

Nr. 46
Monika Hagedorn-Saupe (Hrsg.)

Wissen durch Vernetzung –
Kulturgutdigitalisierung in
Deutschland und Europa
Tagungsband – Berlin 2007

*Knowledge by Networking –
Digitising Culture in
Germany and Europe
Conference Proceedings –
Berlin 2007*

Berlin 2008

MITTEILUNGEN und BERICHTE

aus dem

**Institut für
Museums-
forschung**

Redaktion / *Editorial*: Monika Hagedorn-Saupe, Frank von Hagel,
Astrid B. Müller

Übersetzung / *Translation*: Keith Allen

Die Veranstaltung

Wissen durch Vernetzung / *Knowledge by Networking*

wurde in einer Kooperation zwischen dem *Institut für Museumsforschung der Staatlichen Museen zu Berlin – Preußischer Kulturbesitz* und den Partnern von *MICHAEL (Multilingual Inventory of Cultural Heritage in Europe)* organisiert und von EUBAM, der interministeriellen Arbeitsgruppe zu europäischen Angelegenheiten für Bibliotheken, Archive, Museen und Denkmalpflege, sowie der Deutschen Forschungsgemeinschaft unterstützt.

Weitere Unterstützung erfolgte durch:

Adlib Information S, Gallery Systems New York – Berlin, IMDAS-Pro, JOANNEUM RESEARCH Forschungsgesellschaft mbH, Graz, Land Software-Entwicklung, Oberasbach, Mikro-Univers GmbH / Zeuschel GMBHsystems, OCLC PICA GmbH, robotron*Daphne – Web-basiertes Museumsmanagement-System, startext Unternehmensberatung GmbH, zetcom AG – MuseumPlus

Wir danken den Partnern und Unterstützern der Veranstaltung.

Mitteilungen und Berichte aus dem Institut für Museumsforschung

ISSN 1436-4166 Nr. 46

In dieser Reihe werden aktuelle Forschungsergebnisse, Arbeitsberichte und Handreichungen zur Museumsforschung publiziert. Sie ergänzt damit die „Materialien aus dem Institut für Museumsforschung“ und wird interessierten Fachleuten auf Anfrage kostenlos zur Verfügung gestellt.

Eine Liste aller lieferbaren Publikationen des Instituts für Museumsforschung befindet sich am Ende dieses Heftes.

Institut für Museumsforschung
Staatliche Museen zu Berlin –
Preußischer Kulturbesitz
In der Halde 1
14195 Berlin (Dahlem)
Telefon (030) 8301 460
Telefax (030) 8301 504
e-mail: ifm@smb.spk-berlin.de

Preface

The international conference *Knowledge by networking – Digitising Culture in Germany and Europe* took place in the frame of the German Presidency of the Council of the *European Union* in 2007. The conference focus was on networking and co-operational aspects of digitisation projects and initiatives throughout Europe.

The conference was attended by about 400 participants from 26 countries – professionals from museums, libraries, archives, cultural organisations, ministries, and IT-experts. The *Institute for Museum Research* on behalf of the *Stiftung Preußischer Kulturbesitz* together with the partners of *MICHAEL (Multilingual Inventory of Cultural Heritage in Europe-Project)* were the organisers.

The conference was held in German and English language (simultaneous translation offered). In order to ease access for our international readers we provide translation of these papers which were given in German language.

Further information about the conference (presentations, biography, and a ready to print version of this volume) is available on CD-ROM.

The conference was supported by the *DFG (German Research Foundation)* and several sponsors. Full acknowledgement is given on the website (<http://www.knowbynet.de>).

Prof. Monika Hagedorn-Saupe

Einleitung

Die internationale Konferenz *Wissen durch Vernetzung – Kulturgutdigitalisierung in Deutschland und Europa* fand im Rahmen der deutschen Ratspräsidentschaft der *Europäischen Union* 2007 statt. Im Zentrum der Konferenz stand die EU-weite Vernetzung und Zusammenarbeit bei Digitalisierungsprojekten und -initiativen.

Die Konferenz wurde von circa 400 Teilnehmern aus 26 Ländern besucht. Dabei handelte es sich um Fachleute aus Museen, Bibliotheken, Archiven, Kulturorganisationen, Ministerien und IT-Experten. Die Veranstaltung wurde vom Institut für Museumsforschung für die *Stiftung Preußischer Kulturbesitz* gemeinsam mit den Partnern von *MICHAEL (Multilingual Inventory of Cultural Heritage in Europe)* organisiert.

Die Konferenz fand in deutscher und englischer Sprache statt und wurde simultan übersetzt. Die Beiträge werden in der hier vorgelegten Publikation in der Sprache publiziert, in der sie vorgetragen wurden. Um unseren internationalen Lesern die Lektüre zu erleichtern, wurde den deutschen Beiträgen eine englische Übersetzung hinzugefügt.

Weitere Informationen zur Konferenz (Präsentationen, Biographien und eine druckfertige Fassung dieser Publikation) befinden sich auch auf einer CD-ROM, die im Rahmen des *MICHAEL – Projektes* entstanden ist und auf der Homepage zur Konferenz (<http://www.knowbynet.de>).

Die Konferenz wurde von der *DFG (Deutschen Forschungsgemeinschaft)* und weiteren Sponsoren unterstützt.

Prof. Monika Hagedorn-Saupe

Content / Inhalt

Openings / Begrüßungen

KLAUS-DIETER LEHMANN.....P. 7 / S. 10

HERMANN SCHÄFER.....P. 14 / S. 16

Session 1

Digitising Culture in Germany – from digital projects to digital Collections

Kulturgutdigitalisierung in Deutschland –

von einzelnen Digitalisierungsprojekten zu digitalen Sammlungen

MANFRED THALLER

E pluribus unum. Digitising Cultural Heritage within a federal structure.....P. 20

RALF GOEBEL

Research Support of the German Research Foundation (DFG) –

Funding Initiative, Digitisation.....P. 25

STEFAN ROHDE-ENSLIN

kulturerbe-digital.de: A New Information Platform for

Digitisation – Potential and Perspectives.....P. 27

kulturerbe-digital.de: Eine neue Informationsplattform zur

Digitalisierung – Möglichkeiten und Herausforderungen.....S. 29

MARKUS BRANTL

The Munich Digitisation Centre (MDZ).....P. 31

Das Münchener Digitalisierungszentrum (MDZ).....S. 37

KARL-HEINZ LAMPE

From a Single Project to Knowledge Mapping – Manufacturing Knowledge.....P. 44

MONIKA HAGEDORN-SAUPE

MICHAEL-culture.org: Germany's Participation in a

Common Initiative for Europe's Digital Heritage.....P. 51

MICHAEL-culture.org: Eine gemeinsame Initiative für

Europas digitales Kulturerbe und die deutsche Beteiligung.....S. 55

Session 2

Digitising Culture in European Countries – to merge digital collections

Kulturgutdigitalisierung in europäischen Ländern – digitale Sammlungen verknüpfen

IVÁN RÓNAI

How best can a Ministry promote Digitisation Projects?.....P. 59

MARIA SLIWINSKA

Looking for the Needle in the Haystack: Small, Unique Collections –

How to find hidden gems?.....P. 64

PAUL SMITH

A Single Entrance for Access to Heritage Data

at the French Ministry of Culture.....P. 68

ROSSELLA CAFFO

MICHAEL and the Italian Culture Portal: A Cooperation Model among

National, Regional, and Local Institutions.....P. 71

Session 3

Digitising Culture for all Europeans – Networking for Europe’s knowledge
Kulturgutdigitalisierung für alle Europäer – Das Wissen der Europäer sichtbar machen

MARIUS SNYDERS

The European Digital Library initiative.
Joining up the pieces, strategies and actions at European level.....P. 77

JILL COUSINS

Creating the User’s European Digital Library.....P. 81

DAVID DAWSON

MICHAEL and the European Digital Library
promoting teaching, learning and research.....P. 85

CHRISTOPHE DESSAUX

MICHAEL-culture.org – A Gateway for Collections: Open Access through
Open Source Technology. How MICHAEL’s Technique Works.....P. 88

PETRA RAUSCHENBACH

International Standards for Digital Archives –
3rd European Conference on EAD, EAC and METS.....P. 91
Internationale Standards für digitales Archivgut –
Die 3. Europäische Konferenz über EAD, EAC und METS.....S. 95

MATHIAS JEHN

Conditions for Long-term Preservation of Digital Objects in Europe.....P. 99
Voraussetzungen der Langzeitarchivierung digitaler Objekte in Europa.....S.104

Session 4

Digitising Culture more effectively – Learning from other Academic
Kulturgutdigitalisierung effektiv gestalten – Das Rad nicht neu erfinden

MICHAEL BUCKLAND

Geography, Time, and the Representation of Cultural Change –
Experience from a large collaboration (ECAI)P.109

JENS KLUMP

The Benefits of Cross Linking:
The International Continental Scientific Drilling Program.....P.115

Short Contributions / Kurzbeiträge

MONIKA LINDER

Networking Archaeological Knowledge –
Projects of the German Archaeological Institute.....P.122
Archäologisches Wissen vernetzt –
Projekte aus dem Deutschen Archäologischen Institut.....S.126

JÜRGEN KEIPER, WERNER SUDENDORF

Reconstructing Film Culture: Lost Films and Collaboration on the Web
Die Rekonstruktion von Filmkultur:P.129
Verschwundene Filme und Kollaboration im Netz.....S.133

TERESA MALO DE MOLINA

Digitisation of the Historic Press at the Biblioteca Nacional de España.....P. 136

Klaus-Dieter Lehmann

Opening Remarks

I would like to begin by extending a warm welcome to our two-day international conference, "*Knowledge by Networking: Digitising Culture in Germany and Europe*". I am pleased that so many of you from different parts of Germany and Europe have found your way to Berlin. I would like to thank you in advance for your active participation.

Before offering my own thoughts on today's conference, allow me to steal a glance back to the beginning of what has been for us – at least in my view – a successful half year of the German EU Council Presidency. Here in these rooms, at the beginning of the Council Presidency in February 2007, we hosted a conference on digitisation in culture, a conference that to a certain extent served as the prelude to a whole series of meetings and workshops carried out in Germany during the past half year by various institutions, committees and organisations active in cultural affairs. The motto of our February conference was "*Europe's Cultural and Scientific Heritage in Our Digital Age*". The main thrust of the conference was to point to the techniques and technologies currently used to preserve our digital cultural heritage. Exploring together portals and networks, we discussed how electronic data are located and accessed, learning more about how electronic knowledge resources are presenting and conserving the riches of European culture. Special attention was devoted to the importance of universally valid, binding standards, as these standards allow not only the effective registration, but also the optimal presentation of digital information. Whereas in February the main concern was "what" – the presentation of a selection of the most important digitisation projects in Europe – now in May the focus turns to "how" and above all "what's next".

In February I described libraries, archives and museums (of course the preservation of monuments and historic buildings, along with others, belong in this list) as the "backbone of Europe's cultural infrastructure". I would like to briefly elaborate on that thought here.

Thirty-eight years after the invention of the Internet, we face a rapidly changing world, one in which social, economic and social inter-connections are radically changing, a world increasingly internationalised, where the pace of change is growing ever faster. Knowledge institutions – libraries, archives, museums and monument protection authorities – are massively and fundamentally affected by this development. No longer are permanence, linearity and continuity the constitutive factors, but rather organised concurrence, the fleetingness of media and constant acceleration.

Knowledge institutions must on the one hand serve as repositories of learning, integrating the latest scholarship. They also must communicate knowledge to a community of users subjected to a flood of texts and images. In addition, they have to ensure the integrity of their sources. In other words, they are performing a remarkable balancing act, responding to both the ever faster, global accessibility of digital collections, as well as the need to preserve digital knowledge well into the future. These requirements can hardly be met by individual institutions. They have to be solved together, in networks, through co-operative endeavours, by agreements reached in working groups and similar arrangements.

And here we come to an important facet of our conference: we are asking all speakers to elaborate in their presentations on the work they have carried out in co-operation with others, be it of a technical, organisational or any other nature. We are eager to learn more about organisational networks, as well as the re-use of technologies.

Another facet of the topic of this conference concerns the source material: shall it be referenced by cultural branch, e.g., library, museum or archive, or do our source materials in fact transcend long-familiar cultural divides? For research and teaching, it makes good sense to link sources irrespective of their institutional origin. It also makes good technical sense, as retrieval is easier if resources are dispersed and shared by partners. A still more

important aspect, one that will be significant to this conference, is the creation of “cultural ensembles” through linking. Through the interlinking of networks – networks for the most part of physical separate types of information – unique collections are created and made visible. Especially through database linkages spanning institutions and branches of culture, the available technologies can be used to go beyond the constrictions of a specific technical medium to achieve a new degree of freedom.

Libraries, archives and museums are now all confronting the challenges presented by this situation. The question is: how can we create shared knowledge databases of quality and permanence? Libraries and museums have different approaches to digitisation; approaches which vary as these institutions generally abide by different policies regarding questions of collection and presentation. Whereas museums generally collect the unique and the rare, libraries seek to present comprehensive knowledge – a continuum of the literature as it were. Museums place objects in a specific context, thereby offering objects meaning and interpretation. Through the grouping of objects, they create a selective view of the collection. Libraries do not favour this approach. The more direct, the less altered the view of the collection is, the more interesting this view is for science and scholarship. Not the selection, the grouped profile, but instead a broad offering of collection information, equal in ranking, is the aim of the librarian. Interpretation rests with the user.

With the digitisation of analogue material and in the reciprocal linking of these materials, a new quality emerges as a product of divergent organisational principles. It is not only the case that the structure of digital media fundamentally differs from the analogue. Rather, the digitised museum objects themselves can be supplemented and augmented in numerous ways. Hyperlinks to text, data, language and music define manifold axes of viewing, interpretations and experiences. Library collections are no longer limited to the local physical inventory, instead, they expand to become virtual libraries.

The changes brought about by the advent of virtual collections and complex networking of differing materials and cultural organisations will alter the work of specialists. Acting as editors, perhaps even as censors, future librarians will have to sift through the formidable reservoir of data, choosing materials for inclusion. Because of the need to assign formal, content metadata, they will have to develop from cataloguers to curators. Conversely, museum objects can now be presented as mute items in inventories, with interpretation and functionality emerging from selectable contexts. Libraries, archives and museums should thus rapidly name common fields of application, seeking together appropriate project financing. Joint planning should not be limited to professionals from cultural institutions, but rather embrace technical and organisational experts as well.

In Germany, we face special challenges in this context, as federalism is a guiding principle in German cultural affairs. For this and other reasons, we created *EUBAM*, an inter-ministerial federal/state working group devoted to European matters affecting libraries, archives, museums and monument protection. *EUBAM* has assumed responsibility for professional and strategic consultations regarding a German digitisation strategy. *EUBAM* is also responsible for activities which transcend the boundaries of specific types of cultural organisations, such as museums, libraries and archives.

A large European project which reaches across both cultural organisations and languages is *MICHAEL*. *MICHAEL* is the *Multilingual Inventory of Cultural Heritage in Europe*. Our colleagues from *MICHAEL* are co-organisers of this conference. We wish to extend our thanks to our *MICHAEL* partners for their work in putting together this conference. During the next two days we will be introduced both to the perspectives of project partners in individual European countries and the questions raised by the integration of digital collections at the European level. Enormous efforts and an impressive number of participating institutions are realising together this Europe-wide fusion of data at the collections level. Another large European project is the *European Digital Library*, an initiative of the *European Commission*;

we are fortunate to have the opportunity to learn about this initiative from a representative of the Commission.

Beyond these important initiatives to create networks among cultural organisations (the results of which will be introduced at this conference), I would like to emphasize another, equally important issue, and that is co-operation with universities and scholarly research institutes. As I mentioned, one of the challenges faced by cultural organisations is to integrate the latest scholarly findings in our own canon of knowledge, a canon we in turn must preserve and make available for the long-term. The newest findings of course are coming not only from our own institutions – libraries, archives, museums as well as antiquities and monuments offices – but first and foremost from the universities and special research institutes outside of the formal university structure. It is for this reason not merely obvious, but instead essential that the results of research generated in universities and special research institutes are connected to the objects we are safeguarding and making available to future researchers and interested lay people. For this reason, we also need to factor in examples of “grey literature”, such as university theses or research reports generated by smaller entities.

A further aspect that must be mentioned here is the intelligent integration of scientific knowledge with teaching. Alongside their home academic institutions, we have to become for students and young scholars alike partners in their search for expert knowledge. To this end, we must for example publish in a timely manner the latest insights about our objects – and we need to ensure that these findings, pegged to the objects held in our archives, museums and libraries are preserved for the long-term.

All of us working in libraries, archives, museums and monuments protection live from insights supplied by other scholarly disciplines. In numerous and powerful ways, we have witnessed how essential co-operation with disciplines beyond cultural sciences is for all of us, an insight by no means limited to museum restoration workshops. In order, for instance, to preserve in all their materiality the bibliophilic treasures of our libraries or the oversized architectural drawings of such lasting value to antiquities and monuments offices, knowledge of chemistry, biology, the optical sciences and other fields of scientific inquiry is essential.

Looking beyond our own disciplinary boundaries reveals clearly that the natural scientists have for many years now pursued large – indeed very large – scale networking projects. In the fields of astronomy, as well as the earth and biological sciences, to name only a few examples, many universities and research institutes from many different countries have for many years successfully worked together. The possibilities today offered by the Internet allow specialists working at the most diverse locations on earth to pursue common research projects. These endeavours make clear what the networking of knowledge has enabled – as well as how these networks function.

In this regard, practices in the natural sciences are several years ahead of the humanities and cultural sciences. Can cultural organisations learn from the natural sciences? How and under what conditions does networking lead to new knowledge? This will be the focus of the concluding panel discussion, “Culture Meets Science – What Can We Learn from Colleagues in Other Disciplines”?, to which I especially invite you to attend. Here we will consider the steps we need to take to promote trans-national and interdisciplinary networking. We have invited experts from the worlds of culture and science to contemplate how the near future may look and what cultural organisations in particular may wish for as we go forward.

Digital and digitised objects and documents require more active maintenance and support than print publications. This is not only a result of the ongoing technical transformation, of code or format modifications. More active maintenance and support are also required due to recurrent problems with authenticity, that is, uncertain, or ambiguous, attribution. Various partners with different interests are participating in the production, distribution and availability of digital collections, adding to the complexity of the situation. And various

decisions made by these actors at different points in time considerably influence the entire chain of information.

These partners of course have a clear notion of their own goals and procedures, though they still know little about the goals and procedures of other participating parties. Co-ordinated action is a crucial prerequisite to the needs-based, economical production of digital information. Concerted action is also needed to ensure that we are crafting the right approach to long-term accessibility, a prerequisite to a functioning cultural memory.

In closing allow me to mention the second important partner for this conference, the *German Research Foundation (DFG)*. The Foundation's participation in these proceedings is not limited to sponsorship. As we are all eager to hear, the *German Research Foundation* has formulated a digitisation strategy, a development of major importance, as the Foundation is the leading force for national digitisation research projects in Germany. For its support of this conference I wish to express my sincere thanks to the Foundation.

Ladies and gentlemen, allow me in conclusion to also thank our presenters for their contribution to the success of this conference. I wish all of you an exciting conference and stimulating discussion.

Klaus-Dieter Lehmann

Begrüßung

Meine sehr verehrten Damen und Herren, liebe Kolleginnen und Kollegen! Ich möchte Sie ganz herzlich zu unserer zweitägigen internationalen Konferenz „*Wissen durch Vernetzung – Knowledge by Networking – Kulturgutdigitalisierung in Deutschland und Europa*“ willkommen heißen. Ich freue mich, dass Sie so zahlreich aus Deutschland und Europa den Weg zu uns nach Berlin gefunden haben und möchte mich für Ihre engagierte Teilnahme schon jetzt bedanken.

Bevor ich mit einigen grundlegenden Überlegungen zu dieser Konferenz beginne, lassen Sie mich kurz einen Blick zurück an den Anfang dieses – so meine ich – für uns erfolgreichen Halbjahres deutscher EU-Ratspräsidentschaft werfen. Wir hatten zu Beginn der Ratspräsidentschaft, im Februar 2007 in eben diesen Räumen, bereits eine Konferenz zur Digitalisierungsthematik im Kulturbereich, die gewissermaßen als Auftakt einer ganzen Reihe von Tagungen und Workshops verstanden werden kann, die in diesem ersten Halbjahr hier in Deutschland von verschiedenen Institutionen, Gremien und Verbänden im Kulturbereich durchgeführt worden sind.

Die Tagung im Februar stand unter dem Motto „Europas kulturelles und wissenschaftliches Erbe in einer digitalen Welt“. Dabei ging es besonders darum, aufzuzeigen, mit welchen Techniken und Technologien unser digitales kulturelles Erbe derzeit bewahrt wird. Anhand von Portalen und Netzwerken wurde die Auffindbarkeit und Zugänglichkeit elektronischer Daten diskutiert und dabei elektronische Wissensressourcen vorgestellt, die die Vielfalt und den Reichtum unserer europäischen Kultur aufzeigen und bewahren. Besonderes Augenmerk wurde auf den Aspekt allgemeingültiger, verbindlicher Standards gelegt, mit denen nicht nur die Erfassung sondern auch die Präsentation digitaler Daten effektiv und sinnvoll gestaltet werden kann. Ging es uns im Februar vor allem um das „*Was*“ – die Präsentation einer Auswahl wichtiger Digitalisierungsprojekte in Europa –, so soll dieses Mal der Fokus auf dem „*Wie*“ und vor allem dem „*Wie weiter*“ liegen.

Damals hatte ich die Kulturinstitutionen Bibliotheken, Archiven und Museen – und natürlich gehören auch die Denkmalämter und weitere dazu – als das „Rückrat einer europäischen Kulturinfrastruktur“ bezeichnet. Ich möchte diesen Gedanken hier kurz aufgreifen.

Wir sehen uns im Jahr 38 nach der Erfindung des Internet einer rasanten, die gesellschaftlichen, ökonomischen und sozialen Zusammenhänge radikal verändernden und von einer massiven Internationalisierung begleiteten Entwicklung und einem immer schnelleren Wandel gegenüber. Wissensinstitutionen wie Bibliotheken, Archive, Museen und Denkmalämter sind von dieser Entwicklung grundlegend und massiv betroffen: Es sind nicht mehr Permanenz, Linearität und Kontinuität die bestimmenden Faktoren, sondern organisierte Gleichzeitigkeit, mediale Flüchtigkeit und ständige Beschleunigung. Die Wissensinstitutionen müssen einerseits ihr Wissen bewahren und den aktuellen Forschungsstand integrieren, sie müssen ihr Wissen vermitteln – einer von einer Flut an Bild-Wort-Daten ausgesetzten Benutzerschaft – und sie müssen die Seriosität ihrer Quellen sicherstellen – ein Spagat also zwischen schneller und globaler Verfügbarkeit digitaler Datenbestände einerseits und langfristiger Bewahrung dieses Wissens andererseits. Diese Anforderungen können kaum von den Institutionen allein bewältigt werden. Sie müssen gemeinsam, im Verbund, durch Kooperationen, durch Absprachen in Arbeitsgruppen und dergleichen mehr gelöst werden.

Damit haben wir bereits eine Facette unseres Tagungsthemas benannt: Alle Referenten wurden gebeten, in ihren Vorträgen besonders die Erfahrungen aus der Arbeit an ihren Projekten einzubringen, bei denen Fragen der Kooperation, sei es technischer, organisatorischer oder anderer Art, eine besondere Rolle spielen. Dabei wird es um Netzwerke im organisatorischen Sinne aber zum Beispiel auch um die Nachnutzung von Technologien gehen.

Eine andere Facette des Tagungsthemas stellt die Qualität der Sparten – und materialübergreifenden Bezüglichkeit des Quellenmaterials dar: Es ist für Forschung und Lehre nicht nur rational sinnvoll, die Quellen unabhängig von ihrer Institutionenzugehörigkeit nach Wissenszusammenhängen zu verknüpfen – es ist auch technisch überzeugend, sich für das Retrieval auf viele verteilte Ressourcen zu stützen. Über all dem steht aber der weit wichtigere Aspekt, der ein wesentlicher dieser Tagung sein wird: Durch dieses Verknüpfen entstehen „kulturelle Ensembles“, hier wird Wissen generiert, das erst in der Verknüpfung (zumeist physisch verteilter Informationen) einzelner Wissensbestände entsteht und sichtbar wird. Gerade in sparten- und institutionenübergreifenden Verknüpfungen von Datenbeständen können die verfügbaren Technologien ausgenutzt werden und dann einen neuen Freiheitsgrad – und nicht allein technische Verengung auf ein Medium – darstellen.

Bibliotheken, Archive und Museen sehen sich deshalb vor den Herausforderungen, die diese Situation stellt. Nämlich: wie können wir gemeinsam Wissensdatenbestände von Qualität und Dauer schaffen? Die Digitalisierungsvorhaben der Bibliotheken und der Museen unterscheiden sich voneinander, weil sie unterschiedliche Sammlungs- und Präsentationsprinzipien haben. Während Museen in der Regel das Einzigartige und Seltene sammeln, versuchen Bibliotheken das Umfassende – ein Literaturkontinuum – zu realisieren. Museen stellen ihre Objekte in einen Kontext, geben ihnen eine Bedeutung und Interpretation. Sie schaffen letztlich durch Gruppierung von Objekten eine selektive Sicht auf die Sammlung. Bibliotheken wollen das nicht. Je unmittelbarer, unverstellter die Sicht auf die Bestände ist, umso interessanter sind sie für die Wissenschaft. Nicht die Selektion, das gruppierte Profil, sondern das gleichrangige, breite Informations- und Sammlungsangebot wird gewünscht. Die Interpretation liegt beim Nutzer.

In der Digitalisierung des analogen Materials und in der gegenseitigen Verknüpfung liegt wegen der unterschiedlichen Organisationsprinzipien auch eine neue Qualität. Nicht nur, dass sich die Struktur digitaler Medien grundsätzlich unterscheidet, die digitalisierten Museumsobjekte können jetzt in vielerlei Weise ergänzt und angereichert werden. Hyperlinks zu Texten, zu Daten, zu Sprache, zu Musik definieren vielfache Sichtachsen, verschiedene Interpretationen und Erfahrungen. Bibliothekssammlungen beschränken sich nicht

mehr nur auf den lokalisierten physischen Bestand, sondern erweitern sich zu virtuellen Bibliotheken.

Diese Veränderungen durch virtuelle Sammlungen und durch komplexe Vernetzung unterschiedlicher Materialien und Sparten werden die Arbeit der Spezialisten verändern. Bibliothekare müssen künftig wegen der notwendig werdenden Selektion aus dem erheblich gewachsenen Reservoir zu Editoren, vielleicht sogar zu Zensoren werden. Sie müssen sich wegen der Erstellung von formalen und inhaltlichen Metadaten vom Katalogisierer zum Kurator entwickeln. Umgekehrt können Museumsobjekte als stumme Objekte in Inventaren präsentiert werden, die erst aufgrund der wählbaren Kontexte und der zu erzeugenden Funktionalität ihre Interpretation erfahren. Bibliotheken, Archive und Museen sollten deshalb zügig Anwendungsbereiche von gemeinsamem Interesse benennen und dafür geeignete Projektfinanzierungen suchen. Die gemeinsame Planung sollte nicht nur Fachleute der Kultureinrichtungen, sondern auch technische und organisatorische Spezialisten und Nutzer einbeziehen.

Wir in Deutschland sehen uns dabei mit seiner föderalen Struktur vor besondere Aufgaben gestellt und haben unter anderem deshalb eine interministerielle Bund-Länder-Arbeitsgruppe *EUBAM (Arbeitsgruppe für Europäische Angelegenheiten für Bibliotheken, Archive, Museen und Denkmalpflege)* eingerichtet, die fachliche und strategische Beratung bei der Entwicklung einer deutschen Digitalisierungsstrategie übernimmt, aber auch für spartenübergreifende Aktivitäten zuständig ist.

Ein nicht nur spartenübergreifendes sondern auch multilinguales europäisches Großprojekt im Kulturbereich stellt *MICHAEL (Multilingual Inventory of Cultural Heritage in Europe)* dar. Mit den Projektpartnern von *MICHAEL* wird ja diese Konferenz gemeinsam veranstaltet – ihnen allen einen herzlichen Dank dafür. Uns werden in den nächsten beiden Tagen sowohl die Perspektiven von Projektpartnern aus den einzelnen Ländern vorgestellt werden als auch die Fragen des Zusammenschlusses der digitalen Bestände auf dem europäischen Niveau. Hier wird mit einem enormen Aufwand und einer beeindruckenden Zahl an beteiligten Institutionen der europaweite Zusammenschluss von Daten bis auf Sammlungsebene verwirklicht. Ein weiteres ebenfalls europäischen Großprojekt ist: die Initiative der *Digitalen Europäischen Bibliothek – European Digital Library*, und wir werden sie aus der Sicht der Europäischen Kommission, die diese Initiative initiiert hat, erläutern bekommen, worauf wir sehr gespannt sind.

Neben diesen wichtigen Vernetzungsbestrebungen zwischen den Kultureinrichtungen, deren Ergebnisse wir hier auf der Konferenz vorgestellt bekommen, möchte ich aber noch eine andere, ebenfalls essentielle Facette nicht unerwähnt lassen: es ist dies die Kooperation mit der wissenschaftlichen Forschung an den Universitäten und weiteren Forschungsinstitutionen. Wie bereits erwähnt, ist es eine Herausforderung an unsere Kultureinrichtungen, die neuesten Erkenntnisse in unseren Wissenskanon, den es auf Dauer zu erhalten und verfügbar zu machen gilt, einzubinden. Diese kommen natürlich nicht nur aus unseren Einrichtungen, den Bibliotheken, Archiven, Museen und Denkmalämtern selber, sondern vor allem auch aus der universitären und außeruniversitären Forschung. Es ist daher nicht nur nahe liegend, sondern essentiell, die dort erbrachten Forschungsergebnisse zu den von uns bewahrten Objekten zu verbinden und für die weitere Forschung ebenso wie für den interessierten Laien verfügbar zu halten – also beispielsweise auch „graue Literatur“, wie Abschlussarbeiten an den Hochschulen oder Forschungsberichte kleinerer Einheiten zu berücksichtigen. Ein weiterer Aspekt, der hier erwähnt werden muss, ist die sinnvolle Vernetzung zwischen wissenschaftlicher Erkenntnis und Lehre – wir müssen für Studenten und Nachwuchswissenschaftler neben ihren Universitäten und Hochschulen die nächste Säule ihrer Wissensressource werden. Dazu müssen zum Beispiel auch die neuesten Erkenntnisse zu unseren Objekten zeitnah publiziert werden – und diese wiederum angebunden an die Objekte in den Archiven, Museen und Bibliotheken dauerhaft bewahrt werden.

Dass auch wir in den Bibliotheken, Archiven und Museen – und natürlich längst in der Denkmalpflege – von den Erkenntnissen aus anderen Wissenschaften leben, ist hinlänglich bekannt: Es zeigt sich ja nicht nur in den Restaurierungswerkstätten der Museen in vielfältiger und beeindruckender Weise, wie wesentlich die Zusammenarbeit mit den anderen, nicht kulturwissenschaftlichen Disziplinen für uns ist: Chemie, Biologie, optische Wissenschaften, etc. sind essentiell, um unsere Objekte – zum Beispiel auch die bibliophilen Kostbarkeiten der Bibliotheken oder die übergroßen Pläne, aus denen die Denkmalämter die eine oder andere Erkenntnis ziehen – in ihrer Materialität zu erhalten.

Dieser Blick über die eigenen Wissenschaftsgrenzen hinaus zeigt uns auch ganz deutlich, dass es in den Naturwissenschaften Vernetzungsprojekte im großen und ganz großen Stil schon seit vielen Jahren gibt. In den Biowissenschaften, der Astronomie, der Geoforschung – um nur einige Beispiele zu nennen – arbeiten viele Universitäten und Forschungseinrichtungen aus vielen Staaten seit vielen Jahren erfolgreich zusammen. Die heutigen Möglichkeiten des Internet – und damit auch des Digitalen – erlauben Spezialisten von unterschiedlichsten Orten der Erde gemeinsam an einem Forschungsprojekt zu arbeiten. Hier wird augenfällig, was Vernetzung von Wissen bringen kann, wozu es nutzt und – wie sie funktionieren kann.

Die Praxis in den Naturwissenschaften ist den Geistes- und Kulturwissenschaften hier um einige Jahre voraus. Kann der Kulturbereich von den Naturwissenschaften lernen? Wie und unter welchen Bedingungen führt Vernetzung zu neuem Wissen? Dies soll der Fokus der abschließenden Podiumsdiskussion sein „Kultur und Wissenschaft – wie geht es weiter?“, zu der ich Sie besonders herzlich einladen möchte: Hier wird auch über die nächsten notwendigen Schritte für eine transnationale und interdisziplinäre Vernetzung im Bereich Digitalisierung nachgedacht werden – wir haben dafür Experten aus Kultur und Wissenschaft eingeladen, um mit ihnen gemeinsam nachzudenken, was wir uns im Bereich Digitalisierung wünschen und wie die nahe Zukunft aussehen sollte.

Digitale oder digitalisierte Objekte und Dokumente benötigen eine aktivere Verwaltung und Betreuung als gedruckte Publikationen. Das ist nicht nur die Folge des technischen Wandels, der Änderung von Codierungen und Formaten, es sind auch die auftretenden Probleme der Authentizität, der unsicheren mehrdeutigen Zuordnung. Der Aufwand hat weiterhin damit zu tun, dass an der Produktion, an der Distribution und an der Verfügbarkeit von digitalen Sammlungen über den zeitlichen Zyklus verschiedene Partner mit verschiedenen Interessen beteiligt sind, deren jeweilige Entscheidungen die Informationskette erheblich beeinflussen.

Diese Partner haben zwar eine klare Vorstellung von ihren eigenen Vorgehensweisen und Zielen, aber noch immer eine geringe von denen anderer Beteiligter. Deshalb ist abgestimmtes Handeln eine wesentliche Vorbedingung für eine bedarfsgerechte und ökonomische Produktion und für eine geeignete, vertretbare Langzeitverfügbarkeit – im Sinn eines funktionierenden kulturellen Gedächtnisses.

Bevor ich schließe, möchte ich den zweiten wichtigen Partner für diese Konferenz nicht unerwähnt lassen: es ist die *Deutsche Forschungsgemeinschaft*, die sich nicht nur mit ihrer finanziellen Unterstützung an dieser Konferenz beteiligt, wir werden auch – und das wird uns hier im Saal besonders interessieren – über die Digitalisierungsstrategie der *DFG* ausführlich hören. Die *DFG* ist für die Entwicklung der nationalen Digitalisierungsstrategie von großer Bedeutung. Sie ist die treibende Kraft für nationale Digitalisierungsprojekte für die Forschung in Deutschland. Ich möchte mich an dieser Stelle noch einmal ganz herzlich für ihre Unterstützung dieser Tagung bedanken.

Meine sehr verehrten Damen und Herren, ich darf mich schon jetzt bei allen Referentinnen und Referenten bedanken, die mit ihren Beiträgen zum Gelingen dieser Konferenz beitragen werden. Ihnen allen eine spannende Tagung und anregende Diskussionen.

Hermann Schäfer

Opening Remarks

Professor Lehmann, Ladies and Gentlemen, as a clever sociologist once put it, "the Internet brings people together in a way that allows them to be wherever they want". In this way he enunciated a phenomenon that characterises electronic networking. In an hitherto unknown dimension, the Internet makes possible the linking of knowledge – without demanding physical proximity among the sources of this knowledge. That's why we can say today that no ambitious idea, no project with the aim to come to Europe-wide fruition, can get by without a networking of all the actors. And the networking instrument is the Internet.

As the digitisation of culture in Europe represents such an ambitious, forward-looking project, I am very pleased to welcome you here in the *Berlin State Library* to another digitisation conference taking place in the context of the German EU Council Presidency.

Aspects of today's themes were discussed at our previous digitisation conference in February of this year. At that gathering, also here in Berlin, more than four hundred participants discussed questions related to access and use of electronic scholarly resources; we also gained important insights into how best to preserve Europe's digital cultural heritage. Two additional conferences in April 2007, sponsored by the *German National Library* in Frankfurt and the *German Federal Archives* here in Berlin, have already illuminated areas of emphasis in reference to digitisation strategy. The Frankfurt meeting played a significant role in defining the next steps for Europe's strategy as regards digital cultural infrastructure, whereas the gathering of archivists in Berlin indicated that the three international standards discussed at that meeting, *Encoded Archival Description (EAD)*, *Encoded Archival Context (EAC)* and *Metadata Encoding and Transmission Standard (METS)* should be combined to ensure that archives and their collections attain improved visibility. The need for networks and co-operation among different organisations and resources is by no means limited to the digitisation domain. In the context of the German EU Council Presidency, we organised a conference in May in Bremen on the subject of collections mobility in which the formation of networks was also a topic. The digitisation of cultural assets nonetheless raises a series of questions that point to the special need for combined European expertise. This concentration of expert knowledge is required if we are to seize the enormous opportunities presented by digitisation.

I am convinced that this effort is worthwhile, as the matter at hand, Europe's cultural and scientific heritage, is the basis of Europe's common identity. In the course of each these three conferences it became clear that the topic of networks is so capacious that it should, in the context of the German EU Council Presidency, become the subject of a conference all its own.

At first glance, the situation seems clear. As the sheer number of existing networks and corresponding Internet forums demonstrates, the grouping together of interests via networks has become one – if not the – meaningful form of scholarly exchange in our time. Rapid technological development has made it possible to verify data and to bundle information from various sources across vast distances. The Internet enables us to present various types of information together, allowing us to reach interested parties with virtually no distribution costs. Not only are resources being made available in enormous volumes, but increasingly the resources themselves are directly accessible.

All this is of course taking place in addition to the individual offerings of respective cultural organisations on the Internet and in the "analogue world". Made possible through digitisation, these new possibilities allow us to realise what was long wished and striven for. I restrict myself here to one example, the visionary concepts regarding documentation set out by the Belgian pioneer in this area of knowledge, Paul Otlet. Already at the beginning of the twentieth century Otlet recognised the significance of databases. What this visionary

conceived of nearly one hundred years ago is today increasingly realised by Internet technology. It's not by chance that Otlet is called the "architect of knowledge".

Today's greatly enhanced technological possibilities are a timely answer to the steady growth in user demand during the past decades. What's involved is not merely as processing and propagating more information than hitherto. Rather, what's required is to make accessible this information as well as all previously accessible information much more quickly: in full detail, not merely references to where the information in question is physically located. All this said, the question remains: does the opportunity to interlink knowledge also create the possibility to create knowledge? This is, after all, the claim of this conference with its title "knowledge by networking". In response to this question, allow me to point out that *Web2.0* developments are – as we all know – striking evidence that what we are witnessing is much more than a change in the modalities of information access. The Internet is no longer the unaltered digital rendering of the analogue world. Rather, the Internet serves as a catalyser for the development of true cultural and scholarly added value. Networks create the necessary conditions, advancing the meaningful use of information in instances where individual data collections are enriched with added value. Often, the true value of a data collection only emerges when it has been brought together with sources, facts or other sources of information located elsewhere. This occurs through embedding in a specific context, the expansion of a context or through the interlinking with other, from a content perspective appropriate, units of information. Examples may include ancillary writings, images, or audio or video materials. The list of potentially important sources is nearly endless. In the end, the object of research is the determining factor. Hence, what's needed is the most comprehensive possible interlinking of our digitised heritage, made available to the widest possible audience, for there is really no conclusive way of knowing what types of information may be relevant in the future.

Digital access to archives, museums and libraries allows research on a scale that previously was only possible in isolation, at a high cost and with large investments of time. Further, by virtue of the digital animation of objects, we can now conduct studies and experiments that previously would not have been possible because of well-placed concerns about the integrity of the original.

In my personal experience as longstanding president of the *Stiftung Haus der Geschichte*, I furthermore came to appreciate the value of the Internet in museum work. Digitisation offers us completely new avenues of creative and cultural developments, as for example through virtual museum tours. An evaluation I commissioned of the *Haus der Geschichte* Internet presentation yielded surprising results: for example, the proportion of male to female visitors to the website was more balanced than expected, the age distribution broader than anticipated. Those under twenty years of age constituted eleven percent of users, whereas those over sixty represented five percent. The largest group (eighty percent) were employed individuals between the ages of twenty and sixty. Eight percent of the visitors to the website come from abroad. Forty-six percent of homepage users had never visited the physical museum. All this indicates possibilities offered through digitisation, as digitisation leads to an enormous expansion of users for cultural programmes. In the future, this development will certainly play an important role as regards approaches to cultural education.

The use of digital data requires a basis, such as portals and other initiatives, to support appropriate design and above all the long-term preservation of digital materials. Relevant initiatives will be introduced at this conference. I am very eager to hear your presentations.

No attempt to digitise cultural heritage today is conceivable as a purely national initiative, for national approaches alone would leave unused considerable portions of digitisation potential. For good reasons, then, digitisation is proceeding within a larger European framework.

A central goal of cultural and educational policy ratified by the European Council in Brussels in November 2006 is the creation of the *European Digital Library (EDL)*, a project that will, through a true network, make available in the future an enormous sum of European digital resources. The *EDL* is by no means limited to digital publications; rather, the *EDL* calls for, in equal measure, contributions from archives, libraries, museums and other cultural organisations. Only as a grouping of distributed resources is this "library" imaginable and worthwhile. A clear structure and co-ordination of the contents is clearly needed to ensure that duplication and overlapping of efforts can be kept to a bare minimum. To this end, a considerable degree of co-ordination, planning and other agreements in a network structure are necessary. What's critical is that the *European Digital Library* emerges from local and regional resources already available. Only through the merger of diverse individual projects and initiatives can we reach our goal to offer visitors access to large, rich, diverse holdings.

Regarding the question how such a network shall emerge, what must be considered carefully from the perspective of the participating organisations and state funding agencies, our conference today and tomorrow offers a series of presentations. I hope that this conference makes clear the importance of networking as a requirement of cultural endeavour, and as an increasingly important factor for all learning activities, not least as a means to offer Europe's citizens an ever increasing number of cultural offerings. It is in this sense that we should understand the motto the German Chancellor chose for the EU Council Presidency, "Shaping Europe Together". For only through co-operation and networking are we in a position to make Europe fit for the global challenges of the knowledge society.

Hermann Schäfer

Grußwort

Sehr geehrter Herr Professor Lehmann, meine sehr geehrten Damen und Herren, „Das Internet bringt Menschen so zusammen, dass sie bleiben können, wo sie wollen.“, so formulierte ein kluger Soziologe einmal. Ausgesprochen hat er damit ein Phänomen, das diese Art elektronischer Vernetzung charakterisiert: Das Internet ermöglicht eine nicht gekannte Dimension der Verknüpfung von Wissen, ohne deshalb eine räumliche Nähe zwischen den Quellen des Wissens zu erfordern. Daher können wir heute feststellen: Keine ambitionierte Idee, kein Projekt, das eine europaweite Verwirklichung anstrebt, kommt heutzutage ohne eine Vernetzung der Akteure aus. Das Internet ist hierzu das Instrument.

Und da auch die Kulturgutdigitalisierung in Europa ein solches ambitioniertes und in die Zukunft weisendes Projekt darstellt, freue ich mich sehr, Sie heute hier in der *Staatsbibliothek zu Berlin* zu einer weiteren Digitalisierungskonferenz im Rahmen der deutschen EU-Ratspräsidentschaft begrüßen zu dürfen.

Schon bei unserer vorausgegangenen Digitalisierungskonferenz im Februar dieses Jahres in Berlin wurden Aspekte des heutigen Themenschwerpunktes behandelt. So haben die über 400 Teilnehmer bereits über Fragen der Auffindbarkeit elektronischer Wissensressourcen, ihrer Zugänglichkeit und Nutzbarkeit diskutiert und wichtige Erkenntnisse für die digitale Erhaltung des kulturellen Erbes gewonnen. Auch die Bibliothekskonferenz im April in Frankfurt und die Archivkonferenz im April in Berlin haben bereits Schwerpunktbereiche der Digitalisierungsstrategie beleuchtet. Die europäische Strategie zum Aufbau einer Infrastruktur für das kulturelle Erbe der unterschiedlichen Kulturnationen wurde in Frankfurt maßgeblich vorangetrieben. Auf der Archivkonferenz hat sich gezeigt, dass die drei international diskutierten Standards *Encoded Archival Description (EAD)*, *Encoded Archival Context*

(EAC) und *Metadata Encoding and Transmission Standard (METS)* kombiniert werden sollten. Nur so können Archive und ihre Bestände deutliche und bessere Sichtbarkeit erlangen. Die Notwendigkeit von „Netzwerken“ und das Zusammenwirken verschiedenster Einrichtungen und Ressourcen sind jedoch nicht auf den Bereich der Digitalisierung beschränkt. So haben wir in unserer EU-Präsidentschaft zum Thema „Mobilität von Sammlungen“ im Mai in Bremen eine Konferenz veranstaltet, bei der Netzwerkbildungen ebenfalls ein Thema waren. Die Digitalisierung von Kulturgut wirft allerdings Fragestellungen auf, die im ganz besonderen Maße den Einsatz gebündelten, europäischen Sachverständigen notwendig werden lässt, wenn die enormen Chancen der Digitalisierung auch tatsächlich genutzt werden sollen.

Aber ich bin der festen Überzeugung, dass sich dieser Einsatz lohnt. Denn es geht hier um Europas kulturelles und wissenschaftliches Erbe. Ein Erbe, das die Basis einer gemeinsamen europäischen Identität ist. Es wurde bei allen Konferenzen aber klar, dass das Thema der Netzwerke so umfassend ist, dass es über die drei genannten Konferenzen hinaus im Rahmen der deutschen EU-Ratspräsidentschaft Gegenstand einer eigenen Konferenz werden musste.

Die Sachlage ist auf den ersten Blick eindeutig: In Netzwerken und dazugehörigen Internetforen wird informiert, und diskutiert und allein schon die bloße Menge an bestehenden Netzwerken zeigt deutlich, dass in dieser selbst organisierten Bündelung von Interessen eine – wenn nicht sogar die – wesentliche Form des Wissensaustausches in der heutigen Zeit zu sehen ist. Die rasante technische Entwicklung hat es möglich gemacht, den Nachweis von Informationen und den Nachweis, wo sie zu finden sind, aus unterschiedlichsten Herkünften und über weite Entfernungen hinweg zu bündeln. Das Internet ermöglicht es, Informationen in gemeinsamen Angeboten zu präsentieren und zu Distributionskosten, die gegen Null tendieren, problemlos allen Interessenten zu übermitteln. In umfangreicher Weise werden nicht mehr nur Informationsnachweise, sondern zunehmend auch die Ressourcen selbst direkt zugänglich gemacht.

Und dies alles geschieht natürlich zusätzlich zum fortbestehenden individuellen Angebot der einzelnen Institutionen im Internet oder in der „analogen Welt“. Diese neuen, auf der Digitalisierung basierenden Möglichkeiten erlauben uns das zu verwirklichen, was schon seit langem angestrebt und erwünscht war. Ich nenne an dieser Stelle nur die visionären Konzepte der Dokumentation, die der belgische Pionier auf diesem Gebiet, Paul Otlet, schon zu Beginn des 20. Jahrhunderts ausarbeitete, als er die Wichtigkeit von Datenbanken erkannte. Was er vor bald 100 Jahren nahezu visionär erdachte, kann heute in zunehmendem Maße technisch realisiert werden. Nicht von ungefähr wurde er darum „Architekt des Wissens“ genannt.

Die enorm verbesserten technischen Möglichkeiten antworten zur rechten Zeit auf einen sich über die vergangenen Jahrzehnte stetig steigenden Benutzerbedarf. Und dabei gilt es nicht nur, deutlich mehr Informationen als früher zu verarbeiten und zu verbreiten, sondern auch die an sich und auch schon früher verfügbare Information in erheblich schnellerer Zeit verfügbar zu machen – in vollem Umfang der Inhalte, nicht nur als Hinweise, wo die Information zu finden ist. Und doch stellt sich die Frage: Ist die Möglichkeit, Wissen zu verknüpfen auch eine Möglichkeit, Wissen zu schaffen? Denn genau dies wird ja im Titel dieser Tagung behauptet: „Wissen durch Vernetzung“. Lassen Sie mich hierzu auf folgendes verweisen: Die Entwicklungen des *Web2.0* sind – wie wir alle wissen – ein eindrucksvolles Beispiel dafür, dass es nicht allein um die Modalitäten des Informationszugangs geht. Nicht mehr die inhaltlich unveränderte digitale Wiedergabe der analogen Welt prägt das Internet. Vielmehr wirkt das Internet als Katalysator für die Entwicklung eines echten kulturellen und wissenschaftlichen Mehrwertes.

Netzwerke schaffen hierfür die notwendigen Voraussetzungen: So wird die sinnvolle Nutzung von Information erheblich gefördert, wenn die jeweilige Einzelinformation mit einem „Mehrwert“ angereichert wird. Informationen zeigen ihren vollen Aussagewert oft erst dann, wenn sie mit an anderer Stelle vorhandenen Quellen, Fakten oder sonstigen

Informationen zusammengebracht werden. Dies geschieht durch Einbetten in einen Kontext, die Erweiterung eines solchen Kontextes oder auch durch Verknüpfung mit anderen, sinnvoll zu ihr passenden Informationseinheiten. Dies können zum Beispiel ergänzende Schriften, Bildzeugnisse oder auch Phono- und Filmmaterialien sein. Die Liste potentiell wichtiger Informationen ist nahezu unendlich und bestimmt sich letztlich nach dem jeweiligen Forschungsgegenstand. Notwendig ist daher eine möglichst umfassende Verknüpfung und Zugänglichkeit unseres digitalisierten Kulturgutes. Denn abschließend lässt es sich heute überhaupt noch nicht sagen, welche Informationen zukünftig von Bedeutung sein können.

Der digitale Zugriff auf Archive, Museen und Bibliotheken lässt Forschungen in einem Umfang zu, wie er zuvor nur vereinzelt, mit großen Kosten und mit hohem Zeiteinsatz möglich war. Die digitale Animation von Objekten ermöglicht es darüber hinaus, Studien und Experimente durchzuführen, die aufgrund des notwendigen Schutzes des Originals überhaupt nicht möglich gewesen sind.

Aus meiner persönlichen Erfahrung als langjähriger Präsident der *Stiftung Haus der Geschichte* habe ich zudem das Internet vor allem im Museumsbereich schätzen gelernt. Denn die Digitalisierung eröffnet uns völlig neue Räume der kreativen und kulturellen Entfaltung wie z.B. virtuelle Museumsrundgänge. So hat eine von mir beauftragte Evaluation des Internet-Auftrittes des Hauses der Geschichte überraschende Ergebnisse erbracht: So ist beispielsweise die Verteilung zwischen Männern und Frauen gleichgewichtiger als erwartet und die Altersverteilung breiter als vorab vermutet: die Nutzerquote lag bei den unter 20jährigen bei 11 Prozent, bei den über 60jährigen bei 5 Prozent. Der mit 80 Prozent größte Anteil liegt somit bei der berufstätigen Gruppe im Alter zwischen 20 und 60 Jahren. 8 Prozent der Surfer kommen dabei aus dem Ausland. 46 Prozent der Nutzer der Homepage haben noch nie das Haus selbst besucht. Daran zeigt sich, welche Möglichkeiten die Digitalisierung bietet, die den Nutzerkreis kultureller Angebote enorm erweitert. Dies wird bei Fragen der kulturellen Bildung in der Zukunft mit Sicherheit eine wichtige Rolle spielen.

Diese Nutzung digitaler Daten benötigt aber eine Basis, beispielsweise Portale und Initiativen, die die angemessene Gestaltung und vor allem die Langzeitarchivierung digitalen Materials unterstützt. Entsprechende Initiativen werden auf dieser Tagung auch hierzu vorgestellt – ich bin auf Ihre Beiträge sehr gespannt.

Alle Arbeiten im Bereich der Digitalisierung von Kulturgut sind heute nicht mehr als rein national begrenzte Initiativen denkbar. Denn rein nationale Initiativen würden ein erhebliches Potential der Kulturgutdigitalisierung ungenutzt lassen. Sie finden daher aus gutem Grund innerhalb einer europäischen Vernetzung statt.

Ein zentrales kulturpolitisches Ziel, auf das sich der Rat in Brüssel im November 2006 mit Schlussfolgerungen festgelegt hat, ist die Schaffung einer „*Europäischen Digitalen Bibliothek*“. Mit der *Europäischen Digitalen Bibliothek* entsteht zukünftig ein echtes Netzwerk mit einer enormen Menge europäischer digitaler Ressourcen. Sie erfordert gleichermaßen den Beitrag von Archiven, Bibliotheken, Museen und anderen Kultureinrichtungen und beschränkt sich damit nicht auf den Zugriff auf digitales Schriftgut. Und nur als Bündelung solcher verteilten Ressourcen ist diese „Bibliothek“ denkbar und sinnvoll. Dabei kommt es auf eine gute Struktur und eine Abstimmung der Inhalte an, so dass Doppelarbeiten und Überschneidungen möglichst gering gehalten werden. Ein erhebliches Maß an Koordination, Planung und Absprache im Netzwerk sind hierfür nötig. Entscheidend aber ist, dass die *Europäische Digitale Bibliothek* aus heute bereits vorhandenen lokalen, regionalen Ressourcen entsteht. Nur die Zusammenfassung der vielfältigen Einzelprojekte und -initiativen kann zu dem Ziel führen, einen großen, reichen, vielfältigen Bestand anzubieten.

Für die Frage, wie ein solches Netzwerk entstehen kann, was es aus Sicht der Einrichtungen aber auch der staatlichen Förderer zu beachten gilt, bietet unsere Tagung heute und morgen eine ganze Reihe von Beiträgen. Ich hoffe, dass die Tagung die Vernetzung als Faktor der fachlich-wissenschaftlichen Bedürfnisse zeigt, aber auch als zunehmend wichtigen

Faktor für alle Aktivitäten des Lernens, und nicht zuletzt als Form der Schaffung eines immer dichterem Netzes kultureller Angebote für die Bürger Europas.

