



Paul Klimpel, Jürgen Keiper (Hrsg.)

**Was bleibt?**  
Nachhaltigkeit der Kultur  
in der digitalen Welt

Eine Publikation des  
Internet & Gesellschaft Collaboratory e. V.

**Was bleibt?**

Eine Publikation des Internet & Gesellschaft Collaboratory e. V.



Herausgeber: Paul Klimpel, Jürgen Keiper  
Autorenkontakte, Assistenz: Jana Maire  
Fotos (soweit nicht anders gekennzeichnet): Jürgen Keiper

Lektorat, Satz, E-Book und Produktion: Redaktion und Alltag GmbH

Umschlaggestaltung: Laura Oldenbourg  
Umschlagfotos: Jürgen Keiper  
Vorderseite: Tape Library im Konrad-Zuse-Zentrum für Informationstechnik Berlin  
Rückseite: Digitale Farbefekte bei Codec-Fehler

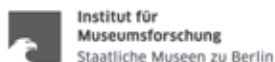
Texte und Fotos stehen (soweit nicht anders gekennzeichnet) unter einer Creative Commons  
Namensnennung - Weitergabe unter gleichen Bedingungen 3.0 Unported Lizenz.



1. Auflage  
September 2013  
Druck: Oktoberdruck  
Verlag: iRights.Media

ISBN-13: 978-3-944-36203-8

Partner der Initiative zur Nachhaltigkeit der Kultur in der digitalen Welt:



Unterstützer: EMC<sup>2</sup>, Google Deutschland

Paul Klimpel, Jürgen Keiper (Hrsg.)

**Was bleibt?**  
Nachhaltigkeit der Kultur  
in der digitalen Welt

Eine Publikation des Internet & Gesellschaft Collaboratory e. V.  
bei iRights.Media

Berlin 2013

# Inhalt

Vorwort des Lenkungskreises	7
<b>Einführung</b>	<b>9</b>
PAUL KLIMPEL Was bleibt?	11
<b>Berliner Appell</b>	<b>15</b>
Berliner Appell	17
JÜRGEN KEIPER Überblick über die Expertendiskussion zur Nachhaltigkeit	21
<b>Problemlage</b>	<b>25</b>
TOBIAS BEINERT/ARMIN STRAUBE Aktuelle Herausforderungen der digitalen Langzeitarchivierung	27
STEFAN WOLF Informationen für die Zukunft – EINE Infrastruktur?	47
ERIC STEINHAUER Wissen ohne Zukunft? Der Rechtsrahmen der digitalen Langzeitarchivierung von Netzpublikationen	61
<b>Nahaufnahmen</b>	<b>83</b>
MAIK STÜHRENBERG Digitale Langzeitarchivierung aus Sicht der IT	85
RALPH GIEBEL Speichertechnologie und Nachhaltigkeit	95
ANDREAS LANGE Die Gaming-Community als Pionier der digitalen Bewahrung	109
JAN FRÖHLICH Farbraum und Bildzustand im Kontext der Langzeitarchivierung	119

<b>Zwischenrufe</b>	<b>127</b>
KATHRIN PASSIG Facebook, Froschlaich und Folianten	129
STEFAN WILLER Kulturelles Erbe und Nachhaltigkeit	139
<b>Der Streit um die Zukunft: Auswahl, Beständigkeit und Vernetzung</b>	<b>155</b>
FELIX SASAKI Nachhaltigkeit und Semantic Web	157
STEFAN GRADMANN Semantische Nachhaltigkeit und Kontrolle	171
JENS BEST Wikipedia und Nachhaltigkeit	179
MICHAEL HOLLMANN Was wollen wir archivieren?	187
JENS CRUEGER Die Dark Ages des Internet?	191
GEORG REHM Eine Strategie zur Förderung der digitalen Langzeitarchivierung	199
<b>Texte, Quellen</b>	<b>219</b>
Textausschnitt aus dem Abschlussbericht der Internet Enquete Kommission des Deutschen Bundestags	221
Vancouver-Deklaration der UNESCO	223
Referenzen, Quellen	231
<b>Biographien</b>	<b>235</b>



## Vorwort des Lenkungskreises

JOHN H. WEITZMANN, PHILIPP S. MÜLLER,  
ULRIKE HÖPPNER, MARIANNE WULFF UND  
MARTIN G. LÖHE

Traditionell haben Gedächtnisinstitutionen wie Archive, Bibliotheken und Museen dafür gesorgt, dass die Zeugnisse unseres Schaffens der Nachwelt erhalten bleiben. Mit dem Übergang in die digitale Welt aber entstehen gänzlich neue Herausforderungen.

Wie sich die Rollen der Akteure ändern und welche neuen Aufgaben entstehen, war Thema der 8. Initiative „Nachhaltigkeit in der Digitalen Welt“ des Internet und Gesellschaft Collaboratory e. V. Die Experten haben sich unter anderem mit unterschiedlichen Strategien digitaler Langzeitarchivierung, dem Medienwechsel und den Rollen von privaten und öffentlichen Akteuren beschäftigt.

Die Treffen der von Paul Klimpel und Jürgen Keiper geleiteten Experten-Gruppe ermöglichten einen sachorientierten, inhaltlichen Diskurs, bei dem die organisatorische Einbindung der Teilnehmer und damit auch die Interessen der verschiedenen Institutionen in den Hintergrund traten. Dieser freie Diskurs war es, der die Arbeit der Initiative auszeichnete und aus dem heraus deutlich wurde, wo gehandelt werden muss. Der Berliner Appell zum Erhalt des digitalen Kulturerbes ist Ausdruck eines Konsenses, der über institutionelle und kulturpolitische Gräben hinweg besteht.

Die Frage, was mit den eigenen privaten Daten auf digitalen Trägern in 10 oder 20 Jahren sein wird, veranschaulicht gut, dass dieses Thema große Relevanz besitzt – für den Einzelnen wie für die Gesellschaft. Mit diesem Band hofft das CoLab, einen Beitrag zum Diskurs über Nachhaltigkeit in der digitalen Welt zu leisten – einem Diskurs, der noch nicht zu Ende ist und fortgesetzt werden muss.





Schrank mit Magnetbändern, Konrad-Zuse-Zentrum für Informationstechnik Berlin

# Einführung



# Was bleibt?

PAUL KLIMPEL

Was bleibt von den hunderten oder gar tausenden digitaler Fotos, die heute bei einer Hochzeit gemacht werden? Fotoalben von Hochzeiten im letzten Jahrhundert finden sich in den meisten Familien – auch wenn damals weniger fotografiert wurde als heute.

Was bleibt von den Objekten der Medienkunst, deren digitale Technik nach kurzer Zeit schon nicht mehr gebräuchlich ist? Was bleibt von den Ergebnissen der vielen Projekte, in denen heute kulturelles Erbe digitalisiert wird? Was wird aus den Sammlungen originär digitaler Objekte? Was bleibt von Filmen, die zunehmend digital in die Kinos kommen?

Was wird aus der großen Masse digitaler Information gesammelt?

Oder werden die Zeugnisse des kulturellen Schaffens und der Wissenschaft, die auf digitalen Trägern festgehalten werden, schon bald verschwinden? Werden die ersten Jahre nach der digitalen Revolution in der Erinnerung späterer Generationen ein „weißer Fleck“ sein?

Traditionell haben Gedächtnisinstitutionen wie Archive, Bibliotheken und Museen dafür gesorgt, dass die Zeugnisse kulturellen Schaffens für die Nachwelt erhalten bleiben. Doch mit dem Übergang zu digitalen Trägern entstehen gänzlich neue Herausforderungen. Diese betreffen die traditionellen Gedächtnisorganisationen ebenso wie die neuen Formen des kulturellen Gedächtnisses, etwa kollaborative Online-Projekte wie Wikipedia oder History Pin. Es werden neue technische und organisatorische Kompetenzen gefordert, wenn es um Fragen der digitalen Langzeitarchivierung geht.

Nur einige Fragen, die sich stellen: Wie ändert sich durch die Digitalisierung die Rolle von Archiven, Bibliotheken und Museen? Welche Aufgaben nehmen heute Rechenzentren und privatwirtschaftlich organisierte Dienstleister wahr? Wer trägt die Verantwortung für die digitale Langzeitarchivierung, wer finanziert sie? Was bedeutet der Medienwechsel für die kleinen und mittleren Museen und

Archive und für die neuen Gedächtnisorganisationen und ihre Rolle in dem Ensemble kultureller Institutionen? Wie lässt sich ein „flüchtiges Medium“ wie das Internet archivieren? Wer entscheidet über Auswahl? Welche Rolle spielt die Semantik für die Nachhaltigkeit?

In Archiven, Bibliotheken und Museen weiß man aus Erfahrung: Bei jedem neuen Medium sind die Anfangsjahre nicht erhalten. Dies war beim Radio so, auch beim Film oder beim Fernsehen. Die Veränderungen durch die Digitalisierung sind sehr viel weitreichender und dringen in jeden Lebensbereich. Die Herausforderungen an die nachhaltige Sicherung sind gewaltig.

Inhalte auf analogen Trägern bleiben grundsätzlich erhalten, sofern man sie nicht aktiv zerstört. Probleme der Erhaltung treten erst nach einer gewissen Zeit auf, in der Regel vergeht mindestens eine Generation, bis sie akut werden. Bei Inhalten auf digitalen Trägern verhält es sich genau umgekehrt: Um sie angesichts häufiger Formatwechsel und der begrenzten Haltbarkeit von Trägermedien erhalten zu können, muss aktiv gehandelt werden.

Wir sind am Scheideweg: Nehmen wir diese Herausforderung an, oder wird sich wiederholen, was wir bei Radio, Film und Fernsehen erlebt haben? Es geht darum, was von den Zeugnissen menschlichen Tuns unserer Zeit – der ersten Jahren der Digitalisierung – in Zukunft erhalten bleibt. „In Zukunft“ heißt in diesem Falle schon in den nächsten 5, 10 oder 20 Jahren. Nachhaltigkeit in der digitalen Welt ist deshalb kein Nischenthema.

Die Herausforderungen sind gewaltig. In den vergangenen Jahren sind bereits wichtige Ansätze formuliert worden, wie diesen Herausforderungen zu begegnen ist. Diese Publikation nun gibt Zeugnis eines Diskurses, der Nachhaltigkeit und digitale Langzeitarchivierung nicht als ein isoliertes technisches, archivarisches, kulturpolitisches, rechtliches oder finanzielles Problem sieht, sondern die Interdependenz verschiedener Aspekte in den Vordergrund stellt. Diese Betrachtung ist notwendig, will man die Rahmenbedingungen schaffen, Informationen und Wissen im Zeitalter der Digitalisierung zu sichern, zu bewahren und zugänglich zu machen.

So wird in diesem Buch die Problemlage aus unterschiedlicher Perspektive beschrieben. Tobias Beinert und Armin Straube erläutern die Sicht der Gedächtnisinstitutionen. Stefan Wolf geht auf das Verhältnis von kulturellem Erbe und dem Wissenschaftsbereich ein und beschreibt die Bemühungen um eine Infrastruktur zur digitalen Langzeitarchivierung. Eric Steinhauer schließlich nimmt eine juristische Perspektive ein. In Nahaufnahmen werden Einzelaspekte beleuchtet: So

erörtert Maik Stührenberg Fragen der digitalen Langzeitarchivierung aus Sicht der IT, und Ralph Giebel zeichnet die Entwicklung der Speichertechnologie nach und geht auf die Rolle der Industrie ein, in einer Welt, in der Informationen und große Datenmengen immer wichtiger werden. In anderen Beiträgen stehen medienspezifische Besonderheiten im Fokus: Andreas Lang beschreibt, wie in offenen, nicht-kommerziellen und kollaborativen Projekten Werkzeuge und Methoden entwickelt wurden, um Videospiele zu erhalten, etwa mit Hilfe von Emulatoren. Jan Fröhlich beschreibt die Herausforderungen der Langzeitarchivierung bei Farbraum und Bildzustand und geht dabei der Frage nach, welchen Charakter und welche Färbung audiovisuelle Inhalte unserer Tage in Zukunft haben werden. Eine gänzlich andere Perspektive bieten Kathrin Passig, wenn sie einen allgemeinen Blick auf Aspekte der Überlieferung nimmt, aber auch Stefan Willer, der den Begriff des kulturellen Erbes kritisch hinterfragt. Kontrovers ringen die Autoren um Fragen der Nachhaltigkeit, der Auswahl und der Semantik des Wissens. Während Felix Sasaki verschiedene Ansätze für die Semantik des Netzes darstellt und ihre Auswirkungen auf die Nachhaltigkeit beleuchtet, positioniert sich Stefan Gradmann klar gegen alle Versuche, Strukturen aufzubauen, die nach seiner Meinung Kontrolle und Zensur im Netz ermöglichen. Jens Best beantwortet Fragen zur Nachhaltigkeit der Wikipedia. Der Präsident des Bundesarchivs, Michael Hollmann, erläutert traditionelle, auf die Relevanz für staatliches Handeln hin orientierte Auswahlkriterien. Jens Cruieger warnt vor dem Verlust von Informationen für die spätere Geschichtsschreibung und entwickelt ein weiteres Verständnis dessen, was zu bewahren ist. Schließlich plädiert Georg Rehm eindringlich für die Entwicklung einer Strategie zur digitalen Langzeitarchivierung.

Die Beiträge entstanden im Rahmen einer Initiative des Internet und Gesellschaft Collaboratory e. V. Zwischen März und Juni 2013 diskutierten Experten Fragen der Nachhaltigkeit von Kultur in der digitalen Welt. Die Initiative fand statt in Kooperation mit dem Institut für Museumsforschung, dem Kompetenznetzwerk zur digitalen Langzeitarchivierung, nestor, und dem iRights.Lab Kultur. Neben einem wertvollen Diskurs über zahlreiche Einzelaspekte konnten sich diese Experten auf den „Berliner Appell“ verständigen, der die gesellschaftliche Bedeutung des Themas hervorhebt und konkrete Handlungsfelder benennt.

Was bleibt also?

Die Erkenntnis, dass gehandelt werden muss, damit etwas bleibt in der digitalen Welt.



Überwachung Stromnetz, Konrad-Zuse-Zentrum für Informationstechnik Berlin

# Berliner Appell





# Berliner Appell

Unsere Gesellschaft ist seit langem von dem Konsens bestimmt, dass Wissen und Kultur zu erhalten sind. Das Wissen unserer Tage wie die Kultur unserer Gesellschaft werden aber zunehmend mittels elektronischer Medien gespeichert und sollen über diese überliefert werden. Die Bewahrung dieses Wissens und dieser Kultur steht auf tönernen Füßen.

Die Bereitschaft, auch in der elektronischen Welt in die Bestandserhaltung zu investieren, ist nicht sehr groß. Die Einsicht in die Notwendigkeit, die Prozesse der digitalen Langzeitarchivierung nachhaltig und dauerhaft zu finanzieren, ist noch nicht ausgeprägt. Die Möglichkeiten der Informations- und Kommunikationstechnologie zur Bestandserhaltung auch von analogen Informationsträgern sind noch lange nicht ausgeschöpft.

Das ist ein Appell zur Nachhaltigkeit in der elektronischen Welt.

## *1. Gefahr des Verlustes*

Digitale Inhalte sind fragil. Sie unterliegen einer rasanten technischen Entwicklung. Es besteht dringender Handlungsbedarf bei der Archivierung, da digitale Inhalte und Informationen sonst unwiederbringlich verloren sind.

## *2. Dauerhafte Aufgabe*

Digitale Langzeitarchivierung ist Teil der Bestandserhaltung für digitale Objekte und muss als eine dauerhafte Aufgabe begriffen werden, die sich nicht in Projekten erschöpft.

### *3. Ausbildung / Organisation*

Digitale Langzeitarchivierung muss ein gezielter Ausbildungs- und Forschungsschwerpunkt an Universitäten und Fachhochschulen werden und auch Eingang in die Curricula von anderen Disziplinen finden.

### *4. Recht*

Der derzeitige Rechtsrahmen behindert vielfach die digitale Langzeitarchivierung. Es müssen eindeutige und verlässliche rechtliche Rahmenbedingungen für die digitale Langzeitarchivierung in all ihren Aspekten geschaffen werden.

### *5. Kosten*

Für die digitale Archivierung besteht eine dauerhafte öffentliche Verantwortung. Digitale Langzeitarchivierung ist kostenintensiv.

### *6. Aufmerksamkeit und öffentlicher Diskurs*

Nachhaltigkeit in der digitalen Welt erfordert eine breite öffentliche Diskussion und starke politische Wahrnehmung. Digitale Langzeitarchivierung ist kein Nischenproblem.

### *7. Langzeitarchivierung und Digitalisierung*

Digitale Langzeitarchivierung sichert und stärkt das Demokratie- und Transparenzversprechen des digitalen Kulturerbes.

### *8. Rollen und Strategie*

Zuständigkeiten und Rollen im Bereich der digitalen Langzeitarchivierung müssen im Rahmen einer nationalen bzw. europäischen Strategie klar bestimmt sein.

### *9. Auswahl*

Auswahlkriterien für die digitale Langzeitarchivierung müssen Teil eines gesellschaftlichen Diskurses sein.

### *10. Reichweite*

Kulturarchivierung ist im Digitalen eine Aufgabe, die in ihrer Bedeutung und ihren Folgen über die Aufgaben der Kulturinstitutionen hinausgeht.

*11. Recherchierbarkeit, Verfügbarkeit und Zugang*

Der Zugang zum digitalen Erbe ist durch neue Technologien so leicht wie noch nie. Er birgt einen großen Nutzen für die Bildung und Forschung. Um den Zugang zu erhalten, sind verlässliche Finanzierungsmodelle notwendig.

*12. Technische Fragen*

Langzeitarchivierung kann nur unter geregelten technischen Rahmenbedingungen nachhaltig erfolgen. Dazu gehören offene und standardisierte Datenformate.

THOMAS BÄHR, TOBIAS BEINERT, WINFRIED BERGMAYER,  
CHRISTOPH BRUCH, JENS CRUEGER, JAN ENGELMANN,  
ELLEN EULER, FREDERIK FISCHER, JAN FRÖHLICH, RALPH GIEBEL,  
STEFAN GRADMANN, MONIKA HAGEDORN-SAUPE,  
MICHAEL HOLLMANN, JÜRGEN KEIPER, PAUL KLIMPEL,  
ANDREAS LANGE, VERENA METZE-MANGOLD, GEORG REHM,  
THORSTEN SIEGMANN, BEATE RUSCH, RALF SCHNEIDER,  
ERIC STEINHAUER, TOBIAS STEINKE, ARMIN STRAUBE,  
MAIK STÜHRENBERG, ANNE VON RAAY, CHRIS WAHL,  
KATRIN WELLER, STEFAN WOLF

*[www.berliner-appell.org](http://www.berliner-appell.org)*



# Überblick über die Expertendiskussion zur Nachhaltigkeit

JÜRGEN KEIPER

Das Besondere der Arbeit im CoLab war das Zusammentreffen von Experten aus unterschiedlichsten Disziplinen. Dies allein garantiert zwar noch keine Interdisziplinarität, sorgte aber für einen erstaunlich offenen Diskurs und eine Atmosphäre, die weniger von Belehren und Erläutern geprägt war als vielmehr von der Vernetzung und Weiterführung von Gedanken. Die Machtansprüche unterschiedlicher Akteure waren damit keineswegs verschwunden, schoben sich aber nicht mehr in den Vordergrund, wodurch der Blick für die vernünftigsten Lösungen geöffnet wurde.

Im Folgenden sollen hier gleichermaßen kursorisch wie abstrakt die wichtigsten Ergebnisse des Internet & Gesellschaft Collaboratory vorgestellt werden. Diese resultieren aus den Referaten, aus den Beiträgen der Teilnehmer und den Gesprächen, die vor und nach den eigentlichen Veranstaltungen geführt wurden.

Im Unterschied zur öffentlichen Diskussion spielte in dieser Diskussion die Haltbarkeit der Datenträger kaum eine Rolle. Mehr noch: Die digitale Langzeitarchivierung wurde fast nie als isoliertes technisches Problem begriffen. Nicht, dass es keine technischen Probleme gäbe, aber die eigentliche Herausforderung, dies wurde schnell klar, ist das Zusammenspiel der vier Themen Technologie, Organisation, Recht und Auswahl. Man kann diese auch in dem Akronym TORA zusammenfassen.

## *Technologie*

Seitens der Technologie kann die Frage der Bitstream Preservation, also des Sicherens des reinen Datenstromes, prinzipiell als gelöst betrachtet werden – wenn

man entsprechend aufwändige, und dies heißt aktuelle und redundante, Technik einsetzt. Deutlich wurde, dass Technologie im Bereich der digitalen Langzeitarchivierung nicht als etwas Statisches im Sinne von „store and forget“ verstanden werden kann, sondern kontinuierlicher Veränderung unterworfen ist. Auch das Anfassen der Objekte zwecks Migration wird sich oft nicht vermeiden lassen. Langzeitarchivierung als Prozess gedacht erfordert aber auch eine völlig andere Kostenstruktur als bisher bekannt. Ähnliches gilt für die Personalstrukturen, welche mit jenen der Informationstechnologie verschmelzen werden: Hochqualifizierte Mitarbeiter und kontinuierliche Weiterbildung werden unverzichtbar.

In der Diskussion angesprochen wurde auch das Problem der geeigneten Archivformate. Hier sind in wichtigen Bereichen wie dem Film noch Fragen offen. Auch die Frage, welche digitale Version denn überhaupt archiviert werden soll, – und damit die Loslösung vom traditionellen Konzept des Originals – tauchte immer wieder als Thema auf.

### *Organisation*

Die Frage der Organisation ist eine zentrale. Denn aufgrund der hohen Kosten ist eine zufällige Archivierung zu teuer. Die Kette von Akquise und Ingest in den Archivworkflow ist nicht für alle Medien gleichermaßen geklärt, zumal wenn sich die digitalen Medien seitens der Vertriebe weg von Erwerbsmodellen (Kauf eines physischen Objektes) und hin zu Nutzungsmodellen (eingeschränkte Bereitstellung von Inhalten) verschieben.

Diskutiert wurde auch die Rolle der Verantwortlichkeiten. So die Frage, ob digitale Langzeitarchivierung nicht auch durch selbstständige, von der Industrie getragene Services jenseits von staatlichen Strukturen abgebildet werden kann.

Offen blieb auch – trotz einer Diskussion der zwei vorliegenden Studien zum Aufbau einer Archivierungsinfrastruktur –, wie zentrale und dezentrale Strukturen, föderale, nationale und europäische Interessen vereinbart werden können.

### *Recht*

Zum Thema Recht lässt sich schlicht zusammenfassend sagen, dass eine Archivierung von digitalen Objekten sehr schnell an massive rechtliche Grenzen stoßen kann – und in der Praxis zahlreicher Gedächtnisorganisationen auch stößt. In der Diskussion wurde aber auch deutlich, dass dies nicht für Archive, Biblio-

theken und Museen gleichermaßen gilt. Hier spielen unterschiedliche Regelungen für unterschiedliche Institutionen eine bedeutende Rolle.

### *Auswahl*

Die Archivierung hat bekanntlich mit einer zentralen inhaltlichen Paradoxie zu kämpfen. Einerseits ist eine Auswahl zu treffen, andererseits gibt es keine verlässlichen Prognosen über zukünftige Interessen. Felix Sasaki erläuterte am Beispiel der Anschläge vom 11. September 2001, wie historische Ereignisse völlig neue Interessen seitens der Politik und der Wissenschaft hervorbringen können und den Rückgriff auf Archivmaterial notwendig machen. Insbesondere bei dem Themenkomplex Webarchivierung wurde diese Paradoxie deutlich.

Dies sind die Kernpunkte der TORA, der vier Felder Technologie, Organisation, Recht und Auswahl. Vieles davon ist nicht neu, sondern wurde auch schon im Rahmen von nestor, dem Kompetenznetzwerk zur digitalen Langzeitarchivierung, benannt.

Jenseits dieser vier Felder haben wir uns zwei weiteren Bereichen zugewandt, die weniger im Zentrum stehen.

