

Digitale Objektdokumentation und Online-Publikation - Verändern die Informationstechnologien das Sammlungsmanagement der Museen?

Jörn Sieglerschmidt, Bibliotheksservice-Zentrum Baden-Württemberg Konstanz

Zusammenfassung

In zunehmendem, vermutlich exponentiell wachsendem Maße werden Inhalte aus dem gesamten Kulturbereich im Netz verfügbar sein. Stellvertretend für andere Projekte sei das von *Google* in Zusammenarbeit mit namhaften Universitäten z. B. der *ivy league* geplante genannt, in dem es um die Digitalisierung der gesamten Buch- und Handschriftenbestände dieser Universitäten geht. Solche Projekte, aus denen offensichtlich erheblicher geldwerter Nutzen erwartet wird, werfen neben den (kultur)politisch-rechtlichen zahlreiche weitere Fragen auf.

- Der Einsatz von technischen Verfahren hat bisher und wird in Zukunft weitere Anstrengungen hinsichtlich der Normierung und Normung der Verfahren erfordern, um anschluss- und kommunikationsfähig zu sein – und zu bleiben. Die Hoffnung, daß intelligente und schnelle Indexierungsverfahren diesen Aufwand erübrigen könnten, hat sich bisher nicht erfüllt. Vielmehr wird es um eine Mischung aus Einsatz wie Entwicklung von Normvokabularien und Indexierung gehen.
- Den Museen stellt sich die Frage, was und wie digitalisiert wird. Das Was öffnet das Problemfeld der Auswahl aus dem Fundus der Kulturgüter, da augenblicklich niemand annimmt, daß eine vollständige Digitalisierung sinnvoll bzw. möglich ist. Hinter dem Wie versteckt sich nicht nur das Problem der Publikationsfähigkeit der Digitalisate, sondern auch der Erhaltung einer Wissensstruktur, die die Vernetzung individueller Objektinformation mit den anderen Wissenswelten nicht zerschneidet, sondern aufbewahrt. Vorhandene Angebote im Netz sollen in Hinsicht auf diese Fragestellung geprüft werden.
- Die Vernetzung und ubiquitäre Verfügbarkeit der Wissensbestände wird zu einer grundsätzlichen Wandlung des wissenschaftlichen Arbeitens, aber auch der Informationsbeschaffung insgesamt führen. Ob gänzlich neue wissenschaftliche Einsichten in den Kulturwissenschaften möglich sind, wie manche hoffen, muß angesichts der erst beginnenden Entwicklung unbeantwortet bleiben. Eines scheint allerdings sicher: Bilder – bewegt und unbewegt – werden eine zunehmende und wichtigere Rolle als bisher spielen.

Technisch, organisatorisch und finanziell stehen die Museen in den nächsten Jahren vor erheblichen Herausforderungen, wenn es darum geht, Optionen auf die Zukunft im Rahmen der internationalen Entwicklung zu wahren.¹

¹ *online* unter: <http://titan.bsz-bw.de/cms/service/museen/publ/>

Einleitung

Die Welt des Wissens, es wird seit einiger Zeit repetiert, verändert sich seit Ende des letzten Jahrhunderts und mit zunehmender Geschwindigkeit durch die Verfügbarkeit von Wissen im Internet. Wenn es um die vielen, nicht selten schlechten und als solche zuweilen wieder beachtenswerten Internetseiten geht,² so ist das Internet sicherlich kein Fortschritt, denn technische Mittel sind immer dafür gut gewesen, menschliche Dummheit zu vervielfältigen und noch wirksamer zu verbreiten. Es geht also um die anderen Angebote, die nicht nur den Anspruch der Seriosität erheben, sondern von einer Mehrheit der Internetnutzer/innen für seriös gehalten werden: im wesentlichen die institutionell gesicherten, traditionellen Agenturen der Wissensaufbewahrung und -vermittlung wie Hochschulen, Bibliotheken, Archive, Museen und andere von den Steuerzahlern alimentierte, Sinn machende Anstalten.