In diesem Sinne ist auch das Motto, das die Bundeskanzlerin für die deutsche EU-Ratspräsidentschaft gewählt hat, zu verstehen: „Europa gemeinsam gestalten“. Denn nur durch Zusammenarbeit und Vernetzung sind wir in der Lage, Europa für die globalen Herausforderungen der Wissensgesellschaft fit zu machen.

Manfred Thaller

E pluribus unum. Digitising Cultural Heritage within in a federal structure

Zusammenfassung

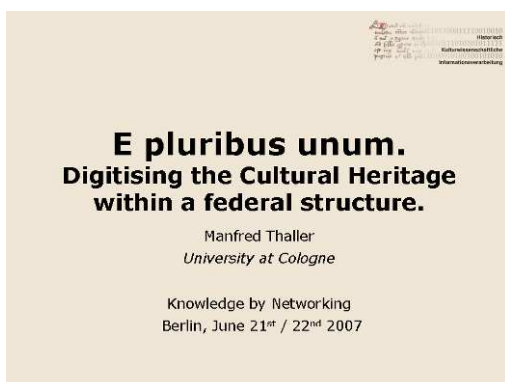
Um Deutschlands Kulturpolitik zu beschreiben, muss man mindestens sechzehn Politiken darstellen, da es keine zentrale Autorität für das Feld gibt. Dies führt zu vielen eigentümlichen Situationen: Eine der jüngeren strategischen Entscheidungen der *Deutschen Forschungsgemeinschaft* geben Deutschland eines der ambitionösesten und systematischsten Programme zur Digitalisierung des kulturellen Erbes in Europa; da die Ergebnisse extrem schlecht sichtbar sind, tendieren jedoch sogar deutsche Kulturpolitiker dazu, sich für die geringe Aktivität des Landes in diesem Bereich zu entschuldigen.

Die Präsentation zeigt den tatsächlichen Reichtum und analysiert die Gründe für die mangelnde öffentliche Sichtbarkeit. Sie verweist anschließend auf Parallelen zwischen einer föderalen Struktur und einer kontinentalen Informationslandschaft, die von unabhängigen nationalen Einheiten errichtet wird. Sie schließt mit einer Reihe von Thesen, was aus den deutschen Erfahrungen gelernt werden kann.

Abstract

Describing German cultural policy, one has to describe at least sixteen policies, as there is no central authority in cultural matters. This leads to many strange situations: Some of the recent strategic decisions of *Germany's research council (DFG)* provide it with one of the most ambitious and systematic funding situations for the digitization of the cultural heritage anywhere in Europe; as the resulting resources are extremely fragmented and poorly visible, even German politicians tend to apologize for the low level of her activities within the area, however.

The presentation starts by showing the richness in reality and analysing the reasons for the lack of its public perception. It continues by drawing parallels between a federal country and a continental information landscape created by independent national players. It closes by some theses what might be learned from the German experience.



A fundamental change ...

Today *reproducing* is much cheaper than *describing*, however.

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007

... changes fundamental assumptions.

Digitising in the cultural heritage domain started in the eighties as advanced activity, implemented by specialist centres, dedicated to the digitisation of a few exalted examples.

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007

... changes fundamental assumptions.

We are discussing *today* the systematic conversion of the complete content of Europe's libraries, archives and museums.

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007

Some constancy?

If even the digitisation of exalted examples requires specialist centres, a more global approach requires centres of excellence, national - or maybe continental?

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007

<http://www.scottishdocuments.com>

Probably the largest European Cultural heritage resource.
Complete coverage of a region (Scotland).
More than 5 million lower upper range manuscript pages.
"Self funding".

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007

<http://www.scottishdocuments.com>

Probably the largest European Cultural heritage resource.
Complete coverage of a region (Scotland).
More than 5 million lower upper range manuscript pages.
"Self funding".

*And not started by any of the strategic initiatives of JISC.
Nor "best practice".*

Manfred Thaller Knowledge by Networking, Berlin, June 21st

Thesis:

Digitisation of the cultural heritage in 2007 is not the exception, but the rule.

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007




 Historisch
 Kulturrechtliche
 Informationsverarbeitung

Germany:

Has projects with extremely high access rates and great international impact.


Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



 Historisch
 Kulturrechtliche
 Informationsverarbeitung



Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



 Historisch
 Kulturrechtliche
 Informationsverarbeitung

Germany:

Has projects which systematically provide access to whole classes of resources.

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



 Historisch
 Kulturrechtliche
 Informationsverarbeitung



Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



 Historisch
 Kulturrechtliche
 Informationsverarbeitung

Germany:

Has the first project world wide, which attempts to apply digitization by scanning automats to early modern prints.

(Munich digitisation centre; first speaker after coffee break.)

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007




 Historisch
 Kulturrechtliche
 Informationsverarbeitung

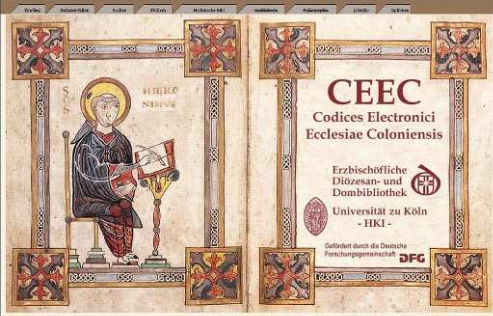
Germany:

Has since 2001 the project which offers world wide the highest number of high resolution medieval codices without access restrictions, which has been copied as starting platform for a systematic national digitisation effort in Switzerland.


Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



 Historisch
 Kulturrechtliche
 Informationsverarbeitung



Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007




 Historisch
 Kulturrechtliche
 Informationsverarbeitung

Germany:

Is one of the first countries with a policy by the National Research Council (DFG), which is not only proclaiming the complete digitization of the holdings of her libraries as a strategic goal, but actually intends to pay for a substantial part of it.

(Next speaker.)

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007




Germany:

Is depicted by its national press as having almost no activities in cultural heritage digitization whatsoever ...

... and has politicians which apologize, while opening conferences on digitization, for Germany being a developing country in this area.


Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



Why?

1. Due to a lack of a clear national responsibility for a information policy on cultural heritage digitisation, that information policy is implemented by individual libraries, archives and museums.


Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



Why?

2. These institutions have an institutional interest in emphasising their profile by claiming to be the only (or almost only) German institution implementing such projects.


Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



Solutions:

Make digitisation the exclusive task of one or a very few national centres, charging them with making their resources visible.


Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



However:

No real discernible differences in the quality of the resources created, or the cost effectiveness achieved, of the ca. top 40 % of Germany's digitisation projects (by institution), as the basic technologies are in the meantime "trivial" in a computer science meaning and ...

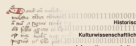
Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



... recapituled thesis:

Digitisation of the cultural heritage in 2007 is not the exception, but the rule.


Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



Consequence:

If you declare the digitisation of cultural heritage material to be the domain of a few centres of competence, you are effectively saying, that the majority of the memory institutions shall remain incompetent in one the fundamental information techniques of the time.


Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



Thesis:

A memory institution which is able to digitize the segment of the cultural heritage it is responsible for, is not especially qualified for the challenges of 2007 ...


Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



Thesis:

... but a memory institution which is *not* able to digitize the segment of the cultural heritage it is responsible for, is significantly unqualified for its mission.

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007




Corollary:

A German national policy for digitisation cannot consist in disqualifying the majority of the memory institutions.

It has instead to encourage them to increase their qualification and provide a strong infrastructure in which their inborn institutional egotism is countered, to promote the sum of the parts together.

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007




Vision:

A highly effective national infrastructure, which virtually combines individual offerings.

It makes no rule, which is not absolutely required, but makes following those, which are, self rewarding.

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007



Thank you for your attention!

Manfred Thaller Knowledge by Networking, Berlin, June 21st 2007

Ralf Goebel

Research Support of the German Research Foundation (DFG) – Funding Initiative, Digitisation

Zusammenfassung

Die *Deutsche Forschungsgemeinschaft* ist die zentrale Förderorganisation für die Forschung in Deutschland. Als politisch unabhängige Organisation unterstützt die *DFG* alle Bereiche der Wissenschaft durch die Finanzierung von Forschungsprojekten. Im Strategiepapier „*Wissenschaftliche Literaturversorgungs- und Informationssysteme – Schwerpunkte der Förderung bis 2015*“ (<http://www.dfg.de/lis/en>) hat die *Deutsche Forschungsgemeinschaft* für den Bereich der digitalen Informationsversorgung die wichtigsten Handlungsfelder der kommenden Jahre formuliert:

Mit der Digitalisierung wissenschaftlicher Quellen (insbesondere von Büchern und handschriftlicher Überlieferung) reagiert die *DFG* auf die zunehmende Erwartung der Wissenschaftler an die Nutzung elektronischer Dokumente direkt an jedem Arbeitsplatz. In den kommenden Jahren werden daher Druckpublikationen von der Erfindung des Buchdrucks an bis zum Ende des 20. Jahrhunderts großflächig digitalisiert (Retrodigitalisierung).

Komplementär wird seit einigen Jahren mit Finanzierung der *Deutschen Forschungsgemeinschaft* für alle Wissenschaftler und Studenten in Deutschland der nationale Zugang zu ausgewählten digitalen Verlagsangeboten ermöglicht (Lizenzierung). Dies umfasst die Koordinierung von Verlagsverhandlungen für digitale Sammlungen und Backfiles im Rahmen eines nationalen Konsortiums (Nationallizenzen), die Fortführung und Erweiterung der Idee der Nationallizenzen auf die kontinuierliche Lizenzierung laufender wissenschaftlicher Zeitschriften und Datenbanken sowie die Erweiterung der Idee der Nationallizenzen im Rahmen internationaler Konsortien.

Vor dem Hintergrund des guten Ausbau-

Abstract

The *German Research Foundation (DFG)* is the central funding agency for German research. Politically neutral, the *German Research Foundation* supports all areas of science and scholarship by funding specific research projects. In the Foundation's Position Paper, "*Scientific Library Services and Information Systems, Funding Priorities Through 2015*" (see more <http://www.dfg.de/lis/en>), the *German Research Foundation* has set the following objectives for information infrastructures:

Through the digitisation of moveable cultural heritage (especially books and other handwritten sources), the Foundation aims to ensure scholars are able to access libraries and other scholarly repositories primarily from their office computers. For this reason, in the coming years publications from the invention of printing to the end of the twentieth century will be retrodigitised.

In recent years the Foundation has – through licensing that students and scholars across Germany enjoy – provided access to publishers' digital offerings. This entails: co-ordination of negotiations with the publishing houses vis-à-vis digital collections and backfiles in the context of a national consortium (national licensing); continuation and expansion of national licensing via uninterrupted licensing of scientific journals and databases; expansion of the national licensing concept to encompass international consortia as well.

In the coming years the Foundation will encourage the growth of a national network of institutional and specialist document repositories. Publication servers in German universities will offer a good starting point for this effort.

Given the good development of publica-

standes von Publikationsservern an deutschen Hochschulen sollen diese in den kommenden Jahren zu einem deutschlandweiten Netzwerk institutioneller und fachlicher Dokumentenrepositorien ausgebaut werden. Auch wird die Frage nach dem künftigen Umgang und der Nutzung wissenschaftlicher Primärdaten einen künftigen Schwerpunkt in der strategischen Ausrichtung der DFG einnehmen. Der Implementierung einer nationalen Hosting- und Archivierungs-Strategie kommt hierbei eine zentrale Bedeutung zu.

Zum Ausbau der wissenschaftlichen Literaturversorgungs- und Informationssysteme in Deutschland will die *Deutsche Forschungsgemeinschaft* im Rahmen ihrer Sonderfördermaßnahme „Digitalisierung“ bis zum Jahr 2012 insgesamt 250 Mio. Euro zur Verbesserung der Informationsangebote zur Verfügung stellen. Die Maßnahme wird eingebunden in eine gemeinschaftliche Initiative der „Allianz der deutschen Wissenschaftsorganisationen“ (Schwerpunktinitiative „Digitale Information“ der Allianz-Partnerorganisationen 11.06.2008 / Priority Initiative „Digital Information“ by the Alliance of German Science Organisations 11.06.2008).

tion servers at German universities; it is intended to form these into a nation-wide network of institutional and specialized document servers in the coming years. Equally, the question of the future handling and use of scientific primary data will be in the focus of the DFG's future strategic direction. The implementation of a national hosting and archival strategy is of central significance in this context.

Beginning in 2008, to expand scholarly information infrastructures in Germany the *German Research Foundation* will – within its special funding scheme “Digitisation” – allocate 250 million Euros until 2012, for the improvement of the information provision. This measure will be part of the joint initiative “Allianz der deutschen Wissenschaftsorganisationen” Schwerpunktinitiative “Digitale Information” der Allianz-Partnerorganisationen 11.06.2008 / Priority Initiative “Digital Information” by the Alliance of German Science Organisations 11.06.2008).

Stefan Rohde-Enslin

kulturerbe-digital.de: A New Information Platform for Digitisation – Potential and Perspectives

Zusammenfassung

kulturerbe-digital.de ist ein Projekt von EUBAM, der Arbeitsgruppe zu europäischen Angelegenheiten für Bibliotheken, Archive, Museen und der Denkmalpflege in Digitalisierungsfragen. Ziel des DFG geförderten Projektes ist der Aufbau einer Internetplattform, die als Multiplikator essentieller Hintergrundinformationen über die Digitalisierung von Kulturgut dienen soll. *kulturerbe-digital.de* wird über deutsche und europäische Förderprogramme und -möglichkeiten informieren, „best practice“-Empfehlungen bündeln und ein Verzeichnis aktueller und abgeschlossener Digitalisierungsprojekte aufbauen. So sollen Erfahrungen aus abgeschlossenen Digitalisierungsprojekten für die Planung künftiger Digitalisierungen nutzbar werden.

kulturerbe-digital.de will für die Digitalisierung Brücken bauen zwischen:

- Projekten, Organisationen und Institutionen in allen Kulturerbe-Sparten,
- Großen und kleinen Organisationen und Projekten,
- abgeschlossenen, laufenden und geplanten Projekten, sowie den relevanten Proponenten in Wissenschaft, Forschung und Politik und nicht zuletzt auch zwischen der deutschen und der europäischen Ebene.

Abstract

kulturerbe-digital.de is a project of EUBAM. EUBAM's focus is matters of European concern to libraries, archives, museums and other organisations charged with cultural heritage and digitisation. *kulturerbe-digital.de* is supported by the German Research Foundation (DFG). The main focus of this project is the development of an Internet presence: *kulturerbe-digital.de* will serve as a clearinghouse of essential background information on the digitisation of cultural assets. *kulturerbe-digital.de* will inform its users about German and European grant programmes and opportunities. It will bundle German and European "best practice" recommendations, enabling one-stop access. *kulturerbe-digital.de* will also provide a directory of ongoing and completed cultural digitisation projects. This will help to ensure that lessons learned in the past are at hand in the planning stage of future digitisation efforts.

kulturerbe-digital.de aims to build bridges between:

- the digitisation actors in different branches of cultural heritage;
- small and large projects;
- yesterday's, today's and tomorrow's digitisation projects and German professionals and their counterparts elsewhere in Europe.

Cultural heritage is going digital. Historical photographs are being scanned; nearly all new photographs are digitally created. Historical sound recordings are being digitised; modern audio recording devices are only able to produce digital sounds. Texts such as dissertations are created and published digitally, television and film are becoming digital. In place of letters e-mails are sent. These are only a few examples of ongoing changes that no one charged with cultural heritage can ignore. Many cultural heritage organisations actively pursue and support digitisation.

In addition to the transformation of content, with the digitisation of historical cultural heritage new possibilities of indexing, administering and presenting are being developed. Much of this occurs as a sideline, in the context of going about one's day-to-day work, while other ideas are realised in the context of formal projects.

The project *kulturerbe-digital.de* sees digitisation in this larger sense, one that, in addition to the media transformation we all are experiencing, encompasses the development, administration and publication of digital and digitised cultural heritage as well.

One of the tasks that *kulturerbe-digital.de* is taking up is to bring together information about different types of projects engaged in digitisation, in one way or another, and to make these projects accessible. The idea behind *kulturerbe-digital.de* is that with similar source materials (texts, graphics, maps, etc.) and similar tasks (collection, transformation, presentation, etc.), co-operation between projects ought to yield synergistic effects. At the same time, information about projects already completed should help those who are planning a digitisation project. Beyond concrete results, each digitisation project also creates a wealth of experience. In the form of a project report, all too often these experiences disappear in dusty drawers. Exemplary answers to the "how", "with what", and "how long", etc. are lost. Recovering these experiences and making them accessible is a further goal of *kulturerbe-digital.de*.

kulturerbe-digital.de is a project of the inter-ministerial Working Group for European matters affecting libraries, archives, museums and monument protection (*EUBAM* – <http://www.eubam.de>). The website of *kulturerbe-digital.de*, whose creation is funded by the *German Research Foundation (DFG)*, shall serve as a source of information for decision-makers in the spheres of policy and administration, assisting them in the development and subsequent adjustment of a national digitisation strategy. Digitisation projects, whether large or small in scale, require financial assistance. *kulturerbe-digital.de* shall also make it easier to find public and private sources of funding. Finally, *kulturerbe-digital.de* offers recommendations on digitisation practice in Europe and beyond. Much of this advice created at the European level actually remains unknown in Germany. By the same token, recommendations and expert knowledge from Germany are heeded only to a limited extent in Europe as a whole.

With its diverse offerings, *kulturerbe-digital.de* will offer new opportunities to connect digitisers, funding agencies, decision-makers and other interested parties. To achieve these goals we are establishing an internet-accessible database with regularly updated information. While the project to develop our Internet-platform will be created within a year, *kulturerbe-digital.de*'s content will remain current for years to come.

kulturerbe-digital.de is all about the digitisation of cultural heritage. The ".de" in the URL reveals that our principal focus is German cultural artefacts. In many places, at many websites and (finally) now also in political circles, the digitisation of cultural assets is increasingly regarded as an important topic; more and more, digitisation is viewed as an important (and thus worthy of funding) objective. For several years now, museums, archives and libraries have been digitising. What experiences have they gathered? In individual digitisation projects, practical advice has been worked out: the task now is to bundle these "guidelines" and to make them centrally accessible! Funds for digitisation schemes came (and come) from different sources, raising the question of how to match specific projects with corresponding funding sources. For these and other questions *kulturerbe-digital.de* aims to serve as a clearinghouse, relaying information that allows users to answer who, what, where and when questions raised in planning and implementing digitisation projects.

Funding for digitisation today also takes form in European contexts, as best practice recommendations are worked out at the European level. Via *kulturerbe-digital.de*, first-hand experience gathered in Paris, Vienna and elsewhere shall become visible and helpful to others. And, for that matter, why shouldn't colleagues in Rome and Madrid profit from lessons learned in Frankfurt and Duderstadt?

Stefan Rohde-Enslin

kulturerbe-digital.de:

kulturerbe-digital.de: Eine neue Informationsplattform zur Digitalisierung – Möglichkeiten und Herausforderungen

Kulturgut wird digital. Historische Fotografien werden gescannt, neue Fotografien fast nur noch digital erstellt. Historische Tonaufnahmen werden digitalisiert, moderne Aufnahmegeräte erzeugen häufig nur noch digitale Töne. Texte, Dissertationen beispielsweise, werden digital erstellt und publiziert. Fernsehen und Kino werden digital. Statt Briefen werden Mails verschickt. Das sind nur einige Beispiele für aktuell stattfindende Veränderungen, denen sich nicht entziehen kann wer Verantwortung für kulturelles Erbe trägt. In vielen kulturerbe-bewahrenden Institutionen wird der Prozess der Digitalisierung aktiv betrieben und unterstützt.

Ergänzend zur Transformation der Inhalte oder Informationen werden mit der Digitalisierung historischer Kulturgüter auch neue Möglichkeiten der Erschließung, Verwaltung und Präsentation entwickelt und gestaltet. Manches geschieht in dieser Hinsicht nebenbei – im Rahmen der laufenden Arbeit –, anderes wird in Form von Projekten verwirklicht.

Das Projekt www.kulturerbe-digital.de sieht „Digitalisierung“ in dem gerade genannten weiteren Sinn, der neben der reinen Medientransformation die Erschließung, Verwaltung und Veröffentlichung digitalen und digitalisierten Kulturerbes einschließt. Eine der Aufgaben, denen sich kulturerbe-digital stellt, ist es Informationen über alle Arten von Projekten, die „digitalisieren“, zusammenzutragen und zugänglich zu machen. Ausgangspunkt ist der Gedanke, dass bei ähnlichen Ausgangsmaterialien (Texten, Grafiken, Karten, etc.) und ähnlichen Aufgabenstellungen (Erschließung, Transformation, Präsentation, etc.), Kooperationen zwischen Projekten synergetische Effekte haben. Gleichzeitig soll die Information darüber, wer bereits Ähnliches durchgeführt hat, jenen helfen, die erst noch dabei sind, ein Digitalisierungsprojekt zu planen. Jedes Digitalisierungsprojekt schafft neben konkreten Ergebnissen einen Fundus von Erfahrungen, die allzu oft, in Projektberichte gekleidet, in Schubläden verschwinden. Beispielhafte Antworten auf Fragen nach dem Wie, Womit, Wie lange, ... gehen verloren. Diese Erfahrungen zugänglich, zu machen ist ein Ziel von *kulturerbe-digital.de*.

kulturerbe-digital.de ist ein Projekt der Arbeitsgruppe für *Europäischen Angelegenheiten für Bibliotheken, Archive, Museen und Denkmalpflege (EUBAM)*. Der Aufbau der Internet-Datenbank wird von der *DFG* gefördert. Es soll auch den Entscheidungsträgern in Politik und Verwaltung eine Informationsquelle sein, die bei der Entwicklung und Fortschreibung einer nationalen Digitalisierungsstrategie helfen kann. Digitalisierungsprojekte, ob im großen oder kleinen Maßstab, brauchen finanzielle Förderung. Die Suche nach öffentlichen oder privaten Förderern soll durch *kulturerbe-digital.de* ebenfalls vereinfacht werden.

Schließlich macht *kulturerbe-digital.de* Empfehlungen zur Digitalisierung, die auf europäischer oder internationaler Ebene erarbeitet wurden, zugänglich. Häufig sind diese Empfehlungen in Deutschland wenig bekannt. Gleiches gilt auch umgekehrt: Empfehlungen und Erfahrungen aus Deutschland werden europaweit nur wenig beachtet.

Mit seinen verschiedenen Angeboten wird *kulturerbe-digital.de* Möglichkeiten für eine Vernetzung von Digitalisierern, Förderern, Entscheidern und Interessierten schaffen. Um dieses Ziel zu erreichen, entsteht eine über das Internet zugängliche Datenbank, die langfristig aktuell gehalten werden soll, auch dann, wenn nach einem Jahr das Projekt „Aufbau einer Internetplattform www.kulturerbe-digital.de“ abgeschlossen ist.

Digitalisieren von Kulturgut – darum geht es bei kulturerbe-digital. Das „.de“ am Ende verrät, dass es vornehmlich um deutsches Kulturgut geht. An vielen Stellen, auf vielen

Internet-Seiten und (endlich) auch in politischen Kreisen – Kulturgut-Digitalisierung wird zunehmend als wichtiges Thema gesehen und mehr und mehr als bedeutendes (und damit förderwürdiges) Ziel angenommen. In Museen, Archiven und Bibliotheken wird schon seit einigen Jahren digitalisiert: Welche Erfahrungen wurden gemacht? In Projekten wurden Handlungs-Empfehlungen für die Digitalisierung erarbeitet: Diese „guidelines“ gilt es zu bündeln und zentral zugänglich zu machen! Das Geld für Digitalisierungsvorhaben kam (und kommt) aus den unterschiedlichsten Töpfen: Von welcher Stelle kann für welches Vorhaben Förderung erwartet werden? Für diese und weitere Fragen will *kulturerbe-digital.de* eine Anlaufstelle sein, will Informationen weitertragen, die helfen, Fragen des Wer, Wie, Was und Wo bei der Planung und Durchführung von Digitalisierungsvorhaben zu beantworten.

Förderung geschieht heute oft auch im europäischen Rahmen, „*best-practice*“-Empfehlungen werden auf dieser Ebene erarbeitet, Erfahrungen in Paris, Wien und anderenorts gesammelt. Auch das soll durch *kulturerbe-digital.de* sicht- und nutzbar werden. Und warum soll man in Rom und Madrid nicht von Erfahrungen aus Frankfurt und Duderstadt profitieren?

Markus Brantl

The Munich Digitisation Centre (MDZ)

Zusammenfassung

Das *Münchener Digitalisierungszentrum (MDZ)* der *Bayerischen Staatsbibliothek (BSB)* kann nunmehr auf 10 Jahre Erfahrungen mit der Retrodigitalisierung zurückblicken. Heute stellt das *MDZ* deutschlandweit das grösste Archiv mit retrodigitalem Content frei online zur Verfügung. Die ursprünglich mit Förderung der *Deutschen Forschungsgemeinschaft* als eines von zwei nationalen Digitalisierungskompetenzzentren errichtete Einrichtung ist seit 2001 nach Auslaufen der Förderung organisatorisch ein Referat der Abteilung für Bestandsaufbau und Erschließung und ein integraler Bestandteil des umfassenden Dienstleistungsspektrums der *BSB* als eine der großen europäischen Forschungsbibliotheken.

Der Vortrag bietet einen kurzen Überblick über das *MDZ* und dessen Rolle im retrodigitalen Bestandsaufbau der *BSB* an der Schwelle zur Massendigitalisierung.

Abstract

The *Munich Digitisation Centre (MDZ)* at the *Bavarian State Library (BSB)* can now look back upon ten years' experience with retro-digitisation. Today, the Centre offers Germany's largest archive of retrodigitised content, free-of-charge and available online. The Centre was initially funded by the *German Research Foundation (DFG)* as one of two national digital competency centres. Upon conclusion of the Foundation's support in 2001, the Centre joined the Library's collections department, becoming an integral part of the Library's service spectrum. The *Bavarian State Library* is one of Europe's pre-eminent research libraries. This presentation offers a brief overview of the *Munich Digitisation Centre*. It describes the Centre's role in the development of digital collections at the *Bavarian State Library* on the cusp of mass digitisation.

Ladies and Gentlemen, dear colleagues, in the following I would like to give you an insight into the work of the *Munich Digitisation Centre* at the *Bayerische Staatsbibliothek*. It has been in operation for 10 years now and currently concentrates primarily on mass digitisation and digital long-term preservation.

I would like to speak about the following points:

- The *Bayerische Staatsbibliothek*: Facts and Figures
- The *Münchener Digitalisierungszentrum* at a Glance
- Technical Infrastructure and Workflow
- Long-term Preservation
- Digitisation Strategy and Collection Development by Retrospective Digitisation

The *Bayerische Staatsbibliothek*: Facts and Figures

In its function as a universal library and internationally renowned research library the *Bayerische Staatsbibliothek* fulfils a worldwide literature supply task. It forms the "Virtual German National Library" together with the *Berlin State Library* and the *German National Library* in Frankfurt and Leipzig.

The library was founded in the year 1558 by Duke Albrecht V. Today – in June 2007 – the *Bayerische Staatsbibliothek* has:

- Over 680 employees;
- a budget volume of more than 43 million Euro;
- 9.1 million volumes, among them
 - 48,000 current journals,
 - 89,000 medieval manuscripts, representing the world's fourth-largest collection, and
 - 20,000 incunabula, placing the *BSB* in the international top position in this field;
- 1.7 million loan requests;
- Germany's most substantial collection of digital resources with over 19 million digital images and metadata as well as an overall volume of 34 terabyte archived data;
- 1.05 million visits to our reading rooms – a figure which proves that the physical form of the book is still in demand, even in the digital age.

The MDZ at a Glance

The *Munich Digitisation Centre (MDZ)*¹ at the *Bayerische Staatsbibliothek* was established in mid-1997, thus at the beginning of internet usage in Germany, as one of two national competence centres,² receiving start-up funding from the *German Research Foundation (DFG)* within the framework of the newly initiated support programme of “retrospective retro-digitisation of library collections”. For the German libraries this support programme was the starting point of a venture into a new world: In addition to their “classic” tasks of collecting, cataloguing, providing and preserving printed publications the libraries took over an additional and completely new function by digitising conventional, copyright-free works and processing them for online use: They became producers of freely accessible electronic contents. The *DFG* support programme was intended to help libraries and archives collect knowledge in this field and develop the relevant competencies.

After the start-up funding by the *DFG* ran out in the year 2001 a political and structural re-orientation of the *MDZ* was unavoidable. Our aim was to securely anchor the *MDZ* in the *Bayerische Staatsbibliothek* as the backbone of digital production. In 2003 these endeavours led to the integration of the *MDZ* in the Department of Acquisition, Collection Development and Cataloguing as the “*Digital Library*” division, thereby perpetuating its organisation and personnel structures. Today the *MDZ* has a staffing of five established posts and more than 25 posts financed with the aid of third-party funds for research and development projects provided by the *German Research Foundation*, the *Free State of Bavaria* and the *European Union*.

The tasks of the *MDZ* encompass among other things the development, testing and putting into operation of new and innovative products and processes all around the multi-faceted topic of the “*digital library*”, in particular:

- retrospective digitisation,
- long-term preservation of digital data and
- setting up virtual subject portals, such as for example the *Bayerische Landesbibliothek Online*³, or acting as the technical provider for the *German MichaelPlus* portal.

The *MDZ* has tested and employed new processes and methods in the generation, administration, storing, long-term preservation and access of digital objects on the internet. Furthermore, the *MDZ* assumes interdisciplinary tasks as a service provider, e. g. for digitisation technology and for development tasks in the framework of the open-source community in the digital library field. This is done in close cooperation with several partners, for

¹ <http://www.muenchener-digitalisierungszentrum.de>

² As the second centre the Göttingen Digitisation Centre was founded at the State- and University Library of Göttingen

³ <http://www.bayerische-landesbibliothek-online.de>

example in the academic field in the execution of digitisation projects with academic and non-academic research institutions, in the technical area in digital long-term preservation in cooperation with the Leibniz Computing Centre (LRZ) in Garching / Munich or with our colleagues from the Göttingen digitisation centre.

Today the *MDZ* has a collection of more than 34 terabyte digital master data, the largest collection on national scale. The data from over 75 projects implemented so far are not only freely accessible on the internet under <http://www.muenchener-digitalisierungszentrum.de> at any desired time, but they are also permanently digitally archived in cooperation with the Leibniz computing centre⁴ and kept retrievable at any desired time, e. g. within the framework of document delivery. The year 2007 brought the start of two major projects, and thereby the start of the real mass digitisation of copyright-free materials:

- The first project started in July 2007. It comprises the digitisation of around 37,000 printed works of the 16th century (1518 to 1600) from the German language area, using automatic book scanners (scanning robots).
- The second project of substantially larger dimensions is the cooperation (public-private partnership) of the *Bayerische Staatsbibliothek* with *Google*, which was made public in March 2007. You are probably curious about the details, which I am going to explain in detail later.

To be able to implement such large-scale projects efficiently and fast, a correspondingly powerful technical infrastructure is required, with a centrally computer-controlled, flexibly adjustable and scalable workflow.

Technical Infrastructure and Workflow

The retrospective digitisation comprises three main processes:

- capture (e. g. scanning, creation of full-text, storing, long-term preservation),
- indexing (e.g. creation of metadata) and
- access on the internet⁵.

The concrete implementation of the processes is generally dependent on the structure of the original and the goal definition in view of the presentation of the work on the internet.

Among the multitude of digitisation process steps – reaching from preparing the materials to providing them on the Internet up to storing them in the long term archive – I would like to introduce two process steps to you: the scanning and the software-controlled, largely automated central workflow of the *MDZ*. In the scanning the *MDZ* adheres to the following principles:

- The object is scanned once in the best possible quality, followed by the long-term preservation of the digital object (“do it once, do it right”).
- Seizing the idea of cross-media publishing, the digital objects are re-used in a broad variety of ways, e. g. for
 - Access via the internet,
 - producing reprint- or facsimile editions, e. g. of valuable manuscripts, and
 - for a later output on microfilm (via COM). We digitise first and subsequently produce a microfilm of the digital copies, if needed.
- We have automated all process steps as comprehensively as possible.

⁴ <http://www.lrz.de>

⁵ Cf. the "Technical Guidelines for Digital Cultural Content Creation Programmes" of 2004 under <http://www.minervaeurope.org/interoperability/technicalguidelines.htm> and the practical guidelines of the DFG for the digitisation of cultural content in the version 03/2007, which are available in an English version under http://www.dfg.de/forschungsfoerderung/formulare/download/12_151e.pdf.

For producing digital images the MDZ uses 14 state-of-the-art book scanners that are gentle on the books, among them two so-called "Graz camera tables" for materials in a critical state of conservation, and two automatic book scanners (scanning robots), enabling us to reproduce all conventional formats up to DIN A0. The MDZ is thus in a position to scan all types of monographs, periodicals, special materials, such as manuscripts and early prints, historical maps, etc., and materials that have to be treated with great care in a gentle manner, with a high resolution and at a 1:1 scale in relation to the original.

To render the execution of digitisation projects more efficient the MDZ has developed the ZEND (Zentrale Erfassungs- und Nachweisdatenbank; central recording and indexing database), a modular electronic publishing system for retrospective digitisation. The objective was that the production of every digital copy of the BSB's collections – no matter if it takes place internally or externally – should in principle follow a uniform, centrally controlled workflow.

The ZEND allows for the following:

- Documents can be processed and presented capturing different amounts of detail, i. e. in any form from image to full text,
- the complete production process can be mapped electronically,
- production takes place according to a standardized workflow taking account of different document types,
- digital images and full-text data can be supplied by an unlimited number of service providers and
- since the system is largely automated, a significant amount of time and expenses can be economized.

The ZEND is based on open-source software modules, such as LAMP (Linux, Apache, MySQL and PHP or Perl) and the XML publishing framework "Cocoon"⁶, thus integrating different components from the areas of document-, web-content- and workflow management to form an ideal overall production environment. The figure of 34 terabyte data which have been processed so far successfully, and above all without disturbance, shows that the ZEND has proven its worth in daily production work. Its open, modular, flexible and freely scalable architecture makes the ZEND suitable for the requirements of both scholarly retrospective digitisation with in-depth indexing and exclusively image-oriented, automated mass digitisation.

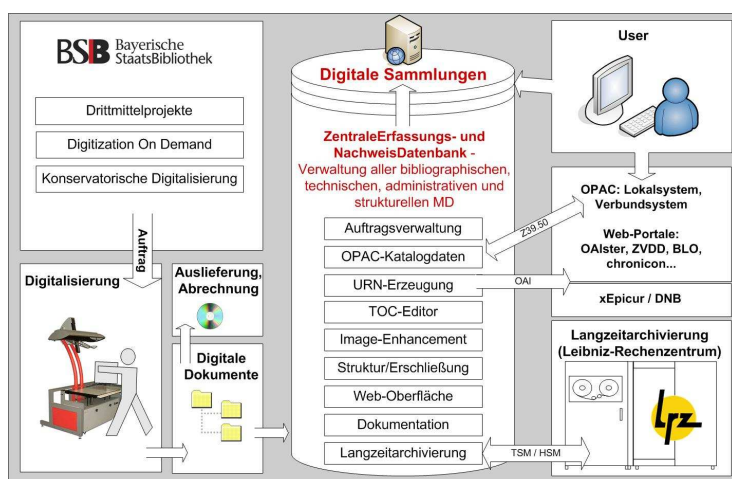


Figure 1:Workflow-Chart
The ZEND supports the XML document formats METS (Metadata Encoding and Transmission Standard) or TEI (Text Encoding Initiative).

Figure 1 gives you an overview of the workflow module centrally controlling the complete production process from preparation to scanning, storage and long-term preservation, online publication and cataloguing, up to the re-use and further use of the product:

In addition to storing the content of a digital object the ZEND also serves to administrate all administrative, bibliographic, structural and technical metadata. The ZEND thus provides a uniform basis for long-term preservation. The ZEND supports the XML document formats METS (Metadata Encoding and Transmission Standard) or TEI (Text Encoding Initiative).

⁶ See <http://cocoon.apache.org>

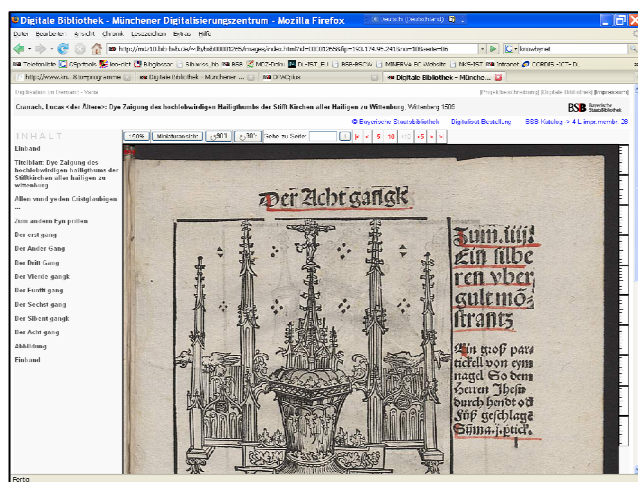


Figure 2: ZEND-Viewer

In addition, currently a variety of interfaces (Z39.50 and OAI with *Dublin Core* as well as *XEpicur-URN*) and tools (online ToC editor for recording the structural data of a digitised object, image enhancement, etc.) are implemented in the ZEND. The figure 2 shows the end product of a title presented in the ZEND viewer.

The left column contains the digital table of contents leading to the individual chapters of the book or to the images. Above the digital image of the book page the various functions of the ZEND viewer can be activated, such as thumbnail view, zooming, rotation, etc.

Digital long-term preservation

Digital long-term preservation and retrospective digitisation form an inseparable unit. This is particularly valid for the digitisation of rare objects of the cultural heritage, since the growing age of an original rapidly renders its digitisation increasingly difficult. These data have to be preserved.

The MDZ has occupied itself with digital long-term preservation since 1999. Since 2003 the MDZ has taken part in the Germany-wide, interdisciplinary cooperation project *nestor*⁷, which is supported by the Federal Ministry of Education and Research. The *German National Library*, the State- and University Library of Göttingen, the Humboldt University of Berlin, the Federal Archive, the Institute for Museum Research and the Distance-Learning University of Hagen are the *BSB's* partners in the *nestor* project. With *nestor* it was achieved to initiate a competence network which goes beyond libraries, archives and museums, and the public is made more strongly aware of the issue of digital long-term preservation.

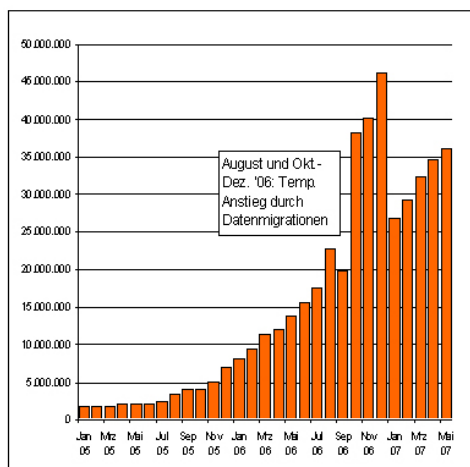


Figure 3: Archived data of the MDZ

Since 2004 the MDZ has cooperated intensively with the Leibniz Computing Centre. Beginning with the migration of the digitisation archive of the MDZ, which at that time contained 3,000 CD-ROMs, into the digital archive of the Leibniz Computing Centre, further projects followed, such as the automated archiving within the framework of the above-mentioned workflow with the ZEND or the setup of a pilot system – supported by the DFG – with an organisational-technical infrastructure for long-term preservation: the library archiving and access system (Bibliothekarisches Archivierungs- und Bereitstellungssystem – BABS)⁸.

Research currently also focuses on the issue of “trustworthy archives”, in which the MDZ plays an internationally leading and coordinating role together with the Humboldt University⁹.

⁷ Cf. the nestor project homepage: <http://www.langzeitarchivierung.de>.

⁸ <http://www.babs-muenchen.de>

⁹ See <http://edoc.hu-berlin.de/series/nestor-materialien/2006-8/PDF/8.pdf>

As shown in figure 3, about 34 terabyte data in 19 million files have accumulated in the digital long-term archive of the *MDZ* up to 1st June 2007.

The expansion of the new scanning technology at the *MDZ* by the introduction of the automatic book scanners which started in July 2007 is expected to bring a significant increase amounting to over 100 terabyte per year.

Digitisation Strategy and Collection Development by Retrospective Digitisation

It is the declared objective of the *Bayerische Staatsbibliothek* to successively digitise all copyright-free works – amounting to about 1.3 million titles, including the manuscripts and incunabula – and to make them freely available on the internet. To achieve this aim the *Bayerische Staatsbibliothek* relies on four support columns:

- third-party funded projects supported e. g. by the *German Research Foundation*, the *Free State of Bavaria* or the *European Union*,
- digitisation on demand,
- digitisation for conservation and
- the already mentioned public-private partnership with *Google*.

I would like to explain the four columns in detail.

Let's take the above-mentioned national mass digitisation project as an example for a third-party funded project. The *German Research Foundation* supports the project aiming at the digitisation of about 37,000 in German speaking countries printed works from the collections of the *Bayerische Staatsbibliothek*. The distinctive feature of the project is the use of a new generation of scanning robots allowing for the careful digitisation of materials of the 16th century, taking account of their state of conservation. The objective of the project is to digitise the above-mentioned number of titles and to make them available on the internet within 24 months. Digital copies are produced in colour and with a high resolution of 300 ppi, so that file sizes of about 30 megabyte can be generated per book page. The partner for long-term preservation of the expected 180 terabyte data is the Federal High-Performance Computing Centre in Munich-Garching, the Leibniz Computing Centre.

Third-party funded projects are mostly initiated taking account of or cooperating with the end users. The *MDZ* thus cooperates with a great number of academic and non-academic institutions in applying for third-party funded projects.

“*Digitisation on demand*” (*DoD*) was introduced as a new service offer after the old, analogously working photographic reproductions service was reorganized into a modern scanning centre. In this respect use is made of the advantage of digital reproduction technology, namely that images once produced can be re-used in many ways (and even to be plotted on microfilm). Thus the orders placed by users can be integrated in the concept of digital collection building. In the concrete case this means: The *BSB* subsidises digitisation by reduced prices and re-uses the digital copy. Even if only parts of works, e. g. of manuscripts, are ordered, frequently the whole work is digitised at the *BSB*'s own expense.

The production of digital copies for the purpose of conservation is based on the decision to prefer digitisation as primary secondary form to the production of microfilms. This does not mean that microfilms were abolished, but that a microfilm is now only produced upon corresponding demand. Thus the funds so far earmarked for the exclusive production of films for the purpose of conservation could be used for digital collection building. The pre-conditions for this are a concept and a technical platform for digital long-term preservation, which are both given in the case of the *Bayerische Staatsbibliothek*. As already mentioned above, the digital long-term preservation forms an integral part of the digitisation process at the *BSB*.

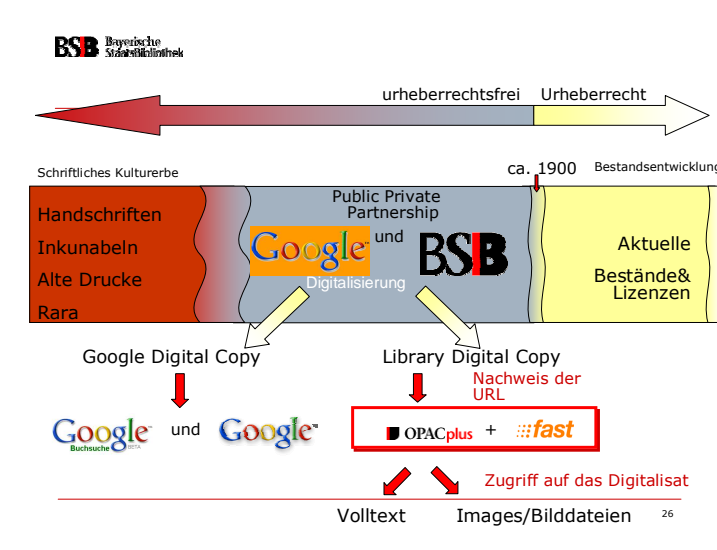


Figure 4: Cooperation with Google

be linked with the BSB's OPAC, with WorldCat etc., and will be freely accessible via the websites of the MDZ and the BSB. The BSB/MDZ will run this offer parallel to the offer by Google Booksearch. The graph shows that the cooperation will concentrate primarily on the BSB's collections from the 17th to the 19th century.

I would like to conclude my brief introduction of the MDZ here. Thank you for your attention!

The fourth column is the cooperation with Google entered into in March 2007. Since the cooperation agreement contains a non-disclosure clause, I must not mention any details. The agreement provides that Google will digitise more than one million copyright-free titles from the holdings of the BSB. The project will require several years, and the digitisation work will be carried out in the Free State of Bavaria. No direct costs are incurred for the BSB, and the library will obtain a copy of all digital master data generated by Google in the course of the digitisation process. The MDZ will subsequently take over these copies, process them, provide them on the internet and archive them in the long term. The copies will

Markus Brantl

Das Münchener Digitalisierungszentrum (MDZ)

Sehr geehrte Damen und Herren, liebe Kolleginnen und Kollegen, im Folgenden möchte ich Ihnen die Arbeit des Münchener Digitalisierungszentrums an der Bayerischen Staatsbibliothek, das seit nunmehr 10 Jahren besteht und sich derzeit primär mit der Massendigitalisierung und der digitalen Langzeitarchivierung beschäftigt, vorstellen.

Ich möchte dies in folgenden Punkten tun:

- Die Bayerische Staatsbibliothek in Zahlen
- Das Münchener Digitalisierungszentrum im Überblick
- Technische Infrastruktur und Workflow
- Langzeitarchivierung
- Digitalisierungsstrategie und retrodigitaler Bestandsaufbau

Die Bayerische Staatsbibliothek in Zahlen

Die Bayerische Staatsbibliothek nimmt in ihrer Funktion als Universal- und international renommierte Forschungsbibliothek einen weltweiten Versorgungsauftrag wahr. Zusammen mit der Berliner Staatsbibliothek sowie der Deutschen Nationalbibliothek in Frankfurt und Leipzig bildet sie die „Virtuelle Deutsche Nationalbibliothek“.

1558 wurde die Bibliothek durch Herzog Albrecht V. gegründet. Heute –Juni 2007 – hat die *Bayerische Staatsbibliothek*:

- über 680 Beschäftigte.
- ein Haushaltsvolumen von 43 Millionen Euro.
- 9,1 Millionen Bände, davon unter anderem
 - 48.000 laufende Zeitschriften,
 - 89.000 mittelalterliche Handschriften, dies ist der viertgrößte Bestand weltweit, und
 - 20.000 Inkunabeln; hier ist die *BSB* weltweit die Nummer eins.
- 1,7 Millionen Ausleihen.
- die bundesweit umfangreichste digitale Sammlung mit über 19 Millionen digitalen Images und Metadaten sowie einem Gesamtvolumen von 34 Terabyte archivierter Daten.
- 1,05 Millionen Lesesaalbesuche, eine Zahl, die belegt, dass die physische Form des Buchs – auch im digitalen Zeitalter – gefragt ist.

Das Münchener Digitalisierungszentrum im Überblick

Das *Münchener Digitalisierungszentrum (MDZ)*¹ an der *Bayerischen Staatsbibliothek* wurde Mitte 1997, also in der Anfangsphase der Internetnutzung in Deutschland, als eines von zwei nationalen Kompetenzzentren² durch eine Anschubfinanzierung der *Deutschen Forschungsgemeinschaft (DFG)* im Rahmen des neu initiierten Förderprogramms „*Retro-spektive Retrodigitalisierung von Bibliotheksbeständen*“ eingerichtet. Dieses Förderprogramm war für die deutschen Bibliotheken der Ausgangspunkt für den Start in eine neue Welt: Durch die Digitalisierung von konventionellen, urheberrechtlich freien Werken und deren Aufbereitung für ein Online-Angebot übernahmen die Bibliotheken – neben ihren „klassischen“ Aufgaben des Sammelns, Erschließens, Bereitstellens und Aufbewahrens gedruckter Veröffentlichungen – eine zusätzliche und völlig neue Funktion: die eines Produzenten frei verfügbarer elektronischer Inhalte. Das *DFG*-Förderprogramm sollte Bibliotheken und Archiven helfen, Wissen in diesem Bereich zu sammeln und einschlägige Kompetenzen aufzubauen.

Nach dem Auslaufen der *DFG*-Anschubfinanzierung im Jahre 2001 war eine politisch-strukturelle Neuorientierung des *MDZ* unausweichlich. Ziel war es, das *MDZ* als Rückgrat der Digitalisierungsproduktion fest in der *Bayerischen Staatsbibliothek* zu verankern. Diese Bestrebungen mündeten 2003 in eine organisatorisch-personelle Verstetigung durch die Verankerung des *MDZ* in der Abteilung Bestandsaufbau und Erschließung als Referat „Digitale Bibliothek“. Heute verfügt das *MDZ* über eine Personalausstattung mit fünf Planstellen und über 25 Stellen, die aus Drittmitteln für Forschungs- und Entwicklungsprojekte seitens der *Deutschen Forschungsgemeinschaft*, des *Freistaats Bayern* und der *Europäischen Union* finanziert werden.

Die Aufgaben des *MDZ* umfassen unter anderem die Entwicklung, Erprobung und Inbetriebnahme neuer und innovativer Produkte und Prozesse rund um das facettenreiche Thema „Digitale Bibliothek“, so besonders für

- die Retrodigitalisierung,
- die Langzeitarchivierung digitaler Daten und
- den Aufbau virtueller Fachportale, wie zum Beispiel die Bayerische Landesbibliothek Online³ oder als technischer Provider für das *MICHAEL*Plus-Portal Deutschland.

¹ <http://www.muenchener-digitalisierungszentrum.de>

² Als zweites Zentrum wurde das Göttinger Digitalisierungszentrum an der Staats- und Universitätsbibliothek Göttingen eingerichtet.

³ <http://www.bayerische-landesbibliothek-online.de>

Dabei wurden und werden neue Verfahren und Methoden bei der Erstellung, Verwaltung, Speicherung, Langzeitarchivierung und Bereitstellung digitaler Objekte im Internet erprobt und eingesetzt. Des Weiteren übernimmt das *MDZ* übergreifende Aufgaben als Serviceprovider, z. B. für Digitalisierungstechnik und für Entwicklungsaufgaben im Rahmen der Open-Source-Community im Bereich Digitale Bibliothek. Dies geschieht in enger Kooperation mit Partnern, so etwa im wissenschaftlichen Bereich bei der Durchführung von Digitalisierungsprojekten mit universitären und außeruniversitären Forschungsinstitutionen, im technischen Bereich bei der digitalen Langzeitarchivierung in Zusammenarbeit mit dem Leibniz-Rechenzentrum (LRZ) in München-Garching oder mit unseren Kollegen vom Göttinger Digitalisierungszentrum.

Heute verfügt das *MDZ* über einen Bestand von über 34 Terabyte an digitalen Masterdaten und nimmt damit national den ersten Platz ein. Die Daten aus mittlerweile über 75 realisierten Projekten sind nicht nur jederzeit frei im Internet unter abrufbar (<http://www.muenchener-digitalisierungszentrum.de>), sondern auch in Kooperation mit dem Leibniz-Rechenzentrum⁴ auf Dauer digital archiviert und für die jederzeitige Rückholung, z. B. im Rahmen der Dokumentlieferung, abrufbar gehalten. Das Jahr 2007 brachte und bringt mit dem Start von zwei Großprojekten den Einstieg in die eigentliche Massendigitalisierung urheberrechtsfreier Materialien:

- Das erste Projekt wird im Juli 2007 starten. Es umfasst die Digitalisierung von rund 37.000 Drucken aus dem deutschen Sprachraum des 16. Jahrhunderts (1518-1600) unter Einsatz von automatischen Buchscannern (Scanrobotern).
- Das zweite von den Dimensionen deutlich größere Projekt ist die im März 2007 offiziell bekannt gegebene Zusammenarbeit (Public-Private-Partnership) der *Bayerischen Staatsbibliothek* mit *Google*, auf deren Details Sie sicher gespannt warten und auf die ich später noch genauer eingehen werde.

Um solche Großprojekte effizient und rasch realisieren zu können, ist eine entsprechend leistungsfähige technische Infrastruktur mit einem EDV-technisch zentral gesteuerten, flexibel anpassbaren und skalierbaren Workflow notwendig.

Technische Infrastruktur und Workflow

Die Retrodigitalisierung besteht aus drei Hauptprozessen:

- Erfassung (z. B. Scannen, Texterfassung, Speicherung, Langzeitarchivierung),
- Erschließung (z. B. Anreicherung mit Metadaten) und
- Bereitstellung im Internet⁵.

Die Prozesse sind in ihrer konkreten Gestalt stets abhängig von der Struktur der Originalvorlage und der Zieldefinition hinsichtlich der Präsentation des Werks im Internet.

Im Folgenden will ich Ihnen – aus der Vielzahl der Prozessschritte beim Digitalisieren, die von der Vorbereitung der Materialien über die Bereitstellung bis zur Langzeitarchivierung reichen – zwei Prozesse, das Scannen und den softwaregestützten, weitgehend automatisierten zentralen Workflow, des *MDZ* vorstellen. Beim Scannen folgt das *MDZ* folgenden Prinzipien:

- Das Objekt einmal in bestmöglicher Qualität scannen und anschließend langzeitarchivieren („do it once, do it right“).

