalogues, repertories, and inventories. Scholars, students, and other users are provided with a first open and free of charge access to cultural heritage information via the internet. Institutions without an own online catalogue will be able to host their data on the BAM server. It is possible that such institutions implement a search form on their websites for the presentation of the specific digital catalogue of the institution, while the data remains within the BAM-Portal database. Supplementary functionalities for the content delivering institutions (e. g. long term preservation) and the users (e.g. e-learning, e-scholarship) are planned, but will not be realized in this phase of the project.

The cultural heritage profiles of the participating institutions will also be presented on the website. This includes information about the particulars, history, collection focus, etc. of the content delivering institutions. In this field there are close relations to the MICHAEL project.⁴² Direct access to data from a type of institution and even to data from a specific institution will be offered by the extended search form and by an guided access to institutions mentioned in the results retrieved.

⁴² <<http://www.michael-culture.org/index.html>>.

The simple Google slot or an extended search form allows for searches over digital collections of heterogeneous provenance and structure. Results are linked back to the originating catalogues of the data in question where detailed information, in as far as it is available, may be accessed. (Fig.1). Digital reproductions of the object, if these exist at all, are also available.

It has to be noted that the cataloguing of objects is done with different levels of granularity in the three participating types of institutions, which results in very different numbers of records available. Museums e. g. have to invest a lot of effort in the cataloguing of their unique resources, while libraries may even take over entire bibliographic records from other libraries. The list of results therefore is somehow unbalanced and will be so in the future. The ranking procedures have to provide for a better ranking of museums and archives, so that these get a chance to be present on the top of the list.

The search engine Lucene, software that collects local data and inserts it into a common database, is able to work effectively by using pre-converted data in formats like EAD, MAB2 and individual museum formats and mappings (Fig. 2). The usability of the CRM for these mappings will be tested during the project phase. By using authority files (see beneath) and stemming tools for the retrieval process search terms will be expanded by either synonyms or additional words that are related to the term searched. Issues of performance do no longer exist within this technical design. Distributed retrieval functionalities was tested during the first phase and can be reactivated when the now implemented procedures will run effectively.

The authority file mainly used in the BAM project is the Schlagwortnormdatei (SWD; Subject Headings Authority File of the German National Library). The SWD provides a controlled vocabulary for all areas of knowledge and will be expanded in the future with special regard to the needs of museums. The sources used for this are to be found in the *Liste der fachlichen Nachschlagewerke zu den Normdateien* (list of the reference works for authority files)⁴³. The SWD is compiled by Die Deutsche Bibliothek and six regional German library organizations.

On the one hand, the SWD traditionally is a tool of librarians. This means that archival and museum jargon is not represented in all respects. Regarding the museums the vocabulary of material and everyday culture is represented poorly. On the other hand the SWD contains a lot of everyday speech that users of the BAM portal are likely to use, because the norms of the determination of terms intend to find – as many as possible – descriptors that are part of everyday speech. Of course, the creation of the SWD is an ongoing process. There will be ample opportunities for archivists and museum specialists to add their specific vocabularies to it.

The SWD may be used in the search option of the BAM Portal (Fig.3). Applying this functionality, users will expand their query with synonyms with hierarchical top-terms being available for retrieval. But the use of authority files is not yet common in German archives and museums. A lot of the information that the BAM project deals with has no relation to the SWD. However, we have tested the automatic indexing of these data and relate the most important results (for roots, composite terms) of this process to the SWD.

The BAM project, like the other ALM projects mentioned, fits well into the plans of the EU. We are strongly interested in cooperation with others, in order to develop the idea of a common portal. First steps are made by projects in the different types of institutions

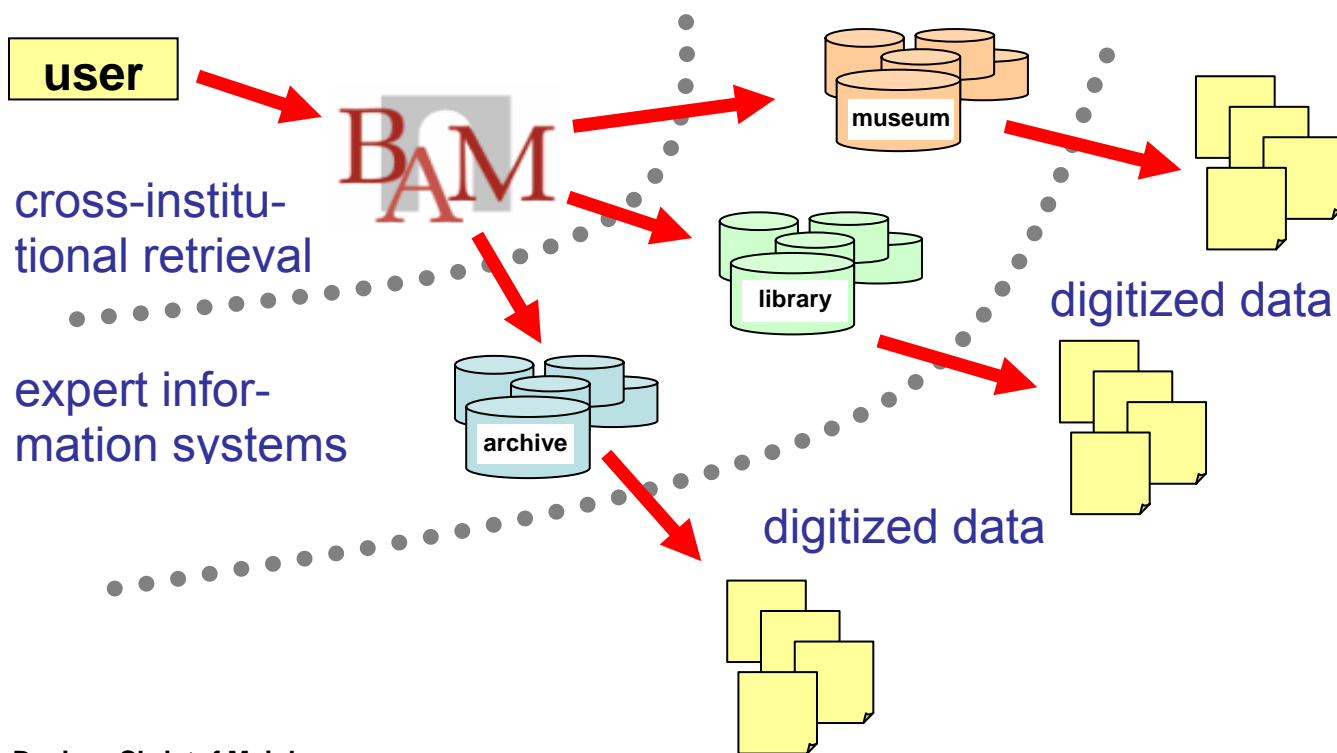
⁴³ <<http://support.ddb.de/swd/listeNSW/index.htm>>.

and in different countries.⁴⁴ Photo archives and other institutions holding cultural heritage objects present a lot of information online. Universities and educational institutions offer much of their knowledge online, too.⁴⁵ They should be integrated into a European portal presenting the diversity and the common traditions of European cultural heritage in all branches of knowledge.

