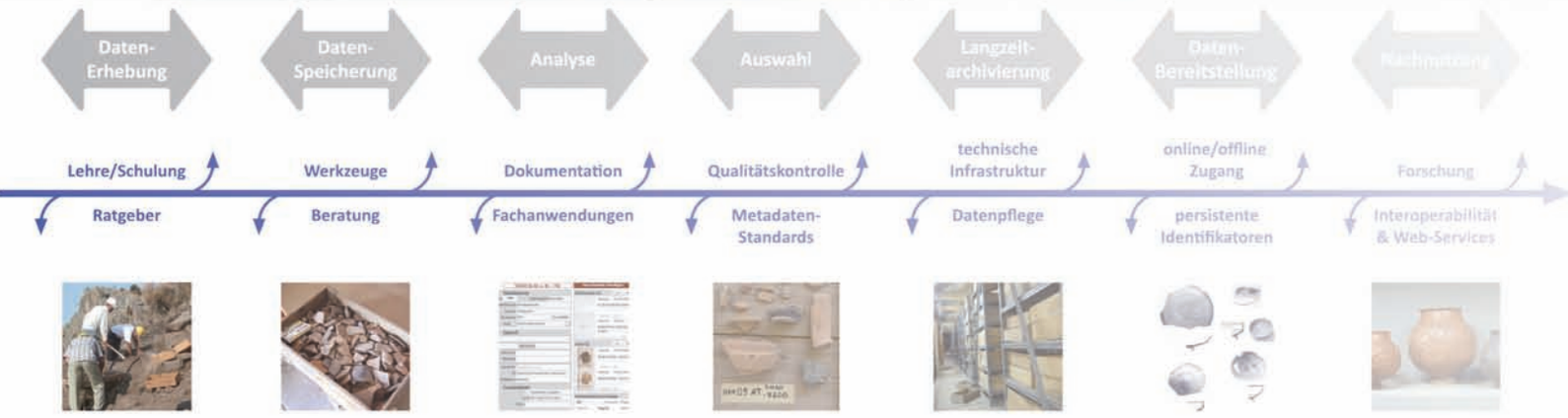
Digitale Archäologie



Analoge Archäologie

Archäologie und Informationssysteme

Niedersachsen

Niedersächsisches Landesamt für Denkmalpflege

Archäologie und Informationssysteme

Vom Umgang mit archäologischen Fachdaten in Denkmalpflege und Forschung

Arbeitshefte zur Denkmalpflege in Niedersachsen 42

Herausgeber:
Stefan Winghart

Redaktion der Reihe:
Dietmar Vonend

Wissenschaftliche Konzeption:
Otto Mathias Wilbertz

Textredaktion:
Josefine Puppe

Gestaltung, Bildredaktion und Satz:
Petra Götting

Gesamtredaktion:
Dietmar Vonend

Titelbild: Collage von Petra Götting

Bibliografische Information der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte bibliografische
Daten sind im Internet über http://dnb.ddb.de abrufbar.

Arbeitshefte zur Denkmalpflege in Niedersachsen 42
ISBN 978-3-8271-8042-1

© Niedersächsisches Landesamt für Denkmalpflege 2013
Nachdruck, auch auszugsweise, nur mit Genehmigung des Herausgebers
ISBN 978-3-8271-8042-1
Druck: Quedlinburg Druck GmbH
Vertrieb: CW Niemeyer Buchverlage GmbH, Hameln

Inhalt

Vorwort
Stefan Winghart 6

EINFÜHRUNG

Die Kommission „Archäologie und Informationssysteme“
Otto Mathias Wilbertz 8

ADeX® Version 2.0 – Standard für den Austausch archäologischer Fachdaten
Reiner Göldner, Ulrich Himmelmann, Thomas Poelmann, Axel Posluschny,
Thomas Richter, Mario Schlapke, Roland Wanninger, Bernd Weidner, Otto Mathias Wilbertz 10

ASPEKTE DER DATENBEREITSTELLUNG

**Freie Daten für freie Bürger –
Ein Essay über archäologische Daten, die Öffentlichkeit und open data**
Ulf Ickerodt 28

Über die Freiheit der Daten
Reiner Göldner 34

Opening access to heritage resources – risk, opportunity or paradigm shift?
Anthony Beck 40

Datenaustausch und Datenabgabe
Otto Mathias Wilbertz 48

THESAURUSFRAGEN

Semantische Interoperabilität
Reiner Göldner 50

Zur Arbeit mit Zeitstufen und Kulturgruppen – „reine Lehre“ und praktische Umsetzung
Otto Mathias Wilbertz 60

Thesauri im Fachinformationssystem des Bayerischen Landesamtes für Denkmalpflege
Markus Ullrich, Roland Wanninger 62

ARCHIVIERUNG UND DOKUMENTATION

Archäologische digitale Daten – Authentizität und Funktionalität
David Bibby, Reiner Göldner 70

Digital curation at the Archaeology Data Service
Catherine Hardman, Julian D. Richards 76

Erstellung von Bildern für Bildarchive
Sigmar Fitting 79

Metadaten in der Archäologie
Evelyn Rauchegger, Utz Böhner 82

Kleiner Ratgeber zur Archivierung digitaler Daten
Reiner Göldner, David Bibby, Andreas Brunn, Sigmar Fitting, Axel Posluschny, Bettina Stoll-Tucker 86

Der Einsatz des Web-Standards SVG in der Archäologie
Ralf Berger 105

archaeoDox – Standardsoftware auf Ausgrabungen in Schleswig-Holstein, der Freien und Hansestadt Hamburg und im Landkreis Harburg
Jörg Räther, Eicke Siegloff 110

PROJEKTE

Die Arbeitsgruppe „Archaeological Archives" des Europae Archaeologiae Consilium und das Culture 2007–2013 Project „ARCHES". Eine kurze Vorstellung.
David Bibby 115

IANUS. Die Konzeption eines nationalen Forschungsdatenzentrums für die Archäologie und die Altertumswissenschaften
Ortwin Dally, Friederike Fless, Reinhard Förtsch, Maurice Heinrich, Felix F. Schäfer 118

Der Workshop 2012 der Kommission Archäologie und Informationssysteme
Otto Mathias Wilbertz 128

Autorenverzeichnis 130

Vorwort

Computer prägen das tägliche Leben in einer Weise, die noch vor 20 Jahren undenkbar erschien. Archäologen mit Computerkenntnissen waren in den Siebziger- und Achtzigerjahren noch seltene Ausnahmen. Seit den Achtzigerjahren Jahren setzten sich in der Archäologie Textverarbeitungs- und Datenbankprogramme mehr und mehr durch und es folgte die weitere Verbreitung von Statistik- und Grafikprogrammen. Seit den Neunzigerjahren fanden zunehmend CAD-Programme und GIS-Systeme in der Bodendenkmalpflege Einsatz, und die rasante Entwicklung des Internets hat auch der Informationstechnologie in der Archäologie einen gewaltigen Schub verliehen. GIS-gestützte und oft Web-basierte digitale Fachinformationssysteme gehören mittlerweile zum Standard in der archäologischen Denkmalpflege und können oft schon mobil abgerufen werden.

Mit den neuen Möglichkeiten haben sich auch neue Chancen und Herausforderungen abgezeichnet. Digitale Inventare, Metadaten-Bestände und GIS-Technologie bieten die Möglichkeiten der flächenhaften Bearbeitung der bisher zumeist isoliert betrachteten Einzelobjekte. Web-basierte Nutzungsmöglichkeiten, Internetportale und digitale Nutzerführung liefern unterschiedliche Angebote zur Fachkommunikation, Vermittlung und Inwertsetzung.

In der föderal organisierten Bodendenkmalpflege in der Bundesrepublik sind es die Forschungsprojekte und vor allem die Landesgrenzen überschreitenden linearen Großprojekte, die deutlich machen, wie wichtig eine enge Zusammenarbeit zwischen den einzelnen Landesarchäologien ist – nicht zuletzt auch beim Austausch von Fachdaten.

Unter den Kolleginnen und Kollegen, die an den verschiedenen Landesämtern mit der Be- und Verarbeitung digitaler Fachdaten betraut sind, kam der Wunsch nach Austausch von Informationen und Erfahrungen sowie zur Vorbereitung von Standards beim Einsatz geografischer Informationssysteme auf.

Auf der Jahrestagung des Verbandes der Landesarchäologen in der Bundesrepublik Deutschland (VLA) am 3. Mai 2004 im niedersächsischen Bad Bederkesa wurde der Vorschlag, eine Kommission zu diesem Thema zu gründen, zunächst kontrovers diskutiert. Neben begeisterter Befürwortung wurde auch die Sinnhaftigkeit einer solchen Kommission infrage gestellt. Zu ambitioniert erschienen die Ziele. Schließlich wurde die Kommission „Archäologie und Informationssysteme“ mit großer Mehrheit ins Leben gerufen. Die Kommission trat erstmals unter Beteiligung von 15 Landesämtern und der Römisch-Germanischen Kommission am 16. März 2005 in Hannover zusammen und hat seitdem in ihrer mehrjährigen Arbeit die von einigen bei der Gründung geäußerten Zweifel durch eine beeindruckende Leistungsbilanz ausgeräumt. Die Gruppe hat – vor allem in ihren aus Experten zusammengesetzten Arbeitsgruppen – wichtige Impulse gegeben und eine informelle Vernetzung der mit Fragen der Informationstechnologie beschäftigten Personen in den Landesarchäologien der Bundesrepublik und ihrem breiten Partnerfeld verwirklicht. Sie dient als Plattform zum länderübergreifenden Austausch von Informationen und Erfahrungen sowie zur Vorbereitung von Standards. Themen sind die Organisation von Fachdaten und Geodaten, das Verhältnis zu Vermessungsbehörden und anderen Partnern mit digitalen Kartenbeständen, die Verfügbarkeit digitaler Karten, Zugriffsberechtigungen, Bereitstellungskosten, Vertragsgestaltung u.v.m. Die Kommission berät den Verband der Landesarchäologen, organisiert in Absprache mit dem Vorstand des VLA zum Beispiel Workshops zu aktuellen Themen, wie zum Beispiel zur Umsetzung der europäischen INSPIRE-Richtlinie zum Datenaustausch und spricht Empfehlungen aus. Der von der Kommission entwickelte Standard für den Austausch und die Weitergabe georeferenzierter archäologischer Daten ADeX (Archäologischer DateneXport) ist ein von allen Bundesländern mitgetragener Erfolg.

Es war der auch von außen an das Gremium herangetragene und vom Verband der Landesarchäologen und den Kommissionsmitgliedern geteilte Wunsch, ein Resümee aus den acht Jahren Kommissionsarbeit zu ziehen. Nachdem die ursprünglich geplante Publikation im Archäologischen Nachrichtenblatt nicht möglich war, bot sich der Druck in der Reihe der „Arbeitshefte zur Denkmalpflege in Niedersachsen“ an. Die Arbeit der Kommission hat in Niedersachsen

ihren Ausgang genommen, die meisten Tagungen und Sitzungen der Kommission und ihrer Arbeitsgemeinschaften fanden aufgrund der verkehrsgünstigen Lage in Hannover statt und der erste und vorletzte Sprecher der Kommission sowie Vertreter der Arbeitsgruppen arbeiten beim NLD. Wir fühlen uns dem Thema auch deshalb besonders verpflichtet, weil das Fachinformationsmanagement seit dem Beginn der Arbeit an der ADAB 1996 bis zum internetgestützten Fachinformationssystem ADABweb eine Schlüsselrolle in der Boden- und Baudenkmalpflege unseres Hauses spielt, die ihren Niederschlag auch in mehreren Auftritten auf der Hannoverschen Computermesse CeBit fand.

So waren wir gerne bereit, diesen wichtigen Sammelband zu publizieren und zu verbreiten. Das Buch ist gleichzeitig auch eine Geste des Dankes an den langjährigen Sprecher der Kommission, Dr. Otto Mathias Wilbertz, der sich auch nach seiner Pensionierung im NLD intensiv in der Kommissionsarbeit engagiert. Wir danken dem Kollegen Wilbertz für seine hervorragende Arbeit und wünschen seinem Nachfolger Dr. Reiner Göldner vom Landesamt für Archäologie Sachsen für die Fortsetzung der Arbeit alles Gute. Die Kommission ist auch weiterhin in unseren Räumen gerne gesehen.

Die Kommission hat nun aus acht Jahren erfolgreicher Arbeit wichtige Ergebnisse, Diskussionen und Anregungen aus ihrer Arbeit zusammengestellt. Diese werden ergänzt durch Beiträge externer Autoren zu sachverwandten Themen. So ist eine Gesamtschau der Situation der Landesarchäologie im Umgang mit der Informationstechnologie, ihren Chancen, aber auch bestehenden rechtlichen Rahmenbedingungen im Umgang mit den Daten entstanden.

Dr. Stefan Winghart,
Präsident des Niedersächsischen Landesamtes für Denkmalpflege

Dr. Henning Haßmann,
Niedersächsischer Landesarchäologe

EINFÜHRUNG

Die Kommission „Archäologie und Informationssysteme“

Otto Mathias Wilbertz

Der Verband der Landesarchäologen in der Bundesrepublik Deutschland (im Folgenden VLA) hat am 3. Mai 2004 auf seiner Jahrestagung im niedersächsischen Bad Bederkesa beschlossen, eine Kommission „Archäologie und Informationssysteme“ ins Leben zu rufen (Wilbertz 2008). Sprecher der Kommission war zunächst Henning Hassmann, auf den 2006 Gregor Schlicksbier folgte. Von diesem übernahm der Verfasser dieser Zeilen 2007 die Sprecherfunktion[1].

Zu den wesentlichen Aufgaben der Kommission gehört die Realisierung der Darstellbarkeit archäologischer Inhalte in digitalen Karten. Als Arbeitsziele der Kommission wurden bei der konstituierenden Sitzung benannt: Herstellung von Interoperabilität, Nutzung von Synergieeffekten, Einbindung in die Geodateninfrastruktur Deutschland (GDI-DE), Empfehlung zur Modellierung archäologischer Ressourcen (Begrifflichkeiten, Datenstrukturen, Thesauri), Schaffung eines bundesweiten Standards zum Datenaustausch, Öffentlichkeitswirksamkeit (innerhalb des Verbandes Machbarkeit und Möglichkeiten digitaler Kartierung aufzeigen, nach außen Ergebnisse präsentieren).

Bei der konstituierenden Sitzung am 16. März 2005 wurde vereinbart, die inhaltliche Arbeit in kleinere Arbeitsgruppen zu verlagern. Als deren erste trat am 2. Mai 2005 die „AG Modellierung“ erstmals zusammen. Es folgten die „AG Rechtliche und wirtschaftliche Aspekte der Datenweitergabe (RWAD)“ am 26. Juli 2007 und die „AG Thesaurusfragen“ am 27. September 2007. Im Rahmen der Tagung des VLA 2008 wurde ein Kolloquium mit dem Titel „Analog und digital – Probleme und Perspektiven der Archivierung und Magazinierung archäologischer Quellen“ durchgeführt (Archäologisches Nachrichtenblatt 14, 2009 [2] passim). Die Beschäftigung mit dieser Thematik führte zur Gründung der „AG Archivierung“, die sich am 16. September 2008 erstmals traf.

Bei den Sitzungen der Kommission steht eine Beteiligung möglichst vieler Landesämter im Vordergrund, sodass um einen Kern kontinuierlich Mitwirkender eine wechselnde Zahl weiterer Personen zusammenkommt. Zu der konstituierenden Sitzung hatten 14 Landesämter, die Stadtarchäologie Köln sowie die Römisch-Germanische Kommission des Deutschen Archäologischen Instituts (RGK) Vertreter entsandt. Bei der Sitzung am 28. November 2011 waren Vertreter von elf Landesämtern und der RGK sowie ein Vertreter der Kommunalarchäologien zugegen. Die Arbeitsgruppen der Kommission setzen sich dagegen aus einem konstanten Personenkreis zusammen, der eine kontinuierliche Arbeit gewährleistet.

Die AG Modellierung begann damit, Informationen über vorhandene Strukturen zu sammeln und die semantischen Grundlagen zu schaffen, um Daten vergleichen und schließlich austauschen oder gemeinsam präsentieren zu können. Hierzu wurde eine einfache Tabellenstruktur in einem von allen beherrschbaren Format gewählt sowie zwei einfache Wertelisten für die Ansprache von Denkmaltyp und Datierung. Für den so geschaffenen Austauschstandard ADeX (=„Archäologischer DateneXport“) wurde eine Wortbildmarke beim Deutschen Patent- und Markenamt eingetragen. Die erste Beispielkartierung zeigte Burgen aus sechs Denkmalämtern. Eine neue, überarbeitete Karte mit Burgen aus nunmehr acht Denkmalämtern und einer ersten chronologischen Differenzierung wurde 2011 von I. Herzog publiziert (Herzog 2011). Die bisherigen Ergebnisse wurden bei der Tagung des Verbandes der Landesarchäologen im Juni 2007 in Esslingen verbandsintern vorgestellt. Im November des gleichen Jahres wurden die Ergebnisse im Rahmen der Internationalen Tagung „Kulturelles Erbe und Neue Technologien“ in Wien innerhalb eines eigens dafür eingerichteten Workshops dargestellt (Wilbertz 2008, vgl. CD-ROM bei den Literaturverweisen).

Eine logische Folge der Realisierung des Austauschstandards war die Beschäftigung mit den Konditionen, zu denen Daten ausgetauscht oder weitergegeben werden beziehungsweise werden sollen. Hiermit beschäftigt sich die AG RWAD, die zunächst eine Fragebogenaktion unter den Landesämtern durchführte. In diesen Prozess fiel die Notwendigkeit, sich als Archäologische Denkmalpflege mit der INSPIRE-Richtlinie der EU (INSPIRE = INfrastructure for SPatial InfoRmation in Europe) auseinanderzusetzen. Bei dieser vom Naturschutz ausgehenden Bestim-

mung, die inzwischen in nationales Recht umgesetzt worden ist, geht es darum, in der öffentlichen Verwaltung vorliegende räumliche Informationen in vereinheitlichter Form unter bestimmten Bedingungen anderen Dienststellen und der Öffentlichkeit zugänglich zu machen. Zu den ersten von INSPIRE betroffenen Objekten gehören Schutzgebiete, und zwar nicht nur Naturschutzgebiete, sondern auch Flächen mit schutzwürdigen archäologischen Objekten (Grabungsschutzgebiete, Denkmalschutzflächen und Ähnliches). Zur Beschäftigung mit dieser Thematik wurde im Herbst 2010 gemeinsam mit dem Sächsischen Landesamt für Archäologie in Dresden eine Klausurtagung durchgeführt. Der Diskussionsprozess ist noch nicht abgeschlossen.

Nachdem im Rahmen der AG Modellierung zwei Wertelisten für Oberbegriffe zu Denkmaltyp und Datierung vereinbart worden waren, ging die AG Thesaurusfragen daran, diese Listen zu überprüfen und nach Möglichkeiten einer Verfeinerung der Begrifflichkeit zu suchen. Die Arbeit begann mit den Datierungsangaben. Schon bald stellte sich heraus, dass für die einzelnen Zeitepochen in den verschiedenen Ländern unterschiedliche Datierungen für ihren Anfang und ihr Ende und damit auch für ihre Dauer gebraucht werden. Daher wurde eine „Zeitstrahl“ genannte Synopse erstellt, welche die unterschiedlichen Zeitansätze mit Jahresangaben korreliert. Sie erhebt nicht den Anspruch einer wissenschaftlich begründeten deutschlandweiten vergleichenden Chronologietabelle, sondern soll als Handreichung zur Orientierung bei Datenaustausch-Vorgängen dienen.

Die AG Archivierung begann mit einer Begriffsklärung – die ursprüngliche Bezeichnung „Langzeitarchivierung“ wurde nach eingehender Diskussion korrigiert: Archivierung nach dem Verständnis der Arbeitsgruppe ist stets unbegrenzt und streng von der (vorübergehenden) Datenspeicherung zu unterscheiden. In diesem Sinne wird der Begriff „Archivierung“ seither in allen Verlautbarungen der AG gebraucht. Im Anschluss daran wurde eine Fragebogenaktion durchgeführt, um die Situation in den Landesämtern verstehen zu lernen.

Am 19. und 20. November 2012 fand in Hannover ein Workshop statt, zu dem das Niedersächsische Landesamt für Denkmalpflege (NLD) und die Kommission „Archäologie und Informationssysteme“ eingeladen hatten. Die technische Organisation lag in den Händen des NLD, das seine Infrastruktur für die Vorbereitung und Durchführung zur Verfügung stellte. Hierfür sei an dieser Stelle ganz herzlich gedankt! Ebenso möchte ich dem Präsidenten des Landesamtes, Herrn Dr. Winghart, dafür danken, dass die vorliegenden Beiträge in der Reihe der Arbeitshefte zur Denkmalpflege in Niedersachsen gedruckt werden konnten.

Das vorliegende Arbeitsheft stellt Themen und Arbeitsergebnisse der Kommission „Archäologie und Informationssysteme“ zusammen. Zu den eigenen Texten aus den Landesarchäologien wurden Gastbeiträge hinzugefügt, die sich mit den gleichen oder verwandten Inhalten beschäftigen und somit den behandelten Themenkreis weiter abrunden sowie ergänzende Entwicklungen vorstellen.

Anmerkung

[1] Am 15.04.2013 übernahm Reiner Göldner die Sprecherfunktion.

Literatur

Herzog 2011: Irmela Herzog, ADeX – grenzenloser Austausch archäologischer Fachdaten. Archäologie in Deutschland 2011 (1) 38.

Museen der Stadt Wien – Stadtarchäologie (Hrsg.), Workshop 12 – Archäologie und Computer 2007 Kulturelles Erbe und Neue Technologien 5.–7. November 2007. Wien 2008. CD-ROM. (Beiträge Dietrich, Göldner, Himmelmann, Weidner, Wilbertz).

Wilbertz 2008: Otto Mathias Wilbertz, Kommission „Archäologie und Informationssysteme“. Archäologisches Nachrichtenblatt 13, 2008 (2) 134–141.

ADeX® Version 2.0 –

Standard für den Austausch archäologischer Fachdaten

Reiner Göldner, Ulrich Himmelmann, Thomas Poelmann, Axel Posluschny, Thomas Richter, Mario Schlapke, Roland Wanninger, Bernd Weidner, Otto Mathias Wilbertz

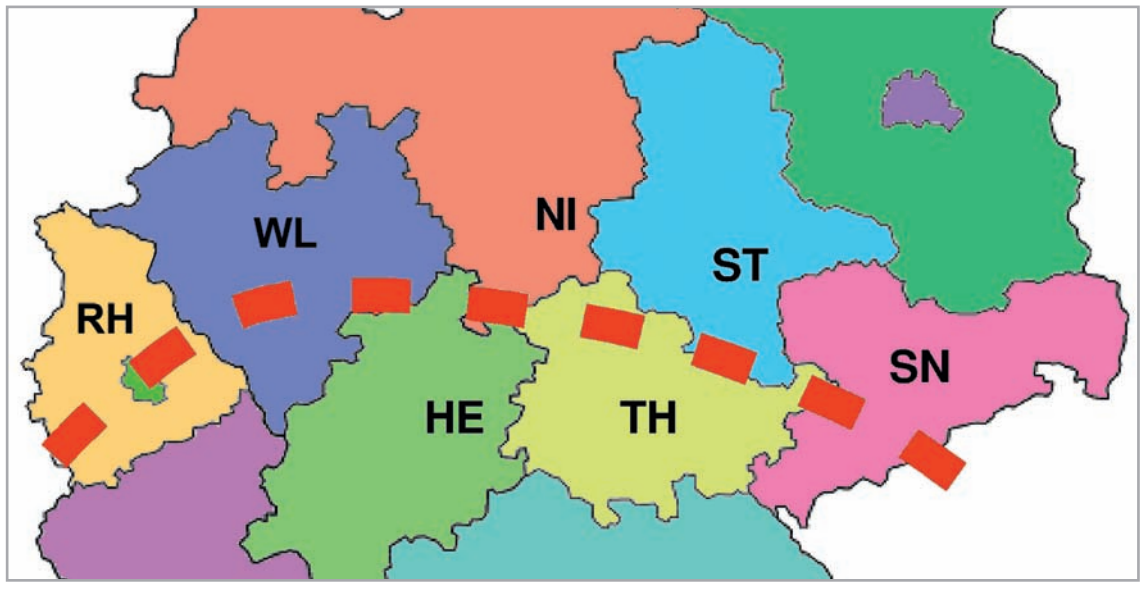
Die Erarbeitung des markenrechtlich geschützten Archäologischen Daten-Export-Standards (kurz ADeX) wurde von der Kommission „Archäologie und Informationssysteme“ beim Verband der Landesarchäologen angeregt und durch die AG Modellierung realisiert. Ziel war die Entwicklung eines bundesweiten Standards für den Datenaustausch zwischen den archäologischen Landesämtern und mit anderen Institutionen.

1 Einführung

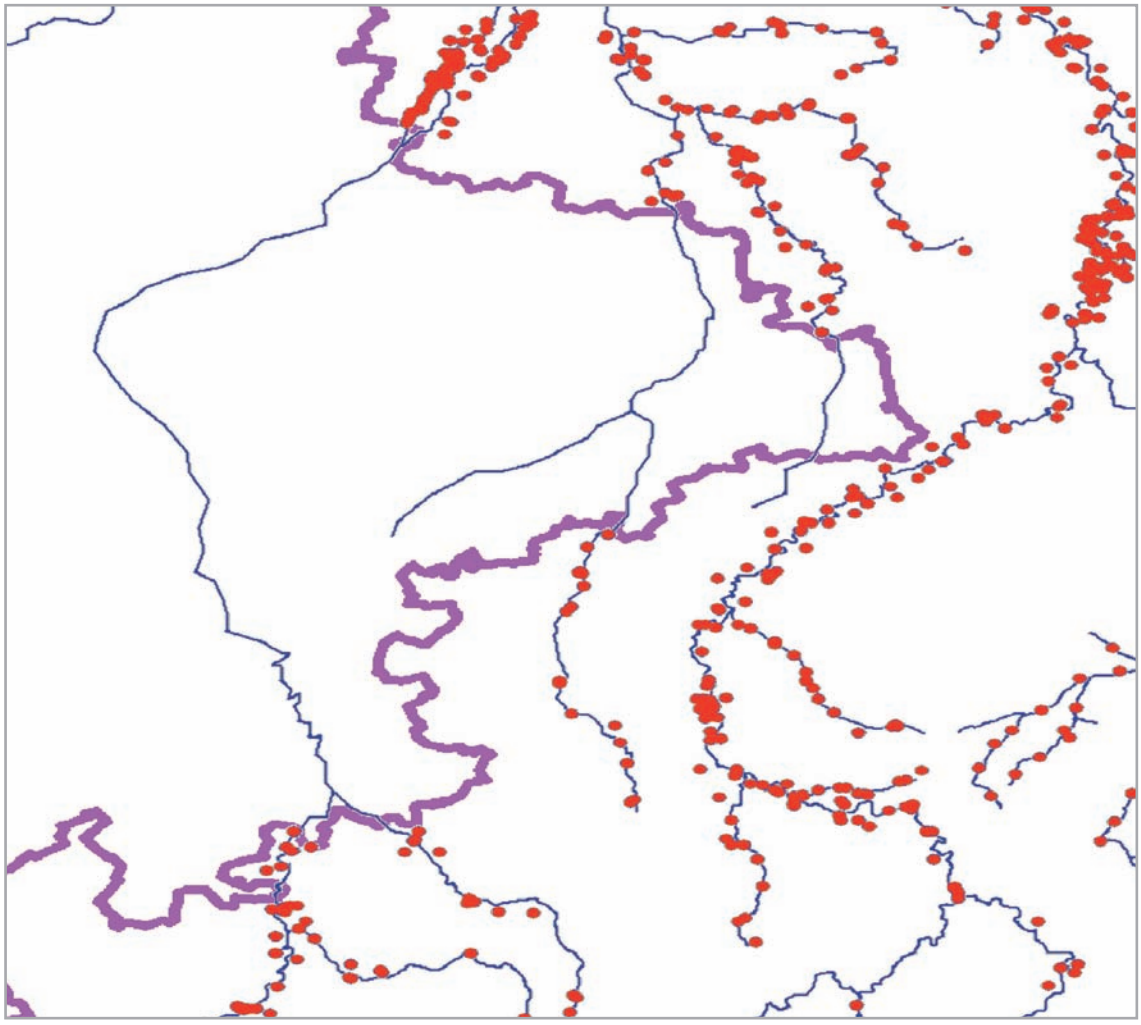
Je weiter wir in die Vergangenheit vordringen, desto weniger haben die betrachteten kulturellen Räume mit den heutigen politischen Grenzen gemeinsam. Archäologische Forschungen können daher nicht an Ländergrenzen aufhören, und auch der Denkmalschutz profitiert vom Blick zum Nachbarn. Durch die Kulturhoheit der Länder ist in Deutschland eine stark differenzierte Welt archäologischer Informationssysteme entstanden, die den Datenaustausch nicht gerade erleichtert.

Die Kommission „Archäologie und Informationssysteme“ des Verbandes der Landesarchäologen widmet sich der Aufgabe, den grenzübergreifenden Austausch archäologischer Daten zu vereinfachen. Ziel ist dabei, den Schutz des Kulturgutes zu gewährleisten und zugleich dem großen Interesse archäologischer Fachanwender und der Öffentlichkeit Rechnung zu tragen. Die Kommission bietet eine Plattform zum Informationsaustausch, leistet konzeptionelle Arbeit in thematischen Arbeitsgruppen (AGs) und macht Vorschläge zur Harmonisierung und Standardisierung archäologischer Fachdaten. Dabei erfolgt jedoch kein Eingriff in die Hoheit der zuständigen Institutionen. Ein Nutzen wird vor allem in den Bereichen Denkmalschutz/Denkmalpflege, Wissenschaft/Forschung und Planung/Verwaltung erwartet.

Die AG Modellierung widmete sich zuerst der Untersuchung der internationalen Standards CIDOC-CRM (http://www.cidoc-crm.org/) und Object-ID (http://archives.icom.museum/object-id/index.html) für Daten des kulturellen Erbes. Eine Übernahme dieser Standards erschien nicht sinnvoll, weil sie nur wenig Gewicht auf kartierbare Daten legen und zum Teil eine sehr hohe Komplexität aufweisen, die der Forderung nach einem möglichst einfachen Austauschformat widersprechen. Dennoch behielt die AG diese Standards auch bei ihren eigenen Überlegungen im Blick. Es folgte der Vergleich der in den verschiedenen beteiligten archäologischen Fachämtern vorhandenen Datenmodelle. Dabei ergab sich eine große Heterogenität, vor allem hinsichtlich der benutzten



1 Länderübergreifendes Bauvorhaben (Erdgastransportleitung Sayda-Werne-Eynatten) und betroffene Fachämter.



2 Kartierung gewässernaher Fundstellen im Grenzgebiet zweier Bundesländer ohne Einbeziehung des Nachbarn.

Begrifflichkeiten (zum Beispiel werden Fundstelle und Fundplatz nicht einheitlich verwendet). Daher wurde der neutrale Begriff „Fläche" eingeführt, der auch Punkt- und Linienobjekte einbezieht, und dazu Archäologieflächen, Untersuchungsflächen und Schutzflächen definiert (siehe dazu Absatz 2).

Aufgrund des Vergleichs der Datenmodelle entwickelte die AG Modellierung in einer ersten Version ein sehr einfaches Austauschformat, welches aus den Modulen Generelles, Georeferenz sowie Typ und Zeit besteht und Archäologieflächen beschreibt. Es umfasst damit sozusagen den kleinsten gemeinsamen Nenner aller Modelle. Diese erste Version des Archäologischen DateneXport-Standards (ADeX) ist aus pragmatischen Gründen bewusst sehr einfach gehalten und kann ohne umfangreichen Programmieraufwand implementiert werden (siehe dazu Absatz 3).

2 Begriffliche Grundlagen

Fast alle archäologischen Informationen haben einen deutlichen geografischen Bezug. Begriffe wie „Fundstelle" oder „Fundplatz" drücken dies sogar namentlich aus, allerdings werden diese Begriffe nicht einheitlich verwendet, sondern besitzen regional unterschiedliche Bedeutungen. Insofern bildet die Verständigung über diese Begriffe eine wichtige Voraussetzung für die Modellierung.

Im Ergebnis dieser Verständigung entstand die hier vorliegende Beschreibung. Sie beinhaltet eine abstrahierte Sicht auf in der Archäologie üblicherweise verwendete Geoobjekte und ist damit für eine regional übergreifende, gemeinsame Modellierung geeignet. Die hier angeführten Definitionen bilden eine Grundlage der Arbeit der AG Modellierung und finden in ihren weiteren Dokumenten Verwendung. Die vorgeschlagenen Begriffe sind ein sprachlicher Kompromiss. Ihre Verwendung hier beziehungsweise im archäologischen Kontext bietet jedoch die notwendige deutliche Abgrenzung zu den herkömmlichen Begriffen.

Allgemein wird hier von Flächen gesprochen, auch wenn diese als Fundpunkte oder Linienobjekte auf der Karte dargestellt werden. Begriffe wie „Stelle", „Platz", „Site" werden hier vermieden, weil sie in verschiedenen Ämtern beziehungsweise archäologischen Institutionen eine feste, aber unterschiedliche Definition haben.

Definition„Fläche":
Eine „Fläche" ist ein im Gelände abgrenzbarer oder lokalisierbarer Bereich.

Die Begriffe „abgrenzbar" und „lokalisierbar" sind hier bewusst nicht weiter qualifiziert, um indirekte und/oder ungenaue Lokalisierungen (wie zum Beispiel „in der Nähe von Adorf") zuzulassen.

In der Archäologie sind die folgenden drei verschiedenen Arten von „Flächen" von zentraler Bedeutung, für sie sollte der Datenaustausch funktionieren:

– Archäologiefläche
– Untersuchungsfläche
– Schutzfläche

Damit sind zwar noch nicht alle in den beteiligten Fachämtern relevanten Arten von „Flächen" abgedeckt, aber es gibt eine gemeinsame Basis. Nachfolgend werden die genannten Flächenarten definiert und näher beschrieben.

2.1 Archäologiefläche

Definition „Archäologiefläche":
Eine Archäologiefläche ist eine Fläche (siehe oben), an der mindestens ein archäologisch qualifiziertes beziehungsweise relevantes Ergebnis vorliegt/ vorgelegen hat oder vermutet wird.

Das heißt, Archäologieflächen umschließen „sicher" nachgewiesene Archäologie oder auch begründeten Verdacht auf beziehungsweise begründete Vermutung von Archäologie – oder die sicher nachgewiesene Abwesenheit von archäologischen Befunden (sogenannte Negativflächen). In der Regel werden verschiedene archäologische Beobachtungen zu einer solchen Archäologiefläche zusammengefasst. Die (wissenschaftliche) Qualität der Archäologiefläche (einfache Fundstelle, komplexe Fundstelle, etc.) beziehungsweise die Sicherheit der Erkenntnis kann man in Attributen festhalten. Archäologieflächen können sich auch auf paläontologisch qualifizierte beziehungsweise relevante Ergebnisse beziehen. Dies lässt sich zwar nicht am Begriff Archäologiefläche ablesen, doch soll dieser Begriff wegen seiner Prägnanz beibehalten werden.

Beispiel: Auf mehreren Luftbildern, die zu verschiedenen Jahreszeiten gemacht wurden, kann man einen Befestigungsgraben und Pfostenspuren zweier Langhäuser der Linearbandkeramik erkennen. Die vollständige Ausdehnung des Befestigungsgrabens ist aus den Luftbildern nicht genau abzulesen. Doch der bearbeitende Archäologe kennt ähnliche Plätze und kann daher die Ausdehnung des Platzes rekonstruieren. Diese Fläche, die die gesamte „Fundstelle" umschließt, ist eine Archäologiefläche.

Natürlich wird nicht in allen Fällen so genau gearbeitet. Gerade bei kleineren Maßnahmen (Lesefunde, Baustellenbeobachtungen) und Altdaten fasst man häufig die Fundstreuungen unterschiedlicher Zeitstellungen zu einer Fläche zusammen und generiert daraus eine Archäologiefläche. Oft ist dies nicht anders möglich, da keine genauen Angaben vorliegen. Das bedeutet, dass es zu einer Archäologiefläche mehrere Ansprachen und Datierungen geben kann. In einem solchen Fall ist die Archäologiefläche häufig mit der Untersuchungsfläche (siehe unten) identisch.

Im Folgenden wird der Idealfall für Archäologieflächen beschrieben, der zwar eher selten anzutreffen, aber für die wissenschaftliche Auswertung erstrebenswert ist: Eine Archäologiefläche umfasst idealerweise den gesamten Bereich, für den archäologische Informationen vorliegen, unabhängig von der heutigen Erhaltung.

Beispiele: Auch eine vollständig zerstörte „Fundstelle" wird als Archäologiefläche gespeichert. Ein teilzerstörter Platz, zum Beispiel die Stadtmauer, ist auch genau eine Archäologiefläche und nicht eine Ansammlung von Flächen mit den erhaltenen Teilen. Im Idealfall werden Archäologieflächen nicht durch Verwaltungsgrenzen zerteilt, das heißt auch, wenn die Archäologiefläche eine solche Verwaltungsgrenze überschreitet, bleibt sie als Einheit erhalten. Mehrere Archäologieflächen können sich überlappen, für eine römische Straße über einem älteren Gräberfeld dürfen (und sollten) zwei getrennte Archäologieflächen erfasst werden.

Sind solche idealen Archäologieflächen vollständig für eine Region erfasst, dann ist die Beantwortung von solchen Fragen möglich, wie: Wie viele merowingerzeitliche Gräberfelder sind in der Region bekannt? Wie viele Siedlungen gleicher Zeitstellung existierten in einer Region? Wie hoch ist der Anteil an befestigten Siedlungen?

2.2 Untersuchungsfläche

Definition „Untersuchungsfläche":
Eine Untersuchungsfläche umschließt den Bereich, der archäologisch beobachtet wurde.

Oftmals ist die im Rahmen einer archäologischen Maßnahme beziehungsweise Aktivität untersuchte beziehungsweise beobachtete Fläche viel umfangreicher als die „Archäologiefläche", für die, wie schon definiert, qualifizierte Ergebnisse vorliegen. Oder es ergibt sich eine umfangreiche Archäologiefläche durch eine große Anzahl von Einzeluntersuchungen. Dies wird vor allem bei nachfolgenden Beispielen deutlich, lässt sich aber auch prinzipiell verallgemeinern.

Beispiel: Bei einer Feldbegehung ist die ganze begangene Fläche die Untersuchungsfläche (also meist das gesamte Feld), nicht nur der Bereich, an dem eine Fundkonzentration nachgewiesen werden konnte. Der Begehende kann einen Teil des Feldes als Untersuchungsfläche definieren, wenn er nur dort gesucht hat. Natürlich kann das Feld auch in unterschiedliche Untersuchungsflächen eingeteilt werden, wichtig ist nur, dass jede Fläche, die archäologisch untersucht wurde, irgendeiner Untersuchungsfläche zugeordnet ist.

Beispiel: Wird der Bau einer Trasse archäologisch begleitet, aber nur an einigen ausgewählten Stellen findet eine Grabung statt, so ist der ganze Ausschachtungsbereich der Trasse eine Untersuchungsfläche.

Beispiel: Das riesige Reihengräberfeld von Krefeld-Gellep ist durch eine große Anzahl von (kleineren) Einzeluntersuchungen dokumentiert.

Es wird jedoch auch vorkommen, dass man die Untersuchungsfläche nicht explizit kennt oder nicht von der Archäologiefläche unterscheiden kann. Dann nimmt man in der Regel eine mit der Archäologiefläche identische Ausdehnung an.

Beispiel: Bei Begehungen mit wenig eindeutigen Fundstreuungen, alten Fundmeldungen oder Meldungen von Laien und bei grober Kartierung sind Untersuchungsfläche und Archäologiefläche häufig identisch.

Wichtig ist, dass man die Lage- beziehungsweise Positionsgenauigkeit als Attribut zur Untersuchungsfläche führt. Untersuchungsflächen sind ergebnisneutral (beziehungsweise archäologieneutral). Die archäologischen Ergebnisse werden als Archäologiefläche geführt, sodass also normalerweise eine oder mehrere Archäologieflächen zu einer Untersuchungsfläche Bezug nehmen (und umgekehrt). Wurde nichts gefunden, so sollte auch dies explizit vermerkt werden. Idealerweise ist also jede Archäologiefläche durch Untersuchungsflächen abgedeckt.

Bei einer Untersuchungsfläche sollte auch festgehalten werden, welche Maßnahme zur Untersuchung der Fläche durchgeführt wurde und mit welcher Intensität beobachtet wurde. Neben den klassischen Maßnahmen wie Grabung und Prospektion könnte auch die Aufarbeitung von historischen Quellen beziehungsweise Altkarten zu Untersuchungsflächen führen.

Wenn ein Platz nur teilweise archäologisch nachgewiesen wurde und die vollständige Ausdehnung durch den Wissenschaftler am Schreibtisch rekonstruiert wird, so sollte auch für diese rekonstruierte Fläche eine Untersuchungsfläche definiert werden.

2.3 Schutzfläche

Definition „Schutzfläche“:
Die Schutzfläche ist eine Fläche, die nach dem jeweiligen Landesgesetz unter Schutz steht oder für die ein Antrag auf Unterschutzstellung erfolgt ist. Eventuell ist dieser jedoch noch nicht (vollständig) bearbeitet worden.

Die Schutzfläche selbst wird, ähnlich wie die Untersuchungsfläche, archäologisch neutral betrachtet. Zu einer Schutzfläche gehört aber ein Verweis auf eine (oder mehrere) Archäologiefläche(n) und Informationen zum Schutzstatus.

Es besteht die Möglichkeit, dass es zu einer Archäologiefläche mehrere Schutzflächen gibt. Dies ist zum Beispiel im Rheinland bei gemeindeübergreifenden Fundstellen erforderlich, da dort die Gemeinden für die Unterschutzstellung zuständig sind.

Im Idealfall bezieht sich eine Schutzfläche nur auf eine Archäologiefläche, aber gerade in Innenstadtbereichen kann es vorkommen, dass mehrere (Teil-)Archäologieflächen zu einer Schutzfläche zusammengefasst werden. Die Ausdehnung einer Schutzfläche kann sich von der Ausdehnung zugehöriger Archäologieflächen unterscheiden (zum Beispiel bei Umgebungsschutz).

Die Denkmalschutzgesetze in den einzelnen Bundesländern sind unterschiedlich. Deshalb kann es nicht die Aufgabe unserer Arbeitsgruppe sein, Standards für die Ausweisung von Schutzflächen zu erarbeiten. Jede beteiligte Institution ist aufgerufen, selbst genau zu definieren, welche Formen des Schutzstatus es in dem Bundesland gibt und wie diese in der Praxis angewendet werden.

Schutzflächen sind nicht zwingend identisch mit Grabungsschutzgebieten.

2.4 Weitere Flächenarten und Beispiele

Eine weitere Flächenart, die ausgetauscht werden kann, sind zum Beispiel die (externen) Planungsflächen. Primär sollte man sich jedoch auf die Flächen beschränken, die von der gesetzlich zuständigen Institution erhoben werden. Die Flächen aus anderen Quellen spielen eine untergeordnete Rolle.

Beispiele: Eine Gemeinde plant ein neues Gewerbegebiet. Das Planungsareal (weitere Flächen) liegt dem Landesamt vor, eine Stellungnahme soll erarbeitet werden. Es wird festgestellt, dass aus diesem Bereich ungenaue Altmeldungen (Archäologieflächen) bekannt sind. Um diese Altmeldungen zu überprüfen, werden Prospektionsmaßnahmen (Untersuchungsflächen) durchgeführt. Diese führen zu einer genaueren Festlegung der Grenzen der Archäologieflächen. Eine bedeutende Fundstelle in diesem Bereich ist bereits teilweise durch eine Straße zerstört, aber der Rest soll erhalten bleiben (Schutzfläche). An dieser Stelle wird eine Umplanung des Gewerbegebietes angestrebt. Die Gemeinde entscheidet sich jedoch für eine Zerstörung des Bodendenkmals, daher wird eine Grabung durchgeführt (Untersuchungsfläche stimmt hier mit der Schutzfläche überein). Die nun vollständig entsorgte Fundstelle bleibt in der Liste der Archäologieflächen verzeichnet. Es wird nur vermerkt, dass sie nicht mehr erhalten ist.

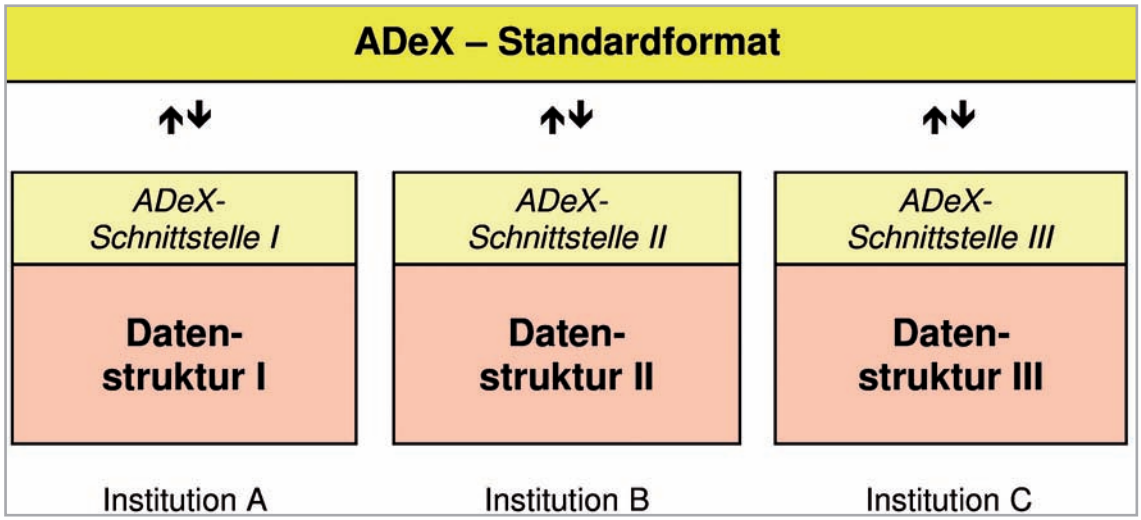
3 ADeX-Modellstruktur

ADeX (Archäologischer DateneXport) ist als bundesweiter Standard für den Datenaustausch zwischen den archäologischen Landesämtern beziehungsweise mit anderen Institutionen ausgelegt.

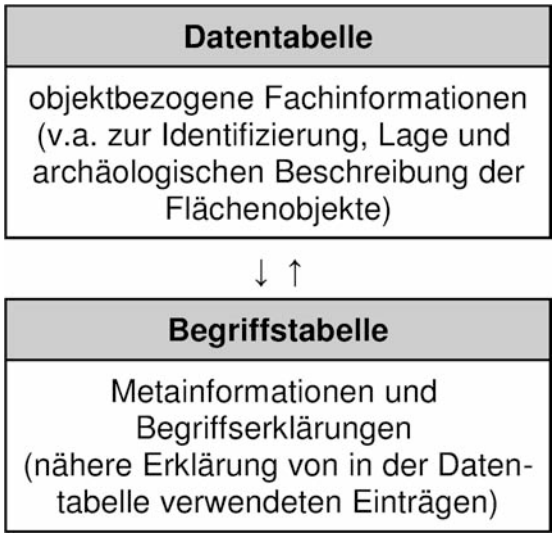
Da die bestehenden Datenmodelle der Denkmalfachbehörden nicht zur Disposition stehen, wird eine Schnittstelle definiert, die dem Export und Import dient und eine einheitliche Sicht auf die archäologischen Informationen schafft (Abb. 3). Dieses Datenaustauschformat bildet, gemeinsam mit Informationen zur technischen Umsetzung, den ADeX-Standard.

Version 2.0 des ADeX-Standards bezieht sich vorerst nur auf Archäologieflächen. Schutz- und Untersuchungsflächen sollen in spätere Versionen integriert werden. Die Datenstruktur selbst ist einfach strukturiert: Neben der Datentabelle mit den objektbezogenen Fachinformationen gibt es die Begriffstabelle mit Metainformationen und Begriffserklärungen.

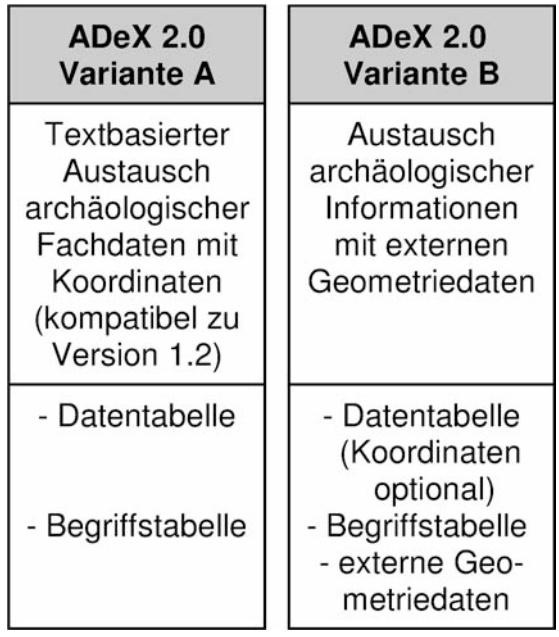
Ab Version 2 erlaubt ADeX den Austausch von genaueren Abgrenzungen von Archäologieflächen in Form von Linien oder Polygonen. ADeX 2 ist abwärtskompatibel, der bisherige textbasierte Standard mit Koordinatenangaben bleibt als Variante A erhalten. Variante B übernimmt weitestgehend die Strukturen von Variante A, hinzu kommen die externen Geome-



3 ADeX-Schnittstelle zur Vereinheitlichung der Sicht auf unterschiedlich strukturierte Daten.



4 Datentabelle und Begriffstabelle.



5 ADeX-Version 2.0, Varianten A und B.

triedaten, die sowohl Punkte als auch Linien und Polygone enthalten können.

Als externe Geometriedaten werden spezifiziert:
SHP: ESRI Shape Format
MIF: MapInfo Interchange File Format.

3.1 Datentabelle

Die Datentabelle ist modular aufgebaut. Das Datenmodell besteht aus den, vorerst unstrukturierten, Modulen Generelle Angaben, Georeferenz und Typ/Zeit.

Das Modul Generelle Angaben umfasst allgemeine Angaben wie den Denkmalnamen und rechtliche Hinweise, das Modul Georeferenz beinhaltet Daten zu Lage und Ausdehnung der Fläche und Typ/Zeit zu Datierung und Ansprache. Zwingend zu füllende Felder sind mit einem Plus-Symbol gekennzeichnet. Bestimmte Felder (TYP_GROB, DAT_GROB) sind mit verpflichtenden Begriffslisten hinterlegt.

Die einzelnen Felder werden nachfolgend beschrieben. Informationen zur technischen Umsetzung folgen in Absatz 3.4.

Generelle Angaben	Georeferenz	Typ / Zeit
+ ADEX_ID	(+) KOO_REFSYS	+ TYP_GROB
– BEZEICHNG	(+) X_KOORD	– TYP_FEIN
+ FLAECH_ART	(+) Y_KOORD	– TYP_ERLAEU
– ERFASS_DAT	– X_VON, X_BIS	+ DAT_GROB
– AENDER_DAT	– Y_VON, Y_BIS	– DAT_FEIN
– ANSPRECHP	+ GENAUIGK	– DAT_ERLAEU
– DAT_QUELLE	– GENAUIGK_T	
– BERECHTIG	– GDE_KENN	
– COPYRIGHT	– GDE_NAME	
– ZUSATZ		

6 ADeX-Module der Version 2.0.

3.1.1 Modul Generelle Angaben

ADEX_ID – AdeX Identifikator
character, unique
(Pflichtfeld, Primärschlüssel, Fremdschlüssel für externe Geometriedaten)
Angabe, die eine eindeutige Identifikation des Datensatzes beim Datennehmer ermöglicht (zum Beispiel Landeskürzel [+ "_" + landesinterne Kennung der Institution, wenn notwendig] + "_" + landesinterne ID [+ "_" + Zähler, wenn notwendig])
Bemerkung: Der Zähler ist nur dann notwendig, wenn in der landesinternen Datenbank die gleiche Kennung für mehrere abzugebende Datensätze vorhanden ist. Zu Landeskürzeln siehe Anlage 1.

BEZEICHNG – Bezeichnung
character
(optional, Freitext)
Beim Datengeber gebräuchliche, leicht verständliche Bezeichnung der Fläche. Häufig vermutlich Gemarkung/Gemeinde/Kreis plus Nr. wie zum Beispiel Hartholzheim 15, gegebenenfalls aber auch Angaben wie Trier 117 (Porta Nigra) beziehungsweise einfach nur Hardenburg.

FLAECH_ART – Flächenart
character
(Pflichtfeld, nur Listenwerte)
Angabe darüber, was hier ausgetauscht wird.
Aktuelle Begriffsliste (V1): Archäologiefläche, später auch: Untersuchungsfläche, Schutzfläche (vgl. Abschnitt 2).

ERFASS_DAT – Erfassungsdatum
character
(optional, Text)
Datum der Datensatzanlage
Format: JJJJ[-MM[-TT]] (J-Jahr, M-Monat, T-Tag, zum Beispiel 1989, 1995-08, 2008-11-03).

AENDER_DAT – Änderungsdatum
character
(optional, Text)
Datum der letzten Änderung am Datensatz
Format: JJJJ[-MM[-TT]] (J-Jahr, M-Monat, T-Tag, zum Beispiel 1989, 1995-08, 2008-11-03).

ANSPRECHP – Ansprechpartner
character
(optional, Kurztext)
Bezeichnung (Name) des Ansprechpartners bei inhaltlichen Fragen (zum Beispiel Bezirksarchäologe, letzter Bearbeiter).

DAT_QUELLE – Datenquelle
character
(optional, Kurztext)
Benennt die Datenquelle beim Urheber, aus welcher der Datensatz stammt. Dient zum Beispiel der Rückverfolgbarkeit von Datensätzen. Wird nur angegeben, wo verfügbar und nötig; zum Beispiel Luftbild-Tabelle, Aktivitäten-Tabelle usw.

BERECHTIG – Berechtigung
character
(optional, Kurztext)
Raum für (datensatzbezogene) Berechtigungsvermerke, Hinweise, Weitergabebeschränkungen etc. (Hinweis: Generelle Berechtigungen werden bilateral ausgehandelt und schlagen sich im Allgemeinen in einer Nutzungsvereinbarung nieder.)

COPYRIGHT – Copyrightinhaber
character
(optional, Kurztext)
Eindeutige Bezeichnung des Copyrightinhabers.

ZUSATZ – zusätzliche Informationen
character
(optional, Freitext)
Dieses Feld bietet die Möglichkeit einer freien Beschreibung von Objekteigenschaften oder dient zum Austausch von zusätzlichen Informationen, die sonst nicht definiert sind. Der Inhalt sollte selbsterklärend sein oder in der Begriffstabelle erläutert werden.

3.1.2 Modul Georeferenz

KOO_REFSYS – Koordinatenreferenzsystem
long integer
(Variante A: Pflichtfeld, nur Listenwerte beziehungsweise Variante B: optional)
EPSG-Codes zur Angabe von Datum und Projektion (siehe auch Anlage 3). Listenwerte entsprechend EPSG. Beachte mögliche Abweichung bei Reihenfolge und Benennung von X- und Y-Koordinate. Siehe Beschreibung bei X_KOORD, Y_KOORD und X_VON, Y_VON, X_BIS, Y_BIS.

X_KOORD, Y_KOORD – Punkt-Koordinaten
double
(Variante A: Pflichtfelder beziehungsweise Variante B: optional)
Koordinaten (Wertepaar) zur (unter Umständen generalisierten) Beschreibung der Lage des Objekts: zum Beispiel Rechtswert + Hochwert oder Ost-Längengrad + Nord-Breitengrad (Werte entsprechend KOO_REFSYS); mathematisches Koordinatensystem angenommen (X: West→Ost, Y: Süd→Nord).

X_VON, Y_VON, X_BIS, Y_BIS – Ausdehnung
double
(optional)
Umschreibendes Rechteck zur (unter Umständen generalisierten) Beschreibung der Lage des Objekts: 4 Werte (2 Koordinatenpaare): links, unten + rechts, oben (Werte entsprechend KOO_REFSYS, Lage beachten: X_VON < X_BIS, Y_VON < Y_BIS); mathematisches Koordinatensystem angenommen (X: W→O, Y: S→N).

GENAUIGK – Genauigkeitswert (Zahl)
real
(Pflichtfeld)
Numerische Angabe zur Genauigkeit der gelieferten Georeferenz (siehe auch Anlage 4, Angaben in Metern (im Sinne eines mittleren Fehlers).

GENAUIGK_T – Genauigkeitsangaben (Text)
character
(optional, Kurztext)
Verbale Angaben zur Genauigkeit der gelieferten Georeferenz, gegebenenfalls als Ergänzung/Erklärung zum Feld GENAUIGK, je nach Datengeber Begriffe aus einer Begriffsliste (Klassen mit Beschreibung in der Begriffstabelle).

GDE_KENN – Gemeindekennzahl
character (8) (nur Ziffern, aber gegebenenfalls mit führender Null)
(optional, nur Listenwerte)
Gemeindezugehörigkeit/Gemeindekennzahl, sollte möglichst mitgeliefert werden, Listenwerte: offizieller Gemeindeschlüsselkatalog.

GDE_NAME – Gemeindename
character
(optional, Kurztext)
Gemeindezugehörigkeit – Gemeindename, sofern verfügbar.

3.1.3 Modul Typ/Zeit

TYP_GROB – Typ grob
character
(Pflichtfeld, nur Listenwerte)
Grobansprache des Typs. Übergeordnete breite Kategorisierung der Fundstelle (zum Beispiel Siedlung, Gräberfeld, unbekannt). Die Begriffe stammen aus der gemeinsamen Typ-Begriffsliste (siehe Abschnitt 3.1.4), sind also bundesweit einheitlich. Vgl. DAT_GROB bei Zeit.

TYP_FEIN – Typ fein
character
(optional, Kurztext)
Feinansprache des Typs. Möglichst detaillierte Beschreibung des Fundstellentyps durch funktionelle oder beschreibende Kriterien, die gegebenenfalls einem hierarchischen Thesaurus entnommen werden (Höhle, Abri, Wasserburg, Hügelgräber). Die Begriffe stammen aus dem jeweiligen Thesaurus des Datengebers, können also von Land zu Land variieren. Vgl. DAT_FEIN bei Zeit.

TYP_ERLAEU – Typ Erläuterung
character
(optional, Freitext)
Datensatzspezifische, nähere (Typ-)Ansprache des konkreten Objekts, sofern diese nicht durch die Felder TYP_GROB und TYP_FEIN abgedeckt werden kann, also zum Beispiel Anmerkungen (siehe auch unter DAT_ERLAEU).
Nicht verwechseln mit Angaben der Begriffstabelle!

DAT_GROB – Datierung grob/Epoche
character
(Pflichtfeld, nur Listenwerte)
Grobdatierung. Übergeordnete breite Datierung der Fundstelle, die der Typ-Begriffsliste (siehe Abschnitt 3.1.5) entnommen werden muss (Vorgeschichte, Neolithikum, Römische Kaiserzeit). Die Begriffe stammen aus der gemeinsamen Typ-Begriffsliste, sind also bundesweit einheitlich. Vgl. DAT_GROB bei Typ.

DAT_FEIN – Datierung fein/Phase
character
(optional, Kurztext)
Feindatierung. Möglichst detaillierte Beschreibung der Datierung durch funktionelle oder beschreibende Kriterien, die gegebenenfalls einem hierarchischen Thesaurus entnommen werden (Altpaläolithikum, Endpaläolithikum, Frühe Bronzezeit, Ha A1, Späte RKZ, Karolingerzeit, 17. Jh.). Die Begriffe stammen aus dem jeweiligen Thesaurus des Datengebers, können also von Land zu Land variieren. Vgl. TYP_FEIN bei Typ.

DAT_ERLAEU – Datierung Erläuterung
character
(optional, Freitext)
Datensatzspezifische, nähere Angaben zur Datierung des konkreten Objekts, sofern diese nicht durch die Felder DAT_GROB und DAT_FEIN abgedeckt werden. Hier zunächst alle Angaben über Qualität der Datierung, Datierungsmethode, Kulturgruppenzugehörigkeit etc. (vgl. TYP_ERLAEU). Nicht verwechseln mit Angaben der Begriffstabelle!

3.1.4 Typ-Begriffsliste

Ein Eintrag aus der Typ-Begriffsliste ist Pflicht für das Feld TYP_GROB.

Folgende Begriffe stehen zur Verfügung:
- Unbekannt
- Erdgeschichtliches Objekt/Natur
- Einzelfund
- Depot
- Grab/Grabfunde
- Siedlung/Befestigung/Produktion
- Siedlung
- Befestigung/Militär
- Produktion/Agrar
- Verkehr/Wasser/Infrastruktur
- Kult
- Sonstiges

Es soll jeweils der genaueste ermittelbare Begriff verwendet werden.

3.1.5 Datierungs-Begriffsliste

Ein Eintrag aus der Datierungs-Begriffsliste ist Pflicht für das Feld DAT_GROB.

Folgende Begriffe stehen zur Verfügung:
- Unbekannt
- Erdgeschichte
- Vorgeschichte
- Steinzeit
- Paläolithikum
- Mesolithikum
- Neolithikum
- Metallzeit
- Bronzezeit
- Vorrömische Eisenzeit
- Frühgeschichte
- Römische Kaiserzeit
- Völkerwanderungszeit
- Mittelalter/Neuzeit
- Mittelalter
- Neuzeit

Es soll jeweils der genaueste ermittelbare Begriff verwendet werden.

3.2 Begriffstabelle

Die Begriffstabelle enthält Erklärungen zu Begriffen, die in den Textfeldern der Datentabelle verwendet werden. Jeder Wert aus der Datentabelle kann in der Begriffstabelle näher erläutert werden.

INST – Institution des Datengebers
character
(Pflichtfeld, Kurztext)
Landeskürzel [+ "_" + landesinterne Kennung der Institution, wenn notwendig]
Zu Landeskürzeln siehe Anlage 1.

FELD – Name des Feldes in der Datentabelle
character
(Pflichtfeld, Kurztext)
Name des Feldes in der Datentabelle.

BEGRIFF – Möglicher Feldinhalt in der Datentabelle
character
(Pflichtfeld, Kurztext)
Möglicher Feldinhalt in der Datentabelle.

ERKLAER – Ausführliche Erklärung des Feldinhaltes
character
(Pflichtfeld, Freitext)
Ausführliche Erklärung des Feldinhaltes.

3.3 Externe Geometriedaten

Um den flachen, einfach strukturierten Aufbau zu erhalten, sind Fachdaten und Geometriedaten 1:1 verknüpft. Für jeden Datensatz der Datentabelle existiert genau ein Grafikobjekt und umgekehrt.

Als externe Geometriedaten werden spezifiziert (inkl. CRS-Definition):

SHP: ESRI Shape Format (Hier werden Punkte, Linien und Polygone in separaten Ebenen gespeichert. Außerdem besteht eine Ebene (Shapefile) aus mehreren Dateien).
MIF: MapInfo Interchange File Format (besteht gegebenenfalls aus 2 Dateien MIF + MID).

Die Verknüpfung von Fachdaten und Geometriedaten wird über ein bei den Geometriedaten angelegtes Feld realisiert: „ADEX_ID" vom Typ character. Optional und in gegenseitiger Absprache können die spezifizierten ADeX-Felder auch in den Attributbereich der Geometriedaten integriert werden. Optional können die Daten auch als Webdienste (WMS/WFS) bereitgestellt werden.

3.4 Informationen zur technischen Umsetzung

Nachfolgende generelle Hinweise sind zu beachten:

- **ADEX_ID** ist Schlüsselfeld.
- **Pflichtfelder** müssen, optionale Felder sollten, soweit möglich, ausgefüllt werden.
- **Freitext** hat in der Regel keine Referenz auf andere Felder oder Tabellen.
- **Kurztext** soll kurzer, prägnanter und lesbarer Text sein, der in der Regel in der Begriffstabelle ausführlicher erläutert wird. Die Angaben stammen gegebenenfalls aus einem Landesthesaurus des Datenanbieters.
- **Nur Begriffsliste**: dieser Eintrag muss mit einem Begriffslisten-Eintrag identisch sein, der für dieses Feld definiert ist (referenzielle Integrität), andere Werte sind nicht erlaubt. Die Begriffe können in der Begriffstabelle näher erläutert werden.
- Die Erklärungen in der **Begriffstabelle** sind nur informatorisch, also nicht streng referenziert.
- Die Feldnamen sind bewusst kurz gewählt (max. 10 Zeichen), um bestimmten GIS-/Datenbankanforderungen (zum Beispiel Shapefiles) zu genügen.
- Die Feldtypen orientieren sich an SQL. ‚Character' meint unformatierten Text (siehe unten: CSV). Sofern bei ‚character' keine Länge angegeben ist, wird von character(255) ausgegangen. Real meint Fließkommazahlen mit einfacher Genauigkeit (circa 7 Stellen), double meint doppelte Genauigkeit (circa 15 Stellen).

und weiterhin

- Der Datenaustausch erfolgt im **CSV-Format** in folgender Konfiguration:
  - erste Zeile enthält die Feldnamen (zwingend)
  - Feldtrennzeichen: „|"
  - Dezimaltrenner Punkt „."
  - Zahlenangaben: ohne Tausender-Trenner
  - Textangaben: ohne Zeilenumbrüche, ohne Textbegrenzungszeichen (kein Memo, keine „|" verwenden)
- Tipps: Feldtrennzeichen bei den Feldnamen (erste Zeile) nicht vergessen. Zum Datenaustausch müssen alle Felder vorhanden sein, sie dürfen je nach Felddefinition aber leer sein (optional). Das Trennzeichen „|" für das CSV-Format muss auch bei leeren Feldern vorhanden sein.

- Bevorzugte Textkodierungen sind UTF8 und Latin-9 (ISO 8859-15).

Die Dateinamen werden nach einheitlichem Muster erstellt. Um Automatismen und Voreinstellungen (zum Beispiel unter Windows) auszuschließen, erhalten die Dateien die Namenserweiterung „.txt“:

- Datentabelle:
  adex2-00_[Landeskürzel]_d_[JJJJ-MM-TT].txt
- Begriffstabelle:
  adex2-00_[Landeskürzel]_b_[JJJJ-MM-TT].txt

Es ist angeraten, auch Metadaten (zum Beispiel die aktuell benutzte Version des ADeX-Formates, CSV-Parameter usw.) beim Datenaustausch mitzuliefern. Dazu kann gegebenenfalls die Begriffstabelle genutzt werden. Bei der Nutzung von Sonderzeichen kann es erforderlich werden, Angaben zu Kodierung beziehungsweise Zeichensatz mitzuliefern.

Ein deutschlandweit einheitlicher Umgang mit den Begriffslisten (Felder TYP_GROB, DAT_GROB) wird angestrebt, kann aber durch die AG Modellierung nicht erarbeitet werden. Hinweis: Datierungsangaben über mehrere Stufen/Epochen/Phasen (Feld DAT_FEIN beziehungsweise DAT_ERLAEU) geben Unsicherheiten oder Spannen wieder! Das Problem kann in Version 1 beziehungsweise 2 nicht gelöst werden.

3.5 Bezug zu internationalen Standards

Bei der Entwicklung von ADeX wurden auf der konzeptuellen Ebene neben den etablierten Datenstrukturen archäologischer Institutionen auch die internationalen Standards „International Core Data Standard for Archaeological Sites and Monuments“ und „CIDOC Conceptual Reference Model (CRM)“ berücksichtigt. Diese Standards sind sehr komplex und stark auf Universalität ausgerichtet, um verschiedenste Bereiche des kulturellen Erbes abdecken zu können. Sie gehen jedoch oft nicht weit genug in die Tiefe und müssen aufwendig individuell erweitert werden, um eine für die Praxis ausreichend detaillierte Modellierung zu ermöglichen.

Eine ausführliche Dokumentation zur CIDOC-CRM-Kompatibilität ist in der AG Modellierung verfügbar.

Das ADeX-Datenmodell lässt sich auf konzeptueller Ebene sowohl mit dem „International Core Data Standard for Archaeological Sites and Monuments“ als auch mit dem „CIDOC Conceptual Reference Model (CRM)“ in Einklang bringen. ADeX geht jedoch an vielen Stellen sehr viel weiter als diese Modelle. Einerseits bietet ADeX auf der konzeptuellen Ebene eine detaillierte fachliche Anpassung an die in Deutschland im Bereich der archäologischen Denkmalpflege etablierten Datenstrukturen. Andererseits wurde großer Wert auf eine möglichst einfache praktische Umsetzung gelegt, die mit Standardmitteln und ohne aufwendige Programmierleistungen auskommen kann.

GDI-DE setzt seit einiger Zeit Maßstäbe beim Aufbau einer deutschen Geodateninfrastruktur. Neben einer Vielzahl genereller Empfehlungen, Vorgaben und technischer Spezifikationen ist für eine funktionierende Geodateninfrastruktur jedoch auch eine fachspezifische Harmonisierung der anzubietenden Datenbestände erforderlich. Diesen Zweck erfüllt ADeX in vorbildlicher Art und Weise, indem es eine einheitliche Sicht auf die bestehenden Datenbestände schafft.