⁴ <http://www.lrz.de>

⁵ Vgl. die „Technischen Richtlinien für Programme zur Erstellung von digitalen Kulturinhalten“ von 2004 unter <http://www.minervaeurope.org/interoperability/technicalguidelines.htm> und die Praxisregeln der DFG zur Digitalisierung von Kulturgut in der Fassung 03/2007 unter http://www.dfg.de/forschungsfoerderung/formulare/download/12_151.pdf, die auch in einer englischen Version zugänglich sind: http://www.dfg.de/forschungsfoerderung/formulare/download/12_151e.pdf.

- Die digitalen Objekte vielfältig, ganz im Sinne des Cross-Media-Publishing, nachnutzen, so z.B., für
 - die Internet-Bereitstellung,
 - die Herstellung von Reprint- oder Faksimile-Ausgaben, z. B. von wertvollen Handschriften, und
 - die Schutzverfilmung von Objekten. Wir digitalisieren zuerst und belichten die Digitalisate anschließend, wenn benötigt auf Mikrofilm aus (Computer output to microform). Hierzu werde ich später noch etwas sagen.
- Alle Prozessschritte, soweit möglich, automatisieren.

Zur Herstellung digitaler Images setzt das MDZ mittlerweile 14 modernste, buchschonende Aufsichtsscanner, darunter zwei sogenannte „Grazer Kameratische“ für konservatorisch schwierige Materialien, und zwei automatische Buchscanner (Scanroboter) ein, welche die Reproduktion aller gängigen Vorlagenformate bis DIN A0 ermöglichen. So ist das MDZ in der Lage, alle Arten von Monographien, Zeitschriften, Sondermaterialien, wie etwa Handschriften und alte Drucke, historische Karten etc., und konservatorisch schwierigste Materialien buchschonend, hochauflösend und im Format 1:1 in Bezug auf das Originalformat zu scannen.

Zur effizienteren Durchführung von Digitalisierungsprojekten hat das MDZ seit 2003 die ZEND (Zentrale Erfassungs- und Nachweisdatenbank), ein modulares Electronic Publishing System für die Retrodigitalisierung, entwickelt. Das Ziel dabei war, dass jedes Digitalisat, das intern oder extern aus dem Bestand der BSB angefertigt wird, grundsätzlich einem einheitlich vorgegebenen, zentral gesteuerten Workflow folgen soll.

Die ZEND erlaubt:

- die Verarbeitung und Präsentation von Dokumenten in unterschiedlicher Erschließungstiefe, also vom Bild bis zum Volltext,
- die elektronische Abbildung des gesamten Produktionsprozesses,
- die Produktion mit einem standardisierten Workflow für diverse Dokumenttypen,
- die Zulieferung von digitalen Bildern und Volltextdaten durch eine unbegrenzte Anzahl von Dienstleistern sowie
- eine deutliche Zeit- und Kostenersparnis durch weitgehende Automatisierung.

Die ZEND basiert auf Open-Source-Software-Modulen, wie LAMP (Linux, Apache, MySQL und PHP bzw. Perl) und das XML-Publishing-Framework „Cocoon“ (<http://cocoon.apache.org/>), und vereint so verschiedene Komponenten aus dem Dokumenten-, Web-Content- und Workflow-Management zu einer idealtypischen Gesamtproduktionsumgebung. Im täglichen Produktionseinsatz hat sich die ZEND, wie die Zahl der bisher rund 34 Terabyte erfolgreich und vor allem störungsfrei verarbeiteten Dateien zeigt, bestens bewährt. Durch ihre offene, modulare, flexible sowie frei skalierbare Architektur ist die ZEND gleichermaßen für die Anforderungen wissenschaftlicher Retrodigitalisierung mit tieferer Erschließung als auch für die zunächst rein bildorientierte, automatisierte Massendigitalisierung geeignet.

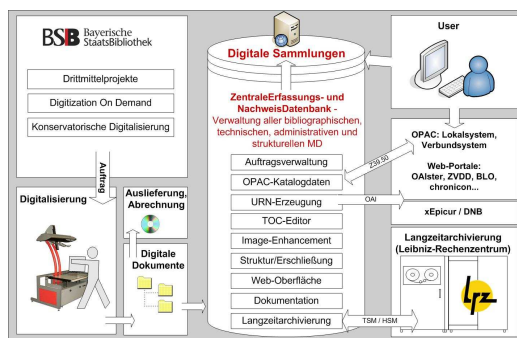


Abbildung 1:Workflow-Chart

Das Workflow-Modul, das zentral den Gesamtproduktionsprozess von der Vorbereitung über das Scannen, Storage und Langzeitarchivierung, Online-Publikation und Nachweis bis hin zur Wieder- und Weiterverwertung steuert, zeigt folgende Grafik im Überblick:

In der ZEND werden alle administrativen, bibliographischen, strukturellen und technischen Metadaten nebst dem Inhalt eines digitalen Objekts verwaltet. So ist durch die ZEND eine einheitliche Basis für die Langzeitarchivierung

gegeben. Die ZEND unterstützt die XML-Dokumentformate METS (Metadata Encoding and Transmission Standard) oder TEI. Daneben sind derzeit verschiedene Schnittstellen (Z39.50 und OAI für Dublin Core sowie XEpicur-URN) und Tools (Online-ToC-Editor zur Erfassung der Strukturdaten eines digitalisierten Objekts, Image Enhancement etc.) in der ZEND implementiert. Der folgende Screenshot zeigt das Endprodukt eines Titels, der im ZEND-Viewer präsentiert wird.

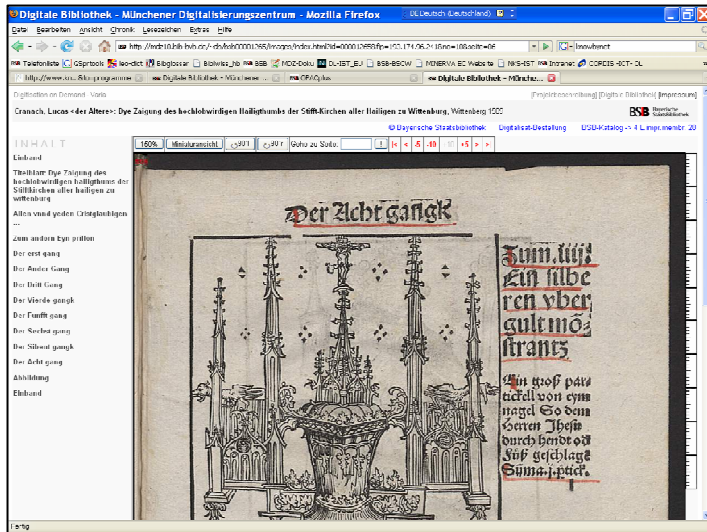


Abbildung 2: ZEND-Viewer

In der linken Spalte befindet sich das digitale Inhaltsverzeichnis, das zu den einzelnen Kapiteln des Buchs bzw. zu den Abbildungen führt. Über dem digitalen Bild der Buchseite können die verschiedenen Funktionen des ZEND-Viewers, wie Miniaturansicht, Vergrößerung, Drehen etc., aufgerufen werden.

Langzeitarchivierung

Die digitale Langzeitarchivierung und Retrodigitalisierung bilden eine untrennbare Einheit, dies in besonderem Maße für die Digitalisierung von älterem Kulturgut, da der Aufwand bei der Digitalisierung mit steigendem Alter einer Originalvorlage rapide in die Höhe schnellte. Es gilt, diese Daten über die Langzeitarchivierung zu erhalten.

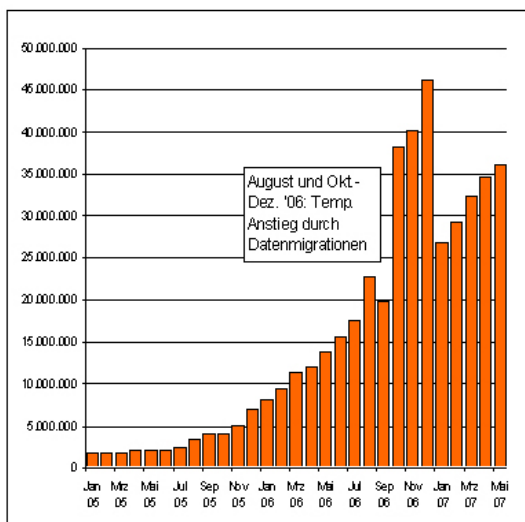


Abbildung 3: Archivierte Daten des MDZ

Seit 1999 beschäftigt sich das MDZ mit der digitalen Langzeitarchivierung; seit 2003 ist das MDZ Partner im bundesweiten, spartenübergreifenden Kooperationsprojekt *nestor*⁶, das vom Bundesministerium für Bildung und Forschung gefördert wird. Partner in *nestor* sind die *Deutsche Nationalbibliothek*, die *Staats- und Universitätsbibliothek Göttingen*, die *Humboldt-Universität Berlin*, das *Bundesarchiv*, das *Institut für Museumsforschung* und die *Fernuniversität Hagen*. Mit *nestor* ist es gelungen, ein über Bibliotheken, Archive und Museen hinausgehendes Kompetenznetzwerk zu initiieren und die Fragestellungen der digitalen Langzeitarchivierung stärker in das Bewusstsein der Menschen zu bringen.

Seit 2004 kooperiert das MDZ intensiv mit dem Leibniz-Rechenzentrum. Beginnend mit der Migration des damals 3.000 CD-ROM umfassenden Digitalisierungsarchivs des MDZ in das digitale Archiv des Leibniz-Rechenzentrums folgten weitere Projekte, wie die automatische Archivierung im Rahmen des oben genannten

⁶ Vgl. die Homepage des Projekts *nestor*: <http://www.langzeitarchivierung.de>.

Workflows mit der *ZEND* oder der Aufbau eines Pilotsystems – unterstützt von der *DFG* – mit einer organisatorisch-technischen Infrastruktur zur Langzeitarchivierung: das Bibliothekarische Archivierungs- und Bereitstellungssystem (BABS)⁷.

Als aktueller Forschungsschwerpunkt sei noch das Thema „Vertrauenswürdige Archive“ genannt, welches das *MDZ* federführend und gemeinsam mit der Humboldt-Universität international betreut⁸.

Im digitalen Langzeitarchiv des *MDZ* sind, wie die folgende Grafik zeigt, bis zum 1. Juni 2007 rund 34 Terabyte in 19 Millionen Dateien aufgelaufen.

Mit dem Ausbau der neuen Scantechnologie am *MDZ* durch die Einführung der automatischen Buchscanner ab Juli 2007 wird eine signifikante Steigerung von mehr als 100 Terabyte im Jahr erwartet.

Digitalisierungsstrategie und retrodigitaler Bestandsaufbau

Es ist das erklärte Ziel der *Bayerischen Staatsbibliothek*, alle urheberrechtsfreien Werke, es sind dies rund 1,3 Millionen Titel inklusive der Handschriften und Inkunabeln, sukzessive zu digitalisieren und frei im Internet zur Verfügung zu stellen. Zur Erreichung des Ziels stützt sich die *Bayerische Staatsbibliothek* mittlerweile auf vier Säulen:

- Drittmittel-Projekte mit Förderung z. B. durch die *Deutsche Forschungsgemeinschaft*, den *Freistaat Bayern* oder die *Europäische Union*,
- Digitization on Demand,
- „Konservatorische Digitalisierung“ und
- die schon genannte Public-Private-Partnership mit *Google*.

Lassen Sie mich im Folgenden die vier Säulen näher erläutern.

Als Beispiel für ein Drittmittelprojekt sei das bereits erwähnte nationale Massendigitalisierungsprojekt genannt. Das Projekt, das die Digitalisierung von rund 37.000 im deutschen Sprachraum erschienenen Drucken aus dem Bestand der *Bayerischen Staatsbibliothek* zum Ziel hat, wird durch die *Deutsche Forschungsgemeinschaft* gefördert. Die Besonderheit des Projekts ist der Einsatz einer neuen Generation von Scanrobotern, die eine konservatorisch schonende Digitalisierung von Materialien des 16. Jahrhunderts erlauben. Ziel dieses Projekts ist es, in 24 Monaten die oben genannte Zahl von Titeln zu digitalisieren und im Internet bereitzustellen. Digitalisiert wird mit einer hohen Auflösung von 300 ppi und in Farbe, sodass pro Buchseite Dateigrößen von rund 30 Megabyte entstehen werden. Partner für die Archivierung der zu erwartenden 180 Terabyte an Daten ist das Bundeshöchstleistungsrechenzentrum in München-Garching, das Leibniz-Rechenzentrum.

Drittmittelprojekte entstehen meist unter Einbeziehung bzw. in Kooperation mit den Endusern. So kooperiert das *MDZ* mit einer Vielzahl von universitären und außeruniversitären Institutionen bei der Beantragung von Drittmittelprojekten.

„Digitization on Demand“ (DoD) entstand als neues Service-Angebot nach der Umstellung der alten, analog arbeitenden Fotostelle auf ein modernes Scanzentrum. Hier wird der Vorteil der digitalen Reproduktionstechnik, einmal produzierte Images vielfältig nachnutzen zu können, genutzt. So können die Benutzeraufträge in das Konzept des digitalen Bestandsaufbaus integriert werden, konkret heißt dies: Die *BSB* subventioniert die Digitalisierung durch reduzierte Preise und nutzt das Digitalisat nach. Auch wenn nur Teile bestellt z.B. von Handschriften, so wird oft in Eigenleistung gleich das gesamte Werk digitalisiert.

Die konservatorische Digitalisierung fußt auf der Entscheidung, der Digitalisierung als primäre Sekundärform den Vorzug vor dem Mikrofilm zu geben. Dies bedeutet nicht, dass der Mikrofilm abgeschafft wurde, sondern dass ein Mikrofilm nur noch nach Bedarf erzeugt

⁷ <http://www.babs-muenchen.de>

⁸ Vgl. <http://edoc.hu-berlin.de/series/nestor-materialien/2006-8/PDF/8.pdf>

wird (z.B. im COM-Verfahren). So konnten die bisher für die reine Sicherheitsverfilmung vorgesehenen Mittel für den Zweck des digitalen Bestandsaufbaus genutzt werden. Dies setzt, wie im Fall der *Bayerischen Staatsbibliothek* gegeben, ein Konzept und eine technische Plattform für die digitale Langzeitarchivierung voraus. Die digitale Langzeitarchivierung ist, wie bereits erwähnt, ein integraler Bestandteil des Digitalisierungsprozesses an der *BSB*.

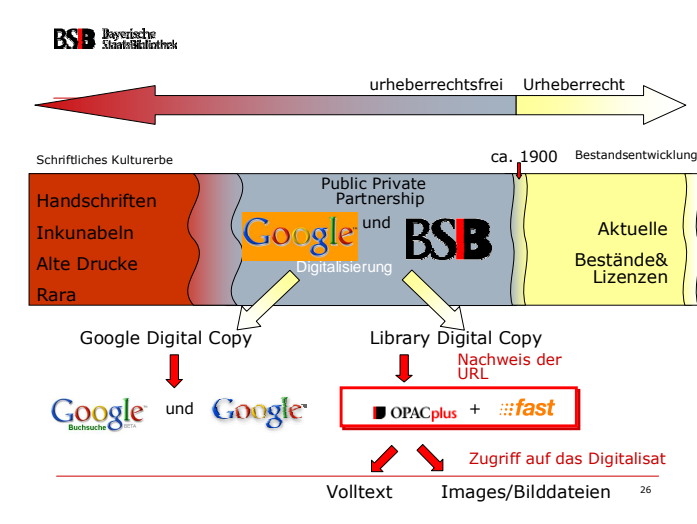


Abbildung 4: Kooperation mit Google

Als vierte Säule ist die im März eingegangene Kooperation mit *Google* zu nennen. Die Kooperationsvereinbarung, von der ich Ihnen keine Details nennen kann, da diese als Geheimhaltungsvertrag abgeschlossen wurde, sieht vor, dass *Google* mehr als eine Million urheberrechtsfreie Titel aus dem Bestand der *BSB* digitalisieren wird. Das Projekt wird mehrere Jahre in Anspruch nehmen, und die Digitalisierung wird im *Freistaat Bayern* durchgeführt werden. Der *BSB* entstehen keine direkten Kosten, und sie bekommt alle digitalen Master-Daten, die durch *Google* im Digitalisierungsprozess erstellt werden, in Kopie. Diese Kopien werden anschließend durch das

MDZ übernommen, verarbeitet, im Internet bereitgestellt und langzeitarchiviert.

Sie werden mit dem OPAC der *BSB*, mit WorldCat etc. verlinkt sowie über die Websites des *MDZ* und der *BSB* frei zugänglich sein. Dieses Angebot von *BSB/MDZ* wird parallel zu dem Angebot von *Google* Buchsuche betrieben werden. Die Zusammenarbeit mit *Google* erläutert die Grafik. Sie macht deutlich, dass sich die Kooperation vor allem auf den Bestand der *BSB* vom 17. bis zum 19. Jahrhundert konzentrieren wird.

Damit endet die kurze Vorstellung des *MDZ*. Ich danke für Ihre Aufmerksamkeit!

Karl-Heinz Lampe

From a Single Project to Knowledge Mapping – Manufacturing Knowledge

Zusammenfassung

Wissensgenerierung ist mehr als nur die Ablage digitaler Daten. Am Beispiel des wissenschaftlichen Projekts „Deutsche Orthopteren Sammlungen“ (DORSA) wird gezeigt, wie mit Hilfe von Ontologien wissenschaftlicher Mehrwert generiert werden kann.

Der Begriff Ontologie stammt aus der Philosophie und bezeichnet die Lehre vom Seienden. Heute wird der Begriff auch in der Informatik genutzt im Sinne einer formalen Spezifizierung von semantischen Konzepten. Durch die formale Spezifizierung können unterschiedliche semantische Konzepte miteinander verbunden werden, was das Verständnis komplexer und vernetzter Systeme ermöglicht. Ontologien können damit die Funktion transdisziplinärer wissenschaftlicher Vermittler spielen; zwischen und innerhalb natur- und kulturwissenschaftlicher Forschungsdisziplinen.

Aufsammlungs- und Beobachtungs-Ereignisse, zum Beispiel, besitzen eine semantische Beziehung, die in der ontologischen Formalisierung sichtbar wird. Unterschiedliche Vorgänge lassen sich in diesem Fall auf ein Modell zurückführen und damit in Beziehung setzen. Am Beispiel eines wohl dokumentierten, genadelten Insekts und der damit verbundenen Wissensbeziehungen wird gezeigt, dass Ontologien nicht nur das Ergebnis einer wissenschaftlichen Aktivität – wie die Neubeschreibung einer Art – beschreiben, sondern auch den Weg, der zu dem Ergebnis geführt hat.

Abstract

The creation of knowledge entails more than the delivery of digital information. Using the example of a scientific project, DORSA (a virtual museum which offers access to the Orthoptera in German museum collections), this presentation will show how, with the assistance of ontologies, scientific "added value" may be generated.

With roots in philosophy, the concept of ontology is now used in the information sciences to describe a formal specification of semantic concepts. Formal specification permits different semantic concepts to be connected to one another, allowing the integration of complex systems of information. Ontologies may in this way play the role of transdisciplinary scholarly interpreters, between and among research disciplines in the natural and cultural sciences.

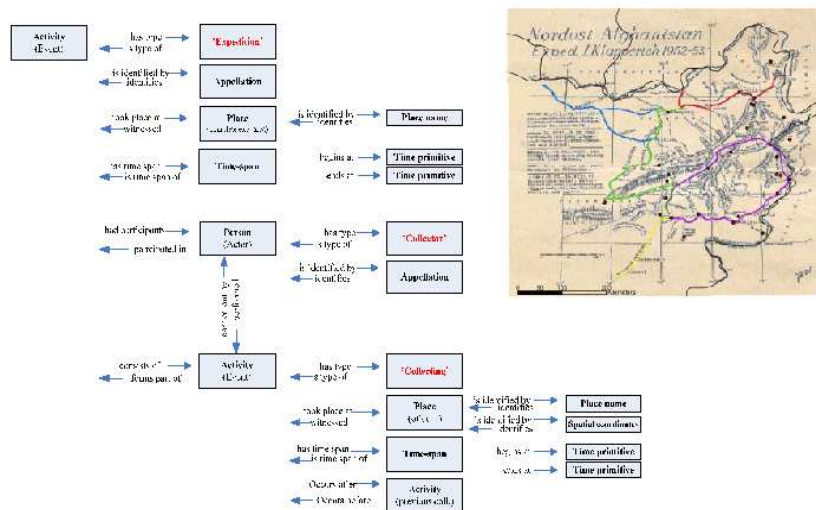
Collection and observation events, for example, possess a semantic relationship that becomes visible through ontological formalisation. Individual procedures can be traced back to a model and, for this reason, be set in relation to one another. Tracing the example offered by one well-documented insect specimen - and various relationships linked to this example - this presentation will show that ontologies not only describe the result of a scientific activity, such as the new description of a species. Ontologies also illustrate the way one comes to the result.

What kind of data do natural history museums have?

Natural History museums are archives of biodiversity. They house millions of specimens collected in space and time, providing first-hand information about geographical and historical presence of organisms, their morphological and other specific characters (anatomical,

People usually see what they want to see. Showing a pinned insect such as in the centre of Figure 1 and asking: "What is it?" the answer often is: "A bee!". A pinned insect, however, obviously is a 'Man-made Object'. It consists of a biological object, the bee, and an information carrier; a pin carrying labels as documents.

For ecologists the white label primarily documents a collecting event, for economists the same label documents an acquisition event and for lawyers it might document a transfer of custody (from nature to a collection). Therefore the white label documents three *CIDOC-CRM* specific events (multiple instantiation in E7, E8 and E10). It also shows the exact location of a particular person at a certain time. Similarly, the red label documents a type creation of a new species and a determination or type assignment. The type creation event has created a document, a species description, and a type that is a new class in a philosophical sense with the appellation *Colletes alini* Kuhlmann 1999. The bee specimen supported the type creation in the taxonomic role of a holotype. Type creations in other disciplines as for example archaeology are often based on the same or similar methodological approach; the terminology, however, is different.



Of course this ontology or knowledge representation can easily be extended (e.g. object identifier and store within a collection in Figure 1) or refined to a more detailed description such as a collecting event or even an expedition (Figure 2). The collecting event is an activity carried out by an actor (which is a person, the "collector", identified by an appellation. It took place at a gathering site (identified either by a

Figure 2: Expedition- and collecting events in terms of *CIDOC-CRM* (identified either by a Place name and/or spatial coordinates) and lasted for a certain time span. An expedition is an activity consisting of continuing collecting activities and participating persons which carry out these activities etc. (Figure 2). In a similar manner collector's itineraries can be mapped. A mapping of semantic relationships in this way makes domain specific information units understandable to everyone. Secondly, by replacing the term "collector" by "observer/photographer" or "painter/drawer" in Figure 2, the same model characterizes observation, photographing and painting or drawing events (instead of a collecting event). Cultural museums for example often house old paintings etc. depicting the animated world in times when no biological voucher specimens were collected, labelled and deposited in a museum of natural history, Digitizing these data in a proper way will complement our knowledge of the history of the animated world. What a fascinating *new* information source for biologists.

Ontologies like these can be used to filter out one's own area of expertise and to identify additional or overlapping sources of information normally covered by other domains.

For example specimen based databases deal with real world objects. They are focused on primary biodiversity data documented on specimen labels; determination, gathering site, date collected, collector's name etc. Secondly, taxonomic authority files, such as 'Fish Base', the 'Hymenoptera Name Server' or the 'Orthoptera Species File', deal with taxonomic names that stand for human concepts, and are concerned with the validity of the respective names. They are primarily focused on published data. Thirdly, geographic authority

files, such as the 'Getty Thesaurus' or the 'Alexandria Digital Library Gazetteer' (ADL), deal with geographic names in a geophysical and geopolitical context. Last but not least other authority files deal with the biography of persons and their identity.

Ontologies can be used as a best-practise guide to improve already existing databases by means of dynamic linking, multiple verification, and semantic enrichment as well as developing new tools via web services by using *GRID* technology. Furthermore it can help to identify other formerly hidden or even obscure information sources from various scientific and scholarly domains.

Improvement of data quality and generating scientific extra value through dynamic linking, multiple verification and semantic enrichment

Within the *Digitised Orthoptera Specimen Access (DORSA)* project, initiated by Klaus Riede and me, a virtual museum of Orthoptera type specimens (grasshoppers, locusts, katydids, crickets) housed in major German museums was created. Thereby the *Orthoptera Species File (OSF)*, a global taxonomic authority file, served and still serves as a taxonomic backbone. The *OSF* was build and is maintained in Philadelphia/USA (<http://osf2.orthoptera.org>). *DORSA* contains full information about 16,000 specimens (incl. 2,300 primary types and 6,700 secondary types). 30,000 images and 11,000 sound records are linked to their respective specimens. Many of the specimens linked to the sound records belong to hitherto unknown species. They can be seen as types of tomorrow. The *DORSA* virtual museum is available through the SYNTAX database infrastructure (<http://www.biologie.uni-ulm.de/syntax>) and major knowledge portals such as *BIOCASE* (<http://www.biocase.org>) or *GBIF* (<http://www.gbif.org>). Figure 3 shows a screenshot of specimen-based multimedia information units, incl. sound-files. The *DORSA* web-based repository provides sufficient detailed information, including diagnostic features (e.g. genitalia), to allow for example taxonomists to narrow down loan requests. As mentioned earlier the *OSF* is a global taxonomic authority file, mainly based on published information. It provides the user with information about the validity of a given taxon name, the source of the original description, the depository place of type specimens and so on. Reciprocal dynamical linking between *DORSA* and *OSF* allow direct access to the respective data, e.g. for validity checks of a given taxon (Figure 4). As you probably know, the validity of a given taxon can change in time.

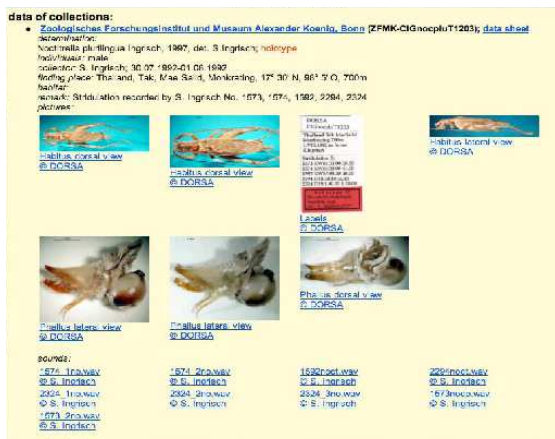


Figure 3: Screenshot of specimen based multimedia information units, including sound files (*DORSA* – *Digitised Orthoptera Specimen Access*)

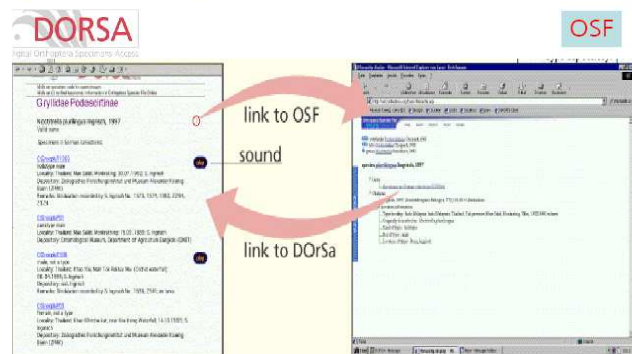



Figure 4: Reciprocal linking between *DORSA* and *OSF*

OSF authority files allow verification of published type data against our *DORSA* specimen data and vice versa (Table 1). Our *DORSA* taxonomist, Sigfrid Ingrisich, detected about 25% of discrepancies between the two sources, such as 545 primary types, which had not been listed in *OSF* or 39 unlabeled, newly recognized primary type specimens within German museums. Besides that, exact data about missing type specimens allow initiation of effective searches to fill these gaps, for example by designating neotypes.



museum	presumption	museum data	validity check				
	taxa with primary types in OSF *	taxa with primary types checked	type data in OSF not confirmed in museum	primary types lost: taxa	primary types not listed in OSF	no taxon entry found in OSF	unlabeled, newly recognized prim. types
Berlin	1093	1272	15	1	221	68	12
Eberswalde	45	79		1	34		
Dresden	66	117	6		55	10	7
Hamburg	142	135	5	55	43	1	
Halle	38	55	4		25	4	16
Bonn	6	44			12	13	
Frankfurt	82	151			39	5	1
Stuttgart	47	112			55	5	3
München	27	54		1	21	1	
Sum	1546	2019	30	57	545	107	39

Table 1: Verification of primary types in *OSF* against *DORSA*

consisted in extraction of basic song parameters, such as carrier frequency and pulse rate (Figure 5). This software module was applied to the entire *DORSA* sound repository and the calculated parameters were then annotated into new columns of the sound file table (semantic enrichment). As a result, important sound parameters now complement the respective sound files. Sound parameters can be evaluated, or you can search for certain parameters by standard SQL queries, and retrieve the respective sound files from the repository. Even complex data-mining is feasible, such as: "Are there any Orthoptera (in our database), from Ecuadorian lowland forests, and singing higher than 10 kHz, etc."

Because gathering sites of *DORSA* specimens became geo-referenced by adding latitude and longitude information, type localities can be visualised by linking to remote available map-servers such as the 'Canadian Biodiversity Information Facility' (*CBIF*) map server or *Google Earth*. Using 'Geographical Information Systems' (*GIS*) distribution maps can be created and intersected with layers such as political boundaries or vegetation zones or recent distribution of tropical rain forests etc. Extraction of special geographical information units such as 'county' or 'district' data allow a geographical semantic enrichment of the respective *DORSA* files; quite similar to the already mentioned enrichment of sound files.

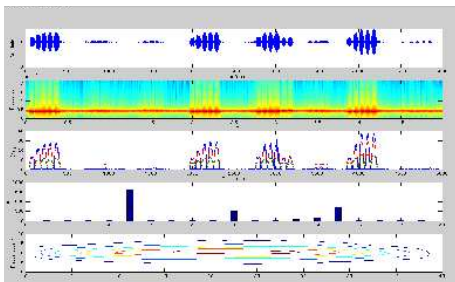


Figure 5: Semantic enrichments of specimen based databases (*DORSA*): Extraction of sound parameters by using MatLab Software: carrier frequency and pulse rate; in cooperation with: Dept. of Neuroinformatics, Ulm University; PhD thesis C. Dietrich

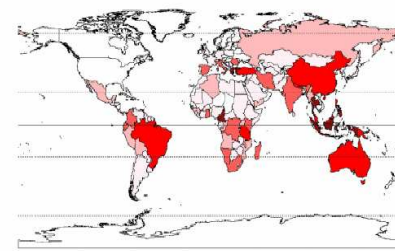


Figure 6: Countries of origin of Orthoptera types in German collections (*DORSA*)

Besides that, verification of geographical *DORSA* data against external geographical authority files can be done in a similar manner to the systematic verification of *DORSA* data against *OSF* data.

Another advantage of geo-referencing specimen data is that one can easily create a map illustrating the countries of origin of type specimens (Figure. 6). As you can see, most type specimens in German collections were collected in tropical countries reflecting Germany's scientific expeditions and colonial history. Publishing type specimen data and

concepts; allowing multiple verification of data, etc.). Thereby *GRID* technology should enable us to merge local, domain specific and global authority files (via web services).

ACKNOWLEDGEMENTS

- Dr Martin Doerr, *ICS Forth*, Heraklion/Crete, Greece and Stephen Stead, Paveprime Ltd, London/UK for checking the Ontologies;
- Dr Siegfried Krause and Dr Dagmar Hirschfelder, *GNM Nuremberg*;
- Dr Klaus Riede and Dr Sigfrid Ingrisich, *ZFMK*, and the *DORSA* data capture team and curators responsible for Orthoptera in the museums in Berlin, Eberswalde, Dresden, Hamburg, Halle, Bonn, Frankfurt, Stuttgart and Munich;
- Dr Christian Dietrich, Dept. of Neuroinformatics, *Ulm University*;
- Dr Bernhard Misof, *ZFMK*;
- Birgit Rach, Dirk Rohwedder and the *BIODAT-* & *BIOTA E15-* data capture teams, *ZFMK*;
- The *German Ministry of Education and Research (BMBF)* for funding the *DORSA-*, *DIG-* & *BIOTA East Africa E15* projects;

References

INGRISCH, S., RIEDE, K., LAMPE, K.-H. and DIETRICH, C. 2004: *DORSA – A “Virtual Museum” of German Orthoptera Collections.* – *Memorie Soc. Entomol. Ital.* 82 (2), 349-356.

LAMPE, K.-H., RIEDE, K. and INGRISCH, S. 2005: Repatriation of knowledge about insects and types through the *DORSA* virtual museum (Digital Orthoptera Specimen Access). *Beiträge zur Entomologie* 55 (2): 477-484.

RIEDE, K. 2004: Biodiversity Informatics in Germany: Ongoing projects and their possible contribution to the Global Taxonomy Initiative (GTI). – In: Junko Shimura (Ed.): *Global Taxonomy Initiative in Asia.* National Institute for Environmental Studies, Japan: 294-300.

Monika Hagedorn-Saupe

MICHAEL-culture.org: Germany's Participation in a Common Initiative for Europe's Digital Heritage

Zusammenfassung

MICHAEL – „Multilingual Inventory of Cultural Heritage in Europe“ ist ein Kulturportal für digitale Bestände und Sammlungen in Europa. Ziel des von der *Europäischen Union* geförderten Projektes ist es, einen zentralen Zugang zum digitalen europäischen Kulturgut zu ermöglichen. Dabei werden digitale Bestände aus Archiven, Bibliotheken und Museen aus vielen europäischen Ländern erschlossen und vernetzt. So werden Informationen zu digitalen Sammlungen und Beständen der bildenden Künste und Naturwissenschaften von der Urzeit bis zur Gegenwart aus verschiedenen Regionen und Kulturkreisen ebenso zugänglich wie Bestände der allgemeinen Geschichte, Technik- und Sozialgeschichte.

Initiatoren waren Frankreich, Italien und Großbritannien, seit 2006 baut auch Deutschland – wie auch weitere europäische Länder sein *MICHAEL*-Portal auf. An der Gründung des deutschen Portals sind insgesamt sieben Institutionen beteiligt, die deutschlandweit digitale Bestände ermitteln und vernetzen: die *Bayerische Staatsbibliothek*, das *Bundesarchiv*, die *Deutsche Nationalbibliothek*, das *Deutsche Museum*, das *Forschungsinstitut Senckenberg*, das *Landesarchiv Baden-Württemberg* und die *Stiftung Preußischer Kulturbesitz*.

Die einzelnen Länderportale werden nach und nach im gesamteuropäischen Portal zusammengeschlossen. Im Dezember 2006 wurde in Rom der erste Schritt, der Zusammenschluss der Portale aus Frankreich, Italien und Großbritannien, demonstriert. Ab 2007 werden weitere Partner ihre Länderportale zuschalten und so einen Beitrag für einen weltweiten zentralen Zugang zum europäischen Kulturerbe leisten.

Abstract

MICHAEL, the “Multilingual Inventory of Cultural Heritage in Europe”, is a cultural portal for digital collections in Europe. Funded by the *European Union*, the goal of this project is to enable centralised access to digitised cultural assets from across Europe. The project involves making accessible and linking together digital collections from many European archives, libraries and museums. The project is bringing together digital collections from the visual arts, the natural sciences (from the distant past to the present, from different regions and cultures) as well as digital sources on the general, technical and social history of European peoples.

France, Italy and Great Britain initiated this project. Starting in 2006, Germany and other European nations began work on their respective parts of the *MICHAEL* portal. Seven institutions are involved in establishing the German component of the *MICHAEL* portal. The German institutions engaged in making accessible and linking together digital collections of culture are: the *Bavarian State Library*, the *German Federal Archives*, the *German National Library*, the *German Museum*, the *Research Institute at Senckenberg*, the *Provincial Archives of Baden-Wuerttemberg* and the *Prussian Cultural Heritage Foundation*.

Bit by bit, the national portals will merge in a common European portal. In December 2006, the first step, the integration of the French, Italian and British portals, was accomplished. Starting in 2007 additional partners will be integrating their national portals as well. The net result of these efforts will be centralised access – access available from anywhere in the world – to Europe's cultural heritage.

Ladies and gentlemen, dear colleagues, as you have already heard from Professor Lehmann (and as stated in the conference materials you received at registration), this conference is made possible through support from the *German Research Foundation* and many other organisations. This conference has been organised in co-operation with our *MICHAEL* partners.

In what follows I would like to introduce *MICHAEL* to all those of you unfamiliar with the project, outlining the project's objectives and, in particular, describing the assignments we are undertaking for the *MICHAEL* project in Germany.

MICHAEL stands for *Multilingual Inventory of Cultural Heritage in Europe*. At present, seventeen European countries are participating in developing this European co-operative project. Rossella Caffo is co-ordinating the project as well as the development of the Italian portal at the behest of the Italian Ministry of Culture. Tomorrow, Rossella will introduce to us the many activities currently underway in Italy.

With *MICHAEL* we have set our sights quite high. If we are nonetheless able to realise our ambitious goal, the portal will constitute a stable platform, promoting access to the whole of Europe's digital cultural heritage. Access in the form of the *European Digital Library* represents at the moment an important objective of European cultural and educational policy, as previously mentioned (we are especially looking forward to hear more about the *EDL* from *European Commission* representative, Marius Snyders). Jill Cousins will also introduce another important *EDL* building block, one made possible also through the support of the *European Commission*. This project has taken as its starting point the portal of European National Libraries.



Figure 1: Screenshots national *MICHAEL*-Portals

ready have the makings for a rather colourful *MICHAEL* round dance. In the upper-left corner you can make out the information portal for the Europe-wide project, in the upper-right hand corner is the German project website covering developments within Germany. In the middle row you see screenshots from the three currently existing national portals from France, Italy and the United Kingdom. At the bottom a screen grab is from the European portal.

But allow me to return to *MICHAEL*. The objective of this project is to enable worldwide access to Europe's cultural heritage. *MICHAEL* aims to link together the diverse digital collections of cultural organisations across Europe, making these collections accessible via a central access portal. *MICHAEL* will facilitate research in digital collections scattered across our continent and thus difficult to trace via individual museums, libraries, archives and other cultural organisations.

Currently in development, the European portal has nonetheless been publicly accessible since the beginning of the year. Here you see the result of a query of natural science collections; 499 such collections were found. We see references to a series of British and Italian collections, and if we scroll down further, references to collections in France as well (see figure 2).

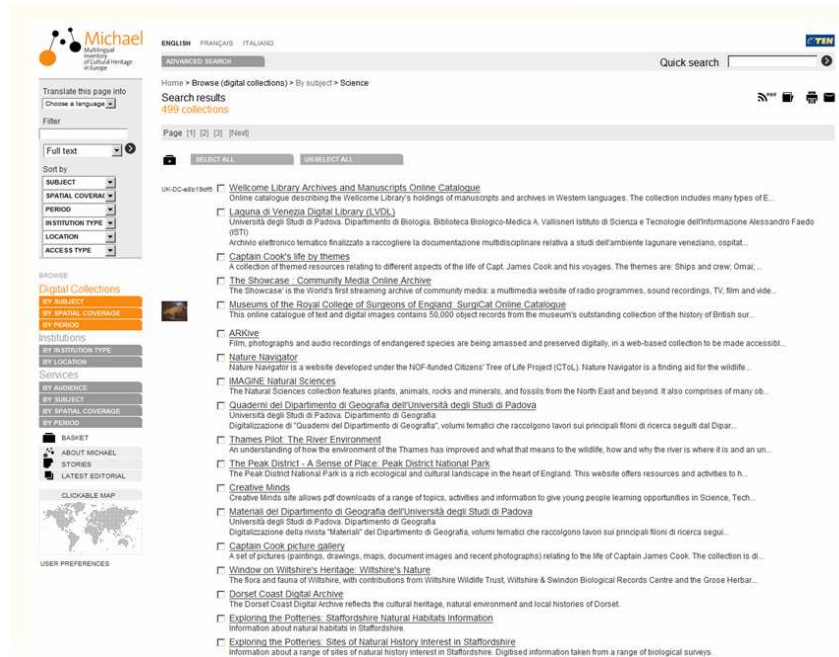


Figure 2: Display of a result page on the *MICHAEL-Portal*

Poland, Portugal, Spain and Sweden. In the meantime, four additional countries have joined the initiative: Belgium, Bulgaria, Estonia and the Slovak Republic, bringing us to a current *MICHAEL* total of eighteen nations.

Our chief aim is to ensure Europe's digitised cultural heritage is made available to Europeans. Access to digital collections of museums, archives and libraries will become faster and easier, available at any time from anywhere so long as one has Internet access. We will provide all Europeans access to digital culture, students and scholars, tourists and "creatives", children and youth, etc.

The focus of our project is not individual digital objects, but instead mixed lots. We have chosen this approach for a variety of reasons. For one, separating objects from collections is not always easy. A digitised journal, for example, consists of many digital objects. Similarly, a collection of digitised objects may only be taken notice of once it has been recorded on a CD. Many institutions are not yet at the stage of digitised individual objects, though they may already possess a digital collections overview. For example, an online researcher may discover that the Museum of Indian Art in Berlin houses important collections from Turfan. However, the museum does not yet possess a digitised image of all the objects in its collections. Nor is the museum in a position to offer online information about all of its objects.

It's clear that providing access to collections is an intermediate step to providing access to objects. The collections bundle individual objects according to thematic connexions chosen by the cultural organisations: such as Asian migrants to Great Britain during the 1950s, registries of local archives in Yorkshire, works of the French artist Charles Hallo, who between 1913 and 1936 produced for different companies railway posters, or the corpus of nineteenth-century Scottish correspondence physically located in Italy.

A special challenge in developing our European server is the multiplicity of languages in Europe. We are not in a position to translate all of the information in the *MICHAEL* project into each and every European language. The portal shall be made available to all interested parties in their own native language, in other words, navigation and general instructions have to be made available in all of the languages of the participating partners. So that each user, in his or her native language, is able to search across all of the collections in the portal, we are also compiling a list of select concepts based on the *UNESCO* Thesaurus. This

The relevant working groups are at the moment intensively engaged in optimising and expanding users' opportunities to browse, scroll and search. We know that, for the national portals alone, this is no easy task, and it is all the more difficult to achieve for the European portal, supplied as it is by many national portals. In the second phase of the project, eleven additional countries joined the original team from France, Italy and the UK. The "new eleven" members are: the Czech Republic, Germany, Hungary, Finland, Greece, Malta, the Netherlands,

master list will be harmonised with lists from each of the partners, in each partner language, allowing a general search in each of the portal's languages. The *UNESCO* Thesaurus is currently available in four languages: for the sub-set of concepts we are using from the *UNESCO* Thesaurus, we are compiling translations of the relevant counterpart terms in the remaining languages. Through the use of the same concept for short descriptions, the classification of the respective collection, one gains access to relevant collections in countries where one does not speak the language. Because we agreed to use common metadata, in other words, the same categories for the formal and informal indexing of collections, it becomes possible to find in one search collections that actually belong together. As you can imagine, moving forward together in such a large consortium requires a good deal of co-ordination and no small amount of time and energy. But it is worth the effort!

Another aspect of the project that was important to us, one that Christophe Dessaux will report on in his presentation tomorrow, was the decision to use open source software, so that we in the future can also develop the software together.

In many countries, the participating *MICHAEL* and *MICHAEL*-plus partners are ministries of culture. Because responsibility for culture rests with the federal states in Germany, our German effort is co-ordinated by *EUBAM*, a German working group of national and state representatives which addresses European matters affecting museums, libraries, archives and monuments protection. In Germany, seven large cultural organisations agreed to join *MICHAEL* as partners: the *German Federal Archives*, the *State Archive of Baden-Württemberg*, the *German National Library*, the *Bavarian State Library*, the *German Museum in Munich*, the *Senckenberg Natural History Museum* in Frankfurt am Main and the *Prussian Cultural Heritage Foundation* (with the *Berlin State Library*, the *State Museums in Berlin* and the *Secret Central Archives*, the Foundation encompasses libraries, museums and archives).

Together, these seven partners are establishing the German portal. And as soon as the *MICHAEL* portal goes online, they will encourage as many other German cultural organisations as possible to make available information about their digital collections so that they, too, can join the German, and therewith European, *MICHAEL* portal.

The division of labour among the German partners ensures that each partner accepts some of the tasks necessary to ensure the timely completion of the overall project. Technical implementation rests with the *Bavarian State Library*, whereas translation of project documentation into German is performed by the *German Museum*. The *State Archive of Baden-Württemberg* and the *German Federal Archives* together co-ordinate digital acquisition among German archives. The *Prussian Cultural Heritage Foundation* assumed responsibility for the translations of the concept lists. The Foundation also, in addition to other tasks, ensured the digital collections from cultural organisations already gathered by *EUBAM* were migrated into the new *MICHAEL* collections module.

We have recently received approval to incorporate into *MICHAEL* the *German-language Central Index of Digitised Prints (zvdd)*. Funded by the *German Research Foundation*, the *zvdd* was established on 1 April 2005 in order to provide a general index and access to digitised library materials. The *zvdd* captures four different levels of information: collections, bibliographic metadata, structural data as well as full-text data.

The Berlin Institute for Museum Research conducts an annual survey of all German museums. The most recent survey requests information about digital collections in these museums, so that by the end of the year we will have up-to-date information which will flow into *MICHAEL*.

In sum, *MICHAEL*'s development is well underway. For this reason, we look forward to learning from today's and tomorrow's joint presentations with a view toward our common goal: establishing the best possible, comprehensive, user-friendly access point to Europe's digital cultural heritage.

Monika Hagedorn-Saupe

MICHAEL-culture.org: Eine gemeinsame Initiative für Europas digitales Kulturerbe und die deutsche Beteiligung

Sehr geehrte Damen und Herren, liebe Kolleginnen und Kollegen, wie Sie bereits von Prof. Lehmann gehört haben und auch den Unterlagen zu entnehmen ist, veranstalten wir diese Konferenz, mit Unterstützung der *Deutschen Forschungsgemeinschaft* und vieler weiterer Partner, in Kooperation mit unseren *MICHAEL* Partnern.

Im Folgenden möchte ich nun für alle diejenigen, denen *MICHAEL* noch kein Begriff ist, vorstellen, was unsere Zielsetzung im Rahmen des Projektes ist und insbesondere, welche Aufgaben wir uns für Deutschland vorgenommen haben.

MICHAEL steht für „*Multilingual Inventory of Cultural Heritage in Europe*“, also für den mehrsprachigen Nachweis von digitalen Sammlungen des Kulturerbebereiches in Europa: Es handelt sich somit um ein europäisches Kooperationsprojekt. Derzeit sind siebzehn europäische Länder an dem Aufbau beteiligt. Koordiniert wird der Aufbau des Portals im Auftrag des italienischen Kultusministeriums von Rossella Caffo. Rossella wird uns morgen die vielen Aktivitäten, die in Italien laufen, vorstellen.

Wir haben uns mit *MICHAEL* ein hohes Ziel gesetzt. Wenn wir es erreichen, wird das Portal aber einen stabilen Baustein bilden für den Zugang zum digitalen europäischen Kulturbereich insgesamt. Dieser stellt ja unter der Losung „*Europäische Digitale Bibliothek*“ im Moment ein wichtiges europäisches kulturpolitisches Ziel dar, wie bereits mehrfach angesprochen (und hierzu werden wir insbesondere morgen von Marius Snyders von der Europäischen Kommission hören). Jill Cousins wird einen weiteren wichtigen Baustein vorstellen, der aufbauend auf dem Portal der europäischen Nationalbibliotheken derzeit durch Förderung der Europäischen Kommission entsteht.



Abbildung 1: Screenshots der *MICHAEL* Webseiten

Aber nun zurück zu *MICHAEL*. Zielsetzung des Projektes ist es, europäisches kulturelles Erbe weltweit zugänglich zu machen. Durch *MICHAEL* sollen die vielfältigen digitalen Sammlungen von Kultureinrichtungen in Europa vernetzt und über ein zentrales Zugangsportale erschlossen werden. *MICHAEL* soll damit die Recherche schwer auffindbarer, sonst weit verstreuter Informationen zu digitalen Sammlungen aus Europas Museen, Bibliotheken, Archiven und anderen Kultureinrichtungen erleichtern.

Wie in Abbildung 1 zu erkennen, gibt es bereits einen bunten *MICHAEL* – Reigen. In der oberen Reihe sehen Sie auf der linken Seite das Informationsportal zum europäischen Projekt, rechts daneben ist die deutsche Projekt-Webseite, die über das Projekt selbst informiert. Screens der bereits existierenden drei nationalen Portale, Frankreich, Italien und UK, sehen Sie in der mittleren Reihe, und das europäische Portal in der unteren.

Das europäische Portal ist seit Anfang des Jahres öffentlich zugänglich. Es ist im Aufbau – hier sehen Sie einen Screenshot, der das Suchergebnis zeigt, das ich erhielt, als ich nach Naturwissenschaftlichen Sammlungen gesucht habe: 499 Sammlungen wurden

gefunden. Wir sehen Nachweise einer Reihe britischer und italienischer Sammlungen und, wenn man etwas weiter herunterscrollt, auch Nachweise aus Frankreich (Abbildung 2).

Die zuständigen Arbeitsgruppen sind zur Zeit intensiv damit beschäftigt, die Möglichkeiten des Browsens, also Blätterns und gezielten Suchens, für die Nutzer zu optimieren und auszubauen. Wir wissen, dass dies schon für nationale Portale keine leichte Aufgabe ist, umso schwieriger stellt sie sich dar für ein europäisches Portal, das sich aus vielen nationalen Portalen speist.

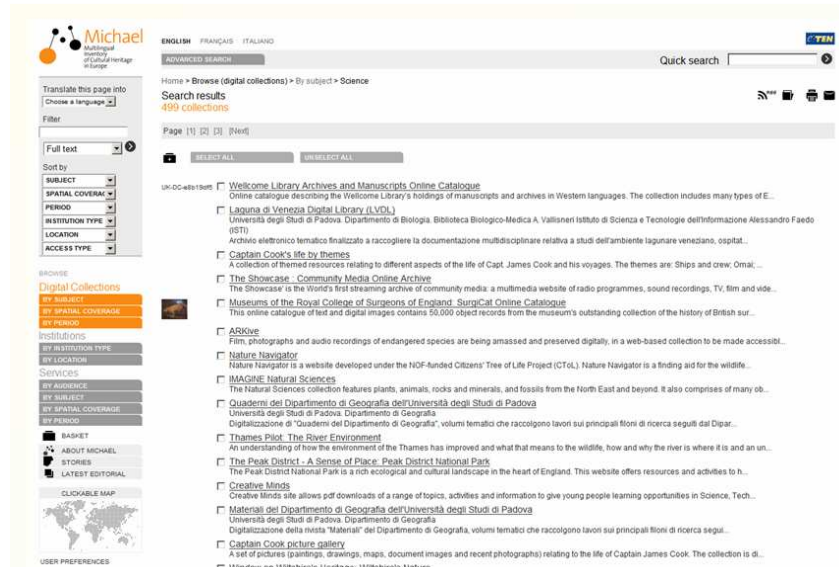


Abbildung 2: Ansicht eines Suchergebnisses auf der MICHAEL-Seite (Screenshot)

einfacher Zugang zu digitalen Beständen in Museen, Archiven und Bibliotheken unabhängig von Zeit und Ort (sofern man Internet-Zugang hat). Und zwar Zugang zum digitalen Kulturerbe für alle Europäer, sowohl für Studentinnen und Studenten wie für Wissenschaftlerinnen und Wissenschaftler, für Touristen wie für „Kreative“, für Kinder und Jugendliche usw.

Im Fokus stehen nicht die einzelnen digitalen Objekte sondern Konvolute – dies aus verschiedenen Gründen: zum einen ist die Trennung zwischen Objekt und Sammlung nicht immer einfach: eine digitalisierte Zeitschrift z.B. besteht aus vielen digitalen Objekten, andererseits wird eine Sammlung digitalisierter Objekte auf einer CD so ebenfalls sichtbar gemacht. Viele Institutionen sind noch nicht auf dem Stand der Einzel-Objekt-Digitalisierung, haben aber bereits eine digitale Beständeübersicht, d.h. ein Forscher erfährt bereits, dass das *Museum für Indische Kunst* hier in Berlin bedeutende Bestände aus Turfan hat, auch wenn das Museum noch nicht von allen seinen einzelnen Objekten ein digitales Foto besitzt und erst recht noch nicht Informationen zu allen seinen Objekten online anbietet.

Klar ist, dass der Zugang zu Sammlungen einen Zwischenschritt bildet zum Zugang zum Objekt. Die Sammlungen bündeln die einzelnen Objekte nach den von den Kultureinrichtungen gewünschten thematischen Zusammenhängen: Migranten aus Asien nach Großbritannien in den 50iger Jahren des letzten Jahrhunderts. Verzeichnisse der Kommunalarchive in Yorkshire, Plakate des Künstlers Charles Halle, der zwischen 1913 und 1936 für verschiedene Firmen Plakate zur französischen Eisenbahn gemacht hat oder der in Italien liegende Korpus Schottischer Korrespondenz aus dem 19. Jahrhundert.

Eine besondere Herausforderung für den Aufbau des europäischen Servers ist die Sprachenvielfalt in Europa. Wir sind sicher nicht in der Lage, alle Informationen in alle

Zu Beginn des zweiten Projektes sind dem ursprünglichen Team Frankreich, Italien und UK elf weitere Länder beigetreten: Deutschland, Finnland, Griechenland, Malta, Niederlande, Polen, Portugal, Spanien, Schweden, Tschechische Republik und Ungarn. Inzwischen gehören vier weitere Länder zur Initiative: Belgien, Bulgarien, Estland und Slowakische Republik, d.h. inzwischen sind achtzehn Länder an MICHAEL beteiligt.