⁴⁴ Beside the above mentioned projects there are many projects of the national libraries, of the national museums and archives: e. g. British Library: <<http://www.bl.uk/collections/toppage.html>>, the National Archive of Germany: <http://www.bundesarchiv.de/bestaende_findmittel/bestaendeuebersicht/index_frameset.html> or Joconde of the Ministère de la culture de France: <<http://www.culture.gouv.fr/documentation/joconde/fr/apropos/presentation-joconde.htm>>.

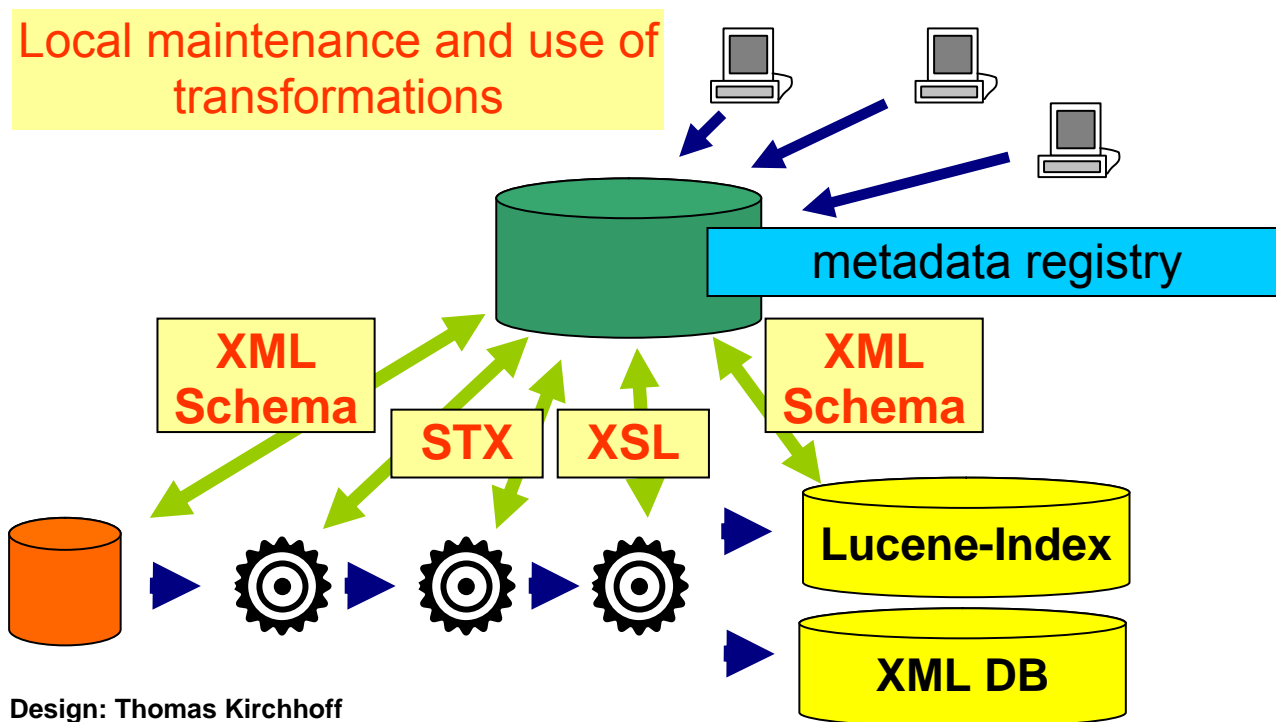
⁴⁵ E. g. Basel Mission Picture Archive: <<http://www.bmpix.org/>>, ECHO project: <<http://echo2.mpiwg-berlin.mpg.de/home>> or the portal of university collections Universeum: <<http://www.universeum.de/>>, to mention only but a few examples.

Fig. 1: Concept of information layers



Design: Christof Mainberger

Fig. 2 Formation of the index in BAM



Design: Thomas Kirchhoff

Fig. 3: Extended BAM Search

BAM Portal für
Bibliotheken Archive Museen

[Start](#) [Erweiterte Suche](#) [Projekt](#) [Teilnehmer](#) [Kontakt](#) [Hilfe](#)

Erweiterte Suche

Bibliothek Archiv Museum
SWD Synonyme Ergebnisse pro Seite: 10

Suchkriterium	Suchtext	Operator
Titel <input type="button" value="v"/>	<input type="text"/>	mit allen Wörtern <input type="button" value="v"/>
Person <input type="button" value="v"/>	<input type="text"/>	mit allen Wörtern <input type="button" value="v"/>
Schlagwort <input type="button" value="v"/>	<input type="text"/>	mit allen Wörtern <input type="button" value="v"/>

Anhang 3: Knowledge organization and multilingual vocabularies

CIDOC Meeting 2007: Managing the global diversity of cultural information
Jörn Sieglerschmidt, Bibliotheksservice-Zentrum Baden-Württemberg (BSZ)

Habet natura scientiarum thesauros innumerabiles, qui nullis aetatibus exhauriri possunt (Jean Bodin).⁴⁶

1. Introduction

Knowledge organization got its name in the context of enhancing the processing of information throughout an institution. In a networked environment knowledge organization should provide for the optimal allocation of information resources to the right person(s), at the right time(s) and places(s), in an expectable and understandable format. In the past knowledge organization was occupied with the classical topics of controlled vocabularies: classifications, thesauri, their theory, development, and usage. The topics have not changed dramatically, but the growing impact of the internet has shifted the focus somewhat to such topics as metadata standards, ontologies, semantic web etc. The politics and ethics of knowledge acquisition and distribution was and is a main topic, too. Here knowledge organization shows strong ties to the social sciences.

The expanding space of internet services has brought together the different language communities, but in most of the cases by neglecting the vernacular language of the internet user. English is most common and seldom perfectly spoken or understood by non-native speakers: so there are modern variants (beside American English e. g. continental English) and dialects (e. g. conference pidgin). Automatic translation produces funny results normally; it works only in very specific environments with a basic vocabulary of around 500 words (e. g. most sciences and applied sciences). So there is a strong need of processing information on multilingual platforms for the cultural heritage (ch) domain. There are many endeavors to meet these needs, but none is convincing until now. Therefore several possible solutions will be discussed below.

Classification is the basis of every naming activity, because giving a thing (nomina), a property (adiectiva), or an action (verba) a name makes it necessary to define the name in relation to the semantic and syntactic environment. Classification is not a basic human need, as one might be inclined to anthropologize this form of intellectual behaviour, but if men or animals use language, they must discriminate the meanings of the signs processed. Thus classification is a transcendental prerequisite of sign processing rather than an innate ability,⁴⁷ although famous scholars like e. g. Noam Chomsky have pleaded for mentalism.

Controlled vocabulary of any sort reflects the intellectual horizon of a certain time, a certain place, certain individuals. So the problem of controlled vocabulary is on the one hand the permanent change of things/thinking and the language expressing it, on the other hand the description of sometimes very specialized collections. This problem

⁴⁶ The nature of knowledge encompasses innumerable thesauri which can at no time be exhausted. Bodinus, Ioannes: *Methodus ad facilem historiarum cognitionem: accurate denuo recusus*. Lyon: Ioannes Mareschallus 1583, 310; this reference relates to the medieval and early modern debates about cosmology and the limited duration of the world: cf. Weichenhan, Michael: 'Ergo perit coelum...'. *Die Supernova des Jahres 1572 und die Überwindung der aristotelischen Kosmologie*. (Boethius 49) Stuttgart: Steiner 2004, here pp. 77-130. The French theologian and philosopher Jean Bodin developed the theory of sovereignty of the state.