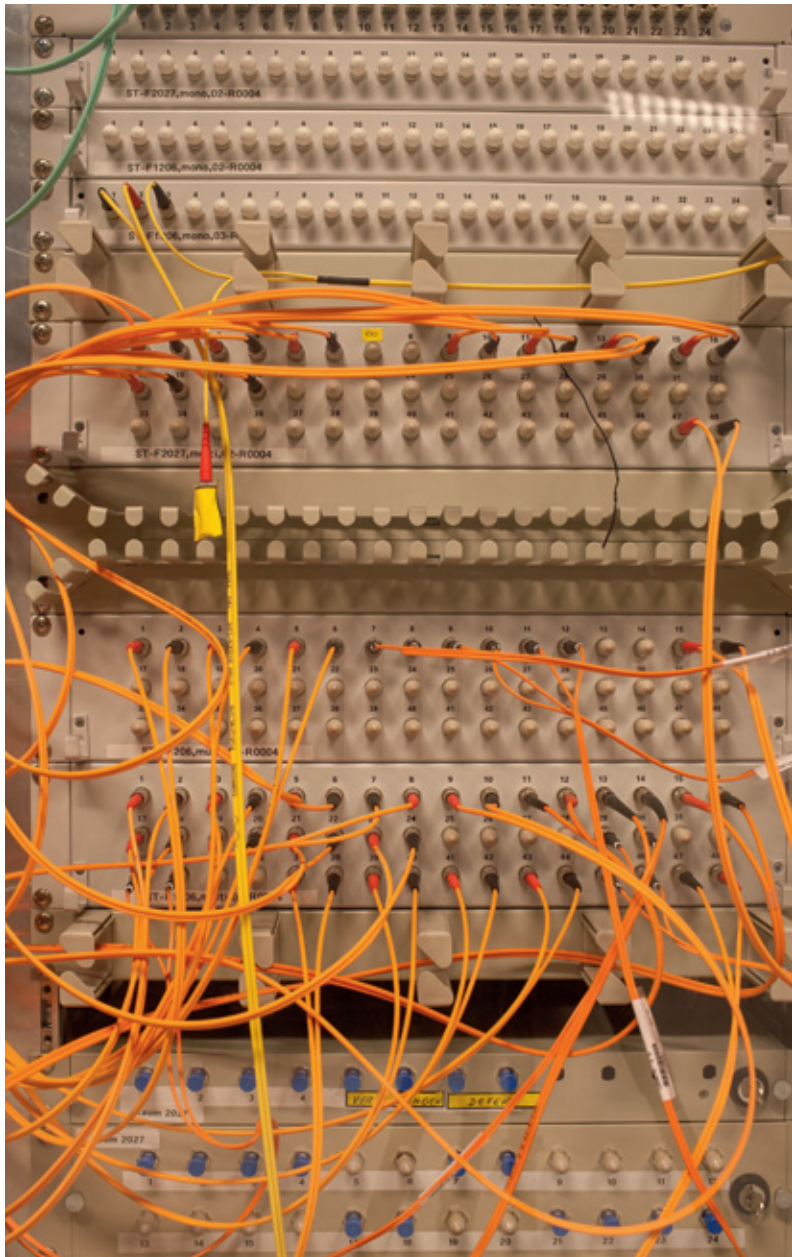
Dies sind einerseits die Randgebiete des Archivierens, nicht von ihrer Bedeutung, sondern von ihrem Stand her. Dies trifft insbesondere auf die Computerspiele zu. Probleme wie die soziale Interaktion von Spielern können als praktisch nicht archivierbar gelten.

Auch zum Film als Objekt der Archivierung wurden kritische Anmerkungen gemacht. Er steht nicht im Fokus des Interesses, obwohl – oder gerade weil – die technischen Probleme sehr groß sind. So existiert keine Pflicht zur Rohdatenabgabe, obwohl dies eine technisch sinnvolle Voraussetzung zur Archivierung wäre. Auch Formatfragen, Fragen des Farbraumes etc. sind mit zahlreichen Problemen behaftet, die andere Medien nicht in diesem Umfang haben.

Wichtig war uns auch ein Blick in die Zukunft der digitalen Medien. Vernetzungskonzepte wie Linked Data stehen vor der Tür. Damit verbunden ist auch ein Wechsel weg vom Fokus auf Objekte und hin zu deren Kontextualisierung. Allerdings sind auch diese Verfahren nicht resistent gegen das Altern.

In der Diskussion um Linked Data wurde deutlich, wie sehr diese Konzepte letztlich auf Sprache basieren und sich damit auch in Abhängigkeit von Sprache befinden. Da auch Sprache altert, sind hier Grenzen für eine langfristige kulturelle Interpretierbarkeit gesetzt.





Lichtwellenleiter, Konrad-Zuse-Zentrum für Informationstechnik Berlin

# Problemlage



# Aktuelle Herausforderungen der digitalen Langzeitarchivierung

TOBIAS BEINERT UND ARMIN STRAUBE

## *Hintergrund*

Der vorliegende Artikel soll die im Rahmen der 8. Initiative „Nachhaltigkeit in der Digitalen Welt“ des Internet und Gesellschaft Collaboratory e. V. von den Teilnehmern gemeinsam erarbeiteten Forderungen des Berliner Appells aus Perspektive der Gedächtnisorganisationen sowie des deutschen Kompetenznetzwerks nestor<sup>1</sup> in allgemein verständlicher Form erläutern und so einen Beitrag zur Verbesserung der Rahmenbedingungen für die digitale Langzeitarchivierung leisten.<sup>2</sup> Aus Sicht der Gedächtnisinstitutionen knüpft der Berliner Appell inhaltlich in vielen Punkten an das bereits 2006 im Namen von nestor veröffentlichte „Memorandum zur Verfügbarkeit digitaler Informationen in Deutschland“<sup>3</sup> an und macht damit bereits auf den ersten Blick deutlich, dass im Bereich der digitalen Langzeitarchivierung nach wie vor massiver Klärungs-, Abstimmungs- und Handlungsbedarf besteht.

Zentral für das Verständnis der Aufgabe – und damit auch die derzeitigen Herausforderungen – der digitalen Langzeitarchivierung aus Sicht der Gedächtnisinstitutionen sind zunächst insbesondere zwei Aspekte, die in der nach wie vor aktuellen Definition von Hans Liegmann und Ute Schwens am besten zum Ausdruck kommen:

1 Vgl. [www.langzeitarchivierung.de](http://www.langzeitarchivierung.de) (Letzter Aufruf: 26.06.2013).

2 Der „Berliner Appell“ der Teilnehmer des CoLab findet sich im vorliegenden Sammelband.

3 nestor: Memorandum zur Langzeitverfügbarkeit digitaler Informationen in Deutschland. 2006. <http://files.dnb.de/nestor/memorandum/memo2006.pdf> (Letzter Aufruf: 26.06.2013).

„Unter Langzeitarchivierung versteht man die Erfassung, die langfristige Aufbewahrung und die Erhaltung der dauerhaften Verfügbarkeit von Informationen. [...] ‚Langzeit‘ bedeutet für die Bestandserhaltung digitaler Ressourcen nicht die Abgabe einer Garantieerklärung über fünf oder fünfzig Jahre, sondern die verantwortliche Entwicklung von Strategien, die den beständigen, vom Informationsmarkt verursachten Wandel bewältigen können.“<sup>4</sup>

Im Folgenden sollen die einzelnen Problemfelder näher beleuchtet werden, und es soll dabei sowohl auf Erfolge als auch auf weiter bestehende Herausforderungen eingegangen werden.

### *Auswahl und Nutzung*

Der Einsatz der notwendigen Ressourcen für die Erhaltung digitaler Informationen kann nur durch den Hinweis auf die aktuelle oder die potentielle zukünftige Nutzung gerechtfertigt werden. Ist dies bei einem aktuellen Nutzungsszenario normalerweise kein größeres Problem, stellt sich bei möglicher Nutzung in der Zukunft eine alte Fragestellung der Gedächtnisorganisationen mit neuer Schärfe: Wie kann heute abgeschätzt werden, was morgen von Interesse ist?

Archive bemühen sich schon seit langem, aus der Flut zeitgenössischer Dokumente eine sinnvolle Auswahl zu treffen, um ein aussagekräftiges Bild unserer Gesellschaft in die Zukunft zu retten.<sup>5</sup> Und auch die Mitarbeiterinnen und Mitarbeiter in Bibliotheken und Museen nehmen beim Aufbau ihrer Sammlungen Auswahlentscheidungen vor. Die Bemühungen der Gedächtnisorganisationen um eine aktive Überlieferungsbildung hatten in der analogen Welt allerdings noch die Möglichkeit, Bestände nachträglich zu ergänzen, da noch nach Jahrzehnten auf Dachböden und in Nachlässen bedeutender Persönlichkeiten interessante historische Zeugnisse aufzufinden waren oder Buchbestände durch antiquarische Käufe erweitert werden konnten. In der digitalen Welt muss die

4 Liegmann, Hans/Schwens, Ute: Langzeitarchivierung digitaler Ressourcen. In: Kühlen, Rainer/Seeger, Thomas/Strauch, Dietmar (Hrsg.): Grundlagen der praktischen Information und Dokumentation. 5., völlig neu gefasste Auflage. München 2004, S. 567. urn:nbn:de:0008-2005110800. [Hervorh. d. Verf.].

5 Einen Überblick bietet Cook, Terry: What is Past is Prologue. A History of Archival Ideas Since 1898 and the Future Paradigm Shift. In: *Archivaria* 43 (1997), S. 17-63.

Sicherung sehr viel schneller und zeitnäher erfolgen. Die Natur digitaler Objekte macht es unmöglich, dass Informationen völlig zufällig und ohne aktives Zutun über lange Zeiträume erhalten bleiben.<sup>6</sup>

Auch wenn die Gedächtnisorganisationen im Rahmen ihrer spezifischen Aufträge und ihrer Möglichkeiten diese Auswahl auch in Bezug auf digitale Informationen treffen, erfordert der schnelle Wandel der Formen und die enorme Steigerung der Menge digitaler Informationen eine übergreifende gesellschaftliche Auseinandersetzung mit der Frage, was unser digitales kulturelles, wissenschaftliches und administratives Erbe ausmacht und wie es bewahrt werden kann. Clifford Lynch fasst treffend zusammen:

„One of the things I talk about nowadays is trying to understand the shape of the overall cultural record and how that shape is changing and where we are succeeding and where we are failing at coming up – as a society – with preservation strategies for deciding what we need to keep and who’s going to keep it and how it’s going to get kept.“<sup>7</sup>

Durch automatisierte Erfassung und Verarbeitung großer Datenbestände erlaubt die Technik grundsätzlich das Anlegen viel größerer Archive, als dies in der Papierwelt möglich und finanzierbar war. Dies ermöglicht zum einen den Umgang mit größeren Datenmengen, zum anderen erhöht es die Chancen, innerhalb solcher „Datenberge“ den zufälligen Dachbodenfund der Zukunft zu machen. Dieser Big-Data-Ansatz kann mit dem Slogan „Collect now, ask later why“ umschrieben werden. Informationen, die nicht zeitnah nach ihrer Entstehung gesichert werden, sind unwiederbringlich verloren, auch wenn man später feststellt, dass ihre Erhaltung sinnvoll gewesen wäre.

Die trotzdem weiterhin nötige Abwägung zwischen Aufwand und potentielltem Nutzen, die der Auswahl zu erhaltender Bestände bzw. Daten zugrundeliegt, folgt dabei neuen Rahmenbedingungen. Digitale Archivierung ist nicht per se billiger als analoge, es ist aber erheblich einfacher, Größeneffekte (economies of scale) zu realisieren und größere Datenmengen zu erhalten. Aber auch

6 Vgl. hierzu den Abschnitt dieses Artikels zu den technischen Herausforderungen der Langzeitarchivierung.

7 Ashenfelder, Mike: Digital Preservation Pioneer: Clifford Lynch. In: The Signal. Digital Preservation. 2013. <http://blogs.loc.gov/digitalpreservation/2013/05/digital-preservation-pioneer-clifford-lynch/> (Letzter Aufruf: 26.06.2013).

wenn die Kosten für zusätzliche Giga- oder Terrabyte an reinem Speicher in einem skalierbaren Langzeitarchivierungssystem gering ausfallen und die Erweiterung der Menge der archivierten Daten meist ein kleineres Problem darstellen als die Unterbringung zusätzlicher Regalkilometer in klassischen Magazinen, entstehen schon bei der Übernahme von Daten oftmals bereits hohe Kosten und erheblicher Personalbedarf. Die Vorbereitung der Daten für spätere, technisch unterstützte Erhaltungsmaßnahmen und die Gewinnung möglichst vielfältiger und aussagekräftiger Metadaten als Anhaltspunkte für künftige Suchanfragen können sehr aufwändig sein.<sup>8</sup> Ebenso sind das laufende Datenmanagement, die Datensicherung sowie die Schaffung neuer Zugriffsmöglichkeiten Kostenfaktoren, die in der digitalen Welt in der Regel in deutlich höherem Maße zum Tragen kommen, als dies bei der klassischen Archivierung der Fall ist.

Illustriert werden können die Probleme der Auswahl an der Debatte um die Archivierung von Twiternachrichten. Natürlich handelt es sich dabei um sehr viel kommunikatives Grundrauschen, aber schon heute interessieren sich Wissenschaftler für die Rezeption bestimmter Ereignisse oder den Verlauf von Debatten auf Twitter. Eine sinnvolle Unterscheidung zwischen beidem kann aber kaum getroffen werden, und die Bewertungen werden sich in Zukunft verschieben. Die Library of Congress hat sich entschieden, den Gesamtbestand der Twiternachrichten zu übernehmen, und steht jetzt in der Praxis vielen Fragen und Problemen gegenüber.<sup>9</sup>

Eine wesentliche Rolle bei der Auswahl erhaltenswerter digitaler Ressourcen können zusätzlich auch private Initiativen spielen. Archive-It<sup>10</sup> und WebCite<sup>11</sup> etwa sind zwei Angebote im Bereich der Webarchivierung, bei denen Nutzer relativ einfach eine Auswahl von Webressourcen archivieren können. Auch Services wie beispielsweise YouTube, Flickr oder Instagram entwickeln sich derzeit faktisch zu einer Form von „Crowd-Archiven“, in denen Millionen Nutzer

8 Vgl. hierzu auch Kapitel 4 in diesem Band.

9 Vgl. Library of Congress: Update on the Twitter Archive At the Library of Congress. 2013. [www.loc.gov/today/pr/2013/files/twitter\\_report\\_2013jan.pdf](http://www.loc.gov/today/pr/2013/files/twitter_report_2013jan.pdf) (Letzter Aufruf: 29.06.2013). In Deutschland gibt es bislang erst wenige Ansätze, auch die Inhalte sozialer Netzwerke zu erhalten. Mit dem Thema setzen sich auch Michael Hollmann und Kathrin Passig in ihren Beiträgen in diesem Band auseinander.

10 Vgl. [www.archive-it.org](http://www.archive-it.org) (Letzter Aufruf: 29.06.2013).

11 Vgl. [www.webcitation.org](http://www.webcitation.org) (Letzter Aufruf: 29.06.2013).

vielfältige kulturelle Äußerungen zugänglich machen und speichern. Allerdings stellt sich hier die Frage nach der Nachhaltigkeit dieser Angebote angesichts der im Hintergrund stehenden kommerziellen Geschäftsmodelle. Dennoch gilt, wenn eine große Anzahl an Akteuren zur Auswahl der zu erhaltenden Quellen beiträgt, kann dies helfen, ein möglichst authentisches und facettenreiches Bild unserer Zeit in die Zukunft retten.

### *Akteure und Strukturen*

In logischer Fortführung ihrer bisherigen Aufgaben erweitern die klassischen Gedächtnisorganisationen, Bibliotheken, Archive und Museen, schrittweise ihre Aktivitäten auf digitale Ressourcen. Bibliotheken erwerben beispielsweise E-Books und schaffen Zugänge zu Zeitschriften, auch wenn diese nur noch digital erscheinen. Archive übernehmen elektronische Akten und setzen sich etwa mit E-Mail-Archivierung und digitalen Fachverfahren der Verwaltungen auseinander. Auch Museen sehen sich mit digitalen Objekten wie Film- und Bildmaterial bis hin zu komplexen 3-D-Modellen konfrontiert. Schon die reine Fortführung bereits bestehender Aufgaben im digitalen Raum stößt oftmals auf technische Probleme, stellt die vorhandene Expertise der Mitarbeiter auf die Probe, muss mit limitierten Ressourcen umgehen und führt auch zu rechtlichen Komplikationen. Viele vor allem kleinere Institutionen sind mit den Aufgaben überfordert.

Die Übernahme von Aufgaben in gänzlich neu entstehenden Feldern konnte deswegen bisher nur von vergleichsweise wenigen Gedächtnisorganisationen – und somit keinesfalls flächendeckend – angegangen werden. Vielfach werden auf Basis von Projekt- oder Fördermitteln erste Lösungsansätze entwickelt und Infrastrukturen aufgebaut. Die Überführung auch sehr erfolgreicher Projekte in den Dauerbetrieb erweist sich aber oftmals als große Hürde. Die Zuwendungsgeber sind vorsichtig, was die Zuweisung neuer Aufgaben an die Gedächtnisorganisationen angeht, nicht zuletzt wegen des sich daraus ableitenden Finanzierungsbedarfs. Probleme gibt es aber nicht nur in neuen Feldern wie der Webarchivierung oder bei den Forschungsdaten, wo der Aufbau von Fähigkeiten und Kapazitäten bei den Gedächtnisinstitutionen bislang noch weit hinter dem Bedarf der wissenschaftlichen Nutzer zurückbleibt. Auch für die schon etwas länger bestehende Herausforderung der Video- und Filmarchivierung gibt es in Deutschland keine Institution, die sich federführend und systematisch um



die vollständige Archivierung bewegter Bilder in analoger oder digitaler Form kümmert.<sup>12</sup> Um Verbesserungen in diesem Bereich bemüht sind das Netzwerk Mediatheken<sup>13</sup> und die nestor AG Media, deren Mitglieder aus unterschiedlichen Kontexten heraus mit audiovisuellen Medien befasst sind.

Insgesamt gesehen, ergibt sich somit bei der Festlegung von Verantwortlichkeiten für die digitale Langzeitarchivierung spartenübergreifend derzeit noch erheblicher Regelungsbedarf. Partiiell werden zwar einzelnen Institutionen bereits Zuständigkeiten durch auf digitale Veröffentlichungen ausgeweitete Pflichtexemplarregelungen, vertragliche Verpflichtungen oder archivgesetzliche Regelungen übertragen, es fehlt hier aber eine übergreifende und verbindlich festgelegte Gesamtkonzeption für eine flächendeckende Aufgabenwahrnehmung.<sup>14</sup>

Während Sammlungsprofile und Zuständigkeiten für Gedächtnisorganisationen schon immer ein wichtiger Punkt waren, wächst die Bedeutung der Vernetzung aller relevanten Akteure (der Stakeholder) im Bereich der digitalen Langzeitarchivierung mit der zunehmenden Komplexität der Aufgaben schnell an. Keine Institution ist heute mehr in der Lage, alle Bereiche der digitalen Langzeitarchivierung konzeptuell und technisch eigenständig zu beherrschen. An Aufgaben wie der Entwicklung von Software und Registries für die Dateiformaterkennung, Standardisierung oder Zertifizierung sind immer eine Vielzahl von Institutionen und Personen beteiligt, die sich damit mit ihren jeweiligen Stärken in die Durchführung und gemeinsame Weiterentwicklung der digitalen Langzeitarchivierung einbringen.

12 Vgl. Kinematheksverbund der Bundesrepublik Deutschland: Studie zu Stand und Aufgaben der Filmarchivierung und zur Verbreitung des nationalen Filmerbes in Deutschland. 2005. [www.bundesarchiv.de/imperia/md/content/abteilungen/abtfa/kinematheksverbund\\_studie\\_zum\\_stand\\_der\\_filmarchivierung.pdf](http://www.bundesarchiv.de/imperia/md/content/abteilungen/abtfa/kinematheksverbund_studie_zum_stand_der_filmarchivierung.pdf) (Letzter Aufruf: 29.06.2013). Mit dem Bundesarchiv, dem Deutschen Filminstitut und der Deutschen Kinemathek existieren zwar einschlägige Organisationen, die aber auch im Verbund nicht über die notwendigen Ressourcen und rechtlichen Kompetenzen verfügen. Zu den technischen Herausforderungen in diesem Bereich sei auf den Beitrag von Jan Fröhlich in diesem Band verwiesen.

13 Vgl. [www.netzwerk-mediatheken.de/](http://www.netzwerk-mediatheken.de/) (Letzter Aufruf: 29.06.2013).

14 Vgl. Schumann, Natascha: Verantwortlichkeiten. In: Neuroth, Heike/Oßwald, Achim/Scheffel, Regine/Strathmann, Stefan/Huth, Karsten (Hrsg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Göttingen 2010, Kap. 3:10ff. urn:nbn:de:0008-2010030552.

Aus Sicht der einzelnen Gedächtnisorganisation gibt es mehrere Ebenen der Zusammenarbeit. Fundamental sind der Austausch und Wissenstransfer zwischen den Gedächtnisorganisationen sowie Absprachen zur Arbeitsteilung. Angesichts der technischen Herausforderungen ist auch die gemeinsame Nutzung von Infrastrukturen sinnvoll. In Deutschland gibt es mittlerweile eine Reihe von sich größtenteils im Aufbau befindlichen Angeboten zur kooperativen Nutzung von Soft- und Hardwareressourcen.<sup>15</sup> Damit können in absehbarer Zeit auch Institutionen kleinerer und mittlerer Größe die von größeren Institutionen betriebenen Langzeitarchivierungsinfrastrukturen für die dauerhafte Sicherung von eigenen Daten einsetzen.

Erste Formen einer neuen Zusammenarbeit zeichnen sich beispielsweise auch im Rahmen der Webarchivierung ab. So bietet das Bibliotheksservice-Zentrum Baden-Württemberg mit SWBcontent<sup>16</sup> einen Service zur Sammlung, Erschließung und Archivierung von Webseiten für Gedächtniseinrichtungen an, ein ähnliches Angebot wird derzeit an der Bayerischen Staatsbibliothek<sup>17</sup> entwickelt. Nachdem im Jahr 2006 der Auftrag der Deutschen Nationalbibliothek auch auf die Archivierung von „unkörperlichen Medienwerken“ bzw. „Netzpublikationen“ ausgedehnt wurde, bemüht sich die Nationalbibliothek schrittweise um eine Umsetzung. Während dies bei den digitalen Entsprechungen gedruckter Medienwerke schon weit vorangeschritten ist, läuft die eigentliche Webarchivierung derzeit erst im Testbetrieb. Mittelfristig wird die Kooperation mit anderen Institutionen bei der Auswahl der zu archivierenden Webseiten angestrebt.<sup>18</sup>

In Deutschland wird die Zusammenarbeit der Institutionen vor allem über nestor, das deutsche Kompetenznetzwerk zur digitalen Langzeitarchivierung, organisiert. Neben dem kollegialen Austausch, zahlreichen fachbezogenen Veröffentlichungen (Praxisratgeber, Handbücher, Expertisen, Stellungnahmen) und Veranstaltungen engagiert sich das Netzwerk in der Qualifizierung der

15 Eine Übersicht über die Dienstleistungs- und Kooperationsangebote in Deutschland bot der nestor-Praktikertag 2013. Die Präsentationen dazu finden sich unter: [www.langzeitarchivierung.de/Subsites/nestor/DE/Veranstaltungen/TermineNestor/praktikertag2013.html](http://www.langzeitarchivierung.de/Subsites/nestor/DE/Veranstaltungen/TermineNestor/praktikertag2013.html) (Letzter Aufruf: 29.06.2013).

16 Vgl. [www.bsz-bw.de/mare/lza/swbcontent.html](http://www.bsz-bw.de/mare/lza/swbcontent.html) (Letzter Aufruf: 01.07.2013).

17 Vgl. [www.babs-muenchen.de/index.html?c=projekte\\_webarchivierung&l=](http://www.babs-muenchen.de/index.html?c=projekte_webarchivierung&l=) (Letzter Aufruf: 01.07.2013).

18 Vgl. [www.dnb.de/DE/Netzpublikationen/netzpublikationen\\_node.html](http://www.dnb.de/DE/Netzpublikationen/netzpublikationen_node.html) (Letzter Aufruf: 08.07.2013).

Mitarbeiter der Gedächtnisorganisationen und bringt sich in Standardisierungsverfahren ein. Seit kurzem wird auch die Zertifizierung vertrauenswürdiger digitaler Langzeitarchive nach DIN 31644 durchgeführt.<sup>19</sup>

All diese Leistungen werden von den nestor-Partnern freiwillig und neben ihren regulären Aufgaben erbracht. Trotz aller Erfolge stößt das Konzept an seine Grenzen. Ein lockeres Netzwerk wie nestor kann letztlich nur sehr geringen Einfluss auf Zuweisung von Zuständigkeiten, auf Ressourcenvergabe und rechtliche Rahmenbedingungen nehmen und eine nationale Strategie und Koordination zur Erhaltung des digitalen Erbes, wie sie schon mehrfach eingefordert wurde,<sup>20</sup> nicht ersetzen. So besteht der bereits 2004 von Ute Schwens und Hans Liegmann formulierte Bedarf, eine übergreifende, belastbare und verbindliche Kooperationsstruktur für digitale Inhalte unter den Gedächtnisinstitutionen schaffen zu müssen, auch weiterhin.<sup>21</sup>

Ein strukturell wichtiger Akteur in der digitalen Langzeitarchivierung ist die Deutsche Forschungsgemeinschaft (DFG). Die Projektförderung hilft beim Aufbau von Expertisen an den Institutionen und bei der Erprobung neuer Ansätze. Doch können hier nur Impulse gegeben werden, die in den Aufbau eines abgestimmten, aber flexiblen nationalen Gesamtsystems der digitalen Langzeitarchivierung münden müssen.

Abschließend sei auch hier auf die zunehmende Bedeutung von Akteuren jenseits der klassischen Gedächtnisorganisationen hingewiesen. Angesichts der langsamen Entwicklungen bei staatlichen Institutionen ist die Rolle privater Initiativen, wie etwa dem Internet Archive,<sup>22</sup> nicht hoch genug einzuschätzen. Von grundlegender Bedeutung ist auch die Erkenntnis, dass digitale Langzeitarchivierung nur ökonomisch durchgeführt werden kann, wenn die Schöpfer

19 Vgl. [www.langzeitarchivierung.de/Subsites/nestor/DE/nestor-Siegel/siegel\\_node.html](http://www.langzeitarchivierung.de/Subsites/nestor/DE/nestor-Siegel/siegel_node.html) (Letzter Aufruf: 29.06.2013).