Unser Hauptziel ist es, Europas digitalisiertes Kulturgut gerade auch den Europäern zugänglich zu machen: Schneller und

europäischen Sprachen zu übersetzen. Das Portal soll den interessierten Nutzern/innen in ihrer eigenen Sprache zur Verfügung stehen, die Navigation und allgemeine Beschreibungen müssen in allen Sprachen der beteiligten Partner vorliegen. Damit die Nutzer/innen über alle Sammlungen in ihrer eigenen Sprache recherchieren können, wird zunächst auf ausgewählte Begriffe des *UNESCO*-Thesaurus aufbauend eine Liste von Begriffen mit ihren jeweiligen Entsprechungen in den Sprachen der Partner im Portal implementiert, über die die Suche erfolgen wird. Der *UNESCO*-Thesaurus liegt bereits in vier Sprachen vor – zu den verwendeten Begriffen wurden die jeweiligen Entsprechungen in den fehlenden Sprachen erarbeitet. Durch die Verwendung der gleichen Begriffe für die kurze Beschreibung, die Klassifizierung der jeweiligen Sammlung, bekommt man Zugang auch zu den relevanten Sammlungen in den Ländern, deren Sprache man selbst nicht spricht.

Dadurch, dass wir uns auf gemeinsame Metadaten, d.h. gleiche Kategorien für die formale und inhaltliche Erschließung der Sammlungen, geeinigt haben, wird es möglich, Zusammenpassendes auch zusammen zu finden. Bei einem solch großen Konsortium kann man sich vorstellen, dass eine gemeinsame Weiterentwicklung eines abgestimmten Prozesses bedarf und dass dies ein nicht zu unterschätzender zeitlicher Aufwand ist – aber die Mühe lohnt sich!

Ein weiterer Aspekt, der für uns wichtig war – und Christophe Dessaux wird uns morgen darüber mehr berichten – war die Entscheidung für den Einsatz von openSource-Software, so dass die Software in Zukunft auch gemeinsam weiterentwickelt werden kann.

In vielen Ländern sind die *MICHAEL* – und *MICHAEL-Plus*-Partner Kultusministerien. Für Deutschland hat *EUBAM*, die Bund-Länder übergreifende Arbeitsgruppe für europäische Angelegenheiten von Museen, Bibliotheken, Archiven und der Denkmalpflege, wegen der verteilten Verantwortlichkeiten für Kultur einen anderen Weg vorgeschlagen. Für Deutschland konnten sieben große Kultureinrichtungen als Partner für das Projekt gewonnen werden: das *Bundesarchiv* und das *Landesarchiv in Baden-Württemberg*, die *Deutsche Nationalbibliothek* und die *Bayerische Staatsbibliothek*, das *Deutsche Museum* in München und das *Naturmuseum Senckenberg* in Frankfurt am Main sowie die *Stiftung Preußischer Kulturbesitz*, die mit der *Staatsbibliothek zu Berlin*, den *Staatlichen Museen zu Berlin* und dem *Geheimen Staatsarchiv* alle drei Kultursparten umfasst. Diese sieben Partner bauen derzeit gemeinsam das deutsche Portal auf und werden, sobald das deutsche *MICHAEL*-Portal online gehen kann, dafür werben, dass möglichst viele deutsche Kultureinrichtungen Informationen über ihre digitalen Sammlungen für das deutsche und damit auch für das europäische Portal bereitstellen.

Die sieben deutschen Partner gehen arbeitsteilig vor: die technische Umsetzung liegt in den Händen der *Bayerischen Staatsbibliothek*, die Übersetzung der Projekt-Dokumente ins Deutsche wird durch das *Deutsche Museum* geleistet, das *Landesarchiv Baden-Württemberg* und das *Bundesarchiv* koordinieren die Erfassung der digitalen Bestände der deutschen Archive – die *Stiftung Preußischer Kulturbesitz* hat die Übersetzungen der Begriffslisten bearbeitet, die digitalen Sammlungen aus dem Kulturbereich, die bereits bei *EUBAM* erfasst wurden, in das Erfassungsmodul eingepflegt usw. So hat jeder Partner einen Teil der anstehenden Aufgaben übernommen.

Inzwischen haben wir auch die Zusage, dass wir die Daten des von der *DFG* geförderten *ZVDD*, des *Zentralen Verzeichnis Digitalisierter Drucke*, für *MICHAEL* bekommen werden. Das *ZVDD* wird seit dem 1.4.2005 aufgebaut, um einen zentralen Nachweis und Zugang zu digitalisierten Bibliotheksmaterialien bereitzustellen. Gesammelt werden dazu umfassende Erschließungsinformationen zu Digitalisaten auf insgesamt vier Ebenen: Sammlungen, bibliografische Metadaten, Strukturdaten sowie Volltextdaten.

In der Befragung des *Instituts für Museumsforschung*, die wir jährlich an allen deutschen Museen durchführen, haben wir im Rahmen der aktuellen Erhebung nach digitalen Sammlungen in deutschen Museen gefragt, so dass uns Ende des Jahres aktuelle Informationen hierzu vorliegen, die dann auch in *MICHAEL* einfließen werden.

Man kann also zusammenfassend sagen: der Aufbau von *MICHAEL* ist in vollem Gang. Insbesondere deshalb hoffen wir, dass wir durch die gemeinsamen Präsentationen und Diskussionen heute und morgen weitere Anregungen erhalten und sehen, wo wir voneinander lernen können und wo wir durch Kooperation unser gemeinsames Ziel, einen umfassenden und Benutzer-freundlichen Zugang zu unserem digitalen kulturellen Erbe in Europa für alle Interessierten, möglichst effektiv erreichen können.

Iván Rónai

How best can a Ministry promote Digitisation Projects?

Zusammenfassung

Wie kann ein Ministerium Digitalisierungsprojekte fördern?

Da öffentliche Mittel derzeit knapp sind, ist es nicht leicht, staatliche Ausgaben für die Digitalisierung von Kulturgütern zu rechtfertigen. Der Vortrag unternimmt den Versuch, die Rolle des Staates und insbesondere die Rolle des ungarischen Kulturministeriums bei der Schaffung digitaler Inhalte und kultureller Serviceleistungen zu beschreiben.

Untersucht werden die regulierenden Funktionen der Behörde und ihre Verantwortung für Finanzierung und Durchführung der Digitalisierung. Anhand verschiedener Tests und Fragenkataloge wird herausgearbeitet, auf welche Weise staatliche Kulturverwaltungen die Digitalisierung von Kulturgut fördern können.

Außerdem werden Erfahrungen aus der Mitarbeit im *MICHAEL-Projekt* vorgestellt.

Abstract

In our time when public resources are scarce, it is not easy to justify state expenditure on cultural digitisation. In this presentation an endeavour is made to describe the role of the government, and hence of the cultural ministry, concerning digital content creation and digital cultural services.

The ministry's regulatory functions, its responsibility in funding and executing the tasks of digitisation are examined. Various tests and questions to be asked are used to determine the ways the ministry can foster cultural digitisation.

The experience earned while working on the *MICHAEL*-Project will be presented.

Currently in Hungary, when so much emphasis is put on the redefinition of the role of the state, along with these efforts I try to go over the exercise of describing the responsibilities of the state in the field of cultural digitisation and cultural content creation. I will use the concept of cultural digitisation in a broad sense as a general term, including the transformation of assets of the cultural heritage into digital format, the digitally born heritage, the digital inventory, cataloguing, registering of cultural objects and the provision with metadata of digital cultural objects, the running of online and offline services, in one, the whole process of making cultural assets and information digitally accessible.

Definitions

In fact, the definition of 'public task' is not available in international literature; instead, 'public services', or 'public goods' are used in English language writings.

Public goods: public goods are goods which are beneficial for society but which are not normally provided by the market given that it is difficult or impossible to exclude

anyone from using the goods (and hence making them pay for the goods). This can be the case of national defence and some types of public broadcasting.¹

Public services: A service that is performed for the benefit of the public or its institutions.

From the definitions above we can conclude the following definition:

Public task: the duty of the state or the community to ensure the availability of public goods and public services.

However, it remains to be clarified why a certain service is performed for the benefit of the public or its institutions and certain ones are not. Another approach to grasp the notion of public tasks is to declare that all tasks that are named in any legal document as the duty of the state or its representatives and institutions are public tasks.

I suggest accepting the following definition of a public task: It is an activity or service which is necessary but

- is not performed by profit-oriented enterprises;
- cannot be performed efficiently or funded by profit-oriented enterprises;
- users cannot buy it on market prices;
- it is difficult or impossible to exclude anyone from using it;
- cost-benefit balance can be measured only with trouble and on the long run;
- the return cannot be expressed in monetary terms.

From the table 1 below we can see that in those columns where we put only "T"s (for "true"), the provision of the activity is in public responsibility. Consequently, in line with our working definition of state tasks, preparing digital inventories of digitised cultural assets and collections, setting up catalogues and registries of digital cultural contents, ensuring long term preservation of digital cultural heritage are apparently duties of the state. The second precondition, whether these activities are necessary, has not been tested yet. Lots of studies and papers have been written about the evidence of the importance of inventorying, cataloguing and preserving digital cultural assets.

	transformation into digital format	digital inventory	cataloguing, registering	metadata	online content services	Long term preservation
is not performed by profit-oriented enterprises	P	T	T	P	P	T
cannot be performed efficiently or funded by profit-oriented enterprises	T	T	T	T	P	T
users cannot buy it on market prices	P	n.a.	n.a.	n.a.	P	n.a.
it is difficult or impossible to exclude anyone from using it	T	n.a.	n.a.	n.a.	F	T
cost-benefit balance can be measured only with trouble and on the long run	P	T	T	T	P	T
the return cannot be expressed in monetary terms	P	T	T	T	T	T
Legend: T = true, F = false, P = partly true, n.a. = not applicable						
Table 1: Duties of digitisation and responsibilities of the public and private sector						

Concerning the transformation into digital format, providing metadata, and operating the online cultural content services, we have a different case. As seen from the table, there is a partial interest of the business sector to perform these tasks. There is a segment of digital cultural assets and services that is marketable. What exactly is this segment, and

¹ STATE AID ACTION PLAN Less and better targeted state aid: a roadmap for state aid reform 2005–2009 (Consultation document) (presented by the Commission) {SEC(2005) 795}

what are the criteria to be part of this segment? To answer these questions let us take the examples of

- full-text books digitisation and online contents services,
- digitisation of museum objects and works of art,
- digitisation of archival material,
- provision of e-learning material,
- online services on culture related information in general (cultural events, programmes, cultural actors, institutions etc).

The most important condition for products and services to be supplied by business is that it contains at least a margin of profit or other benefits that on the long run lead to profit. This is realised if there is a critical mass of users that are willing and able to pay for the products and services. It is evident that the existence of the critical mass of users depends on the level of economic development, the quality of life and the cultural customs and traditions of the people concerned. In some countries people can better afford to pay for services than in others where they rather do it themselves or borrow, or quite often illegally copy. To meet a critical mass of users a number of measures can be implemented. This is where the state, the public sphere must intervene and play a major role. Business itself can and certainly will in due time increase the number of potential users by realising that the internet has drastically changed user habits of consumption. I think, for example, that the online version of newspapers and journals did not result in shrinking profits of publishers. There is probably a synergic effect of digital consumption on traditional markets of cultural goods and services.

Hungary's leading content provider – as they paraphrase themselves – combines content industry services with selling their own digitised works.² The only way this company can survive as an enterprise is getting orders from memory institutions to digitise their holdings. The products they sell bring a very tiny profit due to the small size of the market.

Another newly founded enterprise plans to act as a *mediator* or a *retailer* of digitised educational material, provides services to support the sale of DRM (digital rights management) technology protected digital curricula, acts as a business-based curriculum advice and supports the internet marketplace. According to its business strategy, the company makes available digitised educational material at a price that is advantageous for all actors: users, publishing houses, authors (right holders) and the retailer. Probably, the loss of the publisher due to fewer paper copies sold will be lower than the gain due to paid downloads, because the number of illegal copying will be much smaller.

So, this status quo described in mind, what in effect is the role of the state (see also table 2 at the end of the text)?

- Definitely, it has to regulate:
 - The legal environment: To find the balance between the interests of intellectual right holders and the users, needs on the one hand, to let authors enjoy benefits of their creations so that they remain encouraged to create works of art, and on the other hand to allow the public an equal access to knowledge and cultural goods.

Good examples are designed terminals in a closed network of libraries where access to copyrighted digital cultural objects is free of charge for learning and scientific purposes.

² "Our company builds and publishes databases of cultural valuables, which makes it a pioneer on the Hungarian CD-ROM market. We make these outstanding subjects available on CD-ROM, DVD-ROM, local network, and on internet as well. The broad scale of our products ranges from bibliographic databases for experts, to publications containing the most important values of our national and universal culture, which may enrich the bookshelves of families. We provide books, encyclopaedias, catalogues, and bibliographies for the experts and lovers of history, records, knowledge of a certain country, literature, education, genealogy, and bibliography in a modern way with the possibility of searches." http://www.arcanum.hu/idegennyelvu/inny_index.html

- In order to ensure technical interoperability, the state has to lay down standards, quality principles, requirements of metadata and criteria of authenticity. But in fact, very frequently standards are set by commercial organisations, and only after the need emerges turning them into national or international standards.
- It also has to take up the digitisation, content creation by funding public memory institutions' digitisation activities.
 - This can be done directly by allotting funds in the budgets of state run institutions for digital cultural services (part of the annual budget is assigned to providing internet services and offering their cultural assets in digital form).

In Hungary, although almost every state-run cultural institution has its own website funded from its annual budget, digital content creation (transformation of analogue content into digital) is rarely incorporated into its annual budget.

- Another way of active participation of the state is the release of grant proposals aimed at digitising cultural heritage. In Hungary, the majority of cultural content developments are performed from such funds. This also means that quality and quantity of digitisation is quite accidental, because the existence and the availability of these funds due to a missing long-term policy cannot be planned. Consequently, if once the state has decided to take part in digitisation, it must do it purposefully, based on a strategy and carrying out the activities coming from the action plans of this strategy. In Hungary, we do have an information society strategy, but the need for and the importance of a continuous financing of cultural content development has not been recognised yet.
- Thirdly, the state can act as an initiator to ignite and stimulate the digitisation of cultural heritage until a critical mass is reached, or provide a supplementary funding to make business ready to enter. For example, if the number of digitised books will reach a level when the basic documents can be accessed electronically and users will more and more exploit the advantages of e-books, publishers will become interested in having a bigger market share. Or, if public libraries will have more resources to pay for online content services, business will be more interested in content developments to take advantage of these services.

A good example of state support is the programme of the *Digital Literary Academy*. The objective of the academy is to make accessible an ever widening sphere of top Hungarian literature in a unified database for all interested at any point of the world (<http://www.irodalmiakademia.hu/>). The members of the academy are outstanding contemporary writers, poets. At the outset, there were 39 members, today the number of the members – with 15 posthumous members (writers who died recently) exceeds 70. Every year the members select a new member who becomes eligible to participate in the programme. The members make a contract with the state, in which they give authorization to digitise and make publicly available the digitised version of all their published works. In return, living members are entitled to remuneration in the form of a life-long monthly allowance equal to the fourfold value of the actual minimum salary. After the death of the member the allowance is paid to the heir/s for 8 months.

The role of the state can be illustrated with the *MICHAEL Plus* project. Here, a public organisation is in charge to collect information on cultural collections which are fully or partly digitised, or planned to be digitised. Using a standardised format, every public digitisation project has the duty to forward the information about its digital collection to the *MICHAEL* database. Thus, the *MICHAEL Plus* project as well as the *National Digital Data Archives* projects are state-supported schemes that help digital content development in a coordinated way.

But, what would be, if the state pulled out from digitisation? We would lag behind: The information society is evolving all over the world, without substantial cultural content we would seem like an uncultured nation. Worse, this would result in an uncultured nation, because more and more people, especially the young, get access to cultural values through the internet. What is not in the internet is inexistent for them. In addition, the economy increasingly consists of information society activities. New industries may emerge that are based on digital contents and content creation.

In conclusion, there is a need for state subsidy in cultural digitisation and contents development. The question is the degree, the extent of the financial support. At an earlier stage this support must be greater than later when technology and efficiency will have reached a point that enables business to take over certain parts of the task. Also, with the development of the civil society, NGOs may have a role in cultural digitisation.

Activity	PUBLIC NEEDS TEST	Objective of the task	Output and outcome of task fulfilment	Risks and consequences if the task is not performed
Full text books digitisation	Public	Make available books online in order to foster access to literary and professional works	Easy (occasionally free!) access to scientific and cultural literature and books content	Less equal opportunities to acquire knowledge hence low social mobility
	Business	Make business on online books services by providing access to documents or parts of document by collecting fees on use	Easy access to scientific and cultural literature for those who can afford to pay for it	Poor contents prevail on the net
Digitisation of museum objects	Public	Make available works of art online in order to foster access to cultural heritage	Educational role of museums intensifies, growth of knowledge base of society and research	Less equal opportunities to acquire knowledge hence low social mobility
	Business	Make business on selling digital copies of works of art, foster tourism	Awareness of most outstanding artistic works of humankind	Poor contents prevail on the net
Digitisation of archival material	Public	Make available archival documents online in order to foster access to cultural heritage and historical records	Easy (mostly free!) access to historic, archival records, growth of knowledge base of society and research	Less equal opportunities to acquire knowledge hence low social mobility
	Business	Make business on enabling online search in archival documents	Easy access to archival records for those who can afford to pay for it	Poor contents prevail on the net
Provision of e-learning material	Public	Make the internet an effective learning tool, provide structured knowledge to pupils, students, teachers and LLL	Easy and free access to e-learning material, more equal opportunities for all to learn	Less equal opportunities to acquire knowledge hence low social mobility
	Business	Make business on selling and mediating e-learning material	Easy access to e-learning material for those who can afford to pay for it	Poor contents prevail on the net
Online services, running portals	Public	share knowledge accumulated in memory institutions, better inform people	Up-to-date and easy access to information for all	Lack of information, hence unequal opportunities
	Business	attract people by better informing them to use the site, make business on ads	Up-to-date and easy access to information for all, mixed with advertisements	Lack of information, hence unequal opportunities

Table 2: Digitisation process and public and private responsibilities

Maria Sliwinska

Looking for the Needle in the Haystack: Small, Unique Collections – How to find hidden gems?

Zusammenfassung

Wir alle wissen um die Wichtigkeit nationaler Institutionen wie Nationalbibliotheken, Archive, Museen, Galerien und die umfangreichen Ressourcen, über die sie verfügen. Aber wie viele Millionen von Dokumenten und Objekten diese großen Häuser auch anzubieten haben, kann doch keine Institution der Welt alle unsere Wünsche befriedigen. Oftmals müssen Wissenschaftler, wenn sie auf der Suche nach Quellen sind, in kleine, abgelegene Orte reisen, um wichtige Informationen für ihre Forschungen zu finden.

Wie findet man Informationen über kleine und dennoch wichtige Sammlungen? Im Zeitalter der Information, in dem wir alle bombardiert werden mit Millionen von Nachweisen aus dem Internet, scheint es schwerer, an verwertbare Informationen über kleine Sammlungen zu kommen, als eine Nadel im Heuhaufen zu finden. Sie sind einfach versteckt, übers Internet nicht aufzufinden und damit nicht zugänglich. Und obgleich auch kleine Bibliotheken ihre Materialien katalogisieren und Archive und Museen ihre Sammlungen auf ihrer Homepage beschreiben, ist es doch für Wissenschaftler nicht leicht, diese Informationen zu finden. Es ist einfach unmöglich, in tausenden von kleinen Katalogen zu suchen, um alles unbekanntes Material zu finden, das die Quellengrundlage vervollständigen würde.

Eine gute Lösung bieten die Bibliotheken über die Gesamtkataloge an. Sie sind sehr hilfreich, wenn der oder die Forschende genau weiß, was er oder sie sucht, aber sie reichen nicht aus.

Zusätzliche Unterstützung kommt vom *MICHAEL-Projekt*, das kleine Sammlungen fördert, egal, ob sie bereits online sind

Abstract

We all are aware of the importance of national institutions like national libraries, archives, museums, galleries and the vast resources they offer. However, even though there are millions of items and objects available in a big library, archive or museum, there is no such an institution in the world that can satisfy all our needs. Quite often when looking for materials, researchers have to visit some distant places to discover an important information for their research.

But how to find an information about a small but important collection? In the information era, when we are bombarded by millions of hints from the internet, getting valuable information about a rare small collection is as difficult as finding the needle in the haystack. They are simply hidden, not advertised in the internet – that means not accessible. Even though small libraries have their materials catalogued, and archives and museums have their collections described on their homepages, it is not easy for a scientist to find such information. It's simply impossible to make thousands of searches in so many small catalogues to get unknown materials that make research complete.

A good solution offered by librarians are the union catalogues. Union catalogue is very helpful when a researcher knows what s/he is looking for but is not enough. Additional support comes from the *MICHAEL* project that promotes collections available not only on-line but off-line as well. In this project information about collections from small institutions can be equal or sometimes even more important than that from the big libraries, archives and museums. Information about the

oder nicht: In diesem Projekt werden Informationen über Sammlungen kleiner Institutionen gleichwertig oder gar höherwertig betrachtet als die der großen Häuser. Informationen über Sammlungen, die über ein bekanntes und gut aufgebautes Portal verfügbar sind, vereinfachen die Forschung und vervollständigen deren Quellenkorpus.

Die versteckten Ressourcen werden nicht nur für Forscher leichter sichtbar, die nach ihr suchen, sondern auch für Touristen, die sich inspirieren lassen, kleine Museen und weniger bekannte Monumente zu besuchen, die nicht in den gängigen Reiseführern stehen.

collections available from a trusted and structured source with one access point makes research easier and complete.

The hidden resources became visible not only to the researcher who are looking for them, but can also inspire tourists to visit a number of museums or other attractions that were not known and not advertised in the popular guides.

We all are aware of the importance of national institutions such as national or university libraries, archives, museums, galleries and the vast resources they have collected over the ages and offer to their clients. Everybody has heard about the big institutions with their long histories and can even name at least a couple of them like the *British Museum*, the *Louvre*, the *Uffizi* in Florence, or the *Grünes Gewölbe* in Dresden. Some people also know what kind of collections can be found in the archives.

Professionals know how to locate an institution that collects materials that are useful for their research, or that they just need for pleasure. There are a number of printed guides, catalogues, and inventories. However, among the main contributors to those guides are again big institutions that have the adequate human resources and knowledge of how to use the professional tools for promotion and marketing. Big institutions also have technical possibilities to convert their resources from an analogue to a digital collection, and in addition can advertise this information on their websites that are registered by search engines.

However, even though there are millions of items and objects available in a big library, archive or museum, there is no such institution in the world that can satisfy all our needs. Quite often when looking for materials, researchers have to visit distant places to discover important information for their research. But it costs time and money to travel. So, the growing number of digital collections is welcome at least by those researchers who are familiar with *ICT*.

But, how to discover information about a small but important collection? In the information era when we are bombarded by millions of hits from the internet, to find valuable information about a rare and small collection is as difficult as finding the proverbial needle in the haystack. They are simply hidden, not advertised on internet, or not accessible online. Something that is not available online is rarely used nowadays. But even though small libraries have their materials catalogued, and archives and museums have their collections described on their home pages, it is not easy for a researcher to find such information. It's simply impossible to make thousands of searches in so many small catalogues with just a hope of getting unknown material that can enrich research.

And we can't be even sure that we are able to obtain complete materials, because people used to travel, and so information about them and their works is found in many places. Also, documents have their own travelling records: they were moved for many rea-

sons, so sometimes we can find a real treasure in an unexpected place, as happened to some researchers who just discovered by chance the oldest prints hidden inside another book cover.

Information about the collections available from a trusted and structured source with one access point makes research easier and complete. Then, the hidden resources become visible not only to the researchers who are looking for them, but can inspire tourists to visit a number of small museums or other attractions that were unknown to them, and not described in the popular tourist guides. The hidden gems can breed an epidemic and attract thousands of tourists and millions of 'internauts' like in the case of the Victorian Museum in Tiverton discovered and visited by more than 60 million internauts within only one year after it went on the net.

Small institutions and their collections are valuable for foreign visitors who always feel interested in seeing remains of their ancestors' culture in the countries they visit. Let me recall a couple of Polish examples. One such example is a small *Museum of Mennonites* in Tczew or Nowy Dwór Gdanski with objects of Dutch settlement in that region – not even to be found on the internet. There is also a rare chance that a tourist knows about such a unique book as one of 48 surviving copies of the *Gutenberg Bible* available in a small museum in Peiplin; or about the manuscripts of the Hungarian king Mathias Corvin, available in the public library in Torun and so valuable for Hungarian culture; or about the archival materials of the Czech Amos Comensky, available in Leszno. Tourists are also interested in discovering monuments that are important for their own or their countries' history, like the one from the XIII century in Naumburg, presenting Regelinda, daughter of Boleslaw Chrobry, one of the first Polish kings, who married an aristocrat from Misnia. Such examples, as well as other more modest memorial plaques, can be found in almost every city in the world.

There are many documents available in archives that prove, to our surprise, how similar to the *European Union* of today the situation was in the time of the Middle Ages, when people travelled around Europe to study at the universities or to work. As evidence we can use again a Polish example, related to the *German University of Ingolstadt*, closed more than 200 years ago. This university was one of the most popular among Poles, after the Italian universities in Padua and Bologna. Between 1472 and 1699 828 Poles studied there, and three of them became rectors there. This means that it was much more popular place than the better-known *University in Heidelberg*. Census data from such places can be of the utmost importance for people from all over the world looking for their ancestors' history, not to mention the researchers who can enrich their studies.

Similar information one can find at the cemeteries. Brief information about dates of life – and, sometimes, about occupation – can become the final proof in research when more common sources of information are not available. The style of the tomb can also say a lot about the epoch, and is the so-called small architecture style. Some famous people are so much associated with the place normally connected with their fame that we tend to forget that quite often they have their roots in a different country, or even a different culture. This is the case of the first famous Hollywood actress, Pola Negri, born in Lipno in Poland, where a small museum has been created that can be a surprise even for people who are experts in film history. Maria Curie received the *Nobel Prize* in France, but some documents to her biography can be found in a museum in Warsaw. Similarly, Wit Stwosz from Nürnberg was a creator of fantastic sculpture that can be seen in Krakow's *St. Mary Church*.

Those few examples give us only an illustration of the situation where so many small collections are still hidden, and one can devote one's entire life to complete the data. It was a long time ago, at the end of the XIX century, when two pioneers Henri La Fontaine and Paul Otlet dreamed about a complete bibliography of world literature. They did a lot to reach the goal and established an International Federation for Documentation in Brussels (1895), later moved to The Hague (1930). But an enormous amount of work still did not allow them

Maria Sliwinska

Looking for the Needle in the Haystack:

Small, Unique Collections - How to find hidden gems?

to fully reach their goal. It's time to fulfil their dreams using the available technology and to reveal the hidden gems to the world. But, how doing this in the most efficient way?

A good solution offered by librarians is the union catalogues. A union catalogue is very helpful when a researcher knows what she or he is looking for. But this is not the only solution. People do not always have a clear idea which institution they have to conduct research in, as documents can be found in many different types of historical institutions, as well as those objects that can be found not only in the museums but also in archives and libraries, even, as we all know, the majority of the collections can be found according to specializations. But in a virtual reality it doesn't matter where the real collection is held, as it can be downloaded on our PC located either at home or at work, as easily as water from a tap. Researchers are interested in getting access to the digital collections from one source, with no distinction between type of institution. But someone must organize it for them. "Someone" means us librarians, archivists, custodians in cooperation with IT-specialists. There are many ideas.

I was fortunate to work in the *MINERVA* project group with the goal of coordinating digitisation in Europe in promoting the Lund Principles in my country. A couple of conferences and publications brought good results. Also, the *TEL-ME-MOR*-project, in which the National Libraries have been included, must be mentioned here as a next step forward.

Additional support comes from the *MICHAEL*-project, which promotes collections available not only on-line but off-line as well. In this project, information about collections from small institutions can be equal or sometimes even more important for a user than those from the big libraries, archives and museums that are well known and used in research by many. And last but not least, the *European Digital Library* in The Hague.

We would surprise the founders of the *International Federation for Documentation*, since with the good will of all initiatives mentioned above it has the potential to become the biggest library in the world, available to anybody in 24/7 mode, with structured resources from all kinds of historical institutions and institutions of all sizes, with good guidance to the collections. It requires an enormous effort, but it's worthwhile to join all our forces to obtain results as good as those obtained by those engineers who changed people's lives by supplying water directly to their houses. Water for the human body, knowledge for the human mind, both are equal.

Paul Smith

www.culture.fr/collections: A Single Entrance for Access to Heritage Data at the French Ministry of Culture

Zusammenfassung

Mitte der 1970er Jahre brachte das französische Kulturministerium mit dem Einsatz Computergestützter Datenbanken für die Findmittelverwaltung des architektonischen Kulturerbes eine Pionierleistung auf den Weg. Heute gibt es etwa 250 verschiedene Quellenbestände aus Archiven, Bibliotheken, Museen, aus Archäologie und Denkmalpflege, die entweder direkt vom Ministerium oder unter seiner Federführung von öffentlichen Kultureinrichtungen verwaltet werden.

Im Jahr 2007 hat das Ministerium ein Portal gestartet, um die verschiedenen im Netz vorhandenen Informationsquellen zum kulturellen Erbe Frankreichs zusammen zu binden und einen besseren Zugang zu diesen 'Sammlungen' zu schaffen, ein Portal, das leicht auffindbar und Dank seiner „intuitiven“ Suchmaschine auch für jedermann einfach zu benutzen ist. Gegenwärtig erlaubt die spezielle Suchmaschine Zugriff auf mehr als 30 Ressourcen aus Architektur, Denkmälern, Archiven, Archäologie, Ethnologie, Kartographie, Fotografie, illustrierte Handschriften, Theater, Musik, Tanz, etc – insgesamt ca. drei Millionen, Texte und zwei Millionen Bilder. Der Zugang erlaubt personalisierte Dienste wie E-Mail-Links, Downloads etc.

Abstract

During the 1970s, the French Ministry of Culture pioneered the use of computerised databases in the management of research findings on the architectural heritage. Today, there are some 250 different on-line documentary resources from archives, libraries, museums, archaeological and heritage services, run either directly by the Ministry or by public cultural institutions under its authority.

In 2007, the Ministry launched a portal to federate some of these different sources of cultural information and to give better access to them via a single access (now baptised 'collections') which is easily identifiable and easy to use by the general public. Today, a special search engine covers more than 30 different documentary sources under the Ministry's aegis, dealing with a broad range of cultural interests—architecture, listed monuments, archives, archaeology, ethnology, cartography, photography, illuminated manuscripts, theatre, music, dance, 'primitive' arts, industrial or maritime heritage...— and encompassing three million text notices and two million pictures. The access also provides personalised services such as baskets for downloading notices, e-mail links or recording personal research itineraries.

The French Ministry of Culture was an early adapter where information technology was concerned. A report of 1972 recommended the use of a computerised database to manage the large number of files on the built heritage being produced by the *Inventaire général des Monuments et des Richesses artistiques de la France*, the heritage inventory service created by André Malraux in 1963. This database, *I-ARCHI*, was operational by 1974, indexing the sites identified, documented and photographed by this service, and giving access to the documentary files held in the regions and available on microfilm in the service's Paris headquarters. Other databases soon followed: *Joconde*, a collective catalogue of paintings, objects and sculptures held in France's museums, whether on display or not, and *Leonore*, a database created by the national archives indexing all the files held on personalities who had been awarded the *Légion d'Honneur* during the nineteenth and twentieth centuries. *I-ARCHI*,

renamed *Mérimée* (after one of the country's first inspectors of historic monuments during the 1830s) now comprises more than 200 000 notices, many of them illustrated with digitally generated images or scanned copies of the paper files, and including a specific domain devoted to all of the country's protected monuments, about 143 000 in all. As for *Joconde*, it includes 347 000 notices dealing with works held in 270 different museums throughout the country.

In 1993, the last two databases, *Mérimée* and *Joconde*, were made available to the French public as a whole via the Post Office's Minitel service, a pioneering but now largely forgotten chapter in the early history of electronic networking, totally superseded by the Internet on which the two resources went online in 1996. In the same year, the Ministry launched a plan, and a budget, for digitising cultural documentation (archives, libraries, museums, built and movable heritage) in order to enrich the existing databases, still largely textual. A new database baptised *Mémoire* was also created, assembling graphic and still photographic images culled from the inventory files, from the archives of the historic monuments administration and from archaeological excavations.

In January 2002 another official report on the digital diffusion of heritage and cultural resources made a recommendation for improved interoperability between the different databases and envisaged their simultaneous consultation via a single access on the Ministry's Internet portal. In 2005 this project was given priority in the Ministry's information system master plan. This initiative was in keeping with a constant trend in ministerial policies since Malraux of democratising culture and making cultural resources more easily accessible. It was a top-down initiative, decided at ministerial level and making different services and departments, which, hitherto, had often conducted their information policies in more or less splendid isolation, work together. But, after all, directly or indirectly, all the contents of these different databases were founded on French tax payers' money.

A working group was set up which identified some 250 separate on-line documentary resources, managed directly by the Ministry or by public cultural institutions. These resources used differing technologies and software programmes, and, more often than not, were not at all familiar to the general public. In 2006, then, the '*guichet unique*' project was set up, with its objective of creating a single, user-friendly entrance for accessing digitised cultural information. The project, piloted by the documentation services at the Ministry and bringing together representatives of the different ministerial departments (central administration, archives, architecture and heritage, museums, books and reading, research mission, delegation for the French language and department for information and communication) defined six underlying ambitions for the project

- To offer the general, non-specialist public a single, transversal access for on-line cultural assets, respecting contemporary norms for accessibility
- To federate and improve the diffusion of heritage documentation, the single access making for better overall visibility
- To allow the general public simultaneous access to several different resources, without having to understand the details of how these resources came into existence and the differences in their presentation
- To offer different types of research, simple or advanced
- To display the results in various forms: as a simple list, a screen of thumbnail images or a presentation by documentary category
- To respect the existing technologies used for producing and storing the data.

Based on technical specifications drawn up by the project group, a call for tenders was published in July 2006, and the Sinéqua firm chosen for the development of a semantic search engine based not on the existing organisation of data by fields, but on full-text research with specific functions:

- Morpho-syntactical analysis of words (words in all forms, upper or lower case, with or without accentuation, singular or plural forms)

- Phonetic and spelling variations (*otoportret* is understood as *autoportrait*, or *self-portrait*)
- Synonyms (Le Corbusier/Charles-Edouard Jeanneret)
- Generic and specific terms, using the thesauri developed over the past twenty years for record management



Figure 1: Screenshot <http://www.culture.fr>

The first version of the system was implemented in April 2007, the second version being available in July 2007, covering a total of 32 information sources: databases, static and dynamic web pages and electronic publications. A version 3, to be implemented during the year 2008, will integrate further sources and will comprise a cross-linguistic interface (Spanish and English in addition to French), advanced research capacities (specific thesauri, who, what, when fields) and personalised services such as research history, favourites and a results basket for downloading or e-mailing.

To conclude, the reader is invited to stop reading and pay a visit to the Ministry's portal (<http://www.culture.fr>) and click on the 'collections' tab, or access the site via <http://recherche.culture.fr>. The qualities of the search engine, compared to others, will soon become apparent, and the results displayed

present a certain degree of scientific reliability, guaranteed by the seriousness and long history of the creation of the electronic resources at the French Ministry of Culture. Of course, making digitised cultural resources more readily accessible and increasing the numerical offer (three million notices, today with more than two million images) is neither the last nor the only word in cultural policies, but as other research and digitisation projects enrich the resources with new contents, new electronic articles and new virtual exhibitions, 'Collections' will be there to bring them to your screen.

Rossella Caffo

MICHAEL and the Italian Culture Portal: A Cooperation Model among National, Regional, and Local Institutions

Zusammenfassung

Durch den Impuls des *MINERVA* – Projektes startete der italienische *Minister für Kultur und kulturelles Erbe* ein breites Spektrum an nationalen Digitalisierungsinitiativen; die wichtigsten derzeit laufenden Projekte sind die Realisierung eines italienischen Interface des *MICHAEL* Service und das Italienische Kulturportal.

Die Notwendigkeit, ganz Italien abzudecken und sowohl nationale als auch kleine lokale Kulturinstitutionen mit einzubeziehen, führte dazu, dass das Ministerium die Arbeit beider Initiativen koordinierte und gleichzeitig dezentralisierte, indem alle Landesbehörden des Ministeriums selbst wie auch die Regionen, die Universitäten und das Bildungsministerium beteiligt wurden. Alle genannten Institutionen wurden aufgerufen, digitale Inhalte zu identifizieren und zu produzieren, die nun sowohl für *MICHAEL* als auch für das Italienische Kulturportal zur Verfügung gestellt werden.

MICHAEL und das Italienische Kulturportal stellen integrierte Initiativen dar, weil sie nach den gleichen Standards arbeiten: *MICHAEL* beschreibt die digitalen Sammlungen, während das Kulturportal die einzelnen Gegenstände anbietet – der Ausgangspunkt dabei ist ein und derselbe. Beide Initiativen decken alle kulturellen Bereiche ab: Bibliotheken, Museen, Archive, wie auch Radio (*RAI*) und Film (*Archivio Luce*).

Die von *MICHAEL* gesammelten Daten machen viele, bislang unbekannte digitale Ressourcen sichtbar, von denen die meisten ins Italienische Kulturportal aufgenommen werden. Der integrative Ansatz ermöglichte dem Ministerium, alle kulturellen Bereiche des gesamten Landes zu erfassen. Es leistete damit einen konkreten

Abstract

Under the impulse given by the *MINERVA* project, the Italian *Ministry for Cultural Heritage and Activities* started up a wide range of national digitisation initiatives; the main current initiatives are the implementation of the Italian interface of the *MICHAEL* service and the Italian Culture Portal.

The need to cover the whole territory and involve both national and small cultural institutions led the Ministry to organise the work of both initiatives in a coordinated way and on decentralized basis, involving the territorial offices of the Ministry itself, the Italian regions, the universities through their national coordination and the Ministry of Education. All of these institutions were called to find out and produce digital content that will be made available for both services, *MICHAEL* and the Culture Portal. *MICHAEL* and the Italian Culture Portal are integrated initiatives since they share the same standards; *MICHAEL* describes the digital collections, the Portal the single items. The access point is unique. Both cover all the cultural sectors: libraries, museums, archives, as well as broadcasting (*RAI*) and audiovisual (*Archivio Luce*).

The data gathered by *MICHAEL* highlighted many unknown digital resources, and most of them are being taken into consideration for the Italian Culture Portal.

The integrated approach allowed the Ministry to cover all the cultural fields and the whole national territory, so as to give a concrete contribution to the implementation of the *European Digital Library* since the digital content produced and published is based on open international standards.

Beitrag zum Aufbau der *European Digital Library*, weil der digitale Inhalt nach offenen internationalen Standards produziert und publiziert wurde. Auch die Kulturinstitutionen profitieren hiervon: die Kooperation auf nationaler Ebene erlaubt es ihnen, ein größeres Publikum zu erreichen, ihre Arbeit bekannt zu machen, und erleichtert den Zugang zu ihren Sammlungen und digitalen Aufbereitungen.

The cultural institutions are receiving benefits, too: the cooperation at national level allows them to reach wider audiences, promote their work and the access to their collections and digital content.

The background of the activities of *MICHAEL-IT* and *CulturaItalia* is based on the results of *MINERVA* and the two initiatives are considered as their spin-offs. Under the impulse of *MINERVA*, the Italian *Ministry for Cultural Heritage and Activities* started up a wide range of digitization initiatives, as described in the following figure.

MICHAEL-IT: the Italian portal of MICHAEL

MICHAEL-IT is a multilingual service providing a unique access point to the Italian digital cultural collections following an integrated cross-domain approach that is involving all the sectors of cultural heritage and memory institutions. The *MICHAEL-IT* service in fact allows to search and browse digital cultural information coming from all the cultural heritage institutions (archives, libraries, museums, institutes for administration and safeguarding, centers for research and education, cultural projects...) at different levels and sizes (national, regional, local, large and small, public and private, ...), covering all cultural fields (art, archaeology, architecture, landscape, science, tourism...).

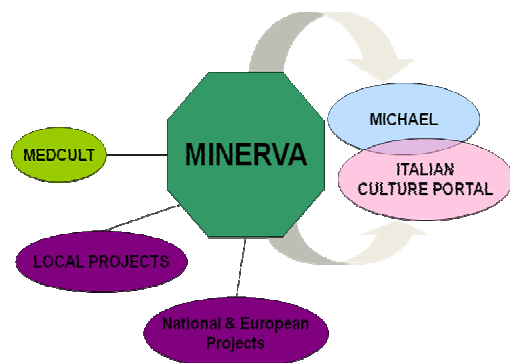


Figure 1: Background of *MICHAEL-IT*, *CulturaItalia* and *MINERVA*

MICHAEL-IT surveys the Italian cultural collections all over the territory: so far, the Italian portal provides access to 2489 digital collections, but I would like to place the attention also on the other numbers related to the project: 1346 institutions, 1356 services and products, 1424 physical collections, 410 projects and programs.

Thanks to the large content base recorded in the Italian instance, *MICHAEL-IT* is considered as one of the most relevant action lines of the Italian cultural portal: it represents the catalogue of the Italian digital collections and in this light it is the most important instrument to discover the existing Italian digital collections that are scattered in many different institutions.

The actors of MICHAEL-IT

I would like to describe now the organizational model we put in place in Italy to develop the work of *MICHAEL-Italy* together with the Italian cultural portal.

Coordination is done by the *Ministry for Cultural Heritage and Activities*, in particular by *OTeBAC* (*Osservatorio Tecnologico per i Beni e le Attività Culturali*), that is the Observatory monitoring technologies for cultural heritage and activities, instituted by the Italian Ministry in 2005. *OTeBAC* activities are carried out by a small group of experts that operate

within the Ministry under my chair. Among these activities, there is the coordination at national level of both *MICHAEL-Italia* and the Italian Cultural Portal *CulturaItalia*.

The Ministry of culture works in coordination with other ministries and in particular with the Ministry of Public Education, in order to explore the opportunities of exploitation of the *MICHAEL* services within the educational sector. As regards the involvement of more than thousand institutions all over Italy, there are three pillars worth to be mentioned:

- the cultural institutions directly depending on the Italian *Ministry for Cultural Heritage and Activities*: these are more than 600, amongst which are 400 museums, 47 state libraries, more than 100 state archives;
- the cultural institutes depending on regions: in Italy we have 20 regions that have competence on local museums, local libraries and local archives. The *Ministry for Cultural Heritage and Activities* established an agreement in order to cover all these regions and now all the Italian regions are working on the census of *MICHAEL* to survey the digital collections that exist on their territory;
- the libraries and museums of the 77 Italian universities.

In order to achieve such a large participation of the Italian cultural institutions, a communication campaign to join the *MICHAEL* service was launched. It has been addressed to all the cultural institutions that wish to give more visibility to their digital collections. The campaign is still running and will continue along the time to maintain the participation into the project active.

CulturaItalia

As already said, the implementation of the *MICHAEL-IT* service is strictly connected with the implementation of the Italian culture portal *CulturaItalia*.

CulturaItalia is a multilingual and integrated point of access to the Italian cultural heritage of all sectors, aggregating the metadata from the existing digital libraries in order to create the Digital Library of the Italian cultural heritage. The development strategy of *CulturaItalia* is based on a tight cooperation among the widest public and private sectors, including: museum, libraries, archives, preservation and administrative offices, national, regional and local institutions as well as universities and other ministries. *CulturaItalia* is also working toward the integration with other initiatives and projects, among which are *MICHAEL-IT* and the regional portals. From the technical point of view, *CulturaItalia* is based on the *MINERVA* and *MICHAEL* recommendations, guidelines, results.

In order to regulate the cooperation amongst the various institutional actors who are providing metadata of their digital content, a set of agreements was established. Up now, agreements were made with the following institutions: all institutes and offices directly depending on the Italian *Ministry for Cultural Heritage and Activities*; the 20 Italian regions, the 77 Italian universities, the *Ministry of University and Research*; the *Library of the Senate of the Italian Republic*; the *conference of the Italian Bishops*; the *Biennale of Venezia*; *BAICR*; *CINECA*; *EXIBART*; *Fratelli Alinari*; *Indire: Istituto Luce*; the *Querini Stampalia: Teche RAI*; *Touring Club Italiano*. The group of the metadata providers is constantly growing.

Contents of *CulturaItalia* is of three different types: indexes of metadata referring to the actual digital cultural objects owned by the single institution; selected and indexed cultural websites; editorial content (focus, short articles, itineraries, interviews...).

Some words about the technical tools. *CulturaItalia* is organized as a unique index of data and metadata harvested from the repositories of different providers, i.e.: ministries' institutions, private institutions, regions and universities. Each provider holds the rights of the digital objects described in the portal and remains responsible for the maintenance and updating of the source of information.

A specific Application Profile (*PicoAp*) has been developed for *CulturaItalia* to describe the digital objects and the tangible resources of all the cultural heritage sectors. Metadata are used to describe the resources both at item and at collection level. *PicoAp* is based on *Dublin Core* metadata standard.



Figure 2: structure of the index is aligned with DC.Culture

Through the metadata harvesting and the implementation of *PicoAp*, *CulturaItalia* is therefore able to aggregate the digital resources of hundreds of Italian "digital libraries". The index of metadata is structured as a tree and gathers references to digital and digitised "objects" of the Italian cultural heritage of all the sectors. These objects are intended in a broad sense: works of art, monuments, digital images, audiovisual, events, web pages... The metadata index of *CulturaItalia* can be considered as the digital library of the Italian cultural heritage. The structure of the index is aligned with DC.Culture, and can be represented as shown in the left figure. The campaign "Aderisci a *CulturaItalia*" (Join *CulturaItalia*) is planned to be launched in autumn 2007, in parallel with the campaign to join *MICHAEL-IT*.

They are both addressed to public and private cultural institutions that wish to share their data and metadata within the national cultural portal. The main objective of the campaign is to increase the number of metadata providers according to a cooperative model.

The current work for the implementation of *CulturaItalia* was carried out by *OTeBAC* and was focused in three directions:

- defining the agreements with the public and private partners interested in joining the project;
- identifying the databases owned by the partners that should be harvested by *CulturaItalia*. In this context *MICHAEL* results were very useful;
- defining the level of data publication through the portal (including the management of IPR, in particular access to and re-use of the content).

The workflow followed by *OTeBAC* and the partners includes also some other technical aspects. As far as the harvesting of the source databases is concerned, there are two ways of action: metadata can be harvested automatically, if data provided are *Dublin Core* compliant from the beginning; on the contrary, and this is the most common case, the data model of the source database must be mapped before with *CulturaItalia* Application Profile.

Further, the provider must install on its servers an *OAI-PMH* compliant repository and the *Ministry for Cultural Heritage and Activities*, through *OTeBAC*, can provide "ready-to-use" technical solutions to help the installation.

Finally, the provider must create a script to generate the XML metadata records, exporting values from the source databases. All this work has a cost: so far the funds for carrying out these activities were provided by the Italian *Ministry for Cultural Heritage and Activities*.

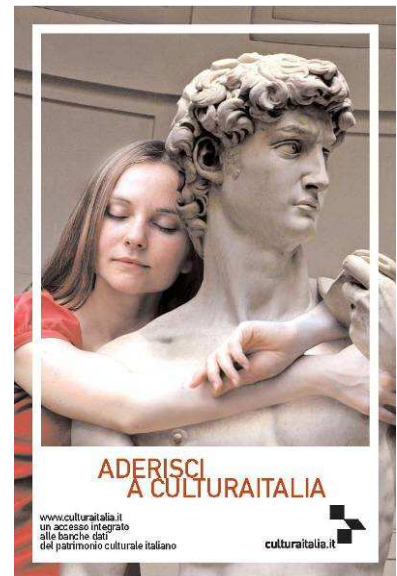


Figure 3: Image choose for the campaign "Aderisci a *CulturaItalia*"

MICHAEL-IT and *CulturaItalia*: similarities and differences

The two projects share the same actors. *MICHAEL-IT*, started in 2004, set up a wide network on the whole Italian territory, involving, as already said, all the cultural heritage sectors. *CulturaItalia*, started in 2005, could benefit from the network set up by *MICHAEL-IT*, taking full advantage of the same organization model.

The two projects share offices, experts and databases. The experts who carried out the census of the digital collections for *MICHAEL-IT* contributed also to the identification of the databases whose metadata are harvested into *CulturaItalia*. We can say that *MICHAEL-IT* "discovered" a lot of databases and provided an overview of the situation of digitization in Italy.

The two projects share the same standards. *MICHAEL-IT* and *CulturaItalia* are both based on open source software platforms and the *MINERVA* recommendations and guidelines are at the basis of these software platforms. Both data models are aligned with *Dublin Core* and metadata are harvested through *OAI-PMH*. Both platforms are based on a distributed architecture, whose technical integration is underway.

Starting from all the above mentioned common points and even if *MICHAEL-IT* and *CulturaItalia* are conceived since the beginning as integrated projects, it is worth to remind that, however, their goals are different. *MICHAEL* describes digital collections; *CulturaItalia* gives access to digital and digitised objects through their metadata. In this light, we can say that *MICHAEL* is the catalogue, while *CulturaItalia* is the "digital library" of the Italian cultural heritage.

Small cultural institutions may have a lot of advantages from the development and integration of *MICHAEL-IT* and *CulturaItalia*: thanks to them, they will be able to reach a wider audience at national and international level, increasing their online profile and the possibility to inform their users about the collections and digital objects they own, as well as making the access to their heritage and digital content (data and metadata) easier.

MICHAEL and the European Digital Library

So far, the most important official document concerning the EU plan towards the *European Digital Library* is represented by the Council's Conclusions on the EC Recommendation.

On the 13th of November 2006, in Brussels, the Council of the Ministers of Culture of the *European Union* adopted the text about *Digitisation and Online Accessibility of Cultural Material, and Digital Preservation* where it is stated that, in order to build the *European Digital Library*, two main pillars are identified:

- *CENL* and the service "*The European Library*" that provides access to the national libraries across Europe,
- *MICHAEL* and the European portal of the digital collections.

The Finnish Presidency worked very hard to reach this point. In this line, we can say that:

- *MICHAEL-IT* is the catalogue of the "*Digital Library of the Italian cultural heritage*;
- *MICHAEL-EU* is the catalogue of the *Digital Library of the European cultural heritage*.

In conclusion, thanks to the use of international standards and guidelines as well as open source software, *MICHAEL-IT* and *CulturaItalia* can be considered as the bricks offered by Italy to the construction of the *European Digital Library*. The concrete contribution that *MICHAEL-IT* and *CulturaItalia* will bring to the *European Digital Library* can be summarized in the following three main points:

- *MICHAEL-IT* and *CulturaItalia* will bring to the *European Digital Library* the catalogue and the index of the Italian digital collections and resources,

- *MICHAEL-IT* and *CulturaItalia* will bring to the *European Digital Library* the involvement of the cultural institutions from all sectors and levels (national, regional and local);
- *MICHAEL-IT* and *CulturaItalia* will bring to the *European Digital Library* the successful experience of its organizational model based on a centralised coordination. This model is very helpful in a situation where many institutions are working in a separate way and on scattered geographical locations. In this context, the capability to offer them a central coordination providing standards, recommendation and help, including training courses organized at national level, is of fundamental importance. The organizational model of the two projects demonstrated to be beneficial for the success of the Italian initiatives and is offered to the *European Digital Library* as a validated reference.

For further information, please refer to the following web resources:

- The *MICHAEL* project web site: <http://www.michael-culture.eu>
- The *MICHAEL* European portal: <http://www.michael-culture.org>
- The *MICHAEL-IT* portal: <http://www.michael-culture.it>
- The Italian Culture Portal: <http://www.culturaitalia.it>

Marius Snyders

The European Digital Library initiative. Joining up the pieces, strategies and actions at European level

Zusammenfassung

Die Europäische Kommission hat die „Europäische Digitale Bibliothek“ als eines ihrer vorrangigen Ziele im *i2010*-Programm benannt, die integrierten Zugang zu digitalen Sammlungen von Museen, Archiven, Bibliotheken und anderen kulturellen Einrichtungen ermöglichen soll. Von besonderer Wichtigkeit für dieses Vorhaben sind gegenwärtig Fragen im Bereich der sektor-übergreifenden Zusammenarbeit hin zum Erzielen einer kritischen Masse und die Sicherstellung von Interoperabilität. Die Kommission nimmt sich dieser Fragen auf verschiedenen Ebenen an und gewährt ihre Unterstützung – politisch, strategisch und technisch.

Abstract

The *European Commission* has designated the “*European Digital Library*” as one of its *i2010* flagship projects, delivering integrated access to the digitised collections of museums, archives, libraries and other cultural organisations. Issues of immediate concern revolve around cross-domain collaboration, creating a critical mass, and ensuring interoperability. The Commission addresses and supports these issues through actions at different levels – political, strategic and technical.

Introduction

In 2005, the *European Commission* adopted a Communication, entitled *i2010*, which sets out a strategy for the Information Society in pursuit of the Lisbon agenda. The *i2010* strategy rests on three pillars:

- Creating a common information space which includes a series of regulatory issues around audiovisual content and services;
- Investments in research and technology and the contribution to competitiveness and growth; and
- Working to make our European society more inclusive.

In the last pillar, three flagship initiatives are embedded one of which targets the digital libraries. The aim of the Digital Libraries flagship is to create common, multilingual access to Europe's distributed cultural heritage. It is taking shape in a series of actions aimed at creating a framework within which various fronts should be tackled – including the networking of Europe's distributed information and knowledge.

Why would we want a flagship like this at European level? Firstly, there is a growing belief all European citizens should be enabled to access their individual and their shared intellectual heritage – and that this access should not be limited by geographic location, by age or by language.

It is also believed that we need to be able to do this not only today but to guarantee this into the future. For that we need new methods, systems and infrastructures for dealing with digital heritage – and that these need not be invented locally but should take shape by a common effort.