Anmerkung

Der Datenaustauschstandard ADeX wurde von der Kommission „Archäologie und Informationssysteme“ beim Verband der Landesarchäologen angeregt und durch die AG Modellierung realisiert, an der folgende Kolleginnen und Kollegen beteiligt sind oder waren:

David Bibby, Landesamt für Denkmalpflege im Regierungspräsidium Stuttgart
Dr. Utz Böhner, Niedersächsisches Landesamt für Denkmalpflege
Dr. Hanns Dietrich, Bayerisches Landesamt für Denkmalpflege
Dr. Reiner Göldner, Landesamt für Archäologie Sachsen
Irmela Herzog, LVR – Amt für Bodendenkmalpflege im Rheinland
Dr. Ulrich Himmelmann, Generaldirektion Kulturelles Erbe Rheinland-Pfalz
PD Dr. Sven Ostritz, Thüringisches Landesamt für Denkmalpflege und Archäologie
Thomas Poelmann, Archäologisches Landesamt Schleswig-Holstein
Dr. Axel Posluschny, Römisch-Germanische Kommission (RGK) Frankfurt a. M.
Thomas Richter, Landesamt für Denkmalpflege und Archäologie Sachsen-Anhalt
Mario Schlapke, Thüringisches Landesamt für Denkmalpflege und Archäologie
Roland Wanninger, Bayerisches Landesamt für Denkmalpflege
Dipl.-Phys. Bernd Weidner, interactive instruments, Bonn
Dr. Otto Mathias Wilbertz, Niedersächsisches Landesamt für Denkmalpflege.

ADeX, Anlage 1 – Landeskürzel

Der ADeX-Identifikator (ADEX_ID) beinhaltet das sogenannte „Landeskürzel". Es soll dazu dienen, unabhängig von internen Nummerierungen der Datengeber, einen global eindeutigen Bezeichner im Sinne eines Datenbank-Schlüssels zu verwenden. Für den Kreis der nachfolgend aufgeführten Institutionen ist dies über die angegebenen Landeskürzel direkt möglich, die Nutzung dieser Kürzel ist Pflicht.

Beim Datenaustausch zwischen hier nicht aufgeführten Institutionen sollte entweder ein unabhängiges Kürzel separat vereinbart werden oder der Datennehmer ist für die Generierung einer eindeutigen ADEX_ID aus den übermittelten Daten selbst verantwortlich.

Liste der festgelegten Landeskürzel:

BW	Baden-Württemberg
BY	Bayern
BE	Berlin
BR	Brandenburg
HB	Bremen
HH	Hamburg
HE	Hessen
MV	Mecklenburg-Vorpommern
NI	Niedersachsen
NW	Nordrhein-Westfalen
RH	(Nordrhein-Westfalen) Rheinland
WL	(Nordrhein-Westfalen) Westfalen-Lippe
RK	(Nordrhein-Westfalen) Köln
RP	Rheinland-Pfalz
SL	Saarland
SN	Sachsen
ST	Sachsen-Anhalt
SH	Schleswig-Holstein
TH	Thüringen
RGK	Römisch-Germanische Kommission
DAI	Deutsches Archäologisches Institut

ADeX, Anlage 2 – Beispieldaten

Die nachfolgenden Abbildungen zeigen Beispieldaten zu ADeX 2.0. Alle angeführten Daten sind fiktiv.

ADEX_ID	NI_111/1111.00001-F	SN_11110-D-01-1	RH_OVP 2001/0111
BEZEICHNG	Dammburg	Sachsenburg	Junckersburg
FLAECH_ART	Archäologiefläche	Archäologiefläche	Archäologiefläche
ERFASS_DAT	1989		2008-11-03
AENDER_DAT		1995-08	2009-04-28
ANSPRECHP	Müller	Meier	Schulze
DAT_QUELLE	ADABweb	DIA	ZADAB-Platz-Modul
BERECHTIG	Weitergabe nur nach Genehmigung durch das NLD	lt. Nutzungsvereinbarung	ADeX-Test
COPYRIGHT	NLD	LfA SN	LVR/RAB
ZUSATZ			
KOO_REFSYS	31467	31468	31466
X_KOORD	3610000	4570000	2590000
Y_KOORD	5840000	5640000	5660000
X_VON	3609990		2589990
Y_VON	5839990		5659990
X_BIS	3610010		2590010
Y_BIS	5840010		5660010
GENAUIGK	10	25	20
GENAUIGK_T	5 - 20 m		bis auf 20 Meter schätzbar
GDE_KENN	3151019		5378012
GDE_NAME	Obernholz	Frankenberg	Kürten
TYP_GROB	Siedlung/Befestigung/ Produktion	Siedlung/Befestigung/ Produktion	Siedlung/Befestigung/ Produktion
TYP_FEIN	Burg	Burg	Burg
TYP_ERLAEU			
DAT_GROB	Mittelalter	Mittelalter	Mittelalter/Neuzeit
DAT_FEIN	Hochmittelalter (11. Jh.)	Mittelalter	Mittelalter bis Neuzeit
DAT_ERLAEU			

A2.1. Datentabelle (aus Darstellungsgründen ist die Tabelle transponiert).

```
ADEX_ID |BEZEICHNG |FLAECH_ART |ERFASS_DAT |AENDER_DAT |ANSPRECHP |DAT_QUEL-
LE |BERECHTIG |COPYRIGHT |ZUSATZ |KOO_REFSYS |X_KOORD |Y_KOORD |X_VON |Y_VON |
X_BIS |Y_BIS |GENAUIGK |GENAUIGK_T |GDE_KENN |GDE_NAME |TYP_GROB |TYP_FEIN |
TYP_ERLAEU |DAT_GROB |DAT_FEIN |DAT_ERLAEU
NI_111/1111.00001-F |Dammburg |Archäologiefläche | 1989| |Müller |ADABweb |Weitergabe nur nach
Genehmigung durch das NLD |NLD | |31467 |3610000 |5840000 |3609990 |5839990 |3610010 |
5840010 |10 |5 - 20 m |3151019 |Obernholz | Siedlung/Befestigung/Produktion |Burg | |Mittelalter |
Hochmittelalter (11. Jh.) |
SN-11110-D-01-1 |Sachsenburg |Archäologiefläche | |1995-08 |Meier |DIA |lt. Nutzungsvereinbarung |
Landesamt für Archäologie SN | |31468 |4570000 |5640000 | | | |25 | | |Frankenberg | Siedlung/Befes-
tigung/Produktion |Burg | |Mittelalter |Mittelalter |
RH_OVP 2001/0111 |Junckersburg |Archäologiefläche | 2008-11-03| 2009-04-28|Schulze |ZADAB-
Platz-Modul |ADeX-Test |LVR/RAB | |31466 |2590000 |5660000 |2589990 |5659990 |2590010 |
5660010 |20 |bis auf 20 Meter schätzbar |5378012 |Kürten |Siedlung/Befestigung/Produktion |Burg | |
Mittelalter/Neuzeit |Mittelalter bis Neuzeit |
```

A2.2. Datentabelle als CSV.

INST	FELD	BEGRIFF	ERKLAER
NI	ANSPRECHP	Müller	Max Müller, Musterinstitut für Archäologie
SN	KOO_REFSYS	31468	RD83 (DHDN-Sachsen), Gauß-Krüger-Projektion 4. Streifen
RH	BERECHTIG	ADeX-Test	Testdatensatz. Nutzung nur im Rahmen der ADeX-Präsentation erlaubt.

A2.3. Begriffstabelle.

```
INST |FELD |BEGRIFF |ERKLAER
NI |ANSPRECHP |Müller |Max Müller, Musterinstitut für Archäologie
SN |KOO_REFSYS |31468 |RD83 (DHDN-Sachsen), Gauß-Krüger-Projektion 4. Streifen
RH |BERECHTIG |ADeX-Test |Testdatensatz. Nutzung nur im Rahmen der ADeX-Präsentation erlaubt
```

A2.4. Begriffstabelle als CSV.

OBJECTID	SHAPE	ADEX_ID	ggf die weiteren ADeX-Felder integrieren
1	Polygon	NI_111/1111.00001-F	
2	Polygon	SN_11110-D-01-1	
3	Polygon	RH_OVP 2001/0111	

A2.5. Geometriedaten als SHP (Tabellenansicht).

MIF (Ausschnitte)	MID
Version … Charset … Delimiter "," CoordSys … Columns 1 ADEX_ID Char(50) Data Region 1 5 3610000.0 5840000.0 3610000.0 5840010.0 3610010.0 5840010.0 3610010.0 5840000.0 3610000.0 5840000.0 Pen … Brush … Center 3610005.0 5840005.0 Region … Region …	„NI_111/1111.00001-F“ „SN_11110-D-01-1” “RH_OVP 2001/0111”
ggf. unter Columns die weiteren ADeX-Felder integrieren	ggf. die weiteren ADeX-Felder integrieren: „NI_111/1111.00001-F“,“Dammburg“,“Archäologiefläche“, … „SN-11110-D-01-1“,“Sachsenburg“,“Archäologiefläche“, … „RH_OVP 2001/0111“,“Junckersburg“,“Archäologiefläche“, …

A2.6. Geometriedaten als MIF/MID (Die Reihenfolge der Datensätze in MIF und MID muss identisch sein!).

ADeX, Anlage 3 – Koordinatenreferenzsysteme

Beschreibung

Das Koordinatenreferenzsystem (engl. Coordinate Reference System – CRS) wird bestimmt durch die Wahl eines Datums (Parametersatz zur Beschreibung des Erdellipsoids) und durch die Art der Projektion (Abbildung in die Ebene).

Als Datum sind zum Beispiel üblich:

- WGS84 (World Geodetic System 1984) – GPS
- ITRS (International Terrestrical Reference System)
- **ETRS89 (European Terrestrical Reference System)**
- DHDN (Deutsches Hauptdreiecksnetz)
- RD83 (Rauenberg Datum 1983) – Sachsen
- PD83 (Potsdam Datum 1983) – Thüringen
- 42/83 (Pulkovo, Krassovsky) – Mecklenburg-Vorpommern, Sachsen-Anhalt

Für das Verständnis ist wichtig, dass sich die Systeme oft nur im Feinen unterscheiden. So ist ETRS89 eine Präzisierung von ITRS, welches sich wiederum an WGS84 orientiert. Ebenso sind RD83 und PD83 Realisierungen des DHDN in Sachsen und Thüringen.

Hinweis: ETRS89 wandert gegenüber WGS84 um circa 2,5 cm/Jahr, zurzeit liegen die Abweichungen aber unter 1 m (genauer Wert von der Lage in Europa abhängig).

Als Projektion werden häufig verwendet

- GK (Gauß-Krüger) – Zylinderprojektion, auf Ellipsoid aufliegend, Streifenbreite 3°
- **UTM (Universal Transversal Mercator) – Zylinderprojektion, teilweise innerhalb des Ellipsoids liegende Streifen von 6° Breite**

Bei Angabe geografischer Koordinaten (Länge, Breite) muss die Projektion nicht angegeben werden. Nach Empfehlung der Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) **soll in Deutschland einheitlich ETRS89 und UTM verwendet werden** (in den Listen fett gedruckt).

EPSG-Codes sind weltweit im GIS-Kontext genutzte Codes für Koordinatenreferenzsysteme, zum Beispiel

- EPSG:4326 für WGS84, geografische Koordinaten
- EPSG:4258 für ETRS89, geografische Koordinaten
- **EPSG:25832 für ETRS89, UTM, Zone 32**
- **EPSG:25833 für ETRS89, UTM, Zone 33**
- EPSG:31466 für DHDN, GK, 2. Streifen
- EPSG:31467 für DHDN, GK, 3. Streifen
- EPSG:31468 für DHDN, GK, 4. Streifen
- EPSG:31469 für DHDN, GK, 5. Streifen
- EPSG:2398 für 42/83, GK, 4. Streifen
- EPSG:2399 für 42/83, GK, 5. Streifen

Die folgenden EPSG-Codes wurden nicht bestätigt, sollen daher nicht mehr benutzt werden: 31462 bis 31465 und 31492 bis 31495.
Weitere Informationen zu den EPSG-Codes findet man im Internet zum Beispiel unter
http://www.epsg.org.

GDI-DE

GDI-DE ist ein gemeinsames Vorhaben von Bund, Ländern und Kommunen für den Aufbau einer länder- und ressortübergreifenden Geodateninfrastruktur Deutschland. Im „Applikationsprofil für Web Map Services" sind Aussagen zu Koordinatenreferenzsystemen getroffen worden, die als Standard dienen sollen. Hier werden folgende Systeme favorisiert:

- **EPSG:25832 für ETRS89, UTM, Zone 32 (Pflicht)**
- EPSG:4326 für WGS84, geografische Koordinaten (Pflicht)
- andere (optional)

Im europäischen Kontext wird auf das CEN-WMS-Profil verwiesen, welches folgende CRS favorisiert:

- EPSG:4258 für ETRS89, geografische Koordinaten
- EPSG:3034 für ETRS89, konforme konische Lambertabbildung (LCC)
- EPSG:3035 für ETRS89, Lamberts flächentreue Azimutalabbildung (LAEA)
- EPSG:3038-3051 für ETRS89, zylindrische transversale Merkator-Abb. (LTM), Zo. 26-39

Anwendung zum archäologischen Datenaustausch

Die Empfehlung zur Wahl eines Koordinatenreferenzsystems für den Datenaustausch orientiert sich

- an den Standards der GDI-DE (Geodateninfrastruktur Deutschland), die ihrerseits in den europäischen und weltweiten Kontext eingebunden sind,
- an der Empfehlung der AdV und
- an den tatsächlich verwendeten Systemen.

Eine individuelle Abstimmung wird angeraten. Zur Kennzeichnung der Koordinatenreferenzsysteme werden EPSG-Codes verwendet. Damit ergeben sich für

den Datenaustausch folgende Koordinatenreferenzsysteme (abnehmende Priorität):

- **EPSG:25832 für ETRS89, UTM, Zone 32 (laut AdV und GDI-DE)**
- **EPSG:4326 für WGS84, geografische Koordinaten (laut GDI-DE)**
- **EPSG:25833 für ETRS89, UTM, Zone 33 (laut AdV)**

sowie bis zur Einführung von ETRS89/UTM in den einzelnen Bundesländern

- EPSG:31466 für DHDN, GK, 2. Streifen
- EPSG:31467 für DHDN, GK, 3. Streifen
- EPSG:31468 für DHDN, GK, 4. Streifen
- EPSG:31469 für DHDN, GK, 5. Streifen
- EPSG:2398 für 42/83, GK, 4. Streifen
- EPSG:2399 für 42/83, GK, 5. Streifen
- sonstige nach individueller Abstimmung

Achtung, zum Teil werden gekürzte Koordinaten verwendet (in Brandenburg fehlt bei UTM Zone 33 zum Teil eine führende „3“, damit die Koordinaten ins ALK/EDRS passen).

ADeX, Anlage 4 – Genauigkeiten

Die Angaben zur Genauigkeit sollen es einerseits ermöglichen, die Qualität der Georeferenzen formal einzuschätzen. Andererseits dienen sie als Orientierung für die kartografische Darstellung.

Die Genauigkeit der Koordinatenangaben erfolgt im Sinne eines mittleren Fehlers. Daraus ergibt sich tendenziell, dass der überwiegende Teil (> 68 %) der Koordinatenangaben innerhalb der Grenzen des angegebenen Fehlerwertes liegen sollte. Die Zahl der Ausreißer, deren Fehler größer als der doppelte mittlere Fehler ist, sollte gering (< 5 %) sein (Bitte die Werte in den Klammern wirklich nur tendenziell betrachten!).

Da mathematisch hergeleitete Fehlerwerte in der Regel nicht vorliegen werden, müssen die Fehler anhand anderer Anhaltspunkte abgeschätzt werden. Dazu werden nachfolgend formale Kategorien eingeführt und mit Fehlerwerten belegt. Auch diese Angaben bitte nur als Anhaltspunkt betrachten, da sie nur empirisch zusammengestellt sind.

Für die Vermessung im Feld kann man ausgehend von der zur Verfügung stehenden Ausrüstung die folgenden Fälle annehmen:

- Kategorie: exakt vermessen
  Beschreibung: mit Tachymeter oder ähnlichem Gerät unter Verwendung von Festpunkten der Landesvermessung und mit Fehlerabschätzung bestimmt; spielt eigentlich nur für Grabungs-/Detailpläne eine Rolle
  → mittlerer Fehler: laut Abschätzung (0,1 m oder besser)
- Kategorie: vermessen
  Beschreibung: in der Regel mit Tachymeter oder ähnlichem Gerät unter Verwendung von Festpunkten der Landesvermessung ohne weitere Fehlerabschätzung bestimmt
  → mittlerer Fehler: 1 m
- Kategorie: grob vermessen
  Beschreibung: manuell bestimmt unter Nutzung von markanten Punkten im Gelände, deren Koordinaten ausreichend bekannt sind
  → mittlerer Fehler: 10 m

Bei Benutzung von GPS ergeben sich, abhängig vom genutzten System, nachfolgende Varianten:

- Kategorie: DGPS HEPS
  Beschreibung: über SAPOS Hochpräzisen Echtzeit-Positionierungs-Service (HEPS) bestimmt
  → mittlerer Fehler: 0,05 m
- Kategorie: DGPS EPS
  Beschreibung: über SAPOS Echtzeit-Positionierungs-Service (EPS) bestimmt
  → mittlerer Fehler: 3 m
- Kategorie: GPS normal
  Beschreibung: mit normalem GPS gemessen, ohne Maßnahmen zur Bestimmung oder Verbesserung der Genauigkeit
  → mittlerer Fehler: 20 m
- Kategorie: GPS verfälscht
  Beschreibung: bei künstlicher Verfälschung des Signals (bis 2. Mai 2000 üblich)
  → mittlerer Fehler: 100 m

Bei Digitalisierung aus modernen topografischen Karten (TK) muss man die Ungenauigkeit der TK und den Digitalisierungsfehler berücksichtigen, insgesamt muss man wohl mit 2 mm Fehler rechnen:

- Kategorie: Flurkarten
  Beschreibung: ermittelt/digitalisiert aus modernen Karten um 1:500, 1:1.000 (zum Beispiel Flurkarten)
  → mittlerer Fehler: 2 m
- Kategorie: DGK5
  Beschreibung: ermittelt/digitalisiert aus modernen Karten um 1:5.000 (zum Beispiel DGK5)
  → mittlerer Fehler: 10 m
- Kategorie: TK10
  Beschreibung: ermittelt/digitalisiert aus modernen Karten um 1:10.000 (zum Beispiel TK10, Orthofoto)
  → mittlerer Fehler: 20 m
- Kategorie: TK25
  Beschreibung: ermittelt/digitalisiert aus modernen Karten um 1:25.000 (zum Beispiel TK25)
  → mittlerer Fehler: 50 m
- Kategorie: TK50
  Beschreibung: ermittelt/digitalisiert aus modernen Karten um 1:50.000 (zum Beispiel TK50)
  → mittlerer Fehler: 100 m
- Kategorie: TK100
  Beschreibung: ermittelt/digitalisiert aus modernen Karten um 1:100.000 (zum Beispiel TK100)
  → mittlerer Fehler: 200 m

Bei Verwendung von historischen Karten, Skizzen, Fotos/Schrägaufnahmen oder anderen ungenauen Vorlagen wird deren Ungenauigkeit maßgeblich. Hat man keine weiteren Angaben zur Hand, kann man den oben angegebenen maßstabsbezogenen Wert mit einem Sicherheitsfaktor belegen (2...5...10):

- Kategorie: historische Karte 1:10.000
  Beschreibung: zum Beispiel sächsische Meilenblätter (um 1800) 1:12.000
  → mittlerer Fehler: 50 m
- Kategorie: historische Karte 1:50.000
  → mittlerer Fehler: 250 m

Falls nur indirekte Georeferenz möglich ist, sollten die durchschnittlichen Bezugsobjekte eigentlich der Größe nach beschrieben werden können, falls nicht, hier einige Anhaltspunkte:

- Kategorie: Flurstück
  Beschreibung: Flurstücks(mittel)punkt
  → mittlerer Fehler: 1 km
- Kategorie: Gemarkung
  Beschreibung: Gemarkungs(mittel)punkt
  → mittlerer Fehler: 5 km
- Kategorie: Gemeinde
  Beschreibung: Gemeinde(mittel)punkt
  → mittlerer Fehler: 10 km

Hinweis zur kartografischen Darstellung: Um Fehldeutungen zu vermeiden, sollten die verwendeten Symbole nicht wesentlich kleiner sein als der mittlere Fehler!

Beispiele aus Sachsen:
- *Gemarkungsausdehnungen von 0,2 bis 15 km*
  → mittlerer Fehler circa 3,3 km
- *Gemeindeausdehnungen von 2 bis 21 km*
  → mittlerer Fehler circa 9,1 km
- *Sächsisches Meilenblatt um 1800, Maßstab 1:12.000*
  → mittlerer Fehler 50m

ASPEKTE DER DATENBEREITSTELLUNG

Freie Daten für freie Bürger – Ein Essay über archäologische Daten, die Öffentlichkeit und open data

Ulf Ickerodt

Das Verhältnis von archäologischer Denkmalpflege zu ihren Daten ist ambivalent und durch ein als individualisiertes Besitzstreben zu bezeichnendes Selbstverständnis gekennzeichnet.[1] Wem gehören die erhobenen Daten? Dürfen Daten überhaupt herausgegeben werden? Kann ein Laie, sei es ein Laienforscher oder ein Raumplaner, mit unseren archäologischen Daten überhaupt etwas anfangen? Im Rahmen solcher Diskussionen wird häufig auch der Verweis auf Sondengänger und Raubgräber als nahezu klassisch zu bezeichnendes Argument bemüht. In dem Tourismusland Schleswig-Holstein ist aber auch eine Vielzahl der eingetragenen Denkmale weiträumig ausgeschildert und wird touristisch beworben. Warum also solche Daten geheim halten?

Um sich in diesem Geflecht aus Fachwissen, praktischen Erfahrungen, rechtlichen Rahmenbedingungen und subjektiven Standpunkten zurechtzufinden, lohnt es, einen Schritt aus der Diskussion herauszutreten und einen Blick von außen zu wagen. Dabei zeigt sich sehr schnell, dass diese als fachspezifisch-archäologisch geführte Diskussion Bestandteil eines übergeordneten Diskurses ist. Dieser koppelt wissenschaftliches Selbstverständnis mit gesellschaftspolitischen Grundannahmen, die wiederum vor dem Hintergrund eines gesellschaftlichen Wandels zu sehen sind. Um dieses alles inhaltlich besser durchdringen zu können, werden im Folgenden drei Untersuchungsbereiche herausgestellt.

Den übergeordneten Bezugspunkt bildet die „Öffentlichkeit", mit der archäologische Denkmalpflege als wissenschaftsorientierte Verwaltungsstruktur in einer Wechselbeziehung steht. Damit sind bereits einige grundlegende Erwartungshaltungen festgelegt. Wohl nicht erst mit der Neuzeit setzt die Erkenntnis ein, dass, abstrakt ausgedrückt, die Teilhabe an Informationen – auch in Demokratien – zum einen ein Machtinstrument und zum anderen in den modernen Massengesellschaften gerade auch deswegen ein Demokratiekriterium sein kann. Dem steht die Erkenntnis der wissenschaftlichen Fachwelt „archäologische Denkmalpflege" gegenüber, dass wissenschaftliche Fakten und Daten und damit Wissen immer auch von politischen Entscheidern oder Wirtschaftsunternehmungen missbraucht werden können. Vor diesem Hintergrund entwickelte sich die ebenfalls berechtigte Forderung der Wissenschaft nach der Kontrolle über die eigenen Daten. Diesem Anspruch entgegnet die öffentliche Hand, als Geldgeber und Finanzier archäologischer Forschung und denkmalpflegerischer Arbeit, mit dem ebenfalls berechtigten Anspruch, den Einsatz von Geldmitteln kontrollieren zu wollen. Den übergeordneten Bezugsrahmen dieser Gesamtdiskussion bildet der mit der Industrialisierung einsetzende gesellschaftliche Umstrukturierungsprozess von einer agrarorientierten zu einer kapitalistisch orientierten Industrie- und Dienstleistungsgesellschaft. Bereits in dieser skizzenhaft zusammenfassenden Einführung deutet sich eine Vielschichtigkeit an, der in der Folge nachgegangen werden soll.

Open data und die Informationsgesellschaft

Vielleicht ist das, was wir heute als Informationsgesellschaft bezeichnen, ein Produkt von sich arbeitsteilig organisierenden Gesellschaften. Sicher haben die modernen Kommunikationsmedien und die Entstehung der IT-Technik zu einem gewaltigen Ausbau von Informationen, deren Vernetzung sowie zum individualisierten Gebrauch dieser Möglichkeiten beigetragen. Ich möchte diese Entwicklung an einem Gegensatzpaar verdeutlichen. Während der Aufwand, archäologische Fachdaten aus herkömmlichen Publikationen zusammenzutragen, beträchtlich war (da mit Bibliotheksgängen, Exzerpieren usw. verbunden), finden sich heute bereits beim einfachen „googeln" vielfältige, weltweite wissenschaftliche Publikationen des 19. Jahrhunderts. Viele der hier enthaltenen Daten sind bereits in unterschiedlichen Datenbanken (oder zumindest in Listen erfasst) und können im Rahmen einfacher Datenbankabfragen analysiert werden. Die Auswertungsmöglichkeiten der modernen archäologischen Denkmalpflege/Forschung wird nur durch die Quantität und Qualität der Daten begrenzt und findet ihren Niederschlag in den gegenwärtigen Forschungsvorhaben.

Diese Entwicklung ist vor dem Hintergrund eines anderen Phänomens zu sehen. Bedenkt man, dass der allergrößte Teil der Bevölkerung im 19. Jahrhundert in Landwirtschaft und Industrie durch massiven Arbeitseinsatz gebunden war, so steht dieser Situation heute

eine breite Masse an Wissenschaftsinteressierten, Laienforschern und Fachwissenschaftlern nebst Forschungs- und forschungsnahen Institutionen gegenüber, deren Tätigkeit durch Eigen- oder Drittmittel getragen wird. Es ist wohl davon auszugehen, dass dies alles die Entwicklung unserer Gesellschaft beeinflusst hat und auch weiter beeinflussen wird. Zumal eine weitere, hier anwendbare These davon ausgeht, dass das zu beobachtende Wachstum an Informationen zu einer Produktivkraft unserer modernen Gesellschaftsform geworden ist. Dieses führt zurück zu der eingangs aufgestellten Behauptung, dass die Teilhabe an Informationen und damit auch an Daten ein Demokratiekriterium ist. Der Grund hierfür ist unter anderem, dass wissensbasierte Partizipation ein grundlegender Qualitätsfaktor demokratisch legitimierter politischer Steuerung ist, die mit einer Verwissenschaftlichung aller Lebensbereiche einhergeht. In diesem Zusammenhang steht einerseits das Streben nach einer wissensbasierten Gestaltung von Umwelt und Gesellschaft bis hin zu der des individuellen Zusammenlebens. Zum anderen wird diese Idee von dem Leitbild getragen, dass Informationsfreiheit zur Entfaltung der offenen Gesellschaft beiträgt und individuelle Bildungs- und Entfaltungschancen vergrößert. Diese Grundausrichtung findet sich in chronologischer Reihe in unterschiedlicher Ausrichtung als Gedanke der Freiheit der Informationen in der *Allgemeine Erklärung der Menschenrechte* vom 10.12.1948 in Artikel 19, in der *Konvention zum Schutz der Menschenrechte und Grundfreiheiten* vom 01.11.1950 in Artikel 10 „Freie Meinungsäußerung", in dem *Internationalen Pakt über bürgerliche und politische Rechte* (ICCPR) vom 19.12.1966 in Artikel 19 (2), in der *Amerikanischen Konvention der Menschenrechte* vom 22.11.1969 in Artikel 13, in der *Banjul Charta der Menschenrechte und Rechte der Völker* (Menschenrechtsschutz in Afrika) vom 27.06.1981 in Artikel 9, im *Übereinkommen über die Rechte des Kindes* (CRC) vom 20.11.1989 in Artikel 13, in der *Internationalen Konvention zum Schutz der Rechte aller Wanderarbeitnehmer und ihrer Familienangehörigen* (MWC) vom 18.12.1990 in Artikel 13, in der *Konvention der Gemeinschaft Unabhängiger Staaten über die Rechte und Grundfreiheiten der Menschen* vom 26.05.1995 in Artikel 11 (1) oder der *Charta der Grundrechte der Europäischen Union* vom 18.12.2000 in Artikel 11 „Freie Meinungsäußerung und Informationsfreiheit".

Dabei wird der Gedanke des Austauschs von Informationen (einschließlich des Schutzes personenbezogener Daten) insbesondere in der genannten *Charta der Grundrechte der Europäischen Union* sowie der der transnationalen Zusammenarbeit in dem ebenfalls bereits genannten *Übereinkommen über die Rechte des Kindes* (CRC) sowie der Schlussakte der *Konferenz über Sicherheit und Zusammenarbeit in Europa* vom 01.08.1975 in Punkt 2 „Informationen" und besonders in der *Erklärung von Rio zu Umwelt und Entwicklung* (Rio-Deklaration) vom 14.06.1992 im Grundsatz 10 betont, auch wenn hier digitale Informationen und Daten nicht explizit behandelt werden. Dieses wurde in anderen Dokumenten und Gesetzen nachgeholt (Göldner S. 34). Hinzu kommen Projekte, wie das sich aus der trilateralen Nordsee Kooperation heraus angestoßene, von der EU über das Interreg IIC Nord-Programm kofinanzierte Projekt Lancewad (12/1999 bis 10/2001; http://www.waddensea-secretariat.org/lancewad/report.html), in dessen Rahmen umfangreiche Datensätze auch auf länderübergreifender Ebene zusammengetragen werden und zum Beispiel für die Öffentlichkeit zur Verfügung gestellt werden sollen.

Vor diesem Hintergrund kann von archäologischer Seite von einem Anspruch der Öffentlichkeit an archäologischen Daten ausgegangen werden. Die Frage wäre dann, um welche Daten es sich handeln dürfte. Am sinnvollsten erscheint es, an dieser Stelle auf Grundsatz 10 der *Erklärung von Rio zu Umwelt und Entwicklung* zurückzukommen. Hier wird klargestellt, dass bei Umweltfragen, und das archäologische Erbe kann mit Fug und Recht als Qualitätsmerkmal einer intakten Umwelt angesehen werden, die Bürger zu beteiligen und dafür angemessen zu informieren sind, um geeignete Entscheidungen herbeiführen zu können. Eine beispielgebende Bedeutung, weil politisch legitimiert und administrativ verankert, kommt dem auf Lancewad folgenden Projekt LancewadPlan (Integrated Landscape and Cultural Heritage Management and Development Plan for the Wadden Sea Region) zu, das ebenfalls durch das Interreg IIIB Nordseeprogramm gegenfinanziert wurde (7/2004 bis 6/2007; http://www.lancewadplan.org/).

Open data und die archäologische Denkmalpflege

Das Konzept der open data sowie alle anderen IT-Möglichkeiten stellen für die archäologische Denkmalpflege eine inhaltliche Herausforderung dar, die mit dem eigenen fachlichen Selbstverständnis und den sich daraus entwickelnden Methoden zu tun hat. Die archäologische Denkmalpflege hat sich als administrativer Arm der universitären archäologischen Forschung sowohl in methodischer als auch in wissenschaftsethischer Hinsicht entwickelt und hat ihre Wurzeln – trotz älterer Vorläufer – unverkennbar im 19. Jahrhundert. In dieser Zeit wurden ihre eigentlich bis heute gültigen inhaltlichen Grundausrichtungen und Fragestellungen angelegt, in deren Fokus archäologische Funde und Befunde als übergeordnete

Ansatzpunkte stehen. Das primäre Ziel ist neben dem wissenschaftlichen Erkenntnisgewinn die Aufgabe, die „Wahrheit" über die menschliche Geschichte auf Basis ihrer wissenschaftlichen Fakten zu schreiben. Wissenstheoretische Basis ist hier der Positivismus. In der archäologischen Forschung bedeutet dieser Ansatz, dass historische Aussagen empirisch am Material überprüft werden können. In praktischer Hinsicht muss die archäologische Forschung Fundstellen und Funde erschließen und neue Forschungsmethoden zu ihrer Auswertung generieren. Dies geschieht zum Teil durch Berufswissenschaftler, im Rahmen akademischer Qualifizierungen oder auch durch Laienforscher, seien es Sammler, Detektorengänger und andere. Die hier im Laufe der Jahrzehnte gewonnenen Fakten werden in Landesaufnahmen erfasst. Dabei werden sie einerseits als absolut und wertneutral angesehen, andererseits ist das Wissen um die Fehlerhaftigkeit dieser Datenbestände ein „alter Hut". Ein wichtiger Aspekt ist hier die Urheberschaft, denn das Wissen um den Datenlieferanten ist neben Datum- und Ortsangabe ein wichtiges Element der Authentizität der erhobenen Daten. Wem gehören aber diese Daten?

Die Antwort erscheint hier relativ einfach, wenn die Daten von Mitarbeitern der Landesarchäologien erhoben werden. Auch wenn der einzelne Mitarbeiter der Urheber seiner Daten ist, so besitzt doch der Arbeitgeber die Verwertungsrechte an diesen Daten. Sie ist auch einfach, wenn die Denkmalfachämter im Rahmen der Raumplanung (zum Beispiel Strategische Umweltprüfung [SUP], Umweltverträglichkeitsprüfung [UVP] etc.) die Übergabe von Daten an Außenstehende (zum Beispiel Grabungsfirmen, externe Fachgutachter) einfordern können. Die Antwort ist im Gegensatz dazu nicht so einfach, wenn Daten durch externe Forschungsinstitutionen, archäologische Museen, Universitäten oder von privater Seite (zum Beispiel Laienforscher) erhoben werden. Hier steht den konkreten rechtlichen Regelungen eine übergeordnete Entwicklung im Wissenschaftsbetrieb gegenüber, die unter dem *Konzept der Kommunalität* bekannt geworden ist. Es geht auf den Soziologen Robert Merton (Social theory and social structure, 1949) zurück und schließt wissenschaftliche Eigentumsrechte an wissenschaftlichem Wissen aus. Wenn wissenschaftliche Fakten/Daten also keinen Besitzer haben, dann steht einem Informationsaustausch nichts mehr entgegen.

Allerdings hat diese geforderte Datenfreiheit auch ihre Schattenseiten und bringt Wissenschaftler unter Umständen in Gewissenskonflikte, wenn die Anwender (zum Beispiel Politiker, Wirtschaft oder gar Sektierer) sich an Daten oder wissenschaftlichen Ergebnissen bedienen. Daher war die wissenschaftliche Welt insbesondere nach den Atombombenabwürfen bemüht, Kontrolle über die eigenen Daten zu bekommen. Diese politische Instrumentalisierung kennt die Archäologie auch. So haben zunächst die drei großen europäischen Nationalstaaten, England, Frankreich und das Deutsche Reich versucht, seit Ende des 18. Jahrhunderts ihr hohes Alter mit Bodenfunden zu belegen. Triebfeder dieser Entwicklung war die Notwendigkeit der jungen Nationalstaaten, starke Identifikationspunkte zu liefern. Vor dem Hintergrund dieser Erfahrungen erscheint es angebracht, die Polarität von freiem Datenzugang und dem Missbrauch dieser Daten bei der Schöpfung neuer Geschichtsmythen im Hinterkopf zu haben.

Open data und die archäologisch-denkmalpflegerische Praxis

Derzeit ringen die unterschiedlichen Institutionen der archäologischen Denkmalpflege um eine gemeinsame Position im Hinblick auf den Austausch und die Freigabe von Daten. Stichwort ist der Begriff „open data", der, durch Politik und öffentlichen Druck gewünscht, den Rahmen der Diskussion vorgibt. Aus archäologischer Sicht müssen im Rahmen der thematischen Erörterung, neben dem zuvor dargestellten Bezugsrahmen, drei analytische Ebenen unterschieden werden.

A) Den übergeordneten Rahmen bildet die juristische Dimension. Diese reicht von der Definition des Begriffs „Datum", das heißt: eine Information außerhalb einer Textform (Bilder, raumbezogene Daten, Statusdaten usw.), bis hin zu Fragen der Zugänglichkeit von Daten, die in unterschiedlichen Gesetzen geregelt sind.

B) Dem gegenüber steht die administrative Ebene. Diese ist in der modernen archäologischen Denkmalpflege durch zwei Unterebenen geprägt.

Die erste Unterebene umfasst die Herstellung von Archivgut und damit Daten zu geleisteter archäologischer Feldarbeit (Prospektion, Ausgrabung usw.), die der Forschung und Öffentlichkeit verfügbar gemacht werden sollen. Leitbilder sind hier die Stichworte Bürgerfreundlichkeit und transparentes Verwaltungshandeln, die sich parallel zu dem Konzept des mündigen Bürgers entwickeln. Hier treffen sich das Recht auf informelle Selbstbestimmung und die Veröffentlichungspflicht der öffentlichen Stellen.

Problematisch ist dieser Bereich immer dann, wenn kommerzielle Interessen ins Spiel geraten und die zur Verfügung gestellten Daten einen Mehrwert erzeugen, der nicht der Allgemeinheit zugutekommt und dessen Ausbeutung vielmehr durch Partikularinteres-

sen geprägt ist. Aus einer forschungsgeschichtlichen Perspektive stehen hier die sogenannten Pfahlbaufischer des 19. Jahrhunderts ihren Nachfolgern, den Raubgräbern und illegalen Sondengängern, gegenüber. Letztere sind nur besser ausgestattet und können, neben eigener Recherche- und Forschungsarbeit, gegebenenfalls auf einen enormen Datenpool zurückgreifen.

Ein anderes, deutlich weniger problematisches Beispiel ist das Verlagswesen. Die einzelnen Verlage nutzen gerne den Datenpool der Denkmalpflegeinstitutionen insbesondere in Form von Bildern zur Wertschöpfung und die Denkmalpflege im Gegenzug die damit einhergehende Öffentlichkeitsarbeit in Büchern, Zeitschriften und in den neuen Medien. Hier stehen kommerzielle Interessen der Informationspflicht einerseits und den Vermarktungsinteressen der informationsführenden Institutionen andererseits gegenüber. Zu diskutierende Aspekte sind hier weiterhin die Frage nach einer Aufwandsentschädigung oder nach zu entrichtenden Gebühren. Das bekannteste Beispiel für die Umsetzung eines wirtschaftlichen Eigeninteresses ist die Vermarktung der sogenannten Himmelsscheibe von Nebra durch das Landesamt für Denkmalpflege und Archäologie mit Landesmuseum für Vorgeschichte/Halle.

Anders liegt der Fall im touristischen Segment. Hier ist es geradezu sinnvoll, archäologische Daten in den Bereich Tourismusplanung, Inwertsetzung und Vermittlung einzusetzen, obwohl dieser Einsatz zwar nicht zum denkmalpflegerischen Kerngeschäft gehört, die Arbeit aber dennoch sowohl zum Erhalt der Denkmale und ihres Raumbezugs als auch einer damit einhergehenden regionalen Wertschöpfung führen kann. Allerdings erscheint es hier fraglich, inwieweit archäologische Grunddaten nebst Raumbezug hierfür ausreichend sind und der Datenveredlung einschließlich der damit einhergehenden Qualitätskontrolle bedürfen.

Die zweite Unterebene umfasst den Aspekt der archäologischen Beteiligung an Raumplanungsprozessen als einem wichtigen Instrument der modernen denkmalpflegerischen Praxis, die der Vermeidung und Minderung von negativen Auswirkungen dient. Auch dieser Bereich profitiert von der freien Zugänglichkeit planungsrelevanter Daten. Grundsatz ist hier aufseiten der archäologischen Denkmalpflege die Darstellung des berechtigten Belangs bei der Wahrnehmung der Trägerschaft öffentlicher Belange, den wiederum die Aspekte Vorsorge-, Kooperations- und Nachhaltigkeitsprinzip prägen.

Das *Vorsorgeprinzip* ist ein Grundsatz des modernen Umweltschutzes und beruht auf frühzeitigem und vorausschauendem Handeln. Es hat zum Ziel, möglichst frühzeitig eventuelle Belastungen und Gefahren zu identifizieren und diese bereits im Vorfeld von Maßnahmen abzuwenden (zum Beispiel BauGB §§ 1 Abs. 5, 1a, BImSchG § 50, ROG § 2 Abs. 2 Nr. 6). Aus Sicht der archäologischen Denkmalpflege soll so zu einem schonenden Umgang mit dem archäologischen Erbe beigetragen werden.

Das *Kooperationsprinzip* ist ein Verfahrensgrundsatz, der im Umweltschutz Beachtung findet und auf der Idee basiert, dass der Staat nicht in erster Linie Zwang zur Verwirklichung der archäologisch-denkmalpflegerischen Ziele auf Basis seiner rechtlichen Möglichkeiten ausüben sollte, sondern im Sinne des Integrierten Küstenzonenmanagements (IKZM) agieren sollte. Dies kann auf unterschiedlichen Ebenen und mit verschiedenen Zielsetzungen erfolgen und ist zum Beispiel in der internationalen Zusammenarbeit in § 9 UVPG geregelt. Auf der Arbeitsebene sollte eine Form von (gegenseitiger) Information und Konsultation gefunden werden, die gleichermaßen Verhandlungen und Abstimmungsprozesse beinhaltet und an deren Ende eine schriftliche Vereinbarung steht und den Vollzug der getroffenen Vereinbarungen regelt. Auf einer übergeordneten Ebene kommt das Kooperationsprinzip in der Anhörung beteiligter Kreise, von Verbänden usw. vor dem Erlass von Rechtsverordnungen und Verwaltungsvorschriften zum Ausdruck.

Das *Nachhaltigkeitsprinzip* ist spätestens seit dem bereits angeführten Gipfel von Rio als Grundsatz im Umgang mit den Umweltressourcen anzusehen. Im Raumordnungsgesetz (§ 1 Abs. 2 Nr. 6) ist dieses Ziel der nachhaltigen Raumplanung als Grundsatz der Raumordnung formuliert und findet sich darüber hinaus auch in anderen Gesetzen.

Vonseiten des an Raumplanungsprozessen beteiligten Partnerfelds wird unabhängig davon die Forderung nach einer gewissen Vollständigkeit sowie der damit verbundenen Integrität des Datenpools erhoben, um zu einer verlässlichen Planungsgrundlage zu kommen. Hierzu muss die archäologische Denkmalpflege allerdings Antworten auf die Anforderungen der unterschiedlichen Kontextualisierungsebenen finden (Interessensgebiet, Pufferzone, Welterbestätte, Fundstelle, eingetragenes Denkmal), deren Qualität zwischen wissenschaftlichem Selbstanspruch und juristischer Überprüfbarkeit anzusiedeln ist.

C) Zu den behandelten juristischen und administrativen Ebenen kommt die wissenschaftliche, die ihrerseits durch unterschiedliche Aspekte geprägt ist. Im Zentrum stehen hier allerdings die Fragen der Datenerhebung und -bewertung, der Datenhaltung sowie die der Datenweitergabe.

Der Aspekt der Datenerhebung und -bewertung ist seinerseits durch unterschiedliche Ebenen geprägt, die unabhängig von ihrer juristischen Komponente die Antworten auf Fragen nach Urheberschaft, Datenqualität und Stabilität sowie Datenzugänglichkeit erfordern. Wichtigste Triebfedern sind hier die Grundauffassung, dass wissenschaftliche Daten Allgemeingut sind beziehungsweise dass wissenschaftliche Fakten keiner Urheberschaft unterliegen können sowie dass sie eine freie Forschung ermöglichen und dadurch für Innovation sorgen. Diese Forderung bildet eine übergeordnete Diskussion ab, die in der sogenannten Berliner Erklärung (vgl. Göldner S. 36) ihren Niederschlag gefunden hat.

Allerdings ist auch diese Ebene mit Fragen nach den qualitativen Eigenschaften von Daten verbunden. Hier steht vor allem die der Authentizität der Daten im Vordergrund. Aber auch die Frage nach der Datenintegrität ist gerade im Hinblick auf die Altdatenbestände und deren Überführung in Datenbanken zu stellen und mit Fragen der Kontextualisierung dieser Daten zu verbinden. Dieses gilt umso mehr, als dass eine Prämisse der archäologischen Forschung die These ist, dass die Fakten/Daten für sich sprechen. Hier schwankt der Wirkkreis zwischen einer unbeabsichtigten Verfälschung der ein Geschichtsbild erzeugenden Datengrundlage bis hin zu einer bewussten, manipulierenden Ausnutzung eben dieser Datengrundlage.

Abschlussbetrachtung

Die neuen Medien haben die archäologische Denkmalpflege in praktischer und methodischer Hinsicht überrollt. Daher ist neben allen technischen und anwendungsorientierten Fragen auch die nach dem Verhältnis der archäologischen Denkmalpflege beziehungsweise Forschung zu dieser Entwicklung zu stellen. Allerdings ist hier weder die ablehnende Haltung der Skeptiker noch das blinde Vertrauen der Befürworter angebracht. Vielmehr erscheint eine grundsätzliche Debatte der Rahmenbedingungen vonnöten, die auf einer gezielten Stärken- und Schwächen-Analyse der juristischen, administrativen und fachlichen Grundlagen basiert. (Die technische Ebene wird an dieser Stelle einmal ausgeklammert!) Diese Diskussion erscheint umso nötiger, als dass ein exponentielles Datenwachstum einem sich stärker ausdifferenzierenden Nutzerkreis gegenübersteht, der vom interessierten Laien über den Raumplaner bis hin zu internationalen Forschern reicht. Dabei muss einerseits von den datenführenden Stellen die Qualitätssicherung vor dem Hintergrund von Stellenabbau und einer zurückgeschraubten finanziellen Ausstattung gewährleistet sein. Andererseits muss ein Missbrauch der Daten in allen Ebenen ausgeschlossen werden können. Um die Dinge nicht zu einfach zu gestalten, muss eine solche Diskussion zusätzlich vor dem Hintergrund der unterschiedlichen adminstrativ-juristischen Grundlagen der jeweiligen Institution erfolgen. Daher wird es bei der Umsetzung des „open data"-Konzepts in der archäologischen Denkmalpflege zwar keinen Königsweg geben, dafür aber bessere und schlechtere Lösungen. Auf diesem Weg bedarf es allerdings auch der Einigkeit über die Diskussionsgrundlage sowie den Diskussionsrahmen. In dieser Hinsicht versteht sich dieser Beitrag als ein Versuch, die unterschiedlichen Facetten und Aspekte in Relation zu bringen, ohne allerdings eine inhaltliche Bewertung vorzunehmen.

Den Rahmen bildet ein Beziehungsgeflecht aus den dargestellten unterschiedlichen Ansprüchen. Grundlegend erscheint mir hier die symbolische Bedeutung des „open data"-Konzepts als gesellschaftspolitischem Leitbild im Rahmen des europäischen und weltweiten Demokratisierungsprozesses, auch wenn die archäologischen Daten in der Gesamtdiskussion sicherlich nur eine untergeordnete Rolle spielen. Diesem kulturhistorisch verankerten Leitbild steht die berechtigte Forderung der archäologischen Denkmalpflege auf Schutz ihrer eigenen Interessen an den sensiblen Daten gegenüber, die zum einen aus der Erfahrung des Missbrauchs der Daten und zum anderen aus der Erfordernis eines verlässlichen Umgangs mit den Daten resultieren. Hier ist vor allem die Datenqualität zu nennen, die immer auch der qualifizierten Präsentation und Bewertung bedarf, wenn die archäologische Denkmalpflege zuverlässiger Partner in Raumplanungsprozessen sein will. Das reine Bereitstellen von Daten erscheint dafür nicht ausreichend und daher auch nicht zielführend. Vergleichbares gilt für den Umgang mit Partnern der universitären und außeruniversitären Forschung sowie mit außerwissenschaftlichen Partnern, insbesondere der Laienforschung, die an der denkmalpflegerischen Arbeit partizipieren wollen. Hier treffen unterschiedliche Vorstellungen zu Datentiefe und Datenqualität aufeinander. So kommt die archäologische Denkmalpflege im Bereich der Raumplanung sicherlich mit einer Kopfdaten- beziehungsweise Metadatenebene aus, deren wissenschaftlich gesehene Minimalaussage im Bereich „hier ist eine Fundstelle" liegt und inhaltlich nicht weiter differenziert werden muss. Dem gegenüber stehen Forschungsansätze, die denkmalpflegerisch relevante Prognosen erstellen wollen, um die Qualität archäologisch-denkmalpflegerischer Aussagen im Rahmen von Raumplanungsprozessen zu verbessern. Hier wird die genannte Metadatenebene sicherlich nicht ausreichen.

Vor diesem zuvor dargestellten Hintergrund wird, insgesamt gesehen, das Verhältnis von archäologischer

Denkmalpflege zu ihren Daten im Detail ambivalent bleiben und mit zunehmender Datentiefe durch ein steigendes als individualisiertes Besitzstreben zu bezeichnendes Selbstverständnis gekennzeichnet bleiben, insbesondere dann, wenn zwischen dienstlich und privat erhobenen Daten unterschieden werden kann. Dessen ungeachtet erscheint hier die Maxime „Altruismus ist der beste Eigennutz" als geeignetes archäologisch-denkmalpflegerisches Leitbild.

Anmerkung

[1] An dieser Stelle möchte ich mich für die spannenden Diskussionen im Kollegenkreis, insbesondere im Rahmen unserer Schleswiger Treffen, bedanken. Mein besonderer Dank gilt Herrn Göldner (Dresden) und Herrn Maluck (Schleswig) dafür, dass sie aus meinen Überlegungen einen lesbaren (so hoffe ich) Text gemacht haben.

Literatur

Bauerochse, A., Haßmann H. und Ickerodt, U. (2007): Kulturlandschaft. administrativ – digital – touristisch. (= Initiativen zum Umweltschutz 67). Berlin: Erich-Schmidt-Verlag.

Ickerodt, U. (2010): My Story – your Story: Contemporary Cultural Heritage Management as Sphere of Social-Interaction. In: T. Bloemers, H. Kars, A. van der Valk und M. Wijnen (Hrsg.), The Cultural Landscape Heritage Paradox. Protection and Development of the Dutch Archaeological-Historical Landscape and its European Dimension. Amsterdam: Amsterdam University Press, 351–362.

Ickerodt, U. (2011): Meine Geschichte – Deine Geschichte: Quo Vadis archäologische Denkmalpflege? Österreichische Zeitschrift für Kunst und Denkmalpflege LXV/3, 240–251.

Über die Freiheit der Daten

Reiner Göldner

Einführung

Wir leben in einer Informationsgesellschaft, Informationen spielen eine bedeutende Rolle in unserem Leben. Soweit ist alles klar, aber was folgt daraus? Die Meinungen sind zwiespältig. Frei zugängliche Informationen beflügeln die Entwicklung enorm, das ist unbestritten. Betrachtet man verschiedene Aspekte der freien Bereitstellung und Nutzung von Informationen jedoch näher, so kann man erkennen, dass es durchaus Vorteile und Nachteile gibt, und dass diese nicht immer gleich verteilt sind. Dies ist bei archäologischen Informationen nicht anders, daher sollen einige dieser Aspekte nachfolgend näher betrachtet werden.

Allgemeine rechtliche Aspekte

Rezipientenfreiheit[1] und Wissenschaftsfreiheit sind Grundrechte und beschreiben zuerst einmal die Freiheit, Informationen zu erlangen und für wissenschaftliche Zwecke zu nutzen:

„Jeder hat das Recht, seine Meinung in Wort, Schrift und Bild frei zu äußern und zu verbreiten und sich aus allgemein zugänglichen Quellen ungehindert zu unterrichten." (GG[2] Art. 5, Abs. 1)

„Kunst und Wissenschaft, Forschung und Lehre sind frei." (GG Art. 5, Abs. 3)

Eine Pflicht, Informationen zugänglich zu machen, ergibt sich daraus jedoch nicht. Veröffentlichungspflichten ergeben sich meist durch spezielle Gesetze, wie zum Beispiel das Umweltinformationsgesetz oder die Geodatenzugangsgesetze. Außerdem gehen deutliche Tendenzen in Richtung Informationsfreiheit[3] (einem weitgehend freien Zugang zu Informationen der öffentlichen Verwaltung), entsprechende Gesetze gibt es inzwischen weltweit.

Die beschriebenen Freiheiten finden jedoch ihre Grenzen in anderen Grundrechten, zum Beispiel im Recht auf informationelle Selbstbestimmung (Datenschutzgesetz) und im Urheberrechtsschutz, welche die Möglichkeit, Informationen zugänglich zu machen, einschränken. Nicht zuletzt gilt das Grundrecht auf Eigentum (Art. 14 GG) und zu den Befugnissen des Eigentümers (BGB[4] § 903) gehört eben auch, über eine Veröffentlichung nach freiem Belieben zu entscheiden, diese also auch abzulehnen.

Informationsfreiheit

„Die Offenlegung aller allgemein verfügbaren Dokumente, die sich im Besitz des öffentlichen Sektors befinden – die nicht nur die Politik, sondern auch die Justiz und die Verwaltung betreffen – ist ein grundlegendes Mittel zur Erweiterung des Rechts auf Wissen, das wiederum ein Grundpfeiler der Demokratie ist." Dieser Auszug aus der Informationsweiterverwendungsrichtlinie[5] der EU fasst die Motivation der Informationsfreiheit gut zusammen.

Aufgrund der Bedeutung transparenten Verwaltungshandelns und der Verfügbarkeit angemessener Informationen zu Themen von öffentlichem Interesse in einer pluralistischen, demokratischen Gesellschaft setzt sich seit einiger Zeit ein Paradigmenwechsel bei der Gewährung von Zugang zu behördlichen Informationen immer weiter durch. Bisher galt zumeist das Prinzip, dass behördliche Informationen grundsätzlich nicht öffentlich sind („Amtsgeheimnis"), solange kein explizit rechtlich geregelter Ausnahmeanspruch besteht. Dieser Anspruch musste dargelegt und begründet werden. Die Informationsfreiheitsgesetze dagegen gewähren grundsätzlich einen freien (voraussetzungslosen) und gleichberechtigten Zugang zu behördlichen Informationen, wenn nicht schützenswerte oder höherwertige Interessen dagegensprechen (zum Beispiel Datenschutz, gegebenenfalls auch Denkmalschutz). Die Nachweispflicht zu einem Ausnahmetatbestand obliegt nun der Behörde.

Bei der Verbreitung des Prinzips der Informationsfreiheit spielt die Europäische Union eine wichtige Rolle. Die grundsätzliche Empfehlung des Europarates Nr. 854 (1979) „on access by the public to government records and freedom of information"[6], präzisiert durch die Empfehlung Nr. R (81) 19 aus dem Jahr 1981[7] und durch die Empfehlung Rec(2002)2 aus dem Jahr 2002[8], regt die Etablierung eines Systems der Informationsfreiheit an: „... invite member states … to introduce a system of freedom of information". Dieser Empfehlung ist bisher eine Vielzahl von Staaten mit der Schaffung von Informationsfreiheitsgesetzen nachgekommen, in Deutschland neben dem Bund (2005) derzeit elf Bundesländer[9]. Die USA haben mit dem „Freedom of Information Act" übrigens schon seit 1966 ein entsprechendes Gesetz.

In einigen Bereichen gibt es spezielle Regelungen und Gesetze, die den Gedanken der Informationsfreiheit widerspiegeln. Seit 1994 gibt es in Deutschland Um-

weltinformationsgesetze (UIG) des Bundes[10] und der Länder. „Zweck dieses Gesetzes ist es, den rechtlichen Rahmen für den freien Zugang zu Umweltinformationen bei informationspflichtigen Stellen sowie für die Verbreitung dieser Umweltinformationen zu schaffen." (UIG Bund, §1). Außerdem regeln seit 2009 Geodatenzugangsgesetze beziehungsweise Geodateninfrastrukturgesetze des Bundes und der Länder den Aufbau einer Geodateninfrastruktur sowie den Zugang zu Geodaten in Deutschland, maßgeblich beeinflusst durch die INSPIRE-Richtlinie.

INSPIRE

Nicht nur in Navigationssystemen, sondern in nahezu allen Lebensbereichen spielen Geodaten eine wachsende Rolle zur Beantwortung räumlich bezogener Fragestellungen. In einer möglichst breiten Verknüpfung und Vernetzung dieser Daten, einer sogenannten Geodateninfrastruktur (GDI), wird ein immenses Informationspotenzial mit positiven Effekten für Politik, Wirtschaft und öffentliches Leben gesehen.

Bezüglich Geodaten spielt die INSPIRE-Initiative[11] der Europäischen Kommission eine wichtige Rolle. Sie stützt sich wesentlich auf die Empfehlungen zur Informationsfreiheit und hat das Ziel, eine Geodateninfrastruktur in der EU zu etablieren. Rechtsverbindlich ist dazu die INSPIRE-Richtlinie[12], welche den Mitgliedsländern verbindliche Vorgaben zum Aufbau der GDI stellt. Sie enthält die umzusetzenden Themenkomplexe für Metadaten und Geodaten (Anhänge 1 bis 3), Angaben zur technischen Umsetzung, Realisierungstermine und anderes mehr. Inzwischen existiert außerdem eine Vielzahl zu berücksichtigender Durchführungsbestimmungen und technischer Spezifikationen.

Auch die INSPIRE-Richtlinie räumt schützenswerten und höherwertigen Interessen Ausnahmerechte ein. Entsprechend Artikel 13 „können die Mitgliedstaaten den Zugang der Öffentlichkeit … beschränken, wenn dieser Zugang auf die internationalen Beziehungen, die öffentliche Sicherheit oder die nationale Verteidigung nachteilige Auswirkungen hätte" sowie, und dies ist für die Archäologie wichtig, „wenn dieser Zugang nachteilige Auswirkungen hätte auf: …" (hier folgen mehrere Tatbestände, deren letzter uns jedoch besonders interessiert) „… den Schutz der Umweltbereiche, auf die sich die Informationen beziehen, wie zum Beispiel die Aufenthaltsorte seltener Tierarten." Dieses Ausnahmerecht kann zweifelsfrei auch auf die Archäologie übertragen werden.

Im Anhang 1 der INSPIRE-Richtlinie werden die primär zur Verfügung zu stellenden Themen genannt. Unter diesen Themen befinden sich die „Schutzgebiete", die auch für den archäologischen Denkmalschutz bedeutsam sind. Schutzgebiete werden wie folgt erläutert: „Gebiete, die im Rahmen des internationalen und des gemeinschaftlichen Rechts sowie des Rechts der Mitgliedstaaten ausgewiesen sind oder verwaltet werden, um spezifische Erhaltungsziele zu erreichen."

Die „Data Specifications on Protected Sites" stellen erstmals auch einen direkten formellen Zusammenhang zwischen archäologischem Denkmalschutz und INSPIRE her. Mögliche Konsequenzen daraus wurden im Verband der Landesarchäologen intensiv beraten. Im „Meinungsbild im Verband der Landesarchäologen zum Thema Archäologischer Denkmalschutz und INSPIRE-Schutzgebiete"[13] liest man dazu: „Der Verband der Landesarchäologen, insbesondere dessen Kommission Archäologie und Informationssysteme …, haben sich intensiv mit dem Thema INSPIRE und dessen Verhältnis zum archäologischen Denkmalschutz beschäftigt. Es wurde eine umfangreiche Sammlung von Kriterien erarbeitet und diskutiert, die in diesem Zusammenhang eine Rolle spielt. Anhand dieser Kriterien können individuelle Entscheidungen zu dem Grad der Beteiligung an INSPIRE sowie gegebenenfalls zur Gestaltung der eigenen Beiträge getroffen werden." (Meinungsbild[14], S. 1) sowie: „Ein Minimalkonsens unter den archäologischen Denkmalschutzbehörden in Deutschland soll bezüglich INSPIRE gewährleisten, dass durch Entscheidung einzelner Behörden beziehungsweise Bundesländer keine Interessen anderer Behörden beziehungsweise Bundesländer beeinträchtigt werden." (Meinungsbild[15], S. 3). Zusammenfassend wird festgestellt: „INSPIRE wird mehrheitlich als Chance betrachtet, um die Wahrnehmung der Archäologie in der Öffentlichkeit zu verbessern und die öffentlichen Belange des Denkmalschutzes besser wahrnehmen zu können." (VLA-Website, Kommission Archäologie und Informationssysteme – AG Rechtliche Aspekte[16]).

Open Data

Der Gedanke des freien Zugangs zu Wissen und Informationen ist alt. Mit dem Internet erlebt er jedoch gerade einen unvergleichlichen Aufschwung. Hier lassen sich, mit vergleichsweise geringem Aufwand, Informationen weltweit verbreiten und veröffentlichen. Portale wie Google® ermöglichen es, die Informationsflut einigermaßen zu beherrschen, und Nachschlagewerke wie Wikipedia bündeln das Wissen vieler freier Autoren und regen zu eigenen Beiträgen an. Mit diesen Mitteln ist der Zugang zu Informationen so einfach wie nie zuvor.

Im Vergleich zur „Informationsfreiheit", die sich an den öffentlichen Sektor der Politik und Verwaltung wendet, richtet sich die Open-Data-Bewegung an alle Besitzer und Anbieter von Informationen. Offene Daten, so die Argumentation, liegen im Interesse der Allgemeinheit, weil sie Forschungen und vorteilhafte Entwicklungen fördern (Wikipedia). Das Rad muss eben nicht neu erfunden werden, wenn der Bauplan frei verfügbar ist. Ein weiterer Vorteil von Open Data sind die relativ einfachen Regeln für die Nutzung, wie sie zum Beispiel in den „Creative Commons"[17]-Lizenzen fixiert sind.

Der Wunsch nach Zugang zu offenen Daten ist gut nachvollziehbar. Die Motivation, offene Daten zu liefern, lässt sich für öffentlich finanzierte Projekte noch recht einfach aufbauen, sie findet jedoch schnell ihre Grenzen in privatwirtschaftlichen Interessen. Insofern hat Open Data seinen Schwerpunkt, neben der Informationsfreiheit des öffentlichen Sektors, vor allem im Bereich des Allgemeinwissens und der Wissenschaft, deren Prinzipien seit jeher auf Öffentlichkeit beruhen. Den positiven Effekten von Open Data stehen aber auch einige Probleme gegenüber, wenn zum Beispiel Rechte Dritter betroffen sind (Urheber-, Marken- und Patentrechte, Datenschutz etc.). Außerdem muss man gegebenenfalls mit mangelnder Qualität rechnen, nicht immer sind die Autoren profunde Kenner der Materie. Außerdem stehen die Daten natürlich auch nicht nur der Nutzung für den rechten Zweck zur Verfügung, sondern auch dem Missbrauch. Bei der Umsetzung von Open Data-Projekten sollte man all diese Aspekte jedoch als Ausnahmetatbestände angemessen berücksichtigen können.

Information als Wert

Es käme wohl keiner ernsthaft auf die Idee, die Branchenführer der Automobilindustrie aufzufordern, ihre neuesten Modellplanungen als Open Data zu veröffentlichen, die Erfolgschancen stünden allzu schlecht. Informationen sind eben auch Werte, die die eigene Position gegenüber der Konkurrenz stärken. Und Verhandlungen werden eben oft genug durch vorhandene beziehungsweise nicht vorhandene Informationen beeinflusst und entschieden. Auch im öffentlichen Sektor tritt dieser Aspekt verstärkt zutage, denn das Prinzip der Wirtschaftlichkeit gilt hier ebenso und wird durch die Bildung von Staatsbetrieben, die einen Teil ihrer Kosten durch eigene Einnahmen selbst erwirtschaften müssen, noch deutlich verstärkt.

Wissenschaftlichkeit

Nachvollziehbarkeit und Überprüfbarkeit, mithin Öffentlichkeit, sind seit jeher wichtige Kriterien der Wissenschaftlichkeit. Nur wer seine Erkenntnisse nachvollziehbar und überprüfbar veröffentlicht, wird wissenschaftlich anerkannt. Dies schließt, neben der Publikation der Ergebnisse, in der Regel auch den Zugang zu Rohdaten und Zwischenergebnissen ein, zumindest, wenn diese nicht anderweitig ermittelt werden können.

Die Freiheit der Wissenschaft wird im Allgemeinen wesentlich durch eine öffentliche Finanzierung oder Förderung gewährleistet. Im Gegenzug erwartet man von den derart betroffenen Projekten mit Recht, dass die sich aus der Forschungsarbeit ergebenden Daten, Informationen und Ergebnisse auch frei und offen zur Verfügung gestellt werden, möglichst als Open Data. Diese Anforderungen spiegeln sich beispielsweise in der „Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen"[18] vom Oktober 2003 wider, der sich inzwischen weit über 300 Institutionen angeschlossen haben. Die Erklärung nennt als Ziel:

„Unsere Aufgabe Wissen weiterzugeben ist nur halb erfüllt, wenn diese Informationen für die Gesellschaft nicht in umfassender Weise und einfach zugänglich sind. Neben den konventionellen Methoden müssen zunehmend auch die neuen Möglichkeiten der Wissensverbreitung über das Internet nach dem Prinzip des offenen Zugangs (Open Access-Paradigma) gefördert werden. Wir definieren den offenen Zugang oder den ‚Open Access' als eine umfassende Quelle menschlichen Wissens und kulturellen Erbes, die von der Wissenschaftsgemeinschaft bestätigt wurden.

Die Vision von einer umfassenden und frei zugänglichen Repräsentation des Wissens lässt sich nur realisieren, wenn sich das Internet der Zukunft durch Nachhaltigkeit, Interaktivität und Transparenz auszeichnet. Inhalte und Software müssen offen zugänglich und kompatibel sein."

Wichtige Forschungsförderer wie beispielsweise die Deutsche Forschungsgemeinschaft, die Max-Planck-Gesellschaft und die Fraunhofer-Gesellschaft haben sich dieser Auffassung von freier Forschung schon in der Berliner Initiative von 2003 angeschlossen. Die Allianz der deutschen Wissenschaftsorganisationen, ein Beratungsgremium aus zehn großen deutschen Wissenschaftsorganisationen[19], widmet sich mit der Schwerpunktinitiative „Digitale Information"[20] diesem Thema:

„Wissenschaftlerinnen und Wissenschaftler mit der bestmöglichen Informationsinfrastruktur auszustat-

ten, die sie für ihre Forschung brauchen, ist das Leitbild der Informationsversorgung. Im digitalen Zeitalter bedeutet das die digitale und für den Nutzer möglichst entgelt- und barrierefreie Verfügbarkeit von Publikationen, Primärdaten der Forschung und virtuellen Forschungs- und Kommunikationsumgebungen. Es gilt daher eine nachhaltige integrierte digitale Forschungsumgebung zu schaffen, in der jeder Forschende von überall in Deutschland auf das gesamte publizierte Wissen und die relevanten Forschungsprimärdaten zugreifen kann." (Leitbild der Allianz-Initiative[21], Juni 2008).

Authentizität von Informationen

Ein interessanter Aspekt bei der Veröffentlichung von Daten beziehungsweise Informationen ist die Authentizität. Authentizität meint, dass die ursprüngliche Intention des Verfassers erhalten ist. Dazu muss beispielsweise bei der Archivierung sichergestellt werden, dass Dokumente nicht verändert oder verfälscht werden können (vgl. Bibby/Göldner S. 70)[22].

An dieser Stelle muss der Unterschied zwischen Daten und Informationen etwas näher beleuchtet werden. Unter Daten versteht man eine Ansammlung von Zeichen, die in der Regel maschinenlesbar sind, also von einem Computer verarbeitet werden können. Informationen erhält man erst aus einer Interpretation dieser Daten durch den Menschen. Dazwischen liegt aber mindestens noch der Prozess der Darstellung der Daten, und dieser Prozess ist leider nicht immer eindeutig festgelegt. Selbst wenn man vereinbarte Datenformate berücksichtigt, so gibt es zumeist verschiedene Programme zur Darstellung. Damit wird die Authentizität aber fraglich. Die nachfolgenden Beispiele sollen dies verdeutlichen.

Während bei reinen Textdaten die Darstellung noch weitgehend unproblematisch ist, so gibt es beispielsweise bei Tabellendaten eine ganze Reihe von Freiheiten. Eine falsche Zuordnung von Werten zu Tabellenspalten führt zwangsläufig zu einer veränderten und höchstwahrscheinlich falschen Interpretation, mithin zu einem Verlust an Authentizität. Auch bei Geodaten kann man diese Problematik finden. Wenn man eine Region im Maßstab 1:100.000 grob umreißt und die bestehende Lageunsicherheit von vielleicht 100 m mit einem 1 mm breiten Polygon darstellt, so enthalten die Geodaten zumeist nur die Stützpunkte des Polygons (zum Beispiel Shapefile). Verwendet man diese Daten in einem anderen Kontext vielleicht im Maßstab 1:500 und achtet nicht auf die Darstellung, so erhält man vermutlich eine dünne Linie, die eine Genauigkeit von wenigen Zentimetern impliziert, die ursprünglich niemals vorhanden war.