Bisher findet zwischen den vielfältigen und inzwischen beeindruckenden Informationsangeboten keine Verknüpfung in der Weise statt, daß Nutzer/innen mit wenigen Anfragen zu verlässlichen Informationen gelangen. Noch immer müssen zahlreiche Seiten einzeln befragt werden. Seit die Portalitis die Cyberwelt befallen hat, scheint ein Weg eingeschlagen zu sein, der Besserung im Sinne der Nutzer/innen verspricht. Es muß allerdings abgewartet werden ob die Heilsversprechen des *semantic web* tatsächlich eintreten oder sich im Nirwana der Cyberwelt verlieren werden.³ Dabei geht es um eine intelligente Verknüpfung von Recherchen und Rechenschritten über Datensammlungen (digitale Objekte z. B. aus Bibliotheken, Archiven und Museen), die jetzt *online* verfügbar sind. Darin läge ein erheblicher Gewinn, der aber voraussetzungsvoll ist. Zu diesen Voraussetzungen gehören:

- eine weitere Entwicklung der Indexierungs- und Suchmaschinen;
- die ständige Pflege und Verbesserung von
 - syntaktischen (Metadaten, Ontologien) und
 - graphisch-semantischen (kontrolliertes Vokabular für die Sacherschließung)Normen, die allerdings die Komplexität der Inhalte in ihrer vernetzten Struktur möglichst weitgehend abbilden sollten;
- die Beschreibung von akustischen und audiovisuellen Internetressourcen in recherchierbarer Textform;
- die weitere Internationalisierung der Normvokabularien, um durch Multilingualität der globalen Verarmung von regionaler Sprache und Kultur entgegenzuarbeiten.

Ziel ist eine flexible Wissensnutzung, die die unendliche Verknüpfung von Inhalten zuläßt und der kreativen Nutzung damit keine Grenzen setzt, also das Wissen ausschöpft in seinen unbegrenzten Möglichkeiten der Kombination und Differenzierung.

Verknüpfung setzt erstens Verständlichkeit voraus. Verständlich heisst auf der Ebene der mündlichen oder schriftlichen Äußerung, daß ich mich an Ausspracheregeln bzw. Schreibkonventionen halte (Ebene der Graphie). Ohne solche kann ein Gegenüber überhaupt nicht erkennen, was ich sagen oder schreiben will. Außerdem muß ich mich an bestimmte Ordnungsregeln in der Anordnung meiner Grapheme bzw. Äußerungen halten, d. h. eine Syntax einhalten, die nach konventionellen Vostellungen verstehbare Äußerungen hervorbringt.

² <<http://muellseite.de/>>; <<http://www.webpagethatsuck.com/>>.

³ <<http://www.w3.org/2001/sw/>>; <<http://www.w3.org/2001/sw/Europe/>>.

Allerdings ergibt die Syntax noch keine Eindeutigkeit bei der Dekodierung einer Information. Verknüpfung setzt daher zweitens Eindeutigkeit voraus. Eindeutigkeit kann durch entsprechende Referenzierung hergestellt werden, d. h. durch die Individualisierung bzw. Präzisierung eines Begriffes mittels Verortung in einem semantischen Netz. Bei Personennamen kann das durch Lebensdaten, Geburts- oder Sterbeorte, Beziehungen zu anderen Personen oder Personengruppen, geschaffene Werke, ausgeübte Berufe, erhaltene Titel usw. geschehen. Bei Orten kann das durch geographische (Montblanc als Teil der Alpen) oder administrativ-politische Zuordnung (Kiel als Landeshauptstadt) geschehen. Bei Sachen könnten wie in entsprechenden Vokabularien und Wörterbüchern vergleichbare, übergeordnete, untergeordnete, synonyme usw. Begriffe angeführt werden, um den individuellen Begriff in seinen semantischen Bezügen eindeutiger zu machen. Daß dieser letzteren Vereindeutigung Grenzen gesetzt sind, wird jedem klar, der sich mit Begriffsuniversen und deren Verknüpfung beschäftigt.