⁴⁷ Batley, Sue: *Classification in theory and practice*. Oxford/New Hampshire: Chandos 2005, 1.

multiplies, of course, if multilingual vocabulary is at stake. Controlled vocabulary must be evaluated therefore according to its ability to be flexible, to accommodate new subjects, and to meet specialized and/ or local needs.⁴⁸ Furthermore: different user communities can have very different needs, especially regarding the depth of hierarchy and detail. And it should be borne in mind, that vocabulary can lose its strength of organizing knowledge, if the professional background – the artistry of concept formation – lacks.

2. Automatic indexing vs. controlled vocabulary

2.1. Automatic indexing

Big amounts of information like e. g. union catalogues, internet sites, databases are prepared for purposes of retrieval by an indexing machine. These machines read the whole text material using the usual separators to produce a list of single words. Certain words like articles and the like are removed. Almost all indexing machines can display stemming devices, i. e. linguistic procedures to reduce inflected forms of words to their stem form respectively, so that different inflected forms of the same word are not treated as different words. Nevertheless multilingual linguistic features are not available by now.

Most of the indexing machines use ranking procedures, in order to make a difference between more and less important parts of the information. The ranking procedures normally use not very sophisticated statistical measures. Nevertheless the ranking procedures produce a useful relevance index.

The big advantage of automatic indexing is the simple fact, that it can be done fast, without any intellectual endeavour (leaving aside the fact that the programming of the procedure is – hopefully – an intellectual endeavour), and for an almost unlimited amount of text. The recall of a search is therefore very high, but the precision very low. Proper names, homonyms, synonyms and the like cannot be identified. The precision can be enhanced by using and decoding metadata standards, in order to mark up certain parts of the information as more important than others. But this makes a difference only for the ranking of the results. The inclusion of controlled vocabulary during the retrieval process increases the recall by including e. g. synonyms of the searched term, but doesn't raise the precision.

2.2. Controlled vocabulary

Thesauri and classifications have a long historical tradition, if one includes encyclopedias, dictionaries and the like as similar forms of knowledge organization.⁴⁹ The following remarks will concentrate on some actual examples of thesauri and classifications with multilingual extensions.

The big advantage of controlled vocabulary: it solves the problems of synonymy, homonymy and conceptual hierarchy. Furthermore a unique identifier provides for a precise reference. So the precision of the search results is very high. The disadvantages of such controlled vocabulary are:

- high amount of intellectual work must be done,
- big institutions (e. g. national libraries) must guarantee the maintenance and development of the vocabulary,
- less flexibility regarding new branches of knowledge,

⁴⁸ *ibid.*, 94f., 107f.

⁴⁹ See the newest account on alphabetical ordering by Küster, Marc Wilhelm: *Geordnetes Weltbild. Die Tradition des alphabetischen Sortierens von der Keilschrift bis zur EDV. Eine Kulturgeschichte.* Tübingen: Niemeyer 2006.

– (normally) the payment of license fees.

The necessity of the institutional background must be stressed. The example of the Universal Decimal Classification (UDC) shows, that the lack or a weakening of institutional support endangers the whole endeavour. Since the Deutsches Institut für Normung (DIN), the German institution maintaining and developing standards, has removed from the development of the German UDC, the systematic development of the German UDC stopped.⁵⁰

Most discussions on controlled vocabulary are devoted to subject matters (what, how) leaving aside other facets of knowledge: proper names of natural persons, legal bodies (who), places (where), and times (when). In the big thesauri of national libraries or some classifications those are included. But the big vocabularies in the different areas are far more elaborated.

Place names are available as authority files in different forms, e. g. geographic or administrative, and different amount of completeness. The only authority file, that tries to combine all relevant information, is the Thesaurus of Geographic Names (TGN). The TGN is supported by the Getty Foundation. All place names included in the TGN cover administrative (in Germany administrative units down to the Landkreise, comparable to counties), geographic (at least the coordinates), in many cases historical information. The vernacular name (beside the American) and the synonyms of other languages (e. g. Firenze, Florence, Florenz) are indicated. So it seems reasonable to concentrate international efforts on the further development of the TGN, especially since some regions of the world are not represented in much detail.

There are many national authority files comprising names of persons and/ or legal bodies. But there exists no database combining the informations from different countries. Even the special area of artist's names, e. g. Union List of Artist Names (Getty Foundation) or Allgemeines Künstlerlexikon, doesn't know such efforts. The national libraries hold big authority files of persons. They exchange title information and with it informations about creators, publishers etc. These institutions should cooperate, in order to produce a common authority file.

Time specifications are very important for the search on ch sites. Nevertheless the use of clear data formats is not very common in the museum community. Only if data input deliver precise data about times and periods (of production, of an event etc.) functionalities of time search, e. g. visual helps like a time bar, could be displayed.

2.2.1. Thesauri

Thesauri are in general very flexible compared with classifications regarding the integration of new knowledge or of whole domains of knowledge. They are very rich in detail and very apt for the ch domain.

2.1.1.1. Subject headings of the national libraries

The endeavours to coordinate the subject headings of the British Library with those of the Bibliothèque National de France and of the German National Library are pending. The MACS project has ended with a fragment of the planned result. The Criss Cross project concentrates on the integration of the DDC and the authority files of the national libraries respectively. The planned multilingual tool for the description of ch objects is not in sight.

⁵⁰ The last, abridged edition based on the Classification décimale universelle of the years 1927 until 1933 was published 1967: Dezimalklassifikation. DK-Handausgabe. Internationale mittlere Ausgabe der universellen Dezimalklassifikation. 2 vol. Berlin/ Köln: Beuth 1967.

2.1.1.2. UNESCO thesaurus

The UNESCO thesaurus covers all branches of knowledge, but with about 7500 descriptors on a very abstract level. The parts of the thesaurus relevant for ch will be translated to many European languages as one of the deliverables of the MICHAEL^{plus} project. As a rough tool of orientation the UNESCO thesaurus is very useful. It should therefore be considered as one source in all efforts of constructing a multilingual vocabulary.

2.1.1.3. Eurovoc

The eurovoc thesaurus has been developed by the European Union and is available in the 21 official languages of the EU. It comprises nearly 7000 descriptors and in some languages more than 10000 non-descriptors. The topics covered are: politics, international relations, European Communities, law, economics, trade, finance, social questions, education and communications, science, business and competition, employment and working conditions, transport, environment, agriculture, forestry and fisheries, agri-foodstuffs, production, technology and research, energy, industry, geography, international organizations. As expectable the focus of the thesaurus lies on economic, technological and administrative affairs. Ch topics are much less represented than in the UNESCO thesaurus and listed under the top term *social affairs*. Nevertheless the thesaurus should be consulted for the structuring of multilingual vocabulary.

2.2.2. Classifications

There are many multilingual and monolingual classifications covering many special branches of knowledge. Here it must be sufficient to discuss two universal classifications claiming to represent all branches of knowledge.