20 Genannt seien hier die Empfehlungen des Wissenschaftsrates zur Weiterentwicklung der Informationsinfrastruktur: [www.wissenschaftsrat.de/download/archiv/2359-12.pdf](http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf) (Letzter Aufruf: 29.06.2013). Ansonsten sei auf den Beitrag von Stefan Wolf in diesem Band verwiesen.

21 Liegmann, Hans/Schwens, Ute: Langzeitarchivierung digitaler Ressourcen. In: Kühlen, Rainer/Seeger, Thomas/Strauch, Dietmar (Hrsg.): Grundlagen der praktischen Information und Dokumentation. 5., völlig neu gefasste Auflage. München 2004, S. 569. urn:nbn:de:0008-2005110800.

22 Vgl. [www.archive.org](http://www.archive.org) (Letzter Aufruf: 29.06.2013).

digitaler Informationen in die Bemühungen zum Erhalt der Informationen einbezogen werden.

*Technische Herausforderungen: Migration, Emulation,  
Datenmanagement*

Auf der technischen Ebene der digitalen Langzeitarchivierung sehen sich alle Akteure vor zwei zentrale Herausforderungen gestellt. Aufgrund der geringen Halbwertszeit von derzeit eingesetzten Datenträgern (Festplattensysteme, Bandspeichersysteme) müssen diese regelmäßig durch aktuelle Speichermedien ersetzt werden, um Datenveränderungen bzw. -verlusten vorzubeugen. Im Bereich Hardware geht es zunächst um eine sichere Speicherung von Bits und Bytes, also den dauerhaften physischen Erhalt von digitalen Objekten (Bitstream Preservation).<sup>23</sup> Zweitens sind die Lesbarkeit und insbesondere die richtige Interpretation dieses Bitstreams durch die rasante Fortentwicklung von Dateiformaten und Softwareanwendungen gefährdet, digitale Objekte in veralteten oder obsoleten Dateiformaten können von heute standardmäßig eingesetzten Anwendungen nicht mehr bzw. nicht mehr vollständig dargestellt werden. Hier bieten sich als Lösungsansätze die Formatmigration, d. h. die Überführung von Daten in einem Datenformat in ein aktuelles, standardisiertes und idealerweise offengelegtes Format, sowie die Emulation, d. h. die Nachbildung älterer Systemumgebungen auf aktuellen Geräten, an.<sup>24</sup>

Um den physischen Erhalt digitaler Objekte sicherzustellen, machen sich mittlerweile zahlreiche größere Gedächtnisinstitutionen wie beispielsweise die Deutsche Nationalbibliothek, die Bayerische Staatsbibliothek oder die Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden das Know-how professioneller Rechenzentren zunutze. Angesichts der in diesen Einrichtungen

23 Zu den Grundlagen der Bitstream Preservation vgl. Ullrich, Dagmar: Bitstream Preservation.

In: Neuroth, Heike/Oßwald, Achim/Scheffel, Regine/Strathmann, Stefan/Huth, Karsten (Hrsg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Göttingen 2010, Kap. 8:3ff. urn:nbn:de:0008-20100305123.

24 Ausführlich hierzu: Funk, Stefan: Migration / Emulation. In: Neuroth, Heike/Oßwald, Achim/Scheffel, Regine/Strathmann, Stefan/Huth, Karsten (Hrsg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Göttingen 2010, Kap. 8:10ff. urn:nbn:de:0008-20100617189 und urn:nbn:de:0008-20100305134.

anfallenden Datenmengen ist die Auslagerung von Datenarchivierung an kompetente Partner eine sinnvolle Lösung.<sup>25</sup> Kleinere Einrichtungen sind dagegen derzeit oftmals noch darauf angewiesen, Speicherstrategien mittels redundanter bzw. verteilter Speicherung auf unterschiedlichen Datenträgern lokal umzusetzen.<sup>26</sup>

Insgesamt gesehen, ist die Bitstream Preservation in deutschen Gedächtnisinstitutionen ein Problem, das erkannt ist und derzeit in der Regel gezielt in Angriff genommen wird und somit zumindest für einen begrenzten Zeitraum als weitgehend gelöst gelten kann. Ob sich diese bestehenden Storage- und Netzwerkinfrastrukturen allerdings auch im Kontext weiterhin rasch anwachsender Datenmengen – Stichwort Big Data – und der damit verbundenen Schaffung neuer Zugriffs- und Nutzungsmöglichkeiten als nachhaltig erweisen können, erscheint heute zumindest fraglich:

„Like almost all engineering problems, bit preservation is fundamentally a question of budgets. Society’s ever-increasing demands for vast amounts of data to be kept for the future are not matched by suitably lavish funds.“<sup>27</sup>

Neben der Bitstream Preservation ist, wie bereits erwähnt, die Erhaltung der Les- und Interpretierbarkeit von Daten die zweite zentrale Herausforderung der Langzeitarchivierung auf technischer Ebene. In der Praxis der Gedächtnisinstitutionen sind allerdings aus den letzten Jahren noch sehr wenig konkrete Anwen-

25 Eine exemplarische Darstellung der Zusammenarbeit zwischen Bibliotheken und Rechenzentren findet sich bei Reiner, Bernd: Sicherung des Weltkulturerbes am Leibniz-Rechenzentrum. In: Akademie Aktuell. Zeitschrift der Bayerischen Akademie der Wissenschaften 3 (2007).

[www.badw-muenchen.de/aktuell/akademie\\_aktuell/2007/heft3/06\\_Reiner\\_LRZ.pdf](http://www.badw-muenchen.de/aktuell/akademie_aktuell/2007/heft3/06_Reiner_LRZ.pdf) (Letzter Aufruf: 29.06.2013), S. 18-22.

26 Vgl. Brantl, Markus/Borghoff, Uwe (Hrsg.): Entwicklung von Geschäfts- und Organisationsmodellen zur Langzeitarchivierung digitaler Objekte aus DFG-geförderten Digitalisierungsprojekten. Studie im Auftrag der Deutschen Forschungsgemeinschaft. München 2009. [www.babs-muenchen.de/content/dokumente/2009\\_04\\_Studie\\_Organisations\\_und-Geschaeftsmodelle.pdf](http://www.babs-muenchen.de/content/dokumente/2009_04_Studie_Organisations_und-Geschaeftsmodelle.pdf) (Letzter Aufruf: 29.06.2013), S. 82.

27 Rosenthal, David: Bit Preservation: A Solved Problem? In: International Journal of Digital Curation 5/1 (2010), S. 134-148, S. 148. Rosenberg weist in diesem Artikel zudem darauf hin, dass bei wachsenden Datenmengen in digitalen Archiven auch Datenfehler in größerem Umfang als bisher auftreten und zu Schwierigkeiten führen werden.

dungsfälle bekannt, in denen eine der oben kurz beschriebenen Erhaltungsstrategien in größerem Rahmen umgesetzt werden musste. Dies lässt sich vor allem dadurch begründen, dass bei der Archivierung von Anfang an auf digitale Masterfiles in langzeitstabilen und weitverbreiteten Dateiformaten gesetzt wurde bzw. vor der Aufnahme ins Archiv eine Umwandlung von seltenen und speziellen in archivierungstaugliche Formate vorgenommen wurde. Dennoch sind die bislang weitgehend experimentellen Forschungsansätze, die sich sowohl für die Migration als auch für die Emulation in den letzten Jahren entwickelt haben, unbedingt fortzuführen und fortlaufend weiterzuentwickeln, um hier den Akteuren für den Fall, dass bestimmte Formate nicht mehr bzw. nicht mehr richtig dargestellt werden können, die entsprechenden Gegenmittel an die Hand geben zu können.

Wenngleich die beiden zentralen Herausforderungen von den meisten Gedächtnisinstitutionen also derzeit recht gut gemeistert werden können, hat die Praxis der letzten Jahre gezeigt, dass der Teufel oftmals im Detail steckt und teilweise sehr spezifische Fragen vielfach schon bei der Aufnahme von digitalen Objekten in ein Langzeitarchiv geklärt werden müssen.<sup>28</sup> Dazu zählen z. B. das richtige Erkennen, Zuweisen und Validieren von Dateiformaten durch die vorhandenen Tools, der Umgang mit in den Objekten verankerten Kopierschutzmaßnahmen oder auch die Loslösung komplexer digitaler Objekte von physischen Datenträgern (z. B. Disketten oder CD-ROMs).<sup>29</sup> Eine genaue technische Analyse der Objekte und die maschinenlesbare Aufzeichnung der so ermittelten Metadaten ermöglichen es, zu einem späteren Zeitpunkt gefährdete Objekte innerhalb eines digitalen Archivs automatisiert zu erkennen und die notwendigen Erhaltungsmaßnahmen (s. o.) durchzuführen. Diese und ähnliche Fragestellungen, die sich im Lebenszyklus eines digitalen Objekts in vielen Fällen nicht nur einmal, sondern oft mehrmals stellen, bedürfen der Bearbeitung durch qualifiziertes Personal. Zudem müssen der sich immer schneller vollziehende technologische Wandel genau

28 Eine gute Darstellung der Komplexität der Aufnahme von Objekten in ein digitales Archiv findet sich bei Altenhöner, Reinhard/Schrimpf, Sabine: Langzeitarchivierung. In: Kuhlen, Rainer/Semar, Wolfgang/Strauch, Dietmar (Hrsg.): Grundlagen der praktischen Information und Dokumentation. 6., völlig neu gefasste Ausgabe. Berlin 2013. [http://files.d-nb.de/nestor/berichte/GdPluD\\_D4.pdf](http://files.d-nb.de/nestor/berichte/GdPluD_D4.pdf) (Letzter Aufruf: 27.06.2013), S. 529-540, S. 535f.

29 Eine exemplarische Liste mit teilweise sehr speziellen technischen Herausforderungen aus der Praxis und möglichen Lösungsansätzen findet sich im Wiki der Open Planets Foundation: <http://wiki.opf-labs.org/display/REQ/Issues> (Letzter Aufruf: 27.06.2013).

beobachtet und die sich dadurch ändernden Rahmenbedingungen und Anforderungen auf der Nutzerseite in die digitale Archivierung einbezogen werden. Neue Daten und neue Medientypen stellen aus Sicht der Archivierung ständig neue Aufgaben: „Each of the vast range of types of digital content that we should be preserving has its own set of problems. There are no one-size-fits-all approaches.“<sup>30</sup>

Unterstützung bei dieser Aufgabe bieten aber zunehmend Systemlösungen für die digitale Langzeitarchivierung und auf spezielle Anforderungen zugeschnittene Software-Tools, die im letzten Jahrzehnt sowohl von öffentlichen Institutionen als auch von kommerziellen Anbietern entwickelt wurden, teilweise auch auf Open-Source-Basis.<sup>31</sup> Auf die Angebote zur kooperativen Nutzung von Langzeitarchivierungssystemen wurde schon im Abschnitt „Akteure und Strukturen“ hingewiesen. Trotz aller Fortschritte wird die technische Seite der digitalen Langzeitarchivierung nie als endgültig gelöst betrachtet werden können, der Betrieb und die Anpassung der Systeme an den sich laufend vollziehenden technologischen Wandel werden eine fortwährende Daueraufgabe für Archive, Bibliotheken, Museen und Forschungseinrichtungen bleiben. Institutionen, die sich der digitalen Langzeitarchivierung verpflichten, müssen eine kontinuierliche Erhaltungsplanung betreiben, die technische Entwicklung laufend beobachten und angemessen auf sie reagieren.

### *Kosten und Finanzierung*

Eine der zentralen Herausforderungen der Nachhaltigkeit in der digitalen Welt ist sicherlich die Schaffung und Etablierung einer breiten und soliden

30 Ashenfelder, Mike: Insights: Interview with David Rosenthal. In: The Signal. Digital Preservation. 2011. <http://blogs.loc.gov/digitalpreservation/2011/06/insights-interview-with-david-rosenthal/> (Letzter Aufruf: 27.06.2013).

31 Eine übersichtliche Einführung zu Systemlösungen und Tools für die digitale Langzeitarchivierung liefert Altenhöfner, Reinhard/Schrimpf, Sabine: Langzeitarchivierung. In: Kuhlen, Rainer/Semar, Wolfgang/Strauch, Dietmar (Hrsg.): Grundlagen der praktischen Information und Dokumentation. 6., völlig neu gefasste Ausgabe. Berlin 2013. [http://files.d-nb.de/nestor/berichte/GdPIuD\\_D4.pdf](http://files.d-nb.de/nestor/berichte/GdPIuD_D4.pdf) (Letzter Aufruf: 27.06.2013), S. 529-540, S. 536ff. Altenhöfner und Schrimpf weisen aber auch darauf hin, dass die Markt- und Entscheidungslage in diesem Bereich aus Sicht der Gedächtnisinstitutionen derzeit durchaus schwierig und unübersichtlich sei, und erwarten eine Konsolidierung des Softwaremarktes erst für einen Zeitraum von fünf bis zehn Jahren.

Finanzierungsgrundlage für das Handeln der Gedächtnisinstitutionen im Feld der Langzeitarchivierung. Viele Einrichtungen sind hier paradoxerweise bislang aufgrund zu gering bemessener bzw. auch noch gänzlich fehlender Etats auf Projekt- und Anschubfinanzierungen zum Aufbau der notwendigen organisatorisch-technischen Infrastrukturen angewiesen oder sehen sich derzeit noch kaum in der Lage, digitale Archivierung in größerem Rahmen zu betreiben. Mit dem Auslaufen von Fördermitteln stehen viele Institutionen vor der oftmals kritischen Schwelle eines Übergangs in einen auch finanziell und personell ausreichend unterfütterten Routinebetrieb.

Bei der Forderung nach einer dauerhaften Finanzierung darf allerdings keinesfalls übersehen werden, dass hier für eine Realisierung zunächst tragfähige Kostenparameter und -modelle erarbeitet werden müssen, die eine seriöse Finanzplanung und eine entsprechende Bewirtschaftung erst ermöglichen. Derzeit arbeiten einer aktuellen Umfrage des APARSEN-Projekts zufolge ca. 75 % der europaweit im Feld der digitalen Langzeitarchivierung tätigen Institutionen noch gänzlich ohne bzw. lediglich mit auf Grundlage von Erfahrungswerten geschätzten Kostenfaktoren.<sup>32</sup> Daher ist die Entwicklung von Kosten- und Geschäftsmodellen für die Langzeitarchivierung schon seit geraumer Zeit im nationalen wie auch im internationalen Umfeld ein Schwerpunkt von Forschungsprojekten, und es konnten einige am Referenzmodell des Open Archival Information System (OAIS) orientierte Ansätze entwickelt werden.<sup>33</sup> Wie die oben stehenden Zahlen belegen, sind die bereits existierenden Kostenmodelle für einen Praxiseinsatz im größeren Umfang allerdings in der Regel noch nicht ausgereift und anwenderfreundlich genug: „Current models are not developed enough for organisations to be confident in their use.“<sup>34</sup> Dass sich diese durchaus ambitionierten Ansätze in der Praxis bis dato nur sehr partiell als erfolgreich erwiesen haben, ist wohl auf mehrere Faktoren zurückzuführen. Erstens ist in vielen Institutionen das konkrete Aufgabenspektrum der digitalen Langzeitarchivierung noch nicht klar definiert, und es kommt daher zu

32 APARSEN: Report on Cost Parameters for Digital Repositories. 2013.

[www.alliancepermanentaccess.org/wp-content/uploads/downloads/2013/03/APARSEN-REP-D32\\_1-01-1\\_0.pdf](http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2013/03/APARSEN-REP-D32_1-01-1_0.pdf) (Letzter Aufruf: 29.06.2013), S. 32.

33 Eine Übersicht und Analyse über die bislang auf internationaler Ebene im Bereich der digitalen Langzeitarchivierung entwickelten Kostenmodellen findet sich ebd., S. 15ff.

34 Ebd., S. 35.



organisatorischen wie inhaltlichen Überschneidungen mit anderen Handlungsfeldern (z. B. Informationstechnologie, Retrodigitalisierung, Benutzung), die auch eine Festlegung und Zuweisung von Kostenfaktoren in der Praxis oftmals erschweren bzw. gänzlich verhindern. Zudem unterscheiden sich auch die in der Langzeitarchivierung agierenden Institutionen hinsichtlich ihres Aufbaus und ihres Aufgabenzuschnitts in vielen Fällen so wesentlich, dass die Entwicklung universell nutzbarer Modelle und vergleichbarer Kostenfaktoren nur begrenzt möglich erscheint bzw. immer eine Anpassung an den konkreten Rahmen eines Langzeitarchivierungssystems notwendig ist. Dies führt wiederum zu einem erhöhten Arbeitsaufwand bei der Umsetzung eines Kostenmodells, der in manchen Fällen nicht als effizient zu beurteilen sein wird: „The more complex the cost model is the more costly and cumbersome is it to set up and operate the cost model itself.“<sup>35</sup>

Als problematisch erweist sich zudem, dass die Ergebnisse der wenigen bislang vorliegenden Fallstudien, in denen existierende Modelle konkret zur Kostenermittlung angewandt wurden, aufgrund der Vertraulichkeit finanzbezogener Daten nicht vollumfänglich öffentlich zugänglich gemacht wurden.<sup>36</sup> Die oft und gern plakativ gestellte Frage „Was kostet die digitale Langzeitarchivierung?“ ist daher aus Sicht von nestor zum jetzigen Zeitpunkt nur in wenigen Fällen zu beantworten und bedarf auf jeden Fall einer differenzierten Betrachtungsweise.

Eine erste Antwort im deutschsprachigen Bereich lieferte allerdings 2012 das im DP4lib-Projekt entwickelte Kostenmodell für die Budgetierung und Verrechnung eines Langzeitarchivierungsdienstes.<sup>37</sup> Mit Hilfe dieses Modells, das in Form einer übersichtlichen Tabelle umgesetzt ist, lassen sich anhand mehrerer Faktoren die Kosten für die Inanspruchnahme des von der Deutschen Nationalbibliothek angebotenen Services zur Langzeitarchivierung der Daten von externen Institutionen nachvollziehen. Bei der Vorstellung des Services wurde ein

35 DP4lib: A Cost Model for a Long-Term Preservation Service. 2013.

<http://aparsen.digitalpreservation.eu/pub/Main/CostModels/DP4lib-Cost-By-Service-Cost-Model.docx> (Letzter Aufruf: 26.06.2013), S. 4.

36 APARSEN: Report on Cost Parameters for Digital Repositories. 2013. [www.alliancepermanence.org/wp-content/uploads/downloads/2013/03/APARSEN-REP-D32\\_1-01-1\\_0.pdf](http://www.alliancepermanence.org/wp-content/uploads/downloads/2013/03/APARSEN-REP-D32_1-01-1_0.pdf) (Letzter Aufruf: 29.06.2013), S. 36f.

37 DP4lib: Kostenmodell für einen LZA-Dienst. 2012. [http://dp4lib.langzeitarchivierung.de/downloads/DP4lib-Kostenmodell\\_eines\\_LZA-Dienstes\\_v1.0.pdf](http://dp4lib.langzeitarchivierung.de/downloads/DP4lib-Kostenmodell_eines_LZA-Dienstes_v1.0.pdf) (Letzter Aufruf: 27.06.2013).

Basispreis von 1,49 Euro pro Gigabyte und Jahr genannt.<sup>38</sup> Wenngleich auch in diesem Modell noch einige Lücken zu schließen sind (z. B. weil die Kostenfaktoren für den Bereich Preservation Planning / Preservation Action noch unbekannt sind) und die Kostenermittlung zu sehr großen Teilen auf Schätzungen beruht, sind die Benennung und die Veröffentlichung eines ersten groben Preisrahmens sicherlich ein wichtiger Schritt hin zu einer substanziellen Kostenplanung und damit auch ein erster Grundstein für die Schaffung von dauerhaften Finanzierungsgrundlagen. Zukunftsweisend ist der Langzeitarchivierungsservice darüber hinaus insofern, als er Institutionen kleinerer und mittlerer Größe die kooperative Nutzung einer bestehenden Infrastruktur ermöglicht und ihnen so den kostspieligen Aufbau und Betrieb eigener technischer Lösungen erspart. Wie bereits erwähnt, befinden sich ähnliche Services auch in anderen Institutionen im Aufbau, sodass sich hier mittelfristig eine auf mehrere tragende Säulen verteilte nationale Gesamtinfrastruktur für die Langzeitarchivierung entwickeln könnte.

#### *Aus- und Fortbildung*

Eng verknüpft mit dem Thema einer dauerhaften Finanzierung ist die Frage, inwiefern für die in den Gedächtnisinstitutionen entstehenden neuen Aufgaben der Langzeitarchivierung geeignetes Personal zur Verfügung steht und auch über Projektzeiträume hinaus langfristig beschäftigt werden kann. Dass hier grundsätzlich ein Bedarf für den Aufbau neuer Qualifizierungsangebote besteht, hält bereits das nestor-Memorandum aus dem Jahr 2006 fest:

„Es muss professionelles Personal zum Einsatz kommen. Die Anforderungen und Aufgaben der digitalen Langzeitarchivierung sind als ein Schwerpunkt in die Aus- und Fortbildung einzubeziehen. Gezielte Fortbildungsangebote sollten sowohl themenspezifisch sensibilisierend wie auch konkret qualifizierend angelegt werden.“<sup>39</sup>

38 Schmitt, Karlheinz: Der Langzeitarchivierungsservice der Deutschen Nationalbibliothek.

2013. <http://files.dnb.de/nestor/veranstaltungen/Praktikertag2013/2013-06-dnb-schmitt.pdf> (Letzter Aufruf: 27.06.2013).

39 nestor: Memorandum zur Langzeitverfügbarkeit digitaler Informationen in Deutschland. 2006.

<http://files.dnb.de/nestor/memorandum/memo2006.pdf> (Letzter Aufruf: 27.06.2013).

Wichtige Schritte konnten dabei in den vergangenen Jahren von nestor in Zusammenarbeit mit assoziierten Hochschuleinrichtungen im Rahmen eines „Memorandum of Understanding“<sup>40</sup> gemacht werden: Mit der kleinen Enzyklopädie der digitalen Langzeitarchivierung<sup>41</sup> und einem Handbuch zu Forschungsdaten<sup>42</sup> wurden zwei umfassende Einführungen veröffentlicht. Daneben stehen mit den von Studierenden für Studierende entwickelten E-Learning-Tutorials sowie den jährlich mit wechselnden thematischen Schwerpunkten stattfindenden nestor-Schools weitere Qualifizierungsangebote zur Verfügung, die gerade Einsteigern in die digitale Langzeitarchivierung eine grundlegende Orientierung ermöglichen und das nötige Basiswissen vermitteln.<sup>43</sup> Auch in vielen informati-onswissenschaftlich orientierten Ausbildungsangeboten werden in den letzten Jahren zumindest die wichtigsten Grundlagen der digitalen Langzeitarchivierung vermittelt.<sup>44</sup>

Dennoch besteht aus Sicht von nestor im Feld Ausbildung / Qualifizierung weiterhin akuter Handlungsbedarf, wie auch eine aktuelle Bestandsaufnahme des von der EU geförderten Projektes DigCurV zur Weiterentwicklung der beruflichen Weiterbildung sowie eines Kerncurriculums für die Langzeitarchivierung bestätigt:

40 nestor: Memorandum of Understanding. Kooperative Entwicklung curriculärer Module zur digitalen Langzeitarchivierung. 2007. <http://nestor.sub.uni-goettingen.de/education/mou.pdf> (Letzter Aufruf: 27.06.2013).

41 Neuroth, Heike/Oßwald, Achim/Scheffel, Regine/Strathmann, Stefan/Huth, Karsten (Hrsg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Göttingen 2010. urn:nbn:de:0008-2010071949.

42 Neuroth, Heike/Strathmann, Stefan/Oßwald, Achim/Scheffel, Regine/Klump, Jens/Ludwig, Jens (Hrsg.): Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme. Version 1.0. Göttingen 2012. urn:nbn:de:0008-2012031401.

43 Vgl. zu den nestor-Schools: [www.nestor.sub.uni-goettingen.de/education/index.php](http://www.nestor.sub.uni-goettingen.de/education/index.php) (Letzter Aufruf: 29.06.2013). Als ausgezeichnetes Beispiel aus dem europäischen Raum sei hier auch das dreitägige Digital Preservation Training Programme des University of London Computer Centre genannt: [www.dptp.org/](http://www.dptp.org/) (Letzter Aufruf: 29.06.2013).

44 Vgl. die Übersicht hierzu unter APARSEN: Data Collection on Digital Preservation Courses. 2013. [www.alliancepermanentaccess.org/wp-content/uploads/downloads/2013/03/D42\\_1-Annex%20I.pdf](http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2013/03/D42_1-Annex%20I.pdf) (Letzter Aufruf: 26.06.2013).

„The participants reported a severe lack of professionals with the appropriate skills for digital preservation and curation both among existing staff of their institutions and potential staff on the labour market. They expressed an urgent need for training, particularly with regard to IT skills and technical digital preservation skills. Generic skills, management skills and the ability to train others have also been indicated as areas with a considerable training need.“<sup>45</sup>

Eine genauere Betrachtung der Ergebnisse des Projektes zeigt darüber hinaus, dass europaweit bereits relativ viele Trainings- und Weiterbildungsmöglichkeiten mit einführendem Charakter angeboten werden,<sup>46</sup> es aber im Bereich einer stärker praxis- und anwendungsbezogenen Kompetenzvermittlung für Fortgeschrittene einen großen Mangel gibt.<sup>47</sup> Allerdings erweisen sich die konkrete Bedarfsermittlung und die Konzeption derartiger weiterführender Schulungsangebote angesichts der oftmals sehr heterogenen und teilweise speziellen Anforderungen in den Gedächtnisinstitutionen derzeit als schwierig.