Bei komplexen Informationen kommt ein weiterer Aspekt hinzu, denn sie verlangen ein ebenso komplexes System der Darstellung, und oft sind neben den reinen Daten noch wichtige Zusatzinformationen erforderlich. Fehlen einzelne Bestandteile, kann es zu folgenreichen Fehlinterpretationen kommen und die Authentizität geht verloren. Aus der Archäologie kommt dazu folgendes Beispiel:

„Oft wird aus der Kartierung bekannter Archäologieflächen der falsche Umkehrschluss gezogen, die verbliebene Fläche sei frei von archäologischen Belangen. Ebenso kann der falsche Eindruck entstehen, das archäologische Kulturerbe ließe sich präzise abgrenzen. Solchen Fehlinterpretationen kann nur durch ein gut abgestimmtes Zusammenspiel von kartografischer Darstellung, Legende und gezielten Zusatzinformationen zu Inhalt und gesetzlichem Rahmen begegnet werden. Der Kontext ist entscheidend!" (Meinungsbild[23], S. 2)

Bei digitalen Informationen reicht also die Forderung nach Vollständigkeit und Unversehrtheit der Daten nicht aus. Es muss vielmehr durch zusätzliche Mittel sichergestellt werden, dass bei der Darstellung und Präsentation der Daten die ursprüngliche Intention des Verfassers, und damit die Authentizität der Informationen, erhalten bleibt.

Archäologisches Erbe

Das archäologische Erbe ist ein gesellschaftliches Erbe. Diese Ansicht wird nicht zuletzt durch das „Europäische Übereinkommen zum Schutz des archäologischen Erbes"[24] (Valetta-Abkommen) deutlich, welches inzwischen in 36 Staaten in Kraft ist. Artikel 1 dieses Übereinkommens besagt:

„Ziel dieses (revidierten) Übereinkommens ist es, das archäologische Erbe als Quelle gemeinsamer europäischer Erinnerung und als Instrument für historische und wissenschaftliche Studien zu schützen."

Auch wenn das Übereinkommen vorrangig materielle Elemente des archäologischen Erbes benennt („Überreste und Gegenstände sowie ... sonstigen Spuren des Menschen"), so muss man archäologische Informationen dem doch als immaterielle Werte gleichsetzen, insbesondere wenn diese Informationen aufgrund unvermeidbarer Zerstörung von materiellen Elementen des Erbes den Rang von Primärquellen einnehmen. Es ist daher nur folgerichtig, dass man auch archäologische Informationen als gesellschaftliches Erbe betrachtet und ihren Austausch, ihre wissenschaftliche Nutzung sowie ihre gesellschaftliche Rezeption unterstützt, jedenfalls solange damit keine

anderweitige Beeinträchtigung des archäologischen Erbes verbunden ist. Artikel 8 des Übereinkommens bemerkt dazu:

„Jede Vertragspartei verpflichtet sich:
- den nationalen und internationalen Austausch von Elementen des archäologischen Erbes für akademisch-wissenschaftliche Zwecke zu erleichtern und gleichzeitig geeignete Schritte zu unternehmen, um zu verhindern, daß der kulturelle und wissenschaftliche Wert dieser Elemente durch die Weitergabe beeinträchtigt wird;
- die zentrale Erfassung von Informationen über bereits laufende archäologische Forschungs- und Ausgrabungsarbeiten zu fördern und zur Aufstellung internationaler Forschungsprogramme beizutragen."

Das archäologische Erbe impliziert also eine deutliche Tendenz zu einem freien Informationszugang, wobei gewisse Einschränkungen zur Verwirklichung von Schutzzielen einbezogen sind.

Denkmalschutz

Der Schutz des archäologischen Erbes wird in Deutschland durch Denkmalschutzgesetze der einzelnen Bundesländer geregelt[25]. Bis auf wenige Ausnahmen ergeben sich Kulturdenkmale dabei nach dem deklaratorischen Prinzip (ipso jure) und müssen also nicht erst benannt oder bekannt gemacht werden, um Denkmaleigenschaft und Schutzstatus zu erlangen. Dies sichert einen umfassenden Schutz des archäologischen Erbes auch dann, wenn es noch nicht umfassend untersucht und dokumentiert ist, sondern sich nur durch Spuren und Hinweise als begründeter Verdacht abzeichnet. Es stehen also neben den schon bekannten auch alle noch nicht bekannten Kulturdenkmale unter Schutz, und diese sind nach einhelliger Meinung der Archäologen die überwiegende Mehrheit.

Eine Herausforderung des deklaratorischen Systems liegt darin, das Schutzgut ausreichend gut zu beschreiben. Ein Kulturdenkmal kann eben nur dann wirkungsvoll geschützt werden, wenn es bekannt ist (hier sind konstitutiv bestimmte Flächen, wie beispielsweise im Naturschutz üblich, deutlich im Vorteil). Außerdem gebieten die Grundsätze der Rechtssicherheit und Rechtsklarheit, dass sich die zu schützenden Objekte nicht nur Fachwissenschaftlern, sondern auch den Adressaten und Anwendern der Rechtsnorm (also der Allgemeinheit) erschließen. Aus diesem Grund fordern auch deklaratorische Denkmalschutzgesetze zur Führung und Veröffentlichung von Denkmallisten auf. Auch wenn diese nicht den Gesamtbestand der geschützten Objekte zeigen können, so unterstützen sie doch immerhin die Rechtsklarheit bei den schon bekannten Objekten des archäologischen Erbes.

Es ist jedoch zu beachten, dass detaillierte Informationen zum archäologischen Erbe sich auch gegen die Ziele des Denkmalschutzes richten können, zum Beispiel bei Denkmalen, die ein gewisses Schatzpotenzial aufweisen: „Die Veröffentlichung von Bodendenkmälern im Internet birgt Gefahren, da diese so für Schatzsucher und Raubgräber einfacher zugänglich sind." (Meinungsbild[26], S. 2). Es ist also stets sorgfältig abzuwägen, welche Informationen zum archäologischen Erbe in welchem Rahmen bekannt gemacht werden.

Resümee

Die Freiheit der Daten, nein besser, die Freiheit von Informationen ist ein hehres Gut und wird in der Gesellschaft allgemein hochgeschätzt. Viele Grundrechte erfordern diese Freiheit. Demokratie, öffentliche Verwaltung und nicht zuletzt die Wissenschaft profitieren von ihr. Auch wenn es gewisse Einschränkungen durch andere Grundrechte wie beispielsweise Datenschutz oder Urheberrechtsschutz gibt und wirtschaftliche Interessen der Informationsfreiheit gelegentlich entgegenstehen, sollten wir nicht von dem Bestreben ablassen, unser Wissen zugunsten der Allgemeinheit mit anderen zu teilen.

Anmerkungen, Quellen, Internet

[1] Der Begriff Informationsfreiheit wird allgemein sowohl für Rezipientenfreiheit (= Freiheit, sich zu informieren, GG Art. 5) als auch für Informationszugangsfreiheit (= freier Zugang zu Informationen) benutzt. Zur Unterscheidung wollen wir im ersten Fall von Rezipientenfreiheit und im zweiten Fall von Informationsfreiheit sprechen.

[2] GG – Grundgesetz der Bundesrepublik Deutschland (http://www.bundestag.de/dokumente/rechtsgrundlagen/grundgesetz/).

[3] Vgl. Anmerkung 1.

[4] BGB – Bürgerliches Gesetzbuch der Bundesrepublik Deutschland (http://www.gesetze-im-internet.de/bgb/index.html).

[5] Informationsweiterverwendungsrichtlinie der EU: „Richtlinie 2003/98/EG des Europäischen Parlaments und des Rates vom 17. November 2003 über die Weiterverwendung von Informationen des öffentlichen Sektors" (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do? uri= J:L:2003:345:0090:0096:DE:PDF).

[6] Empfehlung des Europarates Nr. 854 (1979) „on access by the public to government records and freedom of information". (http://assembly.coe.int/main.asp?Link=/documents/adoptedtext/ta79/erec854.htm).

[7] Empfehlung des Europarates Nr. R (81) 19 „of the committee of ministers to member states on the access to information held by public authorities" aus dem Jahr 1981 (http://www.coe.int/t/dghl/standardsetting/media/doc/cm/rec%281981%29019_EN.asp).

[8] Empfehlung des Europarates Rec(2002)2 „of the Committee of Ministers to member states on access to official documents" aus dem Jahr 2002 (https://wcd.coe.int/ViewDoc.jsp?id=262135).

[9] Informationsfreiheitsgesetze – Übersicht des Bundesbeauftragten für Datenschutz und Informationsfreiheit (http://www.bfdi.bund.de/IFG/Gesetze/Landesgesetze/Landesgesetze_node.html).

[10] UIG des Bundes – Umweltinformationsgesetz vom 22. Dezember 2004 (http://www.gesetze-im-internet.de/uig_2005/ index.html).

[11] INSPIRE – Infrastructure for Spatial Information in the European Community, Internetauftritt (http://inspire.jrc.ec.europa.eu/).

[12] INSPIRE-Richtlinie: „Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft" (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:DE:PDF).

[13] Meinungsbild im Verband der Landesarchäologen zum Thema Archäologischer Denkmalschutz und INSPIRE-Schutzgebiete (http://www.landesarchaeologen.de/fileadmin/Dokumente/Dokumente_Kommissionen/Dokumente_Archaeologie-Informationssysteme/Dokumente_AIS_Rechte/Meinungsbild_Arch%C3%A4ologie-INSPIRE-Schutzgebiete.pdf).

[14] Vgl. Anmerkung 13.

[15] Vgl. Anmerkung 13.

[16] VLA-Website, Kommission Archäologie und Informationssysteme – AG Rechtliche Aspekte (http://www.landesarchaeologen.de/verband/kommissionen/archaeologie-und-informationssysteme/ projektearbeitsgruppen/rechtliche-aspekte/).

[17] Creative Commons – deutscher Internetauftritt (http://de.creativecommons.org/).

[18] „Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen" vom Oktober 2003 (http://oa.mpg.de/files/2010/04/Berliner_Erklaerung_dt_Version_07-2006.pdf beziehungsweise http://oa.mpg.de/lang/de/berlin-prozess/signatoren/).

[19] Partnerorganisationen der „Allianz der deutschen Wissenschaftsorganisationen" sind: Alexander von Humboldt-Stiftung, Deutsche Akademie der Naturforscher Leopoldina, Deutsche Forschungsgemeinschaft, Deutscher Akademischer Austausch-Dienst, Fraunhofer-Gesellschaft, Helmholtz-Gemeinschaft, Hochschulrektorenkonferenz, Leibniz-Gemeinschaft, Max-Planck-Gesellschaft und Wissenschaftsrat. (http://www.allianzinitiative.de/de/partnerorganisationen/).

[20] Schwerpunktinitiative „Digitale Information" der Allianz der deutschen Wissenschaftsorganisationen (http://www.allianz-initiative.de/de/).

[21] Leitbild der Allianz-Initiative (11. Juni 2008) (http://www.allianzinitiative.de/fileadmin/user_upload/keyvisuals/atmos/pm_allianz_digitale_information_details_080612.pdf).

[22] Bibby/Göldner. „Archäologische digitale Daten: Authentizität und Funktionalität". In: Archäologie und Informationssysteme. Arbeitshefte zur Denkmalpflege in Niedersachsen, Bd. 42.

[23] Vgl. Anmerkung 13.

[24] „Europäische Übereinkommen zum Schutz des archäologischen Erbes" (Valetta-Abkommen, revidierte Fassung von 1992, deutsche Übersetzung) (http://conventions.coe.int/Treaty/ger/Treaties/Html/143.htm).

[25] Denkmalschutzgesetze der Bundesländer, Übersicht des Verbandes der Landesarchäologen (http://www.landesarchaeologen.de/denkmalschutzgesetze/).

[26] Vgl. Anmerkung 13.

(Alle Internet-Links mit Stand vom 08.07.2012.)

Opening access to heritage resources: risk, opportunity or paradigm shift?

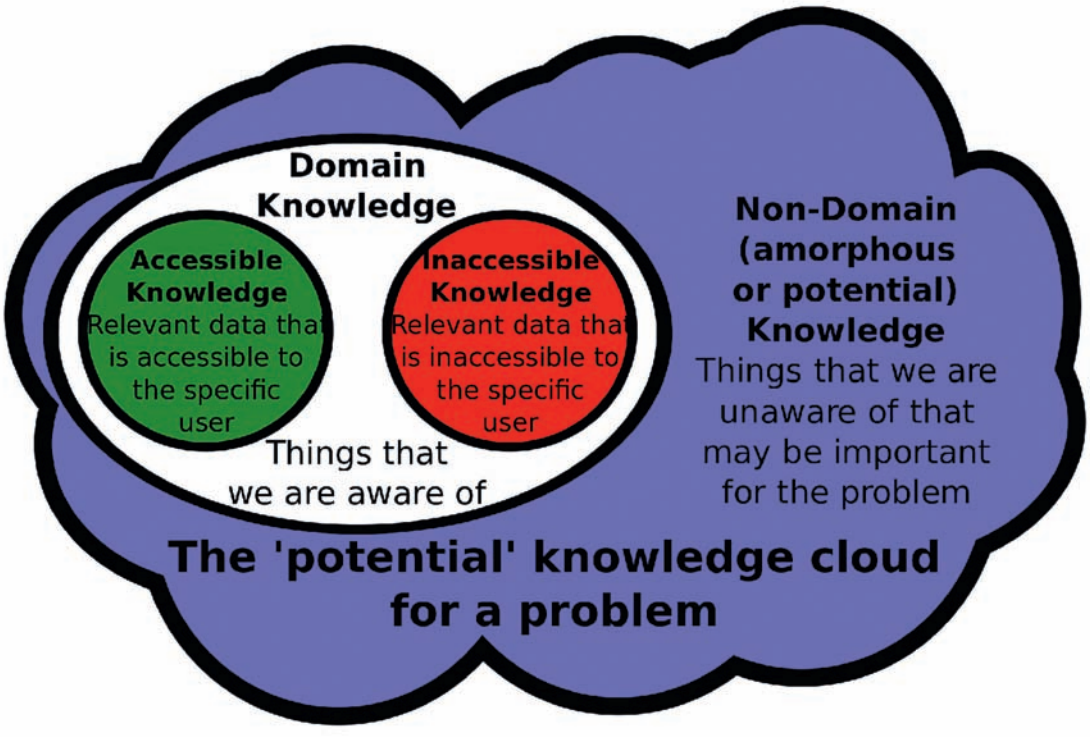
Anthony Beck

Abstract

'Open' has captured the zeitgeist. It is argued that by providing clear and transparent access to knowledge objects (data, theories and knowledge (open access, open data, open methods, open knowledge)) and enhancing collaborative and creative input 'open' approaches have the potential to revolutionise science, humanities and arts. Open approaches require active engagement and are as much a social and organisational issue as a technical one. This requires engagement from a broader societal base. For example, primary data creators need the appropriate incentives to provide access to Open Data – these incentives will vary between different groups: contract archaeologists, curatorial archaeologists and research archaeologists all have different drivers. Equally important is that open approaches raise a number of ethical issues. A UK focus has been taken: the author appreciates there are differences in application outside the UK. This is a shortened version of a paper that is published in the Journal of World Archaeology special issue on open Archaeology in Autumn 2012.

Introduction

Heritage resources within landscapes are at risk from a variety of natural and anthropogenic forces. The loss of some of these resources is inevitable, but the decision making process is a delicate balancing act between different cultural, industrial and economic issues which can be enacted at different scales: from the local to the international. Data, and the subsequent information and knowledge that is derived from an integrated data corpus, should be used effectively to produce realistic policy (i. e. policy which can be enacted and produces the desired impact) and develop effective governance strategies (be they local, regional, national or global). The underlying ethos is that better decisions are made when users can access appropriate parts of the knowledge base in a manner which is relevant to the problem they are solving. The problem is that decisions, policy or research is not enacted with ‚perfect' knowledge. Current approaches mean that research and decision making is undertaken on a sub-set of the corpus of data. This leads to a Rumsfeldian system of knowledge availability (see Figure 1):



1 The 'potential' knowledge cloud for a problem (re-used under a CC-By licence from Anthony Beck: http://commons.wikimedia.org/wiki/File:The_%27potential%27_knowledge_cloud_for_a_problem.svg).

- the known knowns: the accessible knowledge
- the known unknowns: the known but inaccessible knowledge
- the unknown unknowns: the potential knowledge advances gained by integrating all data, collaborating with different domains and future research avenues

Ideally one would reduce the size of the latter two domains so that decisions, be they research, policy or management, can be formed from a position of 'perfect', or 'near-perfect', knowledge with information and data which is both accessible and well understood. The 21st century is likely to see the global transition from data economies, to information economies (Wikipedia 2012a) and finally through to knowledge economies (Wikipedia 2012b). It's postulated that better decisions are made: information and knowledge economies have a better understanding of the nature of any underlying data (its quality, vagaries and uncertainties), have bodies of theory that can be utilised to transform data into information and a better understanding of the 'totality'. It is likely that this will occur through the development of communities of practice (Wenger, McDermott, and Snyder 2002) that comprise of loosely coupled groups who are studying related problems.

The reality is somewhat different, rarely are decisions made on the basis of all the available evidence. The heritage sector is comprised of a network of agencies and stakeholders. They perform a range of functions from curating monuments or undertaking excavations to advising governments and providing public access. Each collects data. Unfortunately access to the corpus of data held by these agencies is fragmented and opaque. Many of the different distributed Historic Environment Information Records (HEIRs) have erected barriers to data sharing in the form of bureaucracy, formats (particularly paper archives), incompatible structures and legal/financial frameworks. The latter point is important as the legal and financial framework includes licences that inhibit downstream sharing and re-use. This severely inhibits knowledge-transfer. Important data resides in silos which effectively block access to different stakeholders on the basis of awareness, privilege (access rights), cost and security. Portals that aggregate content do exist, but the majority only provide access to synthetic derivative, or generalisations. This lack of access to data leads to downstream fragmentation in the data transformation environments (theory, processing methods etc.), the knowledge structuring environments (classification framework: dating, pottery sequences etc.) and interpretations. Stakeholder requirements are not joined up: the systems that facilitate the effective development of diverse 'communties of practice' (Wikipedia 2012c; Wenger, McDermott, and Snyder 2002) that could transform engagement do not exist.

Paradoxically, at a transnational level we are encouraged not just to share our data but to integrate it. This is core to the principles underpinning the EU-wide INSPIRE directive (2012) and is implied in the European Landscape Convention. Interoperability and 'fit-for-purpose' are key terms: data should be easy to integrate across different platforms and it should be made available in a way which the intended end-user can actually use and understand. The issue here is not 'if we should improve access' but 'how we should improve access'. Unfortunately, many of the datasets which fall under the remit of INSPIRE are generalised, or synthetic, and don't reveal the richness of the underlying raw data sources. INSPIRE and other open data initiatives spearheaded by many western governments have demonstrated that there is a real demand for government, and other, data. However, it is as yet unclear how far they will penetrate into the heritage sector.

This paper will consider these issues in relation to opening access to heritage resources. A movement that is being called 'Open Archaeology'.

A vision for Open Archaeology

Imagine a scenario where data is managed in the following way. Fine grained data downloaded from geophysics instruments or collected during, likely to be fully digital, excavations is placed into a virtual 'folder'. This folder synchronises the data with a cloud based repository. This invokes a variety of services which generate description and discovery metadata (essentially archiving the raw data and facilitating discovery). Using ontologies and other semantic web tools the data can interoperate as Linked Data (Wikipedia 2012d) with the corpus of data collected at this and other scales. Persistent Uniform Resource Identifiers (URIs) or Digital Object Identifiers (DOIs) are allocated for each object for long term referencing. The underlying data quality can be evaluated and improved in relation to the network of data, information and knowledge in which the 'new' object exists. When users want to query this resource they can either access it directly (as a service or a download) or utilise mediation interfaces which have been specifically designed to transform the resource into specific data and knowledge products that the user requires (again as a download or a service). The technical terminology may strip this of elegance, but these types of environments are currently under development. They represent a major conceptual shift from depositing 'static' documents, objects and data in archives to developing dynamic, rich and interlinked repositories of knowledge.

If archaeological data is to be viewed as a dynamic resource it is important to consider the relationships between the different data, information and knowledge objects. Archaeology deals with a variety of different data which are integrated to gain insights into the past. For simplicity these have broken these down into three groups:

- Physical observations – measurement and observation data (many of which can not be replicated again) including sensor and laboratory measurements (such as geophysical and soil measurements). This data does not change over time. For example a sherd of pottery may have a location, context, dimensions and can be attributed with form and fabric classification descriptors.

- The structuring knowledge/classification environment – the knowledge (or classification) frameworks which can be attributed to the physical observations. These frameworks reflect clusterings and groupings in the data. This data is dynamic. For example the pottery form and fabric sequence (and their associated dating implications) change when new data is added to the corpus or specialists refine their points of view. A similar position is observed with radiocarbon dates and the body of knowledge that provides calibration for the raw dates. There should be a tight relationship between these relationships and the physical observations which together represent a dynamic knowledge base which can be fed into research, policy, practice and management. Currently this relationship is not formally represented and is normally decoupled.

- Analysis, interpretations and synthesis – a layering of multiple interpretative points of view based on hypotheses and bodies of theory and evidence.

It should be noted that in recent years this contrast has rightly been subject to revision, as various commentators have noted the widely acknowledged fact that all stages of archaeological practice involve theory-laden assumptions, and hence that data collection and interpretation are closely entwined. Irrespective of this broader debate, each of these components require additional metadata that describes the methods of observation, collection and analysis so that subsequent re-users can understand issues pertaining to scale, uncertainty and ambiguity that are essential when datasets are integrated. Being able to link, and therefore integrate and query, the different data resources means that heterogeneous resources can be treated as one. This is close to becoming Linked Data (2012). The Linked Data ethos (Bizer, Heath, and Berners-Lee 2009) underpins the development of the Semantic Web which aims to transform the unstructured, document focussed, web into a 'web of data' (Wikipedia 2012).

This dynamism could have profound implications for archaeological data collection, consumption and engagement. For example, in addition to many other things pottery provides essential dating evidence for archaeological contexts. However, pottery sequences are developed on a local basis by individuals with an imperfect knowledge of the global situation. This means there is overlap, duplication and conflict between different pottery sequences which are periodically reconciled (your Type IIb sherd is the same as my Type IVd sherd and hence the dating range can be refined). This is the perennial processes of lumping and splitting inherent in any classification system. The semantic integration of localised sequences can potentially support more robust pottery frameworks. In addition, as the pottery data is linked (and not decoupled and stale) then updated pottery classifications and dating implications immediately update the dating probability density function for a context or group. One can also automatically reason over the data to find out which contexts, relationships and groups are impacted by a change in the dating sequences either by proxy or by logical inference (a change in the date of a context produces a logical inconsistency with a stratigraphically related group). In a Semantic Web (Wikipedia 2012e) scenario then the logical consistency in the physical and stratigraphic relationships can be automatically verified using reasoning software (such as Prolog). As all of the data is stored as Linked Data, this means that all the primary data archives are linked to their supporting knowledge frameworks (such as a pottery sequence). When a knowledge framework changes the implications are propagated through to the related data dynamically. This means that, in theory, the implications of minor changes in the structuring knowledge environment can be tracked dynamically through to the underlying observations and their impact observed in interpretations. The feedback mechanism implied by Linked Data and the Semantic Web can profoundly alter the way archaeologists, and others, engage with their data and derived resources.

Deposition is no longer the final act of the excavation, or any other, process: rather it is where the dataset can be integrated with other digital resources and analysed as part of the complex tapestry of heritage data. The data does not have to go stale: as the source data is re-interpreted and interpretation frameworks change these are dynamically linked through to the archives. Hence, the data sets retain their integrity in light of changes in the surrounding and supporting knowledge system. This means that any policy,

management, curatorial and research decisions are based upon data that reflects the most-up-to date information and knowledge. This must be an aspiration for a discipline which progresses by re-analysing and re-structuring it's primary data corpus.

The requirements of an 'open' archaeological corpus: licences

Open Content is defined by the Open Knowledge Foundation (OKF 2012a, 2012b) as follows, 'A piece of content or data is open if anyone is free to use, reuse, and redistribute it subject only, at most, to the requirement to attribute and share-alike.' Essentially content is open if the re-user is broadly unfettered in the way they consume the content and share derivatives. What provides this freedom is the licence associated with the resource. The licence describes what can be done with the content or data and how it should be shared and referenced (if at all). OKF describes a number of different licences that conform to the requirements of Open Data and other open content. The two primary issues concern how data or content is referenced (does the licence include a By Attribution (By) clause) and what imposition the data sharer will put on downstream licencing (is a third party user free to do what they want or not). At the very least open resources need licences that encourage re-use.

Ethics and open heritage resources

Open licences provide the mechanisms for effectively sharing resources. However, they do not provide answers to how data and syntheses should be opened or if they should be opened at all. This is the sphere of ethics. Two broad issues will be discussed: the difference between the ethical practices of those that produce and those that consume the corpus of archaeological knowledge.
Archaeological data is normally collected as either part of the research process or as part of the planning (or rescue) process – in advance of construction. Hence archaeological work is normally conducted by academics, government organizations or independent contract units. Each of these bodies operate under legal and professional frameworks. The majority of these frameworks mandate that these organisations 'make this evidence (and any archive generated), ublicly accessible'. Both the National Planning Policy Framework (NPPF: Department for Communities and Local Government 2012) and Research Councils UK (RCUK 2012a) have similar text for the UK. Obviously the term 'publically accessible' is open to a variety of interpretations – however, as the majority of archaeological data are 'born digital' it is likely that the public (taxpayers) will expect digital resources to be made available digitally. This brings us to the ethical consumption of heritage data.

There are two broad ethical issues associated with the consumption of heritage data:
1. Access to data and derivatives that still have resonance for a community or cultural group (indigenous groups, 20th Century war excavations and forensic archaeology).
2. Providing access to data that facilitates looting and the international trade in illegal artefacts.

The distinction between these two points is mainly one of boundaries and certainty. In respect of point 1 it is reasonably clear when an activity may have ethical implications, even in retrospect. Remedial activities can be taken and the corpus of evidence can be redacted or obfuscated for different audiences accordingly. However, point 2 is significantly more ambiguous. The same data which is used to conduct archaeological research or curatorial management can be used to target looting activities (for example, night hawking: the illegal act of metal detecting without the landowners consent or on scheduled sites). In the latter example it is difficult to know what to do – removal of access rights to the archaeological knowledge base may reduce the problem but has a profound impact on legitimate activities. In respect of access to digital data this is effectively the position adopted by many curatorial authorities. Data can be accessed by visiting curatorial facilities but is generally not accessible on-demand externally.

This position is polarised and plays to the lowest common denominator: it is based on the premise that accessible knowledge will inevitably be abused and eschews any of the benefits that data sharing can provide. The Portable Antiquities Scheme (PAS) has done much to improve this situation. In terms of consumption, the data held within the PAS database is publically accessible but in a degraded format - the spatial location and some attributes have been obfuscated. By registering with the PAS different users can access the data at different degrees of granularity including a 'research' grade. This user accreditation framework allows a more nuanced approach to knowledge segmentation which provides a re-use position based upon accreditation and trust. However, it is not, in principle, open. The PAS has an extensive regional and national outreach programme which has resulted in a great deal of engagement and a conduit for citizen science. Through this the PAS has helped to strengthen the relationships between ethical metal detectorists and the archaeological community. In 2007 it was estimated that two thirds of metal detectorists belonging to a club report some of their finds to the PAS.

It is clear that ethical issues are complex. Every stakeholder can produce compelling arguments as to why they should get preferential access to everyone else's data whilst at the same time they should place embargoes on their own. The underlying issue is one of data embargoes and how these are implemented, or not. In this respect the archaeological discipline is in a difficult position. On the one hand the statutory, legislatory and funding frameworks call for public access of the products of archaeological research and practice. On the other hand there are concerns about the ethical implications of sharing any archaeological data as it can be used by looters. These two positions are polarised and obviously conflicting.

It is important that the risks and benefits associated with providing access to heritage data are considered from a pragmatic basis. Some data are sensitive: it can be argued that access to such data should be restricted for reasons of security and ethics (particularly where these issues are clearly defined – such as indigenous archaeologies). It should also be recognised that data sensitivities and risks change over time. A data set which may be deemed sensitive whilst under excavation (at risk whilst it is accessible) is less sensitive once built over (at less risk as it is less accessible). Furthermore, providing access to data does not always mean providing access to raw data: data can be obfuscated by generalising spatial, temporal and other attributes. This implies some form of user accreditation. Whilst not ideal, and definitely not open, access based on user accreditation does mean that different user groups get access to data that might otherwise have been embargoed. By understanding these issues fit-for-purpose systems can be developed that deploy data in a timely and effective way.

Incentives and disincentives for Open Archaeology

This paper has argued that Open Archaeology can have a significantly positive impact on many of the modes of archaeological consumption. The challenge is one of how the environment can be engineered to increase the likelihood of this happening subject to the ethical considerations discussed earlier. This is about incentives and how individuals and organisations embed Open Archaeology principles into their work practice.

As discussed previously the requirements to access the foundational objects for Open Archaeology (data, methods, synthesis etc.) are enshrined within many of the core requirements for academic, curatorial and contract archaeology. One of the problems is that although publically accessible archives are advocated the guidance tends not to stipulate repositories, structures and licences that maximise re-use. Currently the only successful area is open access to synthetic academic publications and excavation, and other, reports. Neither of these provide enough information on source detail and processing methods that allow the scientific re-appraisal of technique nor are they in a format that facilitates re-use of content embedded in the documents.

Academic archaeological researchers funded by the UK government should already be conducting Open Archaeology, however there appears to be a gap between the amount of research commissioned and the number of research outputs. The size of this gap needs quantifying. One of the problems is that the reward system in academia is based upon publication output and not on the output of other, more fundamental, research objects. Hence, publications are a much higher priority to academics in terms of recognition and career advancement. Whilst publications are an important element of research output they have an excessively high weighting which is detrimental to the delivery of other research outputs. It is important that this balance is redressed and that appropriate checks are put in place to confirm that the outputs are in placed in appropriate repositories. The Wellcome Trust is strengthening the manner in which it enforces its open access policy: failure to comply with the policy could result in final grant payments being withheld and non-compliant publications being discounted when applying for further funding (Wellcome Trust 2012).

Contract archaeology has different drivers. Whilst NPPF states that the underlying data should be made publically accessible, the mechanism of deposition does not mean that the resources can be easily reused. The current commercial business process is predicated on delivering reports and depositing 'paper archives' that satisfy both client and curatorial requirements so that the process of excavation can be signed off and the rest of the development can proceed. The problem here is in the nature of the object (data) has moved from a paper, analogue, record to a digital record whilst many curatorial systems are still designed with analogue data in mind. If the surrounding data re-use and analysis environment is generated then it is arguable that the very act of depositing structured data into the corpus is enough to satisfy the requirements of the planning process as this data will be immediately available for review and analysis by multiple stakeholders. This means that the contract unit would no longer be required to produce a handcrafted report that includes textual summaries of data (this kind of synthesis, if necessary, can be built directly from the deposited data itself). Instead one could generate alternative analytical derivatives (for exam-

ple to examine the impact of the work on regional agendas) and alternative synthetic output (for example producing popular synthesis which encourages people to re-use the data). As many large UK contract units are also charities with an educational remit this repositioning would enhance engagement and provide a catalyst for public access to Open data. This should make the excavation process cheaper for the commercial archaeological contractor, and by proxy, the client. The incentives in the commercial framework are more about changing the execution of the business process rather than incentivising individuals.

Although contract and academic archaeologists will be the creators of the majority of data, curatorial archaeologists are key stakeholders and mediators of knowledge. The technologies and approaches described here have the potential to significantly disrupt this sector. Although many benefits have been described the very process of change has social implications and would need effective management. For at least the past decade UK curatorial services at a regional and national level have been significantly cut. This has damaged morale. The introduction of systems that fundamentally change working practice are likely to create issues around role perception, recognition and job security. These will need effective management. Whilst these archaeologists will still produce curatorial and planning advice they will start to base this advice on a variety of data sources rather than just the Historic Environment Information Resource (HEIR) they maintain. This may be seen to erode prestige and relevance: this will also need managing. However, rather than spending time maintaining and updating a dataset based on at least secondary synthesis, curatorial archaeologists will be directly accessing the full corpus of archaeological knowledge to address regional and national policy and research agendas. This has the potential to be more engaging and mean that the sector has more relevance to both planning, research and the public. It could be argued that large scale projects that employ many archaeological contractors, such as the Channel Tunnel Rail Link, could see many benefits by opening the data from these project at an early stage so that the curatorial archaeologists can work more effectively with the consultant, contract and research archaeologists to develop more nuanced strategies and outputs.

Finally there are issues of public engagement and accountability. By making the results of academic and contract archaeology open then the public can engage with the resource more effectively and link planning and research policy actions directly to outcomes (a form of process transparency). A number of incumbent governments have had aspirations to make local and national decision making systems more transparent. There is also another facet, and one which is followed by the author: open is the 'right thing to do'. This is particularly relevant to academics who are funded through the public purse: it's a duty to communicate their research and its social and ethical implications both to policy-makers and to the non-specialist public. There is an increasing view that these types of non-academic engagements are richer, participatory, relationships that go far beyond simple outreach (MORI 2000; Royal Society, RCUK, and Wellcome Trust 2006, 9; Poliakoff and Webb 2007).

Conclusions

The Internet has and is changing society: it has helped to increase expectations of access which in turn has led to more transparent and open positions from governments and institutions. Access to data and derivatives, particularly data collected through public funds, is likely to be significantly enhanced. Importantly this will not just be the provision of data in books, or their digital facsimiles, but the provision of source data that is structured in a way which can be consumed by a multitude of users on the semantic web. This jars against how data and derivatives are currently structured. Traditionally the planning processes uses higher level synthesis (site reports etc.) as a basis for decision making. This essentially decouples the link between the physical observations and the structuring knowledge environment and the available information from which to make decisions. The decoupling of data to generate synthesis is increasingly problematic when one considers we are still dealing with a data explosion producing in the words of English Heritage (1995) 'a large and inaccessible data mountain'. It is essential that data needs to be linked more closely to synthesis, management, curation and policy. Open Archaeology principles help to re-establish these links.

Data is starting to receive the attention it deserves. Initiatives such as Open Data and Open Science are providing access to an unprecedented amount of data (and the algorithms and methods which can be used to transform and synthesise it). International policy initiatives such as INSPIRE are forcing governments to address transnational data and semantic integration and access issues. It is imperative that ways are found for researchers, curators, policy makers and the public to overcome the ethical, legal and technical challenges in providing data and other research objects whilst realizing the potential of all the resource. In addition, in a world which is increasingly reliant on data and associated metrics on provenance and accuracy, we should be providing mechanism whereby providers of such data can be cited and rewarded

appropriately. If decisions are based upon data then those people and organisations that deposit high quality data should be praised and rewarded accordingly.

There will remain areas where data, information and knowledge remains silo-ed. It is probably right that some data remains silo-ed: however, the reasons need to be made clear. It is wrong to silo data on the basis of dogma and worse when this is done in organisations that are publically funded. The benefit should not be calculated on how much the data can be sold for but rather based on the improvements in policy, research and public engagement. Improving access to data will improve the knowledge base which improves research and governance: think provision, not possession (Isaksen 2009).

Opening access to easily re-usable knowledge will start to produce a dynamic knowledge base. The archaeology knowledge base should be, by definition, dynamic: It is predicated on the complex relationship between the corpus of knowledge, theory and classification systems. These are fluid and contain many interlinked dependencies which means that variations in one component can have complex repercussions throughout the knowledge base. The issue then is not about whether it will happen but how we can link this data to enhance its re-use and deal with the organisational and ethical problems: how is access mediated, what facilitates deposition, how re-use communities are created and what incentives are required. It is also important that we articulate the technical, social, organisational, conceptual and cultural issues so that we can deliver appropriate data and knowledge services to the appropriate people so that governance and policy can be enacted more effectively.

If this is the case then we need Open approaches to ensure that we have the capacity to exploit these resources effectively for the full range of stakeholders.

As stated by English Heritage:
Knowledge is the prerequisite to caring for England's historic environment. From knowledge flows understanding and from understanding flows an appreciation of value, sound and timely decision-making, and informed and intelligent action. Knowledge enriches enjoyment and underpins the processes of change (English Heritage 2005).

References

Department for Communities and Local Government. 2012. National Planning Policy Framework. Publication (Legislation and policy). http://www.communities.gov.uk/publications/planningandbuilding/nppf.

English Heritage. 1995. Research Frameworks: Project Report (3rd Draft). London. http://www.eng-h.gov.uk/frameworks/bericht4.pdf.

English Heritage. 2005. Making the Past Part of Our Future. http://www.english-heritage.org.uk/content/publications/publicationsNew/eh-strategy-2005-2010/making-past-part-of-our-future.pdf.

INSPIRE. 2012. 'Inspire > Welcome to Inspire.' http://inspire.jrc.ec.europa.eu/.

Isaksen, Leif. 2009. 'Pandora's Box: The Future of Cultural Heritage on the World Wide Web'.

DocStoc. http://www.docstoc.com/docs/20622714/Pandoras-Box-The-Future-of-Cultural-Heritage-on-the.

MORI. 2000. The Role of Scientists in Public Debate. Wellcome Trust.

OKF. 2012a. 'Open Knowledge Foundation.' http://okfn.org/.

OKF. 2012b. 'Open Definition.' http://opendefinition.org/.

Poliakoff, Ellen, and Thomas L Webb. 2007. 'What Factors Predict Scientists' Intentions to Participate in Public Engagement of Science Activities?' Science Communication 29 (2) (December 1): 242–263. doi:10.1177/1075547007308009.

RCUK. 2012b. 'RCUK Common Principles on Data Policy – RCUK.' http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx.

Royal Society, RCUK, and Wellcome Trust. 2006. Science Communication: Survey of Factors Affecting Science Communication by Scientists and Engineers. http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/publications/2006/1111111395.pdf.

Wenger, Etienne, Richard Arnold McDermott, and William Snyder. 2002. Cultivating Communities of Practice: a Guide to Managing Knowledge. Boston, Massachusetts: Harvard Business School Press.

Wellcome Trust. 2012. 'Wellcome Trust Strengthens Its Open Access Policy.' http://www.wellcome.ac.uk/News/Media-office/Press-releases/2012/WTVM055745.htm.

Wikipedia. 2012a. 'Information Economy.' http://en.wikipedia.org/wiki/Information_economy.

Wikipedia. 2012b. 'Knowledge Economy.' http://en.wikipedia.org/wiki/Knowledge_economy.

Wikipedia. 2012c. 'Community of Practice.' http://en.wikipedia.org/wiki/Community_of_practice.

Wikipedia. 2012d. 'Linked Data.' http://en.wikipedia.org/wiki/Linked_data.

Wikipedia. 2012e. 'Semantic Web.' http://en.wikipedia.org/wiki/Semantic_Web.

Datenaustausch und Datenabgabe[1]

Otto Mathias Wilbertz

Eines der Ziele der Kommission bestand darin, Wege für eine gemeinsame Darstellung georeferenzierter Fachdaten zu suchen. Dies führte zur Entwicklung des Datenexportstandards ADeX (Archäologischer Daten-eXport-Standard), der einen breiten, leicht praktikablen Datenaustausch ermöglicht. Dieser wird gerade für bundesländerübergreifende Projekte der Landesämter für Bodendenkmalpflege[2] als notwendig erachtet, sei es für grenzübergreifende Baumaßnahmen, wie Straßen und Versorgungsleitungen, oder grenzübergreifende Bodendenkmäler, wie zum Beispiel dem Limes, oder aber gemeinsame Forschungsprojekte.

Da sich Datenaustausch und -abgabe jedoch nicht nur auf die Landesämter untereinander beschränken, sondern diverse Partner infrage kommen, galt es, diese zu benennen und Konditionen zur Weitergabe von Daten festzulegen. Hierfür wurde die Arbeitsgruppe „Rechtliche und wirtschaftliche Aspekte der Datenweitergabe" (RWAD) gegründet. Deren Überlegungen seien im Folgenden kurz skizziert.

Generell werden von den Landesämtern sowohl Sach- als auch GIS-Daten abgegeben, ausgetauscht oder auch übernommen. Als Partner dafür kommen die Landesämter untereinander, Partner in der öffentlichen Verwaltung, privatwirtschaftliche Partner, Forschungseinrichtungen, touristische Einrichtungen, aber auch private Interessenten in Betracht. Erwähnt sei, dass Bayern seine archäologischen Denkmale in einem Portal („Bayern-Viewer") öffentlich zur Verfügung stellt.

Da es sich bei Bodendenkmaldaten zum Teil um sensible Daten handelt, muss ein verantwortungsvoller, fach- und sachgerechter Umgang gewährleistet sein. Abhängig von den Partnern, Projekten und Maßnahmen sollten deshalb Nutzungsbedingungen, die Vergabe oder gegebenenfalls auch der Entzug von Zugriffsberechtigungen geregelt sein.

Der Datenaustausch zwischen den Landesämtern wird als selbstverständlich angesehen, da ein angemessener Umgang mit den zum Teil sensiblen Daten vorausgesetzt wird. Eventuell sind projektabhängige Vereinbarungen erforderlich, die Art, Umfang, Verwendungszweck und Dauer der Verfügbarkeit der Daten sowie die zugriffsberechtigten Personen benennen.

Der Datenaustausch zwischen Landesämtern und den Partnern in der öffentlichen Verwaltung gehört zu den Pflichtaufgaben und wird bereits gepflegt. Die gegebenenfalls erforderlichen Nutzungsvereinbarungen entsprechen den zuvor genannten.

Der Datenaustausch zwischen Landesämtern und privatwirtschaftlichen Partnern, wie Planungsbüros oder Grabungsfirmen, ist gemäß den vorliegenden Erfahrungen ähnlich wie mit Partnern in der öffentlichen Verwaltung zu handhaben. Vertragliche Vereinbarungen werden hier als notwendig erachtet, die Art, Umfang und Verwendungszweck der Daten sowie die zugriffsberechtigten Personen, eine zweckgebundene Nutzung und die Verpflichtung, die Daten nicht weiterzugeben, beinhalten sollten. Für die Grabungsfirmen ist die Abgabe ihrer im Rahmen von Grabungen erhobenen Daten an die Landesämter obligat. Hier existieren zum Teil bereits vertragliche Vereinbarungen in den Landesämtern.

Für den Datenaustausch zwischen Landesämtern und Forschungseinrichtungen, wie Universitäten, werden vertragliche Vereinbarungen als notwendig erachtet, die Art, Umfang und Verwendungszweck der Daten sowie die zugriffsberechtigten Personen, eine zweckgebundene Verwendung, die Verpflichtung, die Daten nicht weiterzugeben, Publikationsregelungen/-genehmigungen, die Wahrung der Urheberrechte und des Copyrights beinhalten sollten. Bezüglich des Copyrights besteht die Möglichkeit, die Datensätze beim Export mit einem Vermerk zu versehen. In den Landesämtern existieren ebenfalls zum Teil bereits vertragliche Vereinbarungen.

Vor der Abgabe von Daten von Landesämtern an touristische Einrichtungen stellt sich die Frage, welche Daten überhaupt und mit welcher Lagegenauigkeit abgegeben werden. Als unproblematisch erscheint die Abgabe von Objektinformationen und Bilddaten zu obertägig sichtbaren archäologischen Denkmalen, auch mit Angaben zur genauen Verortung.

Untertägige Bodendenkmäler sollten hingegen – wenn überhaupt – nur mit ungenauer Verortung (zum Beispiel dem Gemeindemittelpunkt als Lageangabe) weitergegeben werden. Die Datenabgabe an private Interessenten (Informationsfreiheitsgesetz) erfolgt nach Einzelfallentscheidung, wobei die Interessen eines möglichst freizügigen Zugangs zu archäologischen Informationen einerseits und des sich aus der besonderen Sensibilität ergebenden spezifischen Schutzinteresses andererseits fallbezogen abgewogen werden.

Die Frage, ob Gebühren beziehungsweise Bereitstellungskosten für die Datenabgabe erhoben werden sollten oder können, wurde in der AG kontrovers diskutiert.

Bei diesem Stand der Diskussion wurde die Arbeitsgemeinschaft mit der EU-Richtlinie „INSPIRE" (Infrastructure for Spatial Information in Europe; „Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft") konfrontriert. Mit dieser Richtlinie besteht mittlerweile die Verpflichtung, auch Schutzgebiete („Protected Sites"), konkret „archäologische Stätten", und somit auch Bodendenkmaldaten öffentlich in Geoportalen zur Verfügung zu stellen. Dies sind Geometrie, Identifikationsnummer beziehungsweise -schlüssel, Unterschutzstellungsdatum sowie -dokument, rechtliche Grundlage, Name und Klassifikation.

Abhängig vom Denkmalschutzgesetz des jeweiligen Bundeslandes sind von dieser Datenbereitstellung Landesämter für Bodendenkmalpflege, aber auch Untere Denkmalbehörden betroffen, je nach Festlegung, wer die Denkmallisten führt. In diesem Zusammenhang erarbeitet die AG „Rechtliche und wirtschaftliche Aspekte der Datenweitergabe" derzeit eine Empfehlung zur Harmonisierung der Metadaten. Da die Daten ohne jeweilige Kenntnis der datengebenden Institution abgerufen werden können, empfiehlt die AG alle oben genannten Einschränkungen und Erklärungen (wie die jeweilige gesetzliche Grundlage, Bedeutung der Eintragung, Nutzungsbeschränkungen und anderes mehr) im Vorfeld darstellbar zu machen. Ob dies über eine Authentifizierung des Datennehmers ermöglicht werden kann oder ob andere Wege möglich sind, ist ebenso zu klären wie die Frage, ob und gegebenenfalls wie auf die Verwendung Einfluss genommen werden kann. Wichtig ist auch darauf hinzuweisen, dass die eingestellten Daten keineswegs sämtliche Fundplätze der Bundesländer darstellen und die zuständigen Behörden bei Planungen weiterhin kontaktiert und eingebunden werden müssen.

Amerkungen

[1] Der Beitrag beruht auf der Arbeit der AG „Rechtliche und wirtschaftliche Aspekte der Datenweitergabe" der Kommission „Archäologie und Informationssysteme". Den Mitgliedern der AG, insbesondere Michaela Aufleger, die das grundlegende Protokoll verfasst und wichtige Anregungen für das endgültige Manuskript beigesteuert hat, sei an dieser Stelle herzlich gedankt.

[2] Mit dem Ausdruck „Landesamt für Bodendenkmalpflege" werden im Folgenden die Dienststellen der Länder bezeichnet, die für die archäologische Denkmalpflege zuständig sind (Landesämter beziehungsweise Abteilungen oder Referate).

THESAURUSFRAGEN

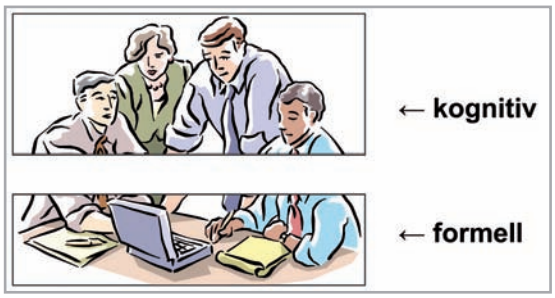
Semantische Interoperabilität

Reiner Göldner

Einführung

„Je weiter wir in die Vergangenheit vordringen, desto weniger haben die betrachteten kulturellen Räume mit den heutigen politischen Grenzen gemeinsam. Archäologische Forschungen können daher nicht an Ländergrenzen aufhören, und auch der Denkmalschutz profitiert vom Blick zum Nachbarn. Durch die Kulturhoheit der Länder ist in Deutschland eine stark differenzierte Welt archäologischer Informationssysteme entstanden, die den Datenaustausch nicht gerade erleichtert." (vgl. S. 10)

Immer wieder ergibt sich die Notwendigkeit, archäologische Fachinformationen auszutauschen beziehungsweise aus verschiedenen Quellen zusammenzuführen. Dies ist für den engagierten Geisteswissenschaftler normalerweise kein Problem, er wird den Text sorgfältig durchlesen und im Kopf seine Schlüsse daraus ziehen. Den Informationstechniker stellt dies jedoch nicht recht zufrieden, er zieht es vor, die Frage formell mit dem PC zu lösen. Außerdem, wer erfasst denn heute noch seine umfangreichen Fundkomplexe in Form von freiem Text? Die Vorteile des PC liegen eben gerade bei formellen Verfahren und strukturierten Daten. Also auch wenn der menschlich-kognitive[1] Ansatz bei kleineren Datenmengen recht erfolgreich sein kann, wird man sich bei größeren Datenmengen schon bald nach einem formellen Verfahren sehnen.



1 Kognitiver und formeller Lösungsansatz.

Die Kommission „Archäologie und Informationssysteme" beim Verband der Landesarchäologen hat sich dieses Themas angenommen, um den Datenaustausch zwischen den archäologischen Landesämtern und mit anderen Fachinstitutionen zu fördern. Dazu wurde das Datenaustauschformat ADeX[2] entwickelt, welches technische und inhaltliche Rahmenbedingungen für einen Austausch von georeferenzierten archäologischen Fachdaten beinhaltet.

ADeX beschränkt sich gegenwärtig stark auf die Struktur der auszutauschenden Fachdaten, da bisher kein einheitliches kontrolliertes Vokabular zur Verfügung steht. Die Bedeutung der verwendeten Strukturelemente (Attribute) wurde einheitlich festgelegt und beschrieben, die Bedeutung der zu transportierenden Inhalte (Attributwerte/Vokabular) bleibt jedoch weitgehend offen und muss vom Nutzer individuell berücksichtigt werden. Die Nutzung von kontrolliertem Vokabular ist für die systematische Erfassung archäologischer Sammlungsbestände sowie die Beschreibung archäologischer Fundstellen natürlich auch im Bereich der Archäologie seit Langem üblich. Wenn man die Situation in Deutschland betrachtet, ist dieses Vokabular jedoch nicht homogen. Durch den stark regional bezogenen Kontext der Forschung, unterstützt durch die Kulturhoheit der Länder, hat sich eine Vielzahl verschiedener und unterschiedlicher Systeme herausgebildet, sodass wir diesbezüglich sehr heterogene Datenbestände vorfinden. Diese Situation ist bezüglich des ADeX-Datenaustausches nicht zufriedenstellend, weil die inhaltliche Ebene des Datenaustausches stark vernachlässigt ist und sich formellen Methoden weitgehend entzieht. Daten unterschiedlicher Herkunft harmonisieren mit ADeX zwar syntaktisch, aber nicht semantisch.

Kann man die benutzten Daten und Vokabulare harmonisieren? Kann man den Datenaustausch mit Semantik anreichern und damit „semantische Interoperabilität" erreichen? Zur Beantwortung dieser Fragen dienen nachfolgende Exkurse in die Themenbereiche Modellierung, Semantik und Harmonisierung. Einige Aspekte dieser Themenbereiche werden dabei verallgemeinert dargestellt. Das Hauptaugenmerk wird jedoch auf Datenbeständen beziehungsweise Datenbanken liegen, weil diese gegenwärtig die größte Bedeutung für die systematische Dokumentation von Sachverhalten der realen Welt besitzen.

Modellierung

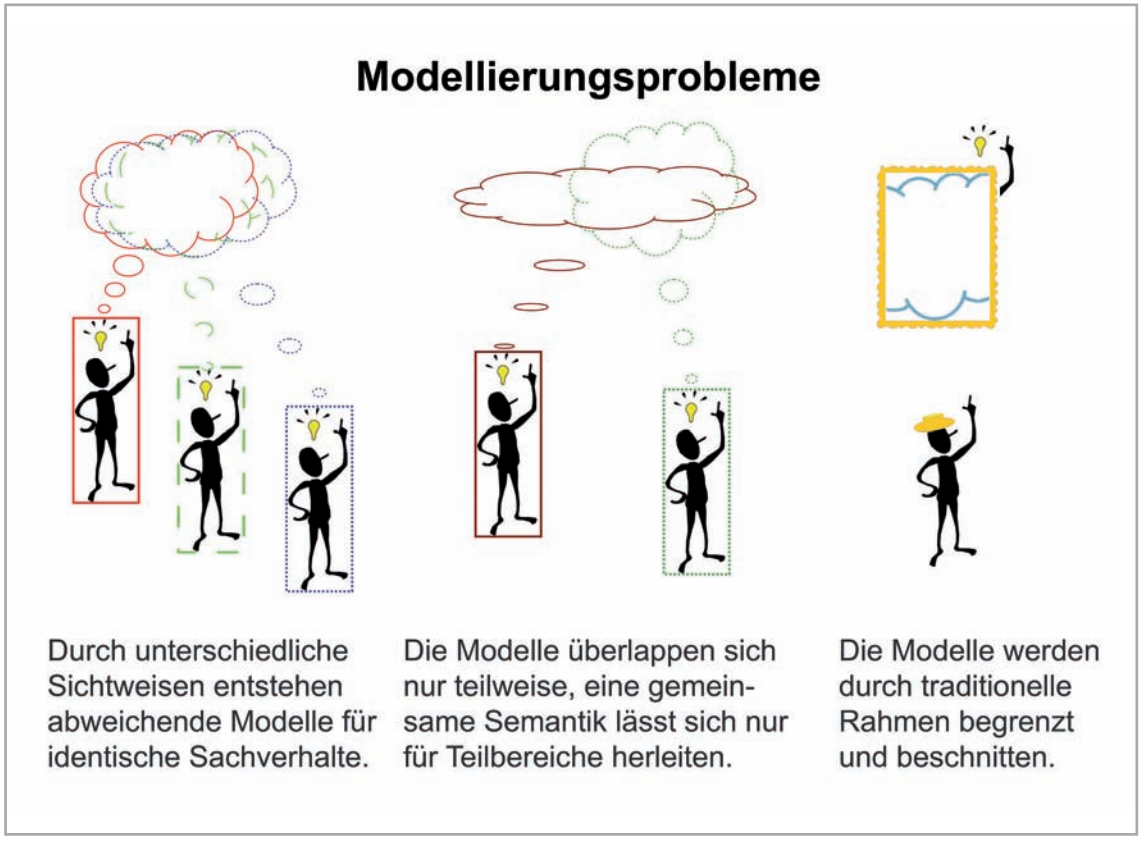
Zur systematischen Beschreibung von Phänomenen der realen Welt dienen Modelle. Solche Modelle sind Abbildungen der Realität, wobei von Unwesentlichem abstrahiert wird und nur für die Betrachtung wesent-

liche Eigenschaften übernommen werden. Werden die Modelle auf wissenschaftlicher Grundlage erstellt, so können die bestehenden wissenschaftlichen Methoden auf das Modell angewandt werden und zu neuen Erkenntnissen führen.

Modelle können für einzelne Objekte oder Sachverhalte, aber auch für größere Mengen von zu beschreibenden Objekten aufgestellt werden (Klassenmodelle). Letzteres ist typisch für Datenbanken und aufgrund deren Bedeutung werden wir uns in diesem Artikel immer wieder auf sie beziehen. Bei Datenbanken gilt das Modell (Datenbankschema) für eine ganze Klasse von Objekten, wobei sich die Eigenschaften des Klassenmodells auf die einzelnen konkreten Objekte vererben. Diese einzelnen konkreten Objekte werden dann wiederum durch Attributwerte näher beschrieben. Und auch diese Attributwerte können wiederum einem Modell entstammen, beispielsweise in Form eines Thesaurus.

Moderne Modelle sind oft theoretischer und formeller Natur, weil sie so computergestützt aufgestellt und verarbeitet werden können. Typische Vertreter dazu sind zum Beispiel Listen, Tabellen oder Datenbanken, in denen die Eigenschaften der zu modellierenden Objekte nach bestimmten Regeln und Vereinbarungen erfasst sind. Für eine gesicherte Interpretation solcher Modelle muss ihre Struktur auch inhaltlich beschrieben werden. Diese Anforderung wird leider allzu oft vernachlässigt, weil man als Bearbeiter die Inhalte als naheliegend und selbsterklärend begreift. In der Praxis wird daher häufig auf der Ebene der Daten diskutiert, wo man eigentlich das Modell betrachten müsste.

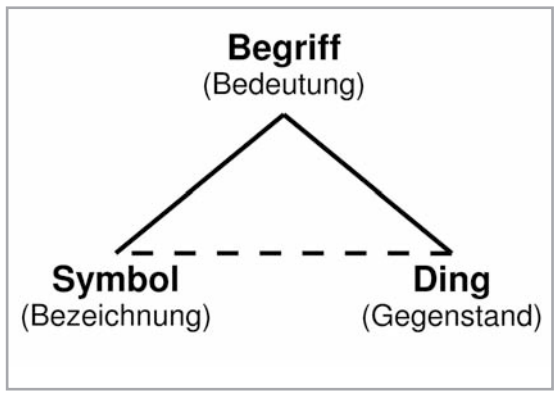
Es gibt einige grundsätzliche Probleme der Modellierung, die sich ungünstig auf die Harmonisierung auswirken können (Abb. 2). Modellierung ist ein subjektiv beeinflusster Prozess, daher werden die Modelle verschiedener Experten zum gleichen Sachverhalt nicht exakt übereinstimmen. Außerdem werden die Modelle in ihrer „Form“ von äußeren Randbedingungen mitbestimmt, für eine Harmonisierung kommen natürlich nur die Teilbereiche in Betracht, die einander überschneiden. Nicht zuletzt werden neue Modelle wesentlich durch ihre Vorgänger bestimmt, wenn der Datenbestand es erfordert oder die Tradition. Dabei treten dann moderne Modellierungsansätze eher in den Hintergrund. Für eine Harmonisierung sollte man diese Probleme so weit wie möglich reduzieren. Auf jeden Fall sollten die benutzten Methoden gegen solche Modell-„Abweichungen“ tolerant sein.



2 Modellierungsprobleme.

Semantik

Die Eigenschaften modellierter Objekte werden meist durch (Mess-)Werte oder Begriffe dargestellt, letzteren wollen wir uns nun etwas näher widmen. Ein „Begriff" ist ein Bedeutungsinhalt, der uns gedanklich ein „Ding" der Realität vermittelt und den wir mit einem „Symbol" (einer Bezeichnung) darstellen. Dies wird durch das semiotische Dreieck (Abb. 3) gut veranschaulicht. Für die Kommunikation und Zusammenarbeit ist es nun entscheidend, dass alle Partner dieselben Begriffe und Bezeichnungen verwenden, also dasselbe Vokabular. Die Fähigkeit von technischen Systemen, zur Zusammenarbeit und zum Austausch von Informationen auf der Ebene von Bedeutungen wird als „semantische Interoperabilität" bezeichnet.



3 Semiotisches Dreieck.

Die Verwendung von gegenseitig abgestimmtem, kontrolliertem Vokabular ist in der Wissenschaft allgegenwärtig. Kontrolliertes Vokabular ist die Sammlung von Begriffen und ihren Bezeichnungen, wobei die Zuordnung von Bezeichnungen zu Begriffen eindeutig ist und keine Homonyme verwendet werden. Wichtige Formen kontrollierten Vokabulars sind:

- Schlagwortliste: (gegliederte) Liste von Schlagworten, Deskriptoren (Bezeichnungen), zum Beispiel Schlagwortnormdatei,
- Glossar: Liste von Bezeichnern (Worten) mit genauen Beschreibungen,
- Thesaurus: kontrolliertes Vokabular, dessen Begriffe über Relationen verknüpft sind (hierarchisch, polyhierarchisch, auch vernetzt), typische Relationen: Ober-/Unterbegriff, Synonym, auch Verwandtschaft.

Auf kontrolliertes Vokabular kann man sich bei der Kommunikation also verlassen. So weit, so gut, aber die Praxis ist oft komplizierter. Da gibt es eben verschiedene wissenschaftliche Schulen, regionale Differenzen und nicht zuletzt auch noch sprachliche Unterschiede. Die daraus resultierenden Vokabulare sind dann, gemeinsam betrachtet, plötzlich inkonsistent und mehrdeutig und damit im Sinne der semantischen Interoperabilität nutzlos. Was tun?

Semantische Modelle

Bei der semantischen Modellierung wird den Modellelementen eine Bedeutung zugewiesen. Der erste Schritt ist dabei immer die Beschreibung der Semantik durch einen Fachexperten. Dies erfolgt mit den fachtypischen Beschreibungsmitteln, zum Beispiel in natürlicher Sprache, mit Abbildungen, Diagrammen usw. Anschließend wird diese Beschreibung durch einen Modellierungsexperten formalisiert und strukturiert. Dieser Prozess kann durch die Anwendung einer Ontologiesprache unterstützt werden und führt damit zu einer Ontologie.

Bezogen auf Datenbanken lassen sich nun zwei Ebenen der semantischen Modellierung unterscheiden, die strukturelle Ebene (Klassenmodell) und die Werteebene (Objektbeschreibung). Die strukturelle Ebene umfasst die Klassenattribute und deren Struktur (es muss also zum Beispiel die Bedeutung der Tabellenspalten festgelegt werden, gegebenenfalls auch die Verknüpfung verschiedener Tabellen). Die Werteebene kommt bei der Verwendung kontrollierten Vokabulars ins Spiel und umfasst vor allem den Typ und gegebenenfalls auch die Struktur des benutzten Vokabulars, also zum Beispiel eines Thesaurus.

Für eine computertechnische Verarbeitung müssen semantische Modelle formalisiert werden (der Computer muss die Sprache verstehen, in der das Modell beschrieben ist). Dies kann durch die Nutzung von Ontologien erreicht werden. Eine Ontologie ist in der Informatik eine formelle Darstellung von Wissen durch Begriffe und Beziehungen mit zugehörigen Regeln für Gültigkeit (Integrität) und Schlussfolgerungen (Inferenz). Man kann Ontologien als hochkomplexe Formen kontrollierten Vokabulars auffassen, ein Thesaurus ist demzufolge eine stark beschränkte Ontologie. Zur Beschreibung von Ontologien nutzt man Ontologiesprachen, wie zum Beispiel:

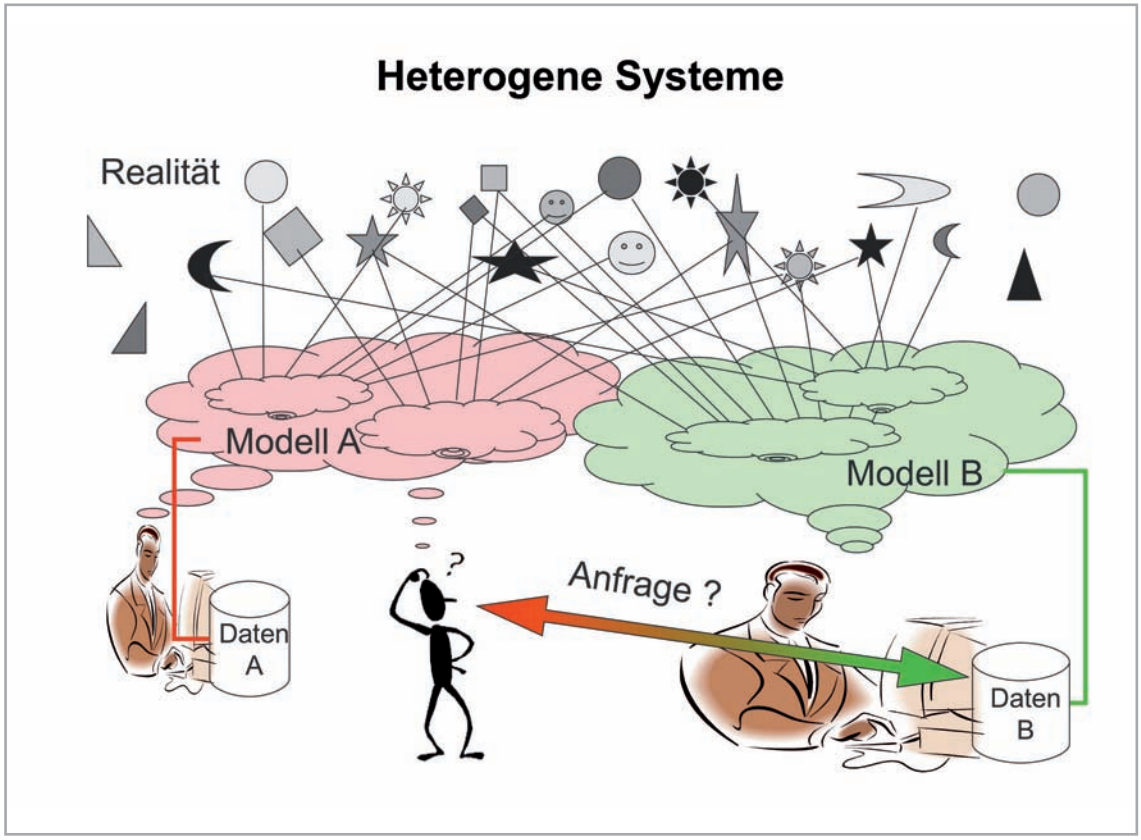
- OWL (Web Ontology Language): Sprache für den Aufbau von Ontologien zum Internet
- CIDOC CRM: ISO-Standard für den Aufbau von Ontologien im Bereich Kulturerbe
- TOPIC Maps: ISO-Standard zur Formulierung von Wissensstrukturen.

Die semantische Modellierung ist ein Grundelement der systematischen Verarbeitung von Daten. Sie ist Voraussetzung dafür, dass die Daten auch außerhalb der originären Umgebung verwendet werden können. Dies bedeutet einerseits Interoperabilität, aber auch Nachhaltigkeit. Semantische Modellierung ist also auch eine wichtige Voraussetzung für die Archivierung digitaler Daten.

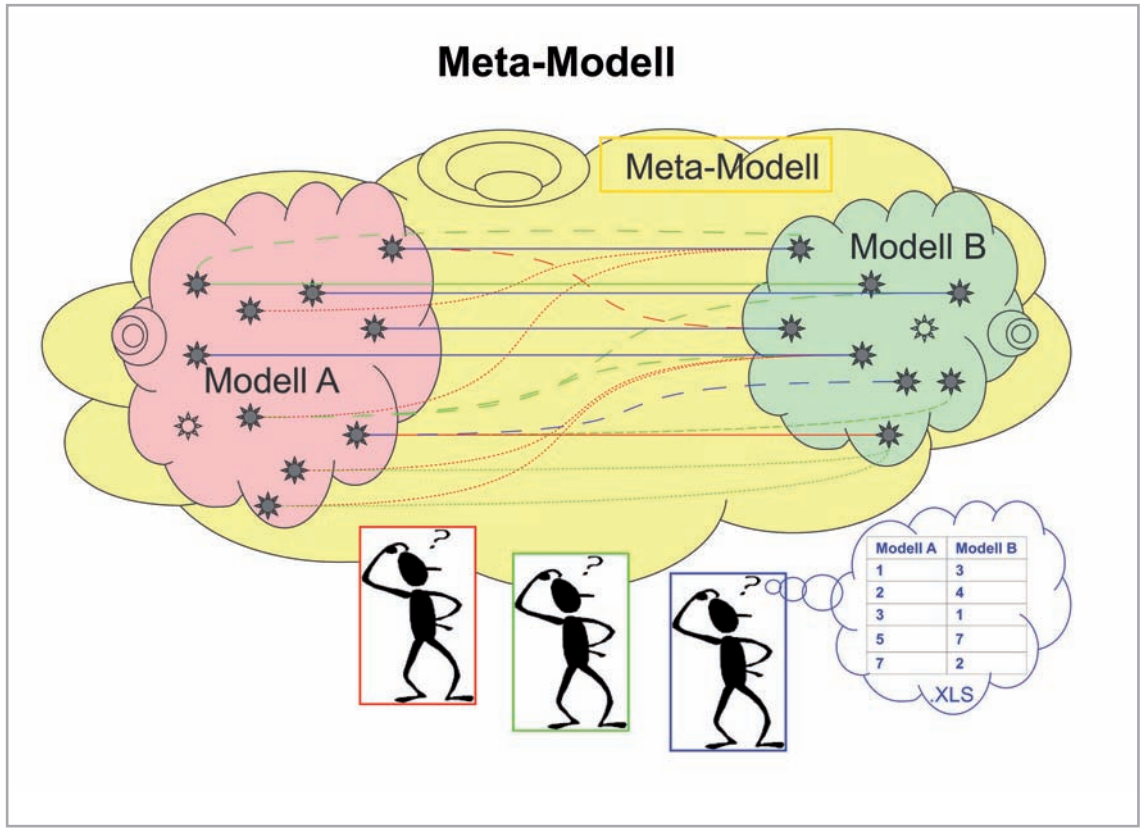
Harmonisierung

Fügt man mehrere Modelle zusammen, so ist es tatsächlich kaum zu vermeiden, dass erst einmal Inkonsistenzen und Mehrdeutigkeiten entstehen. Man bekommt das bekannte Äpfel-Birnen-Problem, die sprichwörtliche Unvergleichbarkeit[3]. Bei Datenbanken betrifft dies dann natürlich auch die Daten (Abb. 4). Wer Modell A im Hinterkopf hat, findet einerseits nicht die richtigen Fragen an Modell B und kann andererseits mit den Antworten nicht viel anfangen.