Wissensrepräsentation zeigte sich in der klassischen Dokumentation: als flache Datenstruktur mit wenigen Verknüpfungen, die meistens als solche auch nicht einzeln zugänglich bzw. recherchierbar waren. Lediglich in begründeten Ausnahmefällen (Werkverzeichnisse einzelner Künstler, Orte als Schauplätze bestimmter Ereignisse usw.) wurden solche Verknüpfungen systematisch hergestellt. In einer vernetzten Umgebung wird sich das zunehmend ändern. So erlauben die seit langem verfügbaren Recherchetechniken komplexe Suchen unter Einschluß boolescher Operatoren und Quantoren (größer/kleiner als usw.). Darüberhinaus wären aber Hinweise und Verknüpfungen mit komplexen Bedeutungsräumen erwünscht. Auf die damit zusammenhängenden Probleme der Abbildung solcher Bedeutungsräume durch die Sacherschließung wird gleich eingegangen.

Zunächst soll die Abbildung komplexer Objekt- und Begriffsbeziehungen behandelt werden. Metadatenformate ebenso wie die in Bibliotheken genutzten Formate RAK und MAB sind flache Datenstrukturen,⁴ d. h. sie beschreiben parallel und gleichrangig angeordnete Informationen zu einem Objekt, deren Zusammenhang sich nicht aus dem Datenmodell ergibt, sondern möglicherweise – wenn überhaupt – erst durch komplexe Rechenschritte. Dabei ist die verarbeitete Information meist nicht komplex, d. h. das Datenmodell umfasst nicht biographische Angaben zu Personen, historische Angaben zu Körperschaften oder Orten. Die in der realen Welt häufig vorhandenen komplexen Informationen können daher garnicht abgebildet werden. Einige für den Museumsbereich entwickelte Datenbanksysteme können inzwischen einen Teil dieser komplexeren Strukturen abbilden. Sinnvoll wäre, abgesehen von anderen guten Gründen, auch unter dem Gesichtspunkt des Datenaustausches ein gemeinsames Referenzmodell, das nicht nur für den Museumsbereich tauglich ist, sondern für Wissensobjekte jeglicher Art. Das *conceptual reference model* (CRM) will ein solches Modell sein, das Verständlichkeit, Eindeutigkeit und syntaktisch – nicht semantisch – beliebige Verknüpfbarkeit von Informationen bietet und damit der Tatsache menschlicher Kreativität Rechnung trägt, d. h. nicht Wege der Verknüpfung abschneidet, die vorhanden bzw. im Datenangebot enthalten sind. Es wird die Aufgabe der nächsten Jahre sein, diese als ISO-Norm angemeldete Ontologie in den

⁴ Zu MAB und RAK: <http://www.ddb.de/index_txt.htm>; gängige Metadatenformate: Dublin Core Metadata Element Set (DCMES, <<http://dublincore.org/>>), Metadata Encoding and Transmission Standard (METS, <<http://www.loc.gov/standards/mets/>>), Encoded Archival Description (EAD, <<http://www.loc.gov/ead/>>), Text Encoding Initiative (TEI, <<http://www.tei-c.org/>>).

Museumsalltag zu übersetzen.⁵ Das BAM-Projekt wird dazu erste Schritte unternehmen und auch fuer die anderen Dokumentationsbereiche Archiv und Bibliothek zeigen, daß das CRM ein angemessenes Werkzeug und Bezugsmodell bei immer gegebenen heterogenen Datenstrukturen darstellt.⁶