Neben einem Ausbau der Möglichkeiten der beruflichen Weiterbildung ist auch eine umfassende Ausbildungsmöglichkeit, z. B. in Form eines eigenständigen Studiengangs, die eine unmittelbare Qualifikation von Berufsanfängern für den Arbeitsmarkt erreicht, aus Sicht der Gedächtnisinstitutionen sicherlich wünschenswert. Eine Ausbildung sollte sich dabei in ihrer Ausrichtung nicht auf klassische Archive und Museen beschränken, sondern auch den in der Wirtschaft und Verwaltung wachsenden Personalbedarf im Feld der digitalen Archivierung berücksichtigen. Entscheidend für den Erfolg eines solchen Angebots ist dabei aus unserer Sicht, dass es gelingt, vor allem stärker Know-how und praxisbezogenes Anwendungswissen der Informationstechnologie zu vermitteln und mit den Inhalten der klassischen Ausbildungen aus dem BAM-Sektor zu verknüpfen. Mittelfristig könn-

45 DigCurV: Report on survey sector training needs. 2012. [www.digcur-education.org/eng/content/download/3321/45923/file/D3-1\\_TrainingNeedsSurvey\\_summary.pdf](http://www.digcur-education.org/eng/content/download/3321/45923/file/D3-1_TrainingNeedsSurvey_summary.pdf) (Letzter Aufruf: 29.06.2013), S. 5f.

46 Eine aktuelle Übersicht findet sich unter APARSEN: Data Collection on Digital Preservation Courses. 2013. [www.alliancepermanentaccess.org/wp-content/uploads/downloads/2013/03/D42\\_1-Annex%20I.pdf](http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2013/03/D42_1-Annex%20I.pdf) (Letzter Aufruf: 26.06.2013).

47 Vgl. DigCurV: Report on survey sector training needs. 2012. [www.digcur-education.org/eng/content/download/3321/45923/file/D3-1\\_TrainingNeedsSurvey\\_summary.pdf](http://www.digcur-education.org/eng/content/download/3321/45923/file/D3-1_TrainingNeedsSurvey_summary.pdf) (Letzter Aufruf: 29.06.2013), S. 14ff.

ten so aus dem Feld der digitalen Langzeitarchivierung wichtige Impulse zu einer übergreifenden Neuorientierung der Ausbildungsmöglichkeiten im Bereich der Gedächtnisinstitutionen vor dem Hintergrund des digitalen Wandels ausgehen.

### *Rechtliche Rahmenbedingungen*

Der Übergang vom analogen zum digitalen Zeitalter wirft für Bibliotheken, Archive und Museen gänzlich neue und bis heute teilweise kaum zu beantwortende Fragestellungen auf. Besonders massive Einschränkungen des Handlungsspielraums ergeben sich für Gedächtnisinstitutionen insbesondere durch das geltende Urheberrecht. Bei den für die Aufnahme und Erhaltung von Objekten in einem digitalen Archiv notwendigen Vervielfältigungen und Bearbeitungen handelt es sich stets um urheberrechtlich relevante Maßnahmen, die entweder vom Rechteinhaber zu genehmigen sind oder aber durch eine gesetzliche Schrankenregelung gedeckt sein müssen.<sup>48</sup> Derartige Regelungen existieren bislang im deutschen Urheberrecht nicht, sodass die nestor-Arbeitsgruppe Recht bereits mehrfach öffentlich darauf hingewiesen hat, dass der derzeitige gesetzliche Rahmen die Gedächtnisinstitutionen zu einer digitalen Amnesie bzw. zur Inkaufnahme von Rechtsverletzungen bei der digitalen Langzeitarchivierung zwingt.<sup>49</sup> Lösungsansätze wie die Schaffung entsprechender Schrankenbestimmungen oder eine am amerikanischen Vorbild orientierte Fair-Use-Klausel für Gedächtnisorganisationen werden in Fachkreisen diskutiert, haben aber bisher nicht das Gehör des Gesetzgebers gefunden.

Daneben können auch persönlichkeitsrechtliche Bestimmungen bei der digitalen Archivierung bzw. der Zugänglichmachung von digitalen Archivinhalten

48 Vgl. u. a. Euler, Ellen: Zur Langzeitarchivierung digital aufgezeichneter Werke und ihrer urheberrechtlichen Einordnung und Beurteilung. In: AfP (2008), S. 474-482; Steinhauer, Eric: Pflichtablieferung von Netzpublikationen. Urheberrechtliche Probleme im Zusammenhang mit der Ablieferungspflicht von Netzpublikationen an die Deutsche Nationalbibliothek. In: K&R (2009), S. 161-166. Zu den rechtlichen Fragen kommen oftmals technische Schwierigkeiten in Form von in den digitalen Objekten verankerten Kopierschutzmechanismen, die eine Langzeitarchivierung in der Regel verhindern.

49 Vgl. nestor: Digitale Langzeitarchivierung als Thema für den 3. Korb zum Urheberrechtsgesetz. Urheberrechtliche Probleme der digitalen Langzeitarchivierung. 2011. [http://files.dnb.de/nestor/berichte/nestor-Stellungnahme\\_AG-Recht.pdf](http://files.dnb.de/nestor/berichte/nestor-Stellungnahme_AG-Recht.pdf) (Letzter Aufruf: 27.06.2013).

problematisch sein, falls z. B. durch die Digitalisierung älterer Zeitungsjahrgänge personenbezogene Ereignisse wieder an das Licht der Öffentlichkeit gelangen, die die Betroffenen längst vergessen glaubten.<sup>50</sup>

Eine zukünftige Herausforderung liegt für Gedächtnisinstitutionen sicherlich auch darin, dass sie viele digitale Objekte gar nicht mehr auf klassische Weise erwerben werden können und so in Form von gespeicherten Dateien in eigenen Archivsystemen vorhalten können. Im Bereich der elektronischen Medien werden zum einen zwischen Anbieter und Institution häufig nur begrenzte Lizenzvereinbarungen über Nutzung eines Angebots abgeschlossen, zweitens setzen viele kommerzielle Content-Produzenten verstärkt auf in sich geschlossene Systeme mit nur temporären Zugriffsrechten oder auch auf rein streamingbasierte Angebote. Gedächtnisinstitutionen werden hier künftig noch stärker als bislang gefordert sein, in den mit diesen Produzenten geschlossenen vertraglichen Nutzungsvereinbarungen praxistaugliche Regelungen für eine digitale Archivierung zu verankern.

Der Rechtsrahmen der digitalen Langzeitarchivierung ist somit insgesamt einer der Bereiche, in dem in den letzten Jahren leider die geringsten Fortschritte zu verzeichnen sind und der damit nach wie vor eines der zentralen Hemmnisse für das Vorankommen von Bibliotheken, Archiven und Museen im Themenkomplex Nachhaltigkeit in der digitalen Welt ist.

### *Fazit*

Ein Großteil der deutschen Gedächtnisorganisationen ist mittlerweile aktiv dabei, den eigenen Handlungsrahmen für die Wahrnehmung der gesellschaftlichen Aufgabe der Sicherung von kulturellem, wissenschaftlichem und administrativem Erbe auch in die digitale Welt zu übertragen. Insgesamt gesehen, wurden in den letzten Jahren einige wichtige Schritte gemacht, das Problembewusstsein, die Fähigkeiten und Kapazitäten in den Institutionen befinden sich in ständigem Wachstum. Erste Standards und Vorgehensweisen wurden national wie international erarbeitet und werden mittlerweile schrittweise in die Praxis umgesetzt. Wenngleich sich das Handeln im Einzelnen dabei sehr heterogen

50 Vgl. Kulturelles Gedächtnis. Wie speichern wir das Internet? Interview mit Eric Steinhauer. In: iRIGHTSinfo. Urheberrecht und kreatives Schaffen in der digitalen Welt. 2012. <http://irights.info/kulturelles-gedchtnis-wie-speichern-wir-das-internet> (Letzter Aufruf: 29.06.2013).

und oftmals individuell gestaltet, bilden sich langsam mögliche Ansätze einer auf mehreren verteilten Leistungsknoten beruhenden nationalen Infrastruktur zur digitalen Langzeitarchivierung heraus.

Trotz der erzielten Fortschritte ergeben sich aus Sicht der Gedächtnisorganisationen in Deutschland derzeit nach wie vor drei zentrale Problemkreise, die eine Reaktion und Anpassung auf den schnellen technologischen Wandel und die neuen Formen digitaler Inhalte verzögern, die aber letztlich gesamtgesellschaftlicher Lösungsansätze bedürfen.

1. Digitale Langzeitarchivierung wird vonseiten der Zuwendungsgeber meist noch nicht als gleichberechtigte Kernaufgabe der Gedächtnisorganisationen gesehen, was sich natürlich auch in der Ressourcenausstattung bemerkbar macht.
2. Die derzeitigen rechtlichen Rahmenbedingungen erschweren die Fortführung der Aufgaben der Gedächtnisorganisationen im Digitalen.
3. Es fehlt eine abgestimmte nationale Strategie und Kooperationsstruktur zur digitalen Langzeitarchivierung.

Auch für Deutschland gilt die treffende Feststellung von Clifford Lynch: „We are still in the middle of a major re-calibration as a society about how we deal with our own memory.“<sup>51</sup> Die notwendige Standortbestimmung und die Schaffung neuer Rahmenbedingungen können die Gedächtnisorganisationen nicht allein durchführen. Es bedarf vielmehr einer breiten Debatte über das digitale kulturelle Erbe in Öffentlichkeit und Gesellschaft. Die 8. CoLab-Initiative zur Nachhaltigkeit in der digitalen Welt war daher ein wichtiger Schritt in diese Richtung – weitere müssen folgen!

51 Ashenfelder, Mike: Digital Preservation Pioneer: Clifford Lynch. In: The Signal. Digital Preservation. 2013. <http://blogs.loc.gov/digitalpreservation/2013/05/digital-preservation-pioneer-clifford-lynch/> (Letzter Aufruf: 26.06.2013).

## Informationen für die Zukunft – EINE Infrastruktur?

STEFAN WOLF

Digitale Daten sind grundlegender Bestandteil der wissenschaftlichen und kulturellen Überlieferung. Diese These ist – beinahe wie in Stein gemeißelt – einleitendes Credo jeder Veröffentlichung und Positionsbestimmung im Bereich der digitalen Langzeitarchivierung. Ein kleines Hindernis in dieser Grundaussage bildet das unauffällige Wörtchen „und“, konstituiert es doch in irgendeiner, nicht näher ausgeführten Weise die Unterscheidung zweier Bereiche, die für Prozesse und Verfahren der digitalen Langzeitarchivierung relevant sind. Dieses „und“ stellt die Frage in den Raum, ob unterschiedliche Verfahren, Herangehensweisen und Anforderungen verlangt sind oder ob durch die Verortung in unterschiedenen institutionellen und interinstitutionellen Zusammenhängen je eigene Anforderungen und Herangehensbedingungen entstehen. Gleichzeitig ist aber die Frage möglich, ob nicht, auch wenn die institutionellen Gegebenheiten unterschiedlich sind, sowohl der Kultur- als auch der Wissenschaftsbereich angesichts der gemeinsamen und gleichartigen Herausforderung zu einem Kooperationsverhältnis finden können.

Der vorliegende Artikel betrachtet hauptsächlich die Strategieentwicklung zur digitalen Langzeitarchivierung im Rahmen der wissenschaftlichen Informationsinfrastruktur in Deutschland. Er tut das im Bewusstsein um den laufenden Einfluss und Austausch mit der internationalen Entwicklung; er versucht aber auch zu zeigen, dass das „und“ wissenschaftlicher und kultureller Überlieferung keinen Widerspruch, sondern ein Kooperationsverhältnis zwischen Wissenschaftsbereich und dem Bereich kultureller Traditionsbildung begründen kann. Gleichzeitig stellt der Artikel die Frage in den Raum, welche gemeinsamen Handlungsmöglichkeiten im Verein von wissenschaftlicher,



kultureller, künstlerischer und bildender Informationsüberlieferung entstehen können.

### *Vielfältigkeit der Wissenschaft*

Ein ins Auge springendes Kennzeichen jeder Betrachtung zur digitalen Langzeitarchivierung im Wissenschaftsbereich ist ihre Vielfältigkeit. Materialbezogen ist kaum eine Gattung denkbar, für die nicht Archivierung relevant ist – zu denken ist an alle Literaturformen vom E-Journal über das E-Book bis zum Digitalisat, von der Hochschulschrift im Open Access bis zur lizenzierten oder gekauften Veröffentlichung mit all den unterschiedlichen Varianten von Verfügbarkeit, Hosting und Speicherung. Audiovisuelles, nicht-textuelles Material in allen Formen gehört ebenso dazu wie die Kommunikationsformen, die durch Social Media auch in den Wissenschaftsbereich getragen werden. Wissenschaftliches Arbeiten ist in vielen Bereichen nicht denkbar ohne Metadaten, die aus Fakten-, Literatur- und Nachweisdatenbanken aller Art stammen können. Hingewiesen sei auf die vielfältigen Formen wissenschaftlicher Roh- und Forschungsdaten, die aus Experimenten, Beobachtungen, statistischen und quantifizierenden Erhebungen stammen können – aber auch aus Archiven und Sammlungen, die üblicherweise dem kulturellen Bereich zugeordnet werden.

Wer diese Materialien näher zu kennzeichnen versucht, kann wiederum nur auf Vielfältigkeit hinweisen: Die physische Existenz der Daten reicht von einmalig bis unzählig dupliziert, die Verfügbarkeit von der freien Zugänglichkeit z. B. im Open Access oder der Ubiquität im World Wide Web bis zur privaten, schützenswerten Zone. Als sicher eingeschätzten Standards kann die Formatgestaltung folgen (PDF/A), sie kann aber auch proprietär sein – folgend je eigenen Bedarfen. Der Nachweis kann isoliert sein, er kann aber auch wie z. B. im Bereich der bibliographischen Daten heute üblich beinahe weltweit vernetzt sein. Auf die disziplinspezifischen Gegebenheiten, Gewohnheiten und Notwendigkeiten sei hingewiesen, gleichzeitig treten fachübergreifende Arbeits- und Traditionsformen in den Blick. Kommerziell auswertbare Forschungsergebnisse werden anderen Verwertungsmodellen unterworfen sein als solche, die eher den Leitlinien geisteswissenschaftlicher Arbeit, einem sachspezifischen Interesse oder dem Aspekt des Reputationsgewinns verpflichtet sind. Dem entspricht, dass besonders auch unter dem grundgesetzlich verbrieften Recht der Freiheit der Wissenschaft die Ergebnisse, Daten und Publikationen der Wissenschaft unter

dem Aspekt ihrer Governance zwischen privat und öffentlich liegen. Dass solche Vielfältigkeit für die Modelle, die Herangehensweise und die Infrastruktur der digitalen Langzeitarchivierung komplexitätssteigernde Konsequenzen hat, liegt auf der Hand.

Ein dritter Aspekt neben den Materialien und ihren Kennzeichen sei erwähnt: Die Teilnehmer selbst am wissenschaftlichen Diskurs, die ja gleichzeitig auch Träger und Nutzer einer Infrastruktur der digitalen Langzeitarchivierung sind, stehen in vielfältigen Beziehungen. Der einzelne Wissenschaftler arbeitet in der Regel an einem Ort, findet aber seine Kooperations- und Diskussionspartner in einem weltweiten Netzwerk, für das eine lokale Bindung zunehmend an Relevanz verliert. Die Wissenschaftler sind eingebunden in ihre institutionellen Kreise als Arbeitgeber, Forschungsgemeinschaften und -gesellschaften, Hochschulen und mit fallender institutioneller Bindewirkung Fachbereiche, Fachgesellschaften und -kreise. Einbindung, Austausch und Theoriebildung erfolgen in Kongressen, offenen, informellen oder gar geschlossenen Diskussionsforen, die Relevanz von Lehrer-Schüler-Beziehungen ist in vielen Disziplinen konstitutiv. Angewiesen ist alle Forschung auf Finanzierung: Private und öffentliche Hand, Hochschulen, Stiftungen, Projekt- und Forschungsförderung, Industrie, Stiftungen treten als Geldgeber auf, gleichzeitig jeweils unterschiedliche Bedingungen setzend – wobei in Deutschland immer auch das Verhältnis von Bundes- und Landesfinanzierung in Betracht gezogen werden muss. Wissenschaftliche Arbeit braucht eine vielfältige, den Anforderungen angepasste Infrastruktur – auch hier treten Rechenzentren aller Größenordnungen, Fachinformationszentren, Bibliotheken, Archive, Museen, Hochschulen auf den Plan.

Spätestens mit dem Hinweis auf die Finanzierung entpuppt sich die offene Frage des Titels dieses Artikels, „EINE Infrastruktur?“, als in gewisser Weise rhetorisch: Wer meint, eine monolithische Infrastruktur könne all diese Ansprüche vereinen, irrt. Zu fragen ist vielmehr, wie ein flexibles, gleichzeitig tragfähiges Netz zur digitalen Langzeitarchivierung aufgebaut werden kann, das zwischen all diesen Bedingungsfaktoren die deutlich und gleichzeitig unterschiedlichen Anforderungen erfüllen kann. Es steht also nicht nur die Frage nach einer „harten“ Informationsinfrastruktur im Raum, sondern die „weichen“ Faktoren in Kooperations- und Geschäftsbeziehungen, in der Kompetenzbestimmung, der Innovationsfähigkeit und der tatsächlichen bzw. unterstellten Leistungsfähigkeit, in der Koalitions- und Kompromissfähigkeit sind mitbestimmend. Die Aufgaben- und Rollendefinition geschieht jetzt, was nicht verwundert, da die

Aufgabe digitale Langzeitarchivierung erst in den letzten Jahren oder den letzten wenigen Jahrzehnten entstanden ist.

### *Stationen der Entwicklung*

Einen bemerkenswerten Auftakt setzte der Workshop „Langzeitverfügbarkeit digitaler Dokumente – Erarbeitung eines ersten kooperativen Konzeptes für Deutschland“ vom 29. und 30. Oktober 2002 in der Deutschen Nationalbibliothek in Frankfurt. Teilnehmer kamen aus dem Bereich der wissenschaftlichen und öffentlichen Bibliotheken, der Archive und Museen, der wissenschaftlichen Fachgesellschaften, der Hochschulrechenzentren, der Forschungsförderung und weiterer interessierter und zu involvierender Kreise. Da die Internetpräsenz dieser Veranstaltung nicht mehr erreichbar ist (dieser Hinweis auf die Notwendigkeit der digitalen Langzeitarchivierung bzw. der Webseitenarchivierung sei gestattet), kann die Abschlusserklärung nur aus einer Spiegelung zitiert werden:

„Die Teilnehmer des Workshops sehen großen Handlungsbedarf bei der Sicherung der Langzeitverfügbarkeit digitaler Objekte. Sie sind der Überzeugung, dass die sich dabei ergebenden Aufgaben als Gemeinschaftsaufgabe arbeitsteilig, kooperativ und ökonomisch gelöst werden müssen. Dabei wird nationaler und internationaler Erfahrungsaustausch einfließen und eine Partizipation an laufenden Entwicklungen erfolgen. In konkreten Kooperationsprojekten sollen unterschiedliche Ansätze und Strategien erprobt und auf ihre Effizienz untersucht werden. Die Ausrichtung auf die Bedürfnisse gegenwärtiger und zukünftiger Nutzergruppen steht dabei im Vordergrund. Pluralität der eingesetzten Methoden und Verfahren und Redundanz der Speicherung sind erwünscht. Synergieeffekte sollen erzielt werden.“<sup>1</sup>

Die Abschlusserklärung spricht die großen Themen schon an: Kooperation, Kosten, nationale Gestaltung in internationaler Einbettung, Speicherung in der nötigen Redundanz und grundlegend die Notwendigkeit, in folgenden Stationen eine Strategie zu entwickeln. Dabei steht diese Abschlusserklärung inhalt-

1 nestor: Abschlusserklärung des Workshops vom 29./30.10.2002. 2002.  
[http://deposit.dnb.de/ep/netpub/94/89/24/972248994/\\_data\\_dyna/\\_snap\\_stand\\_2004\\_08\\_16/www.langzeitarchivierung.de/modules28fb.html?op=modload&name=News&file=article&sid=1&mode=thread&order=0&thold=0](http://deposit.dnb.de/ep/netpub/94/89/24/972248994/_data_dyna/_snap_stand_2004_08_16/www.langzeitarchivierung.de/modules28fb.html?op=modload&name=News&file=article&sid=1&mode=thread&order=0&thold=0) (Letzter Aufruf: 15.08.2013).

lich im Einklang mit der etwas später entstandenen „Charter on the Preservation of Digital Heritage“ vom 15. Oktober 2003, die auch die Notwendigkeit des Handelns bei gleichzeitigem Aufbau von Kooperationsstrukturen und Handlungsplänen betont.<sup>2</sup>

*nestor entsteht*

Aus zwei auf diesen initialen Workshop folgenden Projektphasen entsteht im Sommer 2009 nestor, das deutsche Kompetenznetzwerk zur digitalen Langzeitarchivierung:

„In nestor arbeiten Bibliotheken, Archive, Museen sowie führende Experten gemeinsam zum Thema Langzeitarchivierung und Langzeitverfügbarkeit digitaler Quellen. nestor ist ein Kooperationsverbund mit Partnern aus verschiedenen Bereichen, die alle mit dem Thema ‚Digitale Langzeitarchivierung‘ zu tun haben.“<sup>3</sup>

Als Themen der Arbeit nennt nestor Qualifizierung, Vernetzung und Standardisierung – es geht also um die „weichen“ Faktoren, weniger um die „harte“ Infrastruktur. Aber die teilnehmenden Partner kommen auch aus den Bereichen, die die wissenschaftliche Informationsinfrastruktur mittragen: Auf die involvierten Rechenzentren und die kooperierenden großen deutschen Bibliotheken, Fachinformationszentren und Bibliotheksverbände sei hingewiesen, gleichzeitig auch auf die konstitutive Rolle, die die wissenschaftlichen Fachgesellschaften, insbesondere über ihre IuK-Initiative, beim Aufbau von nestor spielten. 2006 veröffentlicht nestor das „Memorandum zur Langzeitverfügbarkeit digitaler Informationen in Deutschland“:<sup>4</sup> Neben der Betonung der Notwendigkeit des Handelns werden nachhaltige Anstrengungen der Politik, der Rechteinhaber, der Hard- und Softwarehersteller und der Gedächtnisinstitutionen gefordert. Eine nationale Langzeitarchivierungspolicy soll die Rahmenbedingungen vorgeben;

2 UNESCO: Charter on the Preservation of Digital Heritage. 2003. [http://portal.unesco.org/en/ev.php-URL\\_ID=17721&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html) (Letzter Aufruf: 15.07.2013).

3 Vgl. [www.langzeitarchivierung.de](http://www.langzeitarchivierung.de) (Letzter Aufruf: 15.07.2013).

4 nestor: Memorandum zur Langzeitverfügbarkeit digitaler Informationen in Deutschland. 2006. <http://files.dnb.de/nestor/memorandum/memo2006.pdf> (Letzter Aufruf: 15.07.2013).

Vorbilder werden u. a. in der UNESCO-Charta gesehen, aber auch die Strategiepapiere der Deutschen Forschungsgemeinschaft zu Archiven in der Informationsgesellschaft (2003), zur wissenschaftlichen Informationsversorgung (2004) und zum Elektronischen Publizieren (2005) werden genannt; eine Kooperations- und Koordinationsstruktur wird für sinnvoll gehalten, digitale Langzeitspeicher sollen in spezialisierten technischen Systemen aufgebaut werden – eine Aussage über Trägerschaft, Verantwortlichkeit und funktionspezifische Verteilung kann das kurze Memorandum aber nicht machen, denn das Memorandum ist sich bewusst, „eine gemeinsame Diskussion aller Beteiligten ist notwendig“.<sup>5</sup>

*Rahmenkonzept für die Fachinformationsinfrastruktur in  
Deutschland*

2009 erscheint das „Rahmenkonzept für die Fachinformationsinfrastruktur in Deutschland“.<sup>6</sup> Die gemeinsame Wissenschaftskonferenz des Bundes und der Länder in Deutschland (GWK) hatte die Leibniz-Gemeinschaft um die Vorlage eines solchen Gesamtkonzepts gebeten, nachdem zuletzt 2006 ein Bericht zur „Neuausrichtung der öffentlich geförderten Informationseinrichtungen“<sup>7</sup> erstellt und 2008 die gemeinsame Schwerpunktinitiative „Digitale Information“ der Wissenschaftsorganisationen<sup>8</sup> gestartet worden war. Die 2009 befasste Arbeitsgruppe antwortet dann mit einem ersten Schritt, der nur die thematischen Schwerpunkte und Kernaufgaben beschreibt, dies soll dann in einem zweiten Schritt zu einem nationalen Gesamtkonzept erweitert werden. Dabei sollen diese strukturellen Überlegungen auf die wichtigen Akteure im nationalen Rahmen ausgedehnt und die nötigen Schnittstellen und Interaktionsfelder definiert werden; genannt werden zum Beispiel als internationale geisteswissenschaftliche

5 Ebd., S. 5.

6 Arbeitsgruppe Fachinformationsinfrastruktur: Rahmenkonzept für die Fachinformationsinfrastruktur in Deutschland. 2009. [www.gwk-bonn.de/fileadmin/Papers/Rahmenkonzept-WGL.pdf](http://www.gwk-bonn.de/fileadmin/Papers/Rahmenkonzept-WGL.pdf) (Letzter Aufruf: 15.07.2013).