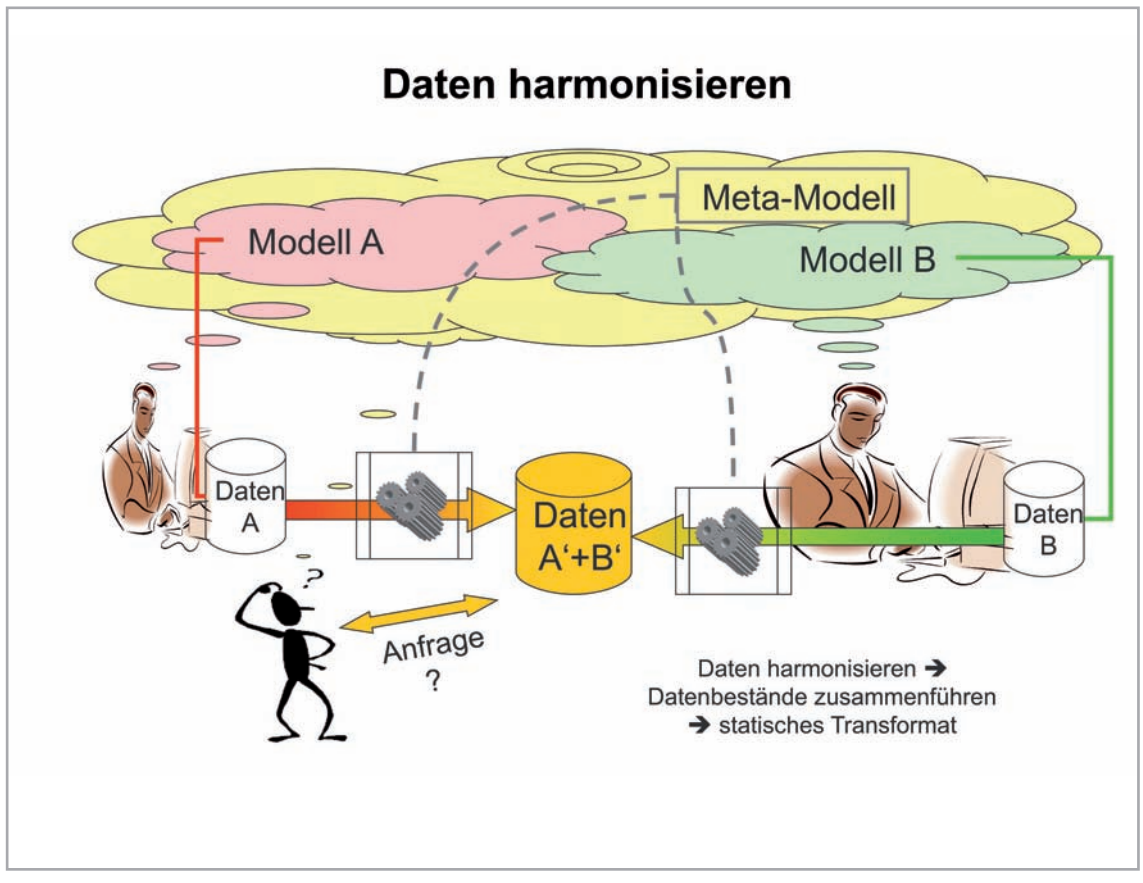
Am besten wäre es natürlich, wenn sich die zu harmonisierenden Modelle schon einer gemeinsamen Sprache bedienen, aber dies ist leider die Ausnahme. Um Inkonsistenzen und Mehrdeutigkeiten zu beseitigen, muss man Modelle und Daten harmonisieren. Modelle und Daten müssen gemeinsam auf dieselbe Art und Weise betrachtet, verglichen und einander angeglichen werden. Dazu bedient man sich am besten eines übergeordneten Meta-Modells, welches die Gemeinsamkeiten und Unterschiede der zusammenzuführenden Modelle beinhaltet (Abb. 5). Dieses Meta-Modell muss nicht konkret ausgebildet sein, seine Nutzung kann intuitiv oder diskursiv sowie mehr oder weniger formell erfolgen (vgl. Abb. 1). Die formellen Methoden haben dabei den entscheidenden Vorteil, dass sie sich technisch umsetzen und damit automatisieren lassen.



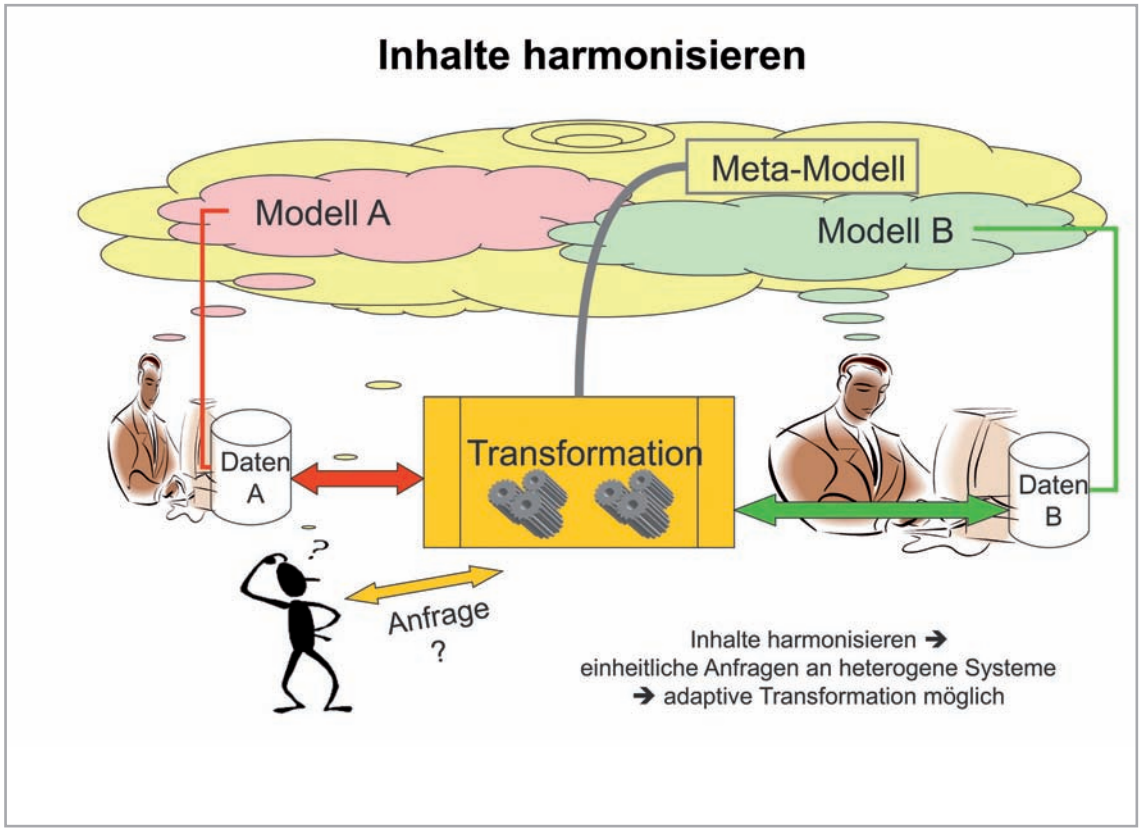
4 Heterogene Systeme.



5 Meta-Modell.



6 Daten harmonisieren.



7 Inhalte harmonisieren.

Für eine semantische Modellierung muss natürlich auch das Meta-Modell Semantik beinhalten. Eine Möglichkeit, um ein solches Meta-Modell formell, also systemtechnisch lesbar aufzustellen, bilden wiederum Ontologien.

Mit Hilfe des Meta-Modells lassen sich nun entweder die Datenbestände zusammenführen (Daten harmonisieren, Abb. 6) oder es lassen sich Transformationen bilden, die auf Anfragen an die nach wie vor heterogenen Systeme konsistente Ergebnisse liefern (Inhalte harmonisieren, Abb. 7). Der zweite Weg ist interessanter, weil er die ursprünglichen Daten nicht verändert und jederzeit adaptiv an neue Fragestellungen angepasst werden kann.

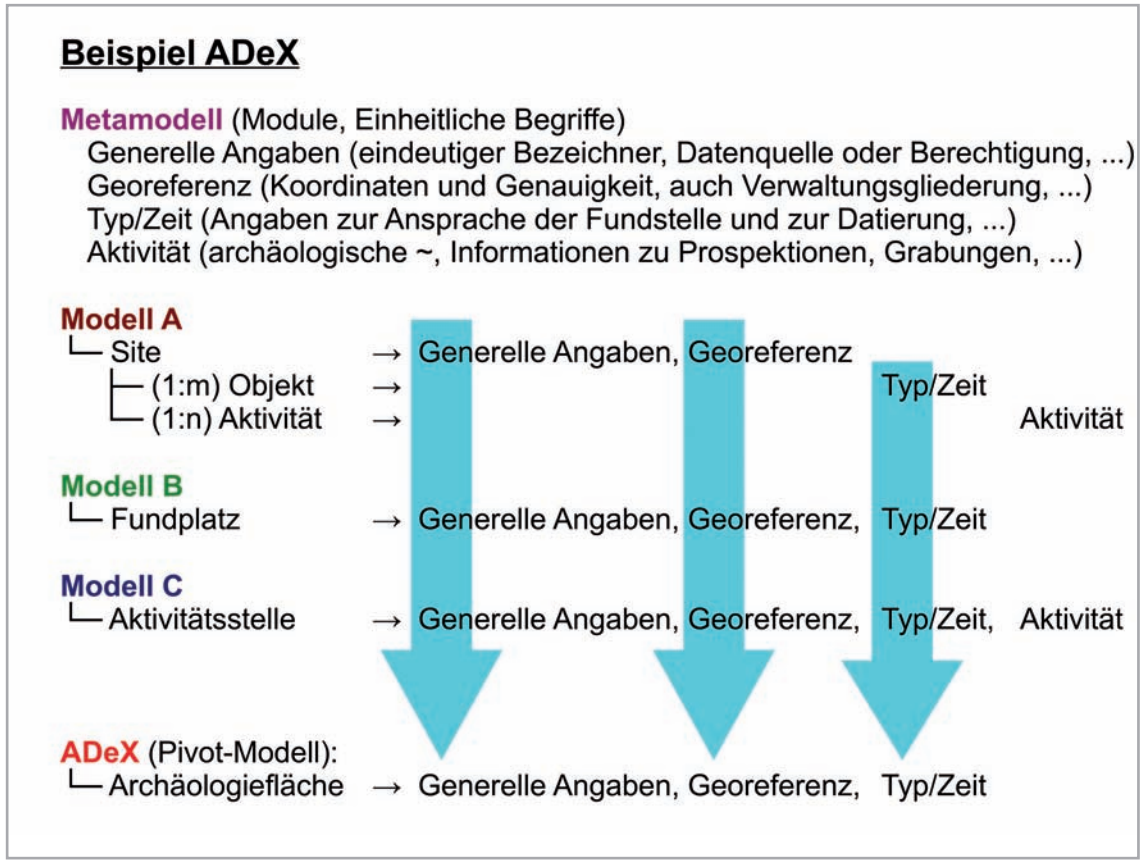
Praxis

Wie erreicht man nun semantische Interoperabilität in der Praxis? Wie harmonisiert man Modelle, Daten oder Vokabulare?

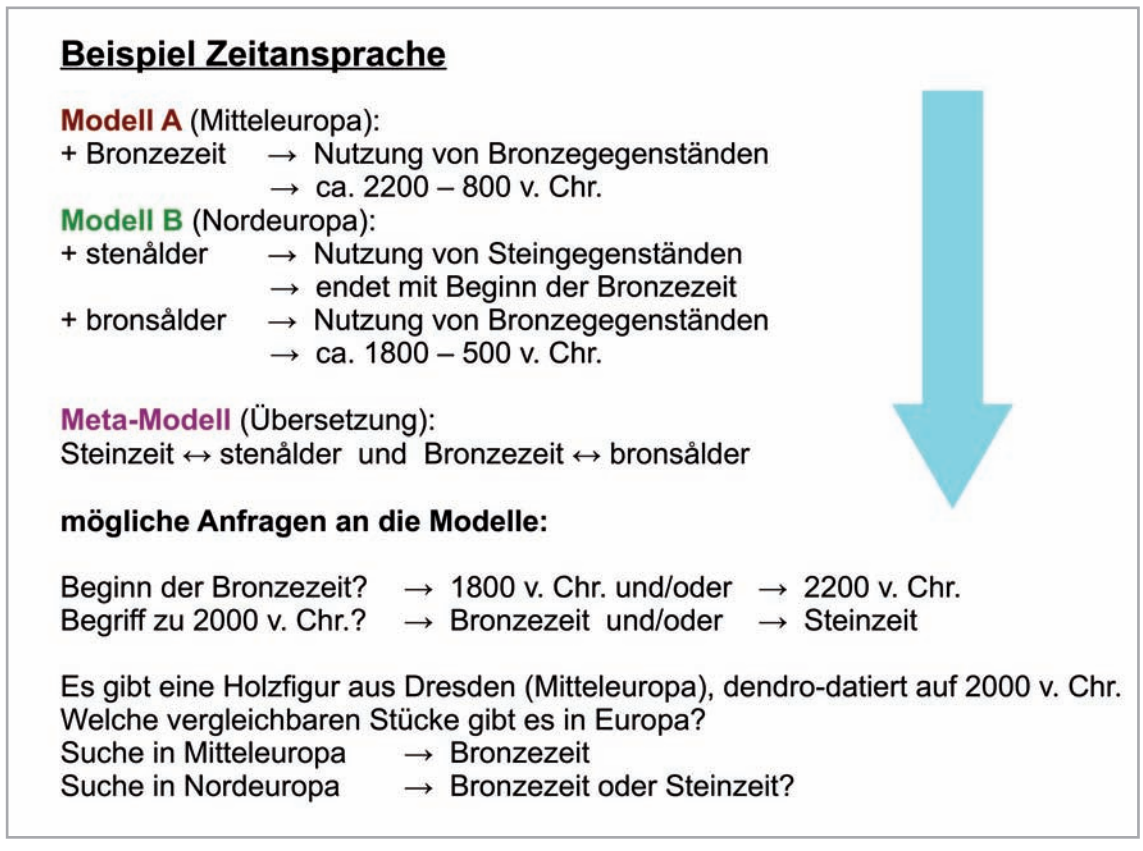
Die einfachste Form, für kleine Datenbestände und übersichtliche Modelle praktikabel, ist die kognitive Umsetzung durch eine engagierte und kreative Person oder Arbeitsgruppe. Man überlegt sich einfach, welche Begriffe synonym zu verwenden sind, gruppiert gegebenenfalls hier und da neu und übersetzt vielleicht in eine einheitliche Sprache. Das Meta-Modell befindet sich dabei im Kopf des Bearbeiters. So oder ähnlich funktioniert die Auswertung wissenschaftlicher Fachartikel, die nicht exakt der gleichen Schule entstammen oder gar in verschiedenen Sprachen abgefasst sind. Jeder Wissenschaftler bedient sich dieser Methode regelmäßig.

Bei größeren Modellen und umfangreicheren Datenbeständen wird man allerdings schnell auf formelle Methoden ausweichen, um nicht den Überblick zu verlieren. Üblicherweise greift man dann zuerst zu Listen, Tabellen oder Thesaurusstrukturen, um beispielsweise Identitäts- und Zugehörigkeitsbeziehungen oder andere einfache, schwach strukturierte Beziehungen darzustellen. Aus Identitätsbeziehungen ergeben sich einfache, leicht anwendbare Transformationen, mit denen Datenbestände besser vergleichbar gemacht werden können, indem beispielsweise Vokabulare maschinell ausgetauscht und vereinheitlicht werden. Bei Datenbanken lassen sich solche Transformationen einfach realisieren, zum Beispiel über die standardmäßig verfügbaren Joins.

Wenn nun allerdings stark strukturierte und differenzierte Beziehungen aller Art eine Rolle spielen, kommt man mit Listen und Tabellen nicht mehr weiter, weil



8 Beispiel ADeX.



9 Beispiel Zeitansprache (Steinzeit =? stenålder, Bronzezeit =? bronsålder).

sich damit die erforderlichen komplexen Beziehungen nicht abbilden lassen. Hier können Ontologien mit ihren umfangreichen Möglichkeiten der Wissensrepräsentation und ihren Schlussfolgerungsmechanismen (Inferenz) weiterhelfen. Dabei sind die Schlussfolgerungsmechanismen entscheidend, denn sie ermöglichen letztendlich den Vergleich und damit die Interoperabilität. Ohne sie bleiben Ontologien rein akademisch. Eine geschickte Modellierung erlaubt auch bei komplexen Modellen automatisierte Transformationen im Datenbestand.

Bezogen auf die Semantik von Datenbanken gelten die angeführten Methoden gleichermaßen für die strukturelle Ebene (Klassenmodell) und die Werteebene (Objektbeschreibung). Dies soll nachfolgend an den beiden Beispielen zu ADeX und zur Zeitansprache deutlich gemacht werden.

ADeX verwendet ein Meta-Modell, in dem vor allem die Begriffe definiert sind. Im Beispiel (Abb. 8) soll dies vereinfachend durch die Module (Generelle Angaben, Georeferenz, Typ/Zeit, Aktivität) dargestellt werden. Dieses Meta-Modell kann nun zur Beschreibung der bestehenden Original-Datenstrukturen angewandt werden und man erhält beispielsweise die Modelle A, B und C als eine vereinheitlichte Sichtweise auf die Originaldaten. Aus diesen Modellen wurde ADeX als Pivot-Modell abgeleitet und damit gilt das Meta-Modell mit seiner Semantik natürlich auch für ADeX. Diese einheitliche Sicht ist Voraussetzung, um Transformationen zu finden, die die jeweilige Originaldatenstruktur in die ADeX-Datenstruktur abbilden. Im Beispiel werden die Module aus Modell B einfach übernommen, bei Modell C trifft man eine Auswahl („Aktivitäten" weglassen) und bei Modell A wird neben der Auswahl noch ein Verbund (Join) von „Generelle Angaben" und „Georeferenz" mit „Typ/Zeit" erforderlich.

Das Beispiel Zeitansprache (Abb. 9) soll Probleme verdeutlichen, die mit unterschiedlich definierten kontrollierten Vokabularen einhergehen. Der Begriff Bronzezeit beschreibt, allgemein und vergröbert, einen Zeitraum, in dem Bronzegegenstände hergestellt und genutzt wurden. Auch wenn man eine ungefähre Vorstellung über diesen Zeitraum hat, so lässt sich dies nur anhand absoluter Zahlen konkretisieren, wenn man sich auf gewisse Regionen beschränkt. Bronzezeit fand eben in Nordeuropa später als in Mitteleuropa statt. Dies führt aber dazu, dass der Attributwert „Bronzezeit" in einem Datensatz aus Nordeuropa gegebenenfalls anders zu bewerten ist als in einem Datensatz aus Mitteleuropa. Es ist also in diesem Fall wichtig, dass die Zeitstellung in ihrem regionalen Kontext betrachtet wird. Dies lässt sich im Text eines wissenschaftlichen Artikels leicht nachvollziehen, in einer überregionalen Sammlung von Datensätzen muss es explizit berücksichtigt, das heißt modelliert werden.

Eine Lösung kann es sein, die Begriffe der benutzten Zeitansprachen regional zu differenzieren und mit absoluten Datierungen zu versehen. Auf die Frage nach dem Beginn der Bronzezeit bekommt man dann eine regional abhängige Antwort. Gleichzeitig wird deutlich, dass man auch die Anfragen sorgfältig formulieren muss. Bei der Suche nach vergleichbaren Stücken zu einer auf 2000 v. Chr. dendrodatierten Holzfigur aus Mitteleuropa kann man einerseits nach Bronzezeit suchen und liegt damit in Nordeuropa um mindestens 200 Jahre später, oder man sucht nach 2000 v. Chr. und kommt in Nordeuropa in den Bereich der Steinzeit. Beide Fragen sind berechtigt!

Übrigens, mit den bei Thesauri üblichen Strukturen kommen wir hier nicht viel weiter. Bilingualität oder Synonyme (Bronzezeit = bronsålder) helfen eben nur dann weiter, wenn die Bezeichner wirklich für denselben Begriff, also für exakt dieselbe Bedeutung stehen, und das ist in unserem Beispiel eben nicht der Fall. Das „Simple Knowledge Organization System" (SKOS) beispielsweise erlaubt in seiner Basisvariante zwar Mehrsprachigkeit, Über- und Unterbegriffe und Assoziationen, aber für die Beschreibung fachbezogener, eigener Beziehungen sind Erweiterungen erforderlich wie zum Beispiel die „eXtension for Labels" (SKOS-XL). Aufgrund seiner wachsenden Verbreitung bei der (standardisierten) Modellierung kontrollierter Vokabulare ist es vermutlich lohnenswert, diesen Ansatz auch hinsichtlich semantischer Interoperabilität weiter zu verfolgen, wobei vor allem den für das Zusammenarbeiten wichtigen Schlussfolgerungsmechanismen entsprechende Aufmerksamkeit zukommen sollte.

Resümee

ADeX wurde mit einfachen formellen Hilfsmitteln (Tabellen) im Wesentlichen kognitiv erstellt. Theoretisch ist es jedoch denkbar, alle Originalmodelle innerhalb einer Ontologie zu modellieren (zum Beispiel mit CIDOC-CRM) und durch einen Vergleichsmechanismus (Schlussfolgerungsmechanismen, Inferenz) ein passendes Pivot-Modell errechnen zu lassen. Die größten Herausforderungen liegen dabei vermutlich in einer möglichst einheitlichen und eindeutigen Modellierung sowie in einem hinreichend toleranten und aussagekräftigen Vergleichsalgorithmus. Eine praktische Realisierung dieses Verfahrens ist mir aber noch nicht bekannt.

Bei den Zeitansprachen wird die Kluft zwischen dem traditionellen kognitiven Ansatz zu formellen Verfahren sehr deutlich. In wissenschaftlichen Texten finden sich immer ausreichend Hinweise zum Kontext benutzter Zeitansprachen, die eine hinreichende Interpretation auch überregional erlauben. Mir ist jedoch kein überregionales Datenmodell bekannt, welches die komplexen Beziehungen von Zeitansprachen umsetzt und auch entsprechende Abfragevarianten anbietet. Hier besteht noch deutlicher Nachholbedarf.

Die Erstellung eines semantisch vermittelnden Meta-Modells und die Implementierung entsprechender Transformationen für eine vereinheitlichte Abfrage sind hochkomplexe Aufgabenstellungen. Herkömmliche Methoden aus dem Datenbank-Bereich versagen dabei schnell. Alternativen stammen aus den Bereichen Knowledge Management/Knowledge Engineering, Ontologien ermöglichen eine formelle Darstellung von Wissen durch Begriffe und Beziehungen. Für Experten stehen anspruchsvolle Werkzeuge zur Verfügung, die sich dem Laien allerdings nur schwer erschließen.

Der vorliegende Beitrag gibt einen Überblick darüber, wie heterogene Datenbestände unter Berücksichtigung semantischer Aspekte harmonisiert werden können und wie dies mit formellen, also computertechnisch umsetzbaren Mitteln unterstützt werden kann. Bei großen Datenmengen und auch bei komplexen Zusammenhängen sind diese formellen Methoden dem rein kognitiven Herangehen überlegen. Für die Umsetzung ist eine enge interdisziplinäre Zusammenarbeit mit Experten der Bereiche Knowledge Management/Knowledge Engineering empfehlenswert.

Heterogene Datenbestände können zusammenwirken, wenn ihre Strukturen gemeinsam mit den Bedeutungen beschrieben und semantisch modelliert werden. Die entsprechenden Methoden dazu existieren. Semantische Interoperabilität ist machbar, probieren wir es aus!

Quellen

ADeX Version 2.0 – Archäologischer Daten-Export, Standard für den Austausch archäologischer Fachdaten. Internet (http://www.landesarchaeologen.de/fileadmin/Dokumente/Dokumente_Kommissionen/Dokumente_Archaeologie-Informationssysteme/Dokumente_AIS_ ADeX/ADeX_2-0_Doku. pdf).

Fraunhofer ISST, Semantische Interoperabilität. Band 1: Eine technische Einführung (http://www.isst.fraunhofer.de/Images/White-Paper_Einführung_Bd1_tcm81-49792.pdf).

W. Gödert, Semantische Wissensrepräsentation und Interoperabilität. In: Information Wissenschaft & Praxis 61 (2010). Dinges & Frick GmbH, Wiesbaden (http://linux2.fbi.fh-koeln.de/reseda/depot/Goedert_SeTIW.pdf).

M. Hall, Automatisierte semantische Harmonisierung von Landnutzungsdaten. In: Strobl, Blaschke, Griesebner (Hrsg.): Angewandte Geoinformatik 2006. Beiträge zum 18. AGIT- Symposium Salzburg. Herbert Wichman Verlag, Heidelberg. (http://work.room3b.eu/wp-content/papercite-data/ pdf/hall2006.pdf).

M. Lang, G. Carver, S. Printz, Standardised Vocabulary in Archaeological Databases. Vortrag zur Tagung: Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton 2012 (Veröffentlichung vorgesehen).

J. Hubrich, Vom Stringmatching zur Begriffsexploration: Das Potential integrierter begrifflicher Interoperabilität. In: Tagungsband der 12. Tagung der Deutschen Sektion der ISKO 2009 in Bonn. Ergon-Verlag, Würzburg. (http://linux2.fbi.fh-koeln.de/crisscross/publikationen/Hubrich_IntegrierteBegrifflicheInteroperabilitaet.pdf).

P. Mayr, B. Zapilko, Y. Sure, Ein Mehr-Thesauri-Szenario auf Basis von SKOS und Crosskonkordanzen. (http://www.ib.hu-berlin.de/~mayr/arbeiten/Oberhof2010.pdf).

Projekt: Grenzüberschreitende Homogenisierung von Geobasisdaten zwischen dem Freistaat Sachsen und der Tschechischen Republik. (http://www.geodat.ioer.info/). Projektbericht (Langfassung) Semantische Harmonisierung: Vergleich und Zuordnung der Datenmodelle. (http://www.geodat.ioer.info/fileadmin/download/Berichte/ AP32-33-34.pdf).

Semiotisches Dreieck, Wikipedia (http://de.wikipedia.org/wiki/Semiotisches_Dreieck).

SKOS – Simple Knowledge Organization System. (http://www.w3.org/2004/02/skos/).

SKOS-XL – SKOS eXtension for Labels (http://www.w3.org/TR/skos-reference/skos-xl.html).

P. Straub, Semantische Interoperabilität – Der Schlüssel zur integrierten Nutzung heterogener Geodaten. In: gis.science – Die Zeitschrift für Geoinformatik 3/2010, S. 93 ff.

K. Woitas, Bibliografische Daten, Normdaten und Metadaten im Semantic Web – Konzepte der Bibliografischen Kontrolle im Wandel. Magisterarbeit, Humboldt-Universität Berlin, 2010. (http://www.scribd.com/doc/50545116).

(Alle Internet-Links mit Stand vom 27.05.2012).

Anmerkungen

[1] In diesem Beitrag wird „kognitiv" nur im Sinne menschlichen Denkens, menschlicher Erkenntnisgewinnung verstanden. Dass auch Computern gelegentlich intelligente oder kognitive Leistungen zugeschrieben werden, spielt hier keine Rolle und wird daher ignoriert.

[2] ADeX ist ein bewusst einfach gehaltenes Datenformat. Es besteht aus nur wenigen essentiellen Attributen und dient dem Austausch von Kerninformationen über archäologisch qualifizierte, abgegrenzte Regionen im Gelände. Der modulare Aufbau von ADeX bietet Flexibilität, es sind Module zu den Bereichen Generelle Angaben, Georeferenz und Typ/Zeit definiert. Dabei werden grobe Typ- und Zeitangaben mit einheitlichen Begriffen belegt, während feinere Ansprachen vorerst aus den unterschiedlichen Terminologielisten der beteiligten Institutionen übernommen werden. Weitere Informationen findet man im Internet unter www.landesarchaeologen.de.

[3] Aber auch das Äpfel-Birnen-Problem ist lösbar, wenn man sich zum Beispiel auf die Vergleichsgrößen Volumen und Gewicht beschränkt.

Zur Arbeit mit Zeitstufen und Kulturgruppen – „reine Lehre" und praktische Umsetzung

Otto Mathias Wilbertz

Beim Aufbau von Datenbanken und Informationssystemen für die Archäologie spielte von Anfang an die Benutzung von feststehenden Begriffen und Formulierungen eine große Rolle. Dies erleichtert entscheidend die Suche nach vergleichbaren Inhalten, zum Beispiel nach einer bestimmten Quellengattung, nach einem Denkmaltyp oder nach einem Epochenbegriff.

Wir sprechen von einfachen Wertelisten, wenn Begriffe vorgegeben sind, die benutzt werden dürfen (andere Begriffe sind nicht zulässig), aber nicht durch definierte (zum Beispiel hierarchische) Relationen zueinander charakterisiert sind. Bei einem Thesaurus handelt es sich um eine Sammlung von Begriffen, die miteinander in Beziehung gesetzt sind. Bei einer Ordnung in Oberbegriffe, die Unterbegriffe umfassen, handelt es sich um einen hierarchischen Thesaurus. Gehört ein Unterbegriff zu mehreren Oberbegriffen, so sprechen wir von einem polyhierarchischen Thesaurus. Einen polyhierarchischen Thesaurus erreicht man zum Beispiel mit objektorientierter Programmierung. Zu den klassischen Anwendungsfällen von Thesaurierung gehört die Klassifizierung von Datierungsangaben. Dies lässt sich am Feld „Grobdatierung" des Datenaustausch-Standards ADeX (Version 1 und 2) zeigen. Die dort definierte Werteliste kann als Ausgangspunkt eines Thesaurus betrachtet werden, zur Verdeutlichung werden die in der Werteliste enthaltenen Begriffe entsprechend ihrer Bedeutung in unterschiedlichen Ebenen angeordnet (Abb. 1).

Unbekannt
Erdgeschichte
Vorgeschichte
- **Steinzeit**
  - **Paläolithikum**
  - **Mesolithikum**
  - **Neolithikum**
- **Metallzeit**
  - **Bronzezeit**
  - **Vorrömische Eisenzeit**

Frühgeschichte
- **Römische Kaiserzeit**
- **Völkerwanderungszeit**

Mittelalter/Neuzeit
- **Mittelalter**
- **Neuzeit**

1 Begriffe der Werteliste für das ADeX-Feld „Grobdatierung", angeordnet nach Ebenen.

Für die folgenden Zusammenhänge werden die Begriffe „Epoche", „Periode", „Phase", „Zeitstufe" zunächst als gleichbedeutend betrachtet. Ebenso werden „Kultur", „Kulturgruppe" und „...-Gruppe" als Synonyme verstanden.

Fragt man nach dem Unterschied zwischen Zeitstufe und Kulturgruppe, so scheint die Antwort ganz einfach: eine Zeitstufe bezeichnet Phänomene, die einem bestimmten Zeitabschnitt zugeordnet werden können. Der Ausdruck Zeit*abschnitt* weist bereits darauf hin, dass es sich um einen Teil einer größeren Einheit handelt, dass es also ein Davor und/oder ein Danach geben muss. Eine Kultur(gruppe) bezeichnet ein Ensemble von Phänomenen, das einem bestimmten regionalen Rahmen zugeordnet werden kann.

Betrachtet man das Verhältnis der beiden Gliederungskonzepte zueinander, so muss man feststellen, dass in einer Zeitstufe mehrere Kulturgruppen nebeneinander vorkommen können. Umgekehrt kann eine Kulturgruppe mehrere Zeitstufen umspannen. Dies würde dafür sprechen, Zeitstufen und Kulturgruppen als zwei eigenständige Attribute zu behandeln und zwei Thesauri vorzuhalten oder zu entwickeln.

Doch so einfach, wie oben angedeutet, ist die Sachlage nicht: Eine chronologische Gliederung und ihre einzelnen Zeitstufen haben einen bestimmten Gültigkeitsbereich, außerhalb dessen eine andere Stufengliederung anzuwenden ist. Das heißt, die Gliederung bezieht sich auf einen regionalen Rahmen, auch wenn dieser noch so weitreichend ist. Und eine Kulturgruppe hat nicht nur einen Bezug zum Raum, sondern auch einen Anfang und ein Ende, das heißt: einen Bezug zur Zeit.

Die Entwicklung der verschiedenen Fachdatenbanken und Fachinformationssysteme in der Archäologie vollzog sich bisher meist von kleinen, einfachen Anwendungen hin zu größeren, komplexeren („bottom-up"); am Anfang stand nicht das theoretische Konzept, sondern die praktische Anwendung. Dem-

entsprechend sind Thesauri auch nicht unbedingt nach den theoretischen Konzepten der Dokumentationswissenschaft, sondern nach dem Bedarf in der Alltagspraxis der Archäologie – hier der archäologischen Denkmalpflege – entstanden.

Beim Versuch der Reformierung bestehender Systeme stößt die Berücksichtigung theoretischer Konzepte auf die eingeführte, mehr oder weniger bewährte Praxis. Diese ist durch unterschiedliche Ausgangssituationen für verschiedene Epochen gekennzeichnet. Für das Neolithikum Mitteleuropas wurden die großen, in weiten Teilen Europas vorkommenden Kulturen (wie zum Beispiel Bandkeramik, Rössen und Michelsberg) auch zur zeitlichen Ordnung des Fundguts und der Befunde herangezogen, das Material innerhalb dieser Komplexe wurde zeitlich geordnet, eine übergeordnete Terminologie spielte kaum eine Rolle.

Wo eine so aufgebaute Chronologie eingeführt und bewährt ist, wo also Bandkeramik immer auch für eine bestimmte Zeit steht, fällt es naturgemäß schwer, zwei Thesauri zu führen, da es keinen Vorteil bedeutet, wie beispielsweise in unserem Fall, in einem Thesaurus den Begriff „Bandkeramik“ als Kultur und in einem zweiten Thesaurus den gleichen Ausdruck oder einen davon abgeleiteten Begriff als Bezeichnung für eine Zeitstufe zu führen. Zwei getrennte Attribute zu führen, bedeutet in diesem Fall eine redundante Datenhaltung ohne Erkenntnis- oder Aussagegewinn.

Andernorts (zum Beispiel im nördlichen Europa) wurden schon früh Epochenbegriffe neben den Bezeichnungen für Kulturen geführt. So entspricht die Trichterbecherkultur den Zeitabschnitten Frühneolithikum und Mittelneolithikum. Andererseits werden dem Mittelneolithikum noch weitere Kulturgruppen zugeordnet.

Für die Metallzeiten wurden schon früh Begriffe für die Zeitstufen neben die Bezeichnungen für Kulturen gestellt (wie zum Beispiel Urnenfelderzeit – Urnenfelderkultur, Hallstattzeit – Hallstattkultur). Die Unabhängigkeit der Gliederungskonzepte mag dadurch belegt werden, dass die Hallstattkultur nicht die Stufen Hallstatt A und Hallstatt B einbezieht, sondern in der Stufe Hallstatt C beginnt.

Aus den geschilderten Beispielen wird klar, dass eine getrennte Führung von Zeit- und Kulturbegriffen in verschiedenen Attributen unter Nutzung getrennter Thesauri in vielen Fällen durchaus sinnvoll ist, so zum Beispiel, um innerhalb des Mittelneolithikums Befunde der Trichterbecherkultur von Fundstellen anderer Kulturgruppen zu unterscheiden oder für eine gemeinsame Kartierung zeitgleicher bronzezeitlicher Gruppen. Ein weiteres Beispiel wäre der Übergang vom Neolithikum zur Bronzezeit mit Schnurkeramik, Glockenbecher- und Aunjetitzer Kultur.

Andererseits ist auch klar, dass eine einfache Dopplung von Einträgen (Beispiel Bandkeramik) nicht weiterhilft. Die Angabe einer Kulturgruppe ist dort wertvoll, wo sie ein Mehr an Information bringt.

Aus diesen beiden Gegebenheiten lassen sich für die Praxis folgende Verfahrensweisen empfehlen: Einerseits sollte die Möglichkeit, Kulturgruppen als selbstständiges Attribut zu führen, gegeben sein; andererseits sollte die Übertragung oder „Übersetzung“ von Chronologiebegriffen (1:1) unterbleiben, da sie keinen Erkenntnisgewinn bringt. Vielmehr sollten die Eintragungen von Kulturgruppen immer eine (meist regionale) Präzisierung der Datierungsangaben darstellen. Das bedeutet, dass der Eintrag einer Kulturgruppe nicht obligatorisch gemacht werden kann.

Für den Datenaustausch bedeutet dies, dass es erstens Systeme geben kann, in welchen Kulturgruppen als eigenständiges Attribut geführt werden, und andere, in denen das nicht der Fall ist. Zweitens bedeutet dies, dass – auch bei Datengebern mit dem Attribut „Kulturgruppe“ – dieses Attribut in einigen oder vielen Fällen keine Werte enthält.

Vor einem Einsatz von Kulturgruppen-Namen beim Datenaustausch empfiehlt sich eine Sammlung der in den Ämtern verwendeten Begriffe mit dem Ziel, zu einer Werteliste zu gelangen. Die Ausdrücke sollten definiert werden hinsichtlich ihrer regionalen und zeitlichen Zuordnung sowie hinsichtlich ihrer Bedeutung (Grabsittenkreis, Werkstattkreis, Leitformen und Formenspektrum). Dabei sollten ähnlich klingende Begriffe gleichen Inhalts angeglichen werden.

Thesauri im Fachinformationssystem des Bayerischen Landesamtes für Denkmalpflege

Markus Ullrich, Roland Wanninger

Einleitung

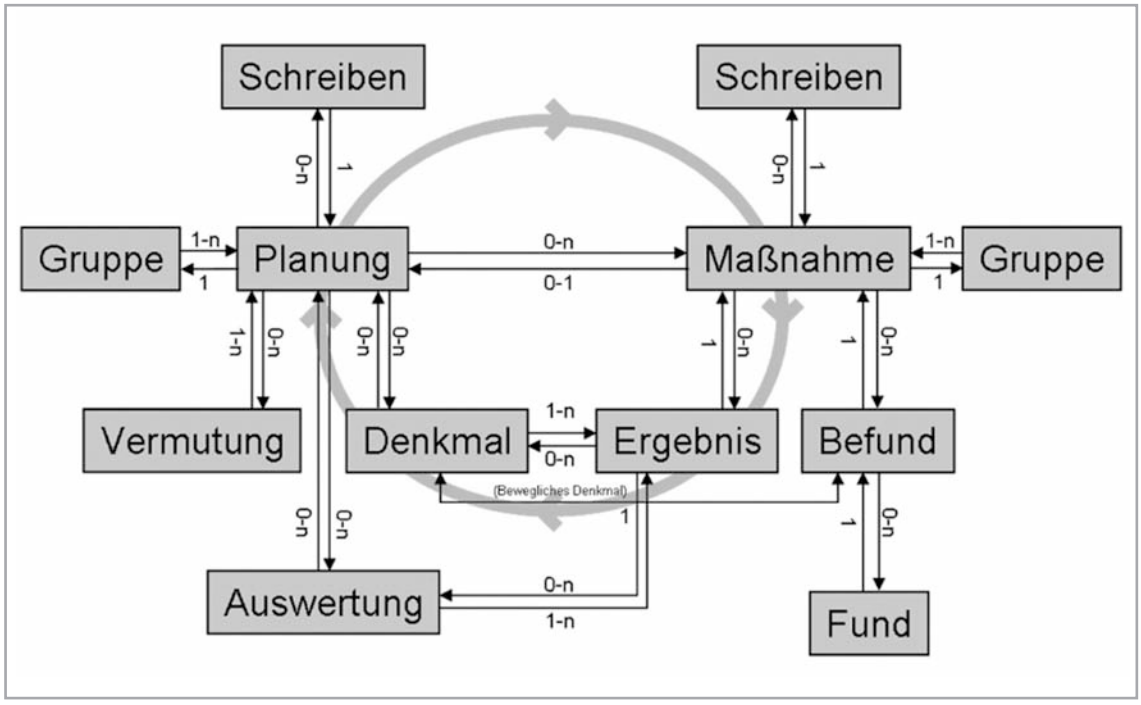
Der Begriff „Thesaurus", im Sinne einer Sammlung zulässiger (Fach-)Termini, ist bereits seit dem 18. Jahrhundert nachweisbar[1]. Im Zeitalter der elektronischen Datenverarbeitung werden solche kontrollierten Vokabulare vor allem bei der computergestützten Suche nach komplexen Inhalten, dem sogenannten „Information Retrieval", verwendet[2]. Die Vorteile liegen dabei auf der Hand: Unterschiedliche Schreibweisen für ein und denselben Terminus werden vermieden sowie Rechtschreibfehler und die Eingabe unzulässiger Begriffe ausgeschlossen. In einer Datenbank laufen Thesaurusabfragen indiziert ab, das heißt über Zahlencodes, die dem jeweiligen Eintrag hinterlegt sind. Somit sind sie im Vergleich zur Freitextsuche wesentlich schneller und benötigen weniger Ressourcen. Die Vokabulare können hierarchisch gegliedert sein, sodass ein Fachbegriff einem Überbegriff zugeordnet sein und gegebenenfalls auch mehrere Unterbegriffe enthalten kann. Bei Bedarf besteht außerdem die Möglichkeit, Synonyme in der Wortliste zu verwalten. Sind ein oder mehrere Begriffe zugleich mehreren Oberbegriffen zugeordnet, spricht man von einem polyhierarchischen Thesaurus. Durch die DIN 1463-1, beziehungsweise deren internationalem Äquivalent ISO 2788, ist die elektronische Verwendung klar definiert. Für die Erfassung der Bodendenkmäler und die Führung der Bayerischen Denkmalliste ist es im Unterschied zu den sicht- und meist begehbaren Baudenkmälern notwendig, alle verfügbaren archäologisch relevanten Informationen zu sammeln und wissenschaftlich zu bewerten[3]. Dementsprechend genießt auch die Rückgewinnung, respektive die Wiedererschließung dieser Fachinformationen, eine hohe Priorität. Schon in den 1980er Jahren wurden deshalb in der Inventarisation des Bayerischen Landesamtes inselartig dBase-Lösungen an den einzelnen Dienststellen eingesetzt. Unter Zuhilfenahme von vordefinierten Listen mit Buchstabenkürzeln, dem sogenannten „Zeit-Typ-Schlüssel", stellte man damals die Informationsrückgewinnung trotz unterschiedlicher Schreibweisen in den Beschreibungen der Fundstellen sicher. Die Kombination „HAGBgh18" bedeutete zum Beispiel „Grabhügelgruppe der Hallstattzeit mit 18 (erkannten) Hügeln". Eine Abfrage von Oberbegriffen, wie etwa dem „Neolithikum", die im Ergebnis automatisch sämtliche untergeordneten Zeitstufen und Kulturen/Gruppen beinhaltete, wie das bei echten hierarchischen Thesauri der Fall ist, war aber nur sehr bedingt möglich.

Mit dem Anschluss des Bayerischen Landesamtes an das Bayerische Behördennetz 2003 wurden die vormals unvernetzt betriebenen Datenbanken auf einem SQL-Server zusammengeführt und die vorhandenen Verschlüsselungen in Wortlisten ohne Hierarchisierung migriert. Unterschiedlichste Informationen fasste man aber weiterhin als Fließtext in einem Datensatz zusammen, unabhängig davon, ob es sich um Fundmeldungen, um die Beschreibung archäologischer Befunde und Funde oder um Angaben zu Ausgrabungen sowie Prospektionsmaßnahmen handelte. Eine systematische Aufschlüsselung der verschiedenen Fachdaten erfolgt erst seit der Einführung des Fachinformationssystems (FIS) der Denkmalpflege im Jahr 2008, bei der die vormalige Datenbank zu einem geografischen Informationssystem (GIS) weiterentwickeltt wurde[4].

Datenmodell des FIS, Teil Bodendenkmäler

Das wesentliche Kennzeichen des Fachinformationssystems gegenüber den vorher in Bayern verwendeten Datenbanken ist die aktivitäts- bzw. maßnahmebezogene Struktur. Sie löste die auf den Fundplatz bezogene Erfassung ab[5]. Ausdehnung, Ansprache und Datierung von Bodendenkmälern sind einem steten Wandel unterworfen und müssen nach aktueller Sachlage überarbeitet werden. Neue Erkenntnisse können sowohl zur Ausweisung eines Denkmals als auch zu einer teilweisen oder gänzlichen Streichung aus der Denkmalliste führen. Im Unterschied zu dieser „Dynamik" der Bodendenkmäler sind Typ, Inhalt und räumliche Ausdehnung der archäologischen Aktivitäten, ebenso wie die daraus resultierenden Erkenntnisse, vergleichsweise konstant. Die Fläche einer Begehung mit dort geborgenem archäologischem Fundmaterial ändert sich auch nach einer Überbauung und der damit einhergehenden Zerstörung nicht. Die „archäologische Information" hat nach wie vor die gleiche Lokalisierung, auch wenn das Bodendenkmal längst nicht mehr existiert. Im Datenmodell des FIS werden diese einzelnen Fachobjekte nun in ihre logischen Beziehungen zueinander gesetzt (Abb. 1)[6].

Bei fast allen Objekten kommen Thesauri zum Einsatz. Dabei handelt es sich sowohl um kurze Wertelisten als auch zum Teil um umfangreiche und hierarchisch gegliederte Vokabulare. Das Datenbankschema aller Thesauri ist aber unabhängig vom jeweiligen Umfang gleich aufgebaut.



1 Vereinfachtes Fachmodell des FIS der Denkmalpflege in Bayern, Teil Bodendenkmäler, mit Darstellung der wichtigsten Entitäten und ihrer logischen Beziehungen im relationalen Datenbankmodell (Grafik BLfD, H. Dietrich).

Fachobjekte im FIS, Teil Bodendenkmäler und verwendete Thesauri

„Maßnahme"
Jegliche archäologisch relevante Erkenntnis stammt aus einer konkreten Aktivität oder Maßnahme[7]. Diese wird zunächst mit einem Typ-Thesaurus kategorisiert (Abb. 2), der sowohl Maßnahmearten aus der praktischen Bodendenkmalpflege als auch aus dem Bereich der Denkmalerfassung und -forschung enthält. Zusätzlich wird verzeichnet, wer eine Maßnahme veranlasst oder durchgeführt hat. Ebenso wird erfasst, wann diese begonnen, beendet oder mitgeteilt wurde und wer diese fachlich betreut hat.

„Ergebnis"
Die Erkenntnisse, die aus einer Maßnahme resultieren, werden in einem oder mehreren Ergebnissen festgehalten[8]. Diese werden als positiv kategorisiert, wenn archäologische Befunde oder Funde daraus stammen. Bei einer Fläche ohne archäologisch relevante Substanz ist das Ergebnis negativ. Besteht Unklarheit, etwa weil eine Untersuchung bauseitig bedingt nicht bis in fundführende Schichten vorgedrungen ist oder diese kolluvial überdeckt sind, ist es unbekannt. Kleinere Wortlisten geben über den Erhaltungszustand der jeweiligen Fläche (erhalten, ausgegraben, unbeobachtet zerstört, Teilerhaltung vertikal) und deren Überdeckung (natürlich überdeckt, überbaut, konservatorisch überdeckt, obertägig sichtbar) Auskunft[9]. Beim Ergebnis findet auch die „wissenschaftliche" Thesaurierung der gewonnen Erkenntnisse mit einer Kombination aus Datierung und Typangabe statt.

„Denkmal"
Die aus den Maßnahmen gewonnenen Erkenntnisse sind die Grundlage für die Erfassung der Bodendenkmäler[10]. Einem Bodendenkmal muss immer mindestens ein Ergebnis zugrunde liegen. Im Fachobjekt „Denkmal" werden auch die Daten für die Veröffentlichung als Geowebdienst und im BayernViewer denkmal geführt[11]. Im Gegensatz zu den Ergebnissen wurde auf eine wissenschaftliche Thesaurierung der Bodendenkmäler verzichtet. Dafür liefern kleinere Wortlisten Angaben zur Gefährdung (nicht gefährdet, Erosion, Landwirtschaft, Bebauung) und zum Verfahrensstand der Eintragung in die Denkmalliste nach Art. 2 Abs. 1 BayDSchG (nicht mitgeteilt, mitgeteilt, Verfahren abgeschlossen, Dissens).

„Planung", „Schreiben" und „Vermutung"
Neben den Maßnahmen, den daraus entstehenden Ergebnissen und den Bodendenkmälern als klassischem Arbeitsbereich der Inventarisation, werden im FIS auch Verwaltungsvorgänge berücksichtigt. Die Beteiligung des Bayerischen Landesamtes als Träger öffentlicher Belange sowie davon unabhängige Genehmigungsverfahren nach Art. 7 Abs. 1 BayDSchG werden als Planung erfasst. Unter Zuhilfenahme eines Planungstyp-Thesaurus werden diese nach Erlaubnisverfahren, nach kommunalen Planungen (wie etwa Flächennutzungsplan, Stadtentwicklungsplan oder

Bebauungsplan), regionale/überregionale Planungen/ Vorhaben, Umweltprüfungsverfahren, Schutzgebiete, Benehmensherstellung/Mitteilung und sonstige Planungen kategorisiert[12]. Über die Planungsgeometrie können gegebenenfalls betroffene Bodendenkmäler ermittelt und in einer Stellungnahme mitgeteilt werden („Schreiben").

Da nach dem Bayerischen Denkmalschutzgesetz eine Erlaubnis für Bodeneingriffe erforderlich ist, auch wenn keine Bodendenkmäler bekannt, sondern nur zu vermuten oder „den Umständen nach an[zu]nehmen" sind (Art. 7 Abs 1. BayDSchG), hat der zuständige Referent der praktischen Bodendenkmalpflege die Möglichkeit, im Rahmen einer Planung Vermutungsflächen anzulegen[13].

Auch bei der Betreuung von Maßnahmen kann ein Schriftverkehr entstehen, der als „Schreiben" im FIS erfasst wird. Außer der Zuordnung zu „Planung" und „Maßnahme" sind die Begriffe im Schreibentyp-Thesaurus jeweils nach Eingang und Ausgang gegliedert[14].

„Auswertung"
Ferner können mithilfe der archäologischen Erkenntnisse, respektive der Ergebnisse im Fachinformationssystem, übergreifende Fragestellungen beantwortet werden. Dafür wurde eigens ein weiteres Fachobjekt definiert. Es dient der Erarbeitung archäologischer Stadt- oder Gemeindekataster, dem Entwerfen von Texten für Denkmaltopografien und anderen Publikationen, der Darstellung der Flächen des archäologischen UNESCO-Welterbes Obergermanisch-Raetischer Limes ebenso wie der Beantwortung wissenschaftlicher Fragen. Der Thesaurus der Auswertungstypen ist nach diesen genannten Zwecken gegliedert.

„Befund" und „Fund"
Aus archäologischen Maßnahmen können Befunde und Funde resultieren. Angesichts der großen Datenmenge ist eine systematische und vor allem vollständige Thesaurierung nicht umsetzbar, weshalb bei den Befunden ganz darauf verzichtet wurde. Die Verschlagwortung eines Fundes ist mithilfe einer Dreierkombination aus Angaben zu Gegenstand, Material und Zeitstellung zwar möglich, doch erfolgt dies nur für ausgewählte Objekte, zum Beispiel solche, die fotografisch und zeichnerisch dokumentiert werden. Dennoch sollen aus Gründen der Datensicherung zukünftig alle aus Grabungen stammenden digitalen Befund- und Fundlisten in genormter Form ins Fachinformationssystem importiert werden.

Lageangaben
Thesauri sind auch bei den Lageangaben im Einsatz. Die Adressen und Flurnummern aller Fachobjekte, die im Fachinformationssystem eine Kartierung besitzen („Maßnahme", „Ergebnis", „Denkmal", „Vermutung", „Auswertung" und „Planung"), werden automatisch über die Geometrie aus dem Datenbestand des Landesamtes für Vermessung und Geoinformation abgeleitet. Dabei wird auf die amtliche bayerische Gemeinde- beziehungsweise Gemarkungskennziffer referenziert, die hier jeweils als Thesaurus-ID fungiert. Die Liste der Gemeinden ist hierarchisch nach den administrativen Einheiten Regierungsbezirk und Landkreis gegliedert, die Gemarkungsliste nach den Gemeinden. Diese beiden Thesauri liegen aber nicht in der Zuständigkeit des Bayerischen Landesamtes für Denkmalpflege, sondern bei der Bayerischen Vermessungsverwaltung.

Thesauri im FIS, zwei Beispiele

Im Folgenden werden zwei im Fachinformationssystem eingesetzte Thesauri exemplarisch vorgestellt: der Thesaurus der Maßnahmetypen und der Zeit-Thesaurus für das Fachobjekt „Ergebnis".

Zur Verschlagwortung des Fachobjekts „Maßnahme" dient ein Thesaurus, der eine Recherche nach Aktivitätstyp ermöglicht (Abb. 2). Neben Maßnahmearten der praktischen Bodendenkmalpflege (Voruntersuchung, Oberbodenabtrag, Ausgrabung, konservatorische Überdeckung usw.) finden sich hier auch Aktivitätsarten der Denkmalerfassung und -forschung (systematische Begehung, Luftbildauswertung, Geophysik, Topografie/Vermessung usw.). Des Weiteren werden auch die Hinweise Externer, die nach Art. 8 Abs. 1 BayDSchG meldepflichtig sind, als Begehung/ Beobachtung erfasst. Gesondert lässt sich die in Bayern wegen des fehlenden Schatzregals problematische Suche mit Metallsonde (Sondenfund) ausweisen.

2 Typ-Thesaurus des Fachobjekts „Maßnahme" mit Markierung des Suchbegriffs „Begehung" (Screenshot Fachinformationssystem BLfD, Stand August 2012).

Der Maßnahmetypen-Thesaurus wurde eigens für das FIS neu definiert. Bereits vor 2008 digital erfasste Aktivitäten, denen diese detaillierten Informationen fehlen, mussten daher bei der Migration als solche gekennzeichnet werden (Übernahme aus IV). Wenn Dokumentationen oder die Archive des Bayerischen Landesamtes es ermöglichen, werden der Typ und die zugehörigen Informationen rückwirkend nachgetragen. Nicht eindeutig lokalisierbare Maßnahmen oder solche mit unvollständigen Angaben werden nach der Prüfung unter dem Typ Aktenstudium subsumiert. Darunter verbirgt sich ein heterogener Datenbestand von Aktivitäten aus den unterschiedlichsten Quellen.

Von den Meldungen externer Mitarbeiterinnen und Mitarbeiter des Bayerischen Landesamtes werden lediglich die Maßnahmen im Fachinformationssystem verarbeitet, die ein positives Ergebnis erbracht haben. Somit gibt der Datenbestand nur einen vergleichsweise kleinen Teil des tatsächlichen Einsatzes ehrenamtlich tätiger Kräfte wieder. Hier werden jedoch regelhaft alle Meldungen aus Begehungen oder Beobachtungen seit 2005 rückwirkend maßnahmebezogen inventarisiert und damit eine der zentralen Aufgaben der bis 2004 analog gedruckten Bayerischen Fundchronik fortgeführt[15].

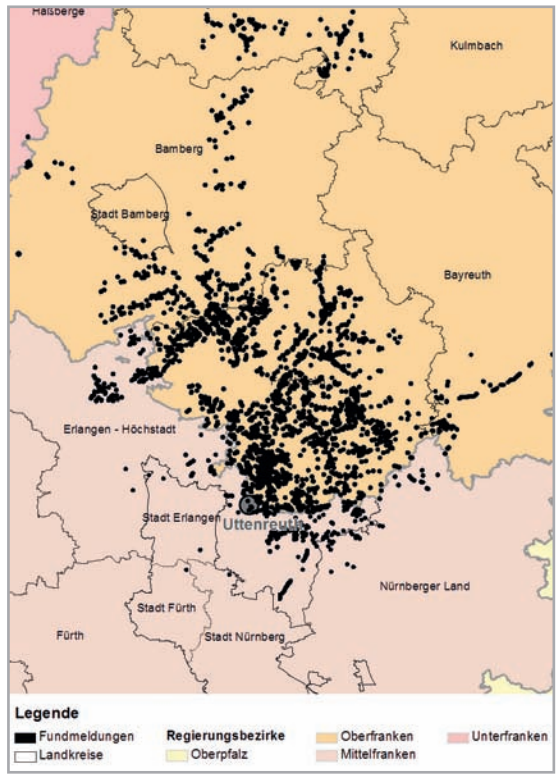
Nur in Ausnahmefällen wird es möglich sein, sämtliche Feld- oder Geländebegehungen eines einzelnen Mitarbeiters im Fachinformationssystem vollständig, das heißt inklusive der Aktivitäten vor 2005, zu erfassen. Dies ist beispielsweise dann sinnvoll, wenn die Denkmallandschaft ganzer Landkreise vornehmlich auf der Aktivität einzelner ehrenamtlicher Kräfte basiert und im Vorgriff auf Stellungnahmen des Amtes zeitnah eine nochmalige Prüfung des Fundmaterials notwendig wird[16]. Als Beispiel kann hier die Kartierung eines Ehrenamtlichen aus dem Grenzgebiet der Regierungsbezirke Ober- und Mittelfranken angeführt werden (Abb. 3)[17].

Im Unterschied zum Maßnahmetypen-Thesaurus musste bei der Entwicklung der Typ- und Zeit-Thesauri für das Fachobjekt „Ergebnis" der vorhandene Altdatenbestand berücksichtigt werden, damit dieser so präzise wie möglich in das Fachinformationssystem migriert werden konnte. Dies hatte zur Folge, dass nicht eindeutige oder nach heutigem Forschungsstand überholte Begriffe in der Regel nur einem Oberbegriff zugewiesen werden konnten. Eine nachträgliche Aufteilung von Daten „per Hand" sollte soweit wie möglich vermieden werden.

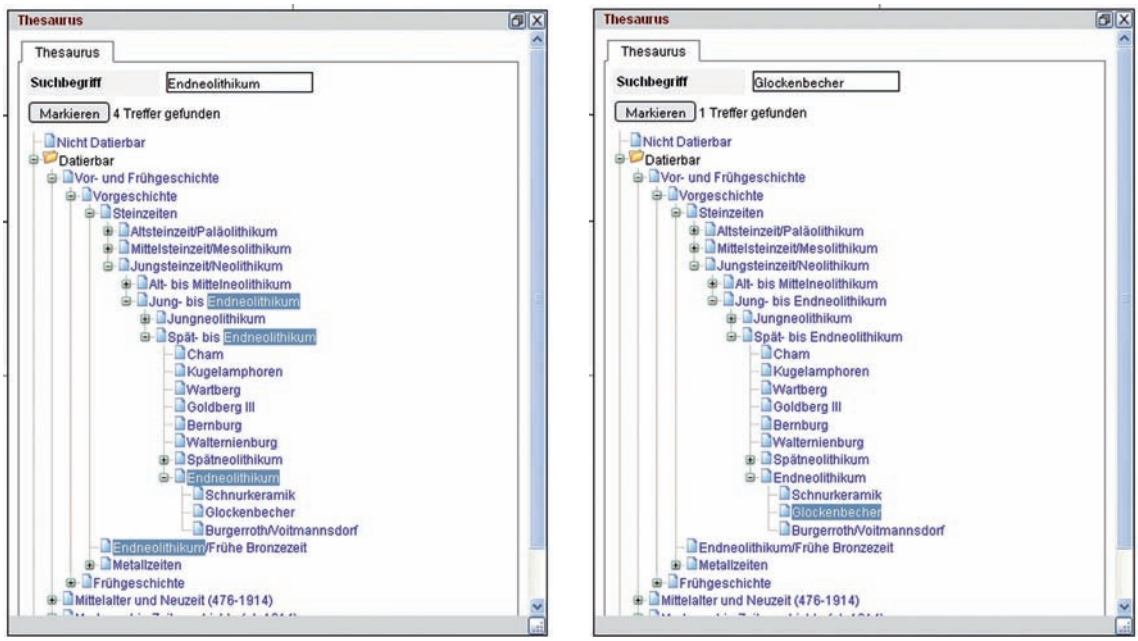
Für die Überarbeitung der Begriffe des „Zeit-Schlüssels" aus der Inventarisationsdatenbank waren mehrere Arbeitsschritte notwendig:

1. Zuordnung des vorhandenen Schlagwortbestandes in Zeitstufen beziehungsweise Epochen einerseits und archäologischen Kulturen beziehungsweise Gruppen andererseits
2. Exakte Definition und absolute Datierung von Zeitstufen und Kulturen beziehungsweise Gruppen
3. Hierarchische Gliederung des Zeitstufen- beziehungsweise Epochen-Thesaurus
4. Zuordnung der Kulturen beziehungsweise Gruppen mithilfe der absoluten Datierungsspanne zu den Zeitstufen

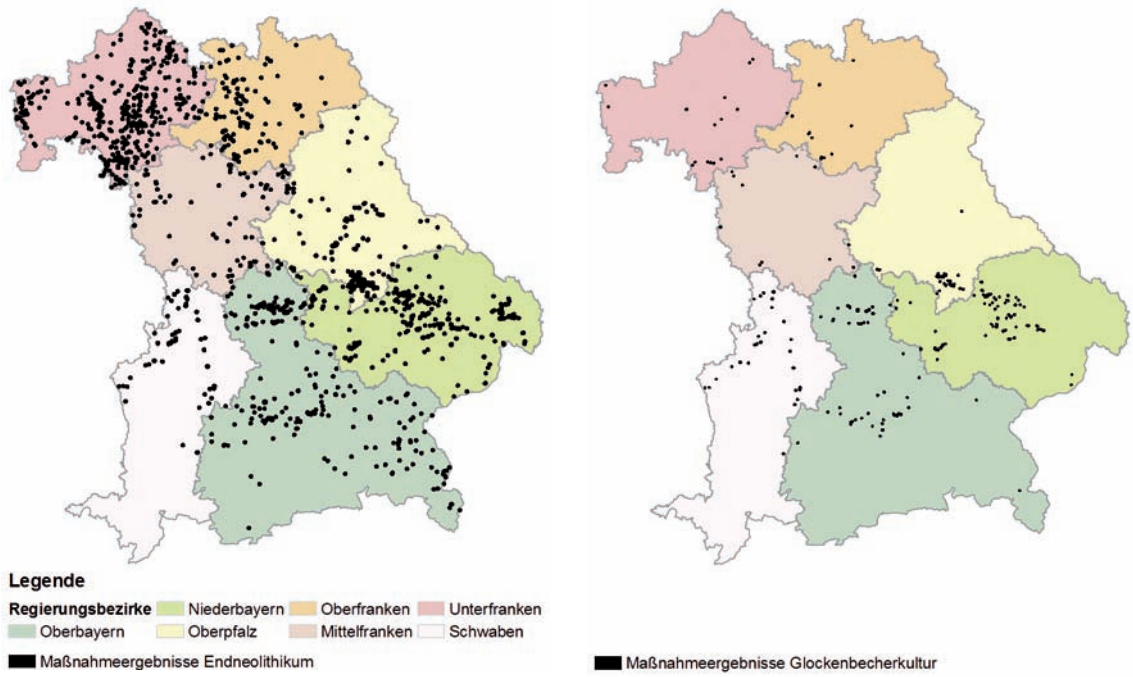
Dieses Vorgehen erübrigte den Aufbau einer eigenen Auswahlliste für Kulturen beziehungsweise Gruppen. Sie sind in den Zeit-Thesaurus mit integriert, sodass die wissenschaftliche Indexierung eines Ergebnisses grundsätzlich nur aus Begriffspaaren von Typ und Zeit besteht. Ein Ergebnis kann aber mehrere Zeit-Typ-Angaben aufweisen[18]. Maßnahmeergebnisse im Bereich einer Höhensiedlung werden zum Beispiel als „Höhensiedlung/Spät- bis Endneolithikum", „Höhensiedlung/Frühe Latènezeit", „Depot/Römische Kaiserzeit" usw. verschlagwortet.



3 Kartierung aller Maßnahmen des Typs Begehung/Beobachtung mit positivem Ergebnis auf Grundlage des Fachinformationssystems, die von einer Einzelperson (Armin Thomschke, Uttenreuth) dem Bayerischen Landesamt gemäß Art. 8 BayDSchG gemeldet wurden. Zur besseren Sichtbarkeit in der Karte wurden alle Fundmeldungen mit einem Radius von 250 m gepuffert (Kartografie R. Wanninger, Stand Juli 2012).



4 Zeit-Thesaurus des Fachobjekts „Ergebnis“ mit Markierung der Suchbegriffe Endneolithikum (links) und Glockenbecher (rechts) (Screenshot Fachinformationssystem BLfD, Stand August 2012).



5 Kartierung aller Maßnahmeergebnisse in Bayern unterhalb der Suchbegriffe Endneolithikum (links) und Glockenbecher (rechts) auf Grundlage des Fachinformationssystems. Zur besseren Sichtbarkeit in der Karte wurden die Ergebnisse mit einem Radius von 1000 m gepuffert (Kartografie R. Wanninger, Stand August 2012).

Am Beispiel der spät- bis endneolithischen Epochen und Kulturen lässt sich das Ineinandergreifen von Zeitstufen einerseits und Kulturgruppen andererseits besonders gut nachvollziehen (Abb. 4). Die Recherche im Zeit-Thesaurus veranschaulicht die verschiedenen Ebenen, denen archäologische Quellen jeweils so exakt wie möglich zugeordnet werden können. Das Jung- bis Endneolithikum wurde unterteilt in Jungneolithikum und Spät- bis Endneolithikum; letzteres wiederum in Spätneolithikum und Endneolithikum. Diesen Epochenbegriffen sind je nach Laufzeit Kulturen beziehungsweise Gruppen zugeordnet[19]. Eine Gruppe wie Cham, die absolutchronologisch sowohl in das Spätneolithikum (3500–2900 v. Chr. nach Thesaurusdefinition) als auch in das Endneolithikum (2900–2100 v. Chr. nach Thesaurusdefinition) datiert, wurde unter dem Oberbegriff Spät- bis Endneolithikum eingereiht[20]. Sollte sich im Zuge des Forschungsfortschritts die Datierungsspanne einer Kultur beziehungsweise Gruppe ändern, ist es leicht möglich, diese mit einer anderen Epochenebene zu verknüpfen, ohne dass die bis dahin zugewiesenen Datensätze überarbeitet werden müssen[21].

Da sowohl Epochen als auch Kulturen beziehungsweise Gruppen wählbar sind, enthalten die Abfrageergebnisse unterschiedliche Aussagen in der Hinsicht, dass die Treffermenge für Endneolithikum nicht mit der Treffersumme aus der Abfrage nach Schnurkeramik, Glockenbecher und dem lokalen Burgerroth/Voitmannsdorf identisch ist. Artefakte, die lediglich allgemein dem Endneolithikum zugewiesen werden können, nicht aber einer bestimmten Kultur oder Gruppe, werden beispielsweise bei einer Suche nach dem Schlagwort Glockenbecher nicht angezeigt. Die Kartendarstellung der Trefferliste einer Schlagwortsuche nach Endneolithikum und Glockenbecher veranschaulicht den unterschiedlichen Charakter der beiden Abfragen (Abb. 5).

Die Datenqualität, vor allem die Genauigkeit der Datierung und die Vollständigkeit der Rechercheergebnisse bei Abfragen im Typ- oder Zeit-Thesaurus, ist vom jeweiligen Stand der Datenüberprüfung abhängig, die im Rahmen des noch laufenden Projektes Nachqualifizierung und Revision der Bayerischen Denkmalliste parallel zur Entwicklung und Einführung des Fachinformationssystems vom Bayerischen Landesamt für Denkmalpflege initiiert wurde. Erst nach dessen Abschluss 2013/14 darf mit einer bayernweit einheitlichen Grundlage für weiterführende Untersuchungen gerechnet werden. Dabei ist auch dann noch zu beachten, dass die Thesaurierung der Maßnahmeergebnisse primär denkmalfachlichen Fragestellungen folgt und nur als Grundlage für wissenschaftliche Abschlussarbeiten an Hochschulen dienen kann, diese aber nicht ersetzt.

Anmerkungen

[1] Aitchison, J. / Stella, D.C. 2004: The Thesaurus: A Historical Viewpoint with a Look to the Future. In: Cataloging & Classification Quarterly 37, no. 3/4, 5–21.

[2] Anderson, J. D. / Pérez-Carballo, J. 2005: Information Retrieval Design. Principles and Options for Information Description, Organization, Display and Access in Information Retrieval Databases, Digital Libraries, Catalogs, and Indexes (St. Petersburg [USA]), 297 ff.

[3] Ullrich, M. 2008: BayernViewer-denkmal – Fachinformationssystem der Denkmalpflege in Bayern. In: Chytrácek, M. / Gruber, H. / Michálek, J. / Rind, M. / Schmotz, K. (Hrsg.), Archäologische Arbeitsgemeinschaft Ostbayern/West- und Südböhmen/Oberösterreich – Archeologická pracovní skupina východní Bavorsko/západní a jižní Cechy/Horní Rakousko. 17. Treffen. 13. bis 16. Juni 2007 in Freistadt. Fines Transire 17, 53–60.

[4] Wanninger, R. 2008a: FIS erfolgreich am Start! Denkmalpflege Informationen 141, 64–66. – Wanninger, R. 2008b: Start des BayernViewer-denkmal am 16. September 2007 in Fürth. In: Greipl, E. J. (Hrsg.), 100 Jahre Bayerisches Landesamt für Denkmalpflege 1908–2008, Band II: Perspektiven (Regensburg 2008), 14 f.