2.2.2.1. Dewey Decimal Classification (DDC)

The Dewey Decimal Classification is the oldest classification scheme, first published in 1876, and is hosted now by Online Computer Library Center (OCLC). The DDC has been translated in many languages and is used by many library institutions worldwide, among them many national libraries like the German. Beside the knowledge organization the main aim of the DDC is an aid for shelving books. The DDC shows very strongly the cultural restrictions of its origin. Fundamental changes of the classification scheme had to be avoided because of the many users, who expected a continuity of the scheme. The DDC allows classification on a more generic level and is a bit inflexible regarding the combination of descriptors. The DDC should be consulted in every multilingual vocabulary project, but it is not very appropriate for purposes of specialized ch collections.

2.2.2.2. Universal Decimal Classification (UDC)

A younger offspring of the DDC, first published 1905-1907 and in a much extended version 1927-1933 as *Classification Décimale universelle*, is the UDC. The UDC was not conceived as shelving aid in first line, but as documentation aid. It comprises features of faceted classifications.⁵¹ The flexibility is possible by using linking signs (+, /, :) and auxiliary tables that meet the needs of the ch domain by combining the main classes and subclasses with extensions giving information about material, persons, places, nationality, form, language, time.⁵² The UDC is available in many languages, but mostly not complete. The new *master reference file* (MRF) is based on the English middle – abridged – edition, comprises some 66000 numbers, and is available in English,

⁵¹ Batley 2005 (see fn. 2), 81f.

⁵² *ibid.*, 92-101.

although English and French terms are available for parts of the MRF. Revisions of the structure of the UDC are discussed, but not yet realized for the same reasons as noted above for the DDC.⁵³ The UDC Consortium maintains and develops the UDC.

3. Ontologies and topic maps

3.1. Ontologies

Ontologies are discussed more frequently, since the semantic web claims to develop automatic devices for the linking of services in the internet. Beside topic maps ontologies are very apt to make the structure of knowledge transparent. It is not necessary to write more about the CRM here,⁵⁴ but it should be clear, that efforts to develop multilingual vocabularies should use the CRM structure as a basis. That all types of controlled vocabulary should conform to the scheme of the simple knowledge organization system (SKOS) is clear likewise.

The screenshot shows the IZEW database search interface. On the left, a topic map for 'genetic engineering' is displayed, with 'genetic engineering' at the center and various related terms like 'bioengineered', 'herbicide', 'application', 'engineering', 'food', 'engeneering', 'cloning', 'gentechnologie', 'issue', 'impact', 'agriculture', 'animal', 'science', 'plant', 'biotechnology', 'risk', 'Genetik', 'génétique', and 'resistance' connected by lines. A legend at the bottom left explains the map's elements: Associations (black square), Translations (green square), Spelling variants (orange square), and Discovery trail (blue square). The main search area shows a search for 'genetic engineering' with 154 results. The first five results are listed with their titles, authors, years, found terms, shelf codes, and locations. The right sidebar contains filters for 'Select Location', 'Refine' (Publication format, Document type, Language, Author, Subject, Thesaurus), and 'DEUTSCH'.

Fig. 1: Associative search with Aquabrowser

⁵³ cf. McIlwaine, I.C.: UDC in the twenty-first century. In: Marcella, Rita/ Maltby, Arthur (eds.): The future of classification. Aldershot: Ashgate 2000, 93-104.

⁵⁴ See my contribution to the Gothenburg meeting last year: pp. 41-49.

3.2. Topic maps

Topic maps are a very interesting device of combining possibly the intellectual work of ontologies, of controlled vocabulary, its common development – or that of folksonomies –, and the techniques of indexing. Topic maps are important as ontologies for the further development of the semantic web, because they define the structure of knowledge processed on the internet. Ontologies and topic maps will help to enhance the retrieval of relevant resources. A practical example is the search function of a German library specialized on issues of ethics in the sciences (with emphasis on life sciences):⁵⁵. On the right side one can find the facets offered, on the left side there are a topic map that can be changed dynamically by the user. After such changes a new list of literature – and a new semantic net - will appear.

4. WordNet

A wordnet serves word sense disambiguation.⁵⁶ It is derived from text corpora and dictionaries. A wordnet consists of synsets, that include one concept and its semantic space defined by the following relations: synonymy (different graphemes, same concept), homonymy (same term, different concepts), antonymy (opposite concepts), hyponymie (subterm), hyperonymie (generic term), meronymy (whole-part-relation), implication (follow-up-relation), causation (causal relationship), association (nearness of concepts). GermaNet comprises about 53000 synsets, among them 38000 nouns, 9000 verbs, and 5500 adjectives.

EuroWordNet has developed a model of the most important and frequent concepts of eight European languages (English, Spanish, Netherlands, Italian, French, German, Czech, Estonian). The interlingual index (ILI) serves as language independent component linking and coordinating the different monolingual wordnets and their synsets respectively.

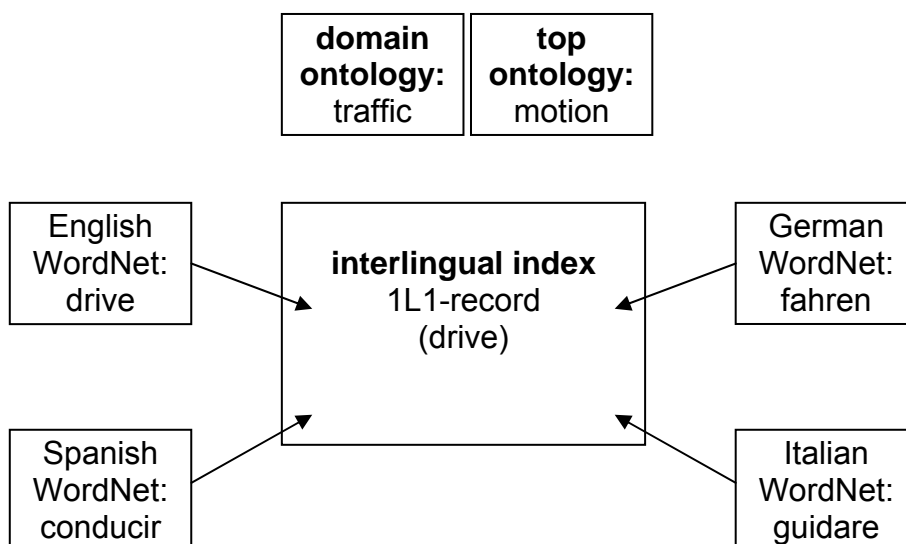


Fig. 2 EuroWordNet-architecture¹

⁵⁵ <http://www.izew.uni-tuebingen.de/lewi/assoziativ_en.html>


⁵⁶ Kunze, Claudia: Lexikalisch-semantische Wortnetze in Sprachwissenschaft und Sprachtechnologie. In: Information Wissenschaft Praxis 57 (2006), 309-314.

Wordnets seem to be very apt to model and structure multilingual vocabulary. There are no applications for the ch domain by now.

5. Folksonomies

There are many projects and efforts to build up folksonomies in the last years, since the Wikipedia has begun to develop procedures of the common and cooperative development of vocabularies. Whereas the community of documentation professionals looked down on these lay endeavors in the beginnings, the attitude has changed since, because the community of contributors to the Wikipedia has shown a remarkable skill. Furthermore the Wikipedia offers the most multilingual vocabulary platform available by now. Although the refinement of articles is very different in the different languages, every language variant offers a translation of a descriptor, perhaps much more information about the concepts linked with the descriptor, partly an elaborated taxonomy.