Selbst bei dem in manchen Bereichen wie z. B. den Bibliotheken üblichen Vorliegen digitaler Daten sind die oben bezeichneten komplexen Recherchewerkzeuge nicht bzw. unzureichend in der Lage, z. B. hierarchische Beziehungen zwischen Begriffen abzubilden. Nur in entsprechenden Datenbankumgebungen ist es möglich poly- und monohierarchische Beziehungen von Begriffen so anzulegen, daß ich Recherchen an jedem beliebigen Punkt der Hierarchie beginnen und verfeinern oder erweitern kann. Diese Hierarchien müssen von Mitarbeiter/innen der Museumsdokumentation aber durch intellektuelle Arbeit hergestellt werden, da die SWD oder geographische Thesauri zwar Anhaltspunkte geben, aber technisch nicht die Hierarchie anbieten. Im Bibliotheksbereich bekannte bzw. verbreitete Klassifikationen wie UDC oder DDC sind für den Museumsbereich bisher nicht tauglich.

So muß ich bei einer Beschlagwortung mit der SWD verwandte Begriffe, Oberbegriffe usw. meist in den Datensatz zugleich eintragen, damit ich diese recherchieren kann. Zwar wird der jeweils zugehörige Oberbegriff im Schlagwortsatz miterfaßt, aber nur in wenigen Lokalsystemen auch angezeigt, von komplexeren Begriffsbeziehungen einmal ganz abgesehen. Thesaurussoftware bietet heute die erweiterten Möglichkeiten einer komplexen Verknüpfung der begrifflichen Welt, wenn auch die Werkzeuge für die Entwicklung semantischer Topographien (*semantic maps*) noch nicht wirklich eingesetzt werden.⁷ Soweit ich mein Wissen in der o. g. genannten Weise ablege und eindeutig referenziere, erzeuge ich erheblich komplexere Wissenräume mit potentiell unendlich vielen Verknüpfungsmöglichkeiten. Voraussetzung dafür ist die Arbeit mit institutionell betreuten Vokabularen, die mir die Beschreibungsinformation liefern. Daraus wird deutlich, dass erstens spezielle Vokabularen aufgebaut und gepflegt werden, zweitens diese in der *online*-Dokumentation genutzt werden müssen. Daß diese Forderung nicht theoretisch ist, sondern in manchen Bereichen bereits zum Alltag gehört, zeigt folgende Äußerung zur Digital Library: "The key aspect for the Digital Library community is the provision of *shared catalogues* which can be published and browsed. This requires the use of common *metadata* to describe the fields of the catalogue, (such as author, title, date, publisher); and common *controlled vocabulary* to allow subject indentifiers to be assigned to publications."⁸

Nun ist bereits häufig eingewandt worden, daß Indexierungsverfahren heute so ausgereift seien, daß der erhebliche intellektuelle Aufwand für die Erstellung von Normvokabularen sich erübrige. Alle empirischen Ergebnisse deuten darauf hin, daß die Indexierung bisher und wohl auf lange Sicht nicht dieselbe Treffsicherheit erreichen kann wie sie mit dem Einsatz von Normvokabularen möglich ist. Insbesondere fehlt es an der Verortung von Begriffen im weiteren Begriffsraum. Vielversprechend erscheint daher ein Weg, der ebenfalls im BAM-Projekt erprobt werden soll: die Verbindung

⁵ Im Moment noch: ISO/CDraft 21127; <<http://cidoc.ics.forth.gr/>>. Es gibt im übrigen mappings von den o. g. Metadatenformaten auf das CRM.

⁶ <<http://www.bam-portal.de/>>.

⁷ <http://www.fbi.fh-koeln.de/institut/labor/Bir/thesauri_new/> gibt hervorragende Auskunft zu verfügbaren Thesauri und Thesaurussoftware; darauf stützt sich auch Leonard Will auf seiner Seite: <<http://www.willpower.demon.co.uk/>>; zu semantic maps vgl. Asmus, Stefan/Stephan, Peter Friedrich: Wissensdesign. Mit neuen Medien Wissen gestalten. <<http://www.wissensdesign.de/>>.