[5] Mückl, F. / Wanninger, R. 2009: Digitale Erfassung flächenhafter Denkmäler – Das Fachinformationssystem der staatlichen Bayerischen Denkmalpflege. In: Mitteilungen des DVW-Bayern e. V. – Gesellschaft für Geodäsie, Geoinformation und Landmanagement, Heft 2/2009, 245–253.

[6] Der wichtigste Ideengeber und somit eigentlicher „geistiger Vater" des Fach- und heute realisierten Datenmodells des FIS ist Dr. Hanns Dietrich, dem an dieser Stelle ausdrücklich dafür gedankt sei. Bei dieser Gelegenheit muss außerdem erwähnt werden, dass wesentliche Anregungen dazu auch aus der AG Datenmodellierung der Kommission „Archäologie und Informationssysteme" stammen. Deshalb soll hier allen Mitgliedern der AG, die ihre Arbeit mit großem Engagement betreiben, ebenfalls gedankt werden.

[7] „Maßnahme" entspricht in seiner fachlichen Definition in etwa der „Untersuchungsfläche" nach dem Austauschformat ADeX. Vgl. Göldner/Himmelmann/Poelmann et al., S. 10, Absatz 2.2, S. 13.

[8] „Ergebnis" entspricht in seiner fachlichen Definition in etwa der „Archäologiefläche" nach ADeX. Vgl. Göldner/Himmelmann/Poelmann et al., Absatz 2.1, S. 12.

[9] Wenn innerhalb einer Maßnahmefläche gravierende Unterschiede bezüglich der Erhaltung noch vorhandener archäologischer Substanz vorliegen, erfordert dies die Ausweisung mehrerer Ergebnisse im FIS. Dagegen kann ein Ergebnis mehrere Zeit-Typ-Thesaurierungen haben. In der Praxis werden sehr häufig unterschiedliche Zeitstellungen zu einem Ergebnis zusammengefasst. Gewissermaßen wird die wissenschaftliche Erkenntnis der Erhaltung „untergeordnet". Der Erhaltungszustand ist vor allem für die denkmalpflegerische Arbeit wichtig. Dem gegenüber steht der Ansatz, jeweils alle Befunde einer Zeitstufe oder Kulturgruppe zu einer „Fundstelle" zusammenzufassen. Vgl. dazu: Wilbertz, O. M. 2004: Archäologische Begrifflichkeit und elektronische Datenverarbeitung. In: Die Kunde N. F. 55, 96.

[10] „Denkmal" entspricht in seiner fachlichen Definition in etwa der „Schutzfläche" nach ADeX. Vgl. Göldner/Himmelmann/Poelmann et al., Absatz 2.3, S. 14.

[11] Wanninger, R. 2011: Öffentliche Präsentation der Denkmäler. In: Bayerischer Landesverein für Heimatpflege e.V. (Hrsg.), Forum Heimatforschung. Ziele – Wege – Ergebnisse, Heft 15, 98–113.

[12] Hierbei handelt es sich um eine erst 2011 nachträglich im Fachinformationssystem eingeführte Gliederung des Thesaurus. Davor waren die Planungstypen in vorbereitende Planungen und verbindliche Planungen aufgeteilt.

[13] „Vermutungsflächen" gehören, wie die „Denkmäler" auch, gemäß der ADeX-Definition zur Kategorie „Schutzflächen". Vgl. Göldner/Himmelmann/Poelmann et al., Absatz 2.3, S. 14.

[14] Die „Gruppierung" von Planungen und Maßnahmen dient dabei im Wesentlichen der leichteren Auffindbarkeit inhaltlich zusammengehöriger Datensätze. Bei den Maßnahmen können das etwa einzelne in zeitlichem Abstand ausgegrabene Parzellen im Rahmen der Erschließung eines Baugebietes sein. Bei den Planungen kann zum Beispiel die Aufstellung eines Flächennutzungsplanes mit den darauf folgenden Flächennutzungsplanänderungen zu einer Gruppe zusammengefasst werden.

[15] Vgl. Haberstroh, J. / Irlinger, W. / Sommer, C. S. 2012: Die Fundchronik für Bayern – „Abgesang" auf ein ambivalentes Produkt der Bodendenkmalpflege. Berichte der Bayerischen Bodendenkmalpflege 52, 2012, 9–17.

[16] In Bayern fehlt eine gesetzliche Regelung zum Eigentum an archäologischen Funden im Denkmalschutzgesetz nach wie vor, sodass BGB § 984 zur Anwendung kommt. Das Fundmaterial aus Begehungen externer Mitarbeiterinnen und Mitarbeiter befindet sich deshalb in den seltensten Fällen in öffentlicher Hand, sondern ist zumeist im Besitz des Finders und/oder des Grundstückseigentümers.

[17] Herrn A. Thomschke, Uttenreuth, danken wir für sein Einverständnis zur Veröffentlichung der Daten. Dem Einsatz von Frau A. Rahm, Bayerisches Landesamt für Denkmalpflege, ist die vollständige Erfassung aller Begehungsergebnisse von Herrn A. Thomschke im Fachinformationssystem zu verdanken.

[18] Wie Anm. 9.

[19] Zu jedem Begriff im Zeit-Thesaurus wird außerdem die Laufzeit hinterlegt. Der Beginn und das Ende einer Epoche oder Kultur sind als Jahreszahl in jeweils einem eigenen Feld gespeichert.

[20] Technisch wäre im Fachinformationssystem auch die Zuordnung eines Terminus zu mehreren Oberbegriffen möglich. Aus Gründen der leichteren Pflegbarkeit wurde in der Bodendenkmalpflege aber auf den Einsatz polyhierarchischer Thesauri verzichtet. In der Bau- und Kunstdenkmalpflege kommt dies dagegen häufiger vor. Die „Autobahnkirche" im Funktions-Thesaurus, der in etwa mit dem in der Archäologie eingesetzten Typ-Thesaurus vergleichbar ist, ist beispielsweise sowohl dem Oberbegriff „Konfessionelle Bauten" als auch der Kategorie „Verkehr" zugeordnet.

[21] Die Hierarchien der Thesauri im FIS werden über ein Feld „Parent-ID" gebildet, das die Thesaurus-ID des jeweils zugehörigen Oberbegriffes enthält. Durch Änderung der Parent-ID kann somit ein Begriff im Thesaurus ohne größeren Aufwand neu zugeordnet werden. Da dieses Datenbankschema bei nahezu allen Vokabularen im FIS zum Einsatz kommt, lässt sich auch die nachträgliche Einführung einer Hierarchieebene in einer zunächst einfachen Werteliste ohne Weiteres bewerkstelligen.

ARCHIVIERUNG UND DOKUMENTATION

Archäologische digitale Daten: Authentizität und Funktionalität

David Bibby, Reiner Göldner

Archäologische Informationen müssen zukünftig digital archiviert werden. Es führt kein Weg daran vorbei. Grabungsberichte, Befundbeschreibungen, Fotodokumentationen, Grabungspläne, Fundstellen- und Denkmalkartierungen, 3-D-Scans von Grabungen und Funden usw.: immer öfter werden sie digital erzeugt und immer seltener greift man zur Bewahrung noch auf analoge Kopien, auf Papier oder Film zurück. Die digitalen Daten bieten eben weit mehr Möglichkeiten, mehr Funktionalität als die analogen Entsprechungen. Inwiefern kann man aber bei digitalen Daten auch nach vielen Jahren und vielen Systemwechseln noch darauf vertrauen, dass sie „echt" sind, „authentisch"?

Klassische Authentizität

In der Literatur der letzten zwei Dekaden über Archivierung von digitalen Daten wird immer wieder versucht zu definieren, was „Authentizität", bezogen auf archivierte oder zu archivierende digitale Daten, bedeutet. Noch immer scheint sich keine endgültige allgemein akzeptierte Meinung gebildet zu haben. Selbst die im Allgemeinen für digitale archäologische Archivierung wichtigen Richtlinien „Kriterienkatalog vertrauenswürdige digitale Langzeitarchive" (Nestor 2008), die sich ausführlich mit einer „konventionellen" Auslegung von „Authentizität", bezogen auf digitale Daten, auseinandersetzen, befassen sich kaum mit dem Spannungsfeld Authentizität-Funktionalität.
„Authentizität" bedeutet allgemein „verbürgt, echt" (Brockhaus), Wikipedia beschreibt die Bedeutung mit „Echtheit im Sinne von ‚als Original befunden'". In der „klassischen" Archivierung beschreibt „Authentizität" die Tatsache, dass die archivierten Dokumente und Gegenstände während der Archivierung original, unverändert und unversehrt bleiben. Sie kommen nicht zu Schaden und der Zustand sollte bei einer Besichtigung/Nutzung auch nach vielen Jahren der Archivierung erkennbar identisch sein mit dem Zustand, in dem sie waren, als sie deponiert wurden. Anders formuliert: Inhalt, Form und Erscheinungsbild sollen unangetastet und integer bleiben. Es muss möglichst gewährleistet sein, dass Dokumente während der Archivierung nicht manipuliert oder gar gefälscht werden. Die Aufgabe des Archivars ist: „to hand on the documents as nearly as possible in the state in which he received them, without adding or taking away, physically or morally, anything: to preserve unviolated, without the possibility of suspicion, every element in them, every quality they possessed when they came to him" (Jenkinson, 1965).

Moderne Authentizität

Es ist selbstverständlich, dass auch digitale Daten möglichst unangetastet bleiben sollen, um als authentisch zu gelten. Aber die Zielsetzung von 1965 beschreibt nicht mehr das ganze Bild: Authentizität in der klassischen Archivierung zielt auf den Erhalt des physischen Originals. Bei der Anwendung dieses klassischen Authentizitätsbegriffs auf digitale Daten gibt es jedoch einige problematische Aspekte:

- Ein digitales „Original" lässt sich „klonen" – das heißt identisch reproduzieren (kopieren). Damit ist Originalität kein Kriterium mehr für Authentizität.

- Ein digitales Dokument bedarf für seine Nutzung (ansehen, lesen usw.) einer bestimmten Systemumgebung, die in die Betrachtung zur Authentizität einbezogen werden muss. Eine veränderte Systemumgebung führt eventuell zu einer veränderten Darstellung und damit zum Verlust von Authentizität im klassischen Sinn.

- Digitale Systeme haben enorm kurze Innovationszyklen, was die klassische „Erhaltungs"-Strategie und damit auch die Authentizität stark einschränkt. Ein Original ist aufgrund der schnellen Erneuerung digitaler Systeme nur kurzzeitig „lebensfähig". Ein längeres „Überleben" geht meist mit Migration einher, was aber physische Veränderung und Verlust von Integrität bedeutet.

Damit wird klar, dass eine unspezifizierte, absolute Authentizität bei der Archivierung digitaler Daten nicht erreichbar sein wird. Was kann man aber erreichen? Schauen wir uns dazu noch einige Gedanken zum Thema Authentizität aus dem IT-Bereich an.

- Wikipedia bietet an: „In der Informationssicherheit bezeichnet Authentizität die Eigenschaften der Echtheit, Überprüfbarkeit und Vertrauenswürdigkeit. Die Überprüfung einer behaupteten Eigenschaft wird als Authentifikation bezeichnet." Und bezieht sich dabei auf das Internet Security Glossary (RFC 4949, S. 29): „The property of being genuine and able to be verified and be trusted."

Hier findet man den ersten Ansatz – es geht um überprüfbare Echtheit und das Zutreffen behaupteter Eigenschaften.

- The National Archives (UK) meinen: „... ein authentisches Dokument ist, was es vorgibt zu sein und frei ... von Abänderungen oder Verfälschungen" (TNA, S. 16). An anderer Stelle wird Authentizität durch Identität (Attribute, die die Einzigartigkeit dokumentieren, zum Beispiel Autor, Erstellungsdatum, Thema und anderes) und Integrität beschrieben. Bezüglich Integrität wird ausgeführt: „Ein Dokument besitzt Integrität, wenn es über den gesamten Zeitraum seiner Lebensdauer hinweg in allen seinen wesentlichen Merkmalen vollständig und unverfälscht ist. Dies bedeutet nicht, dass ein Dokument genau dasselbe sein muss wie das ursprüngliche Erstdokument, um seine Integrität zu demonstrieren. Ein Dokument kann als im Wesentlichen vollständig und unverfälscht betrachtet werden, wenn die für dessen Zweckbestimmung zu kommunizierende Botschaft unverändert ist." (TNA, S. 9).

- In der DIN „Information und Dokumentation – Leitfaden zur Informationsübernahme in digitale Langzeitarchive" finden wir: „Authentizität beinhaltet, dass nur beabsichtigte, dokumentierte Veränderungen im Sinne der Erhaltungsmaßnahmen am Objekt durchgeführt wurden. Maßstab für die Authentizität sind die signifikanten Eigenschaften des Informationsobjekts." (DIN 31645, S. 6).

Der Vollständigkeit halber wollen wir erwähnen, dass Authentizität im IT-Bereich oft auf (personelle) Identität reduziert wird: „Mit dem Begriff Authentizität wird die Eigenschaft bezeichnet, die gewährleistet, dass ein Kommunikationspartner tatsächlich derjenige ist, der er vorgibt zu sein". (BSI). Dabei wird Authentifizierung (als Prozess der Überprüfung von Authentizität) im Sinne von Identifizierung (einer Person) verstanden, zum Beispiel, um Zugang zu einem Computersystem zu bekommen. Wir wollen – insbesondere für unsere archäologischen Daten – diese Reduzierung hier nicht aufnehmen und uns auf eine „allgemeine" Authentizität beziehen.

Bei Authentizität geht es also um behauptete Eigenschaften, um zweckbestimmte wesentliche Merkmale! Es geht nicht um absolut alles. Wir haben die Wahl, welche Eigenschaften wir zur Feststellung von Authentizität ansetzen. Dateigröße, Testsumme, digitales Zertifikat, Schriftgröße und -farbe, Inhalt, Form, Bedeutung, Funktion usw. Welches sind die Echtheitsmerkmale, die langfristig unverändert bleiben müssen? Wie unterscheiden sich Kriterien und Gewichtungen bei verschiedenen Datengattungen? Und wie kann man den Erhalt dieser Merkmale bei Migrationen nachweisen? Diese Fragen bleiben bei der herkömmlichen Archivierung im Hintergrund, für die Archivierung digitaler Daten sind sie jedoch wesentlich.

Beispielsweise Funktionalität

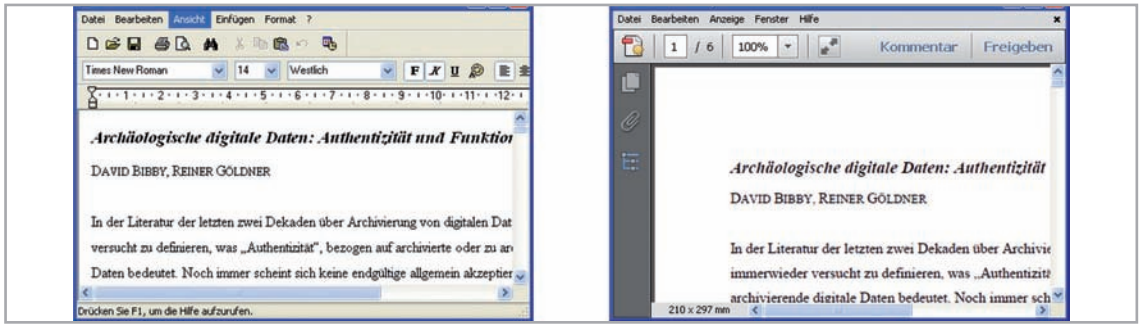
Zum Schutz digitaler Daten verwendet man beispielsweise Hashwerte (Prüfsummen) oder digitale Wasserzeichen. Am gebräuchlichsten sind jedoch digitale Signaturen (zum Beispiel das RSA-Verfahren). Mit diesen Verfahren kann man bei Erfordernis die binäre Unversehrtheit eines Datensatzes nachweisen. Für einfache Textdaten und digitale Fotografien erscheint dies ausreichend. Möglicherweise ist es aber doch nicht so einfach, schauen wir uns dazu einige Beispiele an.

Beispiel 1:
Wenn ein Textdokument mit dem Font Courier New 10 erstellt wurde, wäre eine Ausgabe (zum Beispiel Ausdruck) mit exakt gleichem Wortlaut, aber mit dem Font Times New Roman 12, noch authentisch?

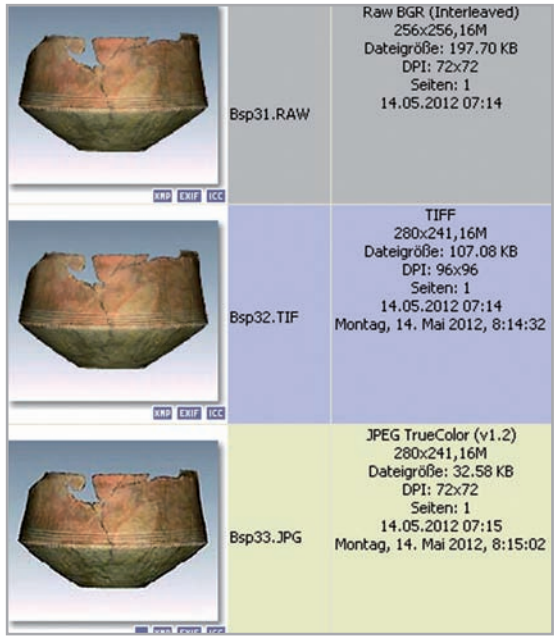
Dieser Satz wurde mit der Schriftart Courier New 10 geschrieben.

Dieser Satz wurde mit der Schriftart Courier New 10 geschrieben.

1 Textpassage in Schriftart Courier (oben) und Times (unten).



2 Präsentation eines Textes: RTF in WordPad (links), PDF in Acrobat Reader (rechts).



3 Kurzinformationen zu einem Bild in verschiedenen Formaten (RAW, TIFF, JPEG).

Beispiel 2:
Wenn ein Textdokument im Rich Text Format (RTF) erstellt wurde, wäre eine Ausgabe im Portable Document Format (PDF) mit exakt gleichem Wortlaut und mit exakt gleichem Font noch authentisch, da sich dabei Größe und innere Struktur der Textdatei und damit Hashcodes und digitale Signaturen ändern?

Beispiel 3:
Wenn eine Bilddatei im RAW-Format aufgenommen, zu einem TIFF entwickelt und anschließend als JPEG gespeichert wurde – wo endet die Authentizität, wenn Dynamikbereich und Farbraum geändert, das Aussehen „verbessert" und die Datei verlustbehaftet gespeichert wird – ohne dass visuell Unterschiede erkennbar sind?

Diese oberflächlich banalen Fragen entpuppen sich beim zweiten Blick als gar nicht so trivial. Abgesehen von den oben genannten Dateigattungen, bei denen eine zufriedenstellende Definition von Authentizität für Langzeitarchivierung von archäologischen Unterlagen ohne allzu großen Aufwand zu erzielen sein dürfte, beinhalten digitale Daten, die archäologischen Interventionen entstammen, zum großen Teil nicht nur textlichen oder bildlichen „Inhalt", sondern auch komplexe Informationsstrukturen und „Funktionalitäten". Deshalb ist das konventionelle Verständnis von Authentizität für archäologische digitale Daten nicht befriedigend. Einige weitere Beispiele sollen dies zeigen.

Beispiel 4:
Grabungsdatenbanken beinhalten nicht nur Datensätze mit Informationen über die Grabung, sondern auch Formulare zur Eingabe und Berichte zur Ausgabe der Informationen sowie Funktionalitäten wie Abfragen, Filter, SQL-Funktionen, Diagrammtools etc., die den Umgang mit den in der Datenbank vorhandenen Informationen erleichtern und eine Grabungsauswertung ermöglichen. Hier ist der Begriff „Authentizität" zu erweitern beziehungsweise anders zu gewichten. Es muss also nicht nur der Sachinhalt der Daten erhalten bleiben, sondern auch und vor allem die Funktionalität – möglicherweise wird in diesem Fall das Erscheinungsbild geopfert werden müssen, um die Funktionalität zu erhalten. Während es zum Beispiel Grund zur Hoffnung gibt, dass ein PDF-A-Dokument, das im Jahr 2005 abgelegt wurde, im PDF-Viewer des Jahres 2020, auch nach möglichen Migrationen zu neueren PDF-Formaten, zumindest im Erscheinungsbild nicht von dem ursprünglich abgelegten Dokument zu unterscheiden ist, wird dies beispielsweise für eine im Jahr 2000 abgelegte, komplex strukturierte „Paradox"-Datenbank sicher nicht der Fall sein. Um den vollen Nutzen dieser Datenbank über die Dauer der (permanenten) Archivierung zu „retten", werden aller Wahrscheinlichkeit nach mehrere „Portierungen" beziehungsweise Migrationen nötig sein.

Es wird sofort klar, dass das ursprüngliche Erscheinungsbild der Datenbank nicht zu erhalten ist. Bei der Frage der Authentizität in diesem Szenario geht es eben nicht vorwiegend um Inhalt und Erscheinungsbild. Es geht nicht darum, dass das Aussehen dessen, was aus dem Archiv herauskommt, identisch sein sollte mit dem, was ursprünglich deponiert wurde, sondern dass Inhalt und Funktionalität erhalten bleiben.

Beispiel 5:
Bei CAD-Grabungsplänen ist die Situation etwas anders. Hier geht es vorwiegend um eine Funktionalität, bestehend aus Inhalt (aufgenommene Objekte, Funde etc.) und Struktur (Aufbau der Ebenen und ihr Verhältnis zueinander). Die Pläne spiegeln vielfältige räumliche und inhaltliche Bezüge der erfassten Befunde, Funde, Schichten, Plana etc. wider. Daraus ergibt sich eine Vielzahl verschiedener Erscheinungsbilder mit unterschiedlichsten Aussagen innerhalb derselben Datei. Eine CAD-Datei ist nicht einfach eine „statische Zeichnung". Dies wäre der Fall, wenn ein spezifischer Zustand einer CAD-Zeichnung als PDF-Dokument „gedruckt" worden wäre. Dann würde sie den Kriterien für Textdaten und Bitmaps (oben) entsprechen. Die CAD-Datei kombiniert Ebenen (Layer), die geometrische – und unter Umständen Sachinformationen – beinhalten können. Durch Zu- oder Abschalten von Layern können in der gleichen Datei ver-

4 Sammelsurium alter und neuer Datenbanken: 1982 dBASE II, 1996 Paradox 7 auf Disketten, 1997 Grabungsdatenbank im LfA Dresden auf Disketten, 2003 Access auf CD, 2008 PostgreSQL im Internet.

schiedene Erscheinungsbilder mit unterschiedlichen Aussagen erzeugt werden. Dies ist ein dynamischer Prozess – eine „dynamische Funktionalität". Bei der Archivierung von CAD-Daten ist es nicht erstrebenswert, nicht einmal wünschenswert, ein Erscheinungsbild zu fixieren. Authentisch zu archivieren bei CAD-Daten ist eben die Fähigkeit dieser Daten, ihr Erscheinungsbild zu ändern!

Beispiel 6:
Fundstellenkataster leben von geografischen Bezügen und werden daher üblicherweise mit Geoinformationssystemen (GIS) bearbeitet. Im Allgemeinen trifft für die grafische Seite der GIS-Daten die Aussage zu CAD-Daten und für die Sachdaten die Aussage (oben) zu Datenbanken zu. Hier geht es sowohl um die Funktionalität der Datenbank wie auch die „dynamische Funktionalität" der grafischen Darstellung der Geometrie. Möglicherweise ist im Fall von GIS-Daten die Situation etwas einfacher, da sich das Dreikomponenten-Shape-Format der Firma Esri für Geometrie und Sachinhalte als Quasi-Standard breit durchgesetzt hat, ohne Anzeichen dafür, dass sich dies mittelfristig ändern wird. Bei GIS-Daten wird eins besonders deutlich: Der Grad der Funktionalität wird nicht nur durch die (fixierte) Struktur der Daten, sondern durch die im benutzten GIS-Programm zur Verfügung stehenden Tools (letztlich also durch die Systemumgebung) bestimmt. Für eine adäquate Nutzung solcher archivierter Daten ist es erforderlich, dass die verfügbare Systemumgebung (inklusive Tools) mit der Systemumgebung der Entstehungsphase funktionell vergleichbar ist.

Beispiel 7:
Die umfassende Eroberung der dritten Dimension (3-D) ist eine noch junge, aber äußerst beliebte Errungenschaft der Archäologie. Während 3-D-Daten geometrisch komplexe Wirklichkeiten abbilden, besitzen sie keine oder nur wenig Funktionalität im Gegensatz zu Datenbanken, CAD- oder GIS-Daten. Insofern haben sie bezüglich ihrer Archivierung mehr mit Bitmaps oder Textdateien gemeinsam. Es ist sogar so, dass der „niedrigste gemeinsame Nenner" für 3-D-Daten eine Punktwolke im ASCII-Format mit den Spalten (x, y, z) oder mit Farbinformation (x, y, z, r, g, b) ist. Eine solche Punktwolkendatei kann zwar sehr groß werden, sie bleibt jedoch einfach strukturiert. Bei Mesh-Daten setzen sich einige wenige Formate durch, zum Beispiel STL, PLY. Falls Fototextur vorhanden ist, werden aktuell Virtual Reality Models erzeugt, die aus mit 3-D-Koordinaten gesteuerten Bitmaps bestehen. Unter Berücksichtigung dieser Tatsachen kommt man zu dem Schluss, dass 3-D-Daten, was Datenauthentizität betrifft, ähnlichen Kriterien unterliegen, wie sie für Digitalbilder angesetzt werden sollen. Im Archiv muss lediglich dafür Sorge getragen werden, dass das, was ursprünglich deponiert wurde, leserlich und im Inhalt und Erscheinungsbild unverändert bleibt.

Resümee

Bei der Archivierung digitaler Daten gilt der Erhalt des originalen Datenformats nicht wirklich als vorwiegender Maßstab für Authentizität. Es geht darum, dass die Inhalte, Erscheinungsbilder, Funktionalitäten und Strukturen der deponierten Daten für den Nutzer erhalten bleiben, und zwar in unterschiedlichen Maßen, abhängig von der Art der Information, dem Erscheinungsbild, der Funktionalität etc., welche die Originaldaten beinhalten. Verschiedene zu archivierende archäologische Digitalunterlagen besitzen verschiedene Charakteristika. Es gilt, möglichst all diese Charakteristika zu erhalten. Wenn allerdings Entscheidungen getroffen werden müssen, ob bestimmte Charakteristika (im Sinne von Form, Funktion, Erscheinungsbild, Inhalt) während des (aktiven) Archivierungsprozesses geopfert werden sollen, muss von vornherein klar sein, welche möglicherweise geopfert werden können und welche gar nicht. Zur Erinnerung: „Ein Dokument kann als im Wesentlichen vollständig und unverfälscht betrachtet werden, wenn die für dessen Zweckbestimmung zu kommunizierende Botschaft unverändert ist." (TNA, S. 9).

In der DIN 31645 (S. 12) findet man dazu: „Digitales Langzeitarchiv und/oder Produzent müssen die signifikanten Eigenschaften der zu übernehmenden Informationsobjekte definieren. Die Beteiligten entscheiden, welche Darstellungsabweichungen eines Informationsobjekts in zukünftigen technischen Umgebungen noch akzeptabel sind und welche nicht. Ein maßgebliches Kriterium sind die Bedürfnisse der Nutzer des digitalen Langzeitarchivs (Zielgruppe)." sowie „Das digitale Langzeitarchiv muss die für jedes Informationsobjekt ausgewählten und eindeutig benannten signifikanten Eigenschaften als Metadaten verzeichnen. Diese Metadaten sind der Gradmesser für die Beurteilung anstehender Erhaltungsmaßnahmen (Konverterprogramme, Emulatoren) und zukünftiger Nutzungsplattformen." In diesem Zusammenhang ist das InSPECT-Projekt erwähnenswert, wo anhand von Audiodateien, E-Mails, Rasterbildern und strukturiertem Text einige signifikante Eigenschaften fachneutral erarbeitet und untersucht wurden. Diese Arbeit wird fortgesetzt und vor allem auf das Fachgebiet der Archäologie erweitert werden müssen!

Um Authentizität im beschriebenen Sinne zu gewährleisten, sind recht umfangreiche Metadaten mit technischem, inhaltlichem und funktionellem Bezug erforderlich, und zwar:

- Metadaten über den Zustand bei der Übernahme ins Archiv und
- laufende Metadaten-Einträge über alle Prozesse, die den Datensatz während des Aufenthalts im Archiv verändern.

Diese Metadaten haben Protokollcharakter (und müssen auch entsprechend abgesichert werden), damit der Archivbestand vertrauenswürdig präsentiert werden kann. In diesem Sinn versteht es sich fast von selbst, dass die Originaldaten bei der Migration nicht durch die neue Version ersetzt werden, sondern erhalten bleiben.

Die Frage ist allerdings, wie umfassend dieser Ansatz in der Praxis realisierbar sein wird. Es ist schwer vorstellbar, Inhalt und Funktionalität (oder gar Semantik) migrierter Dokumente vollständig zu prüfen. So wird man sich wohl oft mit einem bestimmten Grad an Vertrauen begnügen müssen.

Die hier skizzierte Thematik ist zu komplex, als dass sie in dieser Kürze abschließend behandelt werden kann. Dies ist lediglich ein Denkanstoß, der dazu dient, auf die Problematik der Definition(en) von Authentizität bei digitalen Daten in der Archäologie aufmerksam zu machen.

Eine Problematik, die bisher zu wenig Beachtung gefunden hat, die aber dringend angegangen werden muss, um die künftige Integrität, Zugänglichkeit, Nutzbarkeit und nicht zuletzt das Überleben speziell unserer digitalen archäologischen Daten, für die es keine „Haltbarkeitsgrenze" gibt, zu gewährleisten.

Quellen, Literatur, Internet

Brockhaus in 15 Bänden, Bd. 1, Leipzig-Mannheim 1997: authentisch – „verbürgt, echt" (S. 369).

BSI – Bundesamt für Sicherheit in der Informationstechnik. Glossar: Authentizität. (https://www.bsi.bund.de/DE/Themen/Cyber-Sicherheit/Glossar/cs_Glossar_ node.html).

DigiCULT – Technology Challenges for Digital Culture. (http://www.digicult.info).

DIN 31645, Information und Dokumentation – Leitfaden zur Informationsübernahme in digitale Langzeitarchive.

Duranti, Luciana: „Reliability and Authenticity: The Concepts and Their Implications." Archivaria 39 (Spring 1995).

Hirtle, Peter B.: „Archival Authenticity in a Digital Age". Council on Library and Information Resources, Publication 92.
(http://www.clir.org/pubs/reports/pub92/hirtle.html).

InSPECT-Projekt, Final Report. (http://www.significantproperties.org.uk/inspect-finalreport.pdf).

InterPARES Project, all documents. (http://www.interpares.org).

Jenkinson, Hilary: A Manual of Archive Administration. London, 1965.

Nestor: Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Version 2, Frankfurt 2008.
(http://edoc.hu-berlin.de/series/nestor-materialien/8/PDF/8.pdf).

Park , Eun G.: The Role of Authenticity in the Life Cycle of Digital Documents. Graduate School of Library & Information Studies, McGill University. Montreal, 2003.

RFC 4949, Internet Security Glossary, Version 2. IETF. (https://tools.ietf.org/html/rfc4949).

TNA – The National Archives: Generische Anforderungen zur langzeitlichen Erhaltung elektronischer Informationen: 1 Definition der Merkmale authentischer Dokumente.
(www.nationalarchives.gov.uk/documents/de_generic_requirements_vol1.pdf).

Wikipedia: Authentizität.
(http://de.wikipedia.org/wiki/Authentizität).

(Alle Internet-Links mit Stand vom 22.05.2012).

Digital curation at the Archeology Data Service

Catherine Hardman, Julian D. Richards

The problems of curating digital archaeological data sometimes tend to be overlooked in larger considerations of an archaeological archives crisis. Compared with the more tangible storage issues facing tons of bulk finds of pottery or the curation of the paper archive, digital data has been regarded as less of a problem, or even as the solution to the problem (Perrin 2002). After all, digital data consumes less space than either finds or paper, and digitisation can provide easier access.

However, it is now recognised that preservation of digital data introduces a whole new set of problems for the curator: data corruption, hardware redundancy, changes in file formats, software migration, and insufficient documentation. It is inadequate to copy data to DVD and put them in the proverbial ‚box on the shelf'. Digital data require active curation. Digital data preservation has developed as a sub-discipline in its own right with its own vocabulary, standards, and organisations. In the UK, the Digital Curation Centre (DCC) and the Digital Preservation Coalition (DPC) are two such bodies that have been working to highlight the digital preservation issues. The call has also been taken up at the European level with a number of collaborative partnerships working together to develop transnational research infrastructures for digital preservation and access and shared standards.

A survey into the digital preservation of research outputs in Europe (Kuipers and van der Hoeven 2009) found considerable awareness of the problem. Preservation is widely accepted to be very important. 91% of researchers surveyed considered the possibility of re-analysis of existing data as the most important driver for preservation, whilst 87% also believe that public funding is a key reason. The same survey, however, found less agreement about who should pay for the costs of digital preservation of research data, with many respondents providing multiple solutions. 61% of researchers believe that national government should pay for preservation, whilst 41% believe that their own organisation should bear the cost.

The problem is particularly acute in archaeology. Archaeological data are increasingly ‚born digital' and represent the primary record of sites destroyed during the development or fieldwork process. Frequently, there are no paper substitutes. Furthermore, archaeologists create large quantities of data and in a greater variety of formats than most other disciplines. (Austin and Mitcham 2007; Austin et al. 2009). Yet, such data are frequently essential to underpin the excavator's interpretation of a site or to allow its reinterpretation by others.

This issue was recognised at a relatively early stage in the UK. The Archaeology Data Service (ADS) was established in 1996 as part of a wider Arts and Humanities Data Service (AHDS). It was charged with preserving research data created within the Higher Education (HE) sector and making them available for re-use for teaching, learning, and research (Richards 2008). Core funding was initially provided by the Joint Information Systems Committee (JISC): the body responsible for UK HE computing and information infrastructure. The Arts and Humanities Research Board (AHRB, later Council – AHRC) later became the primary research funding body for the sector. This meant that University-based researchers could deposit their data with the AHDS which undertook to use its best efforts to preserve these data free of charge into perpetuity.

Core funding continued to support free deposit from the HE sector. As ADS developed, however, it became apparent that it was important for it to develop a business model which could accommodate the preservation of data emanating from the commercial and public sectors, and that derived from archaeological investigations funded by bodies such as English Heritage (EH). A variety of funding models were considered, including charging the users of the data. This was soon dismissed for a variety of reasons. On pragmatic grounds, it was believed that the administrative overheads of charging large numbers of users would not be cost effective. Also, one of the aims of ADS has always been to encourage re-use, not to put obstacles in the way of re-users. Secondly, the data archived rarely have any great commercial value; there was a strongly expressed view that these data had generally been collected using public funding and, therefore, they should be available free to all at the point of use, an approach which was also in line with the broader convention enshrined within the 1992 Valletta Treaty (1) that ‚the polluter pays.'

Therefore, the ADS adopted a business model whereby the costs of archiving outside the academic sector would have to be covered on a project basis. It was also clear that this would have to be a one-off, or one-time only, charge, levied at the point of deposit. Whilst archiving costs would inevitably be ongoing, it was clear that funders would have to budget for the

full costs of a project when it was commissioned. There was already an existing analogy for this approach in the archives world. Museums had developed a system of box charges, levied when the physical archive was deposited, and based upon its volume.

The principle of a one-off deposit charge implied that all future costs had to be covered up front. The ADS preservation policy is based on migration of files to the latest formats, and it is impossible to predict how many times this might have to be undertaken (ADS 2009). This approach to digital preservation, based upon conversion of files to a sustainable preservation format and a dissemination version at the time of accession has become enshrined within the Open Archival Information System (OAIS) developed by NASA (ISO 14721:2003). The OAIS provides a framework for the understanding and increased awareness of archival concepts needed for long term digital information preservation and access and has been adopted by the ADS.

At an early stage, ADS was funded by EH to undertake the Digital Archiving Pilot Project for Excavation Records (DAPPER) project. The aims of the DAPPER project were to create two exemplar archives to allow ADS to gain an understanding of archiving costs, as well as to investigate the level of interest in re-use of the data. The two projects were the excavations at Eynsham Abbey, by the Oxford Archaeological Unit, and work at the Royal Opera House, undertaken by the Museum of London Archaeology Service (Austin et al. 2001). It was agreed that rather than creating an online interactive interface, access to the data should be via simple downloadable files in open formats. The implication of this was that the majority of the archiving costs were front-loaded.

The figures derived from the small sample of two projects used in the DAPPER project indicated that digital archiving costs would be within 1-5% of the total costs of a project. Economies of scale become a significant factor especially for larger projects. It has also become clear that the digital archiving costs can frequently be offset against savings in publication costs. Digital dissemination may provide the best means of making detailed data and information available, thereby reducing the scale of paper publications and providing a safer option to the inclusion of the CD in the back of the book.

The DAPPER project fed directly into the development of an explicit ADS Charging Policy, now in its 4th edition (ADS 2007). This was first introduced in 1999 and was based on a formula of fixed charges per file. Simply put, files were banded according to number and complexity of format. Images, text, simple 'flat' spreadsheets, and tables cost less than CAD, GIS and relational databases.

Two factors led ADS to undertake a major revision of its Charging Policy in 2007. Firstly, the Big Data project led us to refine our thinking on preservation costs (Austin and Mitcham 2007). The Big Data project recognised that with the introduction of new technologies such as maritime survey, laser scanning and LiDAR, archaeologist's capacity to generate massive data sets was continuing to grow. This was exacerbated by the problem of distinguishing primary from processed data in many of these applications, and by the use of proprietary formats. Secondly, the AHRC announced that it was going to withdraw funding from AHDS and it has encouraged the development of an ADS business plan which aims to cover more costs by full charging for services.

Under the revised Charging Policy, the costs of digital curation are calculated on the basis of four elements, which are categorised in terms of staff time. Examples are provided online (http://archaeologydataservice. ac. uk/advice/collectionsPolicy).

a. **Management and administration:** This reflects the time spent in processing the deposit, including negotiation with the depositor, dealing with rights management issues and deposit licences, and issuing invoices.

b. **Ingest:** This cost reflects the time required to migrate the data to ADS preferred formats; the harmonization of filenames; the creation of delivery and preservation formats and their transfer to offline storage; checksum procedures; and creation of file level and project level metadata.

c. **Dissemination:** The ingest charges include an allowance to cover the creation of a basic archive delivery web page within the ADS catalogue and the delivery of data via simple file download.

d. **Storage and refreshment**. The term 'storage' encompasses the ongoing periodic process of data refreshment. In order to take advantage of technological advances, archives have to periodically upgrade systems. For example, the ADS has progressed through three generations of equipment during its 10+ years of existence, a five year upgrade cycle. This is expensive both in terms of equipment and staff time. The long term cost of storage is often difficult to conceptualize but a dataset maintained for 100 years would go through 20 refreshments based on the five year cycle.

No doubt the ADS Charging Policy will continue to evolve as we gain more practical experience, informed by ongoing studies into preservation costs (Beagrie 2008; 2009). Indeed an opportunity to refine the policy has been recently afforded in the SWORD ARM project (http://archaeologydataservice.ac.uk/research/swordarm) in which we are hoping to achieve on line deposition and some degree of automation which may allow us to reduce ingest element of our costs. Nevertheless, the principle of a one-off deposit charge, passed onto research funding bodies, state heritage organisations, and commercial developers, has been demonstrated to be robust.

References Cited

Archaeology Data Service (ADS)
2007 *Charging Policy*. 4th edition. Julian Richards, Catherine Hardman and Tony Austin, Electronic document, http://archaeologydataservice.ac.uk/advice/chargingPolicy, accessed 11/06/2012.
2009 *Preservation Policy*. Electronic document, http://archaeologydataservice.ac.uk/attach/preservation/PreservationPolicyV1.3.1.pdf,
accessed 11/06/2012.

Austin, Tony, Jonathan Bateman, Stuart Jeffrey, Jen Mitcham, and Kieron Niven (edited by Kieron Niven)
2011 Marine Survey. *A Guide to Good Practice*. Archaeology Data Service/ Digital Antiquity Guides to Good Practice. http://guides.archaeologydataservice.ac.uk/g2gp/RSMarine_Toc accessed on 11/06/2012.

Austin, Tony and Jen Mitcham 2007 *Preservation and Management Strategies for Exceptionally Large Data Formats: 'Big Data'*. http://archaeologydataservice.ac.uk/attach/bigDataDeliverables/bigdata_final_report_1.3.pdf, accessed 11/06/2012.

Austin, Tony, Damian J. Robinson, and Keith A. Westcott
2001 A digital future for our Excavated Past. In *Computing Archaeology for Understanding the Past: CAA 2000*, edited by Zoran Stančič and Tatjana Veljanovski, pp.289-296, BAR International Series 931, ArcheoPress, Oxford.

Beagrie, Neil
2008 *Keeping Research Data Safe*. Electronic document, http://www.jisc.ac.uk/publications/reports/2008/keepingresearchdatasafe.aspx,
accessed 11/06/2012.
2009 *Keeping Research Data Safe 2: The identification of long-lived digital datasets for the purposes of cost analysis.* Electronic document, http://www.jisc.ac.uk/publications/reports/2010/keepingresearchdatasafe2.aspx, accessed 11/06/2012.

Kuipers, Tom and Jeffrey van der Hoeven 2009 *PARSE.Insight: INSIGHT into issues of Permanent Access to Records of Science in Europe. Survey Report.* Electronic document,
http://www.parse-insight.eu/publications.php#d3-4,
accessed 11/06/2012.

Perrin, Kathy 2002 *Archaeological Archives: Documentation, Access and Deposition. A Way Forward.* English Heritage. Electronic document, http://www.english-heritage.org.uk/upload/pdf/archives.PDF, accessed 11/06/2012.

Richards, Julian D. 2008 Managing digital preservation and access: The Archeology Data Service. In *Managing Archaeological Resources: Global context, national programs, local actions, edited by Francis P. McManamon, Andrew Stout and Jodi A. Barnes, pp. 173–194.* One World Archeology 58, Left Coast Press, Walnut Creek, CA.

Erstellung von Bildern für Bildarchive

Sigmar Fitting

Die klassische professionelle analoge Fotografie und die Archivierung der Silberhalogenidfotos stellten einen Prozess höchster Redundanz dar. Von jedem Bildmotiv wurden mindestens zwei Negative belichtet, sorgfältig entwickelt, archivfest fixiert und ausgewässert. Diese Negative wurden, je nach Sicherheitsbedürfnis zusammen oder getrennt, an verschiedenen Orten als Rohmaterial archiviert und mit geeigneten Metadaten versehen. Das eigentliche Bild entstand daraufhin bei der Ausarbeitung im Labor, bei der durch die Auswahl des Ausschnittes und die Anpassung der Belichtung (Nachbelichtung, abwedeln etc.) die endgültige Bildaussage durch den Fotografen festgelegt wurde. In einem letzten Schritt erfolgte dann noch das Retuschieren, Aufziehen und Beschriften der Fotos. In aller Regel erfolgte darauf ein „Back-up" der fertigen Aufnahmen in Form eines Repros oder etwa durch die Mikroverfilmung (Beispiel: Marburger Index) der Archivkartons. Insgesamt also ein sehr aufwendiges Verfahren, welches die Notwendigkeit einer frühen Auslese der zu archivierenden Aufnahmen begründete.

Dieser klassische „Workflow" lässt sich im Grunde sehr gut auf die Archivierung von digitalen Fotografien abbilden.

1. Rohdaten

Nach derzeitigem Stand und Kosten der Fototechnik sollte man nur Aufnahmen akzeptieren, die parallel als Rohdaten und ausgearbeitete Bilddatei abgeliefert werden. Je nach zur Verfügung stehendem Speicherplatz könnte es sehr sinnvoll sein, die RAW-Dateien als Belichtungsserien von mindestens drei Aufnahmen mit einer Schrittweite von 2 EV[1] zu archivieren. Damit eröffnen sich Möglichkeiten, dem begrenzten Dynamikumfang von Digitalfotos etwa durch HDR/DRI-Methoden zu begegnen. Diese RAW-Dateien sollten dann sinnvoll benannt (Datenbank) und auf geeignete Weise mit Metadaten versehen werden und danach möglichst redundant an verschiedenen Orten und auf unterschiedlichen Speichermedien gesichert werden.

Die gleiche Vorgehensweise bietet sich bei Scans von Aufsichts- und Durchsichtsvorlagen an, wobei hier keine Belichtungsserien benötigt werden, da der abzubildende Dynamikbereich auch bei optimal belichteten Diapositiven leicht in einer 48-Bit RGB-Datei gespeichert werden kann. Insbesondere im Bereich der Digitalisierung von Durchsichtsvorlagen bestimmt die eingesetzte Hardware ganz entscheidend die Qualität der Digitalisate (Dichteumfang, Schärfe).

RAW-Dateiformate

Ein sehr hoffnungsvoller Ansatz bot 2004 die Einführung und Offenlegung des DNG-Formates als potentieller Standard für das „digitale Negativ"; dieser Standard konnte sich allerdings bis heute nicht recht durchsetzen. Immerhin gibt es zum Beispiel mit Adobe Camera Raw ein Tool zum Konvertieren der proprietären RAW-Formate.

Nikon etwa hat bis heute noch nicht alle Spezifikationen des eigenen RAW-Formats offengelegt. Bei Canon verändern sich die Spezifikationen von einem Kameramodell zum nächsten. Im High-End-Bereich sieht die Sache nicht wirklich besser aus, hier bleibt die weitere Entwicklung abzuwarten und bis dahin (soweit alle Bearbeitungsmöglichkeiten erhalten bleiben sollen) die kameraeigenen RAW-Dateien zu archivieren oder die RAW-Dateien in das DNG-Format zu konvertieren.

Im Scanbereich sieht es noch düsterer aus: Imacon etwa unterstützt in der neuesten Softwareversion das eigene „3f"-Format nicht mehr oder Vuescan-DNGs lassen sich nur mit Vuescan weiter bearbeiten, wenn man alle Bearbeitungsmöglichkeiten erhalten will. In diesem Bereich bietet sich eher das Speichern einer unkomprimierten 48-Bit Baseline TIFF-Datei an.

TIFF

An dieser Stelle einige Bemerkungen zum TIFF-Format: TIFF hat sich im Bereich der Druckindustrie als Quasi-Standard für Rasterbilder durchgesetzt. Allerdings verfügt das TIFF-Format über außerordentlich viele Optionen und Varianten (Farbtiefe, Komprimierung, Bit-Anordnung, Ebenenkomprimierung, Farbkanäle etc.). Dies führt sehr schnell zu Inkompatibilitäten, ein Ausweg stellt „Baseline TIFF" als Mindestspezifikation für alle „TIFF-fähigen" Programme dar. TIFF-Dateien bieten darüber hinaus beste Möglichkeiten, Metadaten (EXIF, IPTC und XMP) sowie ICC-Profile in die Bilddatei zu integrieren.

Empfehlungen zur Scan-Auflösung

Bei Farb- und Graustufen-Aufsichtsvorlagen hat man bei einer Auflösung von 600 dpi ausreichend Reserven, bei Bitmaps und Strichzeichnungen können je nach Vorlagengröße und gewünschtem Abbildungsmaßstab 1200 bis 2400 dpi sinnvoll sein. Bei Dia- und

Negativfilmen erzielt man jenseits von 2400 bis 3000 dpi (bezogen auf die Negativgröße) schon eine „perfekte Abbildung des Filmkorns".

2. Ausgearbeitete Bilddateien

Dateiformate

Für die Speicherung der fertigen Bilddateien gilt es eine Abwägung der sich teilweise widersprechenden Eigenschaften Speicherplatzbedarf versus Qualität versus Kompatibilität zu treffen. Eine sehr gute Grundlage bietet die oben skizzierte Archivierung der RAW-Dateien, möglichst zusammen mit den entsprechenden „Kochrezepten" für die „Entwicklung" der fertigen Bilder. Diese RAW-Dateien bieten die maximale Qualität bei geringstmöglicher Speichergröße der verlustfrei komprimierten Daten, allerdings zum Preis einer sehr geringen Kompatibilität.

Unter der Voraussetzung, dass die endgültig ausgearbeiteten Bilddateien nicht weiter bearbeitet werden sollen, stellt sich die Frage, ob diese nicht als moderat komprimierte JPG-Dateien, gegebenenfalls nach PDFA-1b exportiert, archiviert werden sollten. Nach heutigem Stand bietet das JPG-Format (sämtliche Digitalkameras mit Ausnahme einiger High-End-Digibacks speichern ausgearbeitete Bilddateien im JPG-Format) die größte Wahrscheinlichkeit, dass diese Dateien auch in vielen Jahren ohne weitere Konvertierungen noch zu öffnen sein werden. Selbstverständlich sollten alle Zwischenstufen, soweit diese gespeichert werden sollen, nicht komprimiert werden. Die Alternative mit der besten Kompatibilität stellt (Baseline)TIFF (gegebenenfalls mit LZW-Komprimierung, wodurch die Kompatibilität wieder eingeschränkt wird) zum Preis eines höheren Speicherplatzbedarfs dar.

Auflösung und Farbtiefe

Nach heutigem Stand der Drucktechnik bietet eine Auflösung von 400 dpi reichlich Reserven. Für eine Abbildung in DIN A4 benötigt man eine Datei von circa 3000 x 4000 Pixeln für ein optimales Ergebnis und 2500 x 3300 Pixeln für ein gutes Druckergebnis. Bei der Farbtiefe reichen 8 Bit, um den optimal möglichen Kontrastumfang eines Druckes auf weißem Papier abzudecken. Diese Bildgrößen entsprechen 12 beziehungsweise 8,5 Megapixel beziehungsweise einer Dateigröße von 36 beziehungsweise 25,5 MByte pro unkomprimierter 8-Bit RGB-Datei.

Farbkanäle und Farbprofil

Es ist sehr sinnvoll, für die gesamte Bildbearbeitung und Archvierung einen RGB-Workflow beizubehalten und die Konvertierung in CMYK dem Druckertreiber, dem RIP-Programm oder der Druckerei zu überlassen. Sehr sinnvoll ist es natürlich, ein ICC-Profil in die Datei einzubetten. Als Standard bieten sich hier Adobe-RGB (von vielen Digitalkameras als Ausgabeprofil unterstützt) oder ECI-RGB (Standard im grafischen Bereich) an. Für die Ausarbeitung im Fotolabor sollte man die Dateien allerdings in das smallRGB-Profil konvertieren.

3. Speicherung und Back-up

Ein wesentlicher Aspekt bei der Auswahl der zu archivierenden Dateien ist die Dokumentation der verwendeten Formate und die Validierung der Dateien gemäß der standardisierten Spezifikationen. Ein dazu geeignetes Tool stellt etwa JHOVE dar. Zum Testen ist die von Harvard University entwickelte Version 1.x am besten geeignet, da sie auch über ein Testtool mit GUI verfügt (http://sourceforge.net/projects/jhove/files/jhove/JHOVE%201.6/jhove-1_6.zip/download).

Im Moment wird von verschiedenen Organisationen an JHOVE 2.x gearbeitet (https://bitbucket.org/jhove2/main/wiki/Home):

„Prinzipiell kann man nie genug über eine archivierte Datei wissen, jedoch kann es durchaus sinnvoll sein, extrahierte Metadaten einmal auf ihre Qualität zu überprüfen und gegebenenfalls für die Langzeitarchivierung nur indirekt relevante Daten herauszufiltern, um das Archivierungssystem nicht mit unnötigen Daten zu belasten. Beispiel für ein solches Tool ist „JHOVE" (das JSTOR/Harvard Object Validation Environment der Harvard University Library, http://hul.harvard.edu/jhove/), mit dem sich auch Formaterkennung und Validierung durchführen lassen. Das Tool ist in Java geschrieben und lässt sich auch als Programmier-Bibliothek in eigene Anwendungen einbinden. Die generierten technischen Metadaten lassen sich sowohl in Standard-Textform als auch in XML mit definiertem XML-Schema ausgeben.

Validierungstools für Dateiformate stellen sicher, dass eine Datei, welche in einem fraglichen Format vorliegt, dessen Spezifikation auch vollkommen entspricht. Dies ist eine wichtige Voraussetzung für die Archivierung und die spätere Verwertung, Anwendung und Migration beziehungsweise Emulation dieser Datei. Das bereits erwähnte Tool „JHOVE" kann in der aktuellen Version 1.1e die ihm bekannten Dateiformate validieren; verlässliche Validatoren existieren aber nicht für alle Dateiformate. Weitverbreitet und gut nutzbar sind beispielsweise XML Validatoren, die auch in XML Editoren wie „oXygen" (SyncRO Soft Ltd., http://www.oxygenxml.com) oder „XMLSpy" (Altova GmbH, http://www.altova.com/XMLSpy) integriert sein können."
(Nestor Handbuch Ver. 2, Kap. 7/S. 16,17).

Eine sehr viel gröbere Erkennung bietet DROID auf Basis der PRONOM Format Registry der National Archives of Great Britain (http://sourceforge.net/projects/droid/files/droid/ bzw. http://www.nationalarchives.gov.uk/PRONOM/Default.aspx).

Back-ups sollten grundsätzlich auf unterschiedlichen Medientypen erstellt werden. Denkbar wäre etwa ein Festplattensystem mit geeignetem Schreibschutz auf der einen Seite und die Ablage auf WORM-Bändern (LTO5, 1,5 TByte Speicherplatz mit einer geschätzten Mindestlebensdauer von 50 Jahren) auf der anderen Seite. Dabei ist natürlich auch auf eine geeignete Speicherung und Archivierung der Informationen zu den verwandten Datentypen und aller Metadaten zu achten (Stand 10/2011).

Eine andere sehr gute Möglichkeit, wichtige Aufnahmen risikofrei für die nächsten 50–100 Jahre aufzubewahren, stellt das Ausbelichten auf Fotopapier und Aufbewahren in altmodischen Archivschränken dar. Auch die immer wieder diskutierte Ausbelichtung auf Dia- oder Mikrofilm könnte eine sinnvolle Back-up-Möglichkeit mit heute noch vertrauter Technik sein.

Anmerkung

[1] EV (= exposure value, also etwa Belichtungswert in der wörtlichen Übersetzung, die korrekte deutsche Bezeichnung ist Lichtwert [LW]). Es handelt sich dabei um einen logarithmischen Wert (Basis 2), gebildet aus dem Quadrat der Blendenzahl, dividiert durch die Belichtungszeit, und steht für verschiedene Blenden/Zeit-Kombinationen mit der Eigenschaft, dass gleich viel Licht auf Film oder Sensor treffen. Die Differenz zwischen zwei aufeinanderfolgenden EVs entspricht immer genau einer Blendenstufe oder der Verdopplung bzw. Halbierung der Belichtungszeit.
Beispiel: (Film/Sensorempfindlichkeit jeweils 100 ASA)
EV 0 entspricht einer Belichtung von 1 sec bei Blende 1
EV 1 -> t=0,5 sec / f=1 oder t=1 sec / f=1,4 … t=8 sec/ f=4 …
EV 2 -> t=1/4 sec / f= 1 oder t=1/2 sec / f = 1,4 oder t=1 sec / f= 2
…
EV -1 -> t=2 sec / f=1
Der Lichtwert ist also ein Maß für die korrekte Belichtung einer bestimmten Motivsituation und kann mit einem Belichtungsmesser bestimmt werden, es sind auch negative Werte möglich.

Metadaten in der Archäologie

Evelyn Rauchegger, Utz Böhner

Die Arbeit der archäologischen Denkmalpflege – wie der Denkmalpflege überhaupt – kann in Maßnahmen oder Aktivitäten gegliedert werden. In Bezug auf den Denkmalbestand empfiehlt es sich, von Denkmalvorgängen zu sprechen. Im Idealfall bilden Denkmalvorgänge den Lebenslauf eines Denkmalobjektes ab, von der Entdeckung bis zur Ausgrabung oder unkontrollierten Zerstörung. Je nach Denkmalvorgang kann es sich bei den Dokumenten zum Beispiel um die Grabungsdokumentationen, eine Zustandsbeschreibung eines obertägig erhaltenen Denkmals, Studio-Fotoaufnahmen von Funden, digitale 3-D-Modelle oder die Ergebnisse naturwissenschaftlicher Analysen handeln.

Während die wissenschaftliche Literatur in Bibliotheken formal und inhaltlich mit Metadaten erschlossen wird und durch Metadatensuchmaschinen recherchierbar ist, verbleiben die archäologischen Dokumente in der Regel in Sammlungen (Ortsakten etc.) und werden kaum durch Metadaten beschrieben. Gerade sie enthalten aber viele primäre Forschungsdaten. Hinzu kommt, dass durch das Fehlen einheitlicher Dokumentationsstandards in der Archäologie die Metadaten zu den Dokumenten nur eingeschränkt vergleichbar sind. Die formale und inhaltliche Vielfalt der Dokumente ist beträchtlich. Sie liegen als digitale Dateien oder analoge Dokumente vor. Es existieren zahlreiche Dateiformate für digitale Dateien und verschiedene physische Trägermaterialien für nicht-digitale Dokumente.

Metadaten sind erforderlich, um diese Vielfalt an Dokumenten zu ordnen, zu gruppieren und als Materialbestand gezielt überblicken und durchsuchen zu können. Aber was genau sind Metadaten?

Metadaten sind strukturierte Beschreibungen von Informationsobjekten. Diese Beschreibungen werden über verschiedene digitale Informationssysteme hinweg nutzbar, indem durch Standardisierung die Struktur und Semantik der Metadaten in den Systemen angeglichen wird. Standardisierung betrifft alle Struktur- und Inhaltsaspekte wie Schemata, Modelle, Formate, Regelwerke oder Thesauri. Metadatenschemata funktionieren vereinfacht dargestellt wie Muster, an denen sich alle individuellen Objektbeschreibungen orientieren. Vorgegeben sind die Datenfelder, die bestimmte Eigenschaften eines Objekts abbilden sollen sowie strukturelle Vorgaben oder Vokabulare, die zu diesem Zweck eingesetzt werden dürfen. Im konkreten Kontext der Denkmalvorgänge sind alle Dokumente, die aus einem Denkmalvorgang entstehen, solche Informationsobjekte und können mit Metadaten beschrieben werden. Metadaten im Allgemeinen können jedoch auch strukturierte Beschreibungen von Objekten sein, die keine Dokumente sind. Abhängig vom Informationsbedarf können völlig verschiedene physische oder digitale Objekte oder sogar Lebewesen als Informationsobjekte betrachtet und mit Metadaten beschrieben werden.

Häufig werden Metadaten in Typen unterteilt, je nachdem welche Aspekte eines Informationsobjekts sie beschreiben. Wie viele Metadatentypen dabei klassifiziert werden und welche Informationselemente in ihnen zusammengefasst werden, ist je nach Definition unterschiedlich.

Administrative oder verwaltende Metadaten beschreiben zum Beispiel die Provenienz und den Aufbewahrungsort eines Informationsobjekts. Informationen zu Nutzungsbedingungen und Urheberrechten können ebenfalls zu den administrativen Metadaten gezählt oder zu einem eigenen Block zusammengefasst werden. Dieser Datenblock wird in der Regel als *rights metadata* oder *use metadata*[1] bezeichnet, da er Metadatenelemente mit Informationen zur Benutzung und Zugänglichkeit von Informationsobjekten enthält. Ähnlich verhält es sich mit Metadaten zur Archivierung beziehungweise digitalen Langzeitarchivierung, die Aussagen über den technischen Entstehungskontext (Software und Hardware) eines digitalen Dokuments treffen oder Datenmigrationen dokumentieren, die aber auch den Erhaltungszustand oder bestimmte Lagerungsbedingungen für physische Archivalien beschreiben können. Diese Angaben können entweder den administrativen Metadaten zugeschlagen werden oder als eigenständiger Block *preservation metadata* gehandhabt werden.

Deskriptive oder beschreibende Metadaten geben durch eine sachliche und formale Erschließung den Inhalt oder das Thema und die Form eines Informationsobjekts wieder[2]. Sie ermöglichen seine Auffindbarkeit und Identifizierung.

Deskriptive Metadaten sind deshalb in der Regel auch diejenigen Metadaten, mit denen die Nutzer von

Informationssystemen bei der Recherche arbeiten, während administrative Metadaten und ihre Subgruppen eher der Verwaltung und Sicherung von Beständen dienen.

Technische Metadaten sind eine besondere Kategorie von Metadaten: sie beschreiben in erster Linie die formalen Eigenschaften digitaler Dateien wie Formate oder Komprimierungsalgorithmen, die für die Benutzung und die Langzeitarchivierung von Belang sind.

Generell ist die Zuordnung konkreter Datenfelder zu den beschriebenen Kategorien durchaus nicht immer eindeutig und wird in der Forschung verschieden interpretiert. Die Ursache dafür liegt in der Multifunktionalität der Metadatenfelder in der Praxis. Deskriptive Metadaten, die sich auf technische Eigenschaften eines Dokuments beziehen, sind nicht nur für die Recherche im Informationssystem oder für die Benutzung interessant, sondern beispielsweise auch für die Archivierung. Grundlegend ist die Unterscheidung in deskriptive und administrative Metadaten, wobei die deskriptiven Metadaten eher nutzer- und rechercheorientiert sind und die administrativen eher der Steuerung von Archivprozessen wie der Benutzung oder Sicherung dienen.

Ein kleines Beispiel verdeutlicht, wie Metadaten unsere Informationswelt im Alltag prägen, ohne dass wir uns dessen immer bewusst sind. So handelt es sich bei den Informationen auf einer Lebensmittelverpackung im Prinzip um Metadaten. Sie beschreiben, welche Zutaten in der Verpackung enthalten sind und bis wann ein Produkt verzehrt werden darf. Daher weiß ein Verbraucher im Ladengeschäft, welche Inhaltsstoffe ein bestimmtes Lebensmittel enthält, ohne die Verpackung zu öffnen und nachzusehen. Oder er kann ein nicht mehr genießbares Produkt anhand des Mindesthaltbarkeitsdatums identifizieren.

Metadaten in der Archäologie betreffen natürlich auch andere Objekte außer den Dokumenten, die hier im Vordergrund stehen. Auch in einem archäologischen Fundmagazin ist es sinnvoll, die Verpackungseinheiten, die Funde mit unterschiedlichen Fundnummern enthalten, mit Metadaten zu versehen. So ist es leichter, sich den Inhalt einer Kiste im Magazin anzeigen zu lassen. Auch muss bei Veränderungen des Standorts nicht jeder Fund einzeln verbucht werden. Informationsobjekt, das heißt, Referenzobjekt für die Metadatenbeschreibung kann sowohl ein individueller Fund als auch eine Verpackungseinheit sein, die mehrere Funde enthält. Die jeweiligen Referenzobjekte müssen bei der Metadatenauszeichnung aber streng getrennt werden. Das Vermengen von Beschreibungsebenen kann nicht nur in einem Fundmagazin unangenehm werden. Das gilt insbesondere für Metadatenformate, die ineinander geschachtelte Beschreibungsebenen besitzen. Der Unterschied zwischen einem Digitalisat und dessen Quelldokument oder zwischen dem Fund und seiner Verpackungseinheit muss stets vergegenwärtigt werden.

Metadaten sollen helfen, Informationsobjekte zu finden und Entscheidungen zu ermöglichen, die ohne strukturierte Beschreibungen wesentlich schwieriger wären. Ein Abstract oder eine inhaltliche Zusammenfassung eines Dokuments hilft bei der Entscheidung darüber, ob ein Dokument für eine bestimmte Fragestellung relevant ist, ohne dass gleich das ganze Dokument gelesen werden muss. Formale Angaben dienen unter anderem dazu, Zugangsvoraussetzungen wie bestimmte Hard- oder Software zu verdeutlichen.

Metadaten dienen aber nicht nur dem *Auffinden* von Informationsobjekten, sondern auch der *Verwaltung* von Archivbeständen, ganz ähnlich wie bei dem Bestandskatalog einer Bibliothek. Administrative Metadaten helfen zum Beispiel, Statistiken über den Bestandsumfang und die Bestandsentwicklung zu führen. Diese Metadaten enthalten beispielsweise für jeden Metadatensatz ein Erfassungs- und Änderungsdatum. Auch rechtliche Aspekte wie Urheber- und Nutzungsrechte für einzelne Dokumente oder spezielle institutionelle Informationsbedürfnisse lassen sich in Metadaten unterbringen. Je nach Wissensgebiet und abhängig von der Beschaffenheit der zu beschreibenden Informationsobjekte existiert eine Vielzahl von sogenannten Metadatenstandards. Das sind Formate und Schemata aber auch relativ abstrakte Datenmodelle oder sehr fachspezifische Thesauri oder Normdaten.

Eine sehr übersichtliche Darstellung bietet die Visualisierung der Metadatenwelt von Riley und Becker[3].

Am Beispiel des Niedersächsischen Landesamts für Denkmalpflege konnten folgende inhaltliche Kategorien für Archivdokumente ermittelt werden, die einen Überblick über das weite Spektrum an Inhalten und auch an Formaten geben:

1. Archäologische Archivalien
    - Begehung
    - Fund-Dokumentation
    - Grabungsdokumentation
        - Ausgrabungsbericht
        - Schnitt-Dokumentation
        - Planum-Dokumentation
        - Profil-Dokumentation
        - Befund-Dokumentation
    - Gelände-Dokumentation

2. Restaurierung
    - Restaurierungs-Dokumentation

3. Naturwissenschaftliche Untersuchungen
    - anthropologische Analyse
    - botanische Analyse
    - faunistische Analyse
    - bodenkundliche Analyse
    - materialkundliche Analyse
    - Röntgenaufnahme; Computertomographie
    - sonstige naturwissenschaftliche Analyse
    - dendrochronologische Datierung
    - Radiokarbon-Datierung
    - sonstige Datierungsverfahren, zum Beispiel Thermolumineszenz-Datierung (TL), Optisch stimulierte Lumineszenz-Datierung (OSL) und Uran-Thorium-Datierung (U/Th)

4. Allgemeine Archivalien
    - Akte (Ortsakten, Objektakten, Sonderakten) [Sammlung]
    - denkmalpflegerische Stellungnahme oder Gutachten
    - geophysikalische Prospektion
    - Luftaufnahme, Satellitenbild, Laserdaten
    - Ausstellungsmaterial (Poster, Flyer usw.)
    - Pressematerial (Zeitungsausschnitte, Pressemitteilungen usw.)
    - Brief

Erforderlich war eine derartige Analyse und Klassifikation der vorhandenen Archivbestände, da für das denkmalpflegerische Archivgut insgesamt keine geeignete Klassifikation existierte, mit der sich die Dokumente in wenigen Gruppen hätten bündeln lassen. Für eine grobe Inhaltserschließung mussten aber Gruppen gebildet und mit Begriffen versehen werden, um später über diese Klassifikation auch gezielt recherchierbar zu werden.

Neben dem archivinternen Informations- und Bestandsmanagement erlauben digitale Metadaten aber auch den Datenaustausch mit Kooperationspartnern oder die Aggregation von Metadaten verschiedener Datenlieferanten in einem gemeinsamen Portal. Voraussetzung dafür ist die *Interoperabilität* der Metadaten. Da von einer Anwendung zur anderen oft verschiedene Metadatenschemata und -formate im Einsatz sind, sind die Anlehnung an übergreifende Standards und eine gute Dokumentation sehr wichtig, um eine qualitativ hochwertige Übertragung von einem Metadatenformat zum anderen zu erlauben (Mapping).

Metadaten sollen die Bedingungen für den Zugang zu Dokumenten verbessern sowie die Nutzung und Verwaltung vorhandener Bestände optimieren. Digitale Fachinformationssysteme, die in der Denkmalpflege stetig an Bedeutung gewinnen, bieten die Möglichkeit, über Metadaten traditionelle, nicht digitale Archivbestände gemeinsam mit digitalen Beständen zu erschließen und mit weiteren fachlichen Informationen zu verknüpfen, und das im Prinzip standortunabhängig. Insbesondere Lagedaten sind von hoher Bedeutung. Die meisten Dokumente weisen direkt oder indirekt aus einer Zuweisung zu einer Fundstelle oder einem Befund Lagedaten auf. In der Forschung und in Planungsvorhaben werden oft Daten über die Lage selektiert. Anhand von Lagekriterien können Archivdokumente somit vergleichsweise einfach zusammen mit den fachlichen Objekten der Forschung beziehungsweise Planung aufgefunden werden.