The following image shows the article about furniture in the Wikipedia. The German version categorizes the concept of furniture in four areas: style, material, function, construction. The newest German thesaurus of furniture is cited. The article is available in 23 languages: Afrikaans, العربية, Български, Brezhoneg, Català, Český, Dansk, Deutsch, English, Esperanto, Español, Suomi, Français, Gàidhlig, Ido, Íslenska, Italiano, Македонски, Nederlands, Polski, Português, Română, Русский, Simple English, Slovenščina, Svenska, Türkçe, Українська.



WIKIPEDIA
Die freie Enzyklopädie

Navigation

- [Hauptseite](#)
- [Über Wikipedia](#)
- [Themenportale](#)
- [Von A bis Z](#)
- [Zufälliger Artikel](#)

Mitmachen

- [Hilfe](#)
- [Autorenportal](#)
- [Letzte Änderungen](#)
- [Spenden](#)

Suche

Werkzeuge

- [Links auf diese Seite](#)
- [Änderungen an verlinkten Seiten](#)
- [Hochladen](#)
- [Spezialseiten](#)
- [Druckversion](#)
- [Permanenlink](#)
- [Artikel zitieren](#)

Andere Sprachen

- [Afrikaans](#)
- العربية
- Български
- Brezhoneg
- Català
- Český
- Dansk
- ?????
- English
- Esperanto
- Español
- Suomi
- Français


Möbel

Dieser Artikel oder Abschnitt weist folgende inhaltlich problematische Lücken auf: *Geschichtliches, bisher nur Kurzdefinition mit Listen*
 Hilf Wikipedia, indem du die fehlenden Informationen [recherchierst](#) und [einfügst!](#)


Der Begriff **Möbel** bzw. Mobiliar (sächlich) (von lat. *mobilis* = beweglich; im Gegensatz zu unbeweglichen Dingen = *Immobilien*) ist der Oberbegriff für Einrichtungsgegenstände in *Wohnungen*, Geschäften oder *Bürräumen*, aber auch im Außenbereich. Ein Möbelstück ist zweckgebunden und dient der Aufnahme von Gegenständen, dem Verrichten von Tätigkeiten, dem Sitzen oder Liegen. Die Einteilung in bestimmte Möbelgruppen ist nicht immer eindeutig und kann nach verschiedenen Kriterien erfolgen.

So können sie nach folgenden Kriterien katalogisiert werden:


- nach der **Stilrichtung**:
 - [Postmoderne](#)
 - [Bauhaus](#)
 - [Jugendstil](#)
 - [Historismus](#)
 - [Gründerzeit](#)
 - [Biedermeier](#)
 - [Barock](#)
 - [Renaissance](#)
 - [Gotik](#)
 - [Romanik](#)
- nach dem verwendeten Material:
 - [Holzmöbel](#)
 - [Plastikmöbel](#)
 - [Pappmöbel](#)
 - [Metallmöbel](#)
 - [Korbmöbel](#)
 - [Polstermöbel](#)
 - [Weichmöbel](#)
- nach ihrer Funktion:
 - [Behältnismöbel](#)
 - [Schrank](#)
 - [Wandschrank](#)
 - [Sitzmöbel](#)
 - [Arbeitsfläche \(Tisch\)](#)
 - [Kleinformöbel](#)
 - [Liegemöbel](#)
 - [Ambient-Möbel](#)
- nach ihrer Konstruktion:



Behältnismöbel Schrank
(Gesellenstück 2006, Hamburg)



Sitzmöbel Bank (Gesellenstück 2005, Hamburg)



Typische Papphocker; Deutscher Evangelischer Kirchentag Hannover 2005

It might be reasonable to cooperate with the Wikipedia in order to use the vocabulary work of so many people for the multilingual access to ch resources, because the furniture is not the only relevant example.

6. The problem of globally unique identifiers

If vocabulary will be merged, in order to build up common authority files, each term must be identified by a unique alphanumerical string. Most big controlled vocabularies have such identifiers for each term, so a matching between two e. g. place name authority files should produce a table paralleling the identifiers of the two files in two columns. If common vocabularies are an aim, there is a need for common registries providing the unambiguity of data, holding and administrating it.

7. Summary

As has been shown there are many ways to get to the same destination. Some are very long, but convincing with regard to the needs of ch domain. None of the offers is so complete, that it could be brought into practical action at this moment. So much work has to be done. Even the biggest European project in the ch domain – *EDLnet* – has found no easy way to multilingual vocabularies. Hopefully the EU funds projects that will bring practical results – as soon as possible. It might be expected that European projects like MultiMatch and STITCH are able to deliver such results.

Anhang 4: Der Objektbezeichnungsthesaurus des BSZ

Jens Lill, Bibliotheksservice-Zentrum Baden-Württemberg

Textfassung der Vorstellung des Objektbezeichnungsthesaurus des BSZ im Rahmen der Informationsveranstaltung zu Fragen der Dokumentation in Museen⁵⁷ am 17. April 2007 im Archäologischen Landesmuseum in Konstanz:

Sehr geehrte Damen und Herren, liebe Kolleginnen und Kollegen,

mein Vorredner hat Ihnen bereits anschaulich nahe gebracht, warum die Sacherschließung in Museen mithilfe von kontrolliertem Vokabular vonstatten gehen sollte. Ich werde Ihnen im Folgenden nun solch ein an unserem Hause erarbeitetes Instrument vorstellen: den Objektbezeichnungsthesaurus des BSZ.

Meinen Vortrag möchte ich dabei in folgende Teile untergliedern:

Im ersten allgemeinen Teil erläutere ich kurz, was ein Thesaurus ist und welche Probleme sich bei dessen Erstellung ergeben können. Dem schließt sich ein Blick auf die momentane Ausgangslage im deutschsprachigen Museumsbereich an, bevor im Hauptteil dezidiert auf den Objektbezeichnungsthesaurus des BSZ eingegangen wird.

Was ist ein Thesaurus?

Thesaurus klingt wie eine Figur aus der griechischen Mythologie, wie eine Mischung aus Theseus und Minotaurus. Der Minotaurus war ein Ungeheuer, halb Mensch, halb Stier, das in einem Labyrinth lebte und schließlich vom Helden Theseus erlegt wurde. Aus dem Labyrinth kam Theseus jedoch nur wieder heraus, weil ihm die kluge Ariadne zuvor ein Wollknäuel in die Hand gegeben hatte, das er beim Hineingehen abwickeln konnte und das ihm so half, später wieder den Weg heraus zu finden.

Der Begriff **Thesaurus** stammt ebenfalls aus dem Griechischen und wird im Sinne von **Wortschatz** übersetzt. Oder um es genauer zu sagen:

„Ein Thesaurus [...] ist eine geordnete Zusammenstellung von Begriffen und ihren [...] Bezeichnungen, die in einem Dokumentationsgebiet zum Indexieren, Speichern und Wiederauffinden dient.“⁵⁸

Zweck eines Thesaurus ist es, Beziehungen zwischen Begriffen und ihren Bezeichnungen herzustellen, um das Begriffsumfeld zu bestimmen und damit die einzelnen Begriffe zu vereindeutigen. Zugleich dient ein Thesaurus der terminologischen Kontrolle in einem Fachgebiet.