Thirdly, we know that digital cultural and scientific content has a creative and economic potential. It contributes to education and research, to new cultural and creative industries, and to new forms of cultural experience and tourism. It attracts new users and new demands and needs for making the content of our archives, libraries and museums not only accessible but also usable and exploitable on the Internet. Cultural and scientific content means potentially big business for Europe. It drives the emergence of media-rich services to use and re-use the material, and commercial interest for such content is shown at a growing scale.

The i2010 – Digital Libraries initiative and actions

In Europe, the different stakeholders and specifically Commission funded projects have been working on digital libraries related initiatives for at least 10 years. This covers both the networking of archives, libraries and museums (and working with digital collections that are not linked to any one particular institution), to develop new ways to access and use digital libraries. What is new today though is the political interest and the interest of the press and general public, whose imagination has been ignited through the increased online possibilities for using and sharing digital music, texts, maps and so on.

The political drive in the Member States started with a letter of the 6 Heads of State to the President of the Commission urging for European action in building a common European digital cultural access point. And that coincided very well with already ongoing plans in the *European Commission* to draft and publish a Communication on digital libraries issues. Since then, we have seen the harnessing of three levels of action – political, strategic and operational – with all three coming together in a visible framework that calls for actions by the *European Commission*, Member States, and the different stakeholder communities.

The 2005 Commission Communication led to a public consultation, followed by a Recommendation of the Commission to Member States in August 2006. That Recommendation sets out a number of actions for Member States to take forward regarding the digitisation and online accessibility of cultural content and the long term preservation. They answer to the legal, financial and technical challenges.

For example, if we are to avoid the creation of a 20th century black-hole in our digital memory, we need to find practical and equitable solutions to IPR and copyrights in the area of orphan works so that these do not disappear from view.

As for the financial side, various elements of digitisation are still expensive and labour intensive. So we need to explore approaches that are better coordinated so unnecessary and redundant digitisation is avoided. Incentives for commercial sponsorship and public-private partnerships also need to be examined and developed further if we are collectively to find the investments needed to make our analogue resources available digitally. Digital libraries are not solely the responsibility of the public sector. Internet media companies, publishers, online bookstores, search engines and technology providers have a crucial role to play, and not necessarily a bad one. Successful collaboration between public and private entities, in particular for digitisation, has shown different forms already, for example in Germany the collaboration of the *Bayerische Staatsbibliothek* with *Google* and the *Universitätsbibliothek Köln* with the *Bertelsmann Stiftung*.

Technically, we need systems that are hospitable to making the digitised cultural items discoverable, accessible and usable as well as longer-term solutions so that we can support new services and keep the resources available over time.

The Commission Recommendation has been taken up and discussed by the Member States, resulting in accompanying Conclusions of the Culture Council in November 2006. This is a significant step: the Commission's Recommendation and the Council Conclusions set out clear actions for Member States and for the Commission. The Council invited the Member States to:

- Reinforce national strategies and targets for both digitisation and digital preservation, through plans, indicators, financial planning and encouragement of public private partnerships;
- Reinforce coordination between and within Member States by setting up coordination mechanisms, share information and common criteria for content selection;
- Contribute to the *European Digital Library*, by building frameworks for stakeholders to bring their content into one, common multilingual access point;
- Monitor and report on progress to the EU;
- Address the legal and other factors conditioning access to digitised and digital material and to improve these by considering how to remove barriers to digitisation progressively for orphan works, out of print and out of distribution materials. Also, by investigating the legal frameworks for digital preservation and by integration of technical standards for digital preservation in the business processes and workflows of organisations.

At the same time, the *European Commission* was invited to act on a number of fronts as well. The Commission:

- Coordinates and brings together stakeholders in such way that the right progress will be made, particularly regarding the *European Digital Library*;
- Supports technical and coordination work through its funding programmes, *eContent-Plus* and *FP7*;
- Proposes solutions to problems such as orphan works, indicating if additional EU-level action is needed;
- Has set up and chairs a Member States group to monitor and steer the process of implementing the actions.

What the Recommendation and Council Conclusions offer is a remarkably clear framework for actions in Digital Libraries. They also present concrete timeframes, aimed at achievements within the *i2010* horizon. And that is a very important message: they indicate actions which are practical and realistic; they support the engagement of the Member States and give a clear commitment to monitoring and reporting on progress in a structured framework. Making this a reality calls upon the concerted effort of a wide and disparate number of stakeholders, acting in different contexts, all contributing to different parts of the puzzle. Pulling together the different pieces of the jigsaw is a significant challenge.

The *European Commission* realises that the impact and results must be measured, should we want to show success and improve on it. Since 2001, since the launch of the Lund action plan, some methods have been practiced already by the Member States and the Commission to monitor digitisation activity. And although these proved somewhat sufficient in terms of monitoring a general level of achievement, with the recent European actions in the context of the Recommendation and the Council Conclusions there is a need to measure more precisely and quantitatively the input and output of the digitisation efforts in the *European Union*.

Member States agreed to gather information about current and planned digitisation efforts and to measure their progress and output. The *European Commission* will support this by its *NUMERIC*-study that will run for two years and that will develop relevant statistical methodologies and indicators for Member States and *Eurostat* to take up and integrate in their cultural data collection exercises.

The European Digital Library

The Council Conclusions highlight also the *European Digital Library*. The *European Digital Library* plays an important role in the *European Commission's* strategy where we realise that it is not enough to talk about the problems, but that we need rapidly to operationalise solutions and services.

The *European Digital Library* is not a library in the traditional sense of the word. The statements and the vision are clear. It will address all types of cultural material and encompass items from all kinds of cultural institutions. So, these come not only from national libraries but from all kinds of cultural organisations at all geographic levels. It clearly extends beyond the boundaries of individual cultural institutions. And the *European Digital Library* should also make more modern works searchable and discoverable.

At the operational front, Viviane Reding, European Commissioner for Information Society and Media, has set as a clear objective that a first version of the *European Digital Library* will be established by the end of 2008. *EDLnet*, a network co-funded under the *eContentplus* Programme and bringing together stakeholders from 25 EU Member States will be at the core of this. It builds upon a large base of existing experience and initiatives and includes individual libraries, archives and museums, umbrella organisations and key-members of projects such as *TEL* and *MICHAEL*.

At the strategic level, the Commissioner has established a High Level Group of multiple stakeholders able to discuss strategies for dealing with contentious and more intractable issues.

Finally, on the technical side, the Framework Programme for Research and Development has continued to stay with digital libraries. Relevant research topics revolve around mass and specialised digitisation concepts, the improvement of factory processes of the digitisation chain, scalability, reduction of costs and digital preservation, and on creative use and experience of cultural content. The Commission has identified the need to network competences, in order to arrive at a virtual European network of competence centres to especially address issues of digitisation — improving processes and to add value to the content, tools, skills and outreach — for all types of content. Other issues revolve around access and digital library research, bridging research and practice, for retrieval, search engines and multilingual access; and further work on interoperability.

Conclusions

Digital Libraries is a challenge that Europe, with the support of many, has readily taken up. There is now a convergence of political will and technological potential that can make a radical shift in our ability to make our cultural knowledge and the resources of archives, libraries and museums digitally accessible. There is a real opportunity to all this. The political framework is clear but it is not an easy job. Contexts — strategically and operationally — are continuously shifting and collaboration, creativity and flexibility are needed to bring the shared ambition for a *European Digital Library* closer to reality every day.

Jill Cousins

Creating the User's European Digital Library

Zusammenfassung

Der Vortrag wird eine kurze Geschichte über die *Europäische Bibliothek (The European Library)* als kollaboratives Unternehmen geben, sowohl in technischer als auch in organisatorischer Hinsicht. Die Infrastruktur dieses Netzwerks war Basis für die Entscheidung des Europäischen Rats, die *Europäische Bibliothek (The European Library)* für die Schaffung der *Europäischen Digitalen Bibliothek (European Digital Library)* zu nutzen.

Der Vortrag konzentriert sich auf die Notwendigkeit, auf dem aufzubauen, was bereits erreicht worden ist, wie etwa auf *MICHAEL*, und technisch wie auch organisatorisch weiter zusammenzuarbeiten, um die Ziele, die die Europäische Union in ihrer Presseerklärung vom August 2006 gesetzt hat, zu erreichen. Schließlich wird die vorgeschlagene „Cross-Domain-Foundation“ für Museen, Archive und Bibliotheken wie auch das thematische Netzwerk *EDLnet* skizziert. Beide zielen darauf, sowohl das menschliche und politische Umfeld wie auch technische und semantische Entwicklungen in den Prototyp einer benutzerfreundlichen „User's *European Digital Library*“ einzubringen, der Ende 2008 ans Netz gehen und „Cross-Domain- Access“ vorführen soll.

Abstract

A brief history of *The European Library* as a collaborative venture both in technical and organisational terms will be given. This networking infrastructure is the basis for the European Council's decision to use *The European Library* to create the *European Digital Library*.

The talk will then focus on the need to use what has already been achieved, such as *MICHAEL*, and to continue to work collaboratively, both technically and organisationally to be able to achieve the goals set out by the *European Union* in their press release of August 2006.

Finally the talk will outline the proposed cross-domain Foundation for Museums, Archives and Libraries and the thematic network *EDLnet* aims of creating both the human and political environment as well as the technical and semantic developments to deliver a “User's *European Digital Library*” prototype demonstrating cross-domain access for launch in the Autumn of 2008.

 <p>Creating the User's European Digital Library</p> <p>Jill Cousins The European Library Knowbynet, Berlin, June 2007</p> <p><small>Project for Bibliothek Europa, The European Library</small></p>	 <p>The European Library Timeline</p> <ul style="list-style-type: none">Started as TEL a project funded by the EU and led by The British Library – now fully operational serviceOwned by CENL, funded solely by the National Libraries, The European Library is live since 17 March 2005Digital Collections from 24 National Libraries - Austria, Britain, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Italy-Rome, Italy Florence, Latvia, Lithuania, Netherlands, Malta, Poland, Portugal, Serbia, Slovakia, Slovenia, Spain, Switzerland34- end 2007 remaining 9 EU & EFTA National Libraries and the Russian State Library.2008 – Adds 8 more Eastern European countries through Swiss Development Agency money and Bulgaria and Romania via TELplus = 44 of 47 national libraries in CENL by end of 2008. <p><small>Project for Bibliothek Europa, The European Library</small></p>
---	---

Organisation - the collaborative network

The diagram illustrates the organizational structure of the European Library. At the top is CENL, which oversees The European Library Management Board, The European Library Executive Group, and The European Library OFFICE. These three entities are interconnected with four working groups: Contacts Working Group, Editorial Working Group, Language Working Group, and Technical Working Group. The connections are shown as a network of lines, indicating a collaborative and interconnected structure.

A screenshot of the European Library website interface. The page features a search bar at the top, navigation tabs (Home, About, Search, etc.), and a main content area with search results and a sidebar with additional information and links.

A screenshot of a search result page for 'Wagner'. The page displays a list of search results on the left and a detailed view of a specific work on the right, including a title, author, and a small image.

A screenshot of a search result page for 'Geometrischer plan des Koeniglichen Thiergartens'. The page features a large, detailed image of a geometric plan of a royal garden, with search results and navigation elements visible around it.

A screenshot of a search result page for 'MARLENE DIETRICH & BETTY BOOP'. The page displays a large, colorful poster for the film 'CENTRO CULTURALE BEATO BERNO ASCONA' featuring the two actresses.

A network diagram illustrating the architecture of the European Library system. It shows a central cloud labeled 'http://SRU' connected to various components: 'Library X server (SRU)', 'Case 3: local SRU/Z39.50 gateway' (Library Y server (Z39.50)), 'Library Z server (Z39.50)', 'Case 1: Central Index TEL server (SRU)', and 'Case 4: Central SRU/Z39.50 gateway'. A 'lohengrin' search engine is also shown connected to the central cloud.

A screenshot of the British Library integrated catalogue. The page shows search results for 'The European Library', including titles, authors, and publication information.

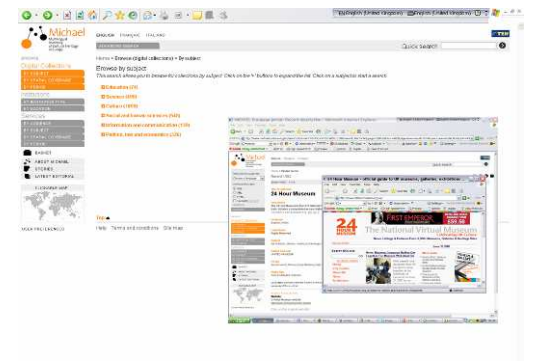
Creating the User's European Digital Library

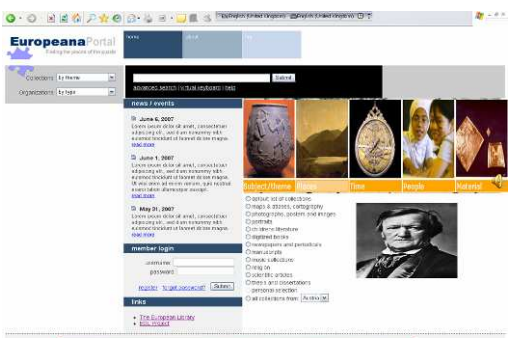
The European Library

EU Vision

- Access** Across Museums, Archives, Libraries initially – then other data or audio or visual repositories in the cultural heritage arena
- Digitisation** Centres of Competence
- Preservation** Centres of Competence

The European Library





By November 2008

So how are we proposing to get there?

- ➔ EDL Foundation
- ➔ Thematic Network EDLnet

The European Library

The EDL Foundation

Purpose:

- ➔ to provide access to Europe's cultural heritage.

The Foundation will:

- ➔ facilitate formal agreement across museums, archives, audio-visual archives and libraries on how to cooperate in the delivery and sustainability of a joint portal.
- ➔ provide a legal framework for use by the EU for funding purposes and as a springboard for future governance.

The European Library

The EDL Foundation

Timings

- ➔ June 29 - Small meeting of major European Associations to discuss concept and first draft of statutes for Foundation
- ➔ Summer 07 – agreement on statutes and cross domain executive committee
- ➔ September 12 - Founder Meeting & Proposal for Name
- ➔ December 07 – Legal Foundation in place

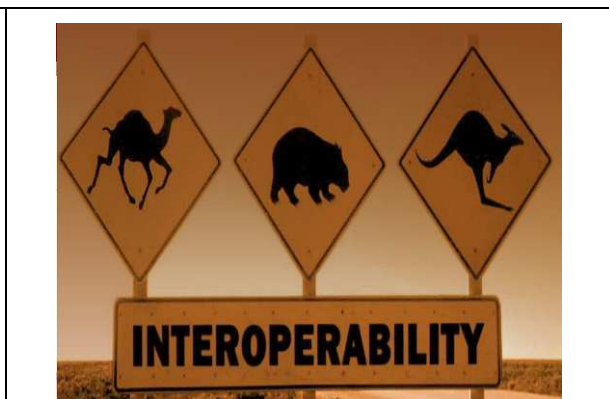
The European Library

EDLnet

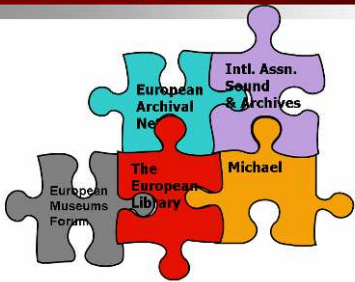

A thematic network for cultural heritage domains
 Led by TEL Office as part of the KB - €1.3 M – 24 m.

- ➔ Starts 1 July 07
- ➔ July 5 – Work Package leaders meeting
- ➔ EDLnet Office, under the KB
- ➔ Thematic Network across 4 domains Museums, Libraries, Archives and Audio-visual Archives
- ➔ Creates a coalition of the willing
- ➔ Creates Organisational Structure for a European Digital Library
- ➔ Tackles domain level interoperability issues
- ➔ Delivers prototypes against stakeholder and user requirements

The European Library **By November 2008!**



<h3>EDLnet Work Packages</h3> <ul style="list-style-type: none"> ➤ Human, Political & Intercommunity Interoperability <ul style="list-style-type: none"> - Governance, Organisational Structure, a Roadmap and eventually the Business Model - Led by MLA UK and ABM Centrum, Sweden ➤ Technical & Semantic Interoperability <ul style="list-style-type: none"> - Prototypes of possible architectures, creating solutions for various interoperability issues of access, metadata schemas, language, search and display - Led by experts in the field ➤ Users for usability <ul style="list-style-type: none"> - Providing recommendations, testing maquette and prototypes with users - Led by Inst. Audiovisuel, France <p>The European Library</p>	<h3>EDLnet Online</h3> <ul style="list-style-type: none"> ➤ www.European Digital Library.eu 'show & tell' Sept. 07 <p>Based on v 1.5 The European Library Different look and feel No Catalogues Access at item level Fully Digitised Material Only Themed Collections</p> <ul style="list-style-type: none"> ➤ EDLnet Thematic Network site ➤ Maquette - functionality & features desired by 4 domains and users - January 2008 ➤ Prototypes 1 and 2 April to August 2008 ➤ Final Prototype November 2008 <p>The European Library</p>
---	--

<h3>Filling up the jigsaw</h3>  <p>The European Library</p>	<h3>European Digital Library</h3>  <p>The European Library</p>
---	---

David Dawson

MICHAEL and the European Digital Library: promoting teaching, learning and research

Zusammenfassung

Der Aufbau der *EDL* braucht die Entwicklung einer soliden technischen Infrastruktur. Die *Europäische Kommission* hat diesen Prozess mit der Entwicklung einer Reihe von Interoperabilitätsprinzipien begonnen. Diese werden die technische Weiterentwicklung leiten und eine Anleitung bereitstellen, wie Museen, Bibliotheken und Archive sich an der *EDL* beteiligen können.

Das *MICHAEL*-Projekt ist ein Kernbestandteil der Infrastruktur der *EDL*, und die Leistungen, die das Projekt bereits jetzt für den Forschungs- und Bildungsbereich erbringt, werden durch einige Fallbeispiele veranschaulicht.

Abstract

Building the *European Digital Library* will need the development of a sound technical infrastructure. The *European Commission* have begun this process, with the development of a set of Interoperability Principles that will guide technical developments, and will provide a route-map showing how museums, libraries and archives can participate in the *European Digital Library*.

The *MICHAEL* project is a core part of the infrastructure for the *EDL*, and the services that the project already provides to the research and educational sectors will be illustrated by a number of case studies.

MICHAEL and the European Digital Library: promoting teaching, learning and research

David Dawson
Museums, Libraries and Archives Council, UK



Researcher

- Associate Professor from Canada

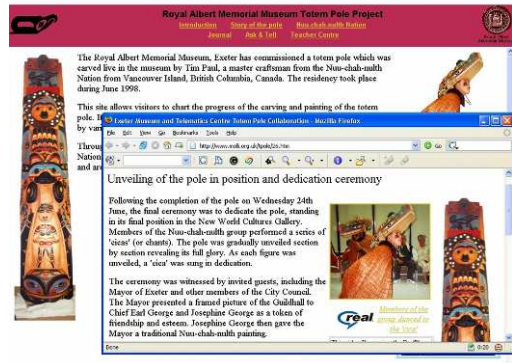


Users of MICHAEL

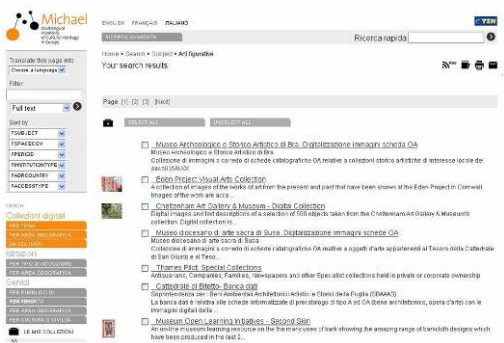
- many different user communities
 - education
 - cultural tourism
 - research
 - 'co-ordination'



The screenshot shows the MICHAEL search results page. The search term 'MICHAEL' is entered in the search bar. The results list several items, including 'MICHAEL: Project Metadata' and 'MICHAEL: Project Metadata'. The interface includes navigation tabs for 'Home', 'About', 'Contact', and 'Help'. The search results are displayed in a list format with various details for each item.



Subjects



What will the EU Digital Library be?

- a common multilingual access point to Europe's distributed
- held in different places by different organisations - digital cultural heritage,
- all types of cultural material
 - texts, audiovisual, museum objects, archival records etc
- targeted at delivering rapidly a critical mass of resources to the users.



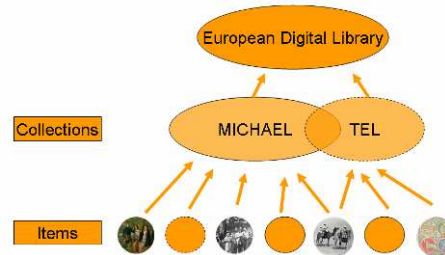


Building blocks

- The European Library (TEL) as a gateway to the collective resources of national libraries across Europe
- The work ongoing in the Michael and Michael Plus projects in describing and linking digital collections of museums, libraries and archives from different Member States and providing access to these collections

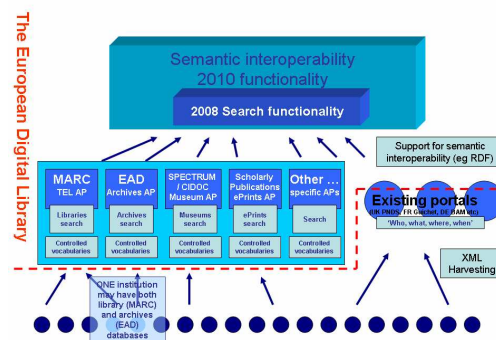


MICHAEL and the European Digital Library



How is the EDL going to work?

- Interoperability Working Group established by the Commission
- Representatives and experts from all sectors
- Discussing high-level issues
- Being taken forward by EDLNet
 - MLA leading workpackage to involve all stakeholders



Licensing Policies

- **Users** need to know what use conditions and rights apply to a resource
- **Administrators** need to know what use conditions and rights apply to a resource
- **Systems** need to know what use conditions and rights apply to a resource
- **Every digital resource has use conditions and rights associated with it:**
 - even if it is freely available
 - even if it is public domain
 - even if it is not obvious what they are
- **Make the use conditions and rights in your resources explicit:**
 - use a recognised machine readable license if you are able to
 - state the use conditions and rights that apply to your resources in the metadata
 - use the basic controlled vocabulary (to be) developed by the European Digital Library, and use it consistently
 - look at existing frameworks for expressing use conditions and rights

A quality website must:

- be transparent, clearly stating the identity and purpose of the website, as well as the organisation responsible for its management
- select, digitise, author, present and validate content to create an effective website for users
- implement quality of service policy guidelines to ensure that the website is maintained and updated at an appropriate level
- be accessible to all users, irrespective of the technology they use or their disabilities, including navigation, content, and interactive elements
- be user-centred, taking into account the needs of users, ensuring relevance and ease of use through responding to evaluation and feedback

transparent · effective · maintained · accessible
 user-centred · responsive · multi-lingual
 interoperable · managed · preserved

Enabling the EDL to promote your content

The EDL will help you to promote your content to new audiences. Get involved! Start planning!

Your content

- **Technical standards**
 - Create your content using open standards
 - use the MINERVA Technical Guidelines
- **Metadata**
 - Use your existing cataloguing standards
 - Prepare to map your data to domain-specific Dublin Core Application Profiles – use the European Metadata Registry
- **Terminologies**
 - Prepare to take the opportunity offered by the semantic web
 - Publish your terminologies and thesauri using SKOS
- **Registries**
 - Make your content visible
 - register your content and services with existing registries - such as TEL and MICHAEL
- ...

Thank you

david.dawson@mla.gov.uk



Christophe Dessaux

MICHAEL-culture.org – A Gateway for Collections: Open Access through Open Source Technology. How MICHAEL's Technique Works

Zusammenfassung

Der *MICHAEL* European Service stellt den Zugang zu vielsprachigen Beständen des digitalisierten kulturellen Erbes in Europa zur Verfügung. Um dies zu gewährleisten, werden Daten von regionalen und nationalen Beständen zusammengetragen, wie sie von den Mitgliedsländern zur Verfügung gestellt werden. Deren Bereitstellung durch *MICHAEL* umfasst eine open source software-Plattform, die auf Standards, einem gemeinsamen Datenmodell und der Organisation der dezentralisierten Architektur des *Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)* basiert, um die Inhalte der einzelnen regionalen und nationalen Institutionen zusammenzuführen.

Der Vortrag konzentriert sich auf die technischen Aspekte des *MICHAEL*-Projekts und seiner Ziele. Die multilingualen Funktionalitäten, die einen so wesentlichen Teil des Service darstellen, werden ebenfalls im Detail vorgestellt.

Abstract

The *MICHAEL* European Service provides access to a multilingual inventory of the digitised cultural heritage in Europe. To achieve this, data is gathered from regional and national inventories provided by participating countries. *MICHAEL* release includes an open source software platform based on standards, a shared data model and an organisation based on a distributed architecture using the *Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)* to gather the contents developed by the separate regional and national instances.

The presentation will focus on the technical aspects of the developments provided by the *MICHAEL* project to achieve their goals. *MICHAEL* multilingual functionalities are a major part of the service and will also be detailed.

MICHAEL is a federation of national portals. It builds on national initiatives in order to give European access to national inventories of digitised collections. At the end of 2007, 20 countries participate in Europe: Belgium, Bulgaria, the Czech Republic, Cyprus, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Latvia, Malta, the Netherlands, Poland, Portugal, Slovakia, Spain, Sweden, the United Kingdom.

The European service provides a European vision of heritage thanks to a single multilingual trans-national search engine. How does it work?

The key issue here is interoperability: to ensure efficient linkage of a federation of several single national instances, *MICHAEL* has created an interoperability framework, which addresses three aspects: 1-semantic, 2-technical, 3-human and organisational.

Semantic interoperability

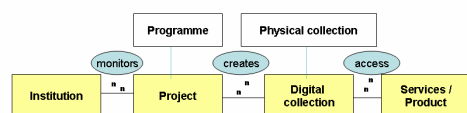
Semantic interoperability relies first on data standardisation. *MICHAEL* has created a common data model used by all partners to provide their information about digital collections. The data model is centred on the description of digital collections, it also gives information about institutions (responsible for the digital collections) and about on-line and off-line

services developed to access these collections. In order to ensure interoperability with other projects, it is aligned with the *Dublin Core* standard and available in XML (XML schema). The data model is complemented with common terminology lists that also implement existing standards (*ISO, UNESCO, Dublin Core, MINERVA* standards...).



Semantic interoperability

A common model for standard data



- centred on the description of **digital collections**
- aligned with Dublin Core standard
- available in XML (XML schema)
- complemented with common terminology lists (ISO, Unesco, Dublin Core, Minerva standards...)



Figure 1: Semantic interoperability

their language that will be integrated in the European interface to perform multilingual searches. In some cases (but this is not mandatory and depends on national choices), national instances can also provide additional human translation of their data, like for example the title¹ and the description of the collections.

The European portal provides multilingual access to the data. A multilingual search is available, using the multilingual dictionaries provided by the partners to extend the query from the language of the user to all the languages available in the portal. A multilingual display of the results is also partly provided to the user: the interface and the controlled vocabulary fields are translated (thanks to the multilingual terminology lists). Free text fields remain in original language but an automated translation is made available on demand using free internet tools.

Technical interoperability

Technical interoperability relies on the use of common tools for national instances and the European portal that communicate with each other through a common standard protocol (*OAI-PMH*).

A set of common software tools have been developed by *MICHAEL* in order to fulfil the functional needs of the project: a production module (*XDEPO*) for data management at national level, developed with an XML database (*eXist*); a platform for developing public interfaces to access the data (*SDX*), used for national portals and for the European portal, developed with Apache/Cocoon. All software developed by *MICHAEL* are based on standard Open Source softwares and are themselves Open Source projects (*CeCill licence*) available on Sourceforge.

The integration of all single national instances in a distributed architecture relies on a harvesting procedure using *OAI-PMH*. The Open Archive Initiative protocol for metadata harvesting (*OAI-PMH*) is a standard protocol for harvesting data. National instances provide their data through repositories compliant with this protocol (data provided both in *Dublin Core* and *MICHAEL* formats). The European portal harvests the repositories to provide a centralised data base of metadata used to process queries at European level through the multilingual interface of the portal.

¹ For example titles are translated in English and Italian in the French instance „Patrimoine numérique“.

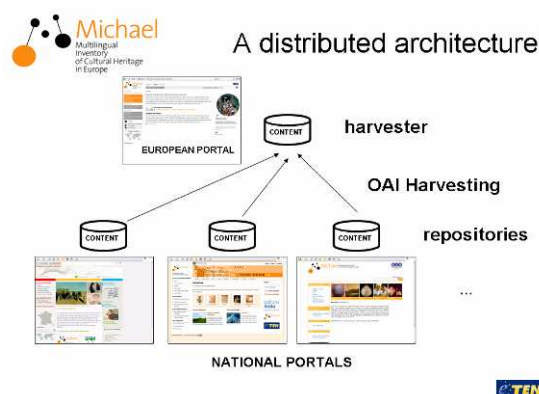


Figure 2: A distributed architecture

Paul Smith of this service in this conference. The same models are in use or planned in other countries: *MICHAEL* national instance is integrated in the *People's Network Discover Service*³ in the UK, in the national cultural portal *CulturaItalia.it*⁴ in Italy.

The European portal also provides an OAI repository and can be harvested by other portals. The re-use of *MICHAEL* information is thus possible for other services (*European Digital Library*, Education, Tourism etc).

This distributed model is also efficient at national level to ensure the integration of *MICHAEL* local instances in national strategies. For example, *Patrimoine numérique* (*MICHAEL* French instance) is harvested by *Collections.fr*², the unique access point to digitised heritage created by the ministry of culture, which provides complementary information at item and collection level [See the presentation made by

Human and organisational interoperability

Interoperability is not only a matter of technical architecture. Organisational and human aspects are critical for the sustainability and the quality of the service provided.

The association *MICHAEL* Culture was created in July 2007 to be the legal structure in charge of sustaining the service and maintaining the portal after the end of European funding (May 2008). It is an international not for profit organisation created under Belgian law (AISBL – association internationale sans but lucratif). All national instances are invited to join to consolidate the consortium of both *MICHAEL* and *MICHAEL+* projects.

The statutes and the internal regulations formalise the commitments of the partners to participate in the service, its maintenance and its developments. The commitment is not only financial (a fee is paid to the association for membership), it includes a quality charter. Semantic and technical requirements are part of this quality charter that is recognised by every *MICHAEL* and *MICHAEL+* partner and has to be accepted by any new member of the portal (and in the association). All partners also commit themselves to provide their data under a *creative commons* licence, which allows re-use of data in other portals for non-commercial purposes (the source of the information – *MICHAEL* project – has to be mentioned).

Perspectives

MICHAEL has been recognised as one of the major stakeholders in the field of access to digitised cultural heritage. It participates to the ongoing process of building the *European Digital Library*, being a partner of *EDLnet* project (starting in July 2007 and in charge of the development of the Europeana prototype due for November 2008) and participating as a Founder and a member of the executive committee of the "Foundation for the *European Digital Library*" (created in November 2007). *MICHAEL* provides a major contribution to Europeana with comprehensive and validated information on digitised collections from 20 European countries. Its interoperability framework facilitates its integration at all level in the ongoing process: informational, technical and organisational.

² <http://www.culture.fr/fr/sections/themes/collections>

³ <http://www.peoplesnetwork.gov.uk/discover/>

⁴ <http://www.culturaitalia.it/>

Petra Rauschenbach

International Standards for Digital Archives – The 3rd European Conference on EAD, EAC and METS (April 24-26th 2007)

Zusammenfassung

Die 3. Europäische Konferenz für digitales Archivgut über *EAD* (*Encoded Archival Description*), *EAC* (*Encoded Archival Context*) und *METS* (*Metadata Encoding and Transmission Standard*) fand vom 24. – 26. April 2007 in Berlin statt.

Sie lieferte nicht nur Informationen zu den Standards, sondern auch die Möglichkeit der Diskussion über bisherige Erfahrungen mit ihrer Nutzung und dem bisherigen Einsatz.

Die Nutzung der Standards ist eine wichtige Voraussetzung zum Austausch von Erschließungsinformationen und Metadaten zu digitalisiertem und digital entstandenem Archivgut für gemeinsame Portale. Der Artikel berichtet über Inhalte und Ergebnisse der Konferenz und deren Umsetzung.

Abstract

The 3rd European Conference for Digital Archives took place from 24 to 26th April 2007 in Berlin. The focus of the conference was *EAD* (*Encoded Archival Description*), *EAC* (*Encoded Archival Context*) and *METS* (*Metadata Encoding and Transmission Standard*).

The conference not only provided information on standards. In our discussions we also addressed hands-on experience using *EAD*, *EAC* and *METS*.

The exchange of collection information and appended metadata – the key prerequisite in establishing a common Internet portal – requires the use of agreed-upon standards. This observation is true of both digitised and *born-digital* materials. In her contribution Petra Rauschenbach reports on this recent international conference, providing us with a sense of the substance, results and next steps to be taken.

Introduction and Overview: Conferences hosted by the *Bundesarchiv* in April 2007

More than 350 participants from 35 nations worldwide followed the invitation of the *Bundesarchiv* to Berlin to join the conference on *EAD*, *EAC* and *METS*. The conference took place together with some other meetings in Berlin:

- the *XV Conference of the European Board of National Archivists (EBNA)*;
- the General Assembly of *EURBICA (European Branch of the ICA)*;
- a meeting of the *DLM Forum members*.¹

Heads and chief-executives of the national archives from the EU member states gathered to the *XV EBNA* conference in Berlin on April 26-27th.² The Minister of State and Federal Government Commissioner for Culture and Media, Mr. Bernd Neumann, welcomed the delegates from 26 EU member states, guests from Switzerland and Liechtenstein, and

¹ Further information (such as reports and minutes of the different conferences and meetings) can be found on the joint conference homepage at: <http://www.instada.eu>.

² A more detailed English conference report can be found at: <http://www.bundesarchiv.de/aktuelles/begegnungen/00178/index.html>.

representatives from the EU Commission as well as from the *International Council on Archives (ICA)*. The conference discussed about ways of how to reinforce international cooperation on disaster prevention management and how to promote a transborder network for this field of cooperation by following the example of the German *NORA*³ central emergency register. Other important subjects of discussion were the cultural importance of the *European Digital Library* project, the benefits of digitized archives and technical approaches for a long-term preservation of born digital records from public authorities.⁴

The 3rd European Conference on EAD, EAC and METS

Leading experts as well as users with enhanced practical experience in employing *EAD*, *EAC* and *METS* discussed at this conference innovative strategies of how to use the three standards for the creation of finding aids and the display of digitized archives on the internet. Many of the presentations were illustrated by examples of practical applications. The conference has been part of a series of meetings which dealt with the implementation of international standards for archives (the *Instada* conference series). The first meeting took place in 1999 in Kew (London) and focused on the use of *EAD*. The second meeting was held in Paris in 2004 and was devoted to the benefits of *EAD* and *EAC*. Now in Berlin, the third conference this year highlighted the advantages of using the three standards *EAD*, *EAC* and *METS* for innovative ways of presenting archival descriptive information together with digitized archives online.

*EAD (Encoded Archival Description)*⁵ is a standard for the encoding of finding aids, especially for use in a networked (online) environment. These online finding aids allow a structural navigation for content as well as a search for single words by browsing texts or using indices. Elements and attributes of *EAD* define typical parts of archival finding aids that can be presented online with the help of XSLT.

*EAC (Encoded Archival Context)*⁶ intends to facilitate the use of different names and notations of persons, organisations or locations by offering explanations of how various forms of these names may occur in archival holdings that can be searched online. *EAC* therefore forms the typical structure of explicit context information that may be part of descriptive information contained in archival holdings guides or finding aids.

*METS (Metadata Encoding and Transmission Standard)*⁷ allows the encoding of descriptive, administrative, and structural metadata of digital objects which may consist of one or multiple files or of a group of digital reproductions and can be linked to online finding aids.

The third European Conference on EAD, EAC and METS was opened with a speech held by the President of the *Bundesarchiv*, Professor Dr. Hartmut Weber, who reflected on the role of the internet for a modern and user-friendly archival service. The first working session dealt with the use of *EAD*. First key note speaker Kris Kiesling (U.S.) explained the development and the widespread use of *EAD* for creating online finding aids. As an example, Claire Sibille demonstrated the use of *EAD* in French archives, while Michael Fox (U.S.) discussed crosswalks between *EAD* and other structural standards such as *ISAD(G)* or *Dublin*

³ *NORA* is a German acronym for *Notfall-Register Archive* (translated: "Emergency Register for Archives") and serving as the name of this central online database; cf. <http://www.bundesarchiv.de/nora/php/registration.html>.

⁴ See for in-depth information the EBNA conferences minutes which have been published online at: <http://www.bundesarchiv.de/imperia/md/content/stab/ebnaprot0407.pdf>.

⁵ For more information see the official *EAD* Internet homepage at: <http://www.loc.gov/ead/>. See also the "EAD Help Pages" (published by the EAD Round Table of the Society of American Archivists) at: <http://www.archivists.org/saagroups/ead/>.

⁶ The standard resulted from the LEAF Project (LEAF is an acronym for "Linking and Exploring Authority Files"); see for further information: <http://xml.coverpages.org/ni2004-08-24-a.html>. The current beta version of *EAC* published by the Ad hoc *EAC* Working Group can be consulted online at: <http://jefferson.village.virginia.edu/eac/>.

⁷ Further information can be found online at the official *METS* Bureau homepage at: <http://www.loc.gov/standards/mets/>.

Core. The quadrilingual online finding aids created in Poland and presented by Hubert Wajs were commonly considered to be a good example of the use of *EAD* in the context of Europe's multilingualism. The application of *EAD* was further illustrated by Petra Rauschenbach, who informed the participants about the digitization of legacy finding aids at the *Bundesarchiv*, and by Amanda Hill, who presented the internet gateway '*Archives HUB*' set up by British university archives.

EAC was the topic of the second working session. It has been used only little so far though *EAC* functions as an important tool for the presentation of information about the originators of archival material. Key note speaker Daniel Pitti (U.S.) explained the development of *EAC* which was created by an international working group of archivists as well as librarians. Highlighting the advantages of information about the originators of archival material, Nils Brübach (Germany) discussed the relationship between *EAC* and *ISAAR(CPF)*, a corresponding standard defined by the *International Council on Archives*. Anke Löbnitz and Andreas Petter presented the current stage of the <daofind> project of the *Bundesarchiv* and introduced methods of how *EAD* and *EAC* may be combined for creating and maintaining online holding guides based on centrally as well as locally administered *EAD* and *EAC* files. Gerald Maier explained the use of international standards in the German *BAM Portal Project*.



Figure 1: Daofind

The third working session focused on *METS*. The session took place together with the *DLM Forum* and discussed how *METS* can be used as a standard for digital preservation and for managing access to digitized reproductions. Key note speaker Merrilee Proffitt underlined the "power of *METS*" for structuring, linking and sharing metadata of digital objects. Markus Enders (Germany) and Krystyna Ohnesorge (Switzerland) presented strategies of how to use *METS* for digital preservation. Angelika Menne-Haritz explained the use of *METS* and the combination of *METS*, *EAD* and *EAC* for the presentation of digitized archives from the *Bundesarchiv* on the internet.

The last working session was concerned with the use of structural standards in online gateways and discussed examples from the UK as well as from Spain and Italy. Bill Stocking presented the portal *A2A* set up by the *TNA* in London which has acted as a pioneer in using *EAD* and now offers cross-searchable information about approximately 10 million archival units from more than 400 British archives. As Blanca Desantes Fernandéz reported, a portal with data from more than 50.000 Spanish-speaking archives worldwide has been created in Spain with the "*Censo Guia*" using *EAD* and *EAG*. According to Stefano Vitali, Italian state archives today provide a comprehensive holdings guide together with many finding aids online, being well experienced in applying native XML databases and *SVG* for a geographically structured online access.⁸

The conference closed with a panel discussion about the need for a gateway to archives in Europe. Being chaired by Angelika Menne-Haritz, vice-president of the *Bundesarchiv*, Martine de Boisdeffre, Director of *Archives de France*, Kris Kiesling, chair of the *SAA EAD Working Group*, José Ramon Cruz Mundet, Director of the *Spanish Archives Administration*, and Frank M. Bischoff, Director of *Marburg Archives School*, joined the panel. All experts stressed the fact that a gateway to archives in Europe is highly recommendable and should be created as soon as possible.

In general, the conference revealed the outstanding importance of the international standards *EAD*, *EAC* and *METS* for archives wishing to offer online-access to their materials independently of time and space. The internet is undoubtedly an excellent medium for

⁸ For Papers of the conference speakers see online at: <http://www.instada.eu>, menu item "Speakers/Presentations".

services generally provided by archives. The key question is now how these three standards may be implemented or combined with each other in order to make archival holdings visible online (rather than questioning whether the standards should be used at all).

All participants received a CD-ROM with a first interim-version of the *MEX* software tools (*MEX: Midosa Editor for XML Standards*) created in the <daofind> Projects of the *Bundesarchiv*. With the support of *The Andrew W. Mellon Foundation*, New York, easy-to-use software tools are being developed for the creation of internet presentations with digitized archives in online finding aids. Current versions of these tools are provided for download at the <daofind>- project homepage.⁹ The software is conceived as a plug-in to the open source *Java* development platform *Eclipse* with user interfaces adapted to the needs of the archival workflow. Currently, there are three sample finding aids available online that are enriched with approximately 30.000 digital images of archival units. The finding aids are integrated into the search engine *MidosaSEARCH* which offers a simultaneous cross searching as well as structural navigation and allows the integration of finding aids into other access points.¹⁰ Thus, the <daofind>- project has already succeeded in creating a pilot application which employs the three standards *EAD*, *EAC* and *METS* for the publishing of archival descriptive information as well as digitized archives online by using an open source software.

New Perspectives: Designing a Reference Model for a Nationwide Internet Gateway to Archives

Over the last years, many archives have started to publish online finding aids for their users.¹¹ The *Bundesarchiv* is currently digitizing many of its archival holdings in order to present these materials online with the help of *METS*.¹² By integrating digitized archives into online finding aids the *Bundesarchiv* seeks to enable users not only to discover some information about the existence of a specific archival unit of description but also to study in detail the content of archival units on the internet. Moreover, the *Deutsche Forschungsgemeinschaft* (*German Research Foundation*) has recently decided to grant a financial support to the *Bundesarchiv* for the development of a reference model for a nation-wide internet gateway to archives by extending the internet portal "*Network SED/FDGB archives*" already established by the *Bundesarchiv* together with five different German federal state archives.¹³ The core of this new internet gateway¹⁴ will be a union finding aid which is designed to allow a cross-simultaneous, comprehensive research for archives down to the unit level in all the online finding aids provided by the participating archival institutions. Again, participating archival services may enrich their online finding aids with digitized reproductions of archival units. Users will thus be able to browse and read archival units virtually in the internet gateway. Technically, the union finding aid is planned to have several interfaces in order to combine metadata from the various IT-systems which are in use at the participating archival institutions. *EAD*, *EAC* and *METS* will serve as the target format as well as the central interchange format to other internet portals. In preparation, the *Bundesarchiv* defined a specific *EAD* profile for the internet gateway. The complete data encoded in *EAD*, *EAC* and *METS* may, for example, be integrated into the new *European Digital Library*. Furthermore, this

⁹ For more information about the project as well as for a download of the current interim version of *MEX* software visit: <http://www.daofind.de>. See also: <http://www.bundesarchiv.de/aktuelles/projekte/00154/index.html?lang=en>.

¹⁰ More detailed information about *MidosaSEARCH* can be found online at: <http://startext.net-build.de:8080/barch2/MidosaSEARCH/Help.htm>.

¹¹ An overview is given online at: <http://www.bundesarchiv.de/aktuelles/fachinformation/00046/index.html?lang=en>.

¹² Cf. note 9.

¹³ For more information, see (in German): <http://www.bundesarchiv.de/aktuelles/pressemitteilungen/00187/index.html>.

¹⁴ This *EAD* profile is published on the homepage of the <daofind> projects; see: <http://www.daofind.de>. For more information, see (in German): <http://www.bundesarchiv.de/aktuelles/begegnungen/00185/index.html>

internet gateway reference model will contain information about the participating archival institutions and an e-learning section for online users.

Members of the *Federal Board of State Archivists (Archivreferentenkonferenz)* gathered to a conference on July, 11th 2007 at the invitation of the *Bundesarchiv* in order to get introduced to the new *EAD* target format designed by the *Bundesarchiv* for the Internet gateway reference model. With this new project the way may be paved for a new quality of archival service provided in the cyber space.

Petra Rauschenbach

Internationale Standards für digitales Archivgut – Die 3. Europäische Konferenz über *EAD*, *EAC* und *METS*, 24.-26. April 2007

Einleitung und Übersicht: Konferenz veranstaltet vom *Bundesarchiv* im April 2007

Über 350 Teilnehmer aus 35 Ländern weltweit folgten der Einladung des *Bundesarchivs* nach Berlin, das ist ein Novum!

Die Konferenz bündelte ein Paket mit mehreren Veranstaltungen:

- *European Board of National Archivists (EBNA)*, Mitgliederversammlung
- *European Regional Branch of ICA (EURBICA)*, Generalversammlung
- *DLM-Forum*, Mitgliederversammlung

Lassen Sie mich kurz auf die *EBNA (European Board of National Archivists)* eingehen. Die Leiterinnen und Leiter der Nationalarchive in den Staaten der *Europäischen Union* trafen sich zu ihrer 15. Tagung. Kulturstaatsminister Bernd Neumann konnte Delegationen aus 26 EU-Staaten und Gäste aus der Schweiz, Liechtenstein sowie Vertreter der EU-Kommission und des Internationalen Archivrates begrüßen. Wichtige Themen waren eine engere Zusammenarbeit beim vorbeugenden Katastrophenschutz und eine grenzüberschreitende Vernetzung nach dem Vorbild des deutschen Notfall-Registers-Archive NORA. Hervorgehoben wurde auch die kulturpolitische Bedeutung der *Europäischen Digitalen Bibliothek* und von digitalisiertem Archivgut sowie die Ansätze zur Langzeitsicherung digitaler Informationen aus Verwaltungsprozessen.

Die 3. europäische Konferenz über *EAD*, *EAC* und *METS*

Führende Experten sowie Anwender mit praktischen Erfahrungen im Einsatz der Standards hatten für Vorträge gewonnen werden können und stellten ein intensives Programm mit grundsätzlichen Erläuterungen zu den Standards und mit praktischen Beispielen aus ihrem Einsatz vor. Die erste dieser Konferenzen fand übrigens 1999 im National Archives in Kew (London) statt und behandelte insbesondere *EAD*. Die zweite Konferenz beschäftigte sich 2004 in Paris mit *EAD* und *EAC*. Auf der dritten Konferenz in Berlin diskutierten die Teilnehmer aktuelle Möglichkeiten der Anwendung und Weiterentwicklung der drei internationalen Standards *EAD*, *EAC* und *METS* für moderne, innovative Formen der Erschließungspräsentation im Internet.

EAD (Encoded Archival Description) wird für die Herstellung und Präsentation von strukturierten Online-Findbüchern verwendet, in denen durch die Gliederung wie in einem Inhaltsverzeichnis navigiert, nach Begriffen gesucht, Indexbegriffe genutzt und der Text

durchblättert werden kann. Die Elemente und Attribute der *EAD*-DTD bezeichnen typische Teile von archivischen Findbüchern. Mit XSLT-Stylesheets können sie in beliebiger Form präsentiert werden.¹

EAC (*Encoded Archival Context*) hilft bei der Nutzung verschiedener Namen und Bezeichnungen von Personen, Organisationen oder Orten, indem verschiedene Formen mit Erklärungen ihrer Verwendung in den Beständen für die Recherche angeboten werden können. *EAC* beschreibt die typische Struktur der expliziten Kontextinformationen in archivischen Erschließungsangaben auf der Ebene der Beständeübersicht oder der Findbücher.²

METS (*Metadata Encoding and Transmission Standard*) erfasst und strukturiert alle erforderlichen Angaben für digitalisiertes Archivgut und unterstützt die Verwaltung digitaler Objekte, die aus einer oder mehreren Dateien oder aus zusammengehörigen Gruppen von digitalen Reproduktionen bestehen können und verknüpft sie mit den zugehörigen Online-Findbüchern.³

Ein Ergebnis der Konferenz möchte ich gleich vorweg nehmen, es entstand quasi im Vorfeld.

Alle Teilnehmer erhielten mit dem gedruckten Programm der Konferenz eine CD mit einer Zwischenversion der im <daofind+>-Projekt des *Bundesarchivs* mit Fördermitteln der *Andrew W. Mellon-Foundation*, New York, erstellten Werkzeuge für den Aufbau einer Internetpräsentation für digitalisiertes Archivgut in Online-Findbüchern, die auch auf der Website des Projektes unter <http://www.bundesarchiv.de/daofind/> bereitgestellt werden.

Der Prototyp des Werkzeuges nutzt die Entwicklungsumgebung Eclipse in einer an die Arbeitsgewohnheiten des archivischen Arbeitsplatzes angepassten Oberfläche. Die Version kann auch heruntergeladen werden. Derzeit können auf der genannten Internetseite 3 in *EAD* kodierte Findbücher mit 30.000 Images eingesehen und genutzt werden. Die Findbücher sind mit der Beständeübersicht und weiteren Online-Findbüchern übergreifend in der Suchmaschine *Midosasearch* recherchierbar.

Damit konnte in einer Pilotanwendung die kombinierte Nutzung der drei internationalen Standards *EAD*, *EAC* und *METS* für einen wesentlich verbesserten Zugang zu Archivgut aus deutschen Archiven getestet und der Prototyp für ein Editionswerkzeug für die Präsentation auf der Basis plattformunabhängiger offener Software entwickelt werden.

Zum Konferenzverlauf:

Nach dem Einführungsvortrag des Präsidenten des *Bundesarchivs*, Prof. Dr. Hartmut Weber, über die Rolle des Internet für die Dienstleistungsqualität der Archive ging es zunächst um *EAD*. Der Strukturstandard für Online-Findbücher wurde erläutert, ebenso seine Entwicklung und seine Verbreitung. Als ein Beispiel für seinen Einsatz in Europa wurden dann die vielfältigen Nutzungen in den französischen Archiven vorgestellt. Ein weiterer Beitrag befasste sich mit den Verbindungen und Entsprechungen mit anderen Strukturstandards wie *ISAD(G)* und *DC*. Die Multilingualität polnischer Online-Findbücher zu den vielsprachigen Beständen des Landes wurde beispielhaft für die Vielsprachigkeit der EU dargestellt. Anschließend wurde über die Nutzung von *EAD* bei der Retrodigitalisierung von Findmitteln am Beispiel des *Bundesarchivs* und im Archivportal Archives HUB der britischen Universitäten informiert.

¹ Weitere Informationen sind auf der offiziellen EAD Homepage zu finden: (<http://www.loc.gov/ead/>). Siehe auch die "EAD Help Pages" (veröffentlicht vom EAD Round Table of the Society of American Archivists) at: <http://www.archivists.org/saagroups/ead/>.

² Der Standard wurde im LEAF Projekt entwickelt (das Akronym LEAF steht für "Linking and Exploring Authority Files"); weitere Informationen unter: <http://xml.coverpages.org/ni2004-08-24-a.html>. Die aktuelle Beta Version von EAC ist von der Ad hoc EAC Working Group unter <http://jefferson.village.virginia.edu/eac/> veröffentlicht worden.

³ Siehe die offizielle METS Homepage: <http://www.loc.gov/standards/mets/>.

Thema der zweiten Sitzung war *EAC*, das bisher sehr viel weniger eingesetzt wird. Es ist, wie bereits erwähnt, ein Werkzeug zur Bereitstellung von Informationen über die Herkunftsstellen von Archivgut. Nach einer Information über die Entstehung in einer von Anfang an internationalen Arbeitsgruppe mit Beteiligung von Bibliotheken wurde die Beziehung von *EAC* zu dem Standard *des Internationalen Archivrates ISAAR(CPF)* erläutert und die unverzichtbare Funktion der Information über Herkunftsstellen für eine Nutzung von Archivgut dargestellt. Die Überlegungen des Projekts <daofind> des *Bundesarchivs* zu einer kombinierten Nutzung von *EAD* und *EAC* für Aufbau und Fortschreibung von Beständeübersichten mit zentralen und dezentralen *EAD*- und *EAC*-Dateien wurden vorgestellt und mit einer Darstellung des *BAM-Portals* wurde eine Brücke geschlagen zu der Abschlussdiskussion über Portale und gemeinsame Zugriffspunkte.⁴



Abbildung 1: Daofind

Archiven für einen übergreifenden Zugriff anbieten. In Spanien ist mit dem Censo Guia ein Portal zu 50.000 spanischsprachigen Archiven weltweit aufgebaut worden, das auf der Basis von *EAD* und *EAC* arbeitet. In Italien wurde eine mehrbändige gemeinsame Beständeübersicht zusammen mit Online-Findbüchern der Staatsarchive mit *EAD* bereitgestellt und hier liegen Erfahrungen in der Nutzung einer nativen XML-Datenbank und von *SVG* für eine geographische Einstiegsseite vor.

Die Sitzung wurde mit einer Podiumsdiskussion mit der Direktorin der französischen Archive Martine de Boisdeffre, der Vorsitzenden der *EAD*-Working Group bei der Society of American Archivists Kris Kiesling, dem Leiter der spanischen Archivverwaltung José Ramon Cruz Mundet und dem Direktor der Archivschule Marburg, Frank M. Bischoff unter der Leitung der Vizepräsidentin des *Bundesarchivs*, Angelika Menne-Haritz, abgeschlossen. In der Diskussion wurde die Notwendigkeit des zügigen Aufbaus des europäischen Archivportals herausgestellt.