In denkmalpflegerischen Fachinformationssystemen werden Informationen zu fachlichen Objekten gesammelt. Metadaten zu Archivdokumenten machen bisher nur einen vergleichsweise kleinen Teil der Informationen in diesen Systemen aus. Aus archäologischen Ausgrabungen resultieren Funde als Ergebnis. Diese werden neben den Dokumenten (zum Beispiel Fotos, Befundpläne, Profile, Berichte) in der Regel als eigenständige Objekte geführt. Um Metadaten verschiedener Objekte optimal gemeinsam nutzen zu können, ist ein möglichst ähnliches Metadatenschema im Hintergrund vorteilhaft. Funde werden als besondere Objektart in der Archäologie angesehen. Bei genauerer Betrachtung weisen Funde aber Ähnlichkeiten bezüglich ihrer deskriptiven Kernmetadaten zu Dokumenten als Informationsobjekten auf. Funde sind physische Objekte, wie es auch Dokumente sein können. Ein Plan besteht ebenso aus einem Material wie ein Fund. Auch ist es möglich, einen Plan wie einen Fund fotografisch abzubilden. Funde können ebenso wie Dokumente Bestandteil einer Sammlung (mehrere Funde je Fundeinheit beziehungsweise mehrere Dokumente je Akte) sein. Es können Metadaten

für einen einzelnen Fund beziehungsweise ein Dokument wie auch für Sammlungen geführt werden. Allerdings gibt es wichtige Unterschiede, die dennoch in einem gemeinsamen Schema untergebracht werden können. Der Finder beziehungsweise Erzeuger eines Fundes ist semantisch nicht identisch mit dem, was viele Metadatenschemata als creator bezeichnen würden. Angaben zu Finder und Funddatum sind aber bei Funden von herausgehobener Bedeutung. Diese Angaben sind als Provenienzinformationen zu verstehen, die theoretisch auch bei Dokumenten vorkommen können, die dort aber in der Regel weniger bedeutsam sind. Es sei denn, es handelt sich zum Beispiel um Nachlässe, wenn etwa der Nachlasser einer Kartierung nicht zugleich der Verfasser ebendieser war. So kann die Archivgeschichte eines Dokuments oder eines Fundes rekonstruiert werden.

Ein Metadatenformat, das mit einem komplexen Datenmodell wie CIDOC-CRM kompatibel ist, erleichtert die Aggregation verschiedener Metadaten unterschiedlicher Objekte mit jeweils variierenden Beschreibungsanforderungen, indem es logische Beziehungen zwischen Objekten, Objektklassen und Eigenschaften zur Verfügung stellt. Die Überlegung, welche Metadaten für Dokumente sinnvoll eingesetzt werden können, steht also nicht isoliert im Raum, sondern es sollen immer Verknüpfungspunkte zu anderen Objekten in denkmalpflegerischen Fachinformationssystemen im Blick bleiben.

Literatur

Gilliland, Anne J. (2008): Setting the Stage. In: Baca, Murtha (Hrsg.): Introduction to metadata. 2. ed. Los Angeles, Calif.: Getty Research Institute.

Voß, Jakob: s.v. Metadaten. In: Umlauf, Konrad; Gradmann, Stefan (Hrsg.) (2012): Lexikon der Bibliotheks- und Informationswissenschaft: LBI. Bd. 2 Lfg. 2, Stuttgart: Hiersemann.

Anmerkungen

[1] Gilliland, Anne J. (2008): Setting the Stage. In: Baca, Murtha (Hrsg.): Introduction to metadata. 2. ed. Los Angeles, Calif.: Getty Research Institute. S. 9.

[2] Voß, Jakob: s.v. Metadaten. In: Umlauf, Konrad; Gradmann, Stefan (Hrsg.) (2012): Lexikon der Bibliotheks- und Informationswissenschaft: LBI. Bd. 2 Lfg. 2, Stuttgart: Hiersemann, S. 602.

[3] Riley, Jenn; Becker, Devin: Seeing Standards: A Visualization of the Metadata Universe. c2009–2010. online verfügbar unter: http://www.dlib.indiana.-edu/~jenlrile/metadatamap/ [Zuletzt geprüft am 02.04.2013].

Kleiner Ratgeber zur Archivierung digitaler Daten

Reiner Göldner, David Bibby, Andreas Brunn, Sigmar Fitting, Axel Posluschny, Bettina Stoll-Tucker

Die AG Archivierung der Kommission „Archäologie und Informationssysteme“ im Verband der Landesarchäologen in der Bundesrepublik Deutschland verfolgt das Ziel, allgemein verfügbare methodische und technische Ansätze zur Archivierung digitaler archäologischer Daten zu begutachten sowie praxisorientierte Empfehlungen, Vorschläge und Tipps zu geben. Die Ergebnisse dieser Arbeit münden in einen Ratgeber, der hier in Ausschnitten dargestellt wird.[1]

1 Einleitung

„Digital“ ist Trend. Digitale Helfer findet man heute für nahezu alle Lebenslagen. Leistungsfähige und preiswerte Systeme, die sich unkompliziert und intuitiv bedienen lassen, finden große Verbreitung. Allerorts wird gescannt und digitalisiert: Bilder, Bücher, Karten, Landschaften, Räume, Museumsobjekte. Und wer fotografiert heutzutage noch mit Film? Auch im Bereich der Archäologie werden herkömmliche Methoden immer häufiger durch digitale Methoden abgelöst. Berichte, Pläne, Fotos usw. existieren oft nur noch als Datei. Viele dieser digitalen Daten sind es aufgrund ihrer Einzigartigkeit und hervorragenden Bedeutung wert, langfristig oder sogar dauerhaft bewahrt – archiviert – zu werden.

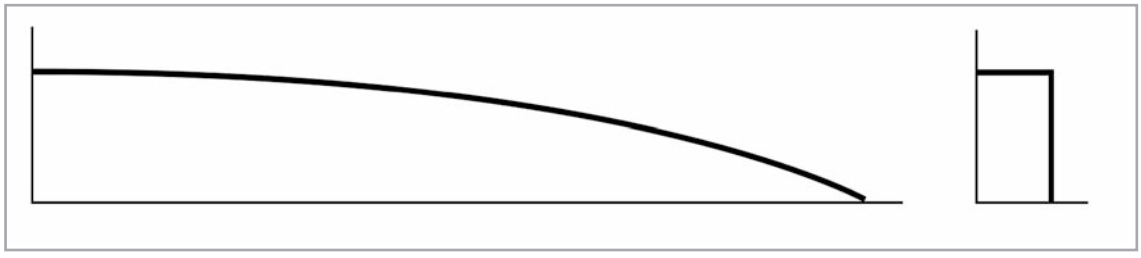
Nun ist allerdings die Informationstechnik heutzutage im Normalfall recht wenig auf Dauerhaftigkeit ausgerichtet. Ganz im Gegenteil bestimmen kürzeste Innovationszyklen das Geschehen und die neue Technik nimmt nur selten Rücksicht auf alte Daten. Wer kann schon heute noch Disketten lesen oder MODs? Wie lange werden CDs noch unterstützt werden? Niemand erwartet noch ernsthaft, Daten lesen zu können, die vor drei oder mehr Systemgenerationen geschrieben wurden. Wie soll man da etwas dauerhaft bewahren? Es ist zwar zu erwarten, dass die Archivierung digitaler Daten in der nächsten Zeit auch als Standard-IT-Thema stärker in den Vordergrund tritt und hoffentlich bald auch preiswerte kommerzielle Lösungen angeboten werden. Bis dahin bleibt allerdings nur, sich mit den Problemen vertraut zu machen und, Schritt für Schritt, nach individuellen Lösungen zu suchen. Mit Kenntnis einiger grundlegender Methoden kann man die Richtung vorgeben und erste Schritte gehen, ohne Datenverluste zu riskieren und auch ohne zukünftig möglichen Archivsystemen vorzugreifen.

Warum lassen sich digitale Daten eigentlich so schlecht aufheben? Zwei wesentliche Problembereiche lassen sich dafür identifizieren: Komplexität und Alterung.

Herkömmliche, analoge Dokumente sind in der Regel direkt und mit bloßem Auge lesbar und erkennbar. Dieses einfache Prinzip hat zu einer weitreichenden Verbreitung geführt und gewährleistet gleichzeitig, dass man Dokumente aus lang zurückliegender Zeit heute noch lesen kann, sofern das Medium erhalten geblieben ist.

Im digitalen Bereich finden wir eine andere Situation vor. Hier erschließt sich der Inhalt oder die Funktionalität nur über ein komplexes System aus Hardware, Software, Datenträger und Datenformat. Sobald sich in diesem komplexen System nach der Datenaufzeichnung etwas ändert, kann man nicht mehr gewiss sein, auf den ursprünglichen Inhalt bzw. die ursprüngliche Funktionalität zurückgreifen zu können. Ohne eine passende Systemumgebung ist der Datenträger mit all seinen Daten wertlos.

1 Komplexität eines analogen Systems (oben) und eines digitalen Systems (unten).



2 Alterungsprozess: links analog (kontinuierlich, langlebig) und rechts digital (abrupt, kurzlebig).

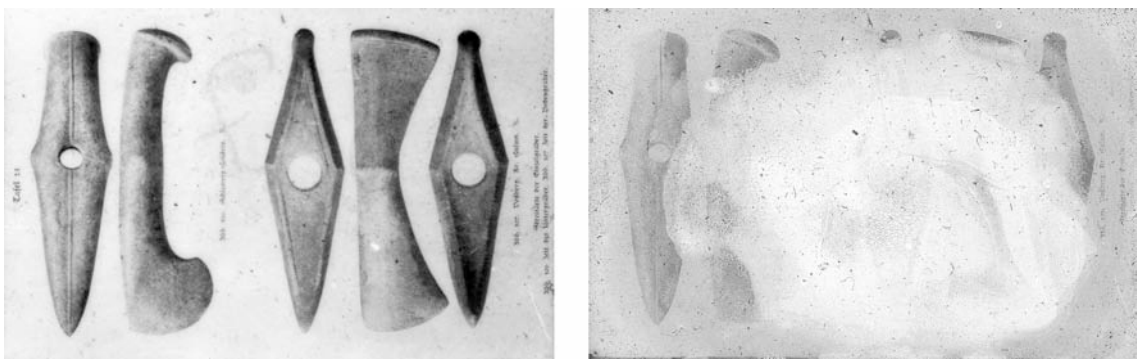
Die Alterung betrifft analoge Dokumente ebenso wie den digitalen Bereich. Allerdings laufen Alterungsprozesse bei analogen Dokumenten meist langsam und kontinuierlich ab, sodass man sie gut vorhersehen und Vorsorge treffen kann. Im digitalen Bereich sind Alterungsprozesse dagegen schnell und abrupt, was im Zusammenhang mit der Systemkomplexität eine Vorhersage äußerst schwierig macht. Das liegt daran, dass digitale Informationstechnologien auf enormen Speicherdichten basieren und nur wenig Platz für Redundanzen lassen. Digitale Daten sind also äußerst fehleranfällig. Hinzu kommen äußerst kurze technische Innovationszyklen von nur wenigen Jahren, wobei Nachhaltigkeit und Bewahrung bisher keine wesentliche Rolle spielen, meist zählt nur das Tagesgeschäft.

Digitale Daten haben aber auch einen entscheidenden Vorteil, sie lassen sich nämlich ohne Informationsverlust kopieren und damit immer wieder erneuern. Dies ist bei analogen Medien nicht möglich. Ein gealtertes, verblasstes Foto hat unwiederbringlich Informationen verloren, und man muss mit gegebenenfalls äußerst aufwendigen Methoden für den Erhalt wenigstens der verbliebenen Informationen sorgen, wenn schon das Medium nicht zu retten ist. Dazu bedient man sich dann übrigens gern der Digitalisierung.

Der Verband der Landesarchäologen erkannte die bestehenden Probleme und regte auf dem Kolloquium des Jahres 2008 die Gründung der AG „Archivierung" innerhalb der Kommission „Archäologie und Informationssysteme" an. Diese AG hat sich zum Ziel gesetzt, allgemein verfügbare methodische und technische Ansätze zur Archivierung digitaler Daten zu begutachten sowie praxisorientierte Empfehlungen, Vorschläge und Tipps zu geben. Durch eine Analyse der Situation in den archäologischen Landesbehörden Deutschlands wurden wichtige Themenbereiche mit dringendem Handlungsbedarf identifiziert. Mit dem vorliegenden Ratgeber soll diesen Zielen im Rahmen der bestehenden Möglichkeiten entsprochen werden.

2 Grundsätze

Archivierung ist nicht immer gleich Archivierung. Der Begriff „Archivierung" wird im traditionellen Archivwesen anders gebraucht, als es im modernen IT-Bereich oft üblich ist. Während „Archivierung" traditionell mit Dauerhaftigkeit einhergeht, gibt man sich im IT-Bereich oft schon mit einigen Jahren Haltbarkeit zufrieden. In diesem Ratgeber wird jedoch auch hinsichtlich der digitalen Daten ausdrücklich die archivarische Sichtweise vertreten, die von einer dauerhaften, also unbegrenzten Erhaltung ausgeht. Im Vergleich zur herkömmlichen Archivierung geht es jedoch nicht nur um Dokumente, sondern um gegebenenfalls komplex strukturierte digitale Inhalte oder sogar um Funktionalitäten.



3 Alterungsprozess analoger Medien: links Kontaktabzug auf Papier (Xerokopie, 1990er Jahre) und rechts Diapositiv (2011).
© Landesamt für Archäologie Dresden.

Das Ziel der digitalen Archivierung ist es, digitale Inhalte oder digitale Funktionalitäten dauerhaft zu erhalten und verfügbar zu machen.

Einige wichtige Anforderungen sollen hier vorab schon einmal erwähnt werden. Der tägliche Grundschutz der IT-Systeme und Daten soll einerseits Vertraulichkeit und Integrität sichern und andererseits für Verfügbarkeit sorgen, also dem Verlust vorbeugen. Die Gefährdungen sind vielfältig und entsprechend umfangreich die zu treffenden Gegenmaßnahmen. Passwörter, Virenschutz und Back-ups seien hier nur beispielhaft genannt, ausführliche Informationen erhält man auf den Seiten des Bundesamtes für Sicherheit in der Informationstechnik (BSI). Die Umsetzung des IT-Grundschutzes ist eine unbedingte Grundlage der Archivierung und erfordert umfangreiche Maßnahmen mit erheblichem Aufwand, deren Umsetzung hier allerdings als gegeben vorausgesetzt wird. In diesem Ratgeber werden also im Wesentlichen nur die Themen angeschnitten, die über die tägliche IT-Grundsicherung hinausgehen.

Entsprechend der auch in Deutschland gültigen ISO 15489 „Information and Documentation – Records Management" ist die Bewahrung der Authentizität (und damit einhergehend die Bewahrung von Zuverlässigkeit, Integrität und Brauchbarkeit) wichtigste Voraussetzung für die dauerhafte Bewahrung von Dokumenten (TNA1, S. 3). Dabei wird unter Authentizität nicht die vollständige Unversehrtheit der Originaldaten verstanden, sondern „nur" die Unversehrtheit der Informationen aus inhaltlicher Sicht und im Sinne der ursprünglichen Zweckbestimmung. In der DIN 31645 (S. 6) findet man dazu: „Authentizität beinhaltet, dass nur beabsichtigte, dokumentierte Veränderungen im Sinne der Erhaltungsmaßnahmen am Objekt durchgeführt wurden." Des Weiteren wird darauf verwiesen, dass Archiv und/oder Produzent entscheiden müssen, welche signifikanten Eigenschaften der zu archivierenden Informationsobjekte erhalten bleiben müssen und welche nicht. Dabei sind die Bedürfnisse der Nutzer der hauptsächliche Maßstab (DIN 31645, S. 12).

Eine Besonderheit der digitalen gegenüber der herkömmlichen Archivierung liegt darin, dass die Dokumente nicht dauerhaft im Original erhalten werden können. Die deshalb erforderlichen Erhaltungsmaßnahmen dürfen jedoch die Authentizität der Dokumente nicht beeinträchtigen.

Eine herausragende Bedeutung für das Verständnis digitaler Archivierung leistet OAIS. OAIS bedeutet „Open Archival Information System". Es ist ein Referenzmodell für ein digitales Archiv und bildet den Standard ISO 14721. Als konsequent logisches Modell ist es unabhängig von jeder Implementierung. Es leistet einen wesentlichen Beitrag zum (einheitlichen) Verständnis digitaler Archivierung.

Auf den Internetseiten der Schweizer Nationalbibliothek (NB) findet man eine gute Kurzbeschreibung zu OAIS, darin heißt es unter anderem: „Das als ISO 14721 verabschiedete Referenzmodell beschreibt ein Archiv als Organisation, in dem Menschen und Systeme mit der Aufgabenstellung zusammenwirken, Informationen zu erhalten und einer definierten Nutzerschaft verfügbar zu machen. Das Modell beschreibt im Detail, wie die von einem Produzenten hergestellte elektronische Information in ein Archivsystem gelangen soll, welche Bearbeitungsschritte für die langfristige Archivierung vorgenommen werden müssen und wie auf die im Archiv gespeicherte Information zugegriffen werden kann." Auch das Schweizer Nationalarchiv hat langjährige Erfahrungen mit digitaler Archivierung und OAIS. Digital erstellte Unterlagen der Schweizer Bundesverwaltung werden dort seit 2012 ausschließlich digital archiviert.

OAIS beschreibt ein digitales Archiv als Organisationseinheit von Mensch und technischen Systemen, das dazu dient, digitale Informationen dauerhaft zu erhalten und verfügbar zu machen. Es wird mittlerweile weltweit als Referenzmodell und Standard akzeptiert und angewandt.

3 Organisation der Archivierung

Bei der Beschreibung der Methoden zur Archivierung digitaler Daten wurde schon deutlich, dass im digitalen Bereich sehr viel stärker als im herkömmlichen analogen Bereich gilt, dass Archivierung eine regelmäßige Betreuung erfordert. Diese Betreuung ist eine komplexe Aufgabenstellung und kann nicht nebenbei erledigt werden. Im Idealfall ist das digitale Archiv eine betriebliche Organisationseinheit, wie jedes andere Archiv auch (Aktenarchiv, Bildarchiv, Kartenarchiv, Bibliothek, Fundarchiv ...). Es erfordert starke Kompetenzen aus den Bereichen Informationstechnik und Archivierung sowie entsprechende personelle und finanzielle Ressourcen.

Digitale Archivierung darf nicht verwechselt werden mit dem Thema Datensicherheit. Während Letzteres zum IT-Tagesgeschäft gehört und nur der unmittelbaren Verfügbarkeit von Daten dient, gehen die Aufgaben der digitalen Archivierung weit über das Repertoire des IT-Ressorts hinaus und können von diesem nicht allein getragen werden!

Komplexe Aufgaben benötigen eine sorgfältige Vorbereitung und Planung. Folgende Punkte spielen dabei eine wichtige Rolle:

- die Verantwortlichkeit für das digitale Archiv muss festgelegt werden,
- es müssen finanzielle und personelle Ressourcen zugewiesen werden,
- es muss formell bestimmt und beschrieben werden, was archiviert werden soll,
- der technische Rahmen mit Erhaltungsmethoden, Datenformaten usw. muss abgesteckt werden,
- es sind Abläufe und Termine zu organisieren und
- die Ergebnisse sind einem Qualitätsmanagement zu unterziehen.

Abseits der großen Archive, die Archivierung als Kernaufgabe betreiben und damit auch der Archivierung digitaler Daten schon erhebliche Aufmerksamkeit widmen, gibt es eine ganze Anzahl von Institutionen, die aus unterschiedlichen Gründen ihre digitalen Unterlagen und Dokumente selbst archivieren wollen oder müssen. Typisch dafür sind Museen oder Fachämter, deren Dokumentation zunehmend digital erfolgt. Hier werden unwiederbringliche Informationen erarbeitet, die zweifellos archiviert werden müssen, nur kommt eine Auslagerung in ein „großes" Archiv in der Regel nicht in Frage, weil der Informationsbestand „lebendig" ist, sich also in steter Bearbeitung und Benutzung befindet.

Für solche kleineren Archive sind Archivierungslösungen erforderlich, die als kleine Bausteine und möglichst ohne große Kosten in die tägliche Arbeit einbezogen werden können. Solche Bausteine lassen sich an bestimmten Stellen aus den „großen" Konzepten und Pilotprojekten isolieren. Diese Bausteine ermöglichen es, den Weg zur Archivierung digitaler Daten „Schritt für Schritt" zu gehen.

Ein entscheidender Schritt ist es, die zu verwendende(n) Archivierungsmethode(n) festzulegen. Diese Entscheidung trifft man anhand der Eigenschaften und Funktionen der zu archivierenden Daten. Im Anschluss sind die wesentlichen Parameter dieser Methode(n) einzugrenzen, also zum Beispiel die Datenformate für die Migrationsmethode. Damit ergeben sich die wesentlichen Trittsteine für die weitere Arbeit. In vielen Fällen wird man damit ein gutes Stück vorwärts kommen, ohne gleich ganz großen Aufwand betreiben zu müssen. Man kann zum Beispiel dafür sorgen, dass schon bei der Datenerzeugung beziehungsweise Digitalisierung archivierungstaugliche Formate benutzt werden. Zumindest wächst damit die Menge der problematischen Daten erst einmal nicht mehr (so stark) an. Vermutlich wird aber auf Dauer nur eine umfassende Strategie für die Archivierung erfolgreich sein. Es bleibt zu hoffen, dass die erforderlichen Bausteine demnächst zu akzeptablen Preisen verfügbar sind.

Seit November 2011 liegt die DIN 31645 „Information und Dokumentation – Leitfaden zur Informationsübernahme in digitale Langzeitarchive" vor. Sie gibt „organisatorische Leitlinien für die Übernahme von Informationen aus einem Informationssystem in ein Langzeitarchiv unter Wahrung der Datenintegrität und Authentizität" und ist damit eine wichtige Planungshilfe für die Übernahme von Informationen in ein Digitalarchiv. „Die Informationsübernahme in ein digitales Langzeitarchiv ist nicht nur ein rein technischer Transfer zwischen zwei Systemen, sondern ein Prozess mit vielen organisatorischen Anforderungen, an dessen Ende die Übernahme der Verantwortung durch das digitale Langzeitarchiv steht.

… Ein Langzeitarchiv ist dabei nicht nur ein technisches System, sondern auch eine organisatorische Einheit oder Institution. Ein digitales Langzeitarchiv kann Bestandteil einer größeren Einrichtung sein, die auch analoge Bestände archiviert. Die Zusammenhänge zwischen digitalen und analogen Beständen sollten gewahrt und bei der Nutzung entsprechend dargestellt werden." (DIN 31645, S. 6).

Auch „The National Archives" verweist auf betriebliche Anforderungen (TNA2, Abschnitt 2):

„Um geeignete Strategien für die langzeitliche Vorhaltung ihrer elektronischen Dokumente entwickeln und umsetzen zu können, müssen die entsprechenden Einrichtungen zunächst einen betrieblichen Orientierungsrahmen erarbeiten, der die geschäftlichen Erfordernisse definiert, die jede Gruppe der vorgehaltenen Dokumente zu erfüllen hat. … Die Dokumentenmanagementstrategie muss Handlungsweisen festschreiben … . Sie hat den Geschäftszweck und die technischen Anforderungen zu erläutern und zu definieren, die für den Erhalt der Dokumente in einer Form sorgen können, die diesem Zweck dient."

Für die Archivierung digitaler Daten sind vielfältige Erhaltungsmaßnahmen erforderlich. Damit diese Maßnahmen die Authentizität der archivierten Informationen nicht beeinträchtigen, ist eine sorgfältige Planung erforderlich, die sich auf eine betriebliche Dokumentenmanagementstrategie stützt. Zur Bewältigung der damit verbundenen umfangreichen organisatorischen Anforderungen sind entsprechende betriebliche Ressourcen bereitzustellen.

4 Archivierungsmethoden

Wenn Archivierung gewährleisten soll, dass digitale Inhalte und Funktionalitäten dauerhaft bewahrt und zur Benutzung erhalten werden, müssen also Maßnahmen gegen den Einfluss der Alterung ergriffen werden. Dies ist integrale Aufgabe der Archivierung in analogen ebenso wie in digitalen Archiven. Nur kann man eben bei Papier viele Jahre abwarten, während man bei magnetischen Datenträgern oft schon nach einigen Monaten wieder aktiv werden muss.

Es gibt verschiedene Methoden zum Erhalt der Benutzbarkeit digitaler Daten, die wichtigsten sind:

- Daten werden als Hardcopy ausgedruckt oder auf Film belichtet und herkömmlich archiviert.
- Man versucht, die verwendeten Computersysteme insgesamt möglichst lange zu erhalten.
- Veraltete Systeme werden auf aktuellen Computern virtuell nachgebaut (Emulation).
- Die Daten werden regelmäßig an die aktuellen Systeme angepasst (Migration).

Eine detailliertere Beschreibung dieser Methoden findet man in Band 14 des Archäologischen Nachrichtenblatts (Göldner).

Betrachtet man die vorgestellten Methoden, so kann man feststellen, dass jede für sich ihre Stärken und Schwächen hat, aber allein nicht ausreichend ist. Insgesamt gesehen bieten Sie jedoch ein gutes Instrumentarium, um die Anforderungen der Archivierung digitaler Daten umfassend zu erfüllen. In der folgenden Tabelle werden wichtige Vor- und Nachteile der vorgestellten Methoden noch einmal zusammengefasst.

Die vorgestellten Erhaltungsmethoden zeigen, dass Archivierung eine kontinuierliche Betreuung erfordert. Dies wird in herkömmlichen Archiven nicht angezweifelt und muss auch im digitalen Bereich akzeptiert werden. Der methodische Ansatz macht außerdem deutlich, dass die Wahl von Datenträger und Datenformat nicht von primärer Bedeutung ist. Nicht die Lebensdauer entscheidet über den Erhalt, sondern die Methodik.

5 Datenträger und Datenformate

Vielen Diskussionen zur Archivierung liegt die Idee zugrunde, möglichst langlebige und haltbare Datenträger zu verwenden. Demzufolge werden Haltbarkeitswerte und zeitliche Fehlerraten ausgiebig diskutiert. Dem Trend dieser Diskussionen soll hier nicht widersprochen werden, jedoch sei angemerkt, dass in Abhängigkeit von der Archivierungsmethode auch noch andere Anforderungen als nur Haltbarkeit eine Rolle spielen. Für die Durchführung einer Migration wird es beispielsweise nicht unerheblich sein, dass die Daten möglichst direkt und ohne manuelle Interaktionen zugänglich sind, ein Stapel CDs wäre dabei also kontraproduktiv. Die Migrations-Methode wird demnach durch einen Onlinezugang zu den Daten stark unterstützt und man wird sich eher für ein Festplattensystem oder eine Tape-Library entscheiden. Bei der Erhaltungs-Methode sieht es dagegen völlig anders aus und man wird auch auf eine lange Lebensdauer des Datenträgers hohen Wert legen.

Was bei den Datenträgern anklingt, gilt ebenso für die Datenformate. Man kann die Diskussion über Vor- und Nachteile nicht von den Archivierungsmethoden trennen.

Es ist zwingend erforderlich, dass man sich über die zu benutzende Archivierungsmethode im Klaren ist, bevor man Datenträger und Datenformate auswählt!

Die vorgestellten Methoden schränken die nutzbaren Datenformate nicht ein, dies ist prinzipiell auch nicht erforderlich. Trotzdem ist es vorteilhaft und vermindert den Aufwand, wenn man

	Hardcopy	Erhaltung	Emulation	Migration
einfache Formate	ja	ja	ja	gut
Dokument mit Layout	ja	ja	ja	möglich
komplexe Strukturen	nein	ja	ja	möglich
Funktionalität	nein	ja	ja	nein
Sonstiges	lange Lebensdauer	aufwändig, nicht dauerhaft	aufwändig	gut beherrschbar

4 Vor- und Nachteile der verschiedenen Archivierungsmethoden.

- stabile Datenformate verwendet, die weitverbreitet, standardisiert und nicht proprietär sind sowie
- einfache Datenformate beziehungsweise Datenstrukturen verwendet, die problemarm zu migrieren sind.

Bezüglich der Tauglichkeit zur Archivierung sollen drei Level unterschieden werden:

- bevorzugt: Diese Formate sind sehr unproblematisch und sollten wenn möglich benutzt werden.
- akzeptiert: Diese Formate sind möglicherweise problematisch. Das heißt, sie erfordern gegebenenfalls einen erhöhten Migrationsaufwand. Sie können benutzt werden, wenn die bevorzugten Formate nicht verfügbar oder ungeeignet sind.
- Kandidaten: Diese Formate sind problematisch hinsichtlich einiger weniger Anforderungen. Sie können in begründeten Einzelfällen in Abstimmung mit dem Archiv benutzt werden, wenn andere Formate nicht verfügbar oder ungeeignet sind. Die Formate sollen in ihrer Entwicklung (zum Beispiel Stabilität, Verbreitung) verfolgt werden.

Details zu den Datenformaten sind in Abb. 10–12 aufgelistet.

5.1 Texte

Unter „Texte" sollen Dokumente verstanden werden, die überwiegend Schriftzeichen enthalten. Man könnte also auch „Schrift-Dokumente" sagen. Für die Archivierung ist von Bedeutung, ob die Texte formatiert sind oder nicht. Unformatierte Texte bestehen nur aus den Zeichen eines Zeichensatzes. Im Gegensatz dazu enthalten formatierte Texte eine Vielzahl von Informationen zur Darstellung von Zeichen: fett, kursiv, Schriftgrad usw., zum Layout (Blattrand, Einzug usw.) sowie auch durchaus zeichenfremde Elemente (Grafiken, Abbildungen usw.).

Einfache Texte werden seit Beginn der Computertechnik unterstützt, sind nach wie vor weitverbreitet und können in der Regel auch über Systemgrenzen hinweg gut gelesen und verarbeitet werden. Es ist auch nicht zu erwarten, dass sich an dieser Situation grundsätzlich etwas ändert. Probleme können sich jedoch durch unterschiedliche Kodierungen beziehungsweise Zeichensätze ergeben, vor allem, wenn diese nicht bekannt sind. Gut verbreitete und damit weitgehend unproblematische Zeichensätze sind zum Beispiel:

- US-ASCII (keine deutschen Sonderzeichen!)
- Unicode UTF-8 (oder UTF-16)
- ISO 8859-1(Latin-1 westeuropäisch).

Aus den meisten üblichen Textverarbeitungsprogrammen heraus lassen sich Dokumente im Textformat, auch unter expliziter Auswahl des Zeichensatzes, speichern. Einfache Texte können unstrukturiert sein (.TXT) oder strukturiert (.XML, .CSV). Strukturierte Texte enthalten Zusatzinformationen, die lesbar in den Text integriert sind. Bei der Archivierung muss nun sichergestellt werden, dass diese Zusatzinformationen interpretiert werden können, was eine zusätzliche Beschreibung der Bedeutung beziehungsweise Struktur dieser Zusatzinformationen erfordert.

Übliche Textverarbeitungsprogramme erzeugen zumeist formatierte Texte. Auch wenn es einige weitverbreitete Formate gibt, so sind diese doch meist nur kurzlebig und daher schlecht zur Archivierung geeignet, da ständige Migrationen erforderlich würden. Abhilfe soll ein normiertes Format schaffen: die ISO-Norm 19005 beschreibt PDF/A explizit für Archivierungszwecke. Ausgehend vom PDF-Format der Firma Adobe wurde hier von allzu modernen und kurzlebigen Features abgesehen, um langfristige Stabilität zu erreichen. PDF/A ist gegenwärtig das einzige Format für formatierte Texte, das weitgehend uneingeschränkt für die Archivierung empfohlen werden kann. Zu berücksichtigen sind allerdings mehrere Versionen von PDF/A, wobei die einfachen Versionen (zum Beispiel PDF/A-1b) nur das Erscheinungsbild des Dokuments abbilden und eine Extraktion des Textes (zum Beispiel copy-paste) nicht erlauben.

Eine Archivierungsvariante ergibt sich, wenn man das Schrift-Dokument als Bild betrachtet und in einem für Bilder geeigneten Format archiviert.

Formate für Texte
bevorzugt: • **TXT** für einfache Texte • **PDF/A** für formatierte Texte
akzeptiert (ggf. problematisch): • ggf. **CSV** für Tabellen (zzgl. Strukturbeschreibung) • ggf. **TIFF** (→ Bilddokumente) für formatierte Texte
Kandidaten (problematisch, weiter prüfen): • ggf. SGML, XML (zzgl. Strukturbeschreibung)

5 Zusammenfassung: Text-Formate zur Archivierung.

5.2 Bilder

Bei Bildern sind die Dateiformate mit einer Vielzahl von Eigenschaften und Attributen verknüpft, die schon bei der Erzeugung der Bilddaten, also beim Scannen, Fotografieren usw. berücksichtigt werden müssen. Dazu gehören zum Beispiel Auflösung, Farbtiefe, Farbraum und vieles andere. Dazu gibt es im Anhang eine ausführliche Beschreibung des „Workflow". Nachfolgend kann der Fokus daher auf die Formate beschränkt bleiben.

Die Menge der oben angegebenen Eigenschaften und Attribute erfordert eine sorgfältige Dokumentation der Formatspezifikationen der archivierten Bilddokumente in den Metadaten. Je detaillierter diese Dokumentation ist, umso leichter lassen sich erforderliche Migrationen planen und durchführen.

Bei der Archivierung von Bildern sollte, in Anbetracht sinkender Preise für Foto- und Speichertechnik, auch darüber nachgedacht werden, Bilder mehrfach redundant zu archivieren. So können die Nachteile einzelner Formate vermieden werden und man kann später bei der Nutzung das zweckmäßigste Format auswählen.

Eine weitere sehr gute Möglichkeit, wichtige Aufnahmen risikofrei für die nächsten 50–100 Jahre aufzubewahren, stellt das Ausbelichten auf Fotopapier und Aufbewahren in herkömmlichen Archivschränken dar. Genau so ernsthaft wird auch immer wieder die Ausbelichtung auf Dia- oder Mikrofilm diskutiert.

RAW-Bildformate

Die besten Möglichkeiten, um möglichst alle „gemessenen" Bildinformationen abzuspeichern, bieten die sogenannten RAW-Bilder (Rohdaten). Diese RAW-Dateien bieten die maximale Qualität bei geringst möglicher Speichergröße der verlustfrei komprimierten Daten, allerdings zum Preis einer sehr geringen Kompatibilität.

Einen sehr hoffnungsvollen Ansatz bot 2004 die Einführung und Offenlegung des DNG-Formates als potentieller Standard für das „digitale Negativ". Dieser Standard konnte sich allerdings bis heute nicht recht durchsetzen. Allerdings gibt es zum Beispiel mit Adobe Camera Raw ein Tool zum Konvertieren der proprietären RAW-Formate.

Nikon etwa hat bis heute noch nicht alle Spezifikationen des eigenen RAW-Formats offengelegt, bei Canon verändern sich die Spezifikationen von einem Kameramodell zum Nächsten. Im High-End-Bereich sieht die Sache nicht wirklich besser aus, hier bleibt die weitere Entwicklung abzuwarten und bis dahin (soweit alle Bearbeitungsmöglichkeiten erhalten bleiben sollen) die kameraeigenen RAW-Dateien zu archivieren oder die RAW-Dateien in das DNG-Format zu konvertieren.

Im Scanbereich sieht es noch düsterer aus, etwa Imacon unterstützt in der neuesten Softwareversion das eigene „3f"-Format nicht mehr oder Vuescan-DNGs lassen sich nur mit Vuescan weiter bearbeiten, wenn man alle Bearbeitungsmöglichkeiten erhalten will.

Aufgrund der fehlenden Kompatibilität können RAW-Formate nur sehr eingeschränkt zur Archivierung empfohlen werden. Es kann jedoch aufgrund der Qualität erforderlich und sinnvoll sein, die RAW-Dateien neben einem für die (Langzeit-)Archivierung tauglicheren Format ebenfalls zu bewahren, zumindest solange eine Verwendung der speziellen Formate noch gegeben ist.

TIFF-Format

Für die Speicherung der fertigen Bilddateien gilt es, eine Abwägung der sich teilweise widersprechenden Eigenschaften Speicherplatzbedarf versus Qualität versus Kompatibilität zu treffen. Dies wird vor allem beim TIFF-Format mit seinen Optionen und Varianten deutlich. TIFF hat sich insbesondere im Bereich der Druckindustrie als Quasi-Standard für Rasterbilder durchge-setzt. Es enthält allerdings außerordentlich viele Optionen und Varianten (Farbtiefe, Komprimierung, Bit-Anordnung, Ebenenkomprimierung, Farbkanäle usw.). Dies führt sehr schnell zu Inkompatibilitäten. Ein Ausweg ist die Mindestspezifikation „Baseline TIFF", die alle „TIFF-fähigen" Programme umsetzen sollten. Leider bleibt bei der Nutzung von Bildverarbeitungsprogrammen zumeist im Dunkeln, welche Speicher-Optionen „Baseline" sind und welche nicht.

Unkomprimierte TIFFs sind vorzuziehen, weil ein Bitfehler nicht gleich zum Totalausfall führt. Wenn die Speichergröße jedoch minimiert werden muss, kann die G4-Komprimierung (für reine Schwarz-Weiß-Bilder) beziehungsweise die LZW-Komprimierung empfohlen werden.

Tiff-Dateien bieten außerdem beste Möglichkeiten, Metadaten (EXIF, IPTC und XMP) sowie ICC Profile in die Bilddatei zu integrieren.

JPEG-Format

JPEG ist normalerweise ein Bildformat mit verlustbehafteter Komprimierung, es gehen also Informationen verloren. Dies ist unter strengen archivalischen Gesichtspunkten schwer hinnehmbar. Man kann sich jedoch bei den derzeit erreichbaren Bildauflösungen auf den Standpunkt stellen, dass diese Verluste in

bestimmtem Maße akzeptabel sind. Dies wird vor allem dann der Fall sein, wenn die Aufnahmen eher einen visuellen Eindruck verschaffen sollen und der exakte „messtechnische" Effekt unbedeutend ist. In solchen Fällen kann eine moderate Komprimierung viel Speicherplatz sparen.

Deutlicher Vorteil von JPEG ist der offene Standard und die hohe Verbreitung. Sämtliche Digitalkameras mit Ausnahme einiger High-End-Digibacks speichern ausgearbeitete Bilddateien im JPEG-Format. Ebenso verarbeiten auch fast alle Bildverarbeitungsprogramme JPEG-Daten. Nach heutigem Stand bietet das JPEG-Format die größte Wahrscheinlichkeit, dass diese Dateien auch in vielen Jahren ohne weitere Konvertierungen noch zu öffnen sein werden.

Einschränkend muss jedoch darauf verwiesen werden, dass durch die verlustbehaftete Komprimierung bei jedem Speichern ein neuer, zusätzlicher Fehler entstehen kann. Insofern sollte man während der Bearbeitungsphase der Bilder das JPEG-Format möglichst vermeiden und nur die fertig ausgearbeitete Bilddatei in JPEG speichern.

Formate für Bild-Dokumente
bevorzugt: • **TIFF unkomprimiert** (TIFF 6.0 Baseline-Formate) • **TIFF G4-komprimiert** (verlustfrei nach ITU/CCITT) nur für Schwarz-Weiß-Bilder • Bearbeitungsverluste vermeiden
akzeptiert (ggf. problematisch): • ggf. **TIFF LZW**-komprimiert • ggf. **JPEG** (wenn Komprimierungsverluste akzeptabel sind, hohe Auflösung erforderlich)
Kandidaten (problematisch, weiter prüfen): • ggf. JPEG2000 (verlustfrei komprimiert, offen aber wenig verbreitet, keine Erfahrung) • ggf. DNG als Rohdatenformat (standardisiert aber nicht weit verbreitet, keine Erfahrungen)

6 Zusammenfassung: Bildformate zur Archivierung.

Das JPEG-Format kann unter denselben Gesichtspunkten auch zur Speicherung von Bildern innerhalb des PDF/A-Formats verwendet werden.

Im Gegensatz zum „normalen" JPEG bietet JPEG2000 eine verlustfreie Komprimierung an. JPEG2000 ist ebenfalls ein offener Standard, aber bei Weitem nicht so verbreitet wie JPEG, und kann daher auch nicht vorbehaltlos empfohlen werden.

5.3 CAD-Daten

Sowohl im CAD-Bereich als auch bei GIS gibt es eine große Heterogenität der Datenformate und nur wenige Standards, die jedoch nicht weitverbreitet sind. In dieser Situation findet man keine Formate, die unproblematisch für die Archivierung geeignet sind, sodass auch keine Empfehlung für ein „bevorzugtes" Archivierungsformat ausgesprochen werden kann. Nachfolgend sollen daher die verbreitetsten Formate betrachtet und hinsichtlich ihrer Archivierung bewertet werden.

Obwohl es unter Anwendung gängiger Programme (ArchäoCAD, TachyCAD, Singularch) mehr oder minder erfolgreich möglich ist, Sachdaten an CAD-Geometrien anzuhängen, ist diese Arbeitsweise nicht sehr üblich (und wird zumeist GIS-Systemen überlassen). In diesem Sinne sollen CAD-Dateien als Grabungszeichnungen von Flächen und Profilen, übergreifenden Grabungszusammenhängen usw. verstanden werden, die keine zusätzlichen Sachdaten enthalten.

Hier wird auf die Archivierung, die Gewährleistung der Lesbarkeit und die inhaltliche Integrität der Daten eingegangen – die Problematik des Austausches mit anderen Systemen (vor allem mit GIS) ist hinlänglich bekannt. Diese Probleme sind aber nicht unbedingt die gleichen, die bei der Frage der Archivierung von CAD-Daten auftauchen. Wenn es um Archivierung geht, ist die Situation sowohl einfach als auch problematisch.

AutoCAD-DWG

Der CAD-Marktführer ist mit Abstand Autodesk mit dem unter Windows laufenden Produkt AutoCAD. Infolgedessen ist das „DWG"(Drawing)-Format von Autodesk zum Quasi-Standard für 2-D- und 3-D-CAD-Zeichnungen geworden. Dieses Format ist völlig proprietär, es ändert sich mit jeder Programmversion (jährlich) und die Dokumentation ist nicht frei erhältlich. Neue Programmversionen sind abwärtskompatibel – zumindest einige Stufen. Um die vollständige Funktionalität von AutoCAD zu gewährleisten, ist man zumeist auf das DWG-Format angewiesen.

AutoCAD-DXF

Das DXF-Format ist ebenfalls ein proprietäres Format von Autodesk, allerdings ist die Dokumentation für DXF frei verfügbar, was zu einer weiten Verbreitung und zu großer Akzeptanz als Zeichnungsaustauschformat geführt hat. Damit ist DXF ein Quasi-Standard. Leider ändert sich auch das DXF-Format mit jeder neuen Programmversion von AutoCAD.

DXF ist ein in ASCII-Zeichen lesbares Abbild der binär abgespeicherten DWG und unterstützt direkt 2-D- und 3-D-Koordinaten. Der Aufbau ist sehr klar, einfach und strukturiert. DXF ist oft der kleinste gemeinsame Nenner vieler Vektorgrafikprogramme, aber meist werden damit nicht alle Funktionen dieser Programme voll unterstützt und es können beim Export Daten oder Funktionalitäten verloren gehen.

Aufgrund des offenen Standards und der weiten Verbreitung ist DXF zur Archivierung recht gut geeignet, mit der Einschränkung der geringen Haltbarkeit und der damit verbundenen Notwendigkeit häufiger Migrationen. Das DXF-Format R12 (1992) ist beispielsweise mit vielen CAD- und GIS-Systemen lesbar und damit für die Archivierung geeignet.

Formate für CAD-Dokumente
bevorzugt: • derzeit ohne Empfehlung
akzeptiert (ggf. problematisch): • ggf. **DXF** (Einschränkungen in der Funktionalität möglich, z.B. DXF R12 ist gut lesbar)
Kandidaten (problematisch, weiter prüfen): • ggf. DWG (stark eingeschränkte Nutzungsdauer, häufige Migrationen erforderlich) • ggf. PDF/A (nur für Ansichten/ Visualisierungen mit stark beschränkter Funktionalität)

7 Zusammenfassung: CAD-Formate zur Archivierung.

Alle weiteren CAD-Formate spielen eine weit untergeordnete Rolle und kommen daher für die Archivierung nicht infrage. Deshalb sollten CAD-Daten möglichst als DWG und DXF gespeichert werden. Etwaige CAD-fremde externe Referenzen (Bitmaps etc.) unterliegen den Empfehlungen für die jeweilige Datengattung.

Das Grundproblem bei der Archivierung von CAD-Daten ist der schnelle Wandel der Versionen, der eine häufige Migration erfordert. Wenn man sich auf längere Migrationszyklen einlässt (und dabei gegebenenfalls mehrere Versionen überspringt), geht dies mit einem Funktionsverlust gegenüber den aktuellen Programmversionen einher. Am sinnvollsten kann man dem begegnen, indem man mehrere Versionen für verschiedene Nutzungszwecke parallel vorhält:

- Eine möglichst stabile und weitverbreitete (gegebenenfalls ältere) Version (Vorschlag DXF) wird zur dauerhaften Archivierung (laufende Migration) benutzt, wobei der wesentliche Funktionsgehalt erhalten bleiben sollte, jedoch spezielle Eigenschaften verloren gehen können.
- Die bei der Erstellung aktuelle Version (Vorschlag: DWG) wird begrenzt aufbewahrt, um eine verlustfreie Weiterbearbeitung zu ermöglichen, solange noch Lesbarkeit besteht (Dies ist in unserem Sinn eigentlich keine Archivierung.).
- Eine weitgehend problemlose Archivierung lässt sich durch die Erstellung von Ansichten (Visualisierungen, Ausdrucke) über Text- oder Bildformate erreichen (Empfehlung PDF/A). Möglicherweise ist auch die Erstellung von Hardcopys sinnvoll. Hierbei ist auf eine sorgfältige Einstellung sämtlicher Ansichten zu achten, damit möglichst viele Aspekte der Originaldaten wiedergegeben werden (gegebenenfalls auch an Listen von Attributen und Sachdaten denken)! Die originale Funktionalität der Daten geht dabei jedoch verloren!

Mit dieser Methode wird sowohl die dauerhafte Archivierung (gegebenenfalls mit inhaltlichen Abstrichen) als auch die vollinhaltliche Nutzbarkeit (zumindest für einen begrenzten Zeitraum) gewährleistet.

CAD-Daten werden umfangreicher Pflege bedürfen, sie unterliegen einem hohen Migrationsaufwand. Diese Migration kann automatisch durchlaufen, muss aber zumindest stichprobenartig manuell auf Datenverluste kontrolliert werden. Dazu ist es erforderlich, dass die CAD-Daten möglichst schnell eine geeignete „Archivierungs- und Pflegeeinrichtung“ erreichen, wo sie entsprechend behandelt werden. Dies könnte eine entsprechend ausgerüstete behördeninterne Einrichtung oder auch eine kompetente externe Stelle sein.

5.4 Geodaten

Im GIS-Bereich unterscheidet man zunächst einmal zwischen Vektordaten und Rasterdaten, wobei die Vektordaten am ehesten GIS-typisch sind und die Rasterdaten oft aus anderen Quellen mit deren typischen Formaten stammen (zum Beispiel Bildformate). Die Verwendung dieser Daten in einem GIS ist höchst komplex und auf die jeweilige Funktionalität direkt bezogen. Eine Standardisierung, wie sie für die Archivierung wünschenswert ist, setzt dieser Funktionalität klare Grenzen und konnte sich daher bisher nur unzureichend durchsetzen. Es gibt jedoch einige Datenformate, die sich als eine Art Quasi-Standard etabliert haben und von vielen GIS importiert und genutzt werden können.

Aufgrund der hohen Komplexität und geringen Standardisierung kann derzeit keine uneingeschränkte Empfehlung für die Archivierung von GIS-Daten ausgesprochen werden. Demzufolge erfordert die Archivierung eine hohe Aufmerksamkeit und einen hohen Migrationsaufwand, ähnlich wie bei CAD-Daten.

AutoCAD DXF

Die meisten GIS unterstützen den Import und Export von CAD-Daten, wobei dem DXF-Format dabei eine herausragende Stellung zukommt. Damit können auch GIS-Daten als DXF archiviert werden, mit allen schon benannten Problemen. Ein Nachteil ist die Tatsache, dass DXF aus sich heraus keine Sachinformationen übermitteln kann, wie sie üblicherweise an GIS-Zeichnungsobjekte angehängt sind. Zusätzliche Informationen, wie zum Beispiel die Beschreibung des Befundes, sind meist nicht ohne Weiteres mit dem Befundzeichnungsobjekt in einem CAD-System verknüpfbar (dies ist meist teurer Spezialsoftware vorbehalten). Aus diesen Gründen ist eine Speicherung von GIS-Vektordaten im DXF-Format wenig empfehlenswert. Die Sachinformationen müssen in jedem Fall zusätzlich archiviert werden, zum Beispiel in einem Tabellenformat.

Tabellenformate

Zumindest die Sachinformationen eines GIS lassen sich zumeist über gängige Datenbankformate importieren und exportieren (TXT, CSV, XLS ..., gelegentlich auch DBF, MDB, XML unter anderem). In Spezialfällen funktioniert dies auch mit den Geometrie-Informationen, zum Beispiel für Punktkoordinaten oder bestimmte textbasierte Repräsentationen der Geometrie (Simple Feature Access bzw. Well Known Text – WKT). In diesen begrenzten Fällen kann eine Archivierung nach den Kriterien von Datenbanken erfolgen. Dabei ist aber darauf zu achten, dass die Punktinformation in einer aus einem GIS zu exportierenden Datei nicht automatisch die beiden benötigten Spalten für das Koordinatenpaar enthält, diese müssen nötigenfalls vorher erst erzeugt und mit den entsprechenden Informationen gefüllt werden.

ADeX

Das Archäologische Datenexportformat ADeX ist in seiner auf Punktkoordinaten beschränkten Version 1.2 uneingeschränkt archivierungstauglich. Bei der Version 2, die auch Linien und Flächen unterstützt, sind zusätzlich die Hinweise zu den Formaten SHP und MIF zu beachten.

ESRI-Shape-Format – SHP

Das gängigste Vektorformat, das heute in GIS-Programmen Verwendung findet, ist das von der Firma ESRI für seine ArcGIS/ArcINFO-Produktschiene entwickelte Shape-Format (SHP). Nahezu alle gängigen GIS-Programme können SHP direkt lesen oder aber zumindest importieren. Auch zahlreiche OpenSource-GIS-Programme (unter anderem gvSIG oder Quantum) verwenden SHP trotz seiner Eigenschaft als proprietäres Format. SHP ist relativ gut dokumentiert, hat sich in den letzten Jahren nicht wesentlich geändert und dadurch eine weite Verbreitung gefunden. Eine besondere Eigenschaft von Shapes ist die Tatsache, dass in einer SHP-Datei nur entweder Punkte, Multipunkte, Linien oder Flächen enthalten sein können. Ein Gewässerlayer mit Quellen (Punkten), Flüssen (Linien) und Seen (Flächen) wird also beim Export in SHP in drei einzelne Shapefiles aufgespaltet. Zudem ist SHP nur „maschinenlesbar", eine Ansicht mittels eines Texteditors fördert (anders als zum Beispiel bei WKT, GML oder MIF) keine Informationen zutage, die auch nach einem System- bzw. Formatwechsel noch verwertbar sind. Shapefiles bestehen immer aus mehreren Dateien, die Geometriedaten, Indexdaten, Sachdaten (im DBF-Format) sowie weitere Informationen enthalten.

Trotz aller Nachteile spricht die weite Verbreitung, die große Akzeptanz und die relativ stabile Versionierung für SHP als Archivformat, sofern auch hier die Grundsätze einer Formatmigrierung nötigenfalls berücksichtigt werden.

MapInfo Interchange Format – MIF

Die Software MapInfo verwendet zum Austausch mit anderen GIS-Programmen das Austauschformat MIF (MapInfo Interchange Format). Es handelt sich um ein reines Textformat, in dem (geo)grafische Informationen abgespeichert werden. Dabei werden neben formatbeschreibenden Informationen alle in einer Datei enthaltenen Zeichnungsobjekte (Punkte, Linien, Flächen) beschrieben und mit Koordinaten sowie dem verwendeten Projektionssystem abgespeichert. Aus der Textdatei sind also die relevanten Informationen auch ohne ein kompatibles GIS-Programm entnehm-

bar und gegebenenfalls anderweitig verwendbar. Ergänzt werden die eigentlichen MIF-Dateien durch gleichnamige MID-Dateien, die die Attribute (das heißt die Datenbankinformationen) der (geo)grafischen Objekte enthalten.

Die Verbreitung als Im-/Exportformat für die in der Archäologie relevanten GIS-Produkte ist mittelmäßig.

Beim Export aus anderen Programmen als MapInfo ist darauf zu achten, dass zum Beispiel IDRISI nicht immer die richtige Georeferenzinformation in die Datei speichert (manchmal Nicht-Welt-Meter statt GK3). Diese kann aber in einem Texteditor (oder bei mehreren Dateien durch die Verwendung des Freeware-Programms BTRWIZ) ersetzt werden.

XML / GML / KML

Außer GoogleEarth, das nicht als GIS, sondern als GIS-Dataviewer mit beschränktem Formatrepertoire bezeichnet werden muss, beherrschen bislang nur wenige gängige GIS-Programme den direkten Im- und Export von XML oder seinen Derivaten GML und KML (ESRI benutzt zum Beispiel XML zur Erzeugung beziehungsweise Speicherung von Metadaten, MapInfo kann über ein Zusatzmodul Daten als KML exportieren). Derzeit ist eine Speicherung in diesen Formaten nicht zu empfehlen, da ein Austausch nicht immer problemlos gewährleistet ist. Da aber die Formate gut dokumentiert, als Textdateien einfach lesbar sind und zukünftig mit einer breiteren Akzeptanz dieser Formate auch bei Softwareherstellern sowie -programmierern zu rechnen ist, ist eine Archivierung zum Beispiel als GML durchaus sinnvoll, sofern das verwendete GIS-Programm einen problemlosen Export gewährleistet.

Sonstige (QGIS, IDRISI, GRASS etc.)

Zahlreiche GIS-Programme sowohl aus dem kommerziellen als auch aus dem OpenSource-Bereich verwenden jeweils eigene Dateiformate, die mehr oder weniger verbreitet und kompatibel sind. Meist bieten diese Programme aber auch die Möglichkeit des Exports in einem der oben genannten Formate. Diese Formate sollten in jedem Fall – zumindest immer zusätzlich – für die Archivierung von GIS-Vektordaten verwendet werden. Nur sie gewährleisten einen einigermaßen reibungslosen Austausch mit anderen Programmen beziehungsweise die langfristig lesbare Speicherung und Archivierung der Informationen.

GPS-Daten

Auch GPS-Daten sind letztlich meist Vektordaten (Tracks, Routes, Waypoints). Sie liegen oft in Formaten vor, die die jeweilige Hardware und die weiterverarbeitende Software vorgibt. Als einigermaßen kompatibel hat sich GPX (GPS Exchange Format) erwiesen, das auch auf einem XML-Schema aufbaut und von zahlreichen Geräten sowie Softwareprodukten gelesen und erzeugt werden kann. Glücklicherweise sind viele Konvertierungsprogramme für die unterschiedlichen GPS-Formate auch frei erhältlich; die umfangreichste Auswahl an unterstützten Formaten bietet dabei GPSBabel.

Selbstverständlich ist auch hier eine Speicherung als GML oder KML oder eine Umwandlung in eines der anderen oben empfohlenen GIS-Formate (SHP, MIF) möglich und sinnvoll.

Geländemodelle

Auch Geländemodelle können außer als Raster- auch als Vektordaten gespeichert werden. Zum Teil sind die Übergänge hier fließend. Einige der am weitesten verbreiteten Formate sind Tabellen, die die entsprechenden Höheninformationen enthalten.

Unregelmäßige Informationen können zum Beispiel als einfache Tabelle mit den Spalten X-Koordinate, Y-Koordinate und Höhe vorliegen und später gegebenenfalls im GIS in regelmäßige Raster umgesetzt (interpoliert) werden. In dieser Form lassen sich auch regelmäßig verteilte Werte (regelmäßige Messraster und vor allem interpolierte Daten) abspeichern. Aus Platzgründen und weil entsprechende Dateien von fast allen gängigen GIS-Programmen gelesen und erzeugt werden können, empfiehlt sich aber die Verwendung des sogenannten nicht proprietären ESRI-beziehungsweise ARC/INFO-ASCII-Grids. Dabei werden nach einem standardisierten Header, der unter anderem Informationen zur Zeilen- und Spaltenzahl des Rasters und zur Zellgröße sowie den Eckkoordinaten enthält, die Höhenwerte (oder sonstigen Rasterinformationen) zeilenweise in der Datei fortgeschrieben.

Da diese Daten anders als diverse Binärformate als in einem Texteditor anzeigbare und editierbare Dateien vorliegen, empfiehlt sich neben der Speicherung beziehungsweise Archivierung in einem GIS-spezifischen Rasterformat (GeoTIFF, DTED usw.) auch ein Export in ein einfaches Tabellenformat, wie oben beschrieben.

GeoTIFF und andere Georeferenzierte Rasterdaten

Rasterdaten benötigen zur Darstellung im GIS eine Georeferenz, das heißt Koordinaten für die Lage und Ausdehnung sowie das benutzte Koordinatenreferenzsystem. Neben verschiedenen proprietären Formaten ist GeoTIFF weitverbreitet. GeoTIFF ist eine spezielle Form des TIFF-Formats. Die normalen TIFF-Optionen bleiben erhalten (siehe Abschnitt „Bilder"), sodass die Dateien theoretisch auch mit normalen Viewern lesbar sind. Allerdings werden oft nicht die

Baseline-Optionen benutzt, sodass doch wieder spezielle Viewer erforderlich werden. Zusätzlich werden die Georeferenz-Informationen in speziellen Tags abgelegt, die mittlerweile von den meisten GIS-Systemen ausgewertet werden können.

Aus der ESRI-Welt (ArcGIS) stammt das Verfahren, eine normale Bilddatei um ein sogenanntes „World-File" zu ergänzen, welches die Georeferenz in lesbarer Form (TXT) enthält, allerdings ohne Koordinatenreferenzsystem. Letzteres muss also separat übermittelt und bewahrt werden. World-Files werden typischerweise für JPEG-, PNG-, GIF- und TIFF-Bilddateien benutzt.

Beide Varianten sind offene Quasi-Standards und weitverbreitet, aber trotzdem nicht immer unproblematisch. Falls angeboten, sollte man ein GeoTIFF im Baseline-Format benutzen und zusätzlich noch ein World-File erstellen; das benutzte Koordinatenreferenzsystem gehört dann in die Metadaten.

Amtliche Landesvermessung

Die Vermessungsämter der deutschen Bundesländer erzeugen Daten in eigenen Formaten beziehungsweise als eigene Modelle (zum Beispiel ALK, AFIS-ALKIS-ATKIS). Diese werden über die sogenannte Einheitliche Datenbankschnittstelle (EDBS) ausgetauscht, die jedoch in gängigen GIS-Programmen bislang keine Rolle spielt.

Geodatabase

Neben der Speicherung in Dateien hat sich, vor allem im Client-Server-Bereich, die Nutzung von Datenbanken mit speziellen Geodatenformaten und räumlichen Abfragemöglichkeiten für die Speicherung und Verarbeitung von Geodaten immer weiter durchgesetzt. Sowohl GIS- als auch Datenbankanbieter haben entsprechende Datenbank-Erweiterungen im Angebot: zum Beispiel Smallworld (VMDS), ESRI (Geodatabase), IBM (DB2 Spatial Extender), Oracle (Spatial), Microsoft (SQL Server spatial types); im Open-Source-Bereich steht PostgreSQL mit der spatial extension PostGIS zur Verfügung.

Obwohl diese Systeme weitgehend proprietär sind, stützen sie sich doch auf langfristig angelegte Datenbankstandards und beinhalten in der Regel schon Migrationsmethoden wie bei Datenbanken üblich. Insofern kann für diesen speziellen Fall von Geodaten auf die Ausführungen im nachfolgenden Abschnitt zu Datenbanken verwiesen und ein hohes Potenzial für die Langzeitarchivierung abgeleitet werden.

Das Grundproblem bei der Archivierung von GIS-Daten ist, wie schon bei CAD-Daten erwähnt, der schnelle Wandel der Versionen, der eine häufige Migration erfordert. Wenn man sich auf längere Migrationszyklen einlässt (und dabei gegebenenfalls mehrere Versionen überspringt), geht dies mit einem Funktionsverlust gegenüber den aktuellen Programmversionen einher. Die Speicherung mehrerer Versionen in unterschiedlichen Formaten ist auch im GIS-Bereich sinnvoll:

- Ein möglichst stabiles und weitverbreitetes Format (Vorschlag SHP) wird zur dauerhaften Archivierung (laufende Migration) benutzt, wobei der wesentliche Funktionsgehalt erhalten bleibt.
- Die bei Beginn der Archivierung für den Betrieb maßgeblichen Dateien (aktuelle Version) werden im Original begrenzt aufbewahrt, um eine verlustfreie Weiterbearbeitung zu ermöglichen, solange noch Lesbarkeit besteht (Dies ist in unserem Sinn eigentlich keine Archivierung.).
- Eine weitgehend problemlose Archivierung lässt sich durch die Erstellung von Ansichten (Visualisierungen, Ausdrucke) über Text- oder Bildformate

Formate für GIS

bevorzugt:

- derzeit ohne Empfehlung

akzeptiert (ggf. problematisch):

- ggf. **SHP** (weit verbreitetes proprietäres, aber offen beschriebenes Binärformat)
- ggf. **ADeX** (für Punktinformationen bzw. in Verbindung mit SHP oder MIF)

Kandidaten (problematisch, weiter prüfen):

- ggf. MIF (proprietäres, aber offen beschriebenes Textformat)
- ggf. *GML* (standardisiert, XML-basiert, weniger verbreitet)
- ggf. KML (proprietär, offen beschrieben XML-basiert)
- ggf. PDF/A (nur für Ansichten/ Visualisierungen mit stark beschränkter Funktionalität)
- ggf. GeoTIFF (besser mit zus. World-File und Angabe zum Koordinatenreferenzsystem)

8 Zusammenfassung: GIS-Formate zur Archivierung.

erreichen (Empfehlung PDF/A), möglicherweise ist auch die Erstellung von Hardcopys sinnvoll. Hierbei ist auf eine sorgfältige Einstellung der möglichen Ansichten zu achten, damit möglichst viele Aspekte der Originaldaten wiedergegeben werden (gegebenenfalls auch an Listen von Attributen und Sachdaten denken)! Die originale Funktionalität der Daten geht dabei jedoch verloren!

Mit dieser Methode wird sowohl die dauerhafte Archivierung (gegebenenfalls mit inhaltlichen Abstrichen) als auch die vollinhaltliche Nutzbarkeit (zumindest für einen begrenzten Zeitraum) gewährleistet.

Auch GIS-Daten werden umfangreicher Pflege bedürfen, da sie einem hohen Migrationsaufwand unterliegen und stichprobenartig manuell auf Datenverluste kontrolliert werden müssen. Dazu ist, ebenso wie bei CAD-Daten, eine leistungsfähige und kompetente „Archivierungs- und Pflegeeinrichtung" erforderlich.

5.5 Datenbanken

Datenbanken sind weitverbreitete, lange und gut eingeführte IT-Systeme. Vor allem Server-Datenbanken sind gut auf einen dauerhaften Betrieb über Jahre hinweg ausgerichtet. Sie sind stark standardisiert und beinhalten die für Versionswechsel erforderlichen Migrationsmethoden. Datenbanken besitzen damit schon jetzt hohes Potenzial für die Langzeitarchivierung. Aktive Datenbanken mit „lebendigen" Daten sind dabei am problemlosesten, weil Betreuung, Wartung sowie Migration zum laufenden Betrieb gehören und damit in der Regel eingeplant und abgesichert sind. Bei inaktiven, „abgelegten" Datenbanken sind jedoch auch Überlegungen zur Archivierung erforderlich.

Bei der Archivierung von Datenbanken ist es wichtig, dass nicht nur die Daten, sondern auch Datenmodelle und Datenbankschemata archiviert werden:

- Datenmodell: Die verbreiteten relationalen Datenbanken benutzen als konzeptuelles Datenmodell meist das ER-Modell (Entity Relationship Model): Entitäten stehen für Objekte der Wirklichkeit, Relationships stehen für Beziehungen zwischen den Objekten. ER-Modelle sind eine ausgezeichnete Modellierungsgrundlage und damit gut zur (langfristigen) Vermittlung des einer Datenbank zugrundeliegenden konzeptuellen Modells geeignet. Sie werden als Diagramm dargestellt (ERD), um die Struktur zu verdeutlichen. Eine rein textbasierte Darstellung der ERD ist nicht bekannt. Daher kommt wohl am ehesten PDF/A für deren Archivierung in Frage.
- Datenbankschema: Das konzeptuelle Datenmodell wird als Datenbankschema technisch umgesetzt. Das Datenbankschema beinhaltet Beschreibungen der Tabellen (zum Beispiel Name), der Attribute (zum Beispiel Datentypen und Länge, Indizes) und Integritätsbedingungen (Primär-/Sekundärschlüssel, Referenzen, Existenz-/Wertebereiche) einer Datenbank. Die Darstellung erfolgt am besten per SQL (DDL, siehe unten).
- Daten: Gut archivierungstauglich sind die „einfachen" Datenformate wie Texte und Zahlen, die als Text darstellbar und damit speicherbar sind. Dagegen sind binäre Formate (BLOBs) ungeeignet, denn sie besitzen keine datenbankseitig standardisierte innere Struktur und können nicht als Text exportiert werden. BLOBs erfordern immer einen expliziten Exportmechanismus in extern lesbare Formate. In Datenbanken sollten also möglichst keine binär kodierten Dokumente oder Bilder integriert werden, sondern nur deren Links zu separat abgelegten Dateien in archivtauglichen Formaten. Daten können extern zum Beispiel als SQL (DML) oder als CSV dargestellt werden.

SQL (Structured Query Language) ist eine standardisierte Sprache zur Definition, Manipulation, Abfrage und Steuerung von relationalen Datenbanken, die schon seit 1979 existiert. Man unterscheidet DDL (Data Definition Language), DML (Data Manipulation Language) und DCL (Data Control Language). SQL ist ein standardisiertes und weitverbreitetes Textformat. Es eignet sich daher gut für den Datenaustausch beziehungsweise die Migration in heterogenen Umgebungen. SQL ist für die Archivierung von Datenbankschema und Daten gut geeignet.

CSV (Comma Seperated Value) ist ein textbasiertes Datenformat zur einfachen Speicherung von Tabelleninhalten. Es ist zwar nicht standardisiert, aber aufgrund seiner großen Verbreitung und Akzeptanz als Quasi-Standard zu betrachten. Eine Beschreibung findet man in RFC 4180 (Request for Comments – techn./org. Dokumente zum Internet). CSV ist für die Archivierung von Tabellendaten gut geeignet.

Desktop-Datenbanken wie zum Beispiel MS-Access sind nicht zur direkten Archivierung geeignet. Ihre internen Strukturen können sich von Version zu Version schnell ändern. Sie müssen, wie oben angegeben, mit Datenmodell, Datenbankschema und Daten als TXT oder ähnlich archiviert werden oder in eine aktive Server-Datenbank mit stabiler Struktur übernommen werden.

Gegebenenfalls ist XML (mit DTD) zukünftig eine Alternative bei der Archivierung von Datenbankschemata und Daten. XML ist textbasiert und daher

gut archivierungstauglich. Gegenwärtig liegen aber noch nicht viele Erfahrungen mit XML-Daten vor, die Thematik sollte also weiter verfolgt werden.

Das Schweizer Bundesarchiv hat einen eigenen Standard „SIARD RDB DATA" etabliert. Tools extrahieren dabei aus relationalen Datenbanken Strukturinformationen sowie Daten und wandeln sie in ein archivierbares Format um. Es wird auch eine Beschreibung der Tabellen und Attribute erwartet. Der Standard basiert auf SQL, CSV und XML.

Übrigens: Das Schweizer Bundesarchiv archiviert ausschließlich die folgenden Formate: TXT, PDF/A, CSV, SIARD, TIFF und WAVE.

Das „kopal"-Projekt (Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen) vertraut in die Langzeitarchivierungsfähigkeit professioneller Datenbanksysteme (IBM). Allgemeine Vor- und Nachteile wurden schon beschrieben. Im Falle von „kopal" werden binäre Dokumentformate als BLOBs gespeichert, womit sich eine Abhängigkeit von proprietären Datenbankschnittstellen ergibt.

Das Deutsche Archäologische Institut (DAI) fordert zur Archivierung von Datenbanken eine Dokumentation (Anforderungen, Datenstruktur als ERD, Tabellen, semantische Beschreibung). Es sollen keine BLOBs, sondern Referenzierungen auf das Verzeichnissystem benutzt werden. Neben der Original-Datenbank soll auch ein softwareunabhängiges Format als XML (+ dokumentierte DTD) oder SQL (DDL) bereitgestellt werden. Weitere Hinweise zur Archivierung von Datenbanken, auch hinsichtlich XML, findet man bei „Nestor".

Formate für Datenbanken
bevorzugt: • dauerhaft betriebene Online-Datenbank
akzeptiert (ggf. problematisch): • ggf. **SQL** (+ Modell, + Schema) • ggf. **CSV** (+ Modell, + Schema)
Kandidaten (problematisch, weiter prüfen): • ggf. PDF/A (nur für Modell, auch für Ansichten/Visualisierungen mit beschränkter Funktionalität)

9 Zusammenfassung: Datenbank-Formate zur Archivierung

Der Betrieb einer Datenbank (Server-DB) beinhaltet in sich schon alle Aspekte der Langzeitarchivierung, vor allem die Migration. Insofern muss eine aktive, laufende Datenbank nicht archiviert werden.