⁸ <http://www.w3.org/2001/sw/Europe/showcase/digital-libraries.html>

beider Verfahren mit dem Ziel die Vorteile derselben zu kumulieren. Indexierungsverfahren sind interessant, weil sie schnell und unaufwendig, inzwischen auch linguistisch bearbeitete Begriffslisten liefern. Die erforderliche Ergänzung des umfangreicheren Begriffsraumes wird durch einen Abgleich mit dem Normvokabular und der Erweiterung der indexierten Begriffe um die dort verfügbaren Begriffsinformationen erreicht. Es ist offen, ob das Verfahren praktisch die Ergebnisse liefert, die theoretisch erwartet werden.

Wissen bleibt - soweit jetzt beschrieben - enzyklopädisch auf einer Beschreibungsebene, die versucht, eindeutige Referenzen zu bieten, um gesagte bzw. geschriebene Dinge und Sachverhalte voneinander unterscheiden zu können. Wer dagegen kausale Zusammenhänge sucht, begibt sich auf die Ebene historisch-literarischen oder naturwissenschaftlichen Erzählens. Hier allerdings eine genaue Grenze zu ziehen, wird schwierig sein. Wichtig bleibt die Zuverlässigkeit, Genauigkeit, Eindeutigkeit der Information auf der Beschreibungsebene, wobei allen klar ist, daß jede Beschreibung auch Interpretation ist. Solche Informationen zu sammeln und verfügbar zu machen, ist die Aufgabe der wissenschaftlichen Mitarbeiter/innen u. a. an Museen.

Die Ausgangsfrage ist damit beantwortet: Die Arbeit in der Museumsdokumentation muß sich erheblich ändern und u. a., denn viele Themen wie Normung bei der Bearbeitung von Bild- und Tondokumenten sind außen vor geblieben, den oben genannten Bedingungen genügen müssen, um in einer zunehmend digitalisierten Informationswelt zukunftsfähig zu bleiben.

Internetquellen:

Archive in Nordrhein-Westfalen: <http://www.archive.nrw.de/>

Bibliothèque nationale de France: <http://gallica.bnf.fr/>

Bildarchiv Foto Marburg: <http://www.fotomarburg.de/index.html>

British Library: <http://www.bl.uk/welcome/researchers.html>

British Museum London: <http://www.thebritishmuseum.ac.uk/compass/index.html>

Clio-Online: <http://www.clio-online.de/>

Codices Electronici Ecclesiae Coloniensis: <http://www.ceec.uni-koeln.de/>

Deutsches Historisches Museum: <http://www.dhm.de/datenbank/>

European Cultural Heritage Online: <http://echo.mpiwg-berlin.mpg.de/home>

Google print: <http://print.google.com/googleprint/library.html>

Harvard University Boston: <http://www.harvard.edu/museums/>

Haus der Geschichte der Bundesrepublik Deutschland: <http://sint.hdg.de/sint/html/suche.html>

h-lists: <http://www.h-net.org/reviews/>

H-Soz-u-Kult: <http://hsozkult.geschichte.hu-berlin.de/>

Louvre Paris: http://www.louvre.fr/llv/commun/home_flash.jsp

MPI Europäische Rechtsgeschichte: <http://www.mpier.uni-frankfurt.de/bibliothek/ressourcen.html>

Opus/DissOnline: http://elib.uni-stuttgart.de/opus/gemeinsame_suche.php

Portal Bibliotheken-Archive-Museen: <http://bam-portal.de>

Prometheus Bildarchiv: <http://www.prometheus-bildarchiv.de/>

Stadtarchiv Duderstadt: <http://www.archive.geschichte.mpg.de/duderstadt/dud-d.htm>

Südwestdeutscher Bibliotheksverbund: <http://swb.bsz-bw.de/>

Universitätsbibliothek Heidelberg: <http://www.ub.uni-heidelberg.de/helios/digi/digilit.html>

Ure Museum of Greek Archaeology: <http://www.rdg.ac.uk/Ure/index.php>

Virtuelles Antikenmuseum/ Universität Göttingen: <http://viamus.uni-goettingen.de/fr/mmdb/k>