Oder, um in der griechischen Mythologie zu bleiben: ein Thesaurus ist der rote Faden der Ariadne, um in einem unübersichtlichen Labyrinth den Überblick zu behalten.

⁵⁷ Weitere Informationen unter <<http://titan.bsz-bw.de/cms/museen/musis/veranstaltungen/euregio-museen2007>> (eingesehen am 26. April 2007).

⁵⁸ DIN 1463, Teil 1, S. 2.

Welche Probleme treten bei der Thesauruserstellung auf?

1. Problem der Synonymie:

Ein Begriff kann durch mehrere Synonyme ausgedrückt werden (z.B. Anrichte, Buffet, Büffet, Kredenz, Speiseschrank). Eine der Aufgaben des Thesaurus ist es, zugunsten eindeutiger Begrifflichkeit zu disambiguieren. Dazu wird eines der Synonyme als Deskriptor (als Vorzugsbegriff) ausgewählt. Die anderen Synonyme des gleichen Begriffs werden als Nicht-Deskriptoren (Synonymverweisung) eingestuft.

Nur die Deskriptoren dürfen zur Indexierung der musealen Objekte verwendet werden. Die Nicht-Deskriptoren, die ebenfalls im Thesaurus vorhanden sind, verweisen die Benutzer auf den zu verwendenden Vorzugsbegriff.

Im Objektbezeichnungsthesaurus geschieht die Festlegung als Deskriptor in Anlehnung an die Schlagwortnormdatei (SWD)⁵⁹ der Deutschen Nationalbibliothek, auf die ich später nochmals gesondert eingehen werde.

2. Problem der Homonymie:

Ein Begriff kann auch mehrere Bedeutungen haben (so kann „Krebs“ sowohl für ein Tier, ein Sternzeichen, ein Sternbild oder eine Krankheit stehen).

Eine weitere Aufgabe des Thesaurus ist es, die Nachteile dieser Homonymie zu beseitigen. Jeder Deskriptor wird dazu in einen Kontext gestellt (Homonymzusatz gemäß SWD), der gewährleistet, dass sein Bedeutungsbezug eindeutig ist. Der Homonymzusatz wird dem Deskriptor als weiterer Qualifikator in spitzwinkligen Klammern angeschlossen, gemäß obigem Beispiel:

- Krebs <Zoologie>,
- Krebs <Sternbild> oder
- Krebs <Medizin>.

Wie ist die Ausgangslage im Museumsbereich?

Gerade für den Bereich der **Bezeichnung** von Museumsobjekten und damit für deren eigentliche Identifizierung gibt es keinen allgemeingültigen, universell einsetzbaren deutschsprachigen Thesaurus bzw. keine Klassifikation, die allen Bedürfnissen aller Museen umstandslos gerecht werden würde.

Gründe hierfür sind zum einen fehlende politische Gegebenheiten bzw. koordinierende Instanzen wie es sie im deutschen Bibliotheksbereich seit langem gibt. Andererseits spielt natürlich auch die Vielfältigkeit der Sammlungsbestände in den Museen (Natur, Kunst, Kultur, Technik, etc.) eine große Rolle. Hinzukommen fachliche Differenzen der verschiedenen Wissenschaftler wie ein Begriff anzusetzen ist und nicht zuletzt eine Tatsache, die ich als ***lapsus indicis*** umschreiben möchte.

Dahinter verbirgt sich der Trugschluss, dass einzigartige Objekte auch einzigartige Objektbezeichnungen haben müssten. Um dies an einem Beispiel weiter zu erläutern: Jedes Bronzemesser, das bei einer Ausgrabung gefunden wird, mag spezielle Eigenschaften besitzen, die es einzigartig machen. Dennoch können alle Bronzemesser als solches indexiert werden. Die Aussage: „Römisches Bronzemesser mit kleiner Klinge und Vollgriff“ ist keine Objektbezeichnung, sondern ein halber Roman! Um die

⁵⁹ Schlagwortnormdatei <<http://www.d-nb.de/standardisierung/normdateien/swd.htm>> (eingesehen am 26. 04. 2007).

Einzigartigkeit von Objekten zu beschreiben, gibt es eine Vielzahl anderer Felder in einem Datenbanksystem.

All diese Gründe haben dazu geführt, dass in vielen Institutionen Wortlisten in Eigenregie erstellt wurden und die Anwendung von Normdaten eher zu den Ausnahmen gehört.

Die Dokumentation von Objekten in Museen findet bisher also meist in frei verwendeter Sprache statt, die oft mehrdeutig ist. Daraus resultierend ergibt sich, dass ein identisches Objekt von Museum zu Museum verschieden benannt sein kann – oder gar hausintern, aufgrund mehrerer Personen, die für die Indexierung zuständig sind. Dies führt in einem regionalen Museumsverbund (oder auch global gesehen hinsichtlich des Aufkommens musealer Fachportale im Internet) zu einer verminderten Qualität im Retrieval.

Eine nachträgliche Vereinheitlichung bei der Sacherschließung von Museen ist deshalb ein zwar schwieriges, aber wichtiges Unterfangen.

Der Objektbezeichnungsthesaurus des BSZ

Der Objektbezeichnungsthesaurus des BSZ basiert auf dem Erschließungsvokabular des **Badischen Landesmuseums Karlsruhe** und der **Reiss-Engelhorn-Museen Mannheim** und spiegelt somit bisher primär das Fachvokabular von kunst- und kulturhistorischen Sammlungen wider. Eine erste Begriffserweiterung im Rahmen von Datenmigrationen ist bereits durch das Landesmuseum Württemberg in Stuttgart sowie die Stiftung Schloss Friedenstein in Gotha erfolgt.

Durch den Einsatz im MusIS-Verbund⁶⁰ und die Zusammenarbeit interessierter Institutionen wird der Thesaurus kontinuierlich um weitere Begriffe ergänzt. So steht in diesem Jahr noch eine Erweiterung um archäologische Termini durch die Stiftung Schloss Friedenstein in Gotha in Zusammenarbeit mit dem dort zuständigen Landesamt für Archäologie in Weimar an.

Um die Erschließungstiefe möglichst allgemein zu halten und um den Thesaurus möglichst museumsübergreifend anbieten zu können, ist an eine Einbindung und Angleichung von vertiefenden Spezialthesauri gedacht (z.B. der bereits existierende Uhrenthesaurus des BSZ), die zusammen mit Spezialsammlungen/-museen entwickelt werden sollen.

Was sind die Ziele des Thesaurus?

Vorrangiges Ziel ist die Verbesserung der Dokumentation von Museumsobjekten und infolgedessen ein qualitativvolleres Retrieval – sowohl institutionell wie auch museumsübergreifend.

Aus diesem Grund soll der Objektbezeichnungsthesaurus einen durchgängigen Bezug zur SWD (Schlagwortnormdatei) erhalten, einem Vokabular, das von der **Deutschen Nationalbibliothek** weiterentwickelt wird, auch in Zusammenarbeit mit dem BSZ und den betreuten Museen.

⁶⁰ MusIS – Landeseinheitliches MuseumsInformationsSystem <<http://titan.bsz-bw.de/cms/museen/musis/>> (eingesehen am 26.04.2007).