Die Konferenz hat die Notwendigkeit deutlich gemacht, internationale Standards im Archiv zu nutzen, um die Zugänglichkeit zu den Unterlagen weiter zu verbessern. Das Internet ist wie geschaffen für die Dienstleistungen, die Archive anbieten. Das große Thema in der nächsten Zeit wird nicht mehr sein, ob man die Standards nutzen sollte, sondern wie man sie einsetzt, wie man die drei diskutierten Standards miteinander kombinieren kann und wie man dadurch die Archive und ihre Bestände deutlich besser sichtbar macht.

Die Rückmeldungen der Teilnehmer an die Veranstalter waren sehr positiv. Sie haben sich im Umweltforum Berlin sehr wohl gefühlt. Etwa 30 Mitarbeiterinnen und Mitarbeiter des *Bundesarchivs* waren aktiv an der Konferenz und an ihrer Organisation beteiligt und haben mitgeholfen, den Gästen einen angenehmen Aufenthalt in Berlin zu bereiten. Diese Konferenz ist bereits zum wichtigen Meilenstein in der internationalen Kooperation für digitale Zugänge zum Archivgut geworden.

⁴ Unter <http://www.daofind.de> finden Sie eine Vielzahl von Materialien und Übersetzungen zu *EAD*, *EAC* und *METS*

Wie werden die Ergebnisse der Konferenz umgesetzt?

Die Archive bauen ihr Angebot an Online-Findbüchern weiter aus. Im *Bundesarchiv* werden darüber hinaus gegenwärtig mehrere Archivbestände digitalisiert, um sie anschließend mit *METS* im Internet zu präsentieren. Mit der Einbindung in die Online-Findbücher und der Abbildung der Klassifikation und Findbucheinträge soll der Benutzer nicht nur einen Hinweis auf das bekommen, was er sucht, sondern er soll auch sehen, was alles vorhanden ist.

Die *DFG* stellt dem *Bundesarchiv* Fördermittel bereit, damit es in den kommenden zwei Jahren wichtige Vorhaben leistet, die in ein zukünftiges deutsches Archivportal einfließen können. Im Zentrum des neuen Projektes steht der Aufbau eines Verbundmittels für Archive, in dem Erschließungsinformationen über die Bestände bis auf die Ebene der bestellbaren Einheiten für eine übergreifende Recherche bereitstehen. Die Archive erhalten außerdem die Möglichkeit, diese Nachweise mit digitalen Reproduktionen zu verknüpfen, so dass direkt in Akten geblättert und gelesen werden kann. In diesem Verbundmittel werden die Archive Erschließungsdaten aus ihren unterschiedlichen EDV-Systemen über eine gemeinsame Schnittstelle zusammenführen. Dazu werden *EAD*, *EAC* und *METS* verwendet, die es ermöglichen, die Daten auch an andere zentrale Zugriffspunkte weiterzugeben. So können sie dann in das geplante Europäische Archivportal ebenso wie in die *Europäische Digitale Bibliothek* integriert werden. Über ein auf der Referenzanwendung aufbauendes Portal sollen außerdem zukünftig alle Erschließungsinformationen, deren Umwandlung in digitale Form ebenfalls von der *DFG* mit einem hohen finanziellen Aufwand gefördert wird, bereitgestellt werden. Neben dem Verbundmittel soll das Referenzportal Informationen zur Erreichbarkeit der Archive anbieten und Hilfestellung geben, wie man Archivgut für die eigenen Fragestellungen auswerten kann.

Die Mitglieder der Konferenz Archivreferenten und Leiter der Archivverwaltungen des Bundes und der Länder sowie Kolleginnen und Kollegen aus den beteiligten Archiven treffen sich am 10. Juli 2007 im *Bundesarchiv*, um das *EAD*-Verwendungsschema des Verbundmittels im Kontext der vorgesehenen Werkzeugentwicklung zu erläutern. Damit beginnt ein innovatives und herausragendes Projekt in der deutschen Archivlandschaft.

Mathias Jehn

Conditions for Long-term Preservation of Digital Objects in Europe

Zusammenfassung

Wie auf der EU-Veranstaltung „Herausforderung: Digitale Langzeitarchivierung. Strategien und Praxis europäischer Kooperation“ am 20. und 21. April 2007 in Frankfurt am Main deutlich wurde, gewinnt die Langzeitarchivierung digitaler Objekte in Europa immer mehr an Bedeutung. Übereinstimmend wurde hier festgestellt, dass der mit dem rasanten technischen Wandel einhergehenden Gefährdung der dauerhaften Nutzbarkeit von digitalen Objekten nicht mit nationalen Alleingängen begegnet werden kann. Gerade die derzeitigen hohen Investitionen in die Digitalisierung und in digitale Inhalte können ins Leere laufen, wenn kein tragfähiges Konzept zur digitalen Langzeitarchivierung vorliegt.

Um aber digitale Langzeitarchivierung und -verfügbarkeit in Europa auf eine technisch und operativ tragfähige Grundlage zu stellen, werden von den Mitgliedsstaaten in Zukunft große Anstrengungen nötig sein. Welcher Art diese sind und welche Chancen in der kooperativen Zusammenarbeit liegen, wurde in Frankfurt in den drei zentralen Themenfeldern: „Aufbau und Absicherung einer Infrastruktur“, „Implementierung von Standards und Zertifizierungsprozessen“ sowie „Konkrete Praxisempfehlungen“ beleuchtet – Ansätze, wie sie auch in dem Projekt *nestor – Kompetenznetzwerk Langzeitarchivierung in Deutschland* umgesetzt werden.

Abstract

The conservation of digital objects in the long term is gaining importance in Europe, as the recent EU conference, “The Challenge: Long-Term Preservation Strategies and Practices of European Partnerships” made clear. Held on 20 and 21 April 2007 in Frankfurt am Main, conference participants unanimously agreed that no one nation alone can tailor a strategy to ensure the long-term availability of digital resources in the present era of rapid technological change. The current major investments Europeans are making in digitisation and digital content could prove futile if we fail to develop a sustainable strategy to preserve digital resources over the long term. The technical and operational basis required to ensure the long-term preservation – and the long-term accessibility – of digital assets across Europe will require Member States to undertake great efforts.

The exact nature of these efforts, and the opportunities associated with enhanced collaboration, were discussed in Frankfurt in three thematic areas: “Development and Safeguarding of Infrastructure”, “Implementation of Standards and Certification Procedures”, and “Practical Recommendations”. The approaches discussed in Frankfurt will be implemented in *nestor*, the German network of expertise in the long-term storage of digital resources.

My paper basically examines the fundamental requirements and conditions for cooperative digital preservation in Europe. I will be addressing all memory organisations here. My deliberations are based both on my experience as Project Coordinator of the *nestor – network of expertise* and also on the results of a conference held in Frankfurt am Main in April during Germany's tenure of the EU Presidency 2007. Carrying the title “*The Challenge: Long-term digital preservation – Strategies and practice of European cooperation*”, the event provided international experts with a forum, hosted by the *Deutsche Nationalbibliothek*, where they

could consult on the current technical and organisational status of digital preservation.¹ At the conference it emerged that digital preservation represents one of the most pressing, sensitive and resource-intensive problem areas faced by most memory organisations, which is why new and cooperative organisational forms need to be found. In the search for solutions, the conference participants formulated not just a number of "declarations of intent". Instead, the focus was primarily on presenting concrete examples. The overriding question was whether we have already identified all possible areas for cooperation in the field of digital preservation in Europe.

Why intensive cooperation now is needed in digital preservation? First, it should be stressed that, even today, concepts such as cooperation and integration can still not be taken for granted within the public realm. There is considerable fear of losing knowledge, disclosing information and losing independence, of potential conflicts, inestimable costs and, not least, of having no guarantee of success. It is therefore important within such a cooperation-based venture to focus upon the mutual benefits to be enjoyed by all institutions involved. The success depends fundamentally on clear agreements being made regarding the organisational framework and on clear communication channels being found. Also, we must recognise that institutions cooperate, but that they also compete with each other. Competing to offer the best ideas is just as important for partnerships. The spotlight must be on the mutual benefits to be had by all institutions involved.

Even if cooperation is not easy to achieve, it is of prime importance for digital preservation. Digital preservation is closely linked with the overriding tasks of the memory organisations such as their social brief, their obligation to provide access to information or the question of national economic performance.

What are the core areas here? At the legal level, copyright is of prime importance. The rights involved in making necessary changes to electronic resources (e.g. migration) need to be clarified on a cooperative basis. Also in need of clarification are the restrictions imposed by *Digital Rights Management* systems (e.g. copy protection). Further points requiring clarification are the conditions for accessing archived objects and for using them. Also, digital objects are not bound to a specific location, rather they can be used via the Web. And then there is the entire subject of finance. Today we still do not know exactly what costs will be involved. It is already apparent; however, that large sums of money will be required for setting up digital preservation systems and also that the amount of financial resources available in the public domain will more and more be limited. Therefore, strategies are required which will permit us to achieve the best possible results despite the limited funds available.

Finally, there are organisational and technical problems, such as the rapid obsolescence of hard and software. There are also problems surrounding the authenticity and integrity of digital objects. Many institutions today are already working on trustworthiness criteria for digital material and for ensuring the high quality of archive storage systems – and how to translate such results into an infrastructure. All of these task areas cannot be addressed by a single institution. Setting up a competence centre also requires a long period of preparation and a high degree of coordination. Besides a financial foundation, it is important to establish a broad-based community bringing together all the digital preservation experts. Above all, it is important to gain the attention of governments, convincing them that the establishment of a network is a venture which is worthy of support.

There are already several good examples of national cooperation projects in the field of digital preservation, such as the *Digital Preservation Coalition (DPC)* in the UK, the *Network of Expertise in long-term STORAGE* and availability of digital Resources in Germany (*nestor*), and the *Pérennisation des Informations Numériques (PIN)* in France. From presentations given at the EU – Conference in Frankfurt it also emerged that the

¹ For further information on the EU Conference see: Mathias Jehn, *Das digitale Gedächtnis der Informationsgesellschaft*, in: BuB 7/8 (2007), pp. 534-537.

Netherlands, Denmark, Sweden, the Czech Republic and Italy are all planning similar projects. The national cooperation projects in these countries are the actual basis for digital preservation. They bundle funds, provide an organisational framework and act as repositories for collecting required areas of action.

As an example of a national network I would like to focus now on *nestor*.² *nestor's* mission statement is: "The aim of the project is to establish a competence network for the preservation and long-term accessibility of digital sources for Germany within a permanent organisational form, and to achieve national and international agreement on carrying out particular tasks." Crucial here, is the fact that the establishment of a national network is directly linked to international cooperation with different partners and activities.

It is of course one thing for an institution to acknowledge the task as a duty, yet something entirely different for it to undertake to carry out the task itself. In response to this necessity, in 2003 the German *Ministry of Research and Education* charged the *nestor competence network for digital preservation and long-term accessibility of digital objects* with organising relevant digital preservation management and coordination tasks, and also with research and development. An important result of the first project phase from 2003 to 2006 was the agreement of common guidelines: in its "Memorandum on the long-term accessibility of digital information in Germany" *nestor* summarised the work needed to be done by political decision-makers, authors, publishers, hard and software producers and cultural and scientific memory organisations in order to define the basic framework for a national digital preservation policy.³ In the first project phase it emerged that the required range of skills for the complex task of "long-term preservation of digital resources" covers a wide spectrum of people working in many institutions, organisations and companies. The main task areas in the current project phase (2006 to 2009) are:

- Expansion of the *nestor* community, predominantly amongst archives, libraries and museums. *nestor* is to set up and promote partnerships, supporting them in resolving the common problem.
- Bundling of identifiable skills in the field of digital preservation: Right at the start of the project a web-based information and communication platform was set up offering a range of interaction, communication and information possibilities on the subject of digital preservation. It should be possible for interested users to find materials in a database without the need for detailed specialist knowledge.
- Standardisation activities and certification processes: *nestor* is to implement standardisation measures aimed at supporting the long-term accessibility of documents, objects etc. for research and scientific purposes.
- Expansion and protection of a cooperative infrastructure: Establishment of contact to initiatives and specialist groups. Specialists from different work areas, relevant institutions and interested non-professionals consult in the competence network via workshops and working groups. In addition, the area of "digital preservation" and scientific and medical technology approaches (eScience and grid computing) should also be integrated.
- Initial, further and advanced training: In the last few years *nestor* has registered extensive demand for initial and further training. Accordingly, work began on drawing

² "nestor" is an acronym derived from the English translation of the project title: "Network of Expertise in long-term STOrage and availability of digital Resources in Germany". The partners in the joint project funded by the German Ministry of Education and Research are: Deutsche Nationalbibliothek, Lower Saxony State and University Library, the Computer and Media Service and University Library of the Humboldt University in Berlin (HUB), Bavarian State Library in Munich, the Hagen Open University, the Institute for Museum Research in Berlin and the Federal German Archive. See: <http://www.langzeitarchivierung.de>.

³ See: <http://www.langzeitarchivierung.de/downloads/memo2006.pdf>.

up proposals for a comprehensive range of training options in conjunction with existing institutions.⁴

Therefore, *nestor* covers a large number and a wide range of aspects and associated tasks which bring together the relevant experts and active project participants, promoting the exchange of information, the development of standards and the exploitation of synergy effects. In its second project phase, *nestor's* overriding goal is to transfer its wide ambit of tasks into a permanent organisational form after the end of the project. With regard to possible plans for the long-term preservation of digital sources in Germany, the *nestor competence network* is therefore developing a permanent cooperation structure for archives and museums which links expert knowledge, supplies information and provides ideas for new projects. The aim is to support memory institutions in their efforts to ensure the preservation and accessibility of all digital resources themselves, to organise the preservation and accessibility of the most important objects in each specialist field and to guarantee the preservation and accessibility of digital archive records.

There is, however, a noticeable sense of unease in many colleagues working in the field when discussion turns to the new media and their implications for archiving. One reason lies doubtlessly in the high level of complexity encountered in many guides and studies on the topics. Studying the vast range of literature available is not sufficient as the basis for an institution to carry out the archiving of electronic documents itself. This was also the conclusion of a study carried out by Margret Hedstrom and Sheon Montgomery for the former *Research Libraries Group*.⁵ Here, it emerged that there is a fundamental link between staff training levels and the preservation of electronic documents. If we distinguish between three knowledge levels, it is apparent that all the top-level experts are to be found in those institutions which already have practical experience in the archiving of electronic documents. 80 percent of the individuals with medium-level knowledge are also to be found in such institutions. Without first becoming involved in the practical implementation of electronic archiving it is therefore almost impossible to obtain the necessary knowledge in this field. It is only possible to gain such experience by actually engaging in the work of electronic archiving. Any national digital preservation competence network must reflect this conclusion. For this reason, at the start of 2007, *nestor*, in collaboration with universities from Germany, Switzerland and Austria, began work on developing a workable proposal for organising and carrying out training. The programmes worked out are aimed at trainees and students, but also at technical staff and researchers. Furthermore, *nestor* is approaching archives and museums in particular to investigate topics such as safeguarding the evidential value of electronic documents, classification standards for digital archive contents or exchange formats. In the training courses, emphasis is placed upon the fact that the long-term preservation of digital resources does not actually entail additional task areas for libraries, archives and museums, rather it involves three of these institutions' traditional core tasks: collecting, classifying and preserving. Ultimately, the overriding priority for any archivist is preventing the loss of memory in the information society.

All of these task areas cannot be covered by a single European institution working in isolation, and need to be tackled on a cooperative basis. Taking the necessary steps, however, requires immediate, concentrated action. Despite in some cases divergent tasks and differing conditions (organisation, legislation), it is important to work on a joint solution. Today there are already several good examples of successful cooperation projects in the field of digital preservation. With regard to technology research, the EU-sponsored research

⁴ The working groups currently active in *nestor* II are: WG Trustworthy Repositories – Certification, WG *nestor* – Media, WG Standards for metadata, transfer of objects to digital repositories and object access, WG Raw Scientific Data, WG Cooperative long-term preservation, WG Interlinking of grid/eScience and long-term preservation.

⁵ Margaret Hedstrom and Sheon Montgomery, *Digital Preservation Needs Member Institutions*, 1998, in: <http://www.rlg.ac.uk/preserv/digpres.html>. See also: Christian Keitel, *Warum, wann und wie – drei Fragen zur elektronischen Archivierung*, in: *Handlungsstrategien für Kommunalarchive im digitalen Zeitalter (Texte und Untersuchungen zur Archivpflege 19)*, Münster 2006, 87-93.

projects *PLANETS* and *CASPAR* represent key developments⁶ e.g. in the implementation of the *Open Archival Information System (OAIS)*⁷. The model succeeds in establishing a uniform language and a universally applicable view of the core functions of a digital archive which goes beyond the confines of individual user communities, archives, data centres, and libraries by delimiting and clearly labelling the function modules, interfaces and types of information objects. It is this universality which makes the model's level of abstraction relatively high. The *Open Archival Information System* defines an information network which regards the archivist and the user as the main components of the digital archive.

In the area of archive certification, too, results have already been published such as the *TRAC Checklist* or the *nestor* "Criteria Catalogue for Trusted Repositories".⁸ The criteria identified describe the organisational and technical requirements of a digital repository and can be applied to a whole range of digital repositories and archives: from university repositories through to major data archives; from national libraries through to external digital archive service providers. The trustworthiness of digital repositories can be tested and assessed on the basis of the criteria catalogues.

There is a need for greater cooperation within Europe in all the areas mentioned above. Most of the current initiatives are project-based and are characterised by their restricted focus and duration. The great challenge for all in the coming years will therefore be to establish and maintain an expandable European infrastructure, which is supported by the member states, for accessing the information on a long-term and trusted basis. The 7th EU *Framework Programme* is pursuing precisely this objective with its research programme.⁹ In her concluding remarks at the Frankfurt conference, Pat Manson, the *European Commission* representative, gave an assurance that the *European Commission* is prepared to provide special support for infrastructure-related applications.

A topical example of an infrastructure created at the European level is the *Alliance for Permanent Access to the Records of Science* founded in 2007 which is purposely interdisciplinary in nature and brings together a large number of communities.¹⁰ Members include: the *European Organisation for Nuclear Research (CERN)*, the *European Space Agency (ESA)*, the *Computational Science & Engineering organisation (CCLRC)* in the UK (now the 3rd Scientific and Technology Facilities Council), the *Joint Information Systems Committee (JISC)*, the UK and Netherlands national libraries, the *STM publishers' association*, the *National Archive of Sweden* and the national competence networks in the UK, Germany and France. This European alliance wishes to create a distributed interdisciplinary organisational structure for safeguarding permanent access to specialist information. Digital preservation is not an end in itself. The activities associated with digital preservation must always be regarded within the context of higher goals. Digital preservation should not be regarded in isolation to an institution's other activities. Rather it is needed to help create a national and European identity. Ultimately, economic success is also dependent upon the ability to preserve data. Not least, it represents a prerequisite if Europe wishes to retain its status as a location for research. In implementing cooperation-based digital preservation, it is important to set pragmatic and realistic goals. At the same time, however, it is also crucial to retain a sense of ambition and aspiration.

⁶ See: <http://www.planets-project.eu/>; <http://www.casparpreserves.eu/>.

⁷ The "Open Archival Information System - OAIS" model passed as ISO 14721 is printed in: <http://public.ccsds.org/publications/archive/650x0b1.pdf>.

⁸ The criteria catalogues can be found at: <http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91>.

⁹ See <http://www.forschungsrahmenprogramm.de/> or http://cordis.europa.eu/fp7/home_en.html.

¹⁰ See: <http://www.alliancepermanentaccess.eu/>.

Mathias Jehn

Voraussetzungen der Langzeitarchivierung digitaler Objekte in Europa

Im Zentrum meines Beitrages steht die Frage nach den grundsätzlichen Erfordernissen und Voraussetzungen kooperativer Langzeitarchivierung in Europa, wobei ich alle Gedächtnisorganisationen einbeziehen möchte. Meine Ausführungen basieren einerseits auf den Erfahrungen als Projektkoordinator des Kompetenznetzwerks *nestor*, andererseits auf den Ergebnissen einer im April 2007 unter dem Dach der EU-Ratspräsidentschaft stattgefundenen Konferenz in Frankfurt am Main. Unter dem Motto: „Herausforderung: Digitale Langzeitarchivierung. Strategien und Praxis europäischer Kooperation“ diskutierten in der *Deutschen Nationalbibliothek* internationale Fachleute über den technischen und organisatorischen Stand der Langzeitarchivierung digitaler Objekte.¹ Was dabei deutlich wurde, ist, dass für meisten Gedächtnisorganisationen die Langzeitarchivierung zu den dringlichsten, sensibelsten und ressourcenintensivsten Problembereichen zählt, weshalb neue und kooperative Organisationsformen gefunden werden müssen. Bei den Lösungswegen formulierten die Konferenzteilnehmer nicht nur „Absichtserklärungen“. Der Fokus lag vielmehr auf der Präsentation von konkreten Beispielen. Kernfrage war, ob wir bereits sämtliche Kooperationsfelder der Langzeitarchivierung in Europa identifiziert haben?

Warum bedarf es einer intensiven Kooperation in der Langzeitarchivierung? Es muss zunächst betont werden, dass im öffentlichen Bereich Begriffe wie Kooperation und Vernetzung immer noch keine Selbstverständlichkeit sind. Dagegen steht beispielsweise die Angst vor möglicher Know-how-Einbuße, die Preisgabe von Informationen, die Sorge um Selbstständigkeit, der kaum kalkulierbare Aufwand und nicht zuletzt ein nicht zu garantierender Erfolg. Wichtig ist daher, dass bei einer Kooperation der wechselseitige Nutzen für die beteiligten Institutionen berücksichtigt wird. Der Erfolg hängt entscheidend davon ab, dass klare Abstimmungen über den organisatorischen Rahmen vereinbart und klare Wege der Kommunikation gefunden werden. Weiter muss man sich bewusst machen, dass Institutionen nicht nur kooperieren, sondern auch konkurrieren. Der Wettbewerb um die besten Ideen ist genauso wichtig für erfolgreiche Kooperationen. Es ist der gemeinsame Nutzen aller beteiligten Institutionen, der im Vordergrund stehen muss.

Auch wenn sich eine Kooperation nicht einfach verwirklichen lässt, ist sie für die Langzeitarchivierung von zentraler Bedeutung. Langzeitarchivierung ist eng mit den übergeordneten Ebenen verbunden, wie z.B. dem gesellschaftlichen Auftrag von Gedächtnisorganisationen, der Bereitstellung eines Zugangs zu Informationen oder die Frage nach dem Wirtschaftsstandort.

Was sind nun die zentralen Kernbereiche bei der digitalen Langzeitarchivierung? Im Bereich des Rechts steht das Urheberrecht an erster Stelle. Kooperativ zu klären sind die Rechte zur Durchführung notwendiger Eingriffe in elektronische Ressourcen (z.B. bei Migration). Zu klären sind vorgegebene Einschränkungen durch *Digital Rights Management Systeme* (z.B. Kopierschutz). Zu klären sind die Bedingungen für den Zugriff auf die archivierten Objekte und für deren Nutzung. Auch sind digitale Objekte nicht an einem Ort gebunden, sondern können über das Web von überall her benutzt werden. Des Weiteren ist der ganze Finanzierungsbereich betroffen. Wir wissen heute noch nicht genau, welche Kosten anfallen werden. Abzusehen ist aber, dass einerseits große Geldsummen für den Aufbau von Langzeitarchivierungssystemen benötigt werden, andererseits aber der finanzielle Spielraum für den öffentlich-rechtlichen Bereich begrenzt sein wird. Es sind daher Strategien nötig, wie wir mit den begrenzten Mitteln die besten Ergebnisse erzielen können.

¹ Zur EU-Konferenz siehe zuletzt ausführlich: Mathias Jehn, Das digitale Gedächtnis der Informationsgesellschaft, in: BuB 7/8 (2007), S. 534-537.

Schließlich sind noch organisatorische und technische Probleme zu nennen: z.B. das rasche Veralten von Soft- und Hardware. Daneben gibt es das Problem von Authentizität und Integrität digitaler Objekte. Viele Einrichtungen arbeiten heute schon an den Kriterien für Vertrauenswürdigkeit des digitalen Materials und für die Qualität von Archivspeichern sowie daran, wie man es das Material in eine Infrastruktur überträgt. All diese Aufgabenfelder sind nicht von einer Institution allein zu leisten. Des Weiteren benötigt man zum Aufbau eines Kompetenzzentrums lange Entstehungszeiten und einen hohen Grad an Abstimmung. Neben der Finanzierungsgrundlage braucht man eine breite Community aus Experten der Langzeitarchivierung. Und vor allem braucht man das Gehör der Politik, die den Aufbau eines Netzwerks als förderungswürdig betrachtet.

Es gibt heute schon einige gute Beispiele für nationale Kooperationen in der digitalen Langzeitarchivierung, wie z.B. die *Digital Preservation Coalition (DPC)* in Großbritannien, *nestor – Kompetenznetzwerk Langzeitarchivierung in Deutschland* oder die *Pérennisation des Informations Numériques (PIN)* in Frankreich. Auf der EU-Konferenz in Frankfurt wurde des Weiteren deutlich, dass die Niederlande, Dänemark, Schweden, die Tschechische Republik und Italien ähnliche Projekte planen. Die dortigen nationalen Kooperationsprojekte sind die eigentliche Basis der digitalen Langzeitarchivierung. Hier werden Finanzmittel gebündelt, der organisatorischen Rahmen gesetzt und die Themengebiete gesammelt.

Als Beispiel für ein nationales Netzwerk möchte ich auf *nestor* zu sprechen kommen.² Das Missionstatement von *nestor* lautet: „Ziel des Projektes ist der Aufbau eines Kompetenznetzwerks zur Langzeitarchivierung und Langzeitverfügbarkeit digitaler Quellen für Deutschland in einer dauerhaften Organisationsform sowie nationale und internationale Abstimmungen über die Übernahme von konkreten Aufgaben.“ Wichtig ist hierbei, dass die Etablierung eines nationalen Netzwerks direkten Bezug auf die internationale Zusammenarbeit mit unterschiedlichen Partnern und Aktivitäten nimmt.

Der Ausgangspunkt des Projekts war die Erkenntnis, dass es das eine ist, die Aufgabe der digitalen Langzeitarchivierung als Pflichtaufgabe anzuerkennen und etwas völlig anderes, sie tatsächlich auch selbst angehen zu wollen. Im Jahre 2003 hat das Bundesministerium für Forschung und Bildung, das Projekt *nestor – Kompetenznetzwerk zur Langzeitarchivierung und Langzeitverfügbarkeit digitaler Objekte* mit der notwendigen Steuerung, Koordination, aber auch mit der Forschung und Entwicklung im Bereich der Langzeitarchivierung beauftragt. Ein wichtiges Ergebnis der ersten Projektphase von 2003 bis 2006 war die Verabschiedung gemeinsamer Richtlinien: *nestor* hat in einem „Memorandum zur Langzeitverfügbarkeit digitaler Informationen in Deutschland“ die notwendigen Anstrengungen von politischen Entscheidungsträgern, Urhebern, Verlegern, Hard- und Softwareherstellern sowie kulturellen und wissenschaftlichen Gedächtnisorganisationen zusammengestellt, um die Rahmenbedingungen einer nationalen Langzeitarchivierungs-Policy abzustecken.³ Auf der Basis der in der ersten Projektphase erreichten Leistungen verteilen sich die notwendigen Fachkompetenzen für den Aufgabenkomplex „Langzeitarchivierung digitaler Ressourcen“ über ein breites Spektrum von Personen, die in vielen Institutionen, Organisationen und Wirtschaftsunternehmen tätig sind. Die wesentlichen Aufgabenpakete für die laufende Projektphase von 2006 bis 2009 sind:

- Ausbau der *nestor*-Community vorrangig aus dem Bereich der Archive, Bibliotheken und Museen. *nestor* bildet und fördert Kooperationen und unterstützt diese bei der Lösung des gemeinsamen Problems.

² *nestor* ist das Akronym der englischen Übersetzung des Projekttitels: „Network of Expertise in long-term storage and availability of digital Resources in Germany“. Die Partner in dem vom BMBF geförderten Verbundprojekt sind: Die Deutsche Nationalbibliothek, die Niedersächsische Staats- und Universitätsbibliothek Göttingen, der Computer- und Medienservice und Universitätsbibliothek der Humboldt-Universität zu Berlin, die Bayerische Staatsbibliothek in München, die Fernuniversität Hagen, das Institut für Museumsforschung in Berlin und das Bundesarchiv. Siehe: <http://www.langzeitarchivierung.de>.

³ Siehe: <http://www.langzeitarchivierung.de/downloads/memo2006.pdf>.

- Bündelung von identifizierbaren Kompetenzen der Langzeitarchivierung: Seit Projektbeginn liegt eine webbasierte Informations- und Kommunikationsplattform vor, die vielfältige Möglichkeiten zur Interaktion, Kommunikation und Information zum Thema Langzeitarchivierung bietet. Das Auffinden und Einbringen der Materialien in einer Datenbank ist dem interessierten Nutzer ohne besondere Detailkenntnisse möglich.
- Standardisierungsaktivitäten und Zertifizierungsprozesse: *nestor* verfolgt Maßnahmen der Normierung, die die Langzeitverfügbarkeit von Dokumenten, Objekten, etc. für Forschung und Wissenschaft befördern.
- Ausbau und Absicherung einer kooperativen Infrastruktur: Kontaktaufbau zu Initiativen und Fachgruppen. Fachleute aus unterschiedlichen Arbeitsbereichen, betroffene Institutionen sowie interessierte Laien tauschen sich im Kompetenznetzwerk mittels Workshops und Arbeitsgruppen aus. Des Weiteren werden die Themen „digitale Langzeitarchivierung“ und naturwissenschaftliche und medizinische Technologieansätze (e-science und Grid-Computing) zusammengeführt.
- Aus-, Fort- und Weiterbildung: In *nestor* wurde in den vergangenen Jahren ein umfangreicher Aus- und Weiterbildungsbedarf erkannt, weshalb damit begonnen wurde, ein Konzept für ein umfassendes Angebot in Zusammenarbeit mit bestehenden Einrichtungen zu entwerfen.⁴

nestor bietet damit eine Vielzahl spartenübergreifender Aspekte und daraus resultierender Teilaufgaben an, die u.a. die kompetenten Experten und aktive Projektnehmer zusammenbringt und den Austausch von Informationen, die Entwicklung von Standards sowie die Nutzung von Synergieeffekten fördert. Für die zweite Projektphase von *nestor* ist es oberstes Ziel, die umfangreichen Aufgaben nach Abschluss des Projekts in einer dauerhaften Organisationsform langfristig zu verstetigen. Hinsichtlich der Überlegungen zur Langzeitarchivierung digitaler Quellen in Deutschland entwickelt das Kompetenznetzwerk *nestor* daher eine dauerhafte Kooperationsstruktur für Bibliotheken, Archive und Museen, die Expertenwissen vernetzt, Informationen vermittelt und neue Projektanstöße gibt. Dadurch sollen die Gedächtnisinstitutionen unterstützt werden, die Bewahrung und Verfügbarkeit aller digitalen Ressourcen selbst zu gewährleisten, die Bewahrung und Verfügbarkeit der wichtigsten Objekte jedes Fachgebiets zu organisieren sowie schließlich die Bewahrung und Verfügbarkeit digitaler Archivalien garantieren zu können.

Doch trotz solcher vorliegenden Erfolge steht man in vielen Bereichen der digitalen Langzeitarchivierung immer noch am Anfang. Bei vielen Kolleginnen und Kollegen verspürt man beispielsweise ein stetiges Unbehagen, wenn die Rede auf die neuen Medien und die sich daraus ergebenden Folgen für die Archivierung kommt. Ein Grund dafür liegt sicherlich in dem hohen Schwierigkeitsgrad, den viele Handreichungen und Studien zu dem Thema haben. Allein das Selbststudium der reichhaltigen Literatur wird nicht ausreichen, um die Archivierung von elektronischen Unterlagen selbst durchzuführen. Das untermauert auch das Ergebnis einer für die ehemalige *Research Libraries Group* angefertigten Studie von Margret Hedstrom und Sheon Montgomery.⁵ Danach besteht ein grundsätzlicher Zusammenhang zwischen der Qualifikation des Personals und der Archivierung elektronischer Unterlagen. Unterscheidet man zwischen drei Kenntnisebenen, so finden sich sämtliche Experten in solchen Institutionen, die bereits praktische Erfahrung in der Archivierung elektronischer Unterlagen sammeln konnten. 80% der Personen mit mittleren Kenntnissen sind gleichfalls in solchen Institutionen angesiedelt. Ohne die praktische Durchführung elektronischer Archivierung können demzufolge kaum die notwendigen Kenntnisse in diesem Bereich aufgebaut

⁴ Die eingerichteten Arbeitsgruppen in *nestor* II sind: AG Vertrauenswürdige Archive – Zertifizierung, AG Media, AG Standards in den Bereichen Metadaten, Transfer von Objekten in digitale Langzeitarchive und Objektzugriff, AG Wissenschaftliche Rohdaten, AG Kooperative Langzeitarchivierung, AG Vernetzung Grid/e-science und Langzeitarchivierung.

⁵ Margaret Hedstrom und Sheon Montgomery, Digital Preservation Needs Member Institutions, 1998, in: <http://www.rlg.ac.uk/preserv/digpres.html>. Siehe auch: Christian Keitel, Warum, wann und wie – drei Fragen zur elektronischen Archivierung, in: Handlungsstrategien für Kommunalarchive im digitalen Zeitalter (Texte und Untersuchungen zur Archivpflege 19), Münster 2006, 87-93.

werden. Diese praktischen Erfahrungen können aber nur dann gemacht werden, wenn man eben mit der elektronischen Archivierung beginnt. Dieser Erkenntnis muss ein Kompetenznetzwerk zur Langzeitarchivierung Rechnung tragen. Deshalb hat *nestor* in Zusammenarbeit mit deutschen, schweizerischen und österreichischen Hochschulen Anfang 2007 begonnen, ein tragfähiges Ausbildungskonzept zur Qualifizierung zu entwickeln und umzusetzen. Die Angebote richten sich sowohl an Auszubildende, Studenten als auch an Praktiker und Forscher. Des Weiteren spricht *nestor* speziell Archive und Museen an, um Themen wie die Sicherung der Beweiskraft elektronischer Unterlagen, Erschließungsstandards für digitales Archivgut oder Austauschformate zu erörtern. In den Fortbildungsveranstaltungen gilt es insbesondere zu betonen, dass die Langzeiterhaltung digitaler Ressourcen im eigentlichen Sinn keine zusätzliche Aufgabe für Bibliotheken, Archive und Museen darstellen darf, sondern den ureigenen Aufgaben: Sammeln, Erschließen und Bewahren entspricht. Schließlich ist die Verhinderung des Gedächtnisverlusts in der Informationsgesellschaft die oberste Priorität jedes Archivars.

Die genannten Aufgabenfelder sind nicht von einer europäischen Institution allein zu leisten und müssen kooperativ angegangen werden. Die Realisierung der notwendigen Schritte erfordert ein sofortiges, konzentriertes Handeln. Trotz teilweise divergierender Aufgabenstellungen und unterschiedlicher Rahmenbedingungen (Organisation, Rechtsgrundlagen) muss an einer gemeinsamen Lösung gearbeitet werden. Heute liegen schon einige gute Beispiele für erfolgreiche Kooperationsprojekte im Bereich der Langzeitarchivierung vor.

Im Bereich der technologischen Forschung sind die von der EU geförderten Forschungsprojekte *PLANETS* und *CASPAR* wichtige Einrichtungen,⁶ etwa bei der Implementierung des *Open Archival Information System (OAIS-Modell)*.⁷ Durch die Abgrenzung und eindeutige Benennung von Funktionsmodulen, Schnittstellen und Typen von Informationsobjekten ist es gelungen, eine einheitliche Sprache und eine über die Grenzen der Anwendergemeinschaften, Archive, Datenzentren und Bibliotheken hinweg geltende allgemeine Sicht auf die Kernfunktionen eines digitalen Archivs zu schaffen. Gerade durch diese Allgemeingültigkeit ist der Abstraktionsgrad des Modells relativ hoch. Das *Open Archival Information System* beschreibt ein Informationsnetzwerk, das den Archivar und den Nutzer als Hauptkomponenten des digitalen Archivs versteht.

Auch für den Bereich der Zertifizierung von Archiven liegen bereits Ergebnisse vor, wie beispielsweise die TRAC Checkliste oder der *nestor* „Kriterienkatalog Vertrauenswürdige Archive“.⁸ Die hier veröffentlichten Kriterien beschreiben die organisatorischen und technischen Voraussetzungen eines digitalen Langzeitarchivs und sind auf eine Reihe digitaler Repositorien und Archive anwendbar, von universitären Repositorien bis hin zu großen Datenarchiven; von Nationalbibliotheken bis hin zu digitalen Archivierungsdiensten Dritter. Anhand der Kriterienkataloge kann die Vertrauenswürdigkeit digitaler Langzeitarchive nun geprüft und bewertet werden.

Für alle erwähnten Bereiche besteht überall in Europa erhöhter Kooperationsbedarf. Die meisten der aktuellen Initiativen sind Projekte und zeichnen sich durch einen begrenzten Fokus und Zeitrahmen aus. Es wird eine große Herausforderung für der kommenden Jahre sein, eine entwicklungsfähige und von den Mitgliedsstaaten getragene europäische Infrastruktur aufzubauen, welche den Zugang zu Informationen dauerhaft und vertrauenswürdig gewährleistet. Das 7. Rahmenprogramm der *Europäischen Union* verfolgt mit seinem Forschungsprogramm dieses Ziel. Die Referentin Pat Manson versicherte als Vertreterin der Europäischen Kommission im Schlusswort der Frankfurter Konferenz, dass die europäische Kommission bereit ist, entsprechende Anträge im infrastrukturellen Bereich besonders zu unterstützen.

⁶ Siehe: <http://www.planets-project.eu/>; <http://www.casparpreserves.eu/>.

⁷ Das als ISO 14721 verabschiedete Referenzmodell „Open Archival Information System – OAIS“ ist abgedruckt in: <http://public.ccsds.org/publications/archive/650x0b1.pdf>.

⁸ Die Kriterienkataloge finden Sie unter: <http://www.crl.edu/content.asp?i1=13&i2=58&i3=162&i4=91>.

Ein aktuelles Beispiel für eine auf europäischer Ebene geschaffene Infrastruktur ist die 2007 gegründete *Alliance for Permanent Access to the Records of Science*, die bewusst interdisziplinär angelegt ist und eine Vielzahl von Communities vereinigt.⁹ Unter den Mitgliedern sind: Die europäische Organisation für Kernforschung (*CERN*), *Europäische Weltraumorganisation (ESA)*, die *Computational Science & Engineering (CCLRC)* in Großbritannien (jetzt der 3. Wissenschaft und Technologie-Service-Rat), *JISC*, die Nationalbibliotheken Großbritanniens und der Niederlande, die *STM Verlags-Vereinigung*, die Nationalarchive von Schweden und die nationalen Kompetenznetzwerke in Großbritannien, Deutschland und Frankreich. Dieses europäische Bündnis möchte eine verteilte interdisziplinäre Organisationsstruktur erreichen, die den dauerhaften Zugang von Fachinformation sichert.

Abschließend möchte ich betonen, dass die digitale Langzeitarchivierung kein Selbstzweck ist. Die Aktivitäten um die Langzeitarchivierung müssen immer unter einer höheren Zielsetzung betrachtet werden. Sie darf nicht losgelöst von den sonstigen Aktivitäten einer Institution stehen. Sie wird vielmehr zur Bildung einer nationalen und europäischen Identität benötigt. Von der Fähigkeit des Erhalts der Daten hängt letztlich auch der wirtschaftliche Erfolg ab. Nicht zuletzt ist sie eine Voraussetzung für den europäischen Forschungsstandort. Bei der Umsetzung kooperativer Langzeitarchivierung ist es zwar wichtig, pragmatische und realistische Ziele anzustreben, dennoch sollte man durchaus ehrgeizig und ambitioniert bleiben.

⁹ Siehe: <http://www.alliancepermanentaccess.eu/>.

Michael Buckland

Geography, Time, and the Representation of Cultural Change: Experience from a Large Collaboration: The Electronic Cultural Atlas Initiative

Zusammenfassung

Geographie, Zeit und die Repräsentation kulturellen Wandels: Erfahrungen einer umfangreichen Zusammenarbeit: Die *Electronic Cultural Atlas Initiative*.

Das Web gestattet Zugang zu zahlreichen Ressourcen verschiedener Sammlungen – sie zu verstehen, hängt vom Kontextwissen ab. Kultur ist ein aktiver, lebendiger Prozess und wird in Gemeinschaften über lange Zeit entwickelt.

Die *Electronic Cultural Atlas Initiative* ist ein großer internationaler informeller Zusammenschluss von Wissenschaftlern, Bibliothekaren, Technikern und anderen Experten, die bessere Methoden für den Umgang mit Orten und Zeiten entwickeln wollen, um unser Verständnis von Kultur zu verbessern. Geographische Suche kann durch das Verbinden von Orten als einem kulturellen Konstrukt mit Räumen als einem physikalischen Konzept optimiert werden; ähnlich können Ereignisse und Epochen mit Zeitabschnitten verknüpft werden. Fragen nach dem WAS, WO, WANN und WER erfordern unterschiedliche Methoden der Repräsentation und Präsentation.

Metadaten sind deskriptiv, konstituieren jedoch auch eine Form der Infrastruktur, wenn Probleme der Interoperabilität gelöst sind. Beispiele für thematische, geographische, zeitliche und biographische Suchhilfen werden vorgestellt, ebenso wie sich wandelnde politische Grenzen.

Abstract

The Web allows access to varied resources from many different collections, but understanding depends on knowing the context. Culture is active and evolves in communities in places over time.

The *Electronic Cultural Atlas Initiative* (ecai.org) is a large international informal collaboration of scholars, librarians, technology specialists and others concerned to improve our ability to understand culture through better methods for handling place and time.

Geographical search can be improved by relating place, a cultural construct, to space, a physical construct; similarly, events and eras can be related to chronological time.

WHAT, WHERE, WHEN and *WHO* require different methods for representation and display. Metadata is descriptive, but also constitutes a form of infrastructure if problems of interoperability are solved. Examples will be shown of topical, geographical, temporal and biographic search support, including changing political borders.

Introduction

The *Electronic Cultural Atlas Initiative* (ECAI, (ecai.org)) is a very informal collaboration among many scholars in several different countries, with librarians, information technology specialists, and others. Administratively, it is based at the *University of California* at Berkeley. The mission of ECAI is to transform education and research through better

understanding of place, time, and culture. We are not creating an atlas. We are developing methods, recommending "best practices," and encouraging collaborations so that everybody can make more effective use of cultural resources by making them more accessible and more interoperable. As better and more interoperable techniques are adopted internationally, a collective atlas will, in effect, begin to emerge.

Recent work has been based on five ideas:

- Understanding requires knowing the context. Nothing has meaning in isolation.
- Finding and using Internet resources should be like using reference books in a library and as easy and as reliable.
- A design goal: Find the context of any museum object, document, or performance: What is related to it in what it is, where it came from, when it originated, and who is associated with it?
- WHAT, WHERE, WHEN and WHO (the four Ws) provide a useful structure.
- Make better use of existing descriptive metadata.

Resources are not available unless somebody collects them. Collected resources are not accessible unless they are cataloged and two problems arise: First, catalogs (and indexes, ontologies, thesauri, etc.) differ, so the vocabulary used in any of them is likely to be unfamiliar. Searches in unfamiliar vocabularies tend to be less efficient and less reliable. Second, the facets *WHAT*, *WHERE*, *WHEN* and *WHO* are different from each other and need different methods, different vocabularies, and different forms of display.

WHAT, WHERE, WHEN and WHO

WHAT

If you are interested in KUNG FU MOVIES, in the *Library of Congress Subject Headings* you must ask for MARTIAL ARTS FILMS. Previously you had to ask for HAND-TO-HAND FIGHTING, ORIENTAL, IN MOTION PICTURES. Description is a language activity, and there are many different languages and dialects. Here are the descriptive terms for 'automobile' in four different dialects used in U.S. federal government databases:

- PASS MOT VEH, SPARK IGN ENG in the U.S. Import/Export statistics is an abbreviated version of "Passenger motor vehicle, spark ignition engine."
- TL 205 (Library of Congress Classification).
- 180/190 (US Patent classification).
- 3711 (Standard Industrial Classification).

Even when a natural language is used it is often difficult to imagine which word will have been used and the meaning can still be unclear even when the proper term has been found. Consider, for example, from the *International Harmonized Commodity Classification System*: "HS 847120 Digital auto data proc mach contng in the same housing a CPU and input & output device." It does not say so, but it means 'computer'.

People understand the need for cross-references *within* a vocabulary, but a network environment raises the far larger problem of cross-references *between* different vocabularies. There is enthusiasm for a "Semantic Web." These examples illustrate the reality of the semantics of the Web.

WHERE

Special, separate treatment is needed for *WHERE*. *Place* is a cultural construct and space is a physical construct. A place name gazetteer, a kind of bilingual dictionary between place and space, is needed. A place name gazetteer gives the name of a place, indicates what kind of place it is (geographical description code or "feature type"), latitude and longitude, and when that name was used. Latitude and longitude are very important. They relate place to space and, thereby, enable us to show the locations of places on maps and to see how

places are related to each other spatially. *WHAT* and *WHERE* can be related to each other when a place name gazetteer and a catalog talk to each other, and, with latitude and longitude from the gazetteer, the catalog can have a map display.

WHAT is *WHERE*? As one example, a search in a library catalog for books on *Folklore* may retrieve many books. Books on *Folklore* are often about folklore in some place. The place names in the catalog records can be passed to a place name gazetteer which can supply the latitude and longitude, so that the interface can display a geographical analysis on a map. One should be able to click on any location to find which books are about folklore in that place (Buckland et al., 2007).

WHERE is *WHAT*? In the other direction, if you are interested in, say, mass transit in the capital cities in South America, one might use the cursor to draw a box on the map around the area of interest, then limit the gazetteer feature type to capital cities. The gazetteer should then list the names of the capital cities in South America, ready for searching in the catalog.

WHEN

When people speak or write about time they often do not use calendar dates. Instead, they mark time by mentioning events or named time periods, e.g. "In reign of Charlemagne...", "during the Weimar Republic," "after Vietnam," or "when I was a student." This practice is culturally situated and it resembles the use of place names, so we designed a "Named Time Period Directory," like a place name gazetteer, and created an example using 2,000 records derived from the chronological subdivisions used in *Library of Congress Subject Headings*.

- A place name gazetteer has: a place name – type of place – latitude and longitude – and when the name was used.
- Our named period directory has a period name – type of period – calendar time – and where it happened.
- Sample entries for the 13th century, organized by country, include:
- China: Yuan dynasty, 1260-1368.
- Czech Republic: Mongol invasion, 1242.
- Denmark: Erik Glipping, 1259-1286.

Who was Erik Glipping? What did he do? To find explanations and context our interface provides icons for links. Clicking on a small blue book generates a live search of the *Library of Congress* catalog, a "federated search" using the *Z39.50* protocol. A live search retrieves an up-to-date set of references, as up-to-date as the library's cataloging. Another symbol sends a query to the *Wikipedia*. The *Wikipedia* uses structured URLs for biographical articles, so one can program the interface to go directly to the right page.

WHO

The *Wikipedia* biography of King Eric V (alias Erik Glipping and Eric Klipping) provides a concise text beginning:

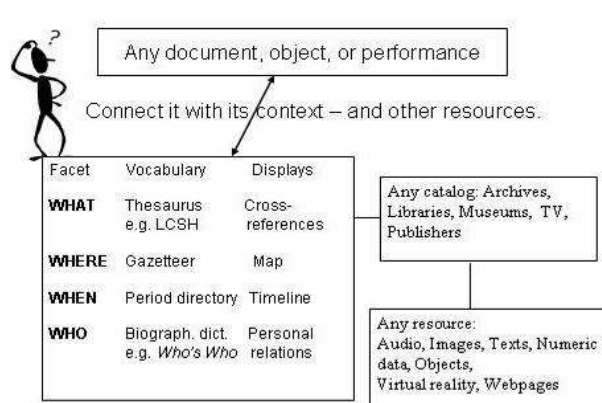
Eric V "Klipping" (1249- November 22, 1286) was King of Denmark 1259-1286) and son of Christopher I. Until 1264 he ruled under the auspices of his mother, the competent Queen Dowager Margaret Sambiria. In this period he was for some time 1261-1262 a prisoner in Holstein after a military defeat . . .

This is helpful, but it may mean little if you do not know the context. Where is Holstein? Who was Margaret Sambiria and what did she do? Fortunately the *Wikipedia* has several links (where underlined) to additional related material in other *Wikipedia* pages. These are internal links. For *knowledge by networking* we need *external* links, dynamic links that will send searches for the latest versions of the best resources on the Internet.

As another example closer to the theme of this conference, consider what a biographical record for the designer of the first desktop search engine would look like:

Emanuel Goldberg, b. Moscow, 1881; son of Grigorii Goldberg; Univ. of Moscow, 1900-04; Ph.D w. Robert Luther, Leipzig Univ., 1906; Assistant, Adolf Miethe, TU Charlottenburg, 1906-07; m. Sophie Posniak, 1907; Prof, Akad. f. graphische Künste, Leipzig, 1907-17; ICA, Zeiss Ikon, Dresden, 1917-1933; Kinamo cine camera, 1921; microdots, 1925; search engine, 1927; Contax 35 mm camera 1932; kidnapped by Nazi SA; refugee in Paris, 1933-37; Laboratory, Tel Aviv, Palestine, 1937; d. 1970.

Several people important in Goldberg’s life are mentioned: Goldberg himself, his father, his dissertation advisor, Luther; his first employer, Miethe; and his wife. Who were they? What did they look like? What did they do? In the spirit of *knowledge by network* one would want links for each name leading to a portrait and a biography wherever they could be found. Various places are named: Moscow, Leipzig, Charlottenburg, Dresden, Paris, and Tel Aviv. Where are these places? What were they like at the time when Goldberg was there? His “life-path” could be displayed as timed movements on a map. Achievements are mentioned: the Kinamo cine camera, microdots, the search engine, and the Contax 35 mm camera. Where are descriptions of these inventions? Do patents exist? Techniques for marking-up biographical texts to link significant names and words to external explanatory resources are under development in a new project, “Bringing lives to light: Biography in context,” led by Professor Ray Larson (Bringing 2007).



Technical comment: In practice, descriptive metadata records are often complex. A library subject heading may have a geographic element and a chronological element, e.g. Architecture – Japan – Edo period, 1600-1868. Records in a place name gazetteer specify the type of feature as well as the name and location. This offers interesting possibilities for building links between the components of different kinds of records. Metadata serve two purposes. The first role of metadata is descriptive, to categorize a document; but, collectively, metadata form structures that can be navigated during search. In this second role descriptive metadata constitutes important infrastructure. These ideas are summarized in Figure 1.

Figure 1: Connecting documents with their contexts

second role descriptive metadata constitutes important infrastructure. These ideas are summarized in Figure 1.

AN EXPERIMENTAL 4W SEARCH INTERFACE

The needs reviewed above have been addressed in an experimental interface developed two years ago. (This description is based on the powerpoint slides presented at the conference (http://www.knowbynet.de/de/download/contributions/buckland_kbn.ppt). The interface has subsequently been redesigned but retains the same basic functionality at <http://ecai.org/imls4w/>.

To overcome the unfamiliarity of vocabulary, interfaces need search term recommender systems. Suppose we see the phrase DJENGIS KHAN and become curious: What is this? – Or who is this? We can ask the interface for a ranked list of *Library of Congress Subject Headings* related to ‘*Djengis Khan*’. This search term recommender system was derived by using statistical techniques to calculate the relative frequency with which words and phrases found in book titles are associated with subject headings assigned to those titles (Buckland et al. 2006)

The recommender service says that DJENGIS KHAN is not an established subject heading, but GENGHIS KHAN, a variant spelling, is. The recommender service lists ten

subject headings that are statistically associated with DJENGIS KHAN. If you select one of the ten headings, the interface can send a search to several resources. In this example four of the recommended headings are for persons somehow associated with Genghis Khan and one is an apparently related period: The tenth heading says MOGUL EMPIRE. What is this? Is MOGUL the same as MONGOL or different? We can ask to see how MOGUL EMPIRE fits in the hierarchical structure of the *Library of Congress Subject Headings*, which shows that the MOGUL EMPIRE was part of the history of India in the 16th to 18th centuries, and that MOGUL is also spelled MUGHAL and MOGHAL, but not MONGOL. It is different. We could click on search links to find more. A live search of the *Library of Congress* catalog found 775 books about the Mogul Empire, the first written in Gujarati. *Google* lists a million links and the *Wikipedia* link leads to a simple, convenient introduction.

The *Wikipedia* article on the *Mughal Empire* explains that it was in Hindustan which included most of India. To look for resources on India in this period we go back to the interface home page where we have four options: (1) We can browse by region, looking for Asia and then South Asia; (2) We can choose a location from a list of states, a list of cities, or a list of countries; (3) We can browse by time period; or (4) We can use a map.

On the map, we zoom to south Asia using the *TimeMap* software developed by colleagues at the *University of Sydney*, Australia (*TimeMap 2000*), which also allows us to set the time period that is of interest. We select the desired period on the time bar, zoom to south Asia, click on India, and the database underlying the interface creates a menu for 15th century India in two parts. The top part has links for India generally, not limited to the 15th century, and currently offers general information about India from the *Wikipedia*, the *CIA's World Factbook*, the *BBC News Country Profiles*, and the *Ethnologue Languages of the World* website, where we could find out about Gujarati. Also a natural history museum website offers details of the animals and plants of India. Note that in each case the interface generates the searches that go directly to the section on India, not just to the home page of website, but the India page. The rest of the page provides links to resources on India in the 15th century. For example, a significant person was Babur, for which the *Wikipedia* has a biography.