7 Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung: Neuausrichtung der öffentlich geförderten Informationseinrichtungen. Abschlussbericht. Heft 138 (2006). [www.blk-info.de/papers/heft138.pdf](http://www.blk-info.de/papers/heft138.pdf) (Letzter Aufruf: 15.07.2013).

8 Vgl. [www.allianzinitiative.de](http://www.allianzinitiative.de) (Letzter Aufruf: 16.07.2013).

Forschungsbibliothek die Bayerische Staatsbibliothek, als Infrastruktureinheit innerhalb einer Forschungsgemeinschaft die Max Planck Digital Library oder der Gemeinsame Bibliotheksverbund als bibliographisches Zentrum. Als Aufgabe der Infrastruktur werden die klassischen Aufgaben im Bereich der Sammlung und Bereitstellung gesehen, aber auch neue Herausforderungen genannt, wie die Unterstützung der notwendigen nahtlosen Vernetzung von Informationen und ihren Publikationen. Gefordert wird von den Einrichtungen der Infrastruktur, dass sie ihr jeweils angestammtes Gebiet erweitern, hin auf die gesamte wissenschaftliche Wertschöpfungskette: Es wird also der gesamte Prozess wissenschaftlichen Arbeitens in all seiner Vielfältigkeit und all seinen Verästelungen in den Blick genommen. Neue Aufgaben, die an die Institutionen der Informationsinfrastrukturen herangetragen werden, werden u. a. erkannt in den Bereichen Hosting, Langzeitarchivierung, Retrodigitalisierung, Forschungsdaten und im Handlungsfeld Standards, Normen und Qualität.<sup>9</sup> Allein die Nennung dieser Bereiche zeigt schon den erheblichen Umfang der Aufgabe – sowohl in organisatorischer Hinsicht, aber auch was den schier unermesslichen Mengenanfall der zu sichernden Daten betrifft.

Hosting wird als eine Infrastruktur zur verlässlichen Sicherung und Dissemination digitaler Objekte beschrieben – reicht also weit über den Bereich z. B. bibliotheksüblicher Publikationen hinaus. Für notwendig wird ein Repositoryum bzw. ein Netz von Repositorien gehalten – angesichts der Komplexität klingt im Hintergrund die Vorstellung einer monolithischen Struktur an. Eine parallele Studie soll die Bedingungen eines nationalen Hostings klären, wobei dazu auch die Frage gehört, über welche Schnittstellen und damit Standards und Normen die Einbindung der Objekte an die nötigen Dienste erfolgt – aus diesem Auftrag erwächst die folgend beschriebene, sogenannte Beagrie-Studie.

Die Langzeitarchivierung für den Erhalt elektronischer Publikationen wird als nicht gelöst betrachtet, der Standardisierungsbedarf wird anerkannt, Migration bzw. Emulation werden als grundlegende Verfahrenskonzepte genannt, organisatorische Konsequenzen werden angedeutet: „Ziel ist die Einbindung von Einrichtungen mit Kernkompetenzen zur Speicherung, Bereitstellung und Erschließung in ein nationales Konzept unter Berücksichtigung der bestehenden

9 Arbeitsgruppe Fachinformationsinfrastruktur: Rahmenkonzept für die Fachinformationsinfrastruktur in Deutschland. 2009. [www.gwk-bonn.de/fileadmin/Papers/Rahmenkonzept-WGL.pdf](http://www.gwk-bonn.de/fileadmin/Papers/Rahmenkonzept-WGL.pdf) (Letzter Aufruf: 15.07.2013), S. 14.

Infrastruktur.“<sup>10</sup> Wenn auch die Verhältnisbestimmung zum Hosting nicht scharf ist und offenbleibt, ob digitale Langzeitarchivierung nicht eher der „doppelte Boden“ ohne eigenen funktionalen Bezug zur direkten Nutzung ist, ist festzuhalten, dass die Punkte der künftigen Diskussion genannt sind: Speicherung, Erschließung und Bereitstellung praktisch aller Datenarten vom sogenannten Kulturgut bis zur wissenschaftlichen Information. Ein Bezug zur im Dezember 2009 von der Bundesregierung beschlossenen Deutschen Digitalen Bibliothek fehlt – und damit ein Bezug zur Verbindung von kultureller und wissenschaftlicher Information.

Wer dieses Papier bewerten will, wird nach seiner gleichzeitigen Relevanz in der Leibniz-Gemeinschaft und in der umgebundenen Informationsinfrastruktur fragen. Gleichzeitig fällt auf, dass 2009 noch keine Institution aus dem Kreis der Leibniz-Gemeinschaft in nestor, dem Kompetenznetzwerk zur digitalen Langzeitarchivierung, Mitglied war, was sich wenig später änderte. Die Verfahrensbeschreibungen zur digitalen Langzeitarchivierung erscheinen wenig konkret, während zugleich schon in etlichen Institutionen erste konkrete Geschäftsgänge (z. B. der Digitalisierungsworkflow an der Bayerischen Staatsbibliothek mit Übergabe der Digitalisate an ein Rechenzentrum zur Archivierung) aufgebaut waren.

### *Beagrie-Studie*

Die für das Rahmenkonzept verantwortliche Arbeitsgruppe Fachinformationsinfrastruktur hatte angeregt, in einer eigenen Studie zu untersuchen, unter welchen Bedingungen und mit welchem Vorgehen die erworbenen digitalen Publikationen dauerhaft über eine entsprechende Infrastruktur in einem koordinierten Vorgehen auf nationaler Ebene verfügbar gehalten werden können. Es geht also im Wesentlichen um den Bereich der lizenzierten Publikationen, um Inhalte aus dem Bereich Open Access, Digitalisierung oder gar Forschungsdaten treten an den Rand. Ziele der Untersuchung waren eine Analyse der existierenden internationalen Ansätze und aktuellen Initiativen, die Berücksichtigung und Priorisierung praktischer Anforderungen, die Erklärung möglicher Anwendungsfälle und eine Empfehlung für eine Strategieumsetzung, die dann vom öffentlichen Auftraggeber zu verfolgen wäre.

10 Ebd., S. 17.

Mit diesem Auftrag entstand die Studie „Dauerhaften Zugriff sicherstellen: Auf dem Weg zu einer nationalen Strategie zu Perpetual Access und Hosting elektronischer Ressourcen in Deutschland“, veröffentlicht im Februar 2010.<sup>11</sup>

Die Analyse in der Studie zeigt: Eine große Lücke und damit Handlungsbedarf bestehen für den Bereich des gesicherten ununterbrochenen Zugangs zu elektronischen Zeitschriften. Zum Zeitpunkt der Studie wird der Zugang meistens über den Verlagsserver ermöglicht gemacht, woran sich etliche offene Fragen anschließen: Was ist z. B., wenn der Verlagsserver ausfällt, der Verlag bankrott geht oder die Zeitschrift nach einer Abbestellung im Eigentum der erwerbenden Institution eigentlich steht, aber aus Lizenzgründen nicht mehr zugänglich ist? Die Studie kommt zum Schluss, dass ein Bedarf an eigenständigen Hostingkapazitäten besteht, und empfiehlt u. a. die Teilnahme an portico, einem US-amerikanischen Non-Profit-Dienst zur elektronischen Archivierung.

Die Studie gelangt zu 30 Empfehlungen im Bereich Hosting und Perpetual Access,<sup>12</sup> worunter genannt seien:

- Nationale Ausrichtung wahren
- Konsensbildung zwischen allen Beteiligten ermöglichen
- Organisations- und Finanzierungsmodell entwickeln
- Referenzarchitekturmodell und entsprechende Austauschformate und Kommunikationsprotokolle entwickeln
- Digitalisierungsrichtlinien der DFG beibehalten und weiterentwickeln
- Dienstleister auswählen und entwickeln, die eine Umsetzung ihrer Perpetual Access-Rechte und lokalen Hosting-Rechte für elektronische Zeitschriften ermöglichen können
- Kriterien, Standards und Zertifizierungsverfahren definieren.

Dem Umfang, der Sorgfalt der Herangehensweise und dem Komplexitätsgrad des untersuchten Gegenstandes kann nur eine beschränkte, thesenartige Bewertung folgen: Gefordert wird eine Beteiligung aller. Angesichts der erkannten

11 Die deutsche Übersetzung des englischen Originals unter Beagrie, Charles: Dauerhaften Zugriff sicherstellen: Auf dem Weg zu einer nationalen Strategie zu Perpetual Access Hosting elektronischer Ressourcen in Deutschland. 2010. [www.allianzinitiative.de/fileadmin/hosting\\_studie\\_d.pdf](http://www.allianzinitiative.de/fileadmin/hosting_studie_d.pdf) (Letzter Aufruf: 17.07.2013).

12 Im Überblick vgl. ebd., S. 15-18.



„weichen“ Faktoren ist das angemessen, aber auch im laufenden Prozess einzu-lösen: Die Studie hat eine durchaus repräsentative Auswahl von Interviewpartnern erreicht, dennoch sind es nicht alle.<sup>13</sup> Fern liegende Angebote (z. T. in den USA) werden analysiert, näher liegende Initiativen (z. B. E-Services in den Universitätsnetzwerken der Schweiz) nicht erwähnt. Die Studie stellt die Topologie der künftigen Archivierungslösungen in den Vordergrund, die Frage der Vertrauenswürdigkeit des Archivs und seiner Abläufe klingt nur am Rande an: Digitale Archive sind nur insofern tragfähig, als die Integrität ihrer Inhalte gesichert und belegt ist – ein bloßes Vertrauen in die referenzierende Institution genügt angesichts der leichten Veränderbarkeit der Inhalte nicht mehr. Letztlich lässt die Studie die Entscheidungen offen, stellt aber die Fragen in den Raum: Wie kommen die Inhalte in die Verfügung der finanzierenden öffentlichen Hand? Ist das gewollt? Wie können die Verfahren und Installationen dauerhaft finanziert werden? In der Gesamtbetrachtung ist aber sicher der Trend weg vom lokalen Hosting zu z. B. portico oder Lockss und weg vom monolithischen System richtig. Eine Zusammenfassung findet sich z. B. auch im Rahmen der Sitzung des Deutschen Bibliotheksverbandes, Sektion IV (wissenschaftliche Bibliotheken) vom 26. und 27. Mai 2010 in Bremen;<sup>14</sup> hier wird festgehalten, dass für das dauerhafte Hosting nur eine öffentlich getragene Infrastruktur in Frage kommt, die keine Monopolstruktur besitzt. Weitere Stellungnahmen treten gleichzeitig dafür ein, den internationalen Kontext einer Lösung stärker zu verfolgen. Da keine Entscheidung zwischen den Möglichkeiten gefunden wurde, wird die Studie in der AG Hosting und Langzeitarchivierung der Kommission Zukunft der Informationsinfrastruktur im Auftrag der gemeinsamen Wissenschaftskonferenz des Bundes und der Länder weiterbearbeitet. Das Ergebnis findet sich im „Gesamtkonzept für die Informationsinfrastruktur in Deutschland“ vom April 2011.<sup>15</sup>

13 Das Bibliotheksservice-Zentrum Baden-Württemberg wurde z. B. nicht im Kreis der Befragten berücksichtigt, obwohl es im Bereich Hosting von E-Publikationen, Digitalisaten und Webseiten zum Zeitpunkt der Befragung über konkrete, teils fast singuläre Erfahrungen verfügte.

14 Deutscher Bibliotheksverband, Sektion IV: Protokoll der Frühjahrstagung am 26. und 27. Mai 2010 in Bremen. 2010. [www.bibliotheksverband.de/fileadmin/user\\_upload/Sektionen/sektion4/Tagungen/2010\\_Protokoll\\_Bremen\\_26.u.27.Mai.pdf](http://www.bibliotheksverband.de/fileadmin/user_upload/Sektionen/sektion4/Tagungen/2010_Protokoll_Bremen_26.u.27.Mai.pdf) (Letzter Aufruf: 17.07.2013).

15 Kommission Zukunft der Informationsinfrastruktur: Gesamtkonzept für die Informationsinfrastruktur in Deutschland. 2011. [www.allianzinitiative.de/fileadmin/user\\_upload/KII\\_Gesamtkonzept.pdf](http://www.allianzinitiative.de/fileadmin/user_upload/KII_Gesamtkonzept.pdf) (Letzter Aufruf: 17.07.2013).

*Gesamtkonzept für die Informationsinfrastruktur in Deutschland*

Hingewiesen sei einleitend nochmals darauf: Die Beagrie-Studie erstreckt sich nur auf elektronische Publikationen – der gesamte weitere Materialbereich, der die Komplexität des Auftrages quasi potenziert, tritt nur am Rande in den Blick. Dies wird nun mit dem Gesamtkonzept nochmals versucht.

Der Erweiterung des Arbeitsauftrages entspricht die Erweiterung des Mitarbeiterkreises: Am Gesamtkonzept sind 135 Mitglieder aus 54 Institutionen beteiligt, erreicht wird auch eine Verzahnung mit der Schwerpunktinitiative „Digitale Information“ der Allianz der Wissenschaftsorganisationen. Die Weiterführung des Arbeitsauftrages durch den Wissenschaftsrat ist von Anfang an mit bedacht. Wieder – wie schon beim Rahmenkonzept – geht es um die Beteiligung der Informationsinfrastruktur an der gesamten Wertschöpfungskette des wissenschaftlichen Diskurses.<sup>16</sup> Die Studie bindet Hosting und Langzeitarchivierung nun fest zusammen, grenzt sie aber auch gleichzeitig voneinander ab:

„Hosting meint in diesem Umfeld das Bereithalten von Inhalten für den unmittelbaren, sofortigen Zugriff. Eng damit verbunden ist das Thema der Langzeitarchivierung: dieser Begriff wird für die Abfolge von Aktivitäten verwendet, die erforderlich sind, um verlässlichen Zugriff auf digitale Bestände so lange wie möglich und unabhängig von Speichermedienausfällen oder (Software-)technischen Veränderungen sicherzustellen.“<sup>17</sup>

Wieder stehen Publikationen im Vordergrund der Betrachtung zu Hosting und Langzeitarchivierung, durch die anderen Handlungsfelder wie nicht-textuelle Materialien, Retrodigitalisierung, Forschungsdaten etc. ist aber klar: Langzeitarchivierung ist eine Querschnittsaufgabe und erstreckt sich auch auf diese Felder. Das Gesamtkonzept fordert deshalb die Schaffung einer nachhaltigen Struktur mit gestaltender und kontrollierender Funktion – im Bereich des Hostings durch die Sicherung des dauerhaften Zugriffs, im Bereich der Langzeitarchivierung durch den Ausbau von nestor.<sup>18</sup> Das ist besonders bemerkenswert, da in der Beagrie-Studie nestor nur in einem Interviewtext am Rande erwähnt wird, während nestor jetzt im Gesamtkonzept eine tragende Rolle zugemessen

16 Graphik vgl. ebd., S. 13.

17 Ebd., S. 27.

18 Vgl. ebd., S. 48.

wird: „Mit nestor besteht im Bereich der Langzeitarchivierung ein Nukleus für eine koordinierende Organisation, die nachhaltig ausgebaut werden sollte.“<sup>19</sup> Als Schwerpunkte für nestor sind benannt Standardisierung, Qualifizierung, Zertifizierung und Vernetzung. Für den Bereich Hosting / Langzeitarchivierung wird Handlungsbedarf zur Sicherung der permanenten Verfügbarkeit erkannt, indem über standardisierte Verfahren und zuverlässige Services die in den Lizenzvereinbarungen geschlossenen Klauseln zum dauerhaften Zugang und zur Eigenverfügung über die Publikationen realisiert werden. Der Nachweis dieser Publikationen ist ebenso dauerhaft und zuverlässig zu gewährleisten, während die Dienste zur Langzeitarchivierung dafür das Fundament bilden. Gleichzeitig müssen die zur Verfügung stehenden Inhalte in die Forschungsumgebungen eingeführt werden sowie die ergänzenden Materialien und Forschungsdaten über persistente Identifier zugreifbar und verfügbar sein. Erkannt wird auch, dass sowohl die Dienste und die Hard- und Software als auch die Organisationsstrukturen erst aufgebaut werden müssen, wobei hier für erste Ergebnisse ein Zeitraum von etwa drei Jahren veranschlagt wird.<sup>20</sup>

Das Gesamtkonzept empfiehlt für alle Handlungsfelder nationale Koordinatoren – für den Bereich des Hostings wird das Fachinformationszentrum Karlsruhe genannt, für die Langzeitarchivierung die Deutsche Nationalbibliothek. Mit dem Gesamtkonzept wurde für 2012 eine Veröffentlichung „Empfehlungen zur Zukunft der Informationsinfrastruktur“ angekündigt.<sup>21</sup> Dieses Papier ist aber noch nicht erschienen. So kann auch – da das Gesamtkonzept noch keine Verwirklichung gefunden hat – auch das Handlungsfeld Hosting / Langzeitarchivierung nur preliminär bewertet werden: Die Erkenntnis, dass eine koordinierende Stelle vorhanden sein muss, setzt sich weiter durch. Zu klären wird für diese Stelle sein, welche Bindewirkung sie mit ihrem Auftrag entfalten kann, in einem Bereich, wo vieles in Entwicklung bzw. zwischen den öffentlichen Akteuren auf freiwilligen Absprachen beruht. Gleichzeitig zeigt aber das laufende und organische Wachstum von nestor, dass hier eine Funktion erkannt wird, in die man sich freiwillig einbringt – über die Sparte der Wissenschaft hinaus aus allen Kultur- und Bildungsbereichen.

19 Ebd., S. B32.

20 Vgl. ebd., S. 34.

21 Gemeinsame Wissenschaftskonferenz: Informationsinfrastruktur. [www.gwk-bonn.de/index.php?id=205](http://www.gwk-bonn.de/index.php?id=205) (Letzter Aufruf: 17.07.2013).

Eine gewisse Fortsetzung fand das Gesamtkonzept in der Ausschreibung der DFG zur Neuausrichtung überregionaler Informationsservices. Hier wurden vier Handlungsfelder benannt, nämlich „Bibliotheksdateninfrastruktur und Lokale Systeme“, „Langzeitverfügbarkeit“, „Electronic Resource Management – das Management digitaler Publikationen“ sowie „Forschungsnahe Informationsinfrastruktur“.<sup>22</sup> Die auf dem Handlungsfeld Langzeitverfügbarkeit eingereichten Projektvorschläge konnten 2012 nicht überzeugen, sodass keine Projektzuteilung auf diesem Gebiet erfolgte. In einer Wiederholung der Ausschreibung am 1. Juli 2013 wurde nun dieses Gebiet nochmals adressiert:

„Die Ausschreibung ‚Langzeitverfügbarkeit im Rahmen der Neuausrichtung überregionaler Informationsservices‘ zielt darauf ab, einen Prozess anzustoßen, in dem überregionale Angebote und Dienstleistungen im Umfeld der Langzeitarchivierung aufgebaut werden. Es soll der Aufbau einer überregionalen, skalierbaren, mandantenfähigen und nach Möglichkeit auch spartenübergreifenden Infrastruktur eingeleitet werden. Diese Struktur soll sämtliche Schritte der digitalen Langzeitarchivierung vom Dateningest bis zur Endnutzerbereitstellung unterstützen.“<sup>23</sup>

Deutlich wird wieder: Der Abschied von einer monolithischen Struktur ist erfolgt; die Ausbildung tragfähiger Kooperations- und Absprachestrukturen ist noch sehr offen, wobei die Komplexität der Aufgabe durchaus benannt ist. Abzuwarten ist, ob und in welcher Form nestor – wie im Gesamtkonzept vorgeschlagen – eine koordinierende und gestaltende Funktion einnimmt.

22 Vgl. Deutsche Forschungsgemeinschaft: DFG bewilligt drei Projekte zur Neuausrichtung überregionaler Informationsservices. 2013. [www.dfg.de/foerderung/info\\_wissenschaft/archiv/2013/info\\_wissenschaft\\_13\\_11/](http://www.dfg.de/foerderung/info_wissenschaft/archiv/2013/info_wissenschaft_13_11/) (Letzter Aufruf: 18.07.2013).

23 Deutsche Forschungsgemeinschaft: Langzeitverfügbarkeit im Rahmen der Neuausrichtung überregionaler Informationsservices. 2013. [www.dfg.de/foerderung/info\\_wissenschaft/info\\_wissenschaft\\_13\\_36/index.html](http://www.dfg.de/foerderung/info_wissenschaft/info_wissenschaft_13_36/index.html) (Letzter Aufruf: 18.07.2013).



# Wissen ohne Zukunft? Der Rechtsrahmen der digitalen Langzeitarchivierung von Netzpublikationen

ERIC W. STEINHAUER

## *Einstieg: Informationsgesellschaft*

Wir leben in einer Wissens- und Informationsgesellschaft. Im Einzelnen sind diese beiden schlagwortartigen Begriffe zwar unscharf und strittig.<sup>1</sup> Einigkeit besteht jedoch darin, dass ein als Wissens- und Informationsgesellschaft bezeichnetes Gemeinwesen durch die Allgegenwart digitaler Kommunikation geprägt ist. Hier kommt dem Internet eine wichtige Rolle zu. Es ist sicher nicht übertrieben, das Internet als den zentralen Wahrnehmungs- und Kommunikationsraum zu verstehen. Dabei liegen die Stärke und Faszination des Internets in der Vernetzung von Inhalten. Wir können dies besonders gut in den sozialen Netzwerken beobachten, in denen Inhalte aus unterschiedlichen Quellen zueinander in Beziehung gesetzt und miteinander geteilt werden. Die Kommunikation bezieht sich auf diese Inhalte und ist auf sie angewiesen.

## *Gedächtnis*

Damit ist so etwas wie ein digitales Gedächtnis angesprochen. Dieses Gedächtnis funktioniert erstaunlich gut. Dank der Allgegenwart des mobilen Internets

1 Vgl. Leidhold, Wolfgang: Wissensgesellschaft. In: Korte, Karl-Rudolf/Weidenfeld, Werner (Hrsg.): Deutschland-TrendBuch. Opladen 2001, S. 441f.; Rösch, Hermann: Informationsgesellschaft. In: Umlauf, Konrad/Gradmann, Stefan (Hrsg.): Lexikon der Bibliotheks- und Informationswissenschaft. Bd. 1. Stuttgart 2011, S. 430f.; Spinner, Helmut F.: Informationsgesellschaft. In: Schäfers, Bernhard/Zapf, Wolfgang (Hrsg.): Handwörterbuch zur Gesellschaft Deutschlands. Opladen 1998, S. 313.

können wir jederzeit und von jedem Ort aus auf einen immensen Fundus an Informationen zugreifen. Der Grad der Informiertheit, den wir dadurch erreichen, ist kaum noch zu übertreffen. Allerdings – und hier beginnen die Probleme – sind das Internet, vor allem aber diese Art der intensiven Kommunikation noch recht jung. Wenn wir hier von Gedächtnis sprechen, so ist damit ein kommunikatives oder ein Funktionsgedächtnis gemeint: Es speichert und erinnert Sachverhalte einer lebendig erlebten Gegenwart.<sup>2</sup>

In den Memory Studies, einem wichtigen Teilgebiet der Kulturwissenschaften,<sup>3</sup> wird dieses Gedächtnis vor allem mit der mündlichen Überlieferung durch Erzählen und Berichten in Verbindung gebracht. Dabei reicht das kommunikative Gedächtnis so weit zurück wie die Erinnerung noch lebender Menschen, also maximal 100 Jahre. Alles, was darüber hinausgeht, ist entweder mythisch oder vergessen oder im *kulturellen* Gedächtnis als einem *Speichergedächtnis* enthalten.<sup>4</sup>

Dieses kulturelle Gedächtnis findet seinen Niederschlag in relativ dauerhaften Zeugnissen der Kultur, hier vor allem in lesbaren Medien, an erster Stelle in Büchern und anderen gedruckten Materialien.

Wir sind auf dieses Speichergedächtnis zwingend angewiesen, wenn wir uns auf Sachverhalte außerhalb des zeitlichen Erfahrungshorizontes lebender Personen beziehen wollen.<sup>5</sup> An dieser Stelle kann die Allgegenwart des Internets in der Wissens- und Informationsgesellschaft mit der Zeit zu einem ernstem Problem werden, wenn immer mehr Inhalte ausschließlich in digitaler Form kommuniziert und gespeichert werden. Digitale Inhalte – kurz Digitalia – haben nämlich die unangenehme Eigenschaft, der sinnlichen Wahrnehmung nicht unmittelbar zugänglich zu sein. Sie müssen durch Software interpretiert und lesbar gemacht werden. Soll dieser Vorgang über lange Zeiträume funktionieren, ist Vorsorge für die Stabilität von Formaten zu treffen oder eine Überführung in neue Formate notwendig. Doch damit nicht genug. Digitalia sind darüber hinaus auch noch

2 Vgl. Levy, Daniel: Das kulturelle Gedächtnis. In: Gudehus, Christiane/Eichenberg, Ariane/Welzer, Harald (Hrsg.): Gedächtnis und Erinnerung. Stuttgart 2010, S. 93f.