Die Schlagwortnormdatei (SWD)

Die SWD ist eine seit 1988 kooperativ geführte Normdatei und versteht sich selbst als Thesaurus nach der DIN 1463⁶¹.

Neben Sachschlagwörtern aller Art und Fachrichtungen enthält sie auch Personen, Körperschaften und Geographika, die für die Sacherschließung in den Bibliotheken genutzt werden. Derzeit befinden sich in diesem terminologisch kontrollierten Wortschatz ca. 750.000 Datensätze⁶² aus allen Wissensgebieten und verzeichnet neben den Ansetzungs- und Verweisungsformen auch zusätzliche strukturierende Elemente. Dazu zählen:

- Die Systematisierung der Begriffe nach einer Grobklassifikation (SWD-Notation) bzw. mit Ländercodes.
- Quellen der Ansetzung (Referenzwerke), Definitionen, Verwendungshinweise, redaktionelle Bemerkungen.

Die SWD hat im gesamten deutschsprachigen Bibliothekswesen normierende Autorität; die Angleichung an internationale Standards ist in verschiedenen Projekten bereits angedacht.

Das BSZ als Betreiber des Südwestdeutschen Bibliotheksverbunds (SWB) fungiert mittlerweile als Fachredaktion für die im MusIS-Verbund betreuten Museen. Seit März 2007 werden Sachschlagwörter aus dem Museumsbereich, die Eingang in den Objektbezeichnungsthesaurus gefunden haben, in der SWD erfasst.

Zeitaufwand für Thesauruserstellung

Der zeitliche Aufwand für die Thesauruserstellung ist nicht zu unterschätzen.

So wurde das Grundgerüst des Objektbezeichnungsthesaurus in einer ersten Phase (Herbst 2005 bis Sommer 2006) im Rahmen eines Werkvertrages⁶³ aus den beiden Basisvokabularen erstellt. In einer zweiten Phase im zweiten Halbjahr 2006 erfolgte dann eine hierarchische Überarbeitung mit Angleichung an die SWD am BSZ. Die Erfassung von noch nicht in der SWD enthaltenen Termini, einhergehend mit einer weiteren Durchsicht der Thesaurushierarchie, wird voraussichtlich Ende 2007 abgeschlossen sein.

Nichtsdestotrotz ist der Objektbezeichnungsthesaurus im MusIS-Verbund schon vereinzelt im Einsatz und erfährt durch Datenmigrationen stetigen begrifflichen Zuwachs.

Wie sieht nun der strukturelle Aufbau des Thesaurus aus?

Der Objektbezeichnungsthesaurus ist durch semantische Relationen strukturiert, die ich Ihnen kurz erläutern möchte:

⁶¹ „Die SWD entspricht damit den Anforderungen, die nach DIN 1463 an einen Thesaurus gestellt werden [...]“. Die Deutsche Bibliothek, in Zusammenarbeit mit dem Bibliotheksverbund Bayern, dem Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen, dem Südwestdeutschen Bibliotheksverbund, dem Verbund der wissenschaftlichen Bibliotheken Österreichs, dem Kunstverbund Florenz, München, Rom, Schweizerische Landesbibliothek Bern und Kooperativer Bibliotheksverbund Berlin-Brandenburg (Hrsg.), Schlagwortnormdatei (SWD), Ausgabe Oktober 2000, Frankfurt a.M., 2000, S. 4.

⁶² Stand: April 2007, Quelle: Bibliotheksservice-Zentrum Baden-Württemberg (BSZ).

⁶³ Kontakt zu Frau Katrin Teichmann, freiberufliche Museologin und wissenschaftliche Dokumentarin: <<http://kultur-gut-bewahren.de>> (eingesehen am 26.04.2007).

1. Äquivalenzrelation:

Die Äquivalenzrelation zwischen Deskriptoren und Nicht-Deskriptoren wird durch den Synonymverweis `Benutze für` gekennzeichnet.

Eine reziproke Verweisung von Nicht-Deskriptor zu Deskriptor in Form von `Zu benutzen` existiert nicht, da die Nicht-Deskriptoren in unserer Dokumentationssoftware (IMDAS-Pro)⁶⁴ nur als Verweise in der Begriffstabelle der Datenbank hinterlegt und nicht als eigentliche Begriffe in den Thesaurus eingearbeitet sind.

Beispiel: Kalebasse, `Benutze für`: Kürbisflasche

2. hierarchische Relation:

Es werden folgende Hierarchierelationen zwischen Deskriptoren unterschieden, die reziprok gelten:

Übergeordneter Begriff: Verweis von einem spezifischen auf einen generischen Term,

Beispiel: Trompete, übergeordneter Begriff: Blechblasinstrument

Untergeordneter Begriff: Verweis von einem generischen auf einen spezifischen Term.

Beispiel: Blechblasinstrument, untergeordneter Begriff: Trompete

Deskriptoren ohne übergeordnete Begriffe sind Leitdeskriptoren (sog. Top Terms).

3. Assoziationsrelation:

Die Assoziationsrelation zwischen Deskriptoren wird durch den Verweis `verwandter Begriff` angegeben.

Beispiel: Sonnenuhr, `verwandter Begriff`: Gnomon <Astronomie>

Daneben ist auf folgende wichtigen Eigenschaften der assoziativen Relation hinzuweisen:

- Sie ist immer wechselseitig.
- Sie ist mit der hierarchischen Relation unvereinbar: sind zwei Deskriptoren durch eine hierarchische Relation miteinander verknüpft, so können sie nicht durch eine assoziative Relation verbunden werden, und umgekehrt.
- Die Deskriptoren unter ein und demselben "Top Term" können nicht durch eine Assoziationsrelation verknüpft werden.

Darüber hinaus gibt es im Objektbezeichnungsthesaurus noch verschiedene Anwendungserläuterungen wie:

⁶⁴ IMDAS-Pro - Integrated Museum Documentation and Administration Programme der Firma Joanneum Research Graz, <<http://www.joanneum.at/index.php?id=353&L=0>> (eingesehen am 26.04.2007).

1. Stützterm

Bei einem Stützterm handelt sich um keinen Deskriptor im eigentlichen Sinn, sondern um eine Hilfskonstruktion, die folglich für die Indexierung von Museumsobjekten nicht ausgewählt werden sollte.

Stützterme sind im Objektbezeichnungsthesaurus als Ganzes in spitzwinkligen Klammern gesetzt und dienen als weiteres formales Kriterium zur Begriffseinordnung.

Beispiel: <Uhr nach Antriebstechnik> oder <Uhr nach Teil>

2. Verweis: keine Objektbezeichnung

Andere Begriffe, die durch ein vorangestelltes Sternchen gekennzeichnet sind (z.B. *Freizeitgestaltung), sind mit dem Verweis `Begriffstyp: keine Objektbezeichnung` belegt und für die Verwendung als Objektbezeichnung ebenfalls nicht empfohlen, da sie zu abstrakt sind.

Auch diese Begriffe haben nur zwecks der hierarchischen Einordnung anderer Termini Eingang in den Thesaurus gefunden.