The *Metropolitan Museum* in New York has a wonderful website at <http://www.metmuseum.org>. This website uses structured URLs in its *Timeline of Art History* ("toah") section at <http://www.metmuseum.org/toah/>. The structured URLs use a simple code so an interface can know and link to the correct URL in the timeline for any given time and place. It is so simple, so easy, so hospitable, and so effective. It is a wonderful design. All museum websites should use this technique! By using a simple table mapping our interface categories to metmuseum toah codes and inserting the appropriate codes into the Met's URL template, the interface can send the searcher directly to the best page in the Metropolitan website, in this case the *South Asia, 1400-1600 A.D.* portion of their timeline, which shows the choices, with 'MUGHAL DYNASTY 1526-1858' in the top on the right. Our interface page also links an architecture website which can be searched by place and time.

Also listed are relevant cultural atlases created with or by *ECAI* affiliates. One is about the ivories of Begram (in present-day Afghanistan, but pre-dating the Moguls); another includes a video of the changing boundaries of the Mogul Empire; a third is about the Silk Road; and a fourth includes a video of the dramatically changing boundaries of the Mongol Empire from 110 - 1400 A.D., created ten years ago by colleagues at the University of Sydney and available at <http://ecai.org/Area/images/mongol.avi>. The red and white lines on it show the different routes of the Silk Road and the yellow line is the changing boundary of the *Mongol Empire*. We can observe the geo-political changes as Genghis Khan unites the Mongols, the Mongols conquer the Tartars, then attack and destroy the Chin Empire, conquer Russia and Korea, strike into Europe, and so on. When technology has advanced, one should, in the future, be able to stop the video at any year, choose any point, and say, "Find knowledge on the Internet about this place at this time!"

Understanding means knowing context. For this we need to develop *geo-temporal systems*, since geo-spatial is not enough, we need search term recommender systems to help deal with unfamiliar vocabularies, and dynamic links to well-maintained resources to find the most up-to-date information.

Acknowledgments

The work described was partially supported by the *Institute of Museum and Library Services* under National Leadership grant entitled "Support for the Learner: What, Where, When and Who." The work was performed by many people including Fredric G. Gey, Ray R. Larson, Kimberly Carl, Linda-Cathryn Everstz, Jeanette L. Zerneke, and Dr Vivien Petras, now at *Informationszentrum Sozialwissenschaften, GESIS, in Bonn.*

References

The Electronic Cultural Atlas Initiative website (ecai.org) provides extensive resources on *ECAI* activities. For more on the work described above see (ecai.org/imls2004).

Bringing lives to light: Biography in context. [Project website] <http://ecai.org/imls2006> visited Aug 29, 2007.

Buckland, M., A. Chen, F. C. Gey and R. R. Larson: 2006. Search Across Different Media: Numeric Data Sets and Text Files. *Information Technology and Libraries* 25, no 4 (Dec 2006): 181-189. Also at <http://www.ala.org/ala/mgrps/divs/lita/ital/252006/number4decembera/buckland.pdf>

Buckland, M. and others. 2007. Geographic search: catalogs, gazetteers, and maps. *College and Research Libraries*. 68, no. 5 (Sept 2007): 376-387, also at <http://www.ala.org/ala/mgrps/divs/acrl/publications/crljournal/2007/crlsept07/Buckland07.pdf>.

TimeMap: Time-based Interactive Mapping. [Website]. 2000. <http://www.timemap.net/> Visited Aug 29, 2007.

Jens Klump

The Benefits of Cross Linking: The International Continental Scientific Drilling Program

Zusammenfassung

Das **International Continental Scientific Drilling Program (ICDP)** untersucht grundlegende geologische und geophysikalische Fragestellungen von globaler Bedeutung. Dazu werden weltweit Daten aus geologischen Bohrungen gesammelt und der wissenschaftlichen Gemeinschaft zur gemeinsamen Analyse zur Verfügung gestellt.

Die **ICDP Operational Support Group** unterstützt die Bohrprojekte sowohl durch wissenschaftliche und technische Beratung und bietet Datenmanagement während der Feldforschung, der Datenanalyse und der Publikationsphase. Das Datenmanagement basiert auf dem **ICDP Drilling Information System**, mit dem Daten vor Ort erfasst und an die zentrale Datenbank des **GeoForschungsZentrum Potsdam (GFZ)** übertragen werden.

Für die wissenschaftliche Auswertung der Bohrdaten in verschiedenen Projektgruppen ist der Austausch und das gemeinsame Nutzen der Daten von entscheidender Bedeutung. Mit dem Onlinezugriff auf die **Scientific Drilling Database (SDDB)** haben **ICDP** und das **GFZ** eine Plattform für die Verbreitung von Daten aus Bohrprojekten geschaffen.

Der Schwerpunkt der **SDDB** liegt auf solchen Daten aus Bohrprojekten, die Grundlage einer wissenschaftlichen Veröffentlichung waren. Alle Datenveröffentlichungen in der **SDDB** sind durch **Digital Object Identifier (DOI)** eindeutig zitierbar und zugänglich. Darüber hinaus werden in der **SDDB** auch Verweise auf weitere Literatur gegeben, die die Daten interpretiert, zukünftig werden auch die Proben, aus denen die Daten erhoben wurden, nachgewiesen. So entsteht ein Netzwerk aus Daten, wissenschaftlicher Analyse, Literatur

Abstract

The **International Continental Scientific Drilling Program (ICDP)** addresses fundamental scientific problems of global importance as an element of geological and geophysical research programs. It seeks geological sites from around the world and involves the international community of scientists to optimize the results from drilling.

Besides scientific and technical consulting to **ICDP** projects the **ICDP Operational Support Group (OSG)** offers data management in the field, and during the analysis and publication phases of scientific drilling project. Operational data management is based on the **ICDP Drilling Information System (DIS)**, which allows data capture in the field and transfer to a central database at **GeoForschungsZentrum Potsdam (GFZ)**.

Since the start of **ICDP**, data sharing has played an important part in **ICDP** projects and the **OSG** has facilitated dissemination of data within project groups and encourages the sharing of data to promote scientific progress. With the online **Scientific Drilling Database (SDDB)** **ICDP** and **GFZ** **Potsdam** created a platform for the dissemination of data from drilling projects.

The thematic focus of **SDDB** is on data that were used as the basis of a publication. All data publications in **SDDB** are citeable and accessible through **Digital Object Identifiers (DOI)**. In addition, **SDDB** maintains a record of links from its datasets to the literature, where the data are interpreted and will in future also reference the samples from which the data were derived. This network of data, literature and physical samples can be shared with other data providers and

und Proben, das – zusammen mit weiteren Datenportalen – eine globale wissenschaftliche Datenanalyse ermöglicht. Völlig neuartige Mehrwertdienste, die auf die semantische Struktur dieses Netzwerkes zurückgreifen, sind nun denkbar.

portals to enhance discovery of data through discipline specific catalogues. Novel added-value services may be developed by exploiting the semantic nature of this network.

Introduction

Projects in the *International Scientific Continental Drilling Program (ICDP)* produce large amounts of data. Since the start of *ICDP*, data sharing has played an important part in *ICDP* projects and the *ICDP Operational Support Group* has facilitated dissemination of data within project groups (Conze et al., 2007). Some of these data later became the basis of scientific publications, while others remained unpublished. However, in most cases the data sets themselves were not available outside of the respective projects. With the online *Scientific Drilling Database (SDDB)* (<http://www.scientificdrilling.org>) *ICDP* and *GeoForschungsZentrum Potsdam (GFZ)* created a platform for the dissemination of data from drilling projects (Klump and Conze, 2007). The thematic focus of *SDDB* is on data that were used as the basis of a publication, but some published datasets are data publications in their own right.

To publish data requires that data are citeable. This means, a mechanism is needed that ensures that the location of the referenced data on the internet can be resolved at any time. In the past, this was a problematic issue because URLs are short-lived, many becoming invalid after only a few months. Data publication on the internet therefore needs a system of reliable pointers to a web publication to make these publications citeable. To achieve this persistence of identifiers for their conventional publications many scientific publishers use *Digital Object Identifiers (DOI)* (Brase, 2004). *GFZ Potsdam* is a member of the project "Publication and Citation of Scientific Data" (*STD-DOI*), which is funded by the *German Research Foundation*. In this project the *German National Library for Science and Technology (TIB Hannover)*, together with *GFZ Potsdam*, *Alfred Wegener Institute (AWI) Bremerhaven*, *University of Bremen* and the *Max Planck Institute for Meteorology* in Hamburg set up a system to assign *DOIs* to data publications.

Data management in scientific drilling projects

Data management in scientific drilling projects of the *International Continental Scientific Drilling Program (ICDP)* and the *Integrated Ocean Drilling Program (IODP)* pursues two main goals: on the one hand, to capture of drilling and scientific data as early as possible, and on the other hand, manage the long-term storage and dissemination of these data. The data capture in both *ICDP* projects and *IODP-Mission Specific Platform (MSP)* expeditions has three phases: fieldwork phase, analysis phase, and publication phase. Drilling, curation, logging, and basic scientific data are captured at the drill site during the fieldwork phase. This phase is followed by the analysis phase during which detailed measurements, descriptions, images and analytical data are captured within a laboratory setting. The data are subsequently transferred to the long-term data storage system. During the fieldwork and analysis phases data are, in most cases, only accessible by the scientists involved in the project.

In the publication phase of a scientific drilling project analytical data become available and their interpretations are published in the literature. The publication platform for

public data from *ICDP* drilling projects is the *Scientific Drilling Database (SDDB)* (Klump and Conze, 2007). Some of these data can be imported directly from the *DIS*, but for most data *ICDP* has to rely on the authors to supply the data. Since most drilling project track their publications for reporting purposes the *SDDB* editorial staff has an overview of publications related to a specific project. The authors are then contacted and asked for their cooperation in the data publication process. At the present stage of the development of *SDDB* the right to upload data is still restricted to the *SDDB* operators. However, a prototype of the *SDDB* data upload assistant already exists and is being tested for usability. The aim is to make the *SDDB* data upload assistant as easy to use as possible so that authors can upload data themselves. Linking data to associated publications and to physical sample material is part of the editorial process

Publication of scientific data is labour intensive and two points are crucial to make the system scaleable: (1) the publication process has to be automated as far as possible and duplication of work, e.g. metadata editing, has to be avoided, and (2) authors have to be motivated to participate in the data publication process. However, the process of scientific data publication is still hampered by structural deficits which will be discussed in the section below.

Data publication today

On 22 October 2003, a group of leading research institutions and research funding institutions published the '*Berlin Declaration on Access to Knowledge in the Sciences and Humanities*' in order to "[...] promote the Internet as a functional instrument for a global scientific knowledge base and human reflection and to specify measures which research policy makers, research institutions, funding agencies, libraries, archives and museums need to consider." (Berlin-Declaration, 2003). The *Berlin Declaration* has since been signed by 242 scientific bodies worldwide. The OECD Governments have also recognised the importance of open access to knowledge. This new policy has been formulated in the '*Communiqué on Science, Technology and Innovation for the 21st Century*' issued at the OECD ministerial meeting, 29-30 January 2004 (OECD, 2004). It has been followed by a recommendation on access to research data to the OECD Council (OECD, 2006). This 'soft legislation' has to be transferred into national legislation by the OECD member states.

Knowledge, as published through scientific literature, is the last step in a process originating from primary scientific data. These data are analysed, synthesised, interpreted, and the outcome of this process is published as a scientific article. The *Berlin Declaration* and the *OECD Ministerial Communiqué* look at the outcome of this process. Because scientific knowledge is ultimately derived from data, we wish to examine more closely the beginning of this process, the issues of data sharing and data publication.

Some organisations encourage scientists to share data freely and even make data sharing a part of their funding policy (e.g. NIH, 2003). In addition, cases of scientific misconduct in recent years have highlighted the importance of making scientific data available. As a consequence, the *German Research Foundation*, and other science organisations, adopted '*Recommendations for Good Scientific Practice*' as part of their policy. They require that institutes archive data, which were used as a basis of a publication, on safe storage media for a minimum duration of ten years (DFG, 1998). Besides being a matter of common sense and good scientific conduct, thorough documentation of experiments also makes economic sense (Alexander et al., 2004).

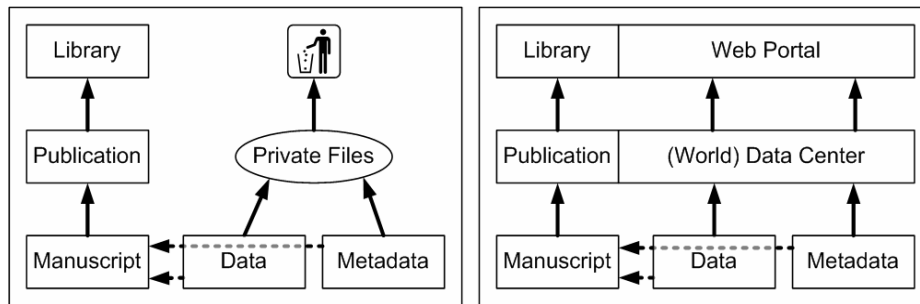


Figure 1: (left) Schematic depiction of the flow of scientific information from research to published library resources as currently practiced (modified after Helly et al., 2003). (right) Potential approach based on data publication by data centres and content syndication to scientific web portals, which could also be library catalogues. Inter-linking publications and their underlying data will create new scientific products with added value. The dashed lines from data and metadata to the manuscript reflect the limited publication of these sources in our conventional scientific journals.

secure institutional repositories, and effectively are lost (Figure 1, left). This lack of access to scientific data is an obstacle to interdisciplinary and international research. It causes unnecessary duplication of research efforts, and the verification of results becomes difficult, if not impossible (Dittert et al., 2001). Large amounts of research funds are spent every year, while already existing data remain underutilised (Arzberger et al., 2004).

Publication and citation of scientific data

The screenshot shows the 'Scientific Drilling Database' interface. The main content area displays a 'Dataset Description' for a study by Heim, Birgit, Oberhänsli, Hedi, Fietz, Susanne, Kaufmann, Hermann (2006). The description includes a citation, DOI (10.1594/GFZ.SDDB.1043), title, and abstract. There are buttons for 'Download Citation (EndNote)' and 'Show in Google Earth'. A 'Related Identifier' section lists a reference to a 2005 paper. An 'Activities' table provides detailed metadata for the dataset, including latitude, longitude, elevation, date, program, expedition, platform, and gear.

Figure 2: Screenshot of a data citation in SDDB. Note the buttons for download of citation into a reference manager and for the visualization of sampling locations in Google Earth. The Digital Object Identifier of this dataset is doi:10.1594/GFZ.SDDB.1043

into the catalogue of the German National Library of Science and Technology (TIB) (Brase, 2004).

In the project STD-DOI, the German National Library of Science and Technology TIB acts as a registration agency for persistent identifiers. For every data publication, it requests a set of metadata to be incorporated into the library catalogue. The data sources are the participating World Data Centers in Germany, WDC-MARE (Bremen / Bremerhaven), WDC Climate (Hamburg), WDC-RSAT (Oberpfaffenhofen), and the proposed WDC-TERRA (Potsdam). The data centres act as registration agents for scientific and technical data DOIs.

Despite advances in the policies regarding access to data, only a very small proportion of the original research data are published in conventional scientific journals. Existing policies on data archiving notwithstanding, in today's practice data are primarily stored in private files, not in

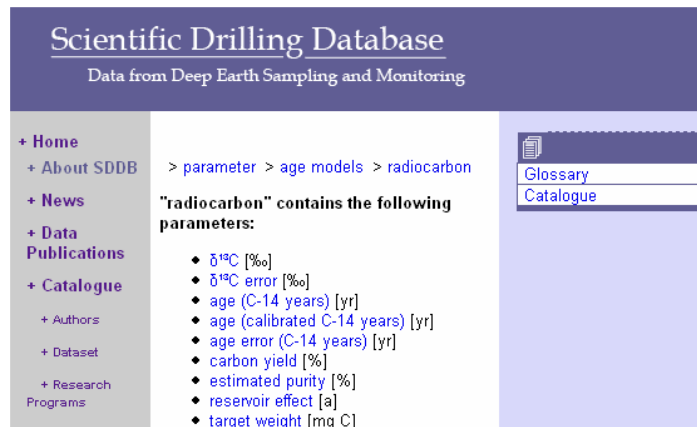
To improve access to data and to create incentives for scientists to make their data accessible, the German CODATA group initiated a project on publication and citation of scientific data which was funded by the German Research Foundation (DFG) for the periods 2003-2005 and 2006-2008 (STD-DOI, 2003).

This project uses persistent identifiers (both DOI and URN) to identify datasets available in a digital format. The identifier is resolved to the valid location (URL) where the dataset can be found. This approach meets one of the prerequisites for citability of scientific data published online. In addition, the data publications are included

These data centres are also responsible for technical quality control in their data domains, at the same time they also act as long-term archives. The project participants thus encompass all functions necessary for the publication of scientific data.

On May 1st 2005 the *TIB* became the world's first *DOI* registration agency for scientific primary data, working in cooperation with the *World Data Center Climate (WDCC)* at the *Max Planck Institute for Meteorology Hamburg*, *GeoForschungsZentrum Potsdam*, *World Data Center for Marine Environmental Sciences (WDC-MARE)* at the *Alfred Wegener Institute Bremerhaven* and at the *University of Bremen* and technically advised by the *Research Center L3S Hannover*. Through this project, the foundations have been laid for a system of scientific data publication.

Information discovery through semantic linking



In many online scientific databases, the access to data is through some kind of search interface, which may consist of a simple field for the entry of keywords or may offer more elaborate search criteria. However, in most cases database users do not know the precise contents of a database, yet no database is so perfectly comprehensive to offer datasets matching any search query. Database users would therefore rather "browse" the contents like in a catalogue. In the design of *SDDB* we therefore took a very different approach to other databases and designed the graphical user interface to contain dynamically generated catalogue listings and cross-links, e.g. between datasets and sample material, datasets and authors,

Figure 3: Screenshot of *SDDB* showing a list of analytical parameters from the context of radiocarbon dating of fossil materials. All parameter names and their parent categories are active links that can be used to browse the contents of *SDDB* by analytical parameters. Other categories, like authors of data sets, or instruments used for sampling, can be browsed in a similar way.

datasets and parameters, etc., to allow the database user to browse through the *SDDB* contents (Figure 3).

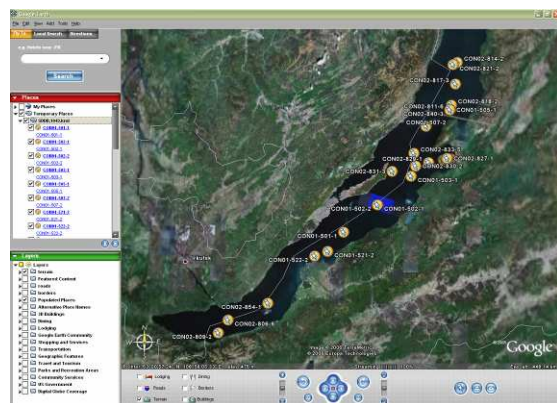


Figure 4: Sampling locations of doi:10.1594/GFZ.SDDB.1043 by (Heim et al., 2006) displayed in the *Google Earth* virtual globe. The place marks at the sampling locations link back to an online description of the fieldwork activities at the respective locations as recorded in *SDDB*.

interactive and act as links back to *SDDB*.

Any sample taken in the field is taken from a geographical and geological context. Not every database user will be familiar with the locality from where the sample was taken. Visualization of this sampling context helps to assess whether the offered datasets are useful to the particular question asked, or which subset of data to choose. At present this geographical visualization primarily shows the sampling positions in their geographical context. Virtual globes, such as *Google Earth*, are useful and intuitive tools for geographical visualization (Butler, 2006; Lyon et al., 2006). To show the sampling locations in *Google Earth*, *SDDB* offers a kml-file for download with every dataset (Figure 2). The kml-file can be automatically imported as *place marks* and viewed in *Google Earth* (Figure 4). The *place marks* are interactive

In a second step we plan to add more specific maps which will be displayed alongside the data or as separate maps in an online geographical information system. This information system will allow the dynamic generation of customised maps through a standardised *Web Map Service*. The elements of this service, i.e. the thematic map layers will be also published as data publications for re-use by other researchers (Heim et al., in press).

Starting in 2008, *ICDP* will also assign identifiers to physical samples taken in the field or in the laboratory to allow unambiguous identification of sample materials and associated data. This identifier is called an *International Geosample Number (IGSN)* and based, like *DOIs*, on the *Handle Service*®. The service for registration of *IGSN* will be run in co-operation with the *System for Earth Sample Registration (SESAR)*, (<http://www.geosamples.org>) at *Lamont Doherty Earth Observatory of Columbia University* (New York, NY) (Lehnert et al., 2006).

Ideally, literature should already reference the materials used and the data derived from these. Since this is not yet done, repositories publishing data and tracking sample material record the literature based on these data and samples in their databases. In the case of the *STD-DOI* project, its metadata profile includes identifiers of related material, e.g. literature interpreting the data, related datasets, or samples from which the data were derived. These metadata can then be used to create ontologies interlinking literature, data and samples.

The challenging task ahead is that of interlinking literature, data and samples with as little editorial work as possible. Keeping the amount of work small is essential to allow the indexing of the back catalogue of already existing works. A key technology to solve this task is the automatic creation of ontologies, which can be generated automatically by text mining applications. These ontologies can be combined with ontologies generated from reference lists and from metadata.

Conclusions

Data sharing is a key strategy in *ICDP's* data management and technical development over the years has aimed to make the data management processes as effective as possible. However, systematic data management is still additional work on the scientists for which they have so far not received the necessary recognition. Therefore, scientists need incentives to share data. A possible incentive may be to make data sharing a proper scientific publication. *ICDP* has implemented the tools necessary for persistent references to published data.

Data will only be re-used if their existence is known. Therefore data publications must be included into library catalogues and scientific portals. They must also be accessible to users that are not aware of the existence of the data or the database. This can be achieved by syndication of metadata among data portals and by semantic linking between literature, data and physical samples from which the data were derived.

The scientific workflow of creating knowledge is slowly adopting the new tools offered by the internet-based information revolution (Berman and Moore, 2006). However, data curation is not at the focus of scientific work and therefore scientists are not willing to invest much of their time in any work related to this task. Therefore it is essential to provide researchers with efficient, easy to use tools that are aligned with the scientific workflow. A high degree of automatisation will take away as much as possible of the workload of data curation and metadata editing from the scientists.

Acknowledgements

The development of the *Scientific Drilling Database (SDDB)* is funded by the *International Scientific Continental Drilling Program (ICDP)*. The project "*Publication and Citation of Scientific and Technical Data*" (*STD-DOI*) is supported by the *German Research Foundation, Libraries and Information Systems (DFG-LIS)*.

Literature

- Alexander, William, Joshua Berlin, Philip Cyr, Andrew Schofield and Leslie Platt (2004), Realities at the leading edge of research – Good practice and proper conduct in research pay off, scientifically and economically, *EMBO Reports*, 5(4), 324-329.
- Arzberger, Peter, Peter Schröder, Anne Beaulieu, Geoffrey C. Bowker, Kathleen Casey, Leif Laaksonen, David Moorman, Peter Uhlir and Paul Wouters (2004), An International Framework to Promote Access to Data, *Science*, 303(5665), 1777-1778. doi:10.1126/science.1095958
- Berlin_Declaration (2003), Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, [online] Available from: <http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>
- Berman, Fran, and Reagan W Moore (2006), Designing and Supporting Data Management and Preservation Infrastructure, *CTWatch Quarterly*, 2(2), 7.
- Brase, Jan, and Jens Klump (2007), Zitierfähige Datensätze: Primärdaten-Management durch DOIs, in *WissKom 2007 : Wissenschaftskommunikation der Zukunft ; 4. Konferenz der Zentralbibliothek, Forschungszentrum Jülich*, 6. – 8. November 2007, Bd. 18, herausgegeben von R. Ball, p. 159-167, Forschungszentrum Jülich, Jülich, Germany.
- Butler, Declan (2006), Virtual globes: The web-wide world, *Nature*, 439(7078), 776-778. doi:10.1038/439776a
- Conze, Ronald, Hans-Joachim Wallrabe-Adams, u. a. (2007), Joint ICDP and IODP Data Management for Scientific Drilling Expeditions, *Scientific Drilling*, 4, 32-33. doi:10.2204/iodp.sd.4.07.2007
- DFG (1998), Regeln guter wissenschaftlicher Praxis, Deutsche Forschungsgemeinschaft. [online] Available from: http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/self_regulation_98.pdf
- Dittert, Nicolas, Michael Diepenbroek and Hannes Grobe (2001), Scientific data must be made available to all, *Nature*, 414(6862), 393. doi:10.1038/35106716
- Heim, Birgit, Jens Klump, Natalie Fagel and Hedi Oberhänsli (2008), Assembly and concept of a web-based GIS within the paleolimnological project CONTINENT (Lake Baikal, Siberia), *Journal of Paleolimnology*, 39(4), 567-584, doi:10.1007/s10933-007-9131-0.
- Helly, John, Hubert Staudigel and Anthony Koppers (2003), Scalable models of data sharing in Earth sciences, *Geochemistry, Geophysics, Geosystems – G (super 3)*, 4(1). doi:10.1029/2002GC000318
- Klump, Jens, and Ronald Conze (2007), The Scientific Drilling Database (SDDb) – Data from Deep Earth Monitoring and Sounding, *Scientific Drilling*, (4), 30-31, doi:10.2204/iodp.sd.4.06.2007.
- Lehnert, Kerstin, Sri Vinayagamorthy, Branco Djapic and Jens Klump (2006), The Digital Sample: Metadata, Unique Identification, and Links to Data and Publications, *EOS, Transactions, American Geophysical Union*, 87(52, Fall Meet. Suppl.), Abstract IN53C-07. [online] Available from: http://www.agu.org/meetings/fm06/fm06-sessions/fm06_IN53C.html
- Lyon, Steve W, Arthur J Lembo, M. Todd Walter and Tammo S. Steenhuis (2006), Internet Mapping Tools Make Scientific Applications Easy, *EOS, Transactions, American Geophysical Union*, 87(38), 386.
- NIH (2003), Final NIH Statement on Data Sharing, National Institute of Health, Bethesda, MD. [online] Available from: <http://grants2.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>
- OECD (2006), Recommendation of the Council concerning Access to Research Data from Public Funding, Organisation for Economic Co-operation and Development, Paris, France.
- OECD (2004), Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 – Final Communiqué, Organisation for Economic Co-operation and Development, Paris, France. [online] Available from: http://www.oecd.org/document/0,2340,en_2649_34487_25998799_1_1_1_1,00.html

Monika Linder

Networking Archaeological Knowledge Projects of the German Archaeological Institute

Zusammenfassung

Zum *Deutschen Archäologischen Institut (DAI)* gehören elf Zweiganstalten im In- und Ausland mit Sitz der Zentrale in Berlin. Während in den Anfängen der Schwerpunkt der Tätigkeit in den Ländern des Mittelmeerraums und des Vorderen Orients lag, geht das Institut heute weltweit seinem Auftrag der Grundlagenforschung zum Verständnis historischer Kulturen nach.

In allen Abteilungen befinden sich sämtliche die Archäologie betreffende Materialgattungen: moderne und historische Grabungsdokumentationen, Nachlässe von Alttertumswissenschaftlern, Materialien zur Institutionengeschichte, Fotosammlungen, Historische und moderne Bibliotheksbestände und nicht zuletzt einige originale Objekte.

Ziel ist es, eine *DAI*-Architektur aufzubauen, die als zentrales Instrument für die Wissenschaftler dient und eine einheitliche Zugangsweise für die diversen Materialgattungen schafft. Hierfür sollen *DAI*-Daten sowohl in einer eigenen objektorientierten *DAI*-Umgebung als auch in Portalen in Kooperation mit anderen Partnern nachgewiesen werden. Bestehende und geplante Strukturen werden vorgestellt, von Virtuellen Katalogen bis hin zu Möglichkeiten der Referenzierung durch das Semantische Netz (*CIDOC-CRM*).

Abstract

The *German Archaeological Institute* consists of eleven branch offices, within Germany and abroad. Head office is in Berlin. In its early years the Institute's focus was on the Mediterranean and Middle East. Today, the Institute's pursues worldwide its mission to conduct basic research on historic cultures.

All divisions of the Institute possess objects from the material categories common to the fields of archaeology: modern and historic documentation of excavations, collections compiled by researchers of antiquities, materials relating to the history of the institution, photographic collections, historic and present-day bibliographic materials and, last but not least, more than a few original objects. Our objective is two-fold: to create an institutional architecture that serves as a central instrument for the scientists and offers an uniform approach to diverse categories of material.

To this end, an object-oriented environment will pool and verify data from the Institute. It will also link to other portals from co-operating institutions. In this presentation existing and planned structures will be introduced, from virtual catalogues to the possibilities of reference over by the Semantic Network (*CIDOC-CRM*).

Ladies and Gentlemen, in this short presentation I am pleased to be able to introduce to you three subjects:

- the history and structure of the *German Archaeological Institute (DAI)*,
- the Institute's existing material categories
- and several networking projects of the Institute.

1) History, Structure, Tasks

On 21 April 1829, a circle of friends in Rome consisting of artists and diplomats founded the *Instituto di corrispondenza archeologica*. Their aim was to publicise and research

monuments to ancient art, epigraphy and topography. As such, the Institute in Rome is part of archaeological history, for its emergence is regarded as a step toward the development of archaeology as a discrete academic discipline. Crown Prince and later King Frederick William IV of Prussia assumed the patronage of the Institute. When Eduard Gerhard, for all intents and purposes the initiator of the Institute, relocated from Rome to Berlin, the directorship was also transferred to Berlin, though the Institute itself remained in Rome. After 1859, Prussia assumed responsibility for the Institute's finances. In 1871, it became a Prussian state institute; in 1874, the Institute became an Imperial institute. That same year, an Athens department was opened as well.



Figure 1: Copyright: D-DAI-ZEN: Archiv Zentrale: Theodor Wiegand in the archeological excavation Didyma

During the twentieth century additional departments were opened in Cairo, Istanbul, Madrid, Baghdad and Teheran. In addition, three national commissions were also set up: the *Greco-Roman Commission* in Frankfurt, the *Commission for Ancient History and Epigraphy* in Munich and the *Commission for the Archaeology of Non-European Cultures* in Bonn. Additional stations were eventually established in Sanaa, Damascus and Ingolstadt. In 1995, a Eurasia department for archaeological research in the area of the Commonwealth of Independent States and neighbouring nations was founded, in which the Teheran department was incorporated as field office. In 1996,

the Baghdad department was consolidated with Damascus and San'a' to create the Orient branch. The DAI is directly subordinate to the German Foreign Office. The Institute's main office is located in the so-called Wiegand House, former place of residence of archaeologist and DAI president (1932-1936) Theodor Wiegand. In 2007, the Institute reached cooperative agreements with the German Protestant Institute of Archaeology which will allow us to organise excavations from Jerusalem and Amman.

2) Categories of Material

a) Documentation of Excavations

The most important types of archaeological documentation are excavation notes with historical and current materials, diagrams, drawings, cartographic results and much more. These sources are incorporated into database systems, inter alia, Geographic Information Systems, allowing us to display different layers, satellite images, vector films or geometric enquiries as you see here from Taymer in the north of Saudi Arabia.

b) Photographic Documentation

An important instrument parallel to the excavation documentation is the routine photographic documentation of excavations and discoveries. From these, as well as through targeted photo campaigns, accessions from completed research projects and the private collections of important scholars, the photographic collections of the individual departments have expanded over the decades. A part of these collections is documented in *ARACHNE*, a database administered by the Research Archive for Ancient Sculpture at the University of Cologne (<http://www.arachne.uni-koeln.de>). In the course of individual projects, portions of

these extensive materials were scanned and processed by scholars; two examples are the digitisation of 40,000 negatives of sculptures from the *DAI's* Rome department and 40,000 glass negatives of historical photographs. Image databases on the Roman Sarkophagcorpus and the architecture and inventory of the *Basilica Aemilia* of the *Forum Romanum* have been developed.

c) Archival Documentation

In addition to excavation and photographic archives, at almost every *DAI* location visitors will find several types of archival material: personal papers and autographs, as well as processed institutional files. For a description of these types using examples of holdings from our headquarters office in Berlin, see http://www.dainst.org/index_6373_de.html.

The headquarters' archive, the Archive for the History of Archaeology, was established on the Institute's one-hundredth anniversary in 1929. The collection's main focus was academic correspondence. Today, the collection consists of more than 250 personal papers from antiquity researchers covering classical and Near Asian archaeologists, architectural historians, ethnographers and Egyptologists. Stored in approximately 850 archival boxes are more than 25,000 scholarly letters, diaries, manuscripts, photographs, glass plate negatives as well as drawings and watercolours. Additional materials pertaining to the history of the Institute include statutes, anniversaries, archaeological conventions in 1939 and 1988, the Institute's building, photographs, films, audio recordings, etc. The processed papers of the Institute have been completely preserved from the period after 1936, prior to that year there are gaps in the collections. Also part of the archive is a Biographical Collection covering the approximately 7700 corresponding and full members of the Institute.

d) Libraries

At each of its locations, the Institute boasts highly specialised library collections, which, taken together, assume a place of prominence within Germany and abroad. (At the moment, the full collection consists of 850,000 published volumes.) Evolved over time, they represent excellent reference libraries for those scholars interested in classical antiquity. To a large extent all branches of archaeology are covered in these departmental libraries: in some cases, for example in Istanbul, these library collections provide for many in-country researchers the only adequate access to pertinent research literature. In addition, the libraries offer numerous opportunities to foster intense contacts to fellow scholars in their home country as well as abroad. A large percentage of the library collections have been captured in the following centralised catalogue (*ZENON-DAI*):

- http://www.dainst.org/index_49_de.html
- http://www.dainst.org/index_3454_de.html

We continue to expand this catalogue's offerings.

In addition, much appreciated by international experts are the Institute's subject bibliographies, maintained by the *DAI* for decades: (http://www.dainst.org/index_6936_de.html)

- *Archaeological Bibliography*
- *Bibliography of the Archaeology of the Iberian Peninsula*
- *Subject catalogue of the Roman-German Commission*
- *Bibliography of the Archaeology of Eurasia* (completed bibliography)

These data are at the moment also being prepared for inclusion in a meta-search project funded by the *German Research Foundation (DFG)*. Entitled Propylaeum, this virtual library devoted to the subject of antiquities research is available here: <http://www.propylaeum.de>

In this endeavour the *German Archaeological Institute* is a co-operation partner of the following special collections libraries:

- The *Bavarian State Library* in Munich
- *Heidelberg University Library*
- *Tuebingen University Library*
- *Humboldt University*.

As the historical collection is also of special significance, several older publications have been digitised and captured in a keyword browser of the *Research Archive for Antique Sculptures* (Stichwerkbrowser) (<http://www.arachne.uni-koeln.de>)

In co-operation with the University Library at the University of Heidelberg, the Institute's Rome department is planning a large digitisation project covering the areas of classical archaeology and Egyptology.

3) Networking Projects

The goal of the *German Archaeological Institute* is the networking of all the above-mentioned materials within and beyond our Institute. One way to reach this goal is through use of the tools of the semantic web; for instance, we are planning to implement *CIDOC-CRM* (<http://cidoc.ics.forth.gr>)

In light of the Institute's plurality of information sources, this decision offers us the long-term prospect of integrative scholarly management, drawing together the data generated in individual projects, making these data sets not only retrievable, but also combinable and visible in novel forms. Not only, as remains sensible and necessary, in the structure of the respective primary data sources (*Arachne*, *ZENON-DAI*, *Stichwerkbrowser*), but rather on a level defined by the structures of *CIDOC-CRM*. As such, the data sets are not merely combined, but rather can be experienced according to a uniform model. This observation applies not only to common searches of diverse sources, but instead to a browsing-based approach oriented toward the contexts of the respective objects and types of information (see my outline on the architecture necessary to achieve this objective over the long term).

In the context of a co-operative project involving the *German Archaeological Institute* and Tufts University outside Boston, the first steps toward this end have been implemented for parts of the *Perseus Digital Library* (<http://www.perseus.org>) and *Arachne*. XML data sets extracted from both information systems, processed via XSLT in RDF, can, through faceted browsing, be viewed and navigated with a Longwell Browser. The first concrete implementation has thus been realised.

Plans are made for an extension of our referencing approach from the world of object data to the world of book data, through the linking of data sets with illustrations of antique statues with the object data in a photo database. Further, in co-operation with the *Perseus Digital Library* we plan to test the *FRBR* (*functional requirements for bibliographic records*) application. (Die *Perseus Digital Library* benutzt bereits diese Applikation, weitere Informationen unter: <http://www.ifla.org/VII/s13/frbr/frbr.pdf>.)

We aspire to undertake additional networking projects involving other archaeological data, for example networking projects where excavation and archival data are the focus. We are continuing our efforts to shape "archaeology in the semantic web".

Monika Linder

Archäologisches Wissen vernetzt – Projekte aus dem Deutschen Archäologischen Institut

Sehr geehrte Damen und Herren, es freut mich, Ihnen in einer Kurzpräsentation die Geschichte und Struktur des *Deutschen Archäologischen Instituts* (<http://www.dainst.org/>), die vorhandenen Materialgattungen sowie einige Projekte und deren Vernetzungen vorstellen zu können.

Geschichte, Struktur, Aufgaben

Am 21. April 1829 gründete in Rom ein Freundeskreis aus Gelehrten, Künstlern und Diplomaten das *Instituto di corrispondenza archeologica*, um die Denkmäler der antiken Kunst, der Epigraphik und der Topographie bekanntzumachen und zu erforschen. Das Institut in Rom ist als solches selbst ein Bestandteil der Archäologiegeschichte, da seine Entstehung als ein Schritt der Herausbildung der Archäologie zu einer eigenständigen akademischen Disziplin gelten kann.

Der preußische Kronprinz und spätere König Friedrich Wilhelm IV. hatte die Schirmherrschaft übernommen, als Eduard Gerhard, der eigentliche Initiator des Instituts, 1832 von Rom nach Berlin übersiedelte, verlagerte sich die Leitung des in Rom fortbestehenden Instituts nach Berlin. Preußen übernahm ab 1859 die regelmäßige Finanzierung des Instituts. 1871 wurde es preußische Staatsanstalt, 1874 Reichsinstitut. Im selben Jahr kam es zur Gründung der Abteilung Athen.



Figure 1: Copyright: D-DAI-ZEN: Archiv Zentrale: Theodor Wiegand auf der Ausgrabung Didyma

Im 20. Jh. wurden die Abteilungen Kairo, Istanbul, Madrid, Bagdad und Teheran sowie drei Kommissionen im Inland (*Römisch-Germanische Kommission* in Frankfurt, *Kommission für Alte Geschichte und Epigraphik* in München, *Kommission für Archäologie Außereuropäischer Kulturen* in Bonn) gegründet. Schließlich wurden weitere Stationen in Sanaa, Damaskus und Ingolstadt eingerichtet. 1995 fand die Gründung der Eurasien-Abteilung für archäologische Forschungen im Gebiet der GUS-Staaten und benachbarter Länder statt, in die die Abteilung Teheran als Außenstelle eingegliedert wurde. Die Abteilung Bagdad wurden

1996 und die Stationen Damaskus und Sanaa zur Orient-Abteilung zusammengeschlossen. Das DAI untersteht heute dem Auswärtigen Amt und hat seinen Hauptsitz im so genannten „Wiegand-Haus“, dem ehemaligen Wohnhaus des Archäologen und DAI-Präsidenten (1932-1936) Theodor Wiegand. Ganz neu seit diesem Jahr wurden Kooperationsvereinbarungen mit dem *Deutschen Evangelischen Institut (DEI)* geschlossen, die von Jerusalem und Amman aus Grabungen organisieren.

Materialgattungen – Grabungsdokumentation

Kern archäologischer Dokumentation sind Grabungsaufzeichnungen mit historischen sowie aktuellen Materialien, Plänen, Zeichnungen, kartographischen Ergebnissen und vielem mehr. Diese werden in Datenbanksysteme eingespeist, u.a. in GIS-Systeme (*Geoinformationssysteme*), also rechnergestützte Informationssysteme, mit denen raumbezogene Daten digital erfasst und redigiert, reorganisiert, modelliert und analysiert werden können, bis hin zu der Möglichkeit dreidimensionaler Darstellung von Bauten und Gelände. So können verschiedene Schichten, Satellitenbilder, Vektordateien oder geometrische Untersuchungen dargestellt werden, wie hier in Taymar im nördlichen Saudi-Arabien.

Photodokumentation

Parallel dazu ist ein wichtiges Instrument die laufende Photodokumentation der Grabungen und Funde. Hieraus sowie durch gezielt durchgeführte Fotokampagnen, über Zuwächse aus abgeschlossenen Forschungsprojekten und aus Nachlässen bedeutender Wissenschaftler sind die Fototheken der einzelnen Abteilungen über Jahrzehnte hin erwachsen. Ein Teil davon ist in der vom *Forschungsarchiv für Antike Plastik* an der Universität Köln, einem Konsortialpartner des DAI, betriebenen Datenbank *Arachne* nachgewiesen (<http://www.arachne.uni-koeln.de>). Im Rahmen einzelner Projekte wurden bereits Teile dieses umfangreichen Materials gescannt und wissenschaftlich bearbeitet, z.B. durch die Digitalisierung von 40.000 Skulpturennegativen des DAI Rom oder von 40.000 Glasnegativen mit historischen Aufnahmen. Es wurde eine Bilddatenbank zum römischen Sarkophagcorpus aufgebaut sowie zur Architektur und zum Inventar der *Basilica Aemilia* am *Forum Romanum*.

Archivdokumentation

Neben den Grabungs- und Photoarchiven finden sich an fast jedem Standort des DAI sich zwei weitere Gattungen an Archivgut: Nachlässe und Autographen sowie Altregistraturen und Institutsakten. Exemplarisch seien sie charakterisiert am Beispiel des Bestandes der Zentrale in Berlin (http://www.dainst.org/index_6373_de.html): Das Archiv der Zentrale, das als *Archiv zur Geschichte der Archäologie* 1929 zum 100. Institutsjubiläum mit dem Schwerpunkt Gelehrtenbriefe gegründet wurde, besteht aus über 250 Nachlässen von Altertumsforschern – Klassischen und Vorderasiatischen Archäologen, Bauforschern, Prähistorikern und Ägyptologen – im Umfang von rund 850 Archivkästen. Darin befinden sich mehr als 25.000 Gelehrtenbriefe sowie Tagebücher, Manuskripte, Fotos, Glasplattennegative, aber auch Zeichnungen und Aquarelle. Zu den Materialien der Institutsgeschichte zählen z.B. Statuten, Jubiläen, Archäologie-Kongresse 1939 und 1988, Institutsgebäude, Fotos, Filme, Tonbänder etc. Die Altregistratur des Instituts ist vollständig ab 1936 erhalten, vorher lückenhaft. Schließlich gehört hierzu die *Biographica-Sammlung* zu den mittlerweile rund 7700 Korrespondierenden und Ordentlichen Mitgliedern des Instituts.

Bibliotheken

Das DAI besitzt an jedem Standort sehr spezialisierte Bibliotheksbestände, die im In- und Ausland eine herausragende Stellung einnehmen (Gesamtbestand zurzeit ca. 850.000 Bände). Historisch gewachsen stellen die Bibliotheken ausgesprochene Referenzbibliotheken für die Klassische Altertumswissenschaft dar. Weitgehend alle Zweige der Archäologie sind in den einzelnen Abteilungen abgedeckt, teilweise, wie in Istanbul, stellen sie im jeweiligen Land für viele Forscher vor Ort den einzigen passenden Zugang zur einschlägigen Forschungsliteratur bereit. Außerdem bieten sie vielfältige Gelegenheiten, mit anderen Forschern aus dem In- und Ausland in intensiven Kontakt zu treten.

Ein großer Teil der Bibliotheksbestände (http://www.dainst.org/index_49_de.html) ist im zentral geführten Katalog *ZENON-DAI* (http://www.dainst.org/index_3454_de.html) nachgewiesen, der kontinuierlich ausgebaut wird. Zusätzlich führt das DAI seit Jahrzehnten

Fachbibliographien, die in der internationalen Fachwelt sehr geschätzt sind (http://www.dainst.org/index_6936_de.html): *Archäologische Bibliographie*, *Bibliographie zur Archäologie der Iberischen Halbinsel*, *Sachkatalog der Römisch-Germanischen Kommission*, *Bibliographie zur Archäologie Eurasiens* (abgeschlossene Bibliographie).

Diese Daten werden derzeit zusätzlich durch ein DFG-gefördertes Projekt für die Metasuche in der Virtuellen Fachbibliothek Altertumswissenschaften *Propylaeum* (<http://www.propylaeum.de/>) aufbereitet. Das DAI ist hier Kooperationspartner der Sondersammelgebietsbibliotheken der *Bayerischen Staatsbibliothek München*, der *UB Heidelberg*, der *UB Tübingen* sowie der *Humboldt-Universität zu Berlin*.

Da auch der Altbestand von herausragender Bedeutung ist, wurden bereits einzelne Alte Drucke digitalisiert und im „Stichwerkbrowser“ des Forschungsarchivs für Antike Plastik nachgewiesen (<http://www.arachne.uni-koeln.de/>). Ein größeres Digitalisierungsprojekt für den Bereich Klassische Archäologie und Ägyptologie ist mit den Beständen der Abteilung Rom in Planung, in Kooperation mit der *UB Heidelberg*.

Vernetzungen

Ziel des DAI ist die Vernetzung all dieser Materialdaten innerhalb und außerhalb des DAI. Eine Möglichkeit bieten hierfür die Werkzeuge des Semantischen Netzes, so ist eine konkrete Implementation des *CIDOC-CRM* (<http://cidoc.ics.forth.gr/>) geplant. Damit ergibt sich für das DAI mit seiner Vielzahl an Datenquellen die langfristige Perspektive eines integrativen Wissensmanagements, das die in einzelnen Projekten erzeugten Daten zusammenführt und nicht nur abfragbar, sondern auch kombinierbar und in neuartiger Form visualisierbar macht. Nicht nur, wie weiterhin sinnvoll und notwendig, in der Struktur der jeweils primären Datenquellen (*Arachne*, *ZENON-DAI*, *Stichwerkbrowser*), sondern darüber hinaus auf einer Ebene, die durch die Strukturen des *CIDOC-CRM* definiert ist und die damit nicht nur alle Daten, wo möglich, kombiniert, sondern auch in einem einheitlichen Muster erfahrbar macht: dies betrifft nicht mehr die gemeinsame Suche in verschiedenen Quellen, sondern etwa ein browsingbasiertes Verfahren, das sich an den Kontexten der jeweiligen Objekte und Informationen orientiert (s. die dafür langfristig notwendige Architektur in der Skizze).

Im Rahmen einer Kooperation des DAI mit der Tufts University, Boston, konnten erste Schritte für Teile der *Perseus Digital Library* (<http://www.perseus.org>) und für *Arachne* umgesetzt werden, so dass sich die aus beiden Informationssystemen extrahierten XML-Daten, via XSLT in RDF aufbereitet, im Longwell-Browser in facettiertem Browsing betrachten und navigieren lassen. Damit ist eine erste konkrete Implementation erreicht.

In Planung befindet sich die Ausweitung der Referenzierung von der Welt der Objektdaten auf die Welt der Buchdaten, z.B. durch die Verknüpfung von Digitalisaten mit Abbildungen antiker Statuen mit den Objektdaten in der Photodatenbank, wie dem Photo der echten Statue. Ebenso in Zusammenarbeit mit der *Perseus Digital Library* soll die Anwendung von *FRBR* (functional requirements for bibliographic records, see <http://www.ifla.org/VII/s13/frbr/frbr.pdf>) getestet werden, die diese bereits praktizieren.

Weitere Vernetzungen mit den anderen archäologischen Daten sind angestrebt, wie den Grabungs- und Archivdaten. Man arbeitet also weiter an der *Archäologie im Semantischen Netz*.

Jürgen Keiper, Werner Surendorf

Reconstructing Film Culture: Lost Films and Collaboration on the Web

Zusammenfassung

Mehr als zwei Drittel aller einmal produzierten und aufgeführten Stummfilme sind verschollen und auch zahlreiche Werke des Tonfilms gelten als verloren. Damit ist ein wesentlicher Teil des visuellen Erbes für das kulturelle Gedächtnis nicht mehr verfügbar. Filmgeschichte aber nur von den überlieferten Filmen her verstehen zu wollen, verwischt die Bedeutung des Mediums für visuelle Traditionen und kulturelle Kontexte

LOST FILMS beginnt eine Rekonstruktion dieses Erbes an ausgewählten Beispielen der deutschen Filmproduktion. Hierzu wurden zwei Strategien gewählt. Einerseits recherchieren wichtige europäische Kinematheken in ihren Beständen und tragen unterschiedlichste Materialien (Photos, Drehbücher etc.) aus ihren Archiven zu diesen Filmen zusammen, andererseits werden diese Dokumente auf einer kollaborativ organisierten Plattform im Web veröffentlicht, um damit eine verloren geglaubte Filmgeschichte öffentlich sichtbar zu machen.

Perspektivisch werden weitere Funktionen entwickelt, welche die gemeinsame Identifikation von verschwundenen Filmen unterstützen. Dies Projekt wurde initiiert von der *Deutschen Kinemathek – Museum für Film und Fernsehen (Berlin)*, wird gefördert durch die Kulturstiftung des Bundes und in Kooperation mit dem *Bundesarchiv – Filmarchiv (Berlin)*, der *Friedrich-Wilhelm-Murnau-Stiftung (Wiesbaden)*, dem *Filmarchiv Austria (Wien)*, dem *Centre national de la cinématographie (Paris)* und dem *Národní filmový archiv (Prag)* durchgeführt.

Abstract

More than two thirds of all silent films have disappeared. A large number of motion pictures with sound, too, are considered lost. A considerable portion of our visual heritage, central to our collective memory, is no longer at our disposal. Understanding film history solely in terms of the films available to us today obscures the full importance of the medium to visual traditions and cultural contexts.

Focusing on specific examples from German film production, *LOST FILMS* has begun a reconstruction of this heritage. We have chosen two strategies. First, important European film museums are bringing together diverse sources (photographs, scripts, etc.) from their respective archives. Second, these materials are published on a web platform. Organised as a collaborative venture, the web platform raises the visibility of a history many believe to have been lost. Additional steps are planned to support the joint identification of vanished films.

This project was initiated by the *Deutsche Kinematik – Museum for Film and Television* in Berlin. The project receives support from the German Federal Cultural Foundation and is carried out in co-operation with the Film Archive of the *German Federal Archives* in Berlin, the *Friedrich Wilhelm Murnau Foundation* in Wiesbaden, the *Austrian Film Archive* in Vienna, the *Centre national de la cinématographie in Paris* and the *Národní filmový archiv* in Prague.

LOST FILMS is a project of the *Deutsche Kinematik – Museum for Film and Television* in co-operation with the *German Federal Archives – Film Archives*, the *Centre national de la*

cinématographie, the *Austrian Film Archive*, the *Friedrich Wilhelm Murnau Foundation* and the *Czech National Film Archive*. The project receives funding from the *German Federal Cultural Foundation*.

The situation is dramatic: in contrast to other media, a large percentage of the twentieth-century's audio-visual works no longer exist. Silent film is particularly impacted. According to estimates, circa 80% of the silent films produced are no longer available. By implication, then, that what we today know as film history is in fact only a small portion of what was actually produced.

This loss of knowledge regarding our audio-visual inheritance pertains not only to silent film. Albeit not to the same degree, some films from the post-war era are also no longer at our disposal. This state-of-affairs often remains unknown, with these circumstances only suddenly emerging in the course of scholarly research or retrospectives devoted to the work of better-known directors.

The reasons for the losses are diverse; a critical appraisal of these reasons would take us well beyond the scope of this presentation. The reasons extend from the physical instability of the material to deliberate destruction to the extraction of raw materials to a misguided archival policy on the part of the film-producing countries. More recent initiatives, such as a mandatory deposit for film productions, are becoming more important in efforts to prevent exactly this type of loss occurring in the future. These efforts are nonetheless counteracted by the development toward digital films. If no satisfactory solutions emerge to address the issues surrounding long-term preservations, we will soon be threatened by a second wave of loss analogous to that suffered by silent films.

Irrespective of preventative measures, film archives in particular must struggle with the fact that a large portion of the materials they seek to collect simply no longer exist. Their task thus becomes to develop strategies to cope with this invisible history.



Image 1: *Das Mädchen ohne Vaterland*, D 1912, Regie: Urban Gad

"Missing Believed Lost" is the title David Meeker chose for what has become a standard work on the problem of lost films. Commissioned in 1992 by the British Film Institute, Meeker derived the title for his book from the military discourse of the Second World War. The phrase describes the situation of those war pilots who did not return from missions; or at least initially, as from time to time one of the pilots shot down would re-appear after a few days or weeks. The situation is similar with film immediately after the publication of Meeker's book. Numerous "lost films" suddenly materialised in the archives, particularly in international film archives. Private collectors, too, reported "lost films" in their holdings as well.

David Meeker's publication was admittedly a drop in the ocean, though his book did shine on a light on the state of archival work at the beginning of globalisation. True, archives have always co-operated across national boundaries, though with the development of networked structures, open databases and transparent archival policies a new situation is emerging. Now, at least, one will someday know with greater certainty about the remainder of films labelled "missing believed lost". While film archives are not quite there yet, new technologies enable important forms of collaboration that can be of terrific value for the subject of lost films.

The initial starting point for our "LOST FILMS" project is research on lost films. From an outsider's perspective, it seems obvious to define first the film productions of a particular era and then to check existing archival collections to confirm their existence. What would

then remain would be a list of missing films. However, this approach has little to do with reality. An accessible catalogue of all films produced in the form of a database does not exist. Nor, for that matter, is there an authoritative, electronic inventory of the relevant archives. And the willingness of archives to present their collections in full transparency is uncommon. The method of choice thus remains to search for films known to be lost, or, to put it somewhat bureaucratically, the development of a positive list of lost films. This approach will not be systematic. Among those films that were historically important, which are considered missing today? Looking backward across time, which films are of special interest to us today that played no special role in the lives of contemporaries? Answering these questions means risking many philological pitfalls, leading ultimately to a strongly subjective approach. For the research we have solicited the opinions of notable archivists, scholars of film and journalists, experts in a position to report on important lost films due to their work. This approach yields a patchwork of the most diverse preferences, but at least one avoids a commitment to one, ultimately dubious, approach. Regardless of what's in the foreground, be it gender, historical preferences, aesthetic innovation or recourse to the great authors of film history, all is blurred through this rather pragmatic approach.