3 Dazu Erll, Astrid: Cultural Memory Studies. In: Erll, Astrid/Nünning, Ansgar: A Companion to Cultural Memory Studies. Berlin u. a. 2010, S. 1-15.

4 Vgl. Assmann, Jan: Communicative and Cultural Memory. In: Erll, Astrid/Nünning, Ansgar: A Companion to Cultural Memory Studies. Berlin u. a. 2010, S. 109-118.

5 Vgl. Assmann, Aleida: Archive und Bibliotheken. In: Gudehus, Christiane/Eichenberg, Ariane/Welzer, Harald (Hrsg.): Gedächtnis und Erinnerung. Stuttgart 2010, S. 165.

hochfragil, da sie entweder unkörperlich und damit flüchtig in Netzen kursieren oder auf Datenträgern aufbewahrt werden, die nach dem gegenwärtigen Stand der Technik eine generationenübergreifende Haltbarkeit, die der von Papier entspricht, nicht besitzen.<sup>6</sup> Digitalia müssen daher in regelmäßigen Abständen auf neue Datenträger umgeschrieben werden.

Formatänderungen und Datenträgerwechsel sind technisch komplexe Vorgänge. Sie stehen bei der Frage der digitalen Langzeitarchivierung meist im Vordergrund.<sup>7</sup> Technische Aspekte sind aber nur eine Seite des digitalen Gedächtnisses.

### *Urheberrecht*

Die andere Seite, die wir uns jetzt näher ansehen wollen, ist eine juristische.<sup>8</sup> Soweit nämlich die betroffenen Digitalia persönliche geistige Schöpfungen sind, was meist der Fall sein wird, unterfallen sie als Werke im Sinne von § 2 Urheberrechtsgesetz (UrhG) dem Urheberrecht, und ihre Vervielfältigung hat bestimmte

- 6 Vgl. Serexhe, Bernhard: Neue Medien – kurzes Gedächtnis? Anmerkungen zum Systemwechsel des kulturellen Gedächtnisses. In: Robertson-v.Trotha, Caroline Y./Hauser, Robert (Hrsg.): Neues Erbe. Aspekte, Perspektiven und Konsequenzen der digitalen Überlieferung. Karlsruhe 2011, S. 78f.
- 7 Vgl. Neuroth, Heike/Oßwald, Achim/Scheffel, Regine/Strathmann, Stefan/Jehn, Mathias (Hrsg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Boizenburg 2009; Umlauf, Konrad: Digitale Langzeitarchivierung. In: Umlauf, Konrad/Gradmann, Stefan (Hrsg.): Lexikon der Bibliotheks- und Informationswissenschaft. Bd. 1. Stuttgart 2011, S. 218.
- 8 Siehe dazu auch Euler, Ellen: Digitale Langzeitarchivierung und Distributed Storage in LUKII. Gutachten zu den Rechtsfragen, die sich im Langzeitarchivierungsprojekt LUKII (LOCKSS und kopal Infrastruktur und Interoperabilität) stellen. 2011. [www.iuwis.de/sites/default/files/LUKII\\_Rechtsgutachten.pdf](http://www.iuwis.de/sites/default/files/LUKII_Rechtsgutachten.pdf) (Letzter Aufruf: 18.06.2013); dies.: Web-Harvesting vs. Urheberrecht. Was Bibliotheken und Archive dürfen und was nicht. In: CR (2008), S. 64-68; dies.: Zur Langzeitarchivierung digital aufgezeichneter Werke und ihrer urheberrechtlichen Einordnung und Beurteilung. In: AfP (2008), S. 474-482; Goebel, Jürgen W./Scheller, Jürgen: Digitale Langzeitarchivierung und Recht. Frankfurt/M. 2004; Heckmann, Jörn/Weber, Marc Philipp: Elektronische Netzpublikationen im Lichte des Gesetzes über die Deutsche Nationalbibliothek (DNBG). In: AfP (2008), S. 269-276; Steinhauer, Eric W.: Pflichtablieferung von Netzpublikationen. Urheberrechtliche Probleme im Zusammenhang mit der Ablieferungspflicht von Netzpublikationen an die Deutsche Nationalbibliothek. In: K&R (2009), S. 161-166.



rechtliche Vorgaben zu beachten. Da sowohl das für die digitale Langzeitarchivierung unverzichtbare Umkopieren als auch mögliche Formatänderungen Vervielfältigungen darstellen, liegt die gar nicht zu überschätzende Relevanz des Urheberrechts für diese Vorgänge auf der Hand. Und hier stellt sich die Lage – in grober Skizze – wie folgt dar:

Jede Vervielfältigung eines urheberrechtlich geschützten Werkes greift in das Verwertungsrecht des Urhebers aus § 16 UrhG bzw. entsprechende Nutzungsrechte von Rechteinhabern ein. Wenn der Urheber bzw. die Rechteinhaber die für die Langzeitarchivierung erforderlichen Vervielfältigungen – sei es in Form von Nutzungsrechten, sei es in sonstiger Form – gestatten, ergeben sich keine Probleme. Die Vervielfältigung ist rechtmäßig.

Kommt es aber zu keiner Verständigung über die Zulässigkeit von Vervielfältigungen, können Maßnahmen der Langzeitarchivierung nur auf Grundlage urheberrechtlicher Schrankenbestimmungen vorgenommen werden.

Dabei kann § 44a UrhG als möglicherweise einschlägige Schrankenbestimmung sogleich ausgeschlossen werden, da diese Norm jedenfalls nur vorübergehende und eben keine auf langfristige Bewahrung angelegte Kopien gestattet,<sup>9</sup> mag die Formulierung im Gesetz, dass eine zulässige Vervielfältigung „Teil eines technischen Verfahrens“ ist, auf den ersten Blick für das im Wesentlichen technische Problem der digitalen Langzeitarchivierung auch verlockend einschlägig erscheinen.

#### *Digitale Langzeitarchivierung nach § 53 UrhG*

Bleibt also im Wesentlichen nur der labyrinthische § 53 UrhG übrig, um Vervielfältigungen im Rahmen der digitalen Langzeitarchivierung zu rechtfertigen.

Danach ist eine digitale Privatkopie nach § 53 Abs. 1 gestattet, sofern sie nicht von einer offensichtlich rechtswidrigen Vorlage angefertigt wird. Ebenfalls zulässig sind nach § 53 Abs. 2 Satz 1 Nr. 1 UrhG entsprechende Vervielfältigungen für den eigenen wissenschaftlichen Gebrauch, wobei hier näher zu begründen wäre, warum der eigene wissenschaftliche Gebrauch gerade eine auf mehrere Generationen angelegte und damit die eigene Lebenszeit des Schrankenbegünstigten wohl übersteigende Langzeitarchivierung erfordert, können Tote

9 Vgl. Dreier, Thomas. In: Dreier, Thomas/Schulze, Gernot: UrhG. Urheberrechtsgesetz. Urheberrechtswahrnehmungsgesetz, Kunsturhebergesetz. Kommentar. 4. Auflage. München 2013, § 44a, Rn. 4f.

doch schwerlich einen eigenen wissenschaftlichen Gebrauch für sich reklamieren. Selbst wenn man sich hier auf den Gebrauch einer Institution bezieht, was durchaus noch als „eigener wissenschaftlicher Gebrauch“ gewertet wird,<sup>10</sup> so stellt sich schon die Frage, ob Kopien, die im Vorfeld des Gebrauchs an sich liegen, von der Schranke überhaupt erfasst sind oder ob nicht die Kopie selbst dem unmittelbaren Gebrauch dienen muss. Vor dem Hintergrund der regelmäßig geforderten engen Schrankenauslegung ist diese Frage keineswegs haarspalterisch, sondern leider nur allzu berechtigt.<sup>11</sup>

Interessanter ist da schon die Erlaubnis in § 53 Abs. 2 Satz 1 Nr. 2 UrhG, eine Vervielfältigung zur Aufnahme in ein eigenes Archiv herzustellen, denn der Aufbau eines Archivs ist doch genau das, was wir wollen, wenn wir über digitale Langzeitarchivierung reden. Allerdings liegt hier die Tücke im Detail. So muss die Vorlage für die Archivkopie zunächst ein „eigenes Werkstück“ sein. Über das Problem, inwieweit unkörperliche Medienwerke, die ja den Großteil der Digitalia mittlerweile ausmachen, dem Schrankenbegünstigten „eigen“ und inwieweit diese wenig handfesten Geisteswerke „Stücke“ sein können, will ich nicht weiter sinnieren, denn die Archivkopie hat einige unangenehme Beschränkungen zu beachten, die in § 53 Abs. 2 Satz 2 UrhG enthalten und für unsere Fragestellung erheblich interessanter sind.

So dürfen nach Nr. 1 und 2 von Satz 2 – hier spätestens ist übrigens Gelegenheit, den kunstvoll verschränkten Aufbau von § 53 UrhG zu bewundern –, nach Nr. 1 und 2 also dürfen die Vervielfältigung selbst oder ihre Nutzung bloß in analoger Form erfolgen, was für die digitale Langzeitarchivierung schlicht sinnlos ist. Allerdings gilt diese Beschränkung nicht für Archive, die im öffentlichen Interesse tätig sind und keine wirtschaftlichen oder Erwerbszwecke verfolgen. Diesen Einrichtungen ist auch die digitale Kopie zur Aufnahme in ein eigenes Archiv gestattet.<sup>12</sup>

10 Vgl. Dreier, Thomas. In: Dreier, Thomas/Schulze, Gernot: UrhG. Urheberrechtsgesetz. Urheberrechtswahrnehmungsgesetz, Kunsturhebergesetz. Kommentar. 4. Auflage. München 2013, § 53, Rn. 23.

11 Zur Erfordernis einer engen Schrankenauslegung vgl. Dreier, Thomas. In: Dreier, Thomas/Schulze, Gernot: UrhG. Urheberrechtsgesetz. Urheberrechtswahrnehmungsgesetz, Kunsturhebergesetz. Kommentar. 4. Auflage. München 2013, vor § 44a, Rn. 7.

12 Einzelheiten bei Dreier, Thomas. In: Dreier, Thomas/Schulze, Gernot: UrhG. Urheberrechtsgesetz. Urheberrechtswahrnehmungsgesetz, Kunsturhebergesetz. Kommentar. 4. Auflage. München 2013, § 53, Rn. 28.

Halten wir hier kurz inne. Die digitale Langzeitarchivierung scheint zwar nicht im Fokus des Gesetzgebers zu stehen, ist aber Einrichtungen gestattet, die im öffentlichen Interesse tätig werden. So jedenfalls kann man die Archivschränke in § 53 Abs. 2 Satz 1 Nr. 2 UrhG verstehen und damit der Ansicht sein, dass immerhin Bibliotheken und Archive, mithin öffentliche Gedächtnisinstitutionen aus urheberrechtlichen Gründen nicht gehindert sind, Maßnahmen der digitalen Langzeitarchivierung durchzuführen.

Diese Ansicht freilich ist etwas schlicht. Bei der Frage der digitalen Langzeitarchivierung geht es doch um den Aufbau eines digitalen Speichergedächtnisses, das in seiner Dauerhaftigkeit den gedruckten Büchern als den wohl wichtigsten Medien des überkommenen kulturellen Gedächtnisses entsprechen soll. Auch wenn man von der leidigen Formatfrage einmal absieht, die Haltbarkeit von Datenträgern gestattet es nicht, es mit einer einzigen Archivkopie für die Dauer des Urheberrechts an dem betreffenden Werk bewenden zu lassen. Um eine zuverlässige Speicherung zu gewährleisten, ist vielmehr die regelmäßige Migration auf einen neuen Datenträger nötig. Die Archivschränke aber gestattet nur Vervielfältigungen zur Aufnahme in ein Archiv, Kopien im Archiv selbst sind nicht im Fokus der Norm. Ob hier eine Auslegung der Schrankenbestimmung helfen kann, ist zweifelhaft.<sup>13</sup>

Wenn man nämlich schon in § 44a UrhG für Vervielfältigungen ohne eigenständige wirtschaftliche Bedeutung – der Gesetzgeber spricht sogar selbst davon! – eine eigene Schrankenbestimmung zu erlassen für notwendig hielt, dann sollte man sich über die Möglichkeiten einer erweiternden Auslegung von Schrankenbestimmungen keine allzu großen Illusionen machen. Die jüngsten Gerichtsverfahren gegen Hochschulen wegen §§ 52a und 52b UrhG zeigen zudem, welche komischen, den Erfordernissen einer digital arbeitenden Wissenschaft nicht unbedingt entgegenkommenden Auslegungen Schrankenbestimmungen durch Gerichte erfahren können.<sup>14</sup>

13 Ausführlich dazu Euler, Ellen: Das kulturelle Gedächtnis im Zeitalter digitaler und vernetzter Medien und sein Recht. Bad Honnef 2011, S. 163-172.

14 Dazu Steinhauer, Eric W.: Bemerkungen zum „Kröner-Urteil“ des OLG Stuttgart. In: Recht, Bibliothek, Dokumentation 42 (2012), S. 103-129; ders.: Die Reichweite der Unterrichtsschranke in der Hochschullehre. Zur Klage des Krönerverlages gegen die FernUniversität in Hagen. In: K&R (2011), S. 311-315; ders.: Der Standort des elektronischen Leseplatzes und seine Nutzung in den Räumen der Bibliothek. In: ZGE 2 (2010), S. 55-74.

Hält man sich dies vor Augen, ist man nicht gut beraten, auf dem sandigen Fundament der Archivschränke eine kostspielige Infrastruktur für die digitale Langzeitarchivierung aufzubauen. Unklares Recht ist hier ein ernstes Investitions- und Planungsrisiko!

Indes, das Recht ist nicht in allen Punkten unklar. Klar und deutlich etwa drückt sich der Gesetzgeber in § 53 Abs. 5 UrhG aus. Danach ist eine Archivkopie von Datenbankwerken unzulässig; sie wird lediglich für den eigenen wissenschaftlichen Gebrauch sowie den Unterrichtsgebrauch unter Einschränkungen gestattet. Für Digitalia im Internet, die in Form von Datenbankwerken nicht unbedingt selten auftreten, ist das eine ziemlich klare, aber keine gute Aussage. Sekundiert wird dieses wenig erfreuliche Ergebnis von § 87c Abs. 1 UrhG, der auch für einfache Datenbanken eine Archivkopie nicht kennt, dafür freilich die Privatkopie sowie, hier § 53 Abs. 5 UrhG entsprechend, Vervielfältigungen für den wissenschaftlichen und den Unterrichtsgebrauch.

Als raffinierter Höhepunkt der klaren Aussagen des Gesetzgebers dürfen die technischen Schutzmaßnahmen nicht vergessen werden. Soweit es um Digitalia geht, die in Netzen kursieren, besteht von Rechts wegen keine Möglichkeit, technische Schutzmaßnahmen für Vervielfältigungen aufzuheben, § 95b Abs. 3 UrhG.<sup>15</sup>

So weit also zu den Möglichkeiten, die das Urheberrecht zur Langzeitarchivierung von Digitalia eröffnet. Es liegt auf der Hand, dass diese Möglichkeiten nicht ausreichen. Außerhalb vertraglicher Vereinbarungen mit den Urhebern bzw. Rechteinhabern aber ist eine rechtlich saubere digitale Langzeitarchivierung im Rahmen der derzeitigen urheberrechtlichen Schranken nicht möglich.

Auf eine Schrankenlösung wird man aber nicht verzichten können, da insbesondere Netzinhalte oft kollaborativ und ohne eindeutige Klärung juristischer Fragen erstellt werden. Es ist in diesen Fällen oft nicht möglich, einen Ansprechpartner zur Einholung aller für die Langzeitarchivierung erforderlichen Rechte ausfindig zu machen.<sup>16</sup> Überdies erscheint es angesichts der Masse an relevanten

15 Dazu kritisch Dreier, Thomas. In: Dreier, Thomas/Schulze, Gernot: UrhG. Urheberrechtsgesetz. Urheberrechtswahrnehmungsgesetz, Kunsturhebergesetz. Kommentar. 4. Auflage. München 2013, § 95b, Rn. 17.

16 Die in §§ 61ff. UrhG geplanten Regelungen über verwaiste Werke werden hier nicht weiterhelfen, da sie für reine Online-Publikationen nicht gelten. Vgl. den Gesetzentwurf der Bundesregierung in BT-Drs. 17/13423.

Digitalia wenig sachgerecht, neben den ohnehin schon hohen technischen Aufwand der Langzeitarchivierung noch einen bürokratischen Aufwand zur Klärung rechtlicher Fragen zu treiben.

### *Sammeln und Vermitteln*

Wenn wir über die urheberrechtlichen Probleme beim Aufbau eines kulturellen Speichergedächtnisses für Digitalia sprechen, dann können wir nicht bei der Frage der Langzeitarchivierung stehen bleiben. Ein Speicher nämlich will gefüllt und will genutzt sein. Bibliothekare sprechen hier in ihren einfachen Worten von Sammeln und Vermitteln.<sup>17</sup>

### *Sammeln*

Digitalia sammelt man nicht einfach so wie Bücher, es sei denn, sie sind, was jedoch immer weniger der Fall ist, auf körperlichen Datenträgern, CD-ROMs etwa, verfügbar. Digitalia sammelt man, indem man sie auf eigene Speichermedien kopiert, also vervielfältigt. Solche Vervielfältigungen sind zweifelsfrei legal, wenn sie auf Grundlage einer Lizenzvereinbarung erfolgen. Allerdings ist nur ein Bruchteil des für das digitale kulturelle Gedächtnis relevanten Materials auf diese Weise verfügbar, etwa weil es von Verwertern entsprechend vermarktet oder auf Grundlage diverser Open-Source-Lizenzen im Internet bereitgestellt wird.<sup>18</sup> Gerade frei zugängliche Internetquellen stehen oft unter keiner Lizenzbestimmung. Was tun?

Der Blick auf die urheberrechtlichen Schranken kann kurz gehalten werden. Abgesehen vom privaten oder eigenen wissenschaftlichen Gebrauch, beides Möglichkeiten, die für die klassischen Gedächtnisinstitutionen wie Archive oder Bibliotheken nicht in Betracht kommen, ist das Sammeln von Digitalia aus dem Internet auf Grundlage urheberrechtlicher Schrankenbestimmung nicht möglich. Die schon erwähnte Archivschränke scheidet hier aus, weil das Ergebnis der Schrankennutzung die Voraussetzung ihrer Anwendung ist, nämlich das eigene Werkstück als allein zulässige Kopiervorlage.

17 Vgl. Gantert, Klaus/Hacker, Rupert: Bibliothekarisches Grundwissen 8. München 2008, S. 11.

18 Zu diesen freien Lizenzen Bartlakowski, Katja/Talke, Armin/Steinhauer, Eric W.: Bibliotheksurheberrecht. Bad Honnef 2010, S. 173-176.

Trotz dieser sehr misslichen urheberrechtlichen Situation hat der Gesetzgeber mit dem Gesetz über die Deutsche Nationalbibliothek (DNBG) im Jahre 2006 dieser Bibliothek die Aufgabe zugewiesen, auch Digitalia als Kulturgut für die Nachwelt zu sichern, also ein digitales kulturelles Gedächtnis anzulegen.<sup>19</sup> Dabei stellte sich der Gesetzgeber ausweislich der Gesetzgebungsmaterialien sogar vor, dass die Deutsche Nationalbibliothek in regelmäßigen Abständen gewissermaßen Sicherungskopien des gesamten deutschsprachigen Internets im Wege des Web-Harvesting erstellt.<sup>20</sup> Das aber ist schlicht illegal.<sup>21</sup> Das Urheberrechtsgesetz enthält keine entsprechende Schranke.<sup>22</sup> Und auch das Nationalbibliotheksgesetz selbst regelt diesen Sachverhalt nicht, obwohl es den Sammelauftrag der Nationalbibliothek explizit auf unkörperliche Medienwerke und damit Digitalia in Netzen erweitert hat. Ansonsten aber ist das Gesetz stockkonservativ. Ganz das sachenrechtlich geprägte Leitbild des traditionellen Pflichtexemplarrechts fortschreibend, müssen nach dem DNBG nicht nur Bücher und Datenträger, sondern auch Netzpublikationen abgeliefert werden. Ähnlich

19 Fundstelle: BGBl. I 2006, S. 1338.

20 Vgl. BR-Drs. 396/05, S. 18: „Die daraus entstehende Fülle von einzelnen digitalen Objekten kann aus wirtschaftlichen Gründen nicht in einer gezielten Auswahl gesammelt werden. Vielmehr sollen für diesen Sammlungszweig automatische Such- und Speicher Methoden für das umfassende Sammeln von Internetseiten zum Einsatz kommen (Web-Harvesting). Hierbei werden die Inhalte von bestimmten Bereichen (zum Beispiel alle Websites mit der Domain ‚de‘) automatisiert eingesammelt.“

21 Im Ergebnis auch Dreier, Thomas. In: Dreier, Thomas/Schulze, Gernot: UrhG. Urheberrechtsgesetz. Urheberrechtswahrnehmungsgesetz, Kunsturhebergesetz. Kommentar. 4. Auflage. München 2013, vor § 44a, Rn. 3.

22 Dies übersieht offenbar Lenski, Sophie-Charlotte: Öffentliches Kulturrecht. Tübingen 2013, S. 273, die – intuitiv – völlig richtig anmerkt, dass Bibliotheken frei zugänglich Inhalte doch „ohne Weiteres“ selbst sammeln können. Aber zwischen einer bloßen gesetzlichen Zuständigkeit und einer rechtlichen Befugnis besteht eben ein Unterschied. Vgl. auch Dreier, Thomas. In: Dreier, Thomas/Schulze, Gernot: UrhG. Urheberrechtsgesetz. Urheberrechtswahrnehmungsgesetz, Kunsturhebergesetz. Kommentar. 4. Auflage. München 2013, § 53, Rn. 27.

„fortschrittliche“ Bestimmungen enthalten die Pflichtexemplarvorschriften der Länder Baden-Württemberg,<sup>23</sup> Hamburg,<sup>24</sup> Thüringen<sup>25</sup> und Sachsen-Anhalt.<sup>26</sup>

Eine spannende Frage ist hier übrigens, ob man bei der Ablieferung der Digitalia eine schlüssig erklärte Nutzungsrechtseinräumung für die schon angesprochenen Maßnahmen der Langzeitarchivierung annehmen kann. Einige Stimmen in der Literatur bejahen dies, wohl aus Barmherzigkeit, damit die entsprechenden gesetzlichen Vorschriften nicht völlig sinnlos sind.<sup>27</sup>

Im Eindruck dieser Diskussion haben die jüngst novellierten Pflichtexemplarvorschriften in Hessen und Nordrhein-Westfalen die Frage der Nutzungsrechte bei Netzpublikationen daher direkt aufgegriffen. Nach § 4a Abs. 3 Satz 3 Hessisches Bibliotheksgesetz müssen die zu sammelnden Medienwerke, zu denen jetzt auch Netzpublikationen gehören, „ohne rechtliche [...] Beschränkungen“ nutzbar sein, was entsprechende Nutzungsrechte der Bibliothek, die bei der Ablieferung schlüssig eingeräumt werden, voraussetzt. Kommt ein Ablieferungspflichtiger seiner Verpflichtung zur Ablieferung nicht nach, so greift nach § 4a Abs. 3 Satz 7 Hessisches Bibliotheksgesetz sogar ein gesetzliches Nutzungsrecht, das einen verwaltungsvollstreckungsrechtlichen Charakter hat.<sup>28</sup> Der nordrhein-westfälische Gesetzgeber spricht in § 4 Abs. 5 Satz 1 des neu gefassten Pflichtexemplargesetzes das Thema Nutzungsrechte bei Netzpublikationen explizit an:

23 § 1a Gesetz über die Ablieferung von Pflichtexemplaren an die Badische Landesbibliothek in Karlsruhe und die Württembergische Landesbibliothek in Stuttgart: „Für digitale Publikationen gelten die Vorschriften dieses Gesetzes entsprechend. Digitale Publikationen sind Medienwerke in unkörperlicher Form, die in öffentlichen Netzen dargestellt werden.“

24 § 1 Satz 2 Gesetz über die Ablieferung von Pflichtexemplaren: „Für digitale Publikationen gelten die Vorschriften dieses Gesetzes entsprechend.“

25 § 12 Abs. 3 Satz 1 Thüringer Pressegesetz: „Für digitale Publikationen gilt Absatz 1 entsprechend.“

26 § 11 Abs. 6 Satz 1 Pressegesetz für das Land Sachsen-Anhalt: „Für digitale Publikationen gilt Absatz 1 mit der Maßgabe entsprechend, dass zur Ablieferung verpflichtet ist, wer den betreffenden Datenträger wie ein Verleger oder gleichgestellter Drucker oder sonstiger Hersteller im Sinne von Absatz 1 verbreitet oder berechtigt ist, die betreffende digitale Publikation öffentlich zugänglich zu machen, und den Sitz, eine Betriebsstätte oder den Hauptwohnsitz in Sachsen-Anhalt hat.“

27 Vgl. Heckmann, Jörn/Weber, Marc Philipp: Elektronische Netzpublikationen im Lichte des Gesetzes über die Deutsche Nationalbibliothek (DNBG). In: AfP (2008), S. 269-276.