3. Zusätzliche Begriffsinformationen

- Begriffsdefinition: möglichst kurze, präzise Bestimmung des Begriffsinhalts
- Erläuterung: die Erläuterung enthält über die Begriffsdefinition hinausgehende Informationen zum Anwendungsbereich des Begriffes
- Quelle: der Begriffsansetzung bzw. der Begriffsdefinition
- Begriffsvorschlag: von welchem Museum ging der Begriffsvorschlag aus
- Redaktionelle Bemerkung: enthält zusätzliche Informationen des BSZ bei z.B. abweichender Begriffseinordnung in anderen Klassifikationen
- SWD-Identnummer: Identnummer des Normdatensatzes der Deutschen Nationalbibliothek
- SWD-Systematiknummer: Notationen der SWD-Sachgruppen

Quantitative Merkmale des Thesaurus⁶⁵

Der Objektbezeichnungsthesaurus enthält aktuell knapp 2.200 Deskriptoren, verteilt auf 44 Top Terms. Insgesamt sind im Objektbezeichnungsthesaurus 37 Stützterme enthalten sowie weitere 37 vom Begriffstyp „keine Objektbezeichnung“. 128 der Deskriptoren sind mit einem Homonymzusatz versehen. Daneben kommen 1.381 Äquivalenzrelationen (Synonyme) und 94 Assoziationsrelationen (verwandte Begriffe) vor. 522 der enthaltenen Deskriptoren sind noch mit keiner SWD-Identnummer versehen und müssen noch in der SWD erfasst werden.

Wie ist die Vorgehensweise bei der Einbindung neuer Termini?

- Begriffsvorschlag erfolgt durch das Museum (optimal wäre mit evt. Synonyme, Begriffsdefinition, Quelle der Definition)
- Abgleich mit SWD erfolgt am BSZ, ob Begriff oder Synonym bereits vorhanden

⁶⁵ Aktualisierter Stand vom Februar 2008.

- → falls dies der Fall: wird der Begriff gemäß der SWD in den Thesaurus aufgenommen
- → falls dies nicht der Fall: Untersuchung des Sprachgebrauchs in gedruckten oder Online-Nachschlagewerken wie Brockhaus, Duden, Wikipedia oder Fachliteratur

- Verortung in Thesaurushierarchie

Als zweiter Schritt erfolgt sodann die Erfassung in der SWD nach den „Regeln für den Schlagwortkatalog“ (RSWK). Mitunter führen die RSWK dazu, dass Termini, v.a. Komposita wie z.B. „Zündholzschachtelhülle“ nicht in die SWD aufgenommen werden dürfen, sondern in Einzelschlagwörtern kombiniert werden müssten, also: „Zündholzschachtel / Hülle“ als Schlagwortkette.

In diesen Fällen verzichtet die Thesaurusredaktion auf eine Referenzierung zur SWD, so dass der Begriff ohne SWD-ID in den Objektbezeichnungsthesaurus eingebunden wird.

Um in Zukunft für alle Termini im Thesaurus eine eindeutige Referenz zu haben (also sowohl für diejenigen mit SWD-ID als auch ohne), ist unsererseits die Einführung einer globalen Begriffs-ID geplant, so dass auch Veränderungen von externer Seite aus nachvollzogen werden können.

Welche Vorteile bietet also die SWD?

- Terminologisch kontrollierter Wortschatz, der ein fächerübergreifendes Vokabular bietet,
- Thesaurus nach DIN 1463 → umgeht das Problem der Homonyme, Polyseme und Synonyme,
- Sie besitzt bereits normierende Autorität im deutschsprachigen Raum und ermöglicht somit die eindeutige Referenzierung der Termini,
- Eine Angleichung an internationale Standards ist in Projekten bereits angedacht → Stichwort: Multilingualität von Thesauri,

Das BSZ fungiert bereits als SWD-Fachredaktion für die Teilnehmermuseen im MusIS-Verbund und versteht sich überdies als Koordinator einer kooperativen Weiterentwicklung kontrollierten Vokabulars für die Museen in Deutschland. Die dahingehenden Bemühungen der Fachgruppe Dokumentation im Deutschen Museumsbund⁶⁶ werden von unserer Seite aus unterstützt.

Zusammenfassung der wichtigsten Punkte:

- der Objektbezeichnungsthesaurus steht allen Museen/Sammlungen offen,
- allgemein gehalten → eine vertiefende Erschließung ergibt sich aus der Anbindung von Spezialthesauri,
- SWD-gestützt (mit all den zuvor genannten Vorteilen),
- sowohl monohierarchische als auch polyhierarchische Version,

⁶⁶ Deutscher Museumsbund <<http://www.museumsbund.de/cms/index.php>> eingesehen am 26.04.2007
→ Fachgruppen → Dokumentation.

- der Thesaurus wird in Zusammenarbeit mit den Museen weiterentwickelt
→ in Spezialsammlungen vor Ort sollen nach Möglichkeit Fachredaktionen gebildet werden,
→ im Rahmen dieser Arbeiten sollte auch eine Liste der maßgeblichen fachlichen Referenzwerke erstellt werden, die die „Liste der fachlichen Nachschlagewerke zu den Normdateien“ der DNB⁶⁷ ergänzen muss.

Zu welchen Konditionen ist der Objektbezeichnungsthesaurus zugänglich?

Für Bildungs- und nicht kommerzielle Zwecke wird der Thesaurus frei verfügbar sein! Versehen mit einer CreativeCommons-Lizenz⁶⁸, die diese freie Nutzung genauer definiert, wird der Thesaurus auf der Online-Plattform „museumsvokabular.de“⁶⁹ veröffentlicht, wo Sie auch andere frei verfügbare Vokabulare finden können.

Ausblick

Folgende Arbeiten werden von uns zeitnah erledigt:

- Erfassung noch ausstehender Termini in der SWD (on demand); parallel dazu die Einarbeitung neuer Begriffe durch Datenmigration bzw. Meldung an die Thesaurusredaktion
- Veröffentlichung des Thesaurus auf **museumsvokabular.de** mit einer **Creative-Commons-Lizenz**, die die weitere Nutzung regelt
- Einrichten von Fachredaktionen in Spezialsammlungen vor Ort; Etablieren eines Redaktionsverfahrens für externe Museen
- Mailingliste oder Newsletter als Informationsinstrument einrichten

Und natürlich würden wir uns über die Nutzung und Mitarbeit bei der Weiterentwicklung des Objektbezeichnungsthesaurus des BSZ Ihrerseits freuen!

Kontakt:

Jens M. Lill
Bibliotheksservice-Zentrum Baden-Württemberg
jens.lill (at) bsz-bw.de
Tel.: +49 (7531) 88-4932

Thesaurusredaktion: thesaurus (at) bsz-bw.de

URL: <http://titan.bsz-bw.de/cms/museen/musis/>

URL des vorliegenden Dokuments:

http://titan.bsz-bw.de/cms/museen/musis/veranstaltungen/Lill_textfassung-OBZ.pdf

⁶⁷ Liste der fachlichen Nachschlagewerke <<http://support.ddb.de/swd/listeNSW/index.htm>> (eingesehen am 26.04.2007).

⁶⁸ Creative Commons Deutschland <http://de.creativecommons.org/cc_erlauterung.html> (eingesehen am 26.04.2007).

⁶⁹ museumsvokabular.de <<http://www.museumsvokabular.de>> (eingesehen am 26.04.2007).