Image 2: Zapatas Bande,
D 1913/14, Regie: Urban Gad

The goal of this project is not a systematic comparison or the development of a list of all missing films. In five or ten years, such a list will emerge as a side effect of digital collection, realised initially in the form of a wiki. At the centre of this project is rather the idea of an interlinked, virtual reconstruction of unaccounted for films.

If we no longer possess the films, the archives can at least make public the materials they do hold regarding missing motion pictures. These include photos and posters, newspaper accounts of the premieres, even entire scripts. With these materials we cannot, of course, "reconstruct" films in the classical sense, though by means of these documents we can say something about these films, offering researchers and the public an opportunity to grapple with these motion pictures. It's entirely possible that in the process film history will be written anew, as we perforce bow to the primacy of what today's actually viewable.

Making archival collections useful for this "virtual reconstruction" is one aspect of this project. Another lies in opening archives' institutional borders through technical collaboration. The willingness of archives to take this path exists. Still missing, however, are applicable technical concepts.¹

Today that has changed. At the core of this project is a database in which digitised documents related to missing films have been integrated. Needless to say, this database is web-based. This means that the participating archives can incorporate and describe individual documents. Dispatching digitised documents per e-mail or DVD is no longer necessary. The collaborative element of this project is not therewith exhausted. The participating institutions can also add remarks on the documents, thereby starting a collaborative annotation process. More important, in my view, is that the collaborative element not only comprises a result of the project, but also how it is funded.

The database is an open source project and is subject to a free general public license. This means that the project partners can continue, free of charge, to use and develop the software upon conclusion of this project. This decision makes it possible to avoid one of the problems of project-specific funding, enabling long-term participation in the project's results.

¹ See Jürgen Keiper, Holger Brocks, Andrea Dirsch-Weigand, Adelheit Stein, Ulrich Thiel: COLLATE – A Web-Based Collaboratory for Content-Based Access to and Work with Digitized Cultural Material. ICHIM (1) 2001: 495-511.

The database used for this project is produced by OpenCollection, first realised by this US firm on behalf of the Museum of the Moving Image in New York City. To be exact, it's not merely one database, but rather a complete system to administer archival objects. Characteristic of this approach is recourse to open standards, including support for the *Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)*, enabling cross-linking of the database in larger configurations and other contingencies.

Further, the database offers the opportunity to incorporate the most diverse types of objects: photographs and sound are just as integrated as PDFs or video files. The database thus makes it possible to present fragments of a film, a feature which helps to dispel cases of doubt. Often enough, it's not clear whether a film fragment at all belongs to a particular film. Resolving this issue can now be transferred to a community of experts.

This example refers us back to the project's starting point. If only a few meters of a feature film are available, can one nonetheless speak of a missing film? What's our judgement if the film has only been passed down to us in an abridged version, or only a black-and-white version exists of what was a colour production? May one in these cases speak of lost films? Unintentionally, the subject of lost films leads us to take a closer look at the generally inadequate situation vis-à-vis the historical transmission of audio-visual materials.

Already at this point in the project one can say that we are making an important contribution to European culture's "intangible heritage". For example, many early films of the famous silent film actress and producer Asta Nielsen are lost. At the *Deutsche Kinemathek*, however, we do have many of Nielsen's personal writings and photos, which at least convey a visual impression of her provocative role play (Nielsen was famous for her unconventional representations of bourgeois roles). This use of photography makes clear the special importance of photos in drawing inferences about lost films.

With all this elation regarding collaborative forms of endeavour – made possible by new technologies and allowing completely new forms of targeted labour – allow me to offer a critical comment.

The possibility to collaborate on this technical level appears to me more than opportune. Internet technology produces tangible project results that either would not have been possible or only possible with a great deal of effort. Nevertheless, there is something inherently technocratic about these approaches, as they tend to manufacture inherent reason, in effect reifying down-to-earth circumstances. These tools are not yet conceived to describe or at least adequately replace the dynamics of social processes, or even work processes. Question of social recognition and authorship are in this context ignored. The success of such large *Web2.0* portals as *Facebook* or *MySpace* underscore that this is possible. We need to tie into this approach the dimension of social recognition if we, over the long term, wish to establish collaboration and "knowledge by networking" as more than a label.

Jürgen Keiper, Werner Surendorf

Verlorene Filme und ihre kollaborative Rekonstruktion

LOST FILMS ist ein Projekt der *Deutsche Kinemathek – Museum für Film und Fernsehen* in Zusammenarbeit mit dem *Bundesarchiv – Filmarchiv*, dem *CNC – Centre national de la cinématographie*, dem *Filmarchiv Austria*, der *Friedrich-Wilhelm-Murnau-Stiftung* und dem *Národní Filmový Archiv*. Das Projekt wird von der *Kulturstiftung des Bundes* gefördert.

Die Situation ist dramatisch: Im Unterschied anderen Medien sind ein Großteil der audiovisuellen Werke des 20ten Jahrhunderts nicht mehr vorhanden. Besonders der Stummfilm ist davon betroffen. Man schätzt, dass ca. 80% der produzierten Stummfilme nicht mehr zugänglich sind. Dies bedeutet aber auch im Umkehrschluss, dass das, was wir heute als Filmgeschichte kennen, nur ein kleiner Ausschnitt der tatsächlichen Produktion ist. Doch der Verlust der Kenntnis des audiovisuellen Erbes betrifft nicht nur den Stummfilm. Obgleich weniger umfassend, sind auch Filme der Nachkriegszeit heute nicht mehr auffindbar. Diese Tatsache bleibt oft lange im Dunkeln. Erst bei wissenschaftlichen Recherchen oder bei den Retrospektiven bekannter Regisseure tritt dieser Umstand plötzlich zutage.

Die Gründe für den Verlust sind vielfältig und ihre Aufarbeitung würde diesen Beitrag sprengen. Sie reichen von der physischen Instabilität des Materials über gezielte Vernichtung zur Rohstoffgewinnung bis zu einer verfehlten Archivierungspolitik seitens der Filmproduktionsländer. Erst in neuester Zeit gewinnen präventive Strategien wie die Initiative zur Pflichtabgabe für Filmproduktionen an Bedeutung, um genau diesen Verlust in Zukunft zu verhindern. Konterkariert werden diese Initiativen durch den technologischen Wandel hin zu digitalen Filmen. Sollten sich hier keine befriedigenden Lösungen im Sinne einer Langzeitarchivierung finden lassen, droht eine zweite Welle des Verlustes analog zu den Stummfilmen.

Unabhängig von den präventiven Maßnahmen haben insbesondere die Filmarchive mit der Tatsache zu kämpfen, dass ein Großteil ihrer eigentlichen Sammlungsobjekte schlicht nicht mehr existiert und Strategien für den Umgang mit dieser unsichtbaren Geschichte zu entwickeln.



Bild 1: Das Mädchen ohne Vaterland,
D 1912, Regie: Urban Gad

„Missing believed Lost“ betitelt David Meeker ein Standardwerk zur Problematik der verlorenen Filme, welches er im Auftrag des BFI im Jahre 1992 publizierte. Der Ausdruck entspringt dem militärischen Diskurs des zweiten Weltkrieges und beschrieb die Situation, dass Kampfflieger von ihrem Einsatz nicht mehr zurückkamen. Zunächst, denn mitunter kam es doch vor, dass die abgeschossenen Piloten nach ein paar Tagen oder Wochen wieder auftauchten. Ähnlich verhielt es sich mit den Filmen, gerade nach der Publikation von Meekers Buch. Zahlreiche „verlorene Filme“ tauchten plötzlich in den Archiven wieder auf, insbesondere in den internationalen Filmarchiven, aber auch Privatsammler meldeten sich und konnten „lost films“ in ihren Archiven benennen.

Die Publikation von David Meeker war freilich nur ein Tropfen auf ein den heißen Stein, doch beleuchtete sie die Situation archivarischer Arbeit zu Beginn der Globalisierung. Zwar haben Archiv schon immer auf internationaler Ebene kooperiert, doch erst mit dem Aufbau vernetzter Strukturen, offener Datenbanken und einer transparenten Archivpolitik wird man zumindest Gewissheit über jenen Restbestand an Filmen erhalten, die als Missing believed Lost gelten. Davon sind die Filmarchive immer noch entfernt, doch haben die neuen Technologien schon im Vorfeld wichtige Formen der Kollaboration eröffnet, die sich wunderbar für das Feld der verlorenen Filme nutzen lassen.

Ausgangspunkt für das Projekt „*LOST FILMS*“, ist zunächst eine Recherche nach verlorenen Filmen. Von außen stehend erscheint es nahe liegend, zunächst die Filmproduktion einer Epoche zu definieren und dann einfach die existierenden Bestände in den Archiven davon zu streichen. Übrig bliebe eine Liste der vermissten Filme. Doch dieses Modell hat wenig mit der Wirklichkeit gemein. Weder gibt es ein zugängliches Verzeichnis aller produzierten Filme in Form einer Datenbank, noch gibt es zuverlässige, elektronische Bestandsverzeichnisse der Archive selbst. Auch die Bereitschaft der Archive, ihre Bestände so transparent zu präsentieren, ist nur vereinzelt anzutreffen. So bleibt methodisch die Recherche nach definitiv verlorenen Filmen das Mittel der Wahl oder, etwas bürokratischer ausgedrückt, der Aufbau einer Positivliste vermisster Filme. Diese wird nicht systematisch sein, dazu fehlen den Archiven schlicht die Voraussetzungen, sondern inhaltlichen Kriterien folgen. Was waren Filme, die historisch besonders wichtig waren und heute als verschollen gelten? Was sind Filme, an denen wir retrospektiv ein besonderes Interesse besitzen, ohne dass sie zeitgenössisch eine besondere Rolle gespielt hätten? Diese Fragen und ihre Beantwortung stecken voller philologischer Fallstricke und führen letztlich zu einem stark subjektiv geprägten Ansatz. Für die Recherche werden namhafte Archivare, Filmwissenschaftler und Journalisten befragt, die über bedeutende, verlorene Filme im Rahmen ihrer Arbeit berichten können. So ergibt sich ein Patchwork unterschiedlichster Präferenzen, das jedoch zumindest eine Festlegung auf einen einzigen und damit letztlich fragwürdigen Ansatz vermeidet. Denn ob nun gender oder historische Präferenzen im Vordergrund stehen, ob ästhetische Innovation oder der Rekurs auf die großen Autoren der Filmgeschichte in den Vordergrund gestellt wird, verwischt sich bei diesem eher pragmatischen Ansatz.



Bild 2: Zapatas Bande,
D 1913/14, Regie: Urban Gad

Auch ist das Ziel dieses Projektes nicht der systematische Abgleich und der Aufbau einer Liste aller vermissten Filme. Diese wird sich in 5 oder 10 Jahren als Nebeneffekt der digitalen Erfassung einstellen und wird hier zunächst in Gestalt eines Wikis realisiert. Im Zentrum steht vielmehr die Idee einer vernetzten, virtuellen Rekonstruktion vermisster Filme. Wenn wir die Filme schon nicht mehr besitzen, so können die Archive zumindest ihre Materialien zu den Filmen öffentlich machen. Es existieren Fotos und Plakate, Zeitungsberichte der Erstaufführung, ja sogar ganze Drehbücher dieser Filme. Mit diesen Materialien können zwar nicht die Filme im klassischen Sinne rekonstruieren, aber wir können vermittels dieser Dokumente etwas über diese Filme erzählen und den Forschern und dem Publikum Gelegenheit geben, sich mit der Geschichte dieser Filme auseinanderzusetzen. Gut möglich, dass dabei Filmgeschichte noch einmal anders geschrieben wird, da diese ja – notgedrungen – dem Primat des Sichtbaren folgte.

Die Sammlungen der Archive nutzbar zu machen für diese „virtuelle Rekonstruktion“ ist die eine Seite dieses Projektes. Die andere besteht darin, die institutionellen Grenzen von Archiven über eine technisch umgesetzte Kollaboration zu öffnen. Die Bereitschaft der Archive, diesen Weg zu gehen, ist da. Noch fehlte es allerdings an den technischen Konzepten hierfür.¹

Dies hat sich heute geändert. Den Kern des Projektes stellt eine Datenbank dar, in welche digitalisierte Dokumente zu diesen vermissten Filmen eingebunden werden. Diese Datenbank ist – natürlich – webbasiert. Dies bedeutet, dass die beteiligten Archive individuell Dokumente einbinden und beschreiben können. Ein Versand digitalisierter Dokumente per Email oder DVD ist nicht mehr notwendig.

¹ Vgl. auch: Jürgen Keiper, Holger Brocks, Andrea Dirsch-Weigand, Adelheit Stein, Ulrich Thiel: COLLATE – A Web-Based Collaboratory for Content-Based Access to and Work with Digitized Cultural Material. ICHIM (1) 2001: S.495-511.

Damit aber erschöpft sich nicht das kollaborative Moment. Die beteiligten Institutionen können ebenfalls Anmerkungen zu den Dokumenten machen und damit einen kollaborativen Annotationsprozess in Gang setzen. Wichtiger aber noch erscheint mir, dass das Moment der Kollaboration nicht nur das Ergebnis dieses Projektes umfasst, sondern auch die bereitgestellten Mittel.

Die Datenbank selbst ist eine OpenSource-Entwicklung und unterliegt der freien Lizenz GPL. Dies bedeutet, dass die Projektpartner diese Software auch nach Projektabschluss frei benutzen und weiterentwickeln dürfen. Damit ist dann auch eine Grundlage gelegt worden, die Problematik projektspezifischer Förderungen zu umgehen und langfristig an den Projektergebnissen partizipieren zu können.

Die Datenbank, welche hier für dieses Projekt benutzt wird, firmiert als OpenCollection und stellt eine Entwicklung dar, die von einer Firma in den USA im Auftrage des *Museum of the Moving Image* (New York) realisiert wurde. Genauer gesagt handelt es sich auch nicht nur um eine Datenbank, sondern um ein komplettes System zur Verwaltung von Archivobjekten. Kennzeichnend auch hier der Rekurs auf offene Standards, die Unterstützung für *OAI-PMH*², also die Vernetzung der Datenbank in größere Zusammenhänge. Außerdem bietet die Datenbank die Möglichkeit, unterschiedlichste Objekttypen zu integrieren: Bilder und Töne werden genauso integriert wie PDF- oder Videodateien. Gerade letzteres ermöglicht eben auch, Fragmente eines Filmes zu präsentieren. Damit lassen sich auch Zweifelsfälle klären. Oft ist es nicht klar, ob ein Filmfragment überhaupt zu einem bestimmten Film gehört. Diese Aufgabe kann damit an eine Fachcommunity übertragen werden.

Dies Beispiel verweist wieder zurück zu dem Ausgangspunkt des Projektes. Sind nur noch ein paar Meter eines Spielfilmes vorhanden, wird man trotzdem von einem verschollenen Film sprechen können? Aber was ist, wenn der Film nur in einer gekürzten Fassung überliefert ist oder von einem Farbfilm nur noch eine Schwarz/Weiß-Kopie existiert. Wird man hier von Lost films sprechen? Das Feld der Lost-Films wirft so unfreiwillig auch einen Blick auf die generell unzureichende Überlieferungssituation audiovisueller Materialien.

Schon jetzt kann man sagen, dass dieses Projekt einen wichtigen Beitrag zum „intangible heritage“ der europäischen Kultur beiträgt. So sind beispielsweise viele frühe Filme der berühmten Stummfilmdarstellerin und -produzentin Asta Nielsen verschollen. Von Nielsen, die bekannt war für ihre unkonventionellen Repräsentationen bürgerlicher Rollen, sind aber beispielsweise in der Deutschen Kinemathek zahlreiche persönliche Bücher und Photos überliefert, die zumindest einen visuellen Eindruck ihrer provokativen Rollenspiele bewahren. Damit wird auch deutlich, dass gerade den Fotos ein besonderer Stellenwert zukommt. Sie sind noch ehesten geeignet, Rückschlüsse auf einen verlorenen Film zuzulassen.

Bei aller Euphorie über die kollaborativen Arbeitsformen, welche diese Neuen Technologien ermöglichen und damit völlig neue Formen des zielorientierten Arbeitens ermöglichen, sei aber auch ein kritisches Wort angefügt.

Die Möglichkeit zur Kollaboration auf dieser technischen Ebene scheint mehr als vernünftig sein. Sie ermöglicht Arbeitsergebnisse – in diesem Fall zur Dokumentation europäischer Kultur – die ohne diese Hilfsmittel nicht oder nur mit sehr viel mehr Aufwand möglich wären. Trotzdem tragen diese Ansätze auch ein technokratisches Moment in sich, weil sie dazu neigen, die inhärente Vernunft von den realen Gegebenheiten zu abstrahieren und absolut zu setzen. Um die Dynamik von sozialen Prozessen, gar Arbeitsprozessen, abbilden oder zumindest zufrieden stellend ersetzen zu können, sind diese Tools noch nicht konzipiert. Fragen der sozialen Anerkennung und Autorenschaft sind hier noch ausgeklammert. Dass dies grundsätzlich möglich ist, zeigen sie Erfolge der großen *Web2.0*-Portale wie *Facebook* oder *MySpace*. An die Dimension sozialer Anerkennung gilt es anzuknüpfen, um auch langfristige Kollaboration und *knowledge by networking* nicht nur als Etikett zu etablieren.

² Open Archives Initiative Protocol for Metadata Harvesting

Teresa Malo de Molina

Digitisation of the Historic Press at the Biblioteca Nacional de España

Zusammenfassung

Das digitale Zeitungsarchiv ist das erste Projekt beim Aufbau der Spanischen Digitalen Bibliothek. Die *Spanische Digitale Bibliothek* hat zum Ziel, das gesamte bibliographische Erbe Spaniens, das in der Nationalbibliothek verwahrt ist, über das Internet zugänglich zu machen.

Bei dieser ersten Phase schafft die Nationalbibliothek digitalen Zugang zu einer Sammlung der historischen Tagespresse, die aus 143 verschiedenen Blättern besteht, die zwischen 1772 und 1933 erschienen sind. Derzeit stehen etwa 500.000 digitalisierte Seiten der digitalen Nutzung online zur Verfügung. Dies ist jedoch nur der Anfang. Das Angebot wird schrittweise die gesamte historische Presse Spaniens abdecken. Dieses Jahr werden Publikationen des 17. und 18. Jahrhunderts aufgenommen, ebenso wie die wichtigsten Schlagzeilen des 19. Jahrhunderts.

Die digitale Sammlung wird Standards für die Erforschung und Befragung des historischen Pressewesens in Spanien setzen. Über die digitalisierten Texte hinaus bietet unser System Informationen über digitalen Zeitungsarchive und -sammlungen des Landes an und vereinfacht damit den Zugang zu Spaniens historischer Zeitungskultur, ein Wissensbestand, der bislang zu guten Teilen unerforscht ist.

Abstract

The Digital Newspaper Archive represents the outset of the *Hispanic Digital Library* project, the aim of which is the public dissemination via Internet of the Spanish Bibliographic Heritage preserved at the National Library.

In this first phase the public is offered access to the digital collection of historic Spanish press contained at the Library, initially composed of 143 titles, published between 1772 and 1933, which represents the possibility of consulting approximately 500,000 digitalised pages. This is just the beginning. The offer of titles will gradually be extended to cover the historic evolution of the Spanish press. During 2007 it is planned on progressively incorporating new publications edited in the 17th and 18th centuries, as well as incorporating press headlines that are essential for the 19th century study.

This digital collection has arisen with the intention of becoming a benchmark for research and consultation of the Spanish historic press. In addition to providing reading and consultation of texts, it offers information about the main digital newspaper archive collections, thus facilitating a knowledge and access to the Spanish newspaper archive heritage, which is today still partly unexplored.



- The Newspapers Digital Library is the beginning of the *Biblioteca Digital Hispánica*, which aims to disseminate through the Internet the Spanish Bibliographic Heritage preserved at the National Library
- In this first stage, public access to the Digital Library of Spanish Old Newspapers is offered, with 143 titles, edited between 1772 and 1933, which means to be able to access 500.000 digitised pages.

Goals

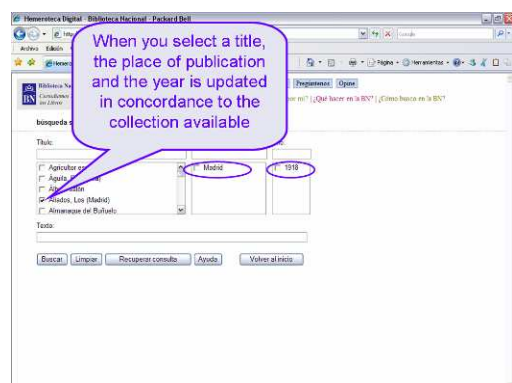
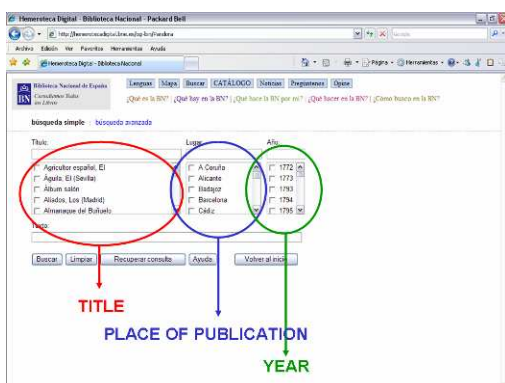
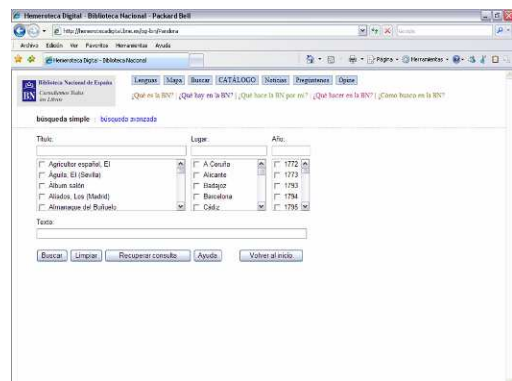
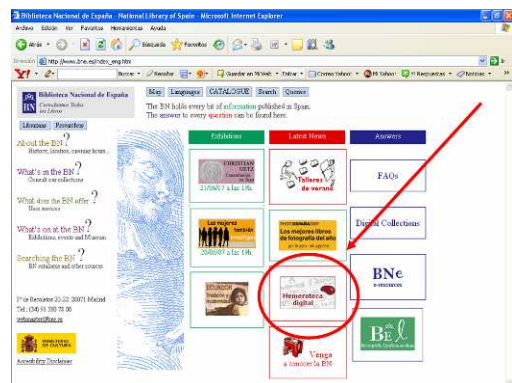
1. To offer public access to the Spanish National Library old newspapers collection
2. To become a reference to research and consultation of the Spanish press
3. To open cooperation channels with other centers to complete and find collections

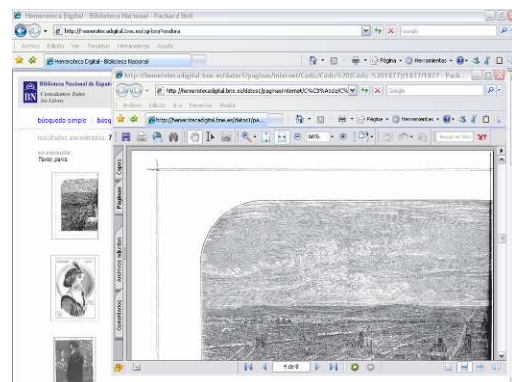
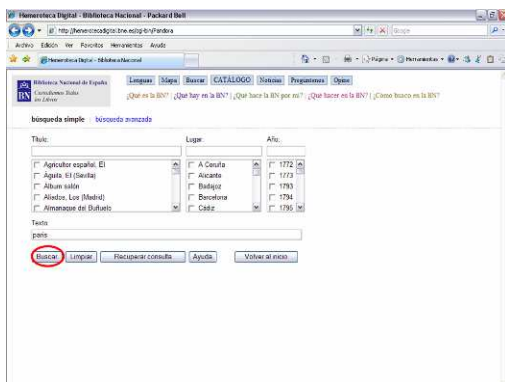
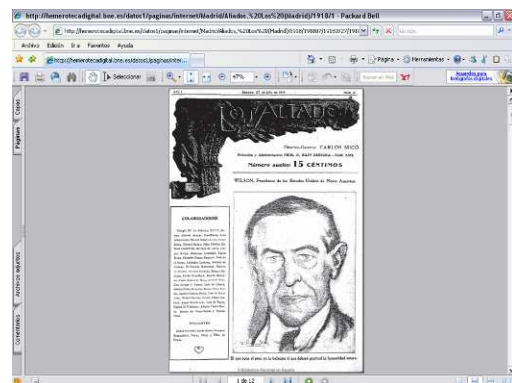
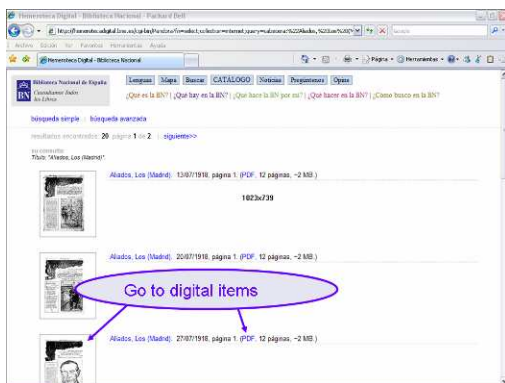
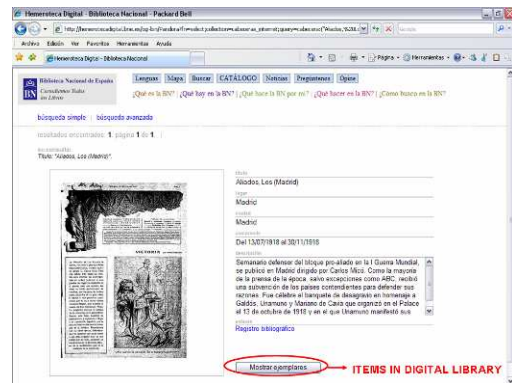
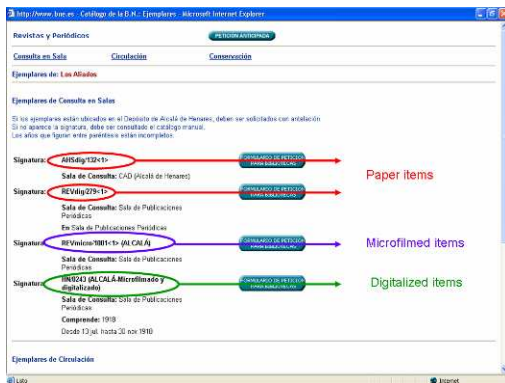
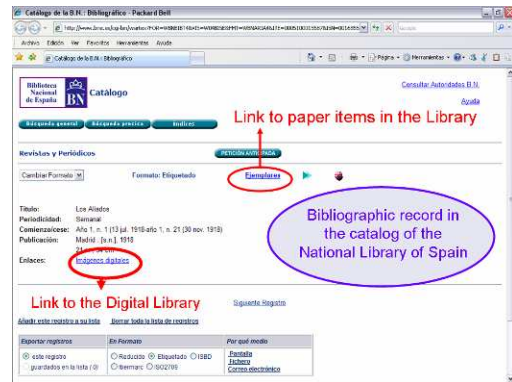
Technical Characteristics

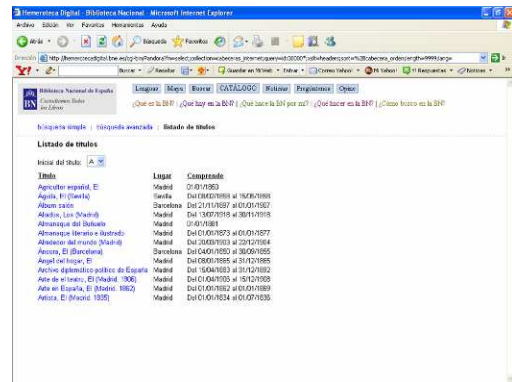
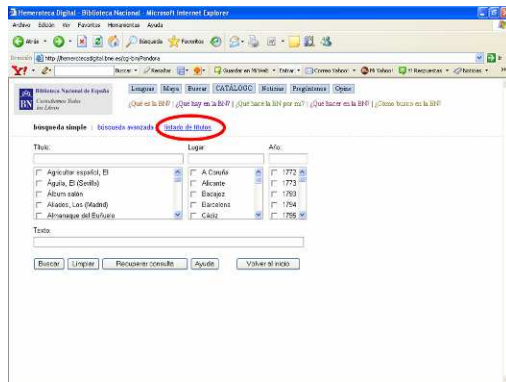
- Following the recommendations of UNESCO's World Memory Program :
 - Minimum resolution of 300 p.p.p.
 - According to the characteristics of publications: RGB colour, grey scales 256 levels or white and black
 - The files format is TIFF Compression CCITT Group IV to white and black and TIFF Compression JPEG to grey scales or colour.
 - OCR reading of each digitised page.
 - PDF files with hidden text and JBIG2 compression to disseminate images on the web
 - All PDF files will have miniatures

Old newspapers like historical resource

- Political Chronicle
- Biographic data and cultural activities
- Socioeconomic research
- Mean of transmission of literary text
- Technical and scientific evolution
- Study of the History of the press







Statistics

Statistics: average data per day

Month	Visitors	Volume	Pages	Files	Access
April	13.416	362,56 GB	92.579	471.546	836.791
May	14.423	351,45 GB	107.749	545.374	893.893
June (1-5)	5.489	112,54 GB	38.158	200.100	341.702
Total	33.328	826,56 GB	238.486	1.217.020	2.072.386

Month	Access	Files	Pages	Visitors
April	27.893	15.718	3.085	447
May	28.835	17.592	3.475	465
June (1-5)	22.089	13.022	2.458	345

Next future

- To increase as much as possible the number of digitized titles and digitized pages:
 - Between 2007-2008 we are going to invest more than 500.000 € to get more than 3 million of new digitized pages
- To integrate this collection – at least at title level – in the OAI Server of Hispanic Digital Library



Veröffentlichungen aus dem Institut für Museumsforschung

Zu beziehen durch: Institut für Museumsforschung, In der Halde 1, 14195 Berlin,
Tel.: (0 30) 8 30 14 60, Fax. (0 30) 8 30 15 04, e-mail: ifm@smb.spk-berlin.de
Vergriffene bzw. durch erweiterte Neuauflagen ersetzte Titel werden nicht mehr aufgeführt
(Stand November 2008)

Materialien aus dem Institut für Museumskunde

- Heft 8: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland samt Berlin (West) für das Jahr 1983*. Berlin 1984 (25 S.)
- Heft 10: *Eintrittsgeld und Besuchsentwicklung an Museen der Bundesrepublik Deutschland mit Berlin (West)*. Berlin 1984 (36 S.)
- Heft 14: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland samt Berlin (West) für das Jahr 1984*. Berlin 1985 (32 S.)
- Heft 16: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland samt Berlin (West) für das Jahr 1985*. Including an English Summary. Berlin 1986 (39 S.)
- Heft 17: *Gutachten zur Änderung der Öffnungszeiten an den Staatlichen Museen Stiftung Preußischer Kulturbesitz*. Erstellt von Hans-Joachim Klein. Berlin 1986 (77 S.)
- Heft 21: *Petra Schuck-Wersig, Martina Schneider und Gernot Wersig: Wirksamkeit öffentlichkeitsbezogener Maßnahmen für Museen und kulturelle Ausstellungen*. Berlin 1993 (119 S.), ISSN 0931-7961 Heft 21
- Heft 23: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland samt Berlin (West) für das Jahr 1987*. Including an English Summary. Berlin 1988 (46 S.), ISSN 0931-7961 Heft 23
- Heft 28: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland samt Berlin (West) für das Jahr 1988*. Including an English Summary. Berlin 1989 (56 S.), ISSN 0931-7961 Heft 28
- Heft 30: *Jane Sunderland, Lenore Sarasan: Was muß man alles tun um den Computer im Museum erfolgreich einzusetzen? Mit einer Einleitung von Christof Wolters*. Berlin 1989 (79 S.), ISSN 0931-7961 Heft 30
- Heft 31: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland samt Berlin (West) mit Besuchszahlenangaben zu den Museen der (ehemaligen) DDR für das Jahr 1989*. Berlin 1990 (64 S.), ISSN 0931-7961 Heft 31
- Heft 32: *Hans-Joachim Klein und Barbara Wüsthoff-Schäfer: Inszenierung an Museen und ihre Wirkung auf Besucher*. Karlsruhe 1990 (141 S.), ISSN 0931-7961 Heft 32
- Heft 33: *Christof Wolters: Wie muß man seine Daten formulieren bzw. strukturieren, damit ein Computer etwas Vernünftiges damit anfangen kann?* Berlin 1991 (133 S., 64 Abb.), ISSN 0931-7961 Heft 33
- Heft 34: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland für das Jahr 1990*. Berlin 1991 (80 S.), ISSN 0931-7961 Heft 34
- Heft 35: *Sigrid Heinze, Andreas Ludwig: Geschichtsvermittlung und Ausstellungsplanung in Heimatmuseen – eine empirische Studie in Berlin*. Berlin 1992. (234 S.), ISSN 0931-7961 Heft 35
- Heft 36: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland für das Jahr 1991*. Berlin 1992 (80 S.), ISSN 0931-7961 Heft 36
- Heft 37: *Petra Schuck-Wersig, Gernot Wersig: Museen und Marketing in Europa. Großstädtische Museen zwischen Administration und Markt*. Berlin 1992 (146 S.), ISSN 0931-7961 Heft 37
- Heft 38: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland für das Jahr 1992*. Berlin 1993 (96 S.), ISSN 0931-7961 Heft 38
- Heft 39: *Bibliographie-Report 1993 zu Museologie, Museumspädagogik und Museumsdidaktik und Besucherforschung*. Berlin 1993 (280 S.), ISSN 0931-7961 Heft 39
- Heft 40: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland für das Jahr 1993*. Berlin 1994 (104 S.), ISSN 0931-7961 Heft 40
- Heft 41: *Monika Hagedorn-Saupe, Annette Noschka-Roos: Museumspädagogik in Zahlen, Erhebungsjahr 1993*. Berlin 1994 (112 S.), ISSN 0931-7961 Heft 41
- Heft 42: *Alexander Geschke: Nutzung elektronischer Bilder im Museum*. Berlin 1995. ISSN 0931-7961 Heft 42
- Heft 43: *Erhebung der Besuchszahlen an den Museen der Bundesrepublik Deutschland für das Jahr 1994*. Berlin 1995 (104 S.), ISSN 0931-7961 Heft 43
- Heft 44: *Annette Noschka-Roos: Referierende Bibliographie zur Besucherforschung*. Berlin 1996 (96 S.), ISSN 0931-7961 Heft 44
- Heft 45: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 1995*. Berlin 1996 (104 S.), ISSN 0931-7961 Heft 45
- Heft 46: *Eintrittspreise von Museen und Ausgabeverhalten von Museumsbesuchern*. Berlin 1996 (145 S.), ISSN 0931-7961 Heft 46
- Heft 47: *Anne Claudel: Bibliographie zum Einsatz des Computers bei Sammlungsmanagement und -dokumentation*. Berlin 1997 (88 S.), ISSN 0931-7961 Heft 47

Heft 48: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 1996*. Berlin 1997 (96 S.), ISSN 0931-7961 Heft 48

Heft 49: *Angelika Costa, Mary Copple, Sebastian Fehrenbach, Bernhard Graf*: Besucherreaktionen zum Katalogverkauf in Ausstellungen. Beispiel: Sonderausstellung "Exil – Flucht und Emigration europäischer Künstler 1933 - 1945". Berlin 1998 (103 S.), ISSN 0931-7961 Heft 49

Heft 50: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 1997*. Berlin 1998 (104 S.), ISSN 0931-7961 Heft 50

Heft 51: *Anne Mikus*: Beispielhafte Konzepte für Museumseigene Publikationen. Produkte, deren Vertrieb und Vertriebspartner. Kurzfassung einer Studie der Staatlichen Museen zu Berlin Preußischer Kulturbesitz. 2. Aufl. Berlin 2000 (100 S.), ISSN 0931-7961 Heft 51

Heft 52: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 1998*. Berlin 1999 (100 S.), ISSN 0931-7961 Heft 52

Heft 53: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 1999*. Berlin 2000 (104 S.), ISSN 0931-7961 Heft 53

Heft 54: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 2000*. Berlin 2001 (104 S.), ISSN 0931-7961 Heft 54

Heft 55: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 2001*. Berlin 2002 (104 S.), ISSN 0931-7961 Heft 55

Heft 56: *Monika Hagedorn-Saupe, Henry Kleinke, Annett Meineke, Sabine Thänert*: Lange Nacht der Museen – eine empirische Untersuchung in Berlin. Berlin 2003 (96 S.), ISSN 0931-7961 Heft 56

Heft 57: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 2002*. Berlin 2003 (96 S.), ISSN 0931-7961 Heft 57

Heft 58: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 2003*. Berlin 2004 (96 S.), ISSN 0931-7961 Heft 58

Heft 59: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 2004*. Berlin 2005 (96 S.), ISSN 0931-7961 Heft 59

Materialien aus dem Institut für Museumsforschung

Heft 60: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 2005*. Berlin 2006 (96 S.), ISSN 0931-7961 Heft 60

Heft 61: *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 2006 Including an English Summary*. Berlin 2007 (96 S.), ISSN 0931-7961 Heft 61

Mitteilungen und Berichte aus dem Institut für Museumskunde

Nr. 1: *Christof Wolters*: Computereinsatz im Museum. Normen und Standards und ihr Preis. Berlin 1994 (38 S.)

Nr. 2: *Jochen Schmitt*: Rechtsfragen des Volontariats. Gutachten, erstattet im Auftrag der Stiftung Preußischer Kulturbesitz Berlin 1994 (24 S.)

Nr. 3: *Organisation und Kosten des Computereinsatzes bei Inventarisierung und Katalogisierung*. Workshop im Konrad-Zuse-Zentrum für Informationstechnik Berlin 18.-19. Oktober 1994. Berlin 1997 (48 S.)

Nr. 4: *Das Institut für Museumskunde der Staatlichen Museen zu Berlin – Preußischer Kulturbesitz*. Kurzdarstellung der Arbeit. Berlin 1995 (20 S.)

Nr. 5: *Monika Löcken*: Wissenschaftliche Volontariate an den Museen in der Bundesrepublik Deutschland. Berlin 1996 (30 S.)

Nr. 6: *Hans-H. Clemens, Christof Wolters*: Sammeln: Erforschen, Bewahren und Vermitteln – Das Sammlungsmanagement auf dem Weg vom Papier zum Computer. Berlin 1996 (75 S.)

Nr. 7: *Zusammenstellung von Eintrittspreisregelungen und Öffnungszeiten ausgewählter Museen in westeuropäischen Großstädten*. Berlin 1996 (48 S.)

Nr. 8: *Workshop zum Sammlungsmanagement – Berlin 29.10.1996*. Friedrich Waidacher: Vom redlichen Umgang mit Dingen – Sammlungsmanagement im System musealer Aufgaben und Ziele. Berlin 1997 (24 S.)

Nr. 9: *Réunion des organisateurs des grandes expositions*. Empfehlungen für die Organisation großer Ausstellungen. Berlin 1996 (34 S.)

Nr. 10: *Regine Scheffel*: Positionspapier zu Tätigkeitsbereich und Berufsbild in der Museumsdokumentation. Berlin 1997 (48 S.)

Nr. 11: *Monika Hagedorn-Saupe, Andrea Pohn*: Mögliche Veränderungen der Öffnungszeiten der Staatlichen Museen zu Berlin. Eine Besucherbefragung. Berlin 1997 (39 S.)

Nr. 13: *Petra Schuck-Wersig, Gernot Wersig, Andrea Pohn*: Multimedia-Anwendungen in Museen. Berlin 1998 (198 S.), ISSN 1436-4166 Nr. 13

Nr. 14: *Kunstmuseen und Urheberrecht in der Informationsgesellschaft*. Dokumentation einer Arbeitstagung der VG Bild-Kunst des Instituts für Museumskunde der Staatlichen Museen zu Berlin-PK und der Kulturstiftung der Länder am 12. Juni 1998. Berlin 1999 (90 S.), ISSN 1436-4166 Nr. 14

Nr. 15: *Friedrich Waidacher*: Museologische Grundlagen der Objektdokumentation. Berlin 1999 (24 S.), ISSN 1436-4166 Nr. 15

- Nr. 16: *Museumsberatung als Beruf? Berliner Herbsttreffen zur Museumsdokumentation. Workshop am 27. Oktober 1998.* Jim Blackaby, Richard Light, John Will. Berlin 2000 (50 S.), ISSN 1436-4166 Nr. 15
- Nr. 17: *Annett Rymarcewicz: Gesundheitsaufklärung in Ausstellungen – ein Besucherforschungsprojekt am Deutschen Hygiene-Museum.* Dresden. Berlin 1999 (35 S.), ISSN 1436-4166 Nr. 17
- Nr. 19: *W. Eckehart Spengler: Thesaurus zu Ackerbaugerät, Feldbestellung – Landwirtschaftliche Transport- und Nutzfahrzeuge – Werkzeuge (Holzbearbeitung), 2. unveränd. Auflage.* Berlin 2000 (92 S.), ISSN 1436-4166 Nr. 19
- Nr. 20: *Museumspädagogik in technischen Museen - Dokumentation des 1. Symposiums 14. bis 17. Juni 1999 in Berlin.* Berlin 2000 (74 S.) ISSN 1436-4166 Nr. 20
- Nr. 21: *Steffen Krestin: Impressionen einer internationalen Tagung – CIDOC 1997 in Nürnberg.* Berlin 2000 (52 S.) ISSN 1436-4166 Nr. 21
- Nr. 22: *Vorabdruck aus Heft 22: Ulrich Lange: Dokumentation aus der Sicht des Trainers.* Workshop Berlin 30.10.2000. Berlin 25/10/2000, ISSN 1436-4166 Nr. 22
- Nr. 23: *Akustische Führungen in Museen und Ausstellungen.* Bericht zur Fachtagung im Filmmuseum Berlin 2001. Berlin 2002 (80 S.), ISSN 1436-4166 Nr. 23 Dokumentation des 2. Symposiums 1. bis 2. Oktober 2000 in Mannheim. Berlin 2002 (44 S.), ISSN 1436-4166 Nr. 24
- Nr. 25: *Methodische Anregungen zu Umweltausstellungen.* Beiträge aus der Veranstaltung der Alfred Toepfer Akademie für Naturschutz (NNA), (67 S.), ISSN 1436-4166 Nr. 25
- Nr. 26: *Science Center, Technikmuseum, Öffentlichkeit. Workshop „Public Understanding of Science“ II. 3. Symposium „Museumspädagogik in technischen Museen“ vom 9. bis 12. September 2001 im Deutschen Museum, München.* (84 S.), ISSN 1436-4166 Nr. 26
- Nr. 28: *Hans Walter Hütter, Sophie Schulenburg: Museumsshops - ein Marketinginstrument von Museen.* Berlin 2004 (121 S.); ISSN 1436-4166 Nr. 28
- Nr. 29: *Anne-Katrin Wienick: Kultursponsoring – eine Untersuchung zur Zusammenarbeit von Berliner Museen und Unternehmen.* Berlin 2004, (153 S.), ISSN 1436-4166 Nr. 29
- Nr. 30: *Bernhard Graf, Astrid B. Müller (Hrsg.): Ausstellen von Kunst und Kulturen der Welt, Tagungsband.* Berlin 2005 (144 S.) ISSN 1436-4166 Nr. 30
- Nr. 31: *Regine Stein u.a.: Das CIDOC Conceptual Reference Model: Eine Hilfe für den Datenaustausch? Berlin 2005 (35 S.),* ISSN 1436-4166 Nr. 31
- Nr. 32: *Hanna Marie Ebert: Corporate Collections. Kunst als Kommunikationsinstrument in Unternehmen.* Berlin 2005 (S. 103), ISSN 1436-4166 Nr. 32
- Nr. 33: *Laura Wittgens: Besucherorientierung und Besucherbindung in Museen*
Eine empirische Untersuchung am Fallbeispiel der Akademie der SMB. Berlin 2005 (128 S.) ISSN 1436-4166 Nr. 33
- Nr. 34: *Petra Helck, Editha Schubert, Ellen Riewe: Absolventenbefragung des Studiengangs Museumskunde an der FHTW Berlin.* Berlin 2005 (51 S.), ISSN 1436-4166 Nr. 34
- Mitteilungen und Berichte aus dem Institut für Museumsforschung**
- Nr. 35: *Petra Schuck-Wersig, Gernot Wersig: Die Staatlichen Museen zu Berlin und ihre Besucher.* Zusammenfassung aus den Jahren 2001-2004. Berlin 2006 (S. 129) ISSN 1436-4166 Nr. 35
- Nr. 36: *Alexandra Donecker: Untersuchungen der Besucherresonanz zur Sonderausstellung „WeltSpielZeug“ im Ethnologischen Museum Berlin.* Berlin 2007 (S. 164) ISSN 1436-4166 Nr. 36
- Nr. 37: *Hannah Bröckers: Das Museum als Event: Museen in der Erlebnisgesellschaft.* Berlin 2007 (S. 94) ISSN 1436-4166 Nr. 37
- Nr. 38: *Stephan Schwan, Helmuth Trischler, Manfred Prenzel (Hrsg.): Lernen im Museum: Die Rolle der Medien.* Berlin 2006 (S. 163) ISSN 1436-4166 Nr. 38
- Nr. 39: *Corina Meyer: Museale Präsentation und Vermittlung von Kunstgewerbe - am Beispiel des Kunstgewerbemuseums Berlin.* Berlin 2007 (126 S.), ISSN 1436-4166 Nr. 39
- Nr. 40: *Charlotte Bristot: Marketing für Museen als systematischer Managementprozess.* Berlin 2007 (238 S.), ISSN 1436-4166 Nr. 40
- Nr. 41: *Brinda Sommer: Gesellschaftliches Erinnern an den Nationalsozialismus: Stolpersteine wider das Vergessen.* Berlin 2007 (137 S.), ISSN 1436-4166 Nr. 41
- Nr. 42: *Judith Kühnle: Professionalisierung der Öffentlichkeitsarbeit von Berliner Galerien.* Berlin 2007 (181 S.), ISSN 1436-4166 Nr. 42
- Nr. 43: *Denis Schäfer: PDA: Mobiles Informationssystem für die Besucherbetreuung im Museum.* Dokumentation und Diskussion ausgewählter Beispiele. Berlin 2008 (88 S.), ISSN 1436-4166 Nr. 43
- Nr. 44: *Claudia Wahl, Theresia Sager, Katja Leikam, Annika Opitz: Qualität im Museum.* Berlin 2008, (73 S.), ISSN 1436-4166 Nr. 44
- Nr. 45: *Paul Klimpel (Hrsg.): Im Schatten der Verwertungsinteressen – Filmarchive, Filmmuseen und das Urheberrecht - Juristisches Symposium der Deutschen Kinemathek Berlin 2007.* Berlin 2008, (82 S.), ISSN 1436-4166 Nr. 45

Nr. 46: *Monika Hagedorn-Saupe (Hrsg.): Wissen durch Vernetzung – Kulturgutdigitalisierung in Deutschland und Europa, Tagungsband – Berlin 2007 / Knowledge by Networking – Digitising Culture in Germany and Europe, Conference Proceedings – Berlin 2007. Berlin 2008 (144 S.), ISSN 1436-4166 Nr. 46*

**Materialien aus dem Institut für
Museumskunde
– Sonderhefte –**

Nr. 1: *Günther S. Hilbert: Vocabulary of Museum Security Terms. Berlin 2000, (284 S.)
ISSN 0931-4641 Sonderheft 1
Nur noch als Online-Katalog verfügbar unter
<http://elib.zib.de/museum/voc/>*

Nr. 2: *nestor/ Institut für Museumskunde: Nicht von Dauer – Kleiner Ratgeber für die Bewahrung digitaler Daten in Museen. Berlin 2004, (52 S.), ISSN 0931-4641 Sonderheft 2*

Nr. 3: *Monika Hagedorn-Saupe, Axel Ermert (Hrsg.): A Guide to European Museum Statistics. Berlin 2004, (203 S.) , ISSN 0931-4641 Sonderheft 3*

Berliner Schriften zur Museumskunde

**Zu beziehen über den Gebr. Mann Verlag,
Berlin**

Bd. 1: *Günther S. Hilbert: Sammlungsgut in Sicherheit. 3. vollständig überarbeitete und erweiterte Auflage. Berlin 2002, ISBN 3-7861-2348-9*

Bd. 2: *Hans-Joachim Klein, Monika Bachmayr: Museum und Öffentlichkeit. Fakten und Daten – Motive und Barrieren. Berlin 1981
ISBN 3-7861-1276-2*

Bd. 5: *Wolfger Pöhlmann: Handbuch zur Ausstellungspraxis von A – Z. Berlin 2007. Zweiter Aufl. in **Berliner Schriften zur Museumsforschung**. ISBN 978-3-7861-1692-9.*

**Zu beziehen über die GWV-Fachverlage,
Wiesbaden:**

Bd. 10: *Andreas Grote (Hrsg.): Macrocosmos in Microcosmo. Die Welt in der Stube. Zur Geschichte des Sammelns 1450 bis 1800. Opladen 1994. ISBN 3-8100-1048-0*

Bd. 11: *Annette Noschka-Roos: Besucherforschung und Didaktik. Ein museumspädagogisches Plädoyer. Opladen 1994, ISBN 3-8100-1049-9*

Bd. 12: *Anne Mikus: Firmenmuseen in der Bundesrepublik. Schnittstelle zwischen Kultur und Wirtschaft. Opladen 1997. ISBN 3-8100-1486-9*

Bd. 14: *Ralf-Dirk Henning : Digitalisierte Bilder im Museum. Technische Tendenzen und organisatorisches Umfeld. Opladen 1996. ISBN 3-8100-1483-4*

Bd. 15: *Petra Schuck-Wersig, Gernot Wersig: Museumsmarketing in den USA. 1999.
ISBN 3-8100-2078-8*

Bd. 16: *Gabriele König: Kinder- und Jugendmuseen. Genese und Entwicklung einer Museumsgattung. Impulse für besucherorientierte Museumskonzepte. Opladen 2002.
ISBN 3-8100-3299-9*

Bd. 17: *Kurt Winkler: Museum und Avantgarde. 2002. ISBN 3-8100-3504-1*

Bd. 18: *Susan Kamel: Wege zur Vermittlung von Religion in Berliner Museen. Black Kaaba meets White Cube. 2004.
ISBN 3-8100-4178-5*

Bd. 19: *Bernhard Graf, Astrid B. Müller (Hrsg.): Sichtweisen. Zur veränderten Wahrnehmung von Objekten in Museen. 2005.
ISBN 3-531-14489-8*

Bd. 20: *Volker Kirchberg: Gesellschaftliche Funktionen von Museen. Makro-, meso- und mikrosociologische Perspektiven. 2005,
ISBN 3-531-14406-5*

Zu beziehen über die G+H Verlage. Berlin

Bd. 21: *Annette Noschka Roos, Walter Hauser, Elisabeth Schepers (Hrsg.): Mit neuen Medien im Dialog mit den Besuchern? 2005
ISBN 3-931768-87-2*

Bd. 22: *Bernhard Graf, Hanno Möbius (Hrsg.): Zur Geschichte der Museen im 19. Jahrhundert 1789-1918. 2006,
ISBN 978-3-931768-928*

Berliner Schriften zur Museumsforschung

Bd. 23: *Heike Riebe: Benchmarking im Museum, ein Managementinstrument zur Qualitätssicherung, 2007, ISBN 978-3-931768-980*

Bd. 24: *Katharina von Chlebowski: Branchenkultur der Kunstmuseen in Deutschland, 2008,
ISBN 978-3-940939-03-6*

Handbuch des Museumsrechts (Einzelbände):

Bd. 1: *Irmgard Küfner-Schmitt: Arbeitsrecht. Opladen 1993. ISBN 3-8100-1018-9*

Bd. 2: *Christian Armbrüster: Privatversicherungsrecht. Opladen 1993. ISBN 3-8100-1008-1*

Bd. 3: *Jochen Laufersweiler, Andreas Schmidt-Rögnitz: Der Erwerb von Museumsgut. Opladen 1994. ISBN 3-8100-1080-4*

Bd. 4: *Rudolf Streinz: Internationaler Schutz von Museumsgut. Opladen 1998.
ISBN 3-8100-1174-6*

Bd. 5: *Gabriele Köhler-Fleischmann: Sozialrecht. Opladen 1994. ISBN 3-8100-1180-0*

Bd. 6: *Gerhard Pfennig: Digitale Bildverarbeitung und Urheberrecht. Eine Einführung in die Museumspraxis. Opladen 1998. ISBN 3-8100-2060-5*

Bd. 7: *Wilhelm Möhle (Hrsg.): Öffentliches Recht. Opladen 1998. ISBN 3-8100-2061-3*

Bd. 8: *Irmgard Küfner-Schmitt, Michael Kulka: Rechtliche Grundlagen der Privatisierung von Dienstleistungen im Museum. Opladen 1998.
ISBN 3-8100-2071-0*

Bd. 10: *Rudolf Gärtner: Versicherungsfragen im Museumsbereich. Opladen 2002.
ISBN 3-8100-3506-8*

ISSN 1436-4166 Nr. 46

S M
B Institut für Museumsforschung
Staatliche Museen
zu Berlin