28 „Frei zugängliche unkörperliche Medienwerke, die der Ablieferungspflicht unterliegen, kann die Bibliothek nach Ablauf der vorstehend genannten Fristen in ihren Bestand übernehmen und im Rahmen ihres gesetzlichen Auftrages nutzen.“

„Mit der Ablieferung eines Medienwerkes auf einem elektronischen Datenträger oder eines Medienwerkes in unkörperlicher Form erhält die Bibliothek das Recht, das Werk zu speichern, zu vervielfältigen und zu verändern oder diese Handlungen in ihrem Auftrag vornehmen zu lassen, soweit dies notwendig ist, um das Medienwerk in die Sammlung aufnehmen, erschließen und für die Benutzung bereitstellen zu können sowie seine Erhaltung und Benutzbarkeit dauerhaft zu sichern.“

Eine eigene Sammlungsbefugnis der Bibliothek bei Säumigkeit der Ablieferung ist in Nordrhein-Westfalen im Gegensatz zu Hessen nicht explizit vorgesehen.

Halten wir hier kurz fest: Pflichtexemplarbibliotheken dürfen, sofern eine gesetzliche Grundlage besteht, Digitalia institutionell sammeln, wenn und soweit sie abgeliefert werden. Daneben dürfen sich aber auch Privatpersonen und Wissenschaftler kleine Sammlungen anlegen.

### *Vermitteln*

Neben der Langzeitarchivierung und dem Sammeln ist schließlich noch das Vermitteln als dritte wichtige Säule des digitalen kulturellen Speichergedächtnisses zu beleuchten. Dass hier vertraglich alles geregelt werden kann, ist trivial. Im Schrankenbereich indes gilt es, eine Bestimmung zu finden, die eine Nutzung der im kulturellen digitalen Gedächtnis niedergelegten Inhalte durch jedermann erlaubt. Die soeben kurz erwähnten kleinen Sammlungen von Wissenschaftlern und Privatpersonen scheiden sofort aus, weil § 53 Abs. 6 UrhG eine Öffnung dieser Sammlungen an außenstehende Dritte untersagt. Für Bibliotheken und Archive im Prinzip interessant hingegen ist § 52b UrhG. Danach darf man sich das digitale kulturelle Speichergedächtnis im Stile einer Mikroformsammlung an Bildschirmen in ausgewählten Pflichtexemplarbibliotheken ansehen und sich von den multimedialen Inhalten eigene Exzerpte machen, vielleicht ein paar Screens abfotografieren oder den Bildschirm auf den Kopierer legen. Für diese komfortable Form der Nutzung, die auch das Ergebnis der Rechtsprechung unserer Zeit ist,<sup>29</sup> ist freilich eine angemessene Vergütung abzuführen. Wobei hier

29 Der BGH hat jüngst in einer Entscheidung zu § 52b UrhG (I ZR 69/11) sein Verständnis für digitales Arbeiten bündig so zum Ausdruck gebracht: „Für ein wissenschaftliches Arbeiten mit Texten ist nach heutigem Verständnis zwar der Ausdruck, nicht aber das Abspeichern wichtiger Textpassagen unerlässlich.“ Zitiert nach GRUR 2013, S. 506.



durchaus die Frage aufgeworfen werden kann, ob diese Vergütungspflicht für Pflichtexemplarbibliotheken im Rahmen ihres Sammelauftrages überhaupt angemessen ist, denn immerhin investieren diese Einrichtungen eine Menge Geld in den Erhalt von Digitalia für die Nachwelt.<sup>30</sup>

Das Problem freilich ist nur ein theoretisches. Von sich aus dürfen Pflichtexemplarbibliotheken, wie wir festgestellt haben, ja gar nichts sammeln.<sup>31</sup> Und was sie abgeliefert bekommen, ist konkludent entsprechend lizenziert, sodass hier eine vertragliche Vereinbarung der Anwendung von § 52b UrhG entgegensteht. Anders gesagt, das Wenige, das gesammelt wird, darf man auch der Öffentlichkeit vermitteln, jedenfalls am stationären Leseplatz im Haus.

### *Urheberrechtliches Fazit*

Dieser Beitrag trägt den Titel „Wissen ohne Zukunft?“. Nachdem wir uns die urheberrechtliche Seite der Langzeitarchivierung angesehen haben, kann man sagen, dass man sich um die Zukunft unserer Digitalia im kulturellen Speichergedächtnis durchaus Sorgen machen muss.

Im Gegensatz zur analogen Welt der gedruckten Bücher etwa unterliegen im digitalen Bereich alle für den Aufbau, die Pflege und die Nutzung des digitalen Gedächtnisses relevanten Handlungen dem Urheberrecht. Es fehlt an einer dem Sachenrecht entsprechenden eigenen Rechtsposition („Eigentum am bedruckten Papier“) von Gedächtnisinstitutionen wie Bibliotheken oder Archiven. Vertragliche Vereinbarungen sind angesichts der Masse des relevanten Materials

30 Vgl. auch Steinhauer, Eric W.: Pflichtablieferung von Netzpublikationen. Urheberrechtliche Probleme im Zusammenhang mit der Ablieferungspflicht von Netzpublikationen an die Deutsche Nationalbibliothek. In: K&R (2009), S. 165; ders.: Stellungnahme zu dem Entwurf eines Gesetzes über die Ablieferung von Pflichtexemplaren in Nordrhein-Westfalen (Pflichtexemplargesetz Nordrhein-Westfalen) (Drucksache 16/179), Stellungnahme (Nordrhein-Westfalen) 16/231, S. 8.

31 Lediglich in Hessen ist eine Sammelermächtigung bei Säumigkeit des Ablieferungspflichtigen vorgesehen. Diese aus kulturstaatlichen Gründen richtige Regelung ist angesichts der fehlenden Gesetzgebungskompetenz der Bundesländer im Bereich des Urheberrechts verfassungsrechtlich gerade noch hinnehmbar. Vgl. Steinhauer, Eric W.: Stellungnahme zu dem Entwurf eines Gesetzes zur Neuregelung des Archivwesens und des Pflichtexemplarrechts (Drucksache 18/6067). In: Ausschussvorlage (Hessen) WKA 18/45, S. 12-14.

keine Lösung. Die vorhandenen urheberrechtlichen Schranken bieten ebenfalls keine geeignete Grundlage für die digitale Langzeitarchivierung. Hier ist der Gesetzgeber gefragt.

Beim Erlass des Nationalbibliotheksgesetzes war die Ausweitung des Sammel- und Sicherungsauftrages der Bibliothek auch auf Digitalia dem Gesetzgeber wichtig. Flankiert werden sollten solche Regelungen aber unbedingt durch entsprechende Vorschriften im UrhG. Die bereits erwähnten Vorschriften im Pflichtexemplarrecht der Länder Hessen und Nordrhein-Westfalen, die Nutzungsrechte ausdrücklich fordern bzw. berücksichtigen, sind hier im Ergebnis eine zwar noch zulässige,<sup>32</sup> aber verfassungsrechtlich bereits beargwöhnte<sup>33</sup> Notwehr der Landesgesetzgebung gegen die Untätigkeit des Bundesgesetzgebers. Es ist an der Zeit, dass die für Gedächtnisinstitutionen in vielen Fällen derzeit bestehende „Rechtspflicht zur Amnesie“ im digitalen Bereich durch eine sachgerechte Novellierung des Urheberrechtsgesetzes endlich beendet wird.

Um jetzt möglichen Gegenstimmen, die sogleich an das enge europäische Korsett im Urheberrecht erinnern werden,<sup>34</sup> zu begegnen: Es gibt bereits auf europäischer Ebene ein überaus interessantes und für unsere Fragestellung einschlägiges Dokument. Es handelt sich um die „Empfehlung der Kommission vom 27.10.2011 zur Digitalisierung und Online-Zugänglichkeit kulturellen Materials und dessen digitaler Bewahrung“.<sup>35</sup> Darin empfiehlt die Kommission „die ausdrückliche und eindeutige Verankerung von Bestimmungen in ihren Rechtsordnungen, die ein mehrfaches Kopieren und Konvertieren digitalen kulturellen Materials durch öffentliche Einrichtungen zum Zwecke der Bewahrung erlauben“, sowie „die Schaffung der erforderlichen Regelungen für die Pflichtexemplarhinterlegung in digitaler Form“. Zu den notwendigen Pflichtexemplarvorschriften rechnet die Kommission auch die Gestattung von Web-Harvesting und des gegenseitigen Austausches von Digitalia zwischen Pflichtexemplarbibliotheken, sofern diese einen eigenen gesetzlichen Sammelausspruch haben.

32 Vgl. Steinhauer, Eric W.: Stellungnahme zu dem Entwurf eines Gesetzes über die Ablieferung von Pflichtexemplaren in Nordrhein-Westfalen (Pflichtexemplargesetz Nordrhein-Westfalen) (Drucksache 16/179), Stellungnahme (Nordrhein-Westfalen) 16/231, S. 10f.

33 Vgl. Abg. Ingola Schmitz (FDP). In: PIPr. (NRW) 16/20, S. 1516f.

34 InfoSoc.RL Art. 9: „Diese Richtlinie lässt andere Rechtsvorschriften insbesondere in folgenden Bereichen unberührt: [...] Schutz nationalen Kulturguts.“

35 EU K(2011) 7579. endg.

Dass die Empfehlungen der Kommission bereits im geltenden richtlichen Rahmen der Europäischen Union umsetzbar sind, zeigt das Beispiel des Österreichischen Mediengesetzes. In Österreich, das genauso wie Deutschland bestehende europäische Vorgaben im Urheberrecht beachten muss, gibt es in § 43 Buchstabe b) bis d) des Österreichischen Mediengesetzes bereits ein recht differenziertes Regelwerk,<sup>36</sup> das sich der deutsche Gesetzgeber als Beispiel nehmen könnte.<sup>37</sup>

*Exkurs: Digitale Langzeitarchivierung – ein schillernder Begriff*

Das Ergebnis der urheberrechtlichen Bestandsaufnahme ist, um es mit einem Wort zu sagen, vernichtend. Da verwundert es doch, wenn man in Plenardebatten des Deutschen Bundestages Sätze wie diese hier hört: „Das Urheberrecht ermöglicht bereits heute die Langzeitdigitalisierung, zum Beispiel durch Archive und Museen. Hier müssen wir nicht gegen irgendwelche Defizite kämpfen.“<sup>38</sup> Will man nicht gleich in billige Politikerschelte verfallen, sollte man genauer hinsehen, was hier eigentlich gemeint ist. Hier geht es nämlich nicht um die Frage der Bestandserhaltung von digitalen Inhalten, sondern um die Bestandserhaltung von analogen Inhalten durch Digitalisierung. Digitale Langzeitarchivierung meint hier also nicht die Langzeitarchivierung von Digitalia, sondern mit Hilfe von Digitalia. Das urheberrechtliche Problem, um das es in diesem Kontext geht, ist das der Digitalisierungsbefugnis. Anders gefragt: Darf ich urheberrechtlich geschützte Werke digitalisieren, um die analogen Originale zu schonen? Die Antwort auf diese Frage fällt, und da stimmt die gerade zitierte Aussage wieder, gar nicht so negativ aus. So dürfen etwa Gedächtnisinstitutionen auf Grundlage der Archivschränke in § 53 Abs. 2 Satz 1 Nr. 2 UrhG ihren gesamten Bestand digitalisieren und im Rahmen von § 52b UrhG auch in dieser

36 Vgl. dazu Berka, Walter/Heindl, Lucie/Höhne, Thomas/Noll, Alfred J.: Mediengesetz<sup>3</sup>. Wien 2012, S. 450-455.

37 Nur kurz hingewiesen sei noch auf den britischen Legal Deposit Libraries Act 2003, der in Section 8 explizit die urheberrechtliche Befugnis zum Einsammeln von Netzpublikationen erteilt und in Section 7 weitere urheberrechtliche Befugnisse der zuständigen Pflichtexemplarbibliothek anspricht: [www.legislation.gov.uk/ukpga/2003/28/contents](http://www.legislation.gov.uk/ukpga/2003/28/contents) (Letzter Aufruf: 19.06.2013).

38 Abg. Reiner Deutschmann (FDP). In: PlPr. 17/155, S. 18595.

Form zur Nutzung bereitstellen.<sup>39</sup> Weiterhin erlauben die neuen Regelungen über verwaiste Werke nicht nur deren Digitalisierung, sondern auch deren öffentliche Zugänglichmachung im Internet.<sup>40</sup> Lediglich die Frage, ob die einmal erstellten Digitalisate selbst dauerhaft erhalten werden dürfen, könnte problematisch sein. Bei § 53 Abs. 2 Satz 1 Nr. 2 UrhG bereiten die dafür notwendigen Vervielfältigungen im Archiv Schwierigkeiten, davon war schon die Rede. Bei den digitalisierten verwaisten Werken aber gestattet die im Gesetz vorgesehene Vervielfältigungsbeugnis auch die für die Langzeitarchivierung notwendigen Maßnahmen. Das ergibt sich eindeutig aus Art. 6 Abs. 1 Buchstabe b) der Richtlinie 2012/28/EU über bestimmte zulässige Formen der Nutzung verwaister Werke,<sup>41</sup> der Vervielfältigungen zum Zweck der „Bewahrung“ ausdrücklich gestattet. In diesem Sinne wird daher auch der neue § 61 UrhG auszulegen sein. Wenn also in der Politik von digitaler Langzeitarchivierung die Rede ist, sollte ganz genau unterschieden werden, um welchen Sachverhalt es sich handelt.<sup>42</sup> Geht es um Bestandsdigitalisierung? Dann sind die Rechtsgrundlagen durchaus tragfähig. Oder geht es um die Sammlung, Bewahrung und Vermittlung von Netzpublikationen? Hier ist nahezu alles problematisch und eine rechtssichere digitale Langzeitarchivierung ebenso unmöglich wie der nachhaltige Aufbau und die Nutzung eines digitalen kulturellen Gedächtnisses.

### *Weitere rechtliche Probleme*

Das Urheberrecht stellt also eine der größten Herausforderungen im Bereich des digitalen kulturellen Gedächtnisses dar. Aber selbst wenn wir hier eine im Ergebnis zufriedenstellende Lösung für die Gedächtnisinstitutionen in einem

39 Vgl. Lutz, Alexander: Zugang zu wissenschaftlichen Informationen in der digitalen Welt. Tübingen 2012, S. 108 (str.).

40 Zu diesen neuen Regelungen vgl. de la Durantaye, Katharina: Die Nutzung verwaister und vergriffener Werke. Stellungnahme zu dem Gesetzentwurf der Bundesregierung. In: ZUM (2013), S. 437-445; Spindler, Gerald: Ein Durchbruch für die Retrodigitalisierung? In: ZUM (2013), S. 349-357.

41 Fundstelle: Abl. EU L 299/5.

42 Ein Beispiel ist hier der Antrag der SPD „Kulturelles Erbe 2.0 – Digitalisierung von Kulturgütern beschleunigen“ (BT-Drs. 17/6296), der die Bundesregierung auffordert, „eine urheberrechtliche Lösung für das Kopieren zur Langzeitarchivierung sowie die Zugänglichmachung und Nutzung der betreffenden Werke und Objekte durch und in Gedächtniseinrichtungen vorzulegen“. Netzpublikationen waren nicht im Fokus.

„Gedächtnis-Korb“ bekommen sollten,<sup>43</sup> die rechtlichen Probleme des digitalen kulturellen Speichergedächtnis sind damit bei Weitem noch nicht gelöst. Ich möchte beispielhaft auf drei Bereiche eingehen, die sich beim Aufbau eines nachhaltigen digitalen kulturellen Gedächtnisses als problemträchtig erweisen.

### *Netzwerke des Sammelns*

Zunächst werden für das digitale kulturelle Gedächtnis, wenn man das soeben zitierte Kommissionsdokument betrachtet, ausschließlich bestimmte Institutionen als zuständig erachtet, vor allem Bibliotheken und Archive. Diese Sichtweise ist nicht unproblematisch. Träger des kulturellen Speichergedächtnisses sind neben diesen Einrichtungen, die unbestritten wichtig sind, eine Fülle weiterer Akteure. Beispielhaft möchte ich hier private Sammler nennen. Neben dem großen Fundus an Kulturgütern in öffentlichen Institutionen existiert ein vielleicht noch größerer, jedenfalls vielfältigerer Schatz in privaten Sammlungen. Beide Bereiche sind aufeinander bezogen.<sup>44</sup> Man denke nur an den Antiquariatsbuchhandel. Bibliotheken können zudem zerstört werden. Die Vielzahl von Sammlungen ermöglicht in einem solchen Fall eine weitgehend lückenlose Überlieferungen von Kulturgütern. Oft sind es private Sammler, die wichtige Kollektionen für das kulturelle Gedächtnis erst zu einer für eine öffentliche Einrichtung interessanten Form zusammenstellen. Keine Gedächtnisinstitution würde Briefkastenwerbung von Edeka archivieren. Aber wenn ein Sammler einem Wirtschaftsarchiv nun ein vollständiges Konvolut solcher Werbezettel aus der Zeit von 1965 bis 2000 anböte? Das wäre fast ein DFG-Projekt! Oder man denke an die vielen Privatarhive der sozialen Bewegungen der 1970er und 1980er Jahre. Diese graue Literatur ist

43 Dazu Euler, Ellen/Steinhauer, Eric W./Bankhardt, Christina: Digitale Langzeitarchivierung als Thema für den 3. Korb zum Urheberrechtsgesetz. Urheberrechtliche Probleme der digitalen Langzeitarchivierung. Stellungnahme der AG Recht im Kompetenznetzwerk nestor. In: Bibliotheksdienst 45 (2011), S. 322-328. Gesetzgeberisches Handeln fordern auch Beger, Gabriele: Langzeitarchivierung und Recht. In: Dreier, Thomas/Euler, Ellen (Hrsg.): Kulturelles Gedächtnis im 21. Jahrhundert. Karlsruhe 2005, S. 83f. sowie der Abg. Jerzy Montag (Bündnis 90/Die Grünen), der in PIPr. 17/246, S. 31647 „die rechtliche Absicherung der digitalen Langzeitarchivierung“ als gesetzgeberisches Desiderat bezeichnet hat.

44 Vgl. beispielhaft Hügel, Hans-Otto: Sammeln. In: ders. (Hrsg.): Handbuch populäre Kultur. Stuttgart 2003, S. 388.

wertvoller Rohstoff für die Forschung und findet sich daher immer öfter in den Gedächtnisinstitutionen wieder. Gesammelt freilich haben andere.

Was aber ist mit bloß digital vorliegenden Inhalten? Könnte ich eine mit Kennerschaft zusammengestellte Sammlung von Fotos und Texten aus dem Internet, alles legale Kopien für den wissenschaftlichen Eigengebrauch, einfach so einer Bibliothek übergeben? Und könnte die Bibliothek damit überhaupt etwas anfangen? Ich will diese Fragen, die wir der Sache nach aus dem Verbraucherschutzrecht und dem Problem der fehlenden digitalen Erschöpfung kennen, nach dem Motto „Was kann ich mit meinen bei iTunes gekauften Musikstücken eigentlich machen?“, nur anreißen, um auf das Problem aufmerksam zu machen, dass im analogen Zeitalter jedermann, der am Kulturleben teilnahm, einen substantiellen Beitrag zur Schaffung des kulturellen Speichergedächtnisses leisten konnte und in sehr vielen Fällen auch geleistet hat und dass im digitalen Bereich mit seiner ungleich größeren Vielfalt der Fokus der Diskussion sich auf wenige institutionelle Akteure verengt. Ich habe aus „sammlungsphilosophischen“ Gründen meine Zweifel, dass das sachgerecht ist.

### *Persönlichkeitsrechte*

Kommen wir zu einem weiteren Problemfeld. Ein Buch eines süddeutschen Kleinverlegers, ein Artikel in der Zeitung und eine Meldung im Fernsehen zur Hauptsendezeit sind allesamt öffentliche Äußerungen, haben aber gleichwohl unterschiedliche Reichweiten. Das Buch nimmt wahrscheinlich niemand wahr, Presse und Fernsehen hingegen erreichen viele Menschen. Allerdings ist das Buch geduldiger und kann noch nach Jahren seine Leser finden, während Äußerungen in den Massenmedien schnell vergessen sind und von neuen Meldungen überlagert werden. Wenn es um Berichte über unvorteilhaftes persönliches Verhalten geht, Straftaten etwa, dumme Plagiate oder eine kleine Affäre, ist es gerade diese Vergesslichkeit der Massenmedien, die aus persönlichkeitsrechtlichen Gründen eine Berichterstattung darüber ermöglicht. Das Bundesverfassungsgericht hat es in der berühmten Lebach-Entscheidung von 1973 daher für unzulässig gehalten, „dass die Kommunikationsmedien sich [...] zeitlich unbeschränkt mit der Person eines Straftäters und seiner Privatsphäre befassen“.<sup>45</sup> Damals war das Fernsehen das Leitmedium.

45 BVerfGE 35, 202, 233.

Was gilt im Internetzeitalter, wenn Google mir über digital verfügbare Medienarchive 20 Jahre alte Skandale wie frische Meldungen serviert? Auch dies möchte ich als Frage stehen lassen. Sie zeigt, dass im digitalen kulturellen Speichergedächtnis persönlichkeitsrechtliche Fragestellungen, durchaus auch post-mortale, eine völlig neue Relevanz bekommen und in bestimmten Erinnerungsbereichen das im Vergleich zum Urheberrecht nicht unbedingt kleinere Problem darstellen werden. Ganz neu sind solche Fragen übrigens nicht. In den 1970er und 1980er Jahren gab es eine intensive Abgrenzungsdebatte zwischen Datenschutz- und Archivrecht.<sup>46</sup> Auch damals ging es um die Alternative Persönlichkeitsrecht und Unverfälschtheit historischer Quellen.

### *Imperative des Vergessens*

Der dritte Problemkreis, den ich ansprechen möchte, hängt eng mit dem Vorhergehenden zusammen, es geht um Imperative des Vergessens.

„Das Netz speichert zu viel, es soll auch wieder vergessen!“ Diese Forderung wird oft erhoben.<sup>47</sup> Tatsächlich könnte ein Urheber ein Werk einfach aus dem Internet löschen. Sofern er keine Nutzungsrechte eingeräumt hat, können Dritte es nicht mehr zugänglich machen. Was aber, wenn das Werk bereits an die Deutsche Nationalbibliothek im Wege der Pflichtablieferung gelangt ist? Könnte jetzt der Urheber nach § 42 UrhG eine Löschung verlangen? Hier wird deutlich, wie anfällig für gezielte Manipulationen und Streichungen das digitale Gedächtnis im Vergleich zum gedruckten ist. Die Vielzahl von nicht mehr manipulierbaren Werkstücken und der Erschöpfungsgrundsatz verhindern hier einen Zugriff des Urhebers auf sein einmal publiziertes und verbreitetes Werk.

Neben dem Urheber, der sein Werk unterdrücken möchte, gibt es noch einen ungleich problematischeren Imperativ des Vergessens. Betrachtet man das Gesetz über die Deutsche Nationalbibliothek und politische Äußerungen zum digitalen Kulturerbe, dann ist man sehr irritiert, wenn man sich näher mit der Praxis des Depublizierens von Webseiten der öffentlich-rechtlichen Rundfunkanstalten beschäftigt. Diese Internetseiten dürfen sicher, schon wegen ihrer hohen Reichweite und der Seriosität ihrer Erstellung, als zeitgeschichtliche Zeugnisse ersten

46 Vgl. Steinmüller, Wilhelm: Datenschutz im Archivwesen. Einige neue Argumente für ein bereichsspezifisches Archivgesetz. In: Der Archivar 33 (1980), S. 175-188.

47 Vgl. dazu Mayer-Schönberger, Viktor: Delete. Die Tugend des Vergessens in digitalen Zeiten, Berlin 2010.

Ranges gelten. Gleichwohl schreibt der Rundfunkstaatsvertrag in §§ 11d und 11f für einen Großteil dieser Inhalte eine bloß befristete Verweildauer vor.<sup>48</sup> Danach werden die Seiten vom Netz genommen.

Und dann? Dann sind die Seiten weg. Die Deutsche Nationalbibliothek fühlt sich unter Verweis auf ihre Sammelrichtlinien nicht zuständig – man kann das durchaus anders sehen! –, und die Rundfunkarchive, wo diese Inhalte möglicherweise noch zu finden sind, gleichen hermetischen Silos und haben keinen klaren Auftrag, ein digitales kulturelles Speichergedächtnis zu schaffen, das allen leicht zugänglich ist.

### *Viele Vorfragen sind ungeklärt*

Die drei exemplarisch angerissenen Problemkreise zeigen: Im digitalen Gedächtnis sind noch viele Fragen offen. Die Probleme rühren nicht nur von der Unkörperlichkeit der Digitalia her, sondern auch von ihrer medialen Ununterscheidbarkeit. Was im analogen Bereich sauberlich getrennt war, ist nun unterschiedslos einfach ein Datenstrom. Im Internet konvergieren Publikation, Kommunikation, Information, Unterhaltung und Business zu einem schwer entwirrbaren Gemisch. In der analogen Welt ist es klar, dass wir Bücher sammeln, Werbeflyer eher nicht und Telefongespräche gar nicht. Im Internet ist das nicht mehr klar. Neben einer angemessenen Rechtsordnung für das digitale kulturelle Gedächtnis brauchen wir auch dessen medien- bzw. kulturwissenschaftliche Theorie. Was soll dieses Gedächtnis eigentlich ausmachen? Und wenn wir es anlegen, sollten wir es vielleicht nur gestuft zugänglich machen? Persönliches nach 100 Jahren, Publiziertes sofort?