Bei der Archivierung von Datenbanken müssen neben den Daten unbedingt auch die Datenmodelle und Datenbankschemata archiviert werden. Geeignete Formate sind SQL und CSV für Daten und Datenbankschema sowie gegebenenfalls PDF/A für das konzeptionelle Datenmodell.

6 Schlussfolgerungen, Empfehlungen

Das Ziel der digitalen Archivierung ist es, digitale Inhalte oder digitale Funktionalitäten dauerhaft zu erhalten und verfügbar zu machen.

Dabei treten jedoch einige Probleme auf. Während analoge Medien in der Regel direkt lesbar sind, können digitale Medien nur indirekt gelesen werden. Digitale Informationen erfordern ein komplexes System aus Hardware, Software, Datenträger und Datenformat, um einem Rezipienten zugänglich gemacht zu werden. Im analogen Bereich wirken Alterungsprozesse langsam, kontinuierlich und damit vorhersehbar. Im digitalen Bereich dagegen wirken spontane und kurzfristige Alterungsprozesse, die durch komplexe Systeme mit relativ geringer Redundanz noch brisanter werden. Vorteilhaft ist jedoch die Möglichkeit, digitale Informationen ohne Verlust kopieren und damit erneuern zu können.

Zur Vorsorge gibt es mehrere Methoden. Die Hardcopy-Methode eignet sich gut für einfache Dokumente, deren Inhalt per Ausdruck vollständig wiedergegeben werden kann. Die Hardcopys werden herkömmlich archiviert. Die Methode ist für komplexe Inhalte sowie für Funktionalitäten nicht geeignet. Die Erhaltungs-Methode eignet sich für Daten, die nur in ihrer originalen Systemumgebung benutzbar sind oder wenn Authentizität im Vordergrund steht. Sie ist sehr aufwendig und außerdem für eine dauerhafte Bewahrung nicht geeignet. Die Emulations-Methode ist aufwendig, zeichnet sich aber dadurch aus, dass sie Originaldaten nutzen kann. Weiterhin bietet sie gute Möglichkeiten, Funktionalitäten zu bewahren. Bei der Migrations-Methode werden alle Systemkomponenten, also auch die Daten, stets aktuell gehalten. Dies gewährleistet einen niedrigen Betreuungsaufwand. Die Aktualisierung der Daten erfordert gegebenenfalls eine Überprüfung der Authentizität.

Noch gibt es keine fertig konfektionierten Digitalarchive zu kaufen. Jedoch sind die Prinzipien, nach denen sie arbeiten müssen, recht klar, sodass man be-

ginnen kann, ein digitales Archiv selbst aufzubauen und zu betreiben. Wenn man berücksichtigt, dass nicht die Lebensdauer über den Erhalt entscheidet, sondern die Methodik, und wenn man akzeptiert, dass Archivierung keine passive Ablagerung ist, sondern ein aktiver Prozess, dann wird digitale Archivierung schon jetzt funktionieren.

Die Archivierung digitaler Daten ist eine komplexe Aufgabenstellung, die Kompetenzen sowohl aus der Informationstechnik als auch aus der Archivierung erfordert! Sorgfältige Planung und Vorbereitung spielen, wie schon bei der herkömmlichen Archivierung, auch im digitalen Bereich eine wichtige Rolle, ebenso angemessene finanzielle und personelle Ressourcen.

Archivierung ist keine passive Ablagerung, sondern ein aktiver Prozess. Nicht die Lebensdauer entscheidet über den Erhalt, sondern die Methodik. Die Art der Datenträger und Datenformate sind dabei nur sekundäre Aspekte. Eine klare Konzeption sowie angemessene finanzielle und personelle Ressourcen bilden die Grundlage für einen erfolgreichen Start in die Archivierung digitaler Daten.

7 Quellen/Verweise

Der „Ratgeber zur Archivierung digitaler Daten" wurde von der Kommission „Archäologie und Informationssysteme" beim Verband der Landesarchäologen angeregt und durch die AG Archivierung realisiert, an der folgende Kolleginnen und Kollegen beteiligt sind oder waren:

Dr. Michaela Aufleger,
LVR – Amt für Bodendenkmalpflege im Rheinland

David Bibby,
Landesamt für Denkmalpflege im Regierungspräsidium Stuttgart

Dr. Utz Böhner,
Niedersächsisches Landesamt für Denkmalpflege

Andreas Brunn,
Landesamt für Denkmalpflege und Archäologie Sachsen-Anhalt

Sigmar Fitting,
Generaldirektion Kulturelles Erbe Rheinland-Pfalz

Dr. Reiner Göldner,
Landesamt für Archäologie Sachsen

Dr. Christoph Grünewald,
LWL-Archäologie für Westfalen

Dr. Axel Posluschny,
Römisch-Germanische Kommission (RGK) Frankfurt a. M.

Dr. Bettina Stoll-Tucker,
Landesamt für Denkmalpflege und Archäologie Sachsen-Anhalt

Dr. Mathias Wilbertz,
Niedersächsisches Landesamt für Denkmalpflege

Bei der Vorbereitung dieses Ratgebers wurden folgende Quellen genutzt, die als Startpunkt für weitere Recherchen wärmstens empfohlen werden und bei deren Autoren wir uns für viele Anregungen bedanken:

ADS – Archaeology Data Service, Guides to Good Practice
(http://guides.archaeologydataservice.ac.uk/).

BAR – Schweizer Bundesarchiv – Polycy Digitale Archivierung.
(http://www.bar.admin.ch/themen/00876/index.html?lang=de&download=NHzLpZeg7t,lnp6I0NTU042l2Z6ln1acy4Zn4Z2qZpnO2Yuq2Z6gpJCDdYB,fmym162epYbg2c_JjKbNoKSn6A--).

BSI – Bundesamt für Sicherheit in der Informationstechnik. IT-Grundschutz.
(https://www.bsi.bund.de/DE/Themen/ITGrundschutz/itgrundschutz_node.html).

BSI, IT-Grundschutz-Kataloge, Baustein B 1.12 Archivierung.
(https://www.bsi.bund.de/ContentBSI/grundschutz/kataloge/baust/b01/b01012.html).

DAI – IT-Leitfaden des Deutschen Archäologischen Instituts .
(http://www.dainst.org/de/project/it-leitfaden).

DIN – Deutsches Institut für Normung.
(www.din.de).

DIN 31645, Information und Dokumentation – Leitfaden zur Informationsübernahme in digitale Langzeitarchive.

Göldner, R. Langzeitarchivierung digitaler Daten. In: Archäologisches Nachrichtenblatt, Band 14, 2/2009. Akademieverlag.

ISO 15489 „Information and Documentation – Records Management".

JHOVE – JSTOR/Harvard Object Validation Environment. (http://hul.harvard.edu/ jhove/).

kopal – Kooperativer Aufbau eines Langzeitarchivs. Deutsche Nationalbibliothek, IBM Deutschland, BMBF. (http://kopal.langzeitarchivierung.de).

Media de Lux – Laserbelichtung auf Farbmikrofilm. (http://www.mediadelux.de/).

Memorandum zur Langzeitverfügbarkeit digitaler Informationen in Deutschland. (http://deposit.ddb.de/ep/netpub/94/89/24/972248994/_data_dyna/_snap_stand_2006_07_04/www.langzeitarchivierung.de/downloads/memo2006.pdf).

NB – Schweizer Nationalbibliothek – OAIS. (http://www.nb.admin.ch/nb_professionnel/01693/01696/ 01876/ 01878/ index.html?lang=de).

nestor – Kompetenznetzwerk Langzeitarchivierung. (http://www.langzeitarchivierung.de).

nestor Handbuch: Eine kleine Enzyklopädie der Langzeitarchivierung. (http://nestor.sub.uni-goettingen.de/handbuch/index. php).

OAIS – Open Archival Information System. Referenzmodell für digitales Langzeitarchiv. (http://public.ccsds.org/publications/archive/650x0b1.pdf).

PDF/A Competence Center. (http://www.pdfa.org).

PRONOM – Online Informationssystem zu Dateiformaten. (http://www.nationalarchives.gov.uk/PRONOM/Default.aspx).
TNA – The National Archives (UK). (http://www.nationalarchives.gov.uk/).

TNA1 – Generische Anforderungen zur langzeitlichen Erhaltung elektronischer Informationen – 1. Definition der Merkmale authentischer Dokumente. (http://www.nationalarchives.gov.uk/documents/de_generic_requirements_vol1.pdf).

TNA2 – Generische Anforderungen zur langzeitlichen Erhaltung elektronischer Informationen – 2. Langzeitliche Erhaltung authentischer und zuverlässiger Dokumente: Erfordernisse hinsichtlich deren Handhabung. (http://www.nationalarchives.gov.uk/documents/de_generic_requirements_vol2.pdf).

Vereinigung deutscher Wirtschaftsarchivare e. V., Arbeitskreis „Elektronische Archivierung". (http://www.wirtschaftsarchive.de/arbeitskreise/fachliche-arbeitskreise/elektronische-archivierung).

Wikipedia. (http://de.wikipedia.org).

(Alle Internetlinks mit Stand vom 27.05.2012).

Anmerkungen

[1] Leider konnten wir nicht alle durch den aktuellen technischen Fortschritt bedingten Änderungen und Ergänzungen einbeziehen, sodass der Ratgeber im Wesentlichen den Stand von 2011/2012 widerspiegelt.

Datei-Spezifikationen (1)		
*	**ADeX**	• Archäologisches Datenexport-Format • offene Spezifikation des Verbandes der Landesarchäologen • einfaches, CSV-basiertes Format zum Austausch von Archäologieflächen
*	**CSV**	• nach RFC 4180 • Kodierung wie TXT • Spezifikation (default): - Kopfzeile verwenden - nur ein Datensatz je Zeile - Trennzeichen , ; oder | • abweichende Spezifikation explizit angeben (z.B. schema.ini) • ggf. inhaltliche Beschreibung erforderlich • ggf. Strukturbeschreibung erforderlich (ER-Modell)
?	**DNG**	• offene Spezifikation der Fa. Adobe zur vom Hersteller unabhängigen Speicherung von RAW-Bilddaten • noch nicht gut verbreitet (Photoshop, Silverfast) → **vorerst nicht nutzen**
?	**DWG**	• proprietäres Format der Fa. Autodesk (Software AutoCAD) • Spezifikation ändert sich mit neuen Releases von AutoCAD → **vorerst nicht nutzen**
*	**DXF**	• proprietäres Datenaustauschformat der Fa. Autodesk • quelloffen, aber nicht genormt • Spezifikation ändert sich mit neuen Releases von AutoCAD • textbasiertes Format, daher einfach zu bearbeiten • weit verbreitet, Quasistandard zum Datenaustausch bei CAD
?	**GeoTIFF**	• spezielle Ausprägung des TIFF-Formats für georeferenzierte Rasterbilder (spezielle Tags für Georeferenz und Koordinatenreferenzsystem) • quelloffen, aber nicht genormt • Quasi-Standard zum Austausch georeferenzierter Rasterdaten • in der Praxis sind die Daten z.T. nicht Baseline konform und sind damit oft nicht in Standard-Viewern darstellbar • stark zunehmende Verbreitung → **vorerst nicht nutzen**
*** = bevorzugtes Format, * = akzeptiertes Format, ? = Kandidat (problematisch)		

10 Dateispezifikationen für potenzielle Archivierungsformate.

Datei-Spezifikationen (2)		
?	**GML**	• standardisiertes XML-basiertes Geodatenformat • offene Spezifikation des Open Geospatial Consortium (OGC) • GML 3.2 ist ISO 19136 und berücksichtigt ISO 19107 • leider noch nicht gut verbreitet → **vorerst nicht nutzen**
*	**JPEG**	• ISO/IEC 10918-1 zertifiziert • verlustbehaftete Komprimierung • einheitlicher Farbraum YCbCr (CCIR 601) • bei hoher Auflösung visuell keine Verluste bemerkbar • sehr weit verbreitet, gute Akzeptanz
?	**JPEG 2000**	• ISO 15444 zertifiziert • speziell: verlustfreie Komprimierung benutzen! • noch nicht gut verbreitet (PlugIn für Photoshop, med. DICOM) → **vorerst nicht nutzen**
?	**KML**	• proprietäres Geodatenformat der Fa. Google • XML-basiertes Geodatenformat, offener OGC-Standard • neben Google-Maps/Earth nicht weit verbreitet → **vorerst nicht nutzen**
?	**MIF/MID**	• Geodaten-Austauschformat der Fa. MapInfo Corporation • quelloffen, aber nicht genormt • textbasiert (2 Dateien für Geometrie- und Sachdaten) • neben MapInfo wenig verbreitet → **vorerst nicht nutzen**
***	**PDF/A**	• nach ISO 19005-1 • min. Konformitätsgrad 1b (weniger streng), 1a empfohlen • mgl. Adobe-PlugIn verwenden oder Format validieren
*	**SHP**	• proprietäres Geodatenformat der Fa. ESRI (Software ArcGIS) • quelloffen, aber nicht genormt • Satz von 3...9 Dateien für Geometrie- und Sachdaten u.a. • weit verbreitet, Quasi-Datenaustauschstandard bei GIS
*	**SQL**	• standardisierte Datenbanksprache zur Definition, Manipulation und Steuerung relationaler Datenbanken • ISO/IEC 9075 • textbasiert,für Datenexport und Migration geeignet • weit verbreitet
*** = bevorzugtes Format, * = akzeptiertes Format, ? = Kandidat (problematisch)		

11 Dateispezifikationen für potenzielle Archivierungsformate.

Datei-Spezifikationen (3)		
***	**TIFF**	• Version 6, Baseline Format • proprietäre Erweiterungen werden ignoriert (Adobe-Enhancements u.a.) • Komprimierung - unkomprimiert - SW-Bilder auch G4 (verlustfrei nach ITU/CCITT) • keine Mehrseiten-Bilder (Multipage-TIFF), keine Ebenen • 24...48 Bit Farbe, 8...16 Bit SW • Auflösung entspricht Inhalt
*	**TIFF-LZW**	• wie TIFF, aber mit LZW-Komprimierung (die Komprimierung war zeitweilig lizenziert, ist nun aber frei)
***	**TXT**	• für unstrukturierten und unformatierten Text • Kodierung/Zeichensatz: - US-ASCII (keine dt. Sonderzeichen!) - Unicode UTF-8 (oder UTF-16) - ISO 8859-1(Latin-1 westeuropäisch) • bei strukturierten Texten → CSV, XML • bei formatierten Texten → PDF-A, TIFF
?	**XML**	• vgl. Extensible Markup Language (XML) 1.0 (Fifth Edition), W3C Recommendation 26 November 2008 (http://www.w3.org/TR/2008/REC-xml-20081126/) • textbasiert, hierarchisch strukturiert • XML-Formate benötigen i.d.R. eine spezielle Präsentation (adäquate Software oder Spezialwissen erforderlich) • inhaltlich-strukturelle Beschreibung erforderlich • ggf. spezielle Sprache separat betrachten, z.B. XML für Text + Layout → ODF ISO/IEC 26300 (Open Office) → **vorerst nicht nutzen**
*** = bevorzugtes Format, * = akzeptiertes Format, ? = Kandidat (problematisch)		

12 Dateispezifikationen für potenzielle Archivierungsformate.

Der Einsatz des Web-Standards SVG in der Archäologie

Ralf Berger

Der vorliegende Artikel beschreibt den Einsatz von SVG (Scalable Vector Graphics) in der GIS-Komponente der ADABweb, dem Denkmalpflege-Informationssystem der Länder Baden-Württemberg und Niedersachsen. Hier werden Informationen zu Lage, Ausdehnung und Denkmaleigenschaften von Boden- und Baudenkmalen grafisch aufbereitet und auf dem Hintergrund von zum Beispiel topografischen oder Luftbild-Karten dargestellt.

Die Geometrien archäologischer Objekte lassen sich in der Regel als Flächen, Punkte oder Linien darstellen. Die Schutzflächen von Grabhügelfeldern sind Beispiele für Flächengeometrien. Objekte mit punktförmigen Geometrien sind zum Beispiel Einzelfunde oder Objekte mit unsicherer Ortsangabe. Deiche und Landwehren sind zwar eigentlich flächenhafte Objekte, werden aber häufig als Linien dargestellt. Für diese Arten von Geometrien eignen sich Vektorgrafikformate (anders als Rastergrafiken), da sie über einen großen Maßstabsbereich ohne Qualitätsverlust – natürlich abhängig von der Qualität der Erfassung – dargestellt werden können.

Bei der Konzeption der ADABweb wurden die folgenden Anforderungen an die GIS-Komponente entwickelt:

- eine geeignete Menge von Grafikelementen und die einfache Generierbarkeit aus den in der Datenbank gespeicherten Geometrien (zum Beispiel durch ein XML-basiertes Format),

- eine geeignete Menge von Darstellungsattributen sowie die Möglichkeit, gleichartigen Objekten gemeinsame Attribute zuweisen zu können,

- eine wesentliche Anforderung des Niedersächsischen Landesamtes ist die Möglichkeit, Punkt- und Liniensignaturen (zum Beispiel Wellenlinien für Deiche) zur Darstellung archäologischer Objekte zu verwenden,

- die Möglichkeit, Grafiken aus anderen Quellen, zum Beispiel topografische Hintergrundkarten, einbinden zu können,

- neben der Darstellung im Webbrowser werden Werkzeuge zur Umwandlung in Rastergrafikformate benötigt,

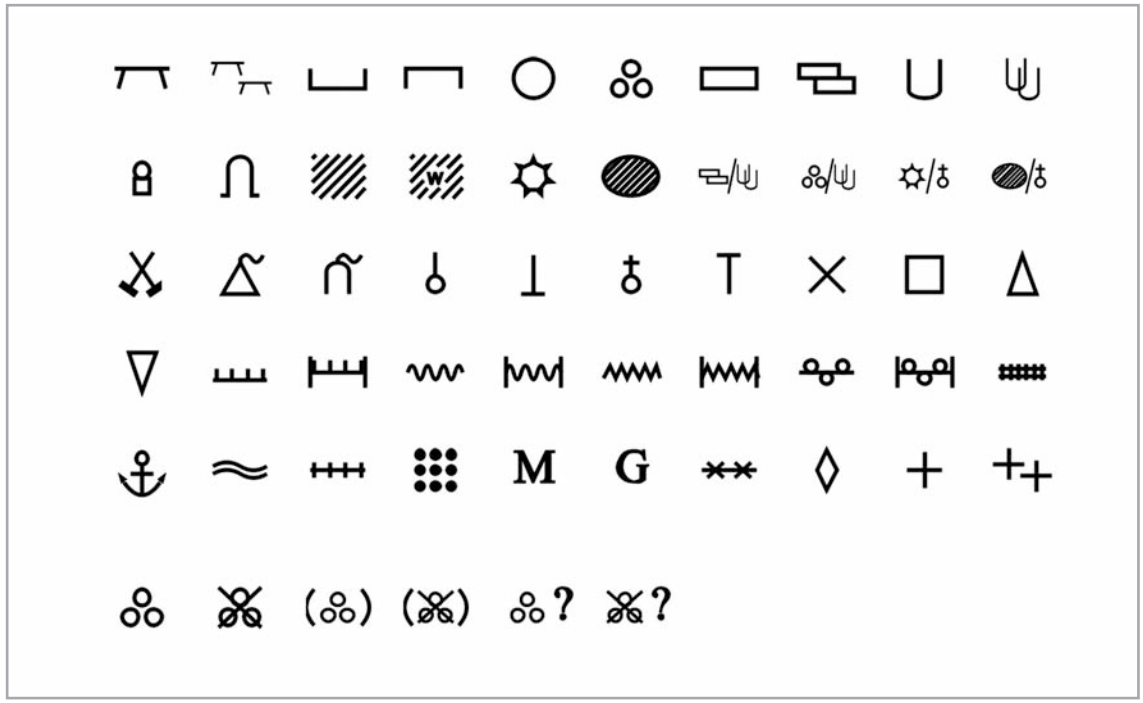
- die einfache Erreich- und Manipulierbarkeit der Grafikelemente, um die Grafik interaktiv ohne Serverinteraktion verändern zu können; zum Beispiel zur Änderung der Darstellung (Farben, Liniendicken) oder zum Bearbeiten von Geometrien.

SVG ist ein XML-basiertes 2-D-Vektorgrafik-Format, das vom World Wide Web Consortium (W3C) für den Einsatz in Web-Anwendungen entwickelt wurde. Es erfüllt diese Anforderungen, wie im Folgenden gezeigt wird.

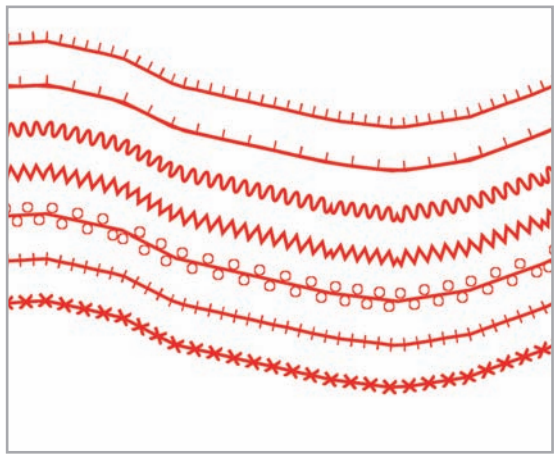
Grafikelemente

SVG kennt drei Arten von grafischen Elementen: Vektor-Grafik-Elemente, (Raster)-Bilder und Text. Bilder können Raster-Bilddateien (zum Beispiel JPEG und PNG) oder andere SVG-Dateien sein; sie können aus einer lokalen Datei stammen oder über das Web bereitgestellt werden, zum Beispiel durch einen Web-MapService (WMS). Als elementare Vektor-Grafik-Elemente stellt SVG zur Verfügung: Rechteck, Kreis, Ellipse, Linie und Polygon.

In der ADABweb werden die Fachdaten in einer PostgreSQL-Datenbank gespeichert; dabei werden die Geodaten mittels der Erweiterung PostGIS bearbeitet – beides OpenSource-Produkte, die inzwischen seit vielen Jahren gepflegt und weiterentwickelt werden. Konzeptuell ist ADABweb nicht auf ein Datenbank-Management-System festgelegt; tatsächlich basierten die ersten Versionen der ADABweb auf Microsoft-SQL-Server- beziehungsweise Oracle-Datenbanksystemen. Diese boten jedoch eine Unterstützung für Geodaten zum Zeitpunkt der Entwicklung der GIS-Komponente der ADABweb entweder gar nicht (SQL-Server) oder nur gegen erheblichen finanziellen Aufwand. PostGIS bietet eine umfangreiche Menge von Funktionen zur Manipulation geografischer Objekte sowie räumliche Indices, mit denen sich Objekte innerhalb eines gegebenen Gebiets effizient auffinden lassen. Außerdem gibt es Funktionen zum Import/Export von Shape-Dateien für den Austausch von Geo-Daten mit anderen GIS-Systemen und zur Ausgabe von Objektgeometrien im SVG-Format oder KML (Google Maps).



1 Der Signaturkatalog des Niedersächsischen Landesamtes für Denkmalpflege (NLD); darunter exemplarisch die Grabhügel-Signatur in allen Ausprägungen: erhalten, zerstört, transloziert, transloziert und zerstört, fraglich, fraglich und zerstört.



2 Die Linien-Signaturen, die sich dem Kurvenverlauf anpassen.

Darstellung

Die Darstellung der Grafikelemente in SVG wird – wie bei HTML selbst – per Cascaded Style Sheets (CSS) unter anderem für die folgenden Attribute festgelegt: Farbe (für Umrandung und Fläche), Liniendicke, Deckkraft, Schraffur für Flächen, Linienmuster (Strichlierung) für Linien und Flächenumrandungen sowie Schriftart und -größe für Beschriftungen. Änderungen der Werte dieser Stil-Attribute werden vom Browser unmittelbar dargestellt, sodass die Darstellung der Grafik ohne Neu-Generierung der SVG-Datei durch den Server vom Benutzer angepasst werden kann.

Grafikelemente können zu Gruppen zusammengefasst werden, deren Darstellung – insbesondere die Sichtbarkeit – für diese gemeinsam festgelegt werden kann. So können Fachdaten-Layer definiert werden, für die aus der Datenbank gleichartige Objekte wie zum Beispiel Ausdehnungs- und Schutzflächen der Archäologie oder Objekte der Baudenkmalpflege ausgewählt werden und die gemäß einer vom System oder dem Benutzer vorgegebenen Konfiguration dargestellt werden.

Signaturen

SVG ermöglicht es, sogenannte „Symbole" zu definieren, das sind Vorlagen für Grafikelemente, die aus den oben genannten einfachen Grafikelementen sowie zuvor definierten Symbolen zusammengesetzt werden. Auf diese Weise können Vorlagen für Signaturen archäologischer Objekte (zum Beispiel für Grabhügel, Landwehren usw.) entworfen werden. Für das Niedersächsische Landesamt für Denkmalpflege wurde ein Katalog mit 53 Punkt-Signaturen implementiert (Abb. 1). Für jede Signatur existieren Varianten, die das Objekt als „fraglich", „zerstört" oder „transloziert" markieren. Außerdem gibt es sieben Liniensignaturen zur Darstellung von linearen Objekten wie zum Beispiel Gräben; hier werden entweder zusätzlich zum Linienverlauf Symbole ausgegeben (zum Beispiel die alternierend an der Linie verlaufenden Kringel bei Wallhecken) oder die Linie selbst wird

nicht gezeichnet, sondern dient nur zur Definition des Verlaufs einer Wellenlinie, zum Beispiel bei der Liniensignatur für Deiche (Abb. 2) Diese Eigenschaft von SVG ist auch heute noch einzigartig.

Hintergrund-Grafiken

Hintergrundkarten können in SVG mit dem Image-Element in die Darstellung eingebunden werden. Sie werden durch WebMapServices (WMS) zum Beispiel vom jeweiligen Landesamt für Geoinformation bereitgestellt. Hierbei handelt es sich zunächst einmal um topografische Karten in den verschiedenen Maßstäben (1:200.000 – 1:25.000 bis hinunter zur Amtlichen Karte 1:2.500). Zusammen mit den Signaturen der archäologischen Objekte und gegebenenfalls Beschriftungen ergibt sich eine Darstellung wie beispielhaft in Abbildung 3 gezeigt.

Daneben können natürlich auch andere Hintergrundkarten wie zum Beispiel historische Flurkarten, Ortho-Luftbilder und LIDAR-Daten (eine plastische Ansicht der Erdoberfläche, die Erhebungen und Vertiefungen hervorhebt) eingebunden werden. Hintergrund-Layer, die keine flächendeckende Darstellung liefern (zum Beispiel Gewann- und Straßennamen sowie Hausnummern), werden in einem Format eingebunden, das Transparenz ermöglicht (zum Beispiel PNG), sodass darunterliegende Karten-Layer sichtbar bleiben. Mehrere Hintergrund-Layer können durch geeignete Einstellung des Deckkraft-Attributs übereinander dargestellt werden. So entsteht zum Beispiel durch Kombination einer topografischen Karte mit einem LIDAR-Layer eine dreidimensional anmutende Ansicht (Abb. 4).

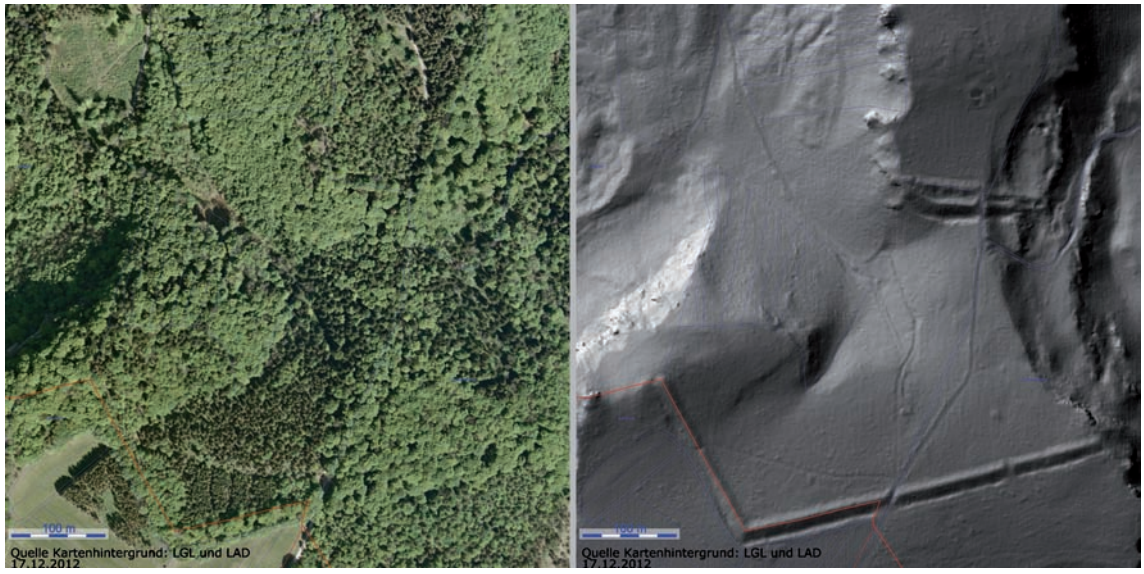
Austausch

Aus einer SVG-Datei kann nicht nur mit einem Webbrowser eine Bildschirmdarstellung erzeugt werden, sondern auch mit geeigneter Software eine Rastergrafik, zum Beispiel JPEG, PNG oder auch PDF. ADABweb verwendet dazu Apache Batik, eine freie Software der Apache Software Foundation.

Interaktivität

Natürlich reicht die Darstellung der Fach- und Geodaten im Webbrowser allein nicht aus. In SVG wird Interaktivität durch JavaScript realisiert. Alle Grafikelemente (einfache sowie komplexe und Gruppierungen) können über das Document Object Model (DOM) angesprochen und modifiziert beziehungsweise neu hinzugefügt werden. Das Document Object Model ist eine vom W3C entwickelte Spezifikation zum Zugriff auf Elemente in einem XML-Dokument.

3 Durch Liniensignaturen dargestellte Landwehren in der Nähe von Ganderkesee; Quelle: NLD und Landesamt für Geoinformation und Landentwicklung Niedersachsen (LGLN).



4 Vergleich Ortho-Luftbild zu LIDAR-Hintergrund einer bronzezeitlichen Befestigungsanlage auf der Zollernalb; Quelle: Landesamt für Denkmalpflege Baden-Württemberg und Landesamt für Geoinformation und Landentwicklung Baden-Württemberg.

So werden eine Reihe interaktiver Werkzeuge möglich, mit der die Darstellung angepasst und auch Geometrien erfasst werden können. Da sind zunächst einmal die üblichen Navigationshilfen, mit denen der Kartenausschnitt verschoben und vergrößert beziehungsweise verkleinert dargestellt werden kann. Mit einem Klick in die Karte können aber auch – je nach aktiviertem Modus – das Datenblatt eines Fachobjekts oder topografische Informationen abgerufen werden: die Koordinate (Gauß-Krüger, WGS84, UTM), der Raumbezug (Flurstück, Gemarkung, Gemeinde, Kreis, gegebenenfalls Anschrift) sowie die Nummern der TK25- und DGK5-Kartenblätter.

Die DOM-Unterstützung durch SVG ermöglicht auch die Entwicklung eines Werkzeugs, mit dem Punkt-, Linien- und Flächen-Geometrien durch Hinzufügen, Verschieben und Löschen von einzelnen Punkten oder durch Übernahme von Geometrien aus geeigneten Hintergrundlayern („Pick-Funktion") erzeugt oder verändert werden. Aus diesem Geometrie-Werkzeug entstehen Werkzeuge zur Bearbeitung von Objektgeometrien, zur Objektauswahl sowie zum Vermessen von Linienzügen und Flächen.

Entwicklung

Das W3C hat bereits 2001 die erste Spezifikation verabschiedet. Daher lag die Erwartung nahe, dass die Browser-Hersteller diesen Standard dann auch in ihren Produkten implementieren. Tatsächlich unterstützte jedoch zu diesem Zeitpunkt noch kein Webbrowser SVG direkt; stattdessen konnten SVG-Grafiken nur durch ein von Adobe entwickeltes Plug-in im Webbrowser dargestellt werden. Für alle gängigen Browser außer dem Microsoft Internet Explorer (IE) sind in der Folge auch entsprechende Implementierungen erfolgt. So wird SVG vom Mozilla Firefox seit der im November 2005 erschienenen Version 1.5 unterstützt, später mit deutlich verbesserter Standard-Konformität ab der Version 3.0 (Juni 2008).

Microsoft dagegen hat im Internet Explorer bisher das eigene Vektorgrafik-Format VML unterstützt, das im Standardisierungsprozess mit einem Konkurrenz-Format (von Adobe und anderen) zu SVG entwickelt wurde. Erst der IE9 unterstützt das SVG-Format direkt. In allen vorhergehenden IE-Versionen funktioniert aber das Adobe Plug-in; erstaunlich, wenn man bedenkt, dass Adobe dessen Entwicklung nach der Veröffentlichung einer Betaversion 2003 eingestellt hat. So konnte die GIS-Komponente der ADABweb zunächst nur im IE (Version 6 aufwärts) mit dem Adobe Plug-in verwendet werden, was jedoch in der Praxis keine starke Einschränkung darstellte, da für die Anwender in den Landesämtern der IE als Standardbrowser vorgesehen war (und noch ist). Aktuelle Zugriffsstatistiken zeigen einen Anteil des Internet Explorers von etwa 60% (vor allem IE7 mit wachsendem Anteil des IE8), knapp 30% für den Firefox und circa 5% für Opera. Erst mit einer allmählich einsetzenden Öffnung des Zugangs zur ADABweb in Niedersachsen in den letzten Jahren ergibt sich die Notwendigkeit, andere Browser stärker zu berücksichtigen. Da diese aber allesamt SVG unterstützen, ergeben sich hier allenfalls kleinere Inkompatibilitäten im Bereich der Interaktivität.

Fazit

Von den Nutzern werden natürlich immer wieder zusätzliche Funktionen nachgefragt. Zu den am häufigsten genannten Weiterentwicklungen gehört die Möglichkeit, eigene Vektor- oder Rasterdaten in die Kartierung einbinden zu können, um Objektgeometrien in die ADABweb übernehmen zu können. Dies können georeferenzierte Rasterkarten (zum Beispiel Grabungskarten) sein, von denen dann die Geometrie mit dem Kartierungseditor nachgezeichnet werden kann. Oder es handelt sich um Vektordaten wie Shape-Dateien oder GPS-Tracks, deren Geometrien mittels der oben beschriebenen Pick-Funktion in die ADABweb übernommen werden sollen.

Die Entscheidung zum Einsatz von SVG (wie bei jeder Technologie) wird natürlich immer wieder hinterfragt; insbesondere in der Zeit, in der die Unterstützung des nicht mehr weiterentwickelten Adobe Plug-ins auszulaufen drohte und der Internet Explorer SVG noch nicht unterstützte. Die Alternativen zeigen entweder regelmäßig Sicherheitsprobleme und müssten auf den Client-Systemen häufig aktualisiert werden, was aber in großen IT-Domänen in der Regel die Nutzer nicht dürfen und daher unterbleibt (Flash). Oder sie sind proprietäre Produkte (MS Silverlight) beziehungsweise es fehlen wichtige Eigenschaften wie zum Beispiel die Fähigkeit zur effektiven Definition eines Signaturen-Katalogs. Dank der inzwischen flächendeckenden Unterstützung durch die Webbrowser hat sich die Entscheidung für SVG als tragfähig erwiesen.

Literatur

SVG-Recommendation des W3C:
http://www.w3.org/TR/SVG11/

archaeoDox – Standardsoftware auf Ausgrabungen in Schleswig-Holstein, der Freien und Hansestadt Hamburg und im Landkreis Harburg

Jörg Räther, Eicke Siegloff

Der primäre Leitgedanke der Bodendenkmalpflege liegt bekanntermaßen im Erhaltungsgebot archäologischer Fundstellen. In der Realität ist dies eher selten umsetzbar, denn nur allzu oft fordern höherrangige Interessen eine Sicherung des archäologischen Erbes in Form von Rettungsgrabungen. Da eine Ausgrabung eine – wenn auch kontrollierte – Zerstörung archäologischer Substanz bedeutet, muss der umfassenden und lückenlosen Dokumentation von Grabungsbefunden und Funden sowie der dauerhaften Aufbewahrung des Archivgutes als Grundlage späterer wissenschaftlicher Bearbeitungen die allerhöchste Sorgfalt beigemessen werden. Der technologische Fortschritt ging natürlich auch an diesem Arbeitsfeld nicht vorüber, sodass seit nunmehr etlichen Jahren digitale Dokumentationsverfahren verstärkt ihre Anwendung finden, dass analoge Verfahren zwar in ein Schattendasein gedrängt scheinen, beispielsweise verbale Beschreibungen aber niemals ersetzbar sein werden. Der Fokus digitaler Dokumentation liegt vor allem auf dem Einsatz immer ausgefeilterer und damit hochtechnisierter Verfahren. Als Kehrseite dieser Errungenschaften potenziert sich hierdurch allerdings auch das zu archivierende Datenvolumen, ein Problem, das oftmals in den Hintergrund verdrängt wird. Um der immer weiter steigenden Datenflut verwalterisch Herr zu werden beziehungsweise zu bleiben, bieten sich besonders Datenbanksysteme an, deren Einsatz prinzipiell schon im Bereich der Feldtätigkeit beginnen sollte. Dazu, wie solch eine Grabungsdatenbank auszusehen hat und welche Funktionen sie erfüllen soll, hat jeder, der schon einmal mit einer oder mehreren gearbeitet hat, eigene Vorstellungen. Plant man eine Datenbank, so ist für das Datenmodell der Einsatzzweck entscheidend. Eine Grabungsdatenbank für die Bodendenkmalpflege, in welcher von der Fundmeldung bis zur mehrjährigen archäologischen Untersuchung eine Vielzahl unterschiedlicher Maßnahmen abgebildet werden soll, wird sich deshalb in einigen Punkten von einer Grabungsdatenbank eines speziellen Forschungsprojektes unterscheiden müssen. Zum einen muss angestrebt werden, möglichst homogene Datenbestände und Dateiablagen zu erzeugen, zum anderen muss die Datenerfassung aber flexibel genug sein, damit die Anwender auf die jeweiligen Bedingungen vor Ort reagieren können und nicht durch das Hilfsmittel Grabungsdatenbank eingeschränkt werden.

Die seit 2010 von derbrill IT-Service in Zusammenarbeit mit dem Archäologischen Landesamt Schleswig-Holstein (ALSH) und seit 2011 mit dem Archäologischen Museum Hamburg (AMH), zuständig für die Bodendenkmalpflege der Freien und Hansestadt Hamburg und des niedersächsischen Landkreises Harburg, entwickelte Grabungsdatenbank archaeoDox basiert auf dem Open-Source Datenbankmanagementsystem PostgreSQL. Das Entwicklungsziel war neben der Erfassung und Verwaltung archäologischer Daten auch deren Langzeitarchivierung. Hierbei setzt archaeoDox auf XML als Standardspeicherformat. ArchaeoDox umfasst mit archaeoDox-Client, archaeoDox-Connect und archaeoDox-Server drei Komponenten, welche es gestatten, den Einsatz der Software an die personellen und räumlichen Ressourcen unterschiedlicher Anwender anzupassen. ArchaeoDox-Client ermöglicht den Einzelplatzbetrieb auf einem Grabungsnotebook oder Desktop-PC und dient der Datenerfassung. Mittels archaeoDox-Connect lassen sich die Einzelplatzinstallationen von archaeoDox sowohl mit einem zentral betriebenen archaeoDox-Server als auch zu einem Ad-hoc-Netzwerk untereinander verbinden. In beiden Fällen besteht so die Möglichkeit der kollaborativen Datenerfassung und -bearbeitung durch den gemeinsamen Zugriff auf einen zentralen Speicherort. Darüber hinaus stellt archaeoDox-Server Werkzeuge zur Administrierung der Datenbank und Schnittstellen für den Datenaustausch mit anderen Datenbanksystemen bereit. Hieraus ergeben sich in der Praxis zahlreiche mögliche Einsatzszenarien. So ist im Archäologischen Landesamt Schleswig-Holstein derzeit jeder Grabungsleiter im Außeneinsatz mit einem archaeoDox-Client zur Datenerfassung ausgestattet. Bereits während der Maßnahme zur Datensicherung beziehungsweise endgültig nach deren Abschluss werden die aufgenommenen Grabungsdaten an einen zentralen Dateneingang im ALSH übergeben. Dort werden die zur Maßnahme gehörigen Dateien in einem festgelegten und nach Aktivitäten sortierten Ablagesystem auf einem File-Server gespeichert, während die lokale Grabungsdatenbank des Grabungsnotebooks mit dem archaeoDox-Server synchronisiert wird. Die Dateien und die Grabungsdatenbank als zugehöriger Aktenplan stehen dann den zuständigen Dezernenten und weiteren Bearbeitern zentral zur Verfügung. Im Archäologischen Museum Hamburg werden die Daten einer Maßnahme häufig nicht allein vom örtlichen Grabungsleiter, sondern von mehreren Mitarbeitern erfasst. So erfolgt die Befundaufnahme in der Regel durch einen Wissenschaftler, während die technischen Objekte, wie zum Beispiel Fotolisten, welche

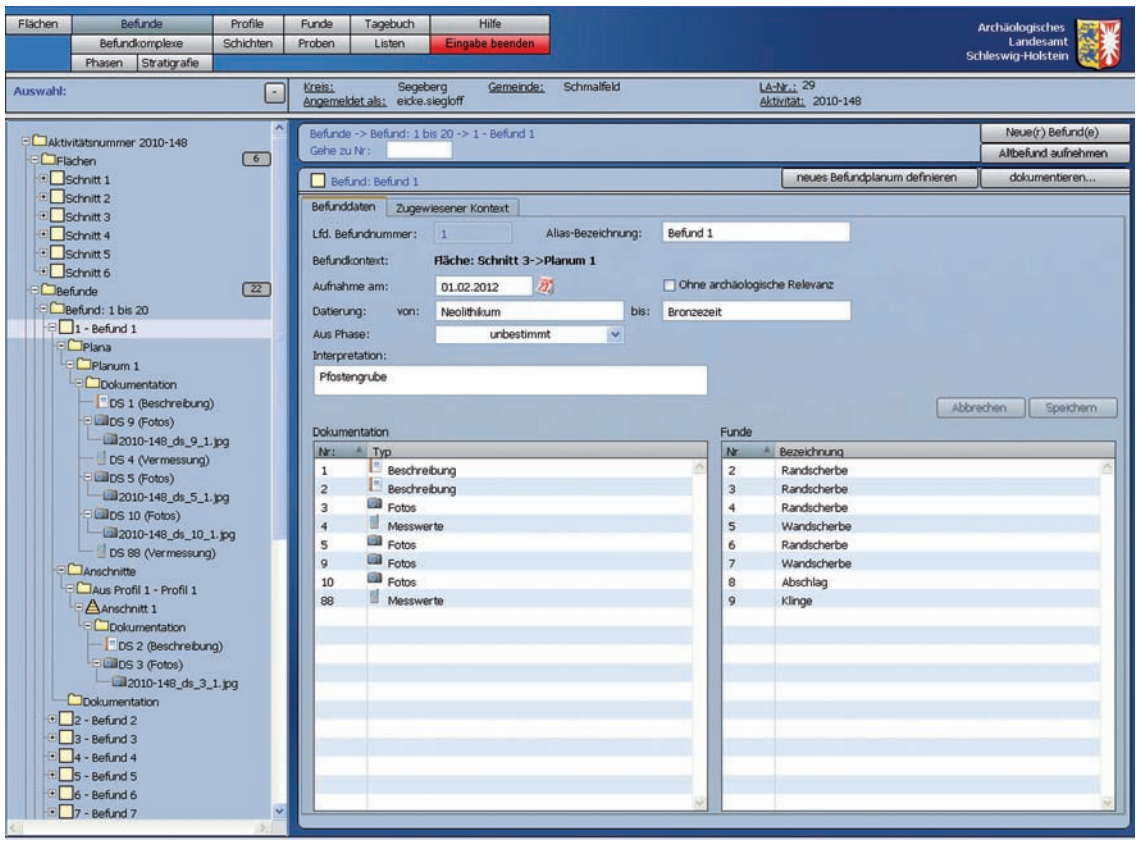
nicht mehr separat geführt zu werden brauchen, wie bisher durch die Grabungstechniker erfasst werden. Die Schnellinventarisation der Funde, teilweise durch studentische Hilfskräfte, erfolgt wiederum häufig erst nach Beendigung der Maßnahme und erfolgter Reinigung der Funde an einem der Büroarbeitsplätze im AMH. Bei der Datenerfassung im AMH erfolgt dabei der Zugriff auf die Datenbank mittels archaeoDox-Client direkt auf dem archaeoDox-Server, was den Vorteil hat, dass allen Bearbeitern ständig die aktuelle Version der Daten zur Verfügung steht, ohne dass bei einem Wechsel des Bearbeiters oder des Büroarbeitsplatzes zuerst die Daten von Client und Server synchronisiert werden müssten. Auf der zur Zeit durch das AMH durchgeführten Ausgrabung in der Harburger Schloßstraße, welche mit Bürocontainern als temporäre Außenstelle eingerichtet ist, können alle Bearbeiter jederzeit per WLAN auf die lokale Grabungsdatenbank zugreifen.

Sowohl in Schleswig-Holstein als auch in Hamburg erhalten alle archäologischen Maßnahmen eines Kalenderjahres von der einfachen Fundmeldung bis hin zur Ausgrabung zu ihrer eindeutigen Identifikation eine aus dem Jahr der Maßnahme und einer fortlaufenden Nummerierung bestehenden Aktivitätsnummer, die auf übergeordneter Ebene vergeben wird (s. u.). Diese Aktivitätsnummer wird für die zu dokumentierende Maßnahme in archaeoDox eingegeben und ordnet als Primärschlüssel alle Objekte einer Aktivität eindeutig zu. Zusätzlich werden in den archaeoDox-Kopfdaten weitere Informationen wie Art der Maßnahme, Gemarkung und Flurstück etc. aufgenommen. Alle in einer Aktivität erfassten Objekte werden innerhalb ihrer jeweiligen Objektgruppe automatisch fortlaufend nummeriert, da nur so die Eindeutigkeit der Objekte gewährleistet werden kann. Hierbei spielt es keine Rolle, ob es sich bei der jeweiligen Aktivität um eine Notgrabung, Baustellenbeobachtung etc. und bei den Objekten um Grabungsschnitte, Befunde, Funde, Profile oder anderes handelt. Zugleich besteht aber für jedes Objekt die Möglichkeit, eine parallel geführte Alias-Bezeichnung, welche das Objekt im Zusammenhang mit der konkreten Maßnahme unter Umständen besser erklärt, wie zum Beispiel „*Baggerplanum*" zusätzlich zur vom System zugewiesenen festen Bezeichnung „Planum 1" zu vergeben.

Für die Erfassung der Grabungsdaten stehen derzeit zwei Module innerhalb von archaeoDox zur Verfügung, von denen das eine auf die Besonderheiten von Flächengrabungen, das andere auf diejenigen von Stadtkerngrabungen ausgerichtet ist. Die Notwendigkeit hierzu ergibt sich aus den im Regelfall ganz andersartigen Charakteristika von Flächen- und Stadtkerngrabungen. Flächengrabungen zeichnen sich häufig durch innerhalb großer Flächen isoliert auftretende Befunde aus, während in Stadtkerngrabungen die Flächen der Bodenaufschlüsse oftmals so klein sind, dass die Grenzen der Befunde vielfach seitlich in den Profilwänden verschwinden und ihre Interpretation damit zum Teil erheblich erschwert ist. Den Modulen Stadtkern- und Flächengrabung liegt als hauptsächlicher Unterschied ein diesen Gegebenheiten angepasster Befundbegriff zugrunde. Während im Modul Flächengrabung zum Beispiel eine Bestattung mit Grube, Sargkante, Leichenschatten usw. als ein Befund aufgefasst wird, wird im Modul Stadtkerngrabung jede differenzierbare Schicht als Befund geführt. Außerdem besteht im Modul Flächengrabung im Unterschied zum Modul Stadtkerngrabung die Möglichkeit, lediglich die Sequenzen mit Befunden ansonsten befundleerer Profile in Form von einzelnen Anschnitten zu dokumentieren, wie es vielfach bei Profilzeichnungen üblich ist. Für die weitere Befundaufnahme gibt es in beiden Modulen die Möglichkeit, Befunde zu Befundkomplexen zu ordnen. Weiterhin können mit einem Zuordnungswerkzeug, welches über eine interne Logikprüfung verfügt, die stratigrafischen Beziehungen der Befunde angegeben werden. Um den Programmieraufwand in diesem Bereich zu minimieren, erfolgt hieraus keine direkte Darstellung einer Harris-Matrix in archaeoDox. Stattdessen wurde entschieden, die stratigrafischen Daten als CSV-Datei zu exportieren und die Harris-Matrix mittels der Software Stratify, welche von Irmela Herzog, LVR-Landesmuseum Bonn, speziell für die Erstellung von Harris-Matrizes entwickelt wurde und kostenlos zur Verfügung steht, darzustellen.

Jedes Objekt einer archäologischen Maßnahme kann, nachdem es innerhalb des archaeoDox-Vorgangs einmalig erstellt wurde, in vielfältiger Weise dokumentiert werden. In inhaltlicher Hinsicht lassen sich im Laufe des Grabungsfortschritts die weiteren Beziehungen eines Objektes dessen Kontext hinzufügen, zum Beispiel dem Befund xy, welcher zunächst in Planum 1 festgestellt wurde, im späteren Verlauf der Grabung auch das Planum 2 oder das Profil A-B. In technischer Hinsicht können zu jedem archäologischen Objekt der Grabung, egal ob Fund, Befund, Probe oder anderes, wiederum eine Reihe an technischen Objekten in Form von Dokumentationsschritten, nämlich Fotos, Zeichnungen, Messwerte, Beschreibungen, Dokumente und GIS-Dateien hinzugefügt werden. Hierbei werden die einzelnen Dokumentationsschritte vergleichbar der Systematik des sogenannten Stellenkartensystems fortlaufend durchgezählt. Auf diese Weise entsteht sukzessiv eine komplette Registratur aller archäologischen und technischen Objekte der jeweiligen zu dokumentierenden archäologischen Aktivität.

Den durch Dokumentationsschritte erfassten technischen Objekten innerhalb der Datenbank entsprechen reale Objekte, welche zum Teil in analoger Form, als Handzeichnung oder Dia-Positiv etc., mit zunehmender Digitalisierung der archäologischen Dokumentationsmethoden aber vornehmlich als Dateien, CAD-Pläne, digitale Fotos usw. vorliegen. ArchaeoDox ist also als eine Art hybrides Dokumentationsmanagementsystem zu verstehen, das sowohl analoges wie auch digitales Archivgut verwaltet. Genau wie bislang für die analoge Dokumentation benötigt man auch für die digitale Dokumentation einen Aktenplan, das heißt eine verbindliche Beschreibung der Ordnerstruktur und der Dateibenennung, welche in einheitlicher Weise für die Dateiablage genutzt werden soll. Damit diese Dateiablage nicht händisch durch den jeweiligen Bearbeiter eingerichtet und gepflegt werden muss, legt archaeoDox bei der Erstellung eines Dokumentationsschritts innerhalb des Hauptordners der Aktivität automatisch einen Ordner für jeden Dokumentationsschritt an, welcher mit der Aktivitätsnummer, einem Kürzel für die Art des Dokumentationsschritts sowie der fortlaufenden Nummerierung der Dokumentationsschritte der Aktivität einheitlich beschriftet wird und zur Aufnahme der zum jeweiligen Dokumentationsschritt gehörigen Dateien dient. Dieser Ordner lässt sich zum einen während dessen Erstellung aus dem jeweiligen Dokumentationsschritt in archaeoDox heraus direkt öffnen und befüllen. Zum anderen werden die Dokumentationsschrittordner und die in diesen enthaltenen Dateien in der Baumansicht der Aktivität auf der linken Seite der Eingabemaske des archaeoDox-Clients (Abb. 1), in welcher alle Objekte in Form eines auf- und zuklappbaren Baumdiagramms dargestellt werden, angezeigt. Bei der Auswahl einer Datei aus einem Dokumentationsschrittordner in der Baumansicht kann archaeoDox den Befehl zum Öffnen der Datei an ein externes Programm weitergeben, im Falle eines CAD-Plans zum Beispiel an das entsprechende CAD-Programm, oder im Falle von Fotos sogar eine Vorschau der Bilder anzeigen. Mit diesen Fähigkeiten ist archaeoDox nicht nur eine bloße Registratur aller Objekte einer archäologischen Ausgrabung, sondern zugleich Dateimanager und -browser für die zugehörigen Dateien. Das Verlassen von archaeoDox und ein Umweg über einen Dateibrowser sind sowohl für die Erzeugung und Pflege der Dateiablage als auch für das spätere Auffinden und Öffnen der Dateien nicht mehr notwendig. Die erzeugte Dateiablage ist nicht nur einheitlich, sondern auch flach. Dies hat gegenüber stark verschachtelten Ordnerstrukturen den Vorteil der besseren Übersichtlichkeit und hilft zugleich, überlange Dateinamen zu vermeiden. Des Weiteren



1 Bildschirmansicht der Grabungssoftware archaeoDox mit dem Übersichtsblatt eines Befundes und dem Übersichts- und Navigationsbaum der Grabung links im Bild.

verfügt archaeoDox über die Fähigkeit, Dateien innerhalb von Dokumentationsschrittordnern nach einem einheitlichen Schema automatisiert umzubenennen, was im Falle von digitalen Fotos und anderen Dokumenten wie beispielsweise PDF-Dokumenten bereits angewendet wird.

Parallel zur Entwicklung von archaeoDox begannen seit Herbst 2009 das Archäologische Landesmuseum Schloss Gottorf (ALM) und das Archäologische Landesamt Schleswig-Holstein zusammen mit der digiCULT Verbund e. G. die Konzipierung eines gemeinsamen webbasierten Datenbanksystems „Archäologische Datenbank Schleswig-Holstein", das hauptsächlich der Verwaltung von Funden und archäologischen Aktivitäten dient, in weiteren Ausbaustufen unter anderem auch ein Leihmodul und ein Restaurierungsmodul vorsieht. Das Archäologische Museum Hamburg ging bereits noch vor der technischen Umsetzung durch die digiCULT Verbund e. G. eine Kooperation mit den beiden schleswig-holsteinischen Institutionen ein, da das inhaltliche Konzept, technisch auf dem CIDOC-CRM Datenmodell basierend und nur von wenigen Maskenanpassungen abgesehen, für die Hamburger Bedürfnisse zugeschnitten schien. Die digiCULT Verbund e. G. hat inzwischen die Umsetzungen so weit für die ersten Module abgeschlossen, dass sowohl in Schleswig-Holstein als auch im AMH bereits produktiv mit den nur geringfügig voneinander abweichenden Systemen gearbeitet wird. Im sogenannten Aktivitätenmodul werden alle Arten von Untersuchungen, Fundmeldungen und Zustandssichtungen von archäologischen Denkmälern als archäologische Aktivitäten verwaltet. Wie bereits erwähnt, dient die Aktivitätsnummer als Primärschlüssel für alle Objekte in archaeoDox. Es werden sowohl laufende als auch bereits weit zurückliegende Maßnahmen hierüber erfasst und mindestens mit einer möglichst genau bestimmten Lokalität, in der Regel jedoch mit einer Fundstelle verknüpft. An der jeweiligen Aktivität hängen die Fundobjekte, die wiederum auf Standorte beziehungsweise Personen verortet werden. In Schleswig-Holstein geschieht dies mittels barcodegestützter Verbuchungen. Es wird also eine hierarchische Kette beginnend mit der Fundstelle, den mit dieser verknüpften Aktivitäten und den wiederum an diesen angehängten Funden gebildet. Da das System von zwei Seiten aus erdacht und in enger Abstimmung konzipiert wurde, nämlich sowohl von musealer als auch bodendenkmalpflegerischer Seite, lässt sich diese hierarchische Kette auch in beiden Richtungen durchleuchten. Aus der fundstellenbasierten Sicht ergibt sich über die Übersicht der angehängten Aktivitäten eine komplette Untersuchungshistorie inklusive der jeweils angehängten Funde bis hin zu deren momentanen Standorten. Zur Zeit werden in Schleswig-Holstein Webservices aufgesetzt, die es ermöglichen sollen, über eine Informationsanfrage im Web-GIS nicht nur alle Informationen zu einer Fundstelle, sondern auch eine Auflistung der mit dieser verknüpften Aktivitäten inkl. Grunddaten anzuzeigen. Aus der fundobjektbasierten Sicht, zum Beispiel im Zentralmagazin des ALM, lässt sich damit im Umkehrschluss für jeden Fund dessen Herkunft und Entdeckungshistorie bis hin zu detaillierten Fundstelleninformationen zurückverfolgen. Das zentrale Modul und somit die Schnittstelle zwischen Bodendenkmalpflege und Museum ist hierbei das Fundobjektmodul. Hierüber werden die Fundobjekte zentral erfasst, wobei die wichtigsten vier Datenfelder Objektbezeichnung, Material, Datierung und Quantität auf einer Grunddatenmaske eingegeben werden. Aus Gründen der einfachen Übersicht und Verschlankung der Datenaufnahme sind vielfältigste erweiterte Daten wie Maße, Herstellungstechnik etc. auf zusätzliche Masken innerhalb dieses Moduls verteilt. Sowohl bei der Entwicklung dieser institutsübergreifenden Archäologischen Datenbank als auch bei der Entwicklung von archaeoDox wurde verstärkt darauf geachtet, dass beide Systeme im späteren produktiven Betrieb interoperabel zusammenarbeiten. Unter der Verwendung von Webservices und international standardisierter Schnittstellen- und Austauschformate werden nunmehr die Grunddaten von Funden aus Ausgrabungen mit archaeoDox erfasst und in die jeweiligen übergreifenden archäologischen Datenbanksysteme von Schleswig-Holstein beziehungsweise Hamburg mittels Webservice eingespeist. Im Anschluss daran können die Fundobjekte direkt auf Lagerbehältnisse und dann auf Standorte verbucht werden. Grundvoraussetzung ist hierbei, dass in beiden Systemen die gleichen Grunddaten erfasst und ein gemeinsames gesteuertes Vokabular verwendet wird. Um eine weitestgehende Einheitlichkeit in der Ansprache aller erfassten Objekte zu gewährleisten, greift archaeoDox, soweit möglich, auf Thesauri zurück. Dies hat gegenüber der Eingabe in Freitextfelder nicht bloß einen Geschwindigkeitsgewinn bei der Eingabe als Vorteil, sondern ist entscheidend für die spätere Durchsuch- und Auswertbarkeit des Datenbestandes und damit für die Qualität der erhobenen Daten. Als eine der nächsten kleineren Entwicklungsschritte von archaeoDox ist deshalb die Implementierung einer Schnittstelle zur ebenfalls von digiCULT Verbund e. G. entwickelten Thesauri-Verwaltung xTree vorgesehen, welche die Erstellung und Administrierung multihierarchischer und multilingualer Thesauri erlaubt. ArchaeoDox nutzt zwar bereits einen in xTree verwalteten Material-Thesaurus, die Implementierung einer Schnittstelle wird aber künftig die Möglichkeit bieten, alle Thesauri in archaeoDox zentral über xTree zu administrieren.

Alle in archaeoDox erfassten Grabungsdaten lassen sich in Form von vorkonfigurierten Listen als Berichte ausgeben. Dies bedeutet zum einen, dass nach Erfassung aller Grabungsdaten lediglich der eigentliche Grabungsbericht noch separat verfasst werden muss. Zum anderen müssen keinerlei Listen für Befunde, Funde, Fotos, Zeichnungen oder ähnliches mehr separat geführt werden, da diese bereits anhand der erfassten Daten und Verknüpfungen sowie der entsprechenden Dokumentationsschritte aus der Datenbank heraus generiert werden und so der umfassende Anhang eines Grabungsberichtes bereits fertig geliefert wird. Auch das Grabungstagebuch des jeweiligen archaeoDox-Vorgangs wird über die Erstellung der Dokumentationsschritte, welche im Tagebuch unter dem jeweiligen Arbeitstag aufgeführt werden, zu einem großen Teil aus bereits an anderer Stelle erfassten Daten generiert. Lediglich die Arbeitszeitenerfassung und gegebenenfalls der Eintrag besonderer Vorkommnisse oder des Wetters muss noch durch den Bearbeiter vorgenommen werden.

Da ein Aufkommen von mehreren Gigabyte an Daten für eine mehrwöchige, digital dokumentierte Grabung kleineren Umfangs inzwischen eher die Regel als die Ausnahme sein dürfte, ist es für die Bodendenkmalpflege unumgänglich, sich mit dem Problem der Datenhaltung, sowohl was die tägliche Arbeit als auch was die Langzeitarchivierung betrifft, auseinanderzusetzen. Der Schwerpunkt der Entwicklung von archaeoDox lag deshalb bislang ganz bewusst auf den Bereichen Datenerfassung und Archivierung sowie Datenablage und Dateimanagement. Eine weitere künftige Fähigkeit von archaeoDox soll die Möglichkeit zum Aufruf von externen GIS-Objekten sein. Für das ALSH und das AMH bedeutet dies die Verknüpfung mit der Fundstellenkartierung Archäologischer Atlas für Schleswig-Holstein und Denkmal-GIS für die Freie und Hansestadt Hamburg. Zugleich ist geplant, vonseiten der Fundstellenkartierung den entsprechenden archaeoDox-Vorgang aufzurufen und anzuzeigen. Damit würden nicht nur, wie bereits oben erläutert, die mit der Fundstelle verknüpften Aktivitäten mit ihren jeweiligen Grunddaten angezeigt werden, sondern man bekäme mit einem Link auf eine geplante archaeoDox-Webanwendung eine browsergestützte Übersicht sämtlicher Detailinformationen zu einer mit archaeoDox dokumentierten Maßnahme angezeigt. Diese Webanwendung soll eine Baumnavigation durch die jeweiligen kompletten Dokumentationen und zugleich eine direkte Zugriffsmöglichkeit auf die zugehörigen Dateien bieten. Mit der grundsätzlichen Möglichkeit, in archaeoDox auch Altgrabungen und andere archäologische Maßnahmen nachträglich zu erfassen, sei es im Rahmen einer Aufarbeitung oder nur durch Erzeugung einer Dateiablage für digitalisierte Akten, ist der Weg hin zu einer digitalen Ortsakte vorgezeichnet. Neben den Modulen Flächen- und Stadtkerngrabung soll ein Modul Denkmalschutz für die vereinfachte Erfassung von Zustandsbesichtigungen hinzukommen. Die Möglichkeiten, ein Modul Bauleitplanung zur Abwicklung und Verwaltung genehmigungsrechtlicher Stellungnahmen zu entwickeln, werden derzeit diskutiert. In diesem Zusammenhang stünde auch ein Ausbau der Tagebuchfunktion auf die betriebswirtschaftliche Abwicklung einer Grabung.

Die umfassendste Erweiterung wird aber im Zuge der Entwicklung eines Auswertungsmoduls stattfinden. Hier sollen nicht nur umfassende Such- und Abfragemöglichkeiten implementiert, sondern auch Möglichkeiten geschaffen werden, aggregierte Daten zur weiteren Modulierung in GIS-Systemen zu exportieren. In diesem Zusammenhang soll auch die Möglichkeit, welche PostgreSQL mit der Erweiterung PostGIS zur Verwaltung von Geometrien bietet, in archaeoDox genutzt werden.

Insgesamt betrachtet baut sich mit der parallelen Entwicklung der verschiedenen Datenbanksysteme für Fundstellen (Geodatenbanken), für Aktivitäten und Funde inklusive Restaurierung und Leihvorgänge (Archäologische Datenbanken) sowie für die Dokumentation von Aktivitäten (archaeoDox) durch die Interoperabilität der Systeme untereinander ein umfassender Datenbankverbund für die jeweiligen Bundesländer auf. Dies wird zukünftig die Möglichkeit bieten, aus diesen drei jeweils autark arbeitenden Systemen mit eigenen homogenen Datenbeständen zentral gebündelt sämtliche Informationen abzufragen.

PROJEKTE

Die Arbeitsgruppe „Archaeological Archives" des Europae Archaeologiae Consilium und das Culture 2007–2013 Project „ARCHES". Eine kurze Vorstellung.

David Bibby

Dieses Projekt wurde mit Unterstützung der Europäischen Kommission finanziert. Die Verantwortung für den Inhalt dieser Veröffentlichung (Mitteilung) trägt allein der Verfasser; die Kommission haftet nicht für die weitere Verwendung der darin enthaltenen Angaben.

Das Thema des Symposiums des EAC (Europae Archeologiae Consilium) in Metz im Jahr 2007 war „archäologische Archive". Während des Symposiums wurde schnell klar, dass Probleme und Sorgen archäologischer Archive sowie auch manche Lösungsansätze nicht national begrenzt sondern pan-europäisch zu behandeln sind. Deshalb entschied sich der EAC-Vorstand während des Symposiums zusätzlich zu den Arbeitsgruppen „Underwater Heritage", „Rural Land Use", „Large Scale Excavations" und „Remote Sensing" eine weitere Arbeitsgruppe innerhalb des EAC zu bilden, die sich allein mit der Problematik der archäologischen Archive auseinandersetzen soll.[1]

Zu den ersten Aufgaben der Arbeitsgruppe gehörte, sich auf eine Definition und die grundlegenden Prinzipien archäologischer Archivierung zu einigen, was wie folgt formuliert wurde:

Archäologische Archive setzen sich zusammen aus sämtlichen beweglichen Objekten/Kulturgut, Daten und Dokumentationen jeglicher archäologischer Beobachtung oder Aktivität.

Grundsätzlich sollte(n)

- ein archäologisches Archiv in einer anerkannten Einrichtung dauerhaft erhalten werden,

- jede archäologische Aktivität zum Ziel haben, ein komplettes, strukturiertes, stabiles und zugängliches Archiv zu erzeugen,

- die Standards für die Erstellung, Verwaltung und Aufbewahrung des archäologischen Archivs, das ein Projekt oder eine Aktivität „verursacht", gleich am Anfang dieser archäologischen Aktivität vereinbart werden.

Aufbauend auf diesen Kernprinzipien hat sich die Arbeitsgruppe zur Aufgabe gemacht, Leitlinien für die Vorbereitung, Erstellung und Ablage von archäologischem Archivgut zu formulieren. Dieses Vorhaben mit dem Akronym **ARCHES** – **A**rchaeolgical **R**esources in **C**ultural **H**eritage, a **E**uropean **S**tandard – wird mit Unterstützung des Culture 2007–2013 Programms der Europäischen Union durchgeführt.

1 Projekthintergrund

Die zunehmende Zahl archäologischer Aktivitäten sowie sich verfeinernde Methoden haben dazu geführt, dass die Menge der Ergebnisse – analoge und digitale Dokumentation, Funde, Proben etc. – einen exponentiellen Zuwachs erfahren hat. Während es allgemeine Übereinstimmung darüber gibt, dass diese heterogenen Ergebnisse archäologischer Interventionen dauerhaft aufbewahrt werden sollen, ist es oft schwierig, passende Ressourcen freizusetzen, die adäquat mit der explodierenden Menge an Objekten und Informationen umgehen können. Während des EAC-Symposiums in Metz 2007 wurde diskutiert, wie verschiedene europäische Länder mit dieser Problematik umgehen. Es wurden folgende wichtige Themen erörtert:

- die Definition eines archäologischen Archivs im Unterschied zu anderen Archiven,

- das Konzept der „ewigen" Aufbewahrung archäologischen Kulturguts,

- juristische Aspekte

- Verantwortlichkeiten und Aufgabenbereiche – Konservierung, Pflege, Archivierungspraxis

- Prozeduren für Auswahl und Aufbewahrung von Funden und Dokumentation

- Lagerungs- und Konservierungsstandards

- die Zugänglichkeit von Archiven.

Während des Symposiums wurde deutlich, dass fast alle Mitgliedsländer des EAC sich mit ähnlichen Problemen konfrontiert sehen. Es gab die einhellige Übereinstimmung, dass die Entwicklung optimaler Verfahren sowie umfassende europäische Archivierungsleitlinien wünschenswert seien. Mit dieser Zielvorgabe wurde die EAC-Arbeitsgruppe „Archäologische Archivierung" ins Leben gerufen. Nach verschiedenen Arbeitsgruppenorientierungstreffen in den Jahren 2008–2010 und einem erfolgreichen Projektantrag bei der Education, Audiovisual and Culture Agency (EACEA) der Europäischen Union, unter Stand 1.2.1 der Kooperationsprojekte, ist diese Arbeit unter dem Titel „ARCHES" seit Juni 2012 in konzentrierter Form angelaufen.

Die Leitlinien sollen alle Aspekte archäologischer Archive und archäologischer Archivierung umfassen – Funde, Proben, schriftliche und bildliche Dokumentation, Fotos und digitale Daten – und sollen einen Workflow von der Entdeckung/ Erzeugung während der archäologischen Aktivität bis zur endgültigen Ablage im Archiv vorschlagen.

2 Projektziele

Die Ziele des ARCHES-Projektes sind:

- gemeinsame Kerninformationen der Projektpartner zum Thema archäologische Archivierung sowie Verantwortlichkeiten und Aufgaben zu identifizieren,

- umfassende Leitlinien über die Erstellung, Vorbereitung und Ablage von archäologischem Archivgut, die auch landesspezifische Besonderheiten berücksichtigen,

- eine umfassende Bibliografie aller relevanten Standards, Werke und Richtlinien aus den Partnerländern zu erstellen,

- die Ergebnisse in Hardcopy sowie webbasiert zu publizieren und zu verbreiten,

- unter Mitwirkung des Archaeological Data Service[2] eine Webpräsenz zu etablieren, die zukunftsgerichtet als Plattform und Forum für das Thema archäologische Archivierung dienen soll.

3 Vorgehensweise

Das Projekt ist in sieben Arbeitsbereiche (Work Packages, WP) aufgeteilt:

- Projektmanagement und Koordination,

- Kernstandarderarbeitung und Umfragen,

- Organisation von Workshops,

- Bibliografie,

- Erarbeitung von anwendungsorientierten Standards,

- Veröffentlichung der Ergebnisse,

- Nachhaltigkeit.

Ein Kernstandardentwurf und Umfragen (WP 2) werden durch eintägige nationale oder regionale Workshops oder Meetings (WP3) ergänzt, die dazu dienen, Informationen über bestehende Verfahrensweisen und Standards zu sammeln. Themenschwerpunkte sind dabei die Projektplanung, die Datenerfassung und der Umgang mit Ausgrabungs- bzw. Projektdokumentationen, mit digitalen Daten und Funden. Weitere Aspekte sind die Analyse und Archivierung der Ergebnisse, die Verpackung von Funden unterschiedlicher Art, die Problematik der ungewissen Lebensdauer digitaler Daten, die Inventarisierung und Archivordnung sowie die Übernahme des zu archivierenden Gutes und die langfristige Pflege und das Management von archäologischen Archiven. Die Umfragen und Workshops sollen bereits existierende erfolgreiche Strategien identifizieren, die in die ARCHES-Leitlinien übernommen werden können. Ebenso werden damit Bereiche sichtbar, wo Strategien fehlen oder so unterschiedlich sind, dass eine neue Orientierung nötig ist. Aus einer Kombination von Informationen aus den Umfragen und Workshops, aus a priori-Wissen[3] und aus der Konsultation mit dem EAC-Vorstand werden im nächsten Arbeitsschritt ein Satz allgemeingültiger Kernprämissen (Core Standards) zusammengestellt sowie Empfehlungen und Leitlinien (auch unter Berücksichtigung regionaler und nationaler Besonderheiten) für die tatsächliche Umsetzung (WP 2, WP 5) abgegeben. Diese werden in Hardcopy in englischer Sprache publiziert. Übertragungen in die jeweiligen Nationalsprachen der

Projektteilnehmer werden auf der ARCHES Homepage zu finden sein. Die Homepage (WP 6, WP7) dient auch als zukunftsgerichtete Plattform für Erneuerungen und Ergänzungen der ARCHES-Leitlinien. Sie stellt Raum für Foren und Diskussionsbeiträge zum Thema „Archäologische Archivierung" zur Verfügung. Die Einrichtung einer leicht zugänglichen Homepage soll die künftige Nachhaltigkeit des Projekts garantieren. Im Rahmen einer Kampagne zur Steigerung des Bewusstseins für das Thema *archäologische Archivierung und archäologische Archive*, zu der auch dieser Beitrag gehört, folgen außer der Webseite weitere Publikationen, Aktionen und Newsletter, Vorträge und Posterpräsentationen des ARCHES-Projekts auf Fachtagungen.

Die Publikation der Kernstandards und Anwendungsleitlinien stellt einen wichtigen Meilenstein dar, da es die erste gezielte internationale Publikation sein wird, die sich allein mit archäologischer Archivierung und archäologischen Archiven befasst. Aber damit nicht genug. Die heutige Welt – auch die Welt der Archäologie und archäologischer Methoden – ändert und entwickelt sich ständig. Es wäre deshalb wünschenswert, dass die Publikation nicht als End-, sondern als Startpunkt verstanden wird: Dies ist der Beginn eines modernen Bewusstseins nachhaltiger archäologischer Archive.

Anmerkungen

[1] Mitglieder sind: Ann Degraeve (Vorsitzende), Ministère de la Région de Bruxelles-Capitale, B; David Bibby (stellvertretender Vorsitzender, lead party manager ARCHES), Landesamt für Denkmalpflege Baden-Württemberg im Regierungspräsidium Stuttgart, D; Duncan Brown, English Heritage, GB; Annika Carlsson, Riksantikvarieämbetet, S; Martin Kuna, Archeologický ústav AV CR, Praha, i.v.v., CZ; Guus Lange, Rijksdienst voor het Cultureel Erfgoed, NL; Sólborg Una Pálsdóttir, Archaeological Heritage Agency, IS; Carla Schulte, Rijksdienst voor het Cultureel Erfgoed, NL; Bettina Stoll-Tucker, Landesamt für Denkmalpflege und Archäologie Sachsen-Anhalt, D; Claire Driver, English Heritage, GB (Sekretariat). Berater: Catherine Hardmann, ADS, GB; Kathy Perrin, GB; Sascha Schmidt, VS-Consulting, D; Adviser, Adrian Olivier, GB. Obwohl die Schweiz keinen Vertreter in der Arbeitsgruppe hat und aufgrund der Nichtmitgliedschaft in der EU an ARCHES nicht teilnehmen kann, wird zur Zeit ein Parallelprojekt in Anlehnung an die ARCHES-Struktur durchgeführt. Hierzu siehe: http://www.archaeoconcept.com/de/projects-2/projects_actuel/#1.

[2] Archaeological Data Service, ADS; http://archaeologydataservice.ac.uk/.

[3] Quasi als „Gerüst" für die Struktur des Projekts dient: Brown, D.H., „Archaeological Archives: A guide to best practice in creation, compilation, transfer and curation", London, 2007.

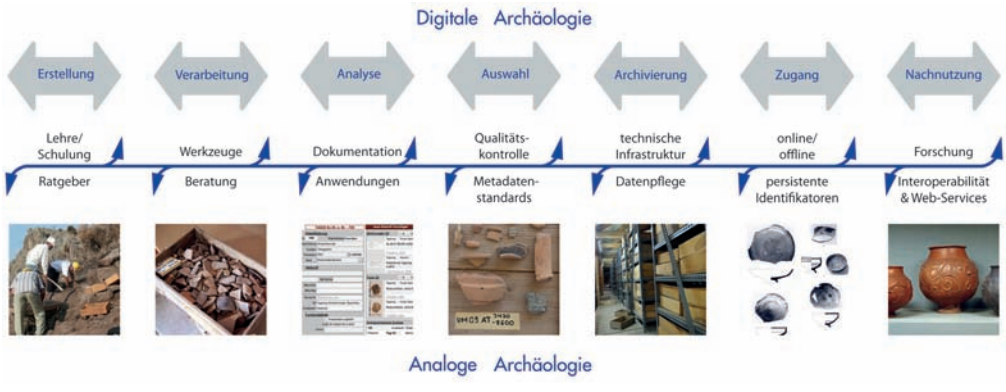
IANUS. Die Konzeption eines nationalen Forschungsdatenzentrums für die Archäologie und die Altertumswissenschaften

Ortwin Dally, Friederike Fless, Reinhard Förtsch, Maurice Heinrich, Felix F. Schäfer

Ausgangslage: Moderne Altertumswissenschaften

Seit der zweiten Hälfte des 19. Jahrhunderts haben die Archäologien und die Altertumswissenschaften einen komplexen Prozess der Ausdifferenzierung erfahren, der sich in der Verankerung der Forschung in spezifischen Institutionen wie den Universitäten, dem Deutschen Archäologischen Institut, den Akademien, den Museen oder den Landesdenkmalämtern artikuliert hat. Befördert durch Großgrabungen, die unter anderem in Kleinasien, Griechenland und dem Vorderen Orient im letzten Viertel des 19. Jahrhunderts einsetzten, sind in rascher Folge neue Arbeitsfelder entstanden, die wiederum als universitäre Disziplinen Gestalt angenommen haben. Die altertumswissenschaftliche Forschung umfasst heute ein breites Spektrum. Zu ihnen gehören im Kern verschiedene Archäologien, Philologien und Geschichtswissenschaften, die jeweils räumlich und/oder zeitlich unterschiedliche Kulturen zum Gegenstand haben. Je nach Quellengattung werden sie durch stärker spezialisierte Disziplinen ergänzt, zum Beispiel durch die Bauforschung, die Dendrochronologie, die Epigrafik oder die Numismatik. Schließlich spielen auch Fächer und Methoden eine Rolle, die eigentlich in anderen Fachdomänen beheimatet sind, aber dennoch wichtige Informationen zu historischen Phänomenen beitragen können; hierzu zählen etwa die Anthropologie, Archäobiologie und Archäozoologie, die Geologie und Geografie, naturwissenschaftliche und medizinische Untersuchungen.

Ergebnisse der Forschung wurden zunächst in analoger Form mit spezifischen Ausprägungen der Wissensaufbereitung langfristig gesichert, zum Beispiel in Bibliotheken, Grabungsarchiven, Gelehrtennachlässen, Corpora, Fototheken, gedruckten Publikationen, Gipsabgüssen und Abklatschen oder Sammlungen in Museen. Heute kommt diese Rolle in einem rasant wachsenden Maße digitalen Daten zu. Je nach Projekten und Fragestellungen werden entweder „(Primär-)Daten" generiert – zum Beispiel bei Feldforschungen, Untersuchungen mit Originalobjekten oder Texteditionen – oder durch Auswertungen, Analysen und Berechnungen neue „(Sekundär-)Daten" mit Erkenntnissen und Interpretationen gewonnen, die dann in der Regel publiziert werden.[1] Beide Prozesse werden zunehmend durch elektronische Systeme, die unmittelbar digitale Daten erzeugen, verarbeiten, visualisieren und bereitstellen, geprägt. In immer mehr Fällen bilden ausschließlich digital vorliegende Informationen („born digital") neben schriftlichen Dokumentationen und materiellen Hinterlassenschaften sogar die einzige Grundlage für moderne Forschungsfragen. Die konkreten Einsatzgebiete von Rechnern in den Altertumswissenschaften sind dabei vielfältig: Datenbanken dienen zur Strukturierung und Verwaltung von Informationen über Fundstellen und Objekten, geografische Informationssysteme helfen bei der Analyse und Darstellung von raumbezogenen Angaben, elektronische Texte werden automatisiert nach sprachlich-syntaktischen Charakteristika durchsucht, Statistikprogramme werten quantitative Beobachtungen aus, 2-D- und 3-D-Rekonstruktionen helfen wissenschaftliche Ergebnisse zu visualisieren und zu überprüfen, digitale Fotos, Zeichnungen und Messungen dokumentieren vergängliche Kontexte. Je nach Sichtweise verdrängen oder ergänzen diese Verfahren immer stärker die bislang etablierten Praktiken und deren in der Vergangenheit zumeist analogen Resultate.



1 Idealisierte Gegenüberstellung analoger und digitaler Arbeitsschritte in der Archäologie.

In solchen Systemen kommt ein generelles Interesse der modernen Forschung zum Ausdruck, das die Untersuchung einzelner Materialgruppen nicht ausschließt, aber den Kontext kultureller Phänomene als vordringliche Grundlage wissenschaftlicher Erkenntnis begreift. Die Wertschöpfung liegt vor allem in der Zusammenschau sinnvoller und logischer Einheiten, die durch automatisierte Verfahren extrahiert und miteinander in Bezug gesetzt beziehungsweise verknüpft werden. Diese Kontexte sind wiederum zunehmend eingebunden in die Betrachtung größerer Zusammenhänge: anspruchsvolle Konzepte (zum Beispiel Raum, Erinnerung und Gedächtnis, Landschaft, Macht, Innovationen), anthropologische Fragestellungen zur Geschichte der Menschheit insgesamt oder auf Regionen beziehungsweise Kulturen bezogene Forschungsschwerpunkte (zum Beispiel Südostasien; Afrika; Pazifik; Germanen; Kelten). Dadurch ergeben sich Schnittmengen unter anderem zur Ethnologie, den Sozial- und den Geschichts-, aber auch den Naturwissenschaften.

Die den Forschungen zugrunde liegenden digitalen Forschungsdaten unterliegen dabei einem sogenannten Lebenszyklus, da sich ihre Natur ständig ändert: Rohdaten werden generiert und ergänzt; vorläufige, dynamische Daten werden verändert, prozessiert und analysiert; statische, finalisierte Daten werden abgelegt und veröffentlicht und bilden (alleine oder angereichert mit weiteren Daten) selbst wieder die Basis für neue Erkenntnisprozesse. Jede Phase stellt dabei spezifische Anforderungen an individuelle oder kollektive Akteure, divergierende Arbeitsabläufe, unterschiedliche Hard- und Software-Ressourcen und rechtliche wie finanzielle Rahmenbedingungen. Der Übergang von einer Phase in eine andere ist dabei fließend und verläuft keineswegs immer zyklisch, wie der abgebildete idealisierte und stark vereinfachte Kreislauf suggeriert.[2]

Diese tief greifenden Veränderungen prägen nicht nur in zunehmendem Maß den Arbeitsalltag von Wissenschaftlern und erfordern eine zunehmende Expertise im Bereich des Forschungsdatenmanagements, sondern bringen auch eine Vielzahl neuer Möglichkeiten und Chancen mit sich. Gleichzeitig werfen sie eine Reihe von offenen Fragen auf und verursachen neue Probleme, die die Forschung gleichermaßen in fachlich-methodischer wie institutionell-organisatorischer Hinsicht betreffen. Die informationstechnische Herausforderung für die Altertumswissenschaften angesichts ihrer Positionierung an einer Schnittstelle zwischen Natur- und Geisteswissenschaften besteht in der Entwicklung von disziplinenübergreifenden, langfristig verfügbaren Systemen und Infrastrukturen, die die Zerstreuung von Informationen nach analogen Ordnungskriterien (Bibliothek, Fotothek, Archiv, Grabungsakten, Publikationen, Objektsammlungen, etc.) aufhebt. Im Speziellen sind zwei Herausforderungen zu konstatieren:

1) Im Verlauf von Ausgrabungen und Oberflächenbegehungen werden historische Befundsituationen irreversibel zerstört. Hierdurch ergeben sich besondere Anforderungen hinsichtlich Vollständigkeit, Nachvollziehbarkeit und langfristigem Zugriff an die Dokumentation solcher Maßnahmen, die zusammen mit den physischen Objekten einmalige Primärquellen mit zeitlich unbegrenzter Gültigkeit darstellen. So wie heute Archäologen noch mit Grabungsunterlagen des 19. Jahrhunderts arbeiten, werden auch künftige Generationen auf die digitalen Daten von heute zurückgreifen wollen und müssen. Das Problem der langfristigen Sicherung von Forschungsdaten betrifft somit zentrale Grundlagen des Fachs.

2) Die zweite Herausforderung besteht in der starken technischen und semantischen Heterogenität der Daten und Quellen, die für die Erforschung vergangener Kulturen relevant sind. Für die Beantwortung komplexer Fragen müssen mehrere Disziplinen mit sehr unterschiedlichen Qualitäten und Quantitäten von analogen Quellen und digitalen Informationen zusammenarbeiten, die jeweiligen Inhalte müssen miteinander verknüpft werden. Erst im Gesamtbestand dieser miteinander verknüpften Daten wird das volle Potenzial eines Forschungspropjekts sichtbar.

2 Vereinfachtes Modell des Data-Life-Cycles.

Archäologen sind einerseits mit der Aufgabe konfrontiert, diese komplexen Arbeitsgebiete in geeigneten technischen Lösungen zur Datenintegration abzubilden. Andererseits wird die Interoperabilität von Daten durch die Vielfalt von Akteuren erschwert, die sich in Deutschland und – bei internationalen Projekten – im Ausland mit der Altertumskunde beschäftigen (Universitätsinstitute, Museen, Akademien, Denkmalfachbehörden der Länder und Kommunen, Forschungseinrichtungen, kommerzielle Firmen und ehrenamtliche Laien) und jeweils eigene institutionelle, rechtliche und finanzielle Rahmenbedingungen aufweisen.

Auch wenn beide Aspekte, die Heterogenität und die Singularität von Daten, nicht alleine für die Altertumswissenschaften spezifisch sind[3], bei Weitem nicht neu sind und in einzelnen Disziplinen bereits Lösungen und Einrichtungen existieren[4], bleibt festzuhalten, dass aufgrund der durchweg hohen Komplexität der Primärdaten und der mit diesen verbundenen vielschichtigen Metadaten, die durchaus ein Spezifikum der Altertumswisenschaften darstellen, bislang eine fachspezifische, zentrale Einrichtung in Deutschland fehlt, die die drängenden Fragen im Umgang mit digitalen Forschungsdaten adressiert sowie diese nachhaltig archivieren und langfristig bereitstellen kann.

Eine Arbeitsgruppe und ein Antrag

Aus der Erkenntnis heraus, dass die skizzierten digitalen Arbeitsprozesse mit ihren Formaten, Metabeschreibungen, Interoperabilitätsfragen und Archivierungsproblemen mit den derzeitigen technischen, personellen und organisatorischen Infrastrukturen in den unterschiedlichen altertumswissenschaftlichen Institutionen und Disziplinen nicht mehr beherrschbar sind, und dass dadurch zentrale Qualitäten digitaler Daten wie die Austauschbarkeit, Verfügbarkeit und Nachnutzbarkeit zur Zeit kaum gesichert werden können, wurde 2008 durch die DFG eine Arbeitsgruppe mit repräsentativen Fachvertretern initiiert. In dieser sind in ihrer heutigen Zusammensetzung (Stand Oktober 2012) beteiligt:

- das Deutsche Archäologische Institut als größtes archäologisches Forschungsinstitut in Deutschland,
- das Exzellenzcluster TOPOI in Berlin und die Graduiertenschule „Human Development in Landscapes" an der CAU Kiel als Vertreter für große Forschungsverbünde und universitäre Nachwuchsförderung,
- der Verband der Landesarchäologen als Vertreter der staatlichen Bodendenkmalpflege,
- das Rechenzentrum der Universität Köln als Vertreter universitärer Infrastrukturen zur Datenzusammenführung, -vorhaltung und -langzeitspeicherung,
- die Staats- und Universitäts-Bibliothek Göttingen als Vertreter für wissenschaftliche Bibliotheken und zentrale Informationseinrichtungen sowie als Konsortialführer wichtiger Infrastrukturprojekte für die Digital Humanities,
- das Curt-Engelhorn-Zentrum Archäometrie in Mannheim als Beispiel für sogenannte An-Institute und als Vertreter für naturwissenschaftliche Forschungen in der Archäologie,
- die Archäologischen Staatssammlungen in München als Vertreter der Museen,
- die Berlin-Brandenburgische Akademie der Wissenschaften als Vertreter der Akademien.

Im April 2010 wurde stellvertretend für die gesamte DFG-Arbeitsgruppe durch Prof. Ortwin Dally (DAI Berlin) und Prof. Frederike Fless (damals TOPOI Berlin, jetzt DAI Berlin) ein Antrag bei der DFG zur „Entwicklung eines Kompetenzzentrums für altertumswissenschaftliche Forschungsdaten" eingereicht. Grundlage dafür waren die zuvor geführten Diskussionen zu Bedarf und Notwendigkeit eines nationalen Zentrums, das langfristig die Fragen und Probleme in den Altertumswissenschaften im Umgang mit digitalen Daten adressiert und maßgeblich zu deren Lösung beiträgt. Mithilfe des Antrags wurde die Förderung einer ersten, dreijährigen Projektphase angestrebt, in der die spezifischen Anforderungen und Rahmenbedingungen formuliert und Konzepte für einen langfristigen Betrieb erarbeitet werden sollen. Die Ergebnisse bilden dann die Basis für eine zweite Förderungsphase, in der eine funktionierende Organisationseinheit aufgebaut und in einen Regelbetrieb überführt werden soll. Erklärtes Ziel ist, dass am Ende des Prozesses ein sich selbst tragendes, unabhängiges Forschungsdatenzentrum mit Dienstleistungen für alle Archäologien und Altertumswissenschaften in Deutschland existiert. Insgesamt soll auf diese Weise eine institutionen-, projekt- und fachübergreifende Infrastruktur für die nachhaltige Archivierung und Bereitstellung von qualitativen und interoperablen digitalen Inhalten geschaffen werden.

Erste Maßnahmen:
Mehrere Arbeitspakete und mehrere Arbeitsgruppen

Nach einer positiven Bewilligung Ende April 2011 durch die DFG konnte zum 1. September des gleichen Jahres das Projekt unter dem neuen Titel „IANUS. Forschungsdatenzentrum Archäologie und Altertumswissenschaften" seine Arbeit aufnehmen[5].

Die beiden beantragten Projektstellen sind logistisch am Deutschen Archäologischen Institut (DAI) in Berlin angesiedelt, das eine koordinierende Rolle in dem

noch jungen Prozess übernommen hat und als Ansprechpartner gegenüber der DFG fungiert. Durch die Einbindung verschiedener Fachvertreter aus unterschiedlichen Institutionen wird jedoch versucht, alle in Deutschland relevanten wissenschaftlichen Stakeholder miteinzubeziehen, um ein möglichst breites Spektrum an Anforderungen, Erwartungen und Expertisen bei der Konzeption und dem Aufbau zu berücksichtigen. Beispielsweise setzt sich die oben genannte DFG-Arbeitsgruppe, die nun das Projekt in der Funktion eines „Wissenschaftlichen Beirats" begleitet und steuert, aus je einem Vertreter aus dem Bereich Museum, Bodendenkmalpflege, Akademie, Universität, An-Institut, Verbundforschung (Exzellenzcluster, Graduiertenschule), unabhängige Forschungseinrichtung, Bibliothek und Rechenzentrum zusammen[6]. Mehrere Vertreter aus dem Kreis der Landesarchäologen sind in einzelnen Arbeitsgruppen (siehe unten) eingebunden, um die dort geführten Überlegungen und Diskussionen mitzugestalten und spezifische Wünsche, Anforderungen und Bedingungen einbringen zu können. Insgesamt arbeiten rund 100 Wissenschaftler aus etwa 40 verschiedenen Einrichtungen und Verbänden im In- und Ausland an dem Aufbau des Forschungsdatenzentrums mit.

Ziel der ersten, konzeptionell geprägten und auf drei Jahre ausgelegten Projektphase ist es, bis Herbst 2014 die zukünftigen Anforderungen und Strukturen, Workflows und Geschäftsgrundlagen eines Forschungsdatenzentrums zu beschreiben und ein nachhaltiges Betriebskonzept zu erarbeiten, welches zum einen durch die Fach-Community mitgetragen wird und zum anderen technisch, politisch und finanziell realisierbar ist. Diese Aufgabe ist in sechs verschiedene Arbeitspakete aufgegliedert, in denen unterschiedliche Arbeitsweisen zum Einsatz kommen: regelmäßige Workshops mit festen Arbeitsgruppen, Gutachten durch externe Fachleute, Expertengespräche in vergleichbaren Einrichtungen, softwaretechnische Realisierung von Demonstratoren, eine Stakeholder-Analyse sowie Vorträge und Publikationen.

Arbeitspaket 1 beinhaltet eine Evaluierung von bereits vorhandenen Forschungsdateninfrastrukturen sowohl im Hinblick auf technische Eigenschaften (zum Beispiel Rechenleistung, Storage-Kapazitäten, Software/Systeme, Sicherheit, Personal etc.) als auch auf ihre fachwissenschaftlichen Aufgaben (zum Beispiel Entwicklungsarbeit, Services, Beratung, Dokumentation, Kooperationen etc.). Da allen Projektbeteiligten bewusst ist, dass Aufbau und Betrieb größerer Serverarchitekturen im Rahmen von IANUS weder finanziert noch sinnvoll aufrechterhalten werden können, wird stattdessen für die (hardware-)technisch geprägte „bitstream preservation" eine Umsetzung mit mehreren Kooperationspartnern angestrebt. Aufgrund dieser Überlegung gilt es, die in den vergangenen Jahren sowohl auf nationaler wie internationaler Ebene angestoßenen und zum Teil verwirklichten Konzepte und Lösungen für die Sicherung, Archivierung und langfristige Bereitstellung von Forschungsdaten hinsichtlich der fachspezifischen Anforderungen zu überprüfen und zu klären, welche existierenden Rechenzentren, Universitätsbibliotheken, Landesarchive oder Initiativen[7] als Anbieter für spezielle Dienstleistungen, Technologien, Infrastrukturen oder Support geeignet sind. Die Beantwortung dieser Fragen wurde an zwei externe IT-Spezialisten an den Universitäten Dortmund und Freiburg vergeben, die hierzu eine schriftliche Expertise verfassen. Die dort formulierten Vorschläge werden in einer eigenen Arbeitsgruppe, die sowohl mit Vertretern aus nichtarchäologischen Infrastruktureinrichtungen und Rechenzentren als auch mit Fachwissenschaftlern besetzt ist, diskutiert, und es werden konkrete Empfehlungen mit Handlungsaufträgen abgeleitet. Auch bei den fachwissenschaftlichen Softwarelösungen, Werkzeugen, Online-Services und Web-Angeboten gibt es bereits mehrere einschlägige Projekte und Anbieter[8], deren Eignung für bestimmte Aufgaben von IANUS diskutiert werden muss.

Wie oben dargelegt stellt die Heterogenität der Daten ein Spezifikum altertumswissenschaftlicher Forschung dar. In Verbindung mit fehlenden fachlichen Standards oder unzureichend angewandten Regelwerken bei der Datengenerierung wird ein Projekt- beziehungsweise institutionenübergreifender Datenaus-

3 Das Logo von „IANUS – Forschungsdatenzentrum Archäologie & Altertumswissenschaften".

tausch erheblich erschwert beziehungsweise das Potenzial vernetzter digitaler Daten bleibt weitgehend ungenutzt. Daher widmet sich ein eigenes Arbeitspaket 2 dem Themenkomplex Interoperabilität von Anwendungen und Informationen. Auch hier kann auf eine ganze Reihe von Entwicklungen zurückgegriffen werden, deren Relevanz für IANUS zu überprüfen ist. Standards, Dienste und Schnittstellen, die primär außerhalb der Altertumswissenschaften in IT-Projekten definiert wurden und die Migration, Standardisierung und Kommunikation von Daten betreffen, sollen unter einer technischen Perspektive betrachtet werden.

Daneben kommen inhaltlich-semantische Aspekte vor allem bei verschiedenen Metadaten-Modellen und kontrollierten Vokabularen zum Tragen, die zum Teil in einzelnen Disziplinen oder Einrichtungen erarbeitet wurden, und deren Anwendbarkeit in einer übergreifenden IT-Dateninfrastruktur zu analysieren ist. Die vielfältigen Fragen und Aspekte zum Datenaustausch und zur Datenqualität werden von einer international besetzten Arbeitsgruppe behandelt, die sich in regelmäßigen Intervallen trifft und Empfehlungen formuliert.

In Arbeitspaket 3 wird der Fokus auf die aktive Einbindung der Fach-Community in den Planungsprozess gelegt, um die zukünftigen individuellen wie institutionellen Nutzer und Dateneigentümer über die Existenz und den Fortschritt des Forschungsdatenzentrums zu informieren und mit diesen frühzeitig die Ziele und Aufgaben sowie Erwartungen an IANUS zu diskutieren. Da nur das DAI über eine zentralisierte Organisationsstruktur verfügt, sind die anderen Bereiche wie Universitäten, Akademien, Landesdenkmalämter und Museen zwar in den verschiedenen Arbeitsgruppen durch einzelne Mitglieder repräsentiert, doch vertreten diese die Fachwelt in Deutschland eher pars pro toto und nicht in der ganzen Breite. Neben Vorträgen und Diskussionen auf verschiedenen Fach- und Verbandstagungen sollen zu diesem Zweck in Zusammenarbeit mit dem 2011 in Bremen gegründeten „Deutschen Verband für Archäologie" (DVA)[9] unterschiedliche Maßnahmen wie eine Stakeholder-Analyse, eine Erhebung des fachlichen Bedarfs, Befragungen zu rechtlichen und finanziellen Rahmenbedingungen (beispielsweise Restriktionen zur Datenweitergabe) und Workshops zu zentralen Aspekten des Forschungsdatenmanagements und der Langzeitarchivierung durchgeführt werden. Dadurch soll einerseits ein breites Bewusstsein für den reflektierten Umgang mit Forschungsdaten geschaffen werden und andererseits durch die Kommunikation der Vorteile und der Notwendigkeit eines solchen Zentrums dessen Akzeptanz und Legitimation in der Fachwelt erhöht werden.

Eng mit dem Arbeitspaket 3 verknüpft ist das vierte Arbeitspaket zum Thema Ausbildung sowie (Weiter)-Qualifizierung von Nachwuchskräften und etablierten Wissenschaftlern. Es ist zu konstatieren, dass zwischen der aktuellen fachlichen (akademischen) Ausbildung und den beruflichen Anforderungen zu elektronischen Arbeitsweisen häufig eine Diskrepanz existiert. Obwohl digitale Techniken und Inhalte besonders für die jüngeren Generationen immer selbstverständlicher werden, ist der richtige, wissenschaftlich adäquate und nachhaltige Umgang mit Daten weitgehend unbekannt und wird auch nur punktuell in den aktuellen Bachelor- und Master-Studiengängen vermittelt. Während der ersten Projektphase soll gemeinsam die Grundfrage diskutiert werden, welche IT-Fähigkeiten und welches IT-Equipment auf welcher Ebene benötigt beziehungsweise erwartet werden. Basierend darauf sollen Bereiche beschrieben werden, in denen IANUS langfristig unterstützend tätig werden kann, um die Ausbildungssituation und Anforderungssituation im Berufsleben besser zu verbinden, etwa durch Lehrangebote, Praxisratgeber, Unterrichtsmaterialien, Online-Tutorials, Summer Schools oder Schulungen. Die gewonnenen Ergebnisse können wichtige Impulse für die Digital Humanities-Ausbildung in Deutschland geben und sollen in die Universitäten zurückgeleitet werden[10].

Die wohl wichtigste Dienstleistung von IANUS wird die Langzeitarchivierung von Forschungsdaten sein. Diese besitzt eine besonders hohe Komplexität und steht deshalb im Mittelpunkt des 5. Arbeitspaketes. In diesem müssen vor allem auf den halbjährlichen Treffen der DFG-Arbeitsgruppe sowie durch praktische Testbeds eine Reihe von Fragen geklärt werden – beispielsweise zur Auswahl und Strukturierung von Daten vor der Übergabe an ein Archiv, zu Workflows und zur Qualitätssicherung, zur technischen und semantischen Validierung von Formaten und Inhalten, zur Verwaltung von Metadaten und Zugriffsrechten sowie zur Bereitstellung archivierter Inhalte für die Nachnutzung durch Dritte. Auch hier kann auf Vorarbeiten und Erfahrungen anderer Projekte zurückgegriffen werden, die diese Aspekte bereits generisch und disziplinenübergreifend adressiert haben[11]. Aufgrund der unterschiedlichen Institutionen, die in IANUS eingebunden werden sollen, und der jeweils unterschiedlichen finanziellen, technischen und rechtlichen Rahmenbedingungen[12], ist davon auszugehen, dass für das Forschungsdatenzentrum ein IT-Konzept mit mehreren vernetzten technischen Partnern angestrebt werden muss. Die daraus resultierenden spezifischen Herausforderungen sollen durch zwei externe Experten thematisiert werden, indem sie die Komponenten existierender technischer Systeme und Ansätze dahingehend evaluieren, inwieweit diese für eine kooperativ verteilte Umsetzung geeignet sind.

In dem letzten Arbeitspaket 6 soll ein Geschäftsmodell und Betriebskonzept entworfen werden, damit IANUS nach dem Ende der DFG-Förderung, die nur den Aufbau, aber nicht den langfristigen Betrieb des Zentrums betrifft, als eigenständige, sich selbst tragende Institution existieren kann. Eine zentrale Frage betrifft dabei die mögliche Rechts- und Organisationsform, die im Rahmen einer Expertise durch zwei Fachjuristen analysiert werden soll, wobei die gesellschaftsrechtlichen Besonderheiten der unterschiedlichen Kooperationspartner (überwiegend Anstalten des öffentlichen Rechts) zu berücksichtigen sind. Ebenso ist zu klären, wie die Aspekte Kosten, Haftung, Verantwortung, Leistungsqualität, Policies, Service-Level-Agreements etc. sinnvoll geregelt werden können. Neben den Rechtsfragen, die IANUS als Organisation betreffen, spielen die juristischen Rahmenbedingungen für die Generierung, Verbreitung und Nachnutzung von altertumswissenschaftlichen Forschungsdaten eine wichtige Rolle, zum Beispiel bezogen auf Zugriffs-, Verwertungs- und Urheberrechte oder hinsichtlich Grenzen, die durch nationale wie internationale Denkmalschutzbestimmungen gesetzt werden. Auch dieser Komplex soll in Form eines Experten-Gutachtens bearbeitet werden. Schließlich ist ein nachhaltiger Finanzierungs- und Kostenplan zu erstellen und zu ermitteln, welche Träger (zum Beispiel Bund, Länder, Drittmittelgeber, Facheinrichtungen, Dateneigentümer, Nutzer usw.) sich in welchem Umfang an laufenden Ausgaben beteiligen wollen und können.

Da sich alle Arbeitspakete in unterschiedlichem Maße überschneiden beziehungsweise gemeinsame Berührungspunkte besitzen, werden in einer weiteren Arbeitsgruppe übergreifende Facetten diskutiert und zueinander in Bezug gesetzt. Als methodisches Vorgehen wurden hierfür Vor-Ort-Begehungen und Expertengespräche mit fachlich oder strukturell vergleichbaren Einrichtungen im In- und Ausland ausgewählt, die mithilfe eines ausgearbeiteten Fragebogens untersucht und evaluiert werden. Ziel eines derartigen Vorgehens ist es, durch den Austausch mit Betreibern bestehender Infrastrukturen von den mehrjährigen Erfahrungen, Lösungen und Problemen zu profitieren sowie gegebenenfalls Potenziale für Synergien und Kooperationen aufzuzeigen. Neben zentralen Fragen zu den jeweiligen Zielsetzungen, Workflows, Standards, Policies, Ressourcen und Geschäftsmodellen werden dabei auch Themen wie Nutzungsstatistiken, User-Feedback, Akzeptanz innerhalb und außerhalb der Fachwelt, Fehlentwicklungen und Perspektiven angesprochen. Aufgrund der fachlichen Nähe sind die wichtigsten Gesprächspartner das 2005 als Teil von DANS gegründete *e-Depot for Nederlandse Archaeology*[13] in Den Haag/Niederlande und der *Archaelogy Data Service*[14], der an der Universität York/Großbritannien beheimatet ist und seit 1996 Dienstleistungen im Bereich Langzeitarchivierung und Online-Bereitstellung von Daten für britische Archäologen anbietet. In Deutschland existiert dagegen eine Reihe von Einrichtungen, die zwar Daten anderer Fachgebiete betreuen, aber hinsichtlich ihrer Einbindung in die für Deutschland spezifischen Rahmenbedingungen wichtig sind: das Datenarchiv für Sozialwissenschaften als Teil des Leibniz-Instituts GESIS in Köln[15], PANGAEA als Langzeitarchiv für Erd- und Umweltdaten in Bremerhaven[16], das Max-Planck-Institut für Psycholinguistic in Nijmegen/Niederlande als Betreiber eines Sprachen-Archives[17] und die Geo-Daten-Infrastruktur-DE, die vom Bundesamt für Kartographie und Geodäsie in Frankfurt koordiniert wird[18].

Ziele:
Künftige Aufgaben und Dienstleistungen

IANUS versteht sich als ein nationales Forschungsdatenzentrum, das verschiedene Dienstleistungen im Umgang mit digitalen Forschungsdaten anbieten und zur Verbesserung der wissenschaftlichen Forschung beitragen will, indem es einen nachhaltigen Zugang zu dokumentierten Daten garantiert und dadurch deren zukünftige Nachnutzung ermöglicht. Es richtet sich dabei vorwiegend an Einrichtungen und Wissenschaftler in Deutschland, die archäologische und altertumskundliche Forschungen im In- und Ausland betreiben, das heißt sowohl an die verschiedenen Ebenen der föderal organisierten Bodendenkmalpflege als auch an universitäre Projekte, an Akademien ebenso wie Museen, an wissenschaftliche Langzeitvorhaben genauso wie an befristete Vorhaben mit kürzeren Laufzeiten, gleichermaßen an große vernetzte Kooperationen wie an individuelle Qualifizierungsarbeiten. Gleichzeitig ist aber auch die breite Öffentlichkeit eingeladen, die Angebote und Inhalte zu nutzen und zum Beispiel für didaktische Zwecke in Schule und Ausbildung zu verwenden.

Die fachliche Relevanz ist dabei weder an einen bestimmten geografischen und zeitlichen Kulturraum gebunden noch auf bestimmte Methoden und Fragestellungen festgelegt. Die Nutzung der Dienstleistungen und Daten von IANUS erfolgt auf einer freiwilligen Basis und kann als Ergänzung zu bestehenden institutionellen IT-Lösungen oder als Ersatz für fehlende Infrastrukturen in eigenen Instituten verstanden werden. Daten, die an das Forschungsdatenzentrum übertragen werden, werden nicht-exklusiven Vereinbarungen unterliegen, sodass die Urheberrechte vollständig bei dem Dateneigentümer verbleiben und dieser seine Inhalte und Erkenntnisse jederzeit anderweitig archivieren, publizieren und verbreiten kann. Auch legt dieser fest, ob und in welchem

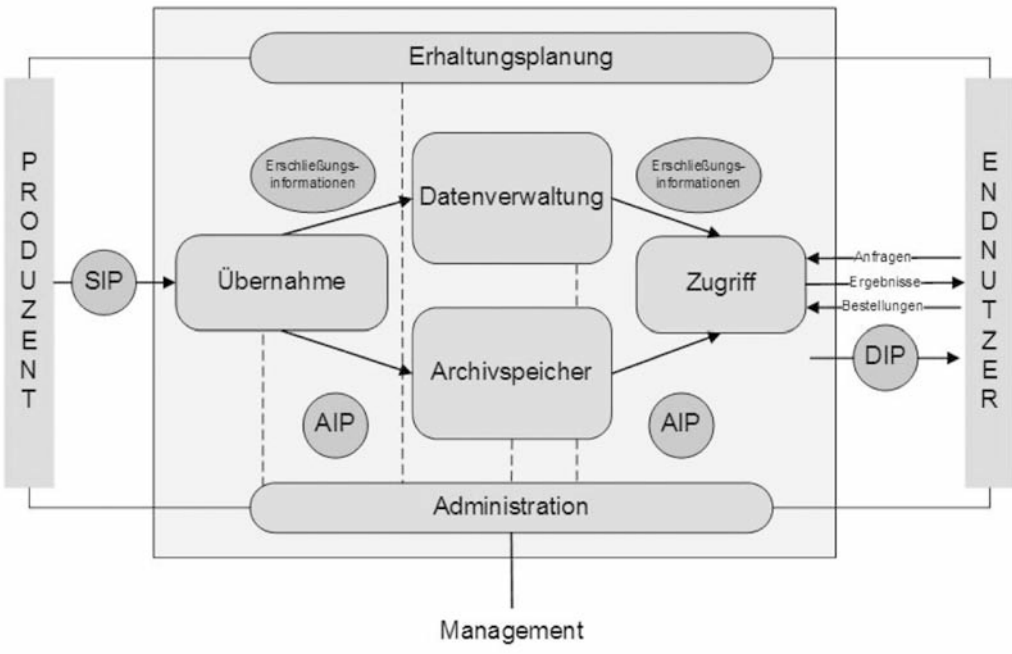
Umfang andere Personen und Gruppen auf die über IANUS bereitgestellten Daten zugreifen dürfen. Lizenzvereinbarungen zum einen zwischen Dateneigentümern und IANUS und zum anderen zwischen IANUS und seinen Nutzern werden den rechtlichen Rahmen bilden, um Fragen zu Haftung, Nutzung, Verantwortlichkeiten und Verpflichtungen verbindlich festzulegen. Aufbau und Funktionsweise von IANUS wird sich dabei an dem OAIS-Referenzmodell orientieren, das zuletzt 2012 als ISO-Standard 14721:2012 für Langzeitarchive veröffentlicht wurde[19].

Das Modell beschreibt und vereinheitlicht zentrale Konzepte und Prozesse für Organisationen, deren Ziel es ist, archivierte Daten für eine vorgesehene Nutzergruppe auf unbestimmte Zeit zur Verfügung zu stellen. Vor allem bei digitalen Dokumenten geht es dabei einerseits um die langfristige technische Interpretierbarkeit von Dateien durch Computer und die Nachvollziehbarkeit der in diesen enthaltenen Inhalte durch Menschen, andererseits um die Bitstream Preservation, also den physischen Erhalt von elektronischen Datenobjekten auf physikalischen Speichermedien[20]. Im OAIS werden hierzu drei verschiedene Information Packages definiert. Das erste beschreibt als Submission Information Package (SIP) die Übergabe von Daten und zugehörigen Metadaten von einem Produzenten an ein Archiv; das Archive Information Package (AIP) bezeichnet die standardisierte, vailidierte und archivierungswürdige Version eines SIP; und zur Bereitstellung der Daten an Endnutzer wird aus einem AIP ein „optimiertes" Dissemination Information Package (DIP) erzeugt.

Auch wenn IANUS sich aktuell noch in der Konzeptions- und Planungsphase befindet und viele Ziele und Anfoderungen noch gemeinsam diskutiert werden, zeichnen sich bereits jetzt bestimmte zentrale Dienstleistungen ab:

Langzeitarchivierung: Statische, weitgehend unveränderliche Daten aus Projekten und Institutionen können – zum Beispiel nach dem Ende einer Förderung – bei IANUS abgegeben werden, wo sie nach definierten Workflows und Standards formal überprüft, semantisch vervollständigt, technisch aufbereitet und ausführlich dokumentiert werden, bevor sie in ein Offline-Archivsystem zur Bitstream Preservation überführt werden.

Datenbereitstellung: Alle oder Teile der langzeitarchivierten Daten werden mittels festgelegter Regeln und Rechte-Rollen-Konzepte über ein Online-Portal und über definierte Schnittstellen zur Nachnutzung für neue Projekte und Forscher zur Verfügung gestellt. Die zugehörigen Metadaten werden im Sinne eines Findbuches durchsuchbar gemacht, wobei auch hier Nutzungseinschränkungen vorliegen können. Die Zugriffsmöglichkeiten auf die Dokumente können dabei individuell gesteuert werden, sodass die Sichtbarkeit von OpenAccess-Publikationen bis hin zu unzugänglichen, geschützten Archiven variieren kann. Um eine eindeutige Referenzierbarkeit von digitalen Ressourcen in IANUS zu ermöglichen, werden Persistente Identifikatoren (PIDs)[21] vergeben.



4 Das OAIS-Referenzmodell mit sechs Funktionseinheiten, drei Informationspaketen und verschiedenen Kommunikationswegen innerhalb eines Archives (Quelle: Siehe Anmerkung 19, S. 127).

Nachweiskatalog: Vorgesehen ist eine zentrale Internetadresse, in der Fundstellen, Objekte, Maßnahmen beziehungsweise Aktivitäten, Forschungsprojekte, Akteure und Archivsammlungen nachgewiesen werden können. Sofern weitere Informationen zu einem Katalogeintrag digital vorliegen und online adressierbar sind, führen Referenzen entweder auf Archivinhalte innerhalb von IANUS oder auf Web-Ressourcen anderer Institutionen. Bei nur lokal vorliegenden analogen oder digitalen Dokumenten ist die Nennung der besitzenden Institution, einer Anschrift und einer zuständigen Kontaktperson denkbar.[22]

Leitfäden und Best Practice Guides: Ein zentrales Informationsportal soll fachspezifische Leitfäden, Standards, Anleitungen, Tutorials, Softwaretools und Beispiele für die Anwendung von IT-Techniken in den Altertumswissenschaften zur Verfügung stellen. Die Informationen werden dabei alle Phasen des Lebenszyklus[23] adressieren und dank des Einsatzes von Wikis, Foren oder ähnlichen Web2.0-Techniken die Möglichkeiten zur kollaborativen Weiterentwicklung geben, zum Beispiel für die Spezifizierung allgemein gehaltener Vorgaben innerhalb einzelner Institutionen und Projekte. Hier sind bereits erhebliche Anstrengungen unternommen worden:

Im Jahr 2009 wurde auf Initiative des DAI von mehreren Autoren ein „Leitfaden zur Anwendung von Informationstechnik in der archäologischen Forschung" verfasst, dessen erster, kürzerer Teil „Mindeststandards" definiert und dessen zweiter Teil als „Praxisratgeber" ausführlichere Erläuterungen zu diesen enthält. Neben verschiedenen internationalen und fachunabhängigen Standards wurden bei der Erstellung unter anderem auch die existierenden Richtlinien der unterschiedlichen Landesämter für (Boden)Denkmalpflege berücksichtigt und in den Text einbezogen[24]. Seit Fertigstellung der ersten Version wird der Leitfaden zurzeit über die Homepage des Deutschen Archäologischen Institutes online zur Verfügung gestellt[25]. Mittelfristig soll er über IANUS publiziert, aktualisiert und inhaltlich im Sinne eines „Guide to Good Practice" fortentwickelt werden[26]. Bislang fehlende Abschnitte zu naturwissenschaftlichen Daten in der Altertumsforschung und zum Umgang mit historischen Texten und Sprachen werden gegenwärtig ausformuliert, um das gesamte Fachspektrum in den Altertumswissenschaften abzubilden.

Projektunterstützung: Laufende Projekte mit einem hohen Anteil an dynamischen, häufig veränderten Daten sollen durch ein Ablagesystem unterstützt werden, das über eigene institutionelle (technisch-rechtliche) Grenzen hinweg die Freigabe von Dateien zur gemeinsamen Nutzung erlaubt und diese zeitnah automatisch synchronisiert.[27] Durch den Zugriff auf ein vollständiges und aktuelles Projektverzeichnis soll das Forschungsdatenmanagement bei mehreren Kooperationspartnern in unterschiedlichen Einrichtungen erleichtert und eine regelmäßige automatische Datensicherung vereinfacht werden.

Weiterbildung/Qualifizierung: Für Nachwuchswissenschaftler an Universitäten und berufstätige Forscher mit abgeschlossenen Qualifikationen sollen Lehrveranstaltungen, Summer Schools oder Schulungen angeboten und Online-Materialien aufbereitet werden, in denen notwendige IT-Kenntnisse für die Erhebung, Kuratierung und Archivierung von digitalen Forschungsdaten vermittelt werden. Parallel dazu sollen inhaltliche Standards definiert werden, die in die Studiengänge der Hochschulen einfließen können, um die Diskrepanz zwischen der Ausbildungssituation und den tatsächlichen Anforderungen im Berufsleben schließen zu helfen.

Abgesehen von diesen grundsätzlichen, längerfristigen Zielen sollen während der ersten, dreijährigen Projektphase neben den verschiedenen informativen, diskursiven und konzeptuellen Arbeitsschritten auch sogenannte Testbeds realisiert werden. Anhand von vier exemplarischen Anwendungsfällen werden prototypische Softwarelösungen entwickelt, um das Forschungsdatenzentrum nicht nur theoretisch voranzutreiben, sondern um auch praktische Erfahrungen zu sammeln, die Nutzbarkeit existierender Lösungen für IANUS zu testen und – zusätzlich zu den bereits fortgeschrittenen IT-Empfehlungen – erste, von Nutzern der Fachcommunity erfahrbare Ergebnisse zu präsentieren.

Zunächst ist in diesem Zusammenhang eine erste, einfache Version des oben skizzierten Nachweiskatalogs geplant, der auf einer zentralen Webseite nicht nur über die zukünftigen Archivbestände von IANUS informiert, sondern auch Auskunft über analoge und digitale Dokumente in anderen Institutionen gibt, zum Beispiel bei Denkmalfachbehörden, Akademien, Universitäten oder Museen. In welchem Umfang Grundinformationen, etwa zu archäologischen Einheiten (Fundstellen, Denkmäler, Objekte, Aktivitäten), wissenschaftlichen Akteuren (Personen, Projekte, Institutionen) oder Archivmaterialien (Dokumentationen, Akten, Berichte), zur Verfügung gestellt werden und in welcher Weise auf diese zugegriffen werden darf, entscheiden dabei die jeweiligen Dateneigentümer. Als eines der Datenmodelle für diesen Nachweiskatalog bietet sich das ADeX-Format an[28], das von der Kommission „Archäologie & Informationssysteme" im Verband der Landesarchäologen Deutschlands zum Austausch von raumbezogenen archäologischen Metadaten zwischen verschiedenen Ämtern und Ländern entwickelt wurde.

Des Weiteren ist in diesem Rahmen der Testbetrieb eines „Sync-&-Share"-Dienstes geplant, der zur Synchronisation von häufig genutzten Projekt-Verzeichnissen und -Dateien, die auf einem oder mehreren zentralen Servern gespeichert werden, dient und gleichzeitig den lokalen Zugriff auf den jeweils aktuellen Datenbestand erlaubt. Dies soll insbesondere auch über die technisch-rechtlichen Grenzen einzelner Institutionen hinweg möglich sein, sodass auch jeweils externe Projektpartner auf gemeinsam relevante Daten zugreifen können. Auf diese Weise soll der einfache Austausch und das nachhaltige Management von Forschungsdaten bei mehreren verteilten Projektpartnern verbessert und auch eine automatische Datensicherung unterstützt werden. Als Alternativen zu dem kommerziellen Anbieter Dropbox[29] existieren derzeit zwei Lösungen, die für diese Anforderung geeignet erscheinen: PowerFolder[30], das voraussichtlich durch das Deutsche Forschungsnetzwerk als Service zur Verfügung gestellt werden wird, und eine von KIT Karlsruhe umgesetzte Adaption von ownCloud im Rahmen des DARIAH-DE-Projektes[31].

Ein drittes Testbed widmet sich ersten Fragen der praktischen Datenarchivierung. Dabei geht es zum einen um Fragen zur Vielfalt der Datenformate, notwendigen Metadaten und erforderlichen Arbeitsschritten, zum anderen aber auch um konkrete Tipps und Tools für Datenproduzenten. Ziel ist es hierbei, einzelnen Projekten oder Individuen erste technische Hilfestellungen zu geben, bevor diese Datenbestände IANUS zur Archivierung anvertraut werden. Dies bedeutet, es müssen bestehende Software-Werkzeuge angepasst beziehungsweise gegebenenfalls neu entwickelt werden, die dann zum Beispiel zur Überprüfung und Migration von Dateiformaten oder zur (semi-)automatischen Erzeugung von Angaben, die die zu archivierenden Dateien beschreiben (Metadaten), eingesetzt werden können.

Dem Aufgabengebiet der Datenbereitstellung ist schließlich der vierte Prototyp zuzurechnen, in dem es um die Vergabe sogenannter Persistenter Identifier[32] gehen wird. Diese „ISBN-Nummern" für online veröffentlichte Digitale Ressourcen stellen eine zentrale Voraussetzung für die Zitierbarkeit von Dateien im Internet dar und sind eine wesentliche Basis für deren Interoperabilität mit anderen Fachsystemen. Grundlegende Anforderung an das Testbed ist folglich die Generierung und Registrierung von PIDs mithilfe etablierter nationaler Service-Anbieter, sodass am Ende eines zu noch definierenden Workflows innerhalb von IANUS einzelne, online-verfügbare Ressourcen eindeutig referenziert werden können.

Sowohl die sechs langfristigen als auch die vier kurzfristigen Aufgaben und Dienstleistungen von IANUS werden als essenziell notwendige und logische Bestandteile einer konsequenten, erfolgreichen und verantwortungsvollen Forschungsdatenspeicherung angesehen. Die langfristige Sicherung und die Bereitstellung von qualitätsgeprüften digitalen Forschungsdaten in den Altertumswissenschaften trägt nicht nur dazu bei, die 1997 von der DFG formulierten „Vorschläge zur Sicherung guter wissenschaftlicher Praxis" umzusetzen[33], sondern können, dank verbesserter Zugangsmöglichkeiten zu diesen, auch maßgeblich dazu beitragen, die fachliche und gesellschaftliche Relevanz der eigenen wissenschaftlichen, denkmalpflegerischen und bildungspolitischen Arbeit zu unterstreichen.

Anmerkungen

[1] Zu den Unzulänglichkeiten der Begriffe „Primär-/Rohdaten" und „Sekundärdaten" siehe J. Klump, Digitale Forschungsdaten, in: H. Neuroth u. a. (Hrsg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3 (2010) Kap.17:104 f. (http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0008-2010071949 [Zugriff am 25.11.2012]).

[2] Abbildung 2 beruht auf dem Model, das vom UK Data Archive publiziert wurde: http://www.data-archive.ac.uk/create-manage/life-cycle [Zugriff am 16.11.2012]. Ein komplexeres Model, das stärker noch den Aspekt der Datenkuratierung berücksichtigt, wurde vom Digital Curation Centre erstellt: http://www.dcc.ac.uk/resources/curation-lifecycle-model [Zugriff am 16.11.2012]. Ausführlich siehe S. Rümpel, Der Lebenszyklus von Forschungsdaten, in: S. Büttner – H.-C. Hobohm – L. Müller (Hrsg.), Handbuch Forschungsdatenmanagement (2011) 25–34.

[3] Zum Beispiel begegnet die Nicht-Reproduzierbarkeit von Daten auch in der Klimaforschung, die unter anderem auf einmaligen Wettermessungen basiert, oder in den Sozialwissenschaften, in denen viele Umfragen nur eine Momentaufnahme wiedergeben.

[4] Siehe hierzu verschiedene Beiträge in H. Neuroth u. a. (Hrsg.), Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme (2012).

[5] http://www.ianus-fdz.de [Zugriff am 25.11.2012].

[6] Ausführlich: http://www.ianus-fdz.de/projects/ zentr-dig-arch/wiki/Zusammensetzung-DFG-Arbeitsgruppe [Zugriff am 25.11.2012].

[7] Vor allem sind hier zu nennen D-GRID (http://www.d-grid.de), TextGRID (http://www.text-grid.de), WissGRID (http://www.wissgrid.de), DARIAH (http://www.dariah.eu) und CLARIN (http://www.clarin.eu) [Zugriff am 25.11.2012].

[8] Zum Beispiel ArcheoInf als GIS-Frontend für heterogene Datenbanken (http://www.archeoinf.de), eAqua mit seinen Tools für text-mining (http://www.eaqua.net), eSciDoc als Fedora basiertes Repositorium der Max Planck Digital Library (http://www.escidoc.org), Propylaeum als virtuelle Fachbibliothek der Altertumswissenschaften (http://www.propylaeum.de) und Arachne als größte deutschsprachige Online-Datenbank zu antiken Objekten (http://www.arachne.uni-koeln.de) [Zugriff am 25.11.2012].

[9] http://idw-online.de/pages/de/news445078. [Zugriff am 25.11.2012].

[10] Siehe hierzu die aktuelle Studie: CCeH, Universität zu Köln (Hrsg.), Digitale Geisteswissenschaften. Erstellt im Rahmen des Projektes DARIAH-DE (2011) (http://www.dig-hum.de/sites/default/files/cceh_broschuereweb.pdf [Zugriff am 25.11.2012]).

[11] Vor allem „kopal – Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen" (http://kopal.langzeitarchivierung.de [Zugriff am 25.11.2012]) und „Nestor – Deutsches Kompetenznetzwerk zur digitalen Langzeitarchivierung" (http://www.langzeitarchivierung.de [Zugriff am 25.11.2012]).

[12] Allein die Landesarchäologien in den 16 Bundesländern besitzen für den Umgang mit Ausgrabungsdaten (analog wie digital) und -objekten jeweils länderspezifische Regelungen, Infrastrukturen und Ressourcen.

[13] http://www.edna.nl [Zugriff am 25.11.2012].

[14] http://archaeologydataservice.ac.uk [Zugriff am 25.11.2012].

[15] http://www.gesis.org/das-institut/wissenschaftliche-abteilungen/datenarchiv-fuer-sozialwissenschaften/ [Zugriff am 25.11.2012].

[16] http://www.pangaea.de/ [Zugriff am 25.11.2012].

[17] http://tla.mpi.nl [Zugriff am 25.11.2012].

[18] http://www.geoportal.de/DE/GDI-DE/ [Zugriff am 25.11.2012].

[19] Nestor-Arbeitsgruppe OAIS-Übersetzung (Hrsg.), Referenzmodell für ein Offenes Archiv-Informations-System. Deutsche Übersetzung, nestor-materialien 16 (2012). (http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0008-2012051101 [Zugriff am 25.11.2012])

[20] D. Ullrich, Bitstream Preservation, in: H. Neuroth u. a. (Hrsg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3 (2010) Kap.8:3-9 (http://nbn-resolving.de/urn/resolver.pl?-urn:nbn:de:0008-2010071949
[Zugriff am 25.11.2012]).

[21] PIDs können als ISBN-Nummern für digitale Ressourcen im Internet verstanden werden, ausführlicher: K. Schroeder, Persistent Identifier (PI) – ein Überblick, in: H. Neuroth u. a. (Hrsg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3 (2010) Kap.9:22-45 (http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0008-2010071949 [Zugriff am 25.11.2012]).

[22] Vgl. ergänzend unten die Angaben zum skizzierten Testbed.

[23] Siehe Anm. 2.

[24] Ein inhaltlicher Abgleich mit dem jüngst erstellten „Ratgeber zur Archivierung digitaler Daten" aus der AG Archivierung im Verband der Landesarchäologen ist vorgesehen.
http://www.landesarchaeologen.de/fileadmin/Dokumente/Dokumente_Kommissionen/Dokumente_Archaeologie-Informationssysteme/Dokumente_AIS_Archivierung/Ratgeber-Archivierung_Vorab-V0.07.pdf [Zugriff am 25.11.2012].

[25] http://www.dainst.org/de/project/it-leitfaden?ft =all [Zugriff am 25.11.2012].

[26] Voraussichtlich ab Januar 2014 unter http://www.ianus-fdz.de/it-empfehlungen.

[27] Als wichtigste Anbieter dieser Sync-&-Share-Lösungen sind PowerFolder, ownCloud und Dropbox zu nennen.

[28] http://www.landesarchaeologen.de/kommissionen-/komm_informationssysteme/ag_is_adex_00.html [Zugriff am 25.11.2012].

[29] http://www.dropbox.com [Zugriff am 25.11.2012]

[30] http://wiki.gwdg.de/index.php/PowerFolder [Zugriff am 25.11.2012].

[31] http://de.dariah.eu [Zugriff am 25.11.2012].

[32] Siehe Anm. 21.

[33] DFG Denkschrift: Sicherung guter wissenschaftlicher Praxis (1998).
(http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf [Zugriff am 25.11.2012]).

Workshop „Archäologie und Informationssysteme“ am 19./20. November 2012 im Niedersächsischen Landesamt für Denkmalpflege

Otto Mathias Wilbertz

Die Kommission „Archäologie und Informationssysteme“ des Verbandes der Landesarchäologen in der Bundesrepublik Deutschland (VLA) veranstaltete gemeinsam mit dem Niedersächsischen Landesamt für Denkmalpflege (NLD) am 19. und 20. November 2012 in Hannover einen Workshop, bei dem Themen aus der Arbeit der Kommission erörtert wurden. Das NLD übernahm die Einladungen, die Organisation vor Ort und stellte seine Infrastruktur zur Verfügung. Der Kommission oblag die inhaltliche Vorbereitung und Strukturierung der Veranstaltung.

Nach mehr als siebenjähriger Tätigkeit wollte die Kommission detaillierte Einblicke in die bisher geleistete Arbeit geben. Daran anknüpfend sollten Leitlinien für ihre weitere Arbeit entwickelt werden.

Eingeladen waren Kolleginnen und Kollegen aus den Mitgliedsinstitutionen des VLA sowie als Gastreferenten Dr. Felix Schäfer vom Deutschen Archäologischen Institut (DAI), der das Projekt „IANUS – Forschungsdatenzentrum Archäologie & Altertumswissenschaften“[1] vorstellte, und Sascha Kuhnt vom Landesamt für Geoinformation und Landentwicklung Niedersachsen (LGLN), Leiter der Arbeitsgruppe Geodatenportal der GDI-NI[2], der über den INSPIRE[3]-Prozess informierte.

Nach der Begrüßung durch den Präsidenten des NLD, Herrn Dr. Winghart, und Grußworten des Vorsitzenden des VLA, Herrn Prof. Dr. Kunow, gab der Sprecher der Kommission, Dr. Wilbertz, einen kurzen einführenden Überblick über die Stationen der Kommissionsarbeit. Anschließend wurden Arbeitsergebnisse der Arbeitsgruppen Archivierung und Thesaurusfragen vorgestellt und diskutiert.

Am Nachmittag des ersten Tages folgten mehrere Impulsreferate zum Thema INSPIRE. An die Ausführungen von S. Kuhnt über die technischen und administrativen Aspekte knüpfte sich eine ausführliche Diskussion.

Anschließend berichteten verschiedene Kollegen über den Stand des INSPIRE-Prozesses in ihren Ländern beziehungsweise Ämtern und R. Göldner schilderte Erfahrungen der spanischen und schottischen Kollegen mit INSPIRE. Das Referat von S. Kuhnt bot vielfältige neue Einblicke, die in der anschließenden Diskussion weiter vertieft werden konnten.

Den Abschluss des ersten Tages bildete das Referat von F. Schäfer über das IANUS-Projekt, das auch in diesem Band ausführlich vorgestellt wird.

1 Der Präsident des Niedersächsischen Landesamtes für Denkmalpflege, Dr. Stefan Winghart, eröffnet den Workshop.

Am zweiten Tag wurde der Datenaustausch-Standard ADeX[4] behandelt. Nach einer kurzen Darstellung der bisherigen Ergebnisse durch R. Göldner erfolgte eine lebhafte Diskussion über die Weiterentwicklung des Standards.

In der anschließenden Pause kamen Moderatoren und Protokollanten der einzelnen Sessionen zusammen, um die Abschlusssession vorzubereiten. In dieser wurden dann die Ergebnisse zusammenfassend betrachtet.

Zum Ende des Workshops wurde seitens des NLD eine Führung durch die Restaurierungswerkstätten mit der Präsentation aktueller Highlights angeboten.

In der Nachbereitungsphase des Workshops entstand ein Ergebnisprotokoll, welches eine wichtige Grundlage der weiteren Arbeit bilden wird. Das Protokoll vereint die im Wesentlichen einvernehmlich erzielten Ergebnisse des Workshops sowie die von der Kommission aufzunehmenden offenen Fragen. Die Kommission wird ihr Arbeitsprogramm dementsprechend neu ausrichten.

2 Sascha Kuhnt, der Leiter der AG Geodatenportal der GDI-NI, beantwortet Fragen im Anschluss an seine Präsentation, hier mit Prof. Dr. Jürgen Kunow, dem Vorsitzenden des Verbandes der Landesarchäologen.

3 Sitzung des Workshops, stehend Dr. Reiner Göldner. (Alle Fotos: H. Haßmann, NLD)

Anmerkungen

[1] Zur Benennung des Projektes wurde kein Akronym, sondern der Name des römischen Gottes IANUS verwendet.

[2] GDI-NI = Geodateninfrastruktur Niedersachsen.

[3] INSPIRE = **IN**frastructure for **SP**atial **I**nfo**R**mation in **E**urope.

[4] ADeX = **A**rchäologischer **D**aten-**eX**port

Autorenverzeichnis

Dr. Anthony Beck
Project Champion, DART Project
School of Computing
University of Leeds
LS2 9JT, UK

Dipl.-Inf. Ralf Berger
interactive instruments GmbH
Trierer Straße 70–72
53115 Bonn

David Bibby
Landesamt für Denkmalpflege im
Regierungspräsidium Stuttgart
Archäologische Denkmalpflege: Zentrale Fachdienste
und Restaurierungswerkstatt
– digitale Archäologie –
Berliner Straße 12
73728 Esslingen

Dr. Utz Böhner
Informationsmanagement und
Fachinformationssystem
Niedersächsisches Landesamt für Denkmalpflege
Scharnhorststraße 1
30175 Hannover

Andreas Brunn
Referent Dokumentation/EDV
Landesamt für Denkmalpflege und Archäologie
Sachsen-Anhalt
– Landesmuseum für Vorgeschichte –
Richard-Wagner-Straße 9
06114 Halle/Saale

Prof. Dr. Ortwin Dally
Generalsekretär des Deutschen Archäologischen
Instituts
Podbielskiallee 69–71
14195 Berlin

Sigmar Fitting
Systemadministrator
Stabsstelle Zentrale Verwaltung und Marketing
Generaldirektion Kulturelles Erbe Rheinland-Pfalz
Schillerstraße 44 – Erthaler Hof
55116 Mainz

Prof. Dr. Friederike Fless
Präsidentin des Deutschen Archäologischen Instituts
Podbielskiallee 69–71
14195 Berlin

Prof. Dr. Reinhard Förtsch
Wissenschaftlicher Direktor
für Informationstechnologien
Deutsches Archäologisches Institut
Podbielskiallee 69–71
14195 Berlin

Dr. Reiner Göldner
Referent Geografisches Informationssystem
Landesamt für Archäologie
Zur Wetterwarte 7
01109 Dresden

Catherine Hardman
Deputy Director
Archaeology Data Service
Department of Archaeology
The King's Manor
University of York
YO1 7EP, UK

Dr. Henning Haßmann
Landesarchäologe
Niedersächsisches Landesamt für Denkmalpflege
Scharnhorststraße 1
30175 Hannover

Maurice Heinrich
Projektkoordination IANUS
Deutsches Archäologisches Institut
Podbielskiallee 69–71
14195 Berlin

Dr. Ulrich Himmelmann
Leiter Sachgebiet EDV/IT-Kommunikation
Stabsstelle Zentrale Verwaltung und Marketing
Generaldirektion Kulturelles Erbe Rheinland-Pfalz
Schillerstraße 44 – Erthaler Hof
55116 Mainz

Dr. Ulf Ickerodt
Archäologisches Landesamt Schleswig-Holstein
Obere Denkmalschutzbehörde
Schloss Annettenhöh
Brockdorff-Rantzau-Straße 70
24837 Schleswig

Dipl.-Ing. Thomas Poelmann
Archäologisches Landesamt Schleswig-Holstein
Geoinformation und IT-Management
Schloss Annettenhöh
Brockdorff-Rantzau-Straße 70
24837 Schleswig

Dr. Axel G. Posluschny M.A.
Projekt-Leiter EU-Projekt ArchaeoLandscapes Europe
Römisch-Germanische Kommission des
Deutschen Archäologischen Instituts
Palmengartenstraße 10–12
60325 Frankfurt/Main

Jörg Räther
Grabungstechnik + IT
Archäologisches Museum Hamburg
Bodendenkmalpflege
Museumsplatz 2
21073 Hamburg

Evelyn Rauchegger
ZBW – Deutsche Zentralbibliothek für
Wirtschaftswissenschaften
Leibniz-Informationszentrum Wirtschaft
Neuer Jungfernstieg 12
24105 Hamburg

Prof. Julian D. Richards
Director Archaeology Data Service
Department of Archaeology
The King's Manor
University of York
YO1 7EP, UK

Dipl.-Phys. Thomas Richter
Sachgebeitsleiter IT/GIS
Landesamt für Denkmalpflege und Archäologie
Sachsen-Anhalt
– Landesmuseum für Vorgeschichte –
Richard-Wagner-Straße 9
06114 Halle (Saale)

Dr. Felix F. Schäfer
Projektkoordination IANUS
Deutsches Archäologisches Institut
Podbielskiallee 69–71
14195 Berlin

Dipl.-Inf. Mario Schlapke
Fachreferent Informatik und Numismatik
Thüringisches Landesamt für Denkmalpflege und
Archäologie
Dienststelle Weimar
Humboldtstraße 11
99423 Weimar

Eicke Siegloff
Archäologisches Landesamt Schleswig-Holstein
Dezernat Denkmalschutz und Inventarisation
Schloss Annettenhöh
Brockdorff-Rantzau-Straße 70
24837 Schleswig

Dr. Bettina Stoll-Tucker
Abteilungsleiterin Übergreifende Fachdienste
Landesamt für Denkmalpflege und Archäologie
Sachsen-Anhalt
– Landesmuseum für Vorgeschichte –
Richard-Wagner-Straße 9
06114 Halle (Saale)

Dr. Markus Ullrich
Bayerisches Landesamt für Denkmalpflege
– Bayerische Denkmalliste und Denkmaltopographie –
Schloss Seehof
96117 Memmelsdorf

Dipl.-Geogr. Roland Wanninger
Bayerisches Landesamt für Denkmalpflege
– Denkmalerfassung und Denkmalforschung –
Fachinformationssystem
Hofgraben 4
80539 München

Dipl.-Phys. Bernd Weidner
interactive instruments GmbH
Trierer Straße 70–72
53115 Bonn

Dr. Otto Mathias Wilbertz
Kommission Archäologie & Informationssysteme des
Verbandes der Landesarchäologen
c/o Niedersächsisches Landesamt für Denkmalpflege
Scharnhorststraße 1
30175 Hannover

Dr. Stefan Winghart
Präsident
Niedersächsisches Landesamt für Denkmalpflege
Scharnhorststraße 1
30175 Hannover