Erst wenn diese Fragen gelöst und in ein stimmiges Konzept gebracht sind, kann auch der Gesetzgeber ein sinnvolles Rechtsregime für Digitalia aufbauen. Am Rande sei erwähnt, dass dabei auch die Rolle der Digitalisate, also der digitalisierten analogen Kulturgüter, mit zu bedenken ist, denn bei aller Unsicherheit bei der Bewahrung von Digitalia werden paradoxerweise Unmengen analoger Kulturgüter digitalisiert, auch und gerade um sie der Nachwelt zu erhalten.<sup>49</sup> Ein weites Feld. Wir lassen es liegen und kommen zum Schluss.

48 Dazu Dreizehnter Zwischenbericht der Enquete-Kommission Internet und digitale Gesellschaft. Kultur, Medien und Öffentlichkeit (BT-Drs. 17/12542), S. 130f.

49 So jedenfalls sieht das auch die Politik. Vgl. nur den Abg. Ansgar Heveling (CDU). In: PIPr. 17/246, S. 31643: „Mithilfe der Digitalisierung unseres wissenschaftlichen und kulturellen Erbes wollen wir sicherstellen, dass dauerhaft Schriften, Filme und Tonträger zugänglich sind.“



*Ausblick*

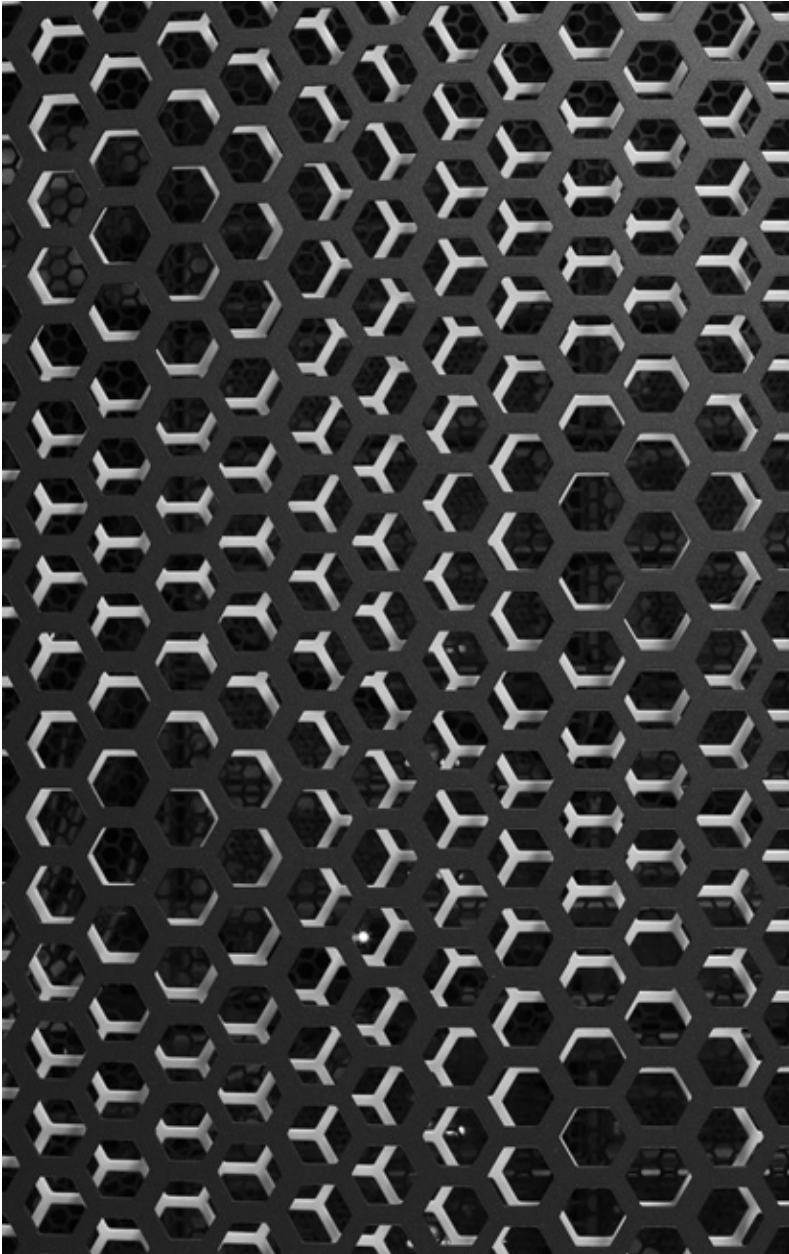
Der Ausblick auf die Zukunft des digitalen kulturellen Speichergedächtnisses ist ausgesprochen düster, allerdings sind in Gestalt der Initiative der Europäischen Kommission freundliche Aufheiterungen am Horizont sichtbar. Des ungeachtet bleibt die Aufgabe bestehen, ein stimmiges Rechtsregime für alle Fragen des digitalen kulturellen Speichergedächtnisses zu entwickeln. Das ist eine große Aufgabe. Konzertierte Bemühungen in diesem Bereich sehe ich derzeit nicht.

Speziell für das Urheberrecht liegt hier die in mittlerer Zukunft wohl wichtigste Bewährungsprobe. Dabei wird es nicht nur darum gehen, ob Digitalia echte, den alten Medien ebenbürtige Kulturgüter oder bloß mediale Wegwerfkonserven sind. Entscheidender wird der Einfluss des Urheberrechts auf das kulturelle Speichergedächtnis sein. Sollte sich in einer Generation herausstellen, dass vor allem das gegenwärtig stark urheber- bzw. verwerterzentrierte Recht des geistigen Eigentums den Aufbau nachhaltiger Speicher- und Sammelstrukturen verhindert hat, wird dies eine Debatte auslösen, die an Heftigkeit die gegenwärtigen Diskussionen um Nutzer- und Wissenschaftsrechte weit übertreffen wird. Zerstörte Gedächtnisspuren lösen stärkste Emotionen aus. Man denke nur an den Weimarer Bibliotheksbrand oder den Einsturz des Kölner Stadtarchivs. Wir betreten dann einen anthropologischen Resonanzboden, der das Urheberrecht in der Form, in der wir es kennen, nicht nur zum Wanken, sondern zum totalen Einsturz bringen könnte. Ob hier der archimedische Punkt der neuen Urheberrechtsordnung liegt, nach dem alle suchen? Wer weiß.

Eines aber ist sicher: Die Rechtsfragen des digitalen kulturellen Gedächtnisses sind vielfältig, bunt und spannend und zudem wissenschaftlich hoch ergiebig. Man sollte, auch und gerade in der Kulturpolitik, mehr daraus machen.<sup>50</sup>

50 Für den Gesetzgeber gilt hier die zutreffende Feststellung der Enquete-Kommission Internet und digitale Gesellschaft. In: Sechster Zwischenbericht der Enquete-Kommission Internet und digitale Gesellschaft. Bildung und Forschung (BT-Drs. 17/12029), S. 45: „Gleichzeitig erfordern Langzeitarchivierungen besondere rechtliche Rahmenbedingungen.“ Dieser Satz kann auch als Fazit des vorliegenden Beitrages gelesen werden.





Verkleidung eines Hochleistungsrechners, Konrad-Zuse-Zentrum für Informationstechnik Berlin

# Nahaufnahmen



# Digitale Langzeitarchivierung aus Sicht der IT

MAIK STÜHRENBERG

## *Einleitung*

Aus technischer Sicht ist die digitale Langzeitarchivierung ein äußerst aufwendiges Unterfangen, ist der technische Wandel doch geradezu sprichwörtlich. Diese schnelle Abfolge von Veränderungen, die schon für den regulären Nutzer nicht nur Vorteile mit sich bringt, stellt Archivare vor große Herausforderungen. Im Folgenden sollen die technischen Aspekte dieser Herausforderungen kurz skizziert werden. Zunächst ist zu unterscheiden zwischen zwei Arten von Problemfeldern, wobei das erste die Hardware betrifft, das zweite die Software. Weiterhin sind verschiedene Prozesse zu unterscheiden, die im Rahmen einer Archivierung eine Rolle spielen: der Import, die Aufbereitung und die eigentliche Speicherung.

## *Import vorhandener Daten*

Digital gespeicherte Informationen sind prinzipiell keiner Alterung unterworfen. Eine Digitalkopie ist vom Original nicht zu unterscheiden, eine Abnutzung oder sonstige Verschlechterung während des Kopiervorgangs findet nicht statt (das klassische Gegensatzpaar ist hier die Analogkopie, beispielsweise von einer Langspielplatte auf eine Compact Cassette, vs. das digitale Kopieren einer CD). Dies betrifft allerdings nur die Daten als solche, d. h. die eigentliche Bitfolge (Bitstream), in der die Informationen digital gespeichert sind. Die Trägermedien dagegen unterliegen sehr wohl alterungsbedingten Qualitätseinbußen. Daher müssen die Daten zunächst von ihrem ursprünglichen Medium kopiert werden. Hier können erste Probleme auftreten, wenn für das Originalmedium ein geeig-

netes Lesegerät fehlt. Die vor einer Dekade noch weitverbreiteten 3,5"-Diskettenlaufwerke befinden sich heute üblicherweise in keinem aktuellen Rechner mehr (ganz zu schweigen von älteren Formaten wie 5,25" oder 8"). Zwar gibt es für 3,5"-Disketten weiterhin Laufwerke, es ist aber davon auszugehen, dass auch sie bald vom Markt verschwinden werden. Für exotischere, aber zu ihren Hochzeiten durchaus weitverbreitete Formate wie ZIP-Medien wird es noch schwerer, passende Lesegeräte zu erwerben. Dazu kommt, dass beide genannten Medien zu den Magnetspeichern gehören, die – je nach Verwendung – eine eher geringe Lebensdauer aufweisen. Aber auch optische Medien sind – trotz des berührungslosen Lesevorgangs – einer Alterung unterworfen, die abhängig von der Zusammensetzung der Medien (bei CD-ROMs beispielsweise vom verwendeten Farbstoff, aus dem die Aufnahmeschicht besteht), aber auch der Lagerung (Lichtstrahlung, Temperatur und Feuchtigkeitsgehalt) ist. Je nach verwendetem Medium stehen verschiedene Fehlerkorrekturverfahren zur Verfügung, die zumindest teilweise das Wiederherstellen von Daten auch von nicht einwandfreien Medien ermöglichen. Ziel des erstmaligen Imports ist daher die Erstellung einer unveränderten Bitstream-Kopie. Zur Sicherstellung kommen üblicherweise verschiedene Verfahren der Überprüfung der Datenintegrität zum Einsatz, z. B. Prüfsummen. Ist eine solche Bitstream-Kopie erstellt, sind auf der Softwareseite verschiedene Aspekte die Datenformate betreffend zu beachten.

### *Datenformate*

Bei den Datenformaten ist zu unterscheiden zwischen Formaten, die abhängig vom ursprünglichen Speichermedium sind (z. B. Videodaten auf einer DVD oder der Aufbau einer Audio-CD gemäß Red-Book-Standard), und denen, die unabhängig davon sind (beispielsweise einzelne Dateien und Verzeichnisse auf einer Daten-CD nach ISO 9660:1988<sup>1</sup>). Im Folgenden soll es um Letztere gehen. Der Aufbau eines importierten Bitstreams wird durch ein solches Datenformat bestimmt. Datenformate werden im Wesentlichen aus zwei Gründen festgelegt: 1. als Ergebnis einer Spezifikation oder 2. auf Grundlage eines Anwendungsprogramms. Das Format kann im ersten Fall – je nach vorliegender Spezifikation – Ausprägung eines (inter-)nationalen Standards sein (De-jure-

1 ISO/IEC JTC 1/SC 23. Information processing – Volume and file structure of CD-ROM for information interchange. International Standard ISO 9660:1988.

Standard), im zweiten Fall – je nach Verbreitung des Anwendungsprogramms – eines De-facto-Standards. In beiden Fällen ist für eine erfolgreiche Archivierung entscheidend, ob die zugrundeliegende Spezifikation proprietär oder offen vorliegt. Ein konkretes Beispiel ist das doc-Dateiformat, das von der Textverarbeitung Microsoft Word bis zur Version 2004 als Standardformat verwendet wird. Es handelt sich um ein binäres Format, d. h., die Informationen liegen kodiert und nicht in einer textuell interpretierbaren Form vor. Im Umkehrschluss ist eine spezielle Software notwendig, um die so kodierten Daten ansehen und verarbeiten zu können. Obwohl das doc-Dateiformat aufgrund der weiten Verbreitung von Microsoft Office einen De-facto-Standard darstellt, können die darin gespeicherten Informationen aufgrund des lange Zeit nicht offengelegten Formats (einige Versionen wurden unter einer restriktiven Lizenz veröffentlicht, andere wiederum nicht) von Drittprogrammen nicht immer so dargestellt werden, wie der ursprüngliche Autor es vorgesehen hat – hier ist also zu unterscheiden zwischen den eigentlichen Informationen und der Präsentation bzw. Formatierung. Dieser Umstand wird noch dadurch erschwert, dass sich hinter der Dateieindung .doc nicht ein einheitliches Format verbirgt, sondern mehrmals interne Änderungen stattgefunden haben, um Merkmale neuerer Softwareversionen abzubilden. Seit der Version 2004 verwendet die Office Suite von Microsoft eine Reihe von neuen, auf XML basierenden Formaten, genannt Office Open XML (ISO/IEC 29500-1:2008<sup>2</sup>, ISO/IEC 29500-1:2011<sup>3</sup>). Datenformate, die auf der offenen Metasprache XML (Extensible Markup Language)<sup>4</sup> aufbauen, haben einen entscheidenden Vorteil: Sie sind textbasiert.<sup>5</sup> Der innere Aufbau einer solchen XML-basierten Auszeichnungssprache wird von einer (üblicherweise offen

- 2 ISO/IEC JTC 1/SC 34. Information technology — Document description and processing languages – Office Open XML File Formats – Part 1: Fundamentals and Markup Language Reference. International Standard ISO/IEC 29500-1:2008.
- 3 ISO/IEC JTC 1/SC 34. Information technology — Document description and processing languages – Office Open XML File Formats – Part 1: Fundamentals and Markup Language Reference. International Standard ISO/IEC 29500-1:2011.
- 4 Bray, Tim/Paoli, Jean/Sperberg-McQueen, C. M./Maler, Eve/Yergeau, François: Extensible Markup Language (XML) 1.0 (Fifth Edition). W3C Recommendation. World Wide Web Consortium (W3C). 2008. [www.w3.org/TR/2008/REC-xml-20081126/](http://www.w3.org/TR/2008/REC-xml-20081126/) (Letzter Aufruf: 15.08.2013).
- 5 Bei einer docx-Datei handelt es sich um ein mit dem bekannten Zip-Algorithmus gepacktes Verzeichnis, das u. a. die entsprechenden XML-Dateien sowie eventuell eingebundene Bilder enthält.



zugänglichen) Dokumentgrammatik bestimmt, die mit Hilfe standardisierter Formalismen (wie XML-Schema; RELAX NG, ISO/IEC 19757-2:2008<sup>6</sup>)<sup>7</sup> definiert wird, während die Syntax der Auszeichnung durch die offen zugängliche XML-Spezifikation festgelegt ist. Beispiele dafür sind die in geisteswissenschaftlichen Korpora recht häufig anzutreffenden TEI Guidelines (P5 2.3.0) oder auch DocBook<sup>8</sup> zur strukturierten Speicherung von technischen Dokumentationen. Während solcherart kodierte Informationen damit neutral gegenüber Anwendungsprogrammen sind, sind sie auch formatunabhängig, d. h., die Darstellung der Inhalte wird üblicherweise über zusätzliche Transformationen bzw. Stylesheets gesteuert, die ebenfalls mit Hilfe offener Standards wie CSS<sup>9</sup> oder XSLT<sup>10</sup> definiert sind. Dieser Umstand bedeutet für eine Lang-

- 6 RELAX NG, ISO/IEC 19757-2. Information technology — Document Schema Definition Language (DSDL) – Part 2: Regular-grammar-based validation. International Standard ISO/IEC 19757-2:2008.
- 7 Thompson, Henry S./Beech, David/Maloney, Murray/Mendelsohn, Noah: XML Schema Part 1: Structures. W3C Recommendation. World Wide Web Consortium (W3C). 2001. [www.w3.org/TR/2001/REC-xmlschema-1-20010502/](http://www.w3.org/TR/2001/REC-xmlschema-1-20010502/) (Letzter Aufruf: 15.08.2013); Fallside, David C.: XML Schema Part 0: Primer. W3C Recommendation. World Wide Web Consortium (W3C). [www.w3.org/TR/2001/REC-xmlschema-0-20010502/](http://www.w3.org/TR/2001/REC-xmlschema-0-20010502/) (Letzter Aufruf: 15.08.2013); Biron, Paul V./Malhotra, Ashok: XML Schema Part 2: Datatypes. W3C Recommendation. World Wide Web Consortium (W3C). 2001. [www.w3.org/TR/2001/REC-xmlschema-2-20010502/](http://www.w3.org/TR/2001/REC-xmlschema-2-20010502/) (Letzter Aufruf: 15.08.2013); Peterson, David/Gao, Shudi (Sandy)/Malhotra, Ashok/Sperberg-McQueen, C. M./Thompson, Henry S.: W3C XML Schema Definition Language (XSD) 1.1 Part 2: Datatypes. W3C Recommendation. World Wide Web Consortium (W3C). 2012. [www.w3.org/TR/2012/REC-xmlschema11-2-20120405/](http://www.w3.org/TR/2012/REC-xmlschema11-2-20120405/) (Letzter Aufruf: 15.08.2013); Gao, Shudi (Sandy)/Sperberg-McQueen, C. M./Thompson, Henry S.: W3C XML Schema Definition Language (XSD) 1.1 Part 1: Structures. W3C Recommendation. World Wide Web Consortium (W3C). 2012. [www.w3.org/TR/2012/REC-xmlschema11-1-20120405/](http://www.w3.org/TR/2012/REC-xmlschema11-1-20120405/) (Letzter Aufruf: 15.08.2013).
- 8 Walsh, Norman: DocBook 5: The Definitive Guide, hrsg. von Richard L. Hamilton. Sebastopol 2012.
- 9 Bos, Bert/Çelik, Tantek/Hickson, Ian/Wium Lie, Håkon: Cascading Style Sheets Level 2 Revision 1 (CSS 2.1) Specification. W3C Recommendation. World Wide Web Consortium (W3C). 2011. [www.w3.org/TR/2011/REC-CSS2-20110607/](http://www.w3.org/TR/2011/REC-CSS2-20110607/) (Letzter Aufruf: 15.08.2013).
- 10 Kay, Michael: XSL Transformations (XSLT) Version 2.0. W3C Recommendation. World Wide Web Consortium (W3C). 2007. [www.w3.org/TR/2007/REC-xslt2-20070123/](http://www.w3.org/TR/2007/REC-xslt2-20070123/) (Letzter Aufruf: 15.08.2013).

zeitarchivierung, dass neben der XML-Instanz auch das entsprechende Stylesheet (und eventuell eine zu verarbeitende Software, vgl. Abschnitt „Software“) gespeichert werden muss, um die ursprüngliche Darstellung zu erhalten.

Auch ist darauf zu achten, dass Textdateien sich hinsichtlich ihrer Zeichenkodierung (ANSI X3.4-1963<sup>11</sup> vs. ISO/IEC 8859-1:1998<sup>12</sup> vs. Unicode 6.2.0<sup>13</sup>) und der internen Kodierung der Steuerzeichen (z. B. für das Zeilenende), aber auch der Byte-Reihenfolge (Little Endian vs. Big Endian)<sup>14</sup> unterscheiden können. Hier kann eine Angleichung an eine einheitliche Repräsentation sinnvoll sein, um den späteren Verwaltungsaufwand zu minimieren. In diesem Sinne ist ein technischer Aspekt der Langzeitarchivierung, die Migration der Daten, nicht nur auf die Umwandlung von einem Datenmedium auf ein anderes, sondern teilweise auch auf die Umwandlung zwischen Datenformaten bezogen.

Während für textuelle Informationen die geschilderten Probleme noch handhabbar sind, bergen multimediale Darstellungen darüber hinaus weitere Probleme für die Langzeitarchivierung.<sup>15</sup>

### *Software*

Anwendungsprogramme können aus verschiedenen Gründen für eine Langzeitarchivierung relevant sein. Wie bereits im letzten Abschnitt erwähnt, ist bei proprietären Binärformaten für die korrekte Darstellung von Inhalt und Form die Verwendung der entsprechenden Software oftmals der einzige Weg, die Informationen in ihrer ursprünglichen Form anzuzeigen.<sup>16</sup> Kommerzielle Programme

11 ASA standard X3.4-196. American Standard Code for Information Interchange.

12 ISO/IEC JTC 1/SC 2. Information technology – 8-bit single-byte coded graphic character sets – Part 1: Latin alphabet No. 1. International Standard ISO/IEC 8859-1:1998.

13 The Unicode Consortium: The Unicode Standard. Version 6.2.0. 2012. [www.unicode.org/versions/Unicode6.2.0/](http://www.unicode.org/versions/Unicode6.2.0/) (Letzter Aufruf: 15.08.2013).

14 Vgl. Blanc, Bertrand/Maaraoui, Bob: Endianness or Where is Byte 0? White Paper. 3B Consultancy. 2005. [www.3b-consultancy.com/endianness05.pdf](http://www.3b-consultancy.com/endianness05.pdf) (Letzter Aufruf: 15.08.2013).

15 Vgl. den Beitrag von Jan Fröhlich in diesem Band.

16 Diese Einschränkung bezieht sich auf den Fall, dass der ursprüngliche Bitstream archiviert wurde. Darüber hinaus ist es problemlos möglich, Inhalt und Form im Rahmen einer Migration in für die Archivierung konzipierten und standardisierten Formaten wie PDF/A (ISO 19005-1:2005, ISO 19005-2:2011, ISO 19005-3:2012) zu speichern und zu präsentieren. ISO/TC

liegen üblicherweise ebenfalls nur als Binärdatei vor, die das Ergebnis des Kompilierens der ursprünglichen Quelldatei(en) darstellt. Ein solches Anwendungsprogramm ist festgelegt auf eine bestimmte Rechnerarchitektur (zu der gebräuchlichsten gehört die x86- bzw. x86-64-Architektur, die allgemein als PC-Architektur tituliert wird) und ein bestimmtes Betriebssystem. So laufen Programme, die für die PowerPC-Architektur der Apple-Macintosh-Rechner bis 2006 kompiliert wurden, nicht mehr auf aktuellen Versionen des OS X-Betriebssystems, da seit dem genannten Datum Apple ebenfalls die x86-Architektur einsetzt. Ebenso laufen Programme, die für Microsoft Windows erstellt wurden, nicht auf Linux-Systemen, obwohl diese üblicherweise die gleiche Architektur einsetzen. Selbst verschiedene Versionen eines Betriebssystems verhalten sich bei der Ausführung bestimmter Anwendungsprogramme unterschiedlich oder verhindern diese sogar, da ursprünglich vorhandene Systembestandteile wie Bibliotheken nicht mehr oder nur in veränderter Form vorhanden sind. Für die Langzeitarchivierung bedeutet dies, dass neben der eigentlich zu archivierenden Datei auch ihr Kontext in Form des entsprechenden Anwendungsprogramms (und teilweise für dieses eine geeignete Betriebssystemumgebung) gespeichert werden muss. Hier bietet es sich an, Emulationen der ursprünglichen Umgebung (Betriebssystem und Anwendungssoftware) einzusetzen.<sup>17</sup> So enthielten frühere Versionen des Betriebssystems OS X bis zur Version 10.5 die Emulationsbibliothek Rosetta, die für die Version 10.6 bei Bedarf nachinstalliert werden konnte und die Ausführung von älteren Programmen ermöglichte. Seit der Version 10.7 ist das nicht mehr möglich. Zur Ausführung von Windows-Programmen unter Linux stehen mit dem quelloffenen WINE<sup>18</sup> und dessen kommerziellen Ableger Crossover<sup>19</sup> ebenfalls Emula-

171/SC 2. Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1). International Standard ISO 19005-1:2005; ISO/TC 171/SC 2. Document management – Electronic document file format for long-term preservation – Part 2: Use of ISO 32000-1 (PDF/A-2). International Standard ISO 19005-2:2011; ISO/TC 171/SC 2. Document management – Electronic document file format for long-term preservation – Use of ISO 32000-1 with support for embedded files (PDF/A-3). International Standard ISO 19005-3:2012.

17 Vgl. den Beitrag von Lange in diesem Band.

18 Vgl. [www.winehq.org/](http://www.winehq.org/) (Letzter Aufruf: 15.08.2013).

19 Vgl. [www.codeweavers.com/products/](http://www.codeweavers.com/products/) (Letzter Aufruf: 15.08.2013).

tionsmöglichkeiten zur Verfügung, diese unterstützen aber einen geringen Teil der Gesamtzahl an Anwendungen. Als Ultima Ratio kann hier der Einsatz von vollständig virtuellen Umgebungen (VM, virtuellen Maschinen) angesehen werden.

Eine Sonderrolle stellen Open-Source-Programme dar. Wie der Name schon suggeriert, ist bei dieser Software nicht nur das Kompilat, sondern auch der Programmquelltext vorhanden (als Beispiel seien hier der Kernel des freien Betriebssystems Linux, das sogenannte GNU-Userland, aber auch Anwendungsprogramme wie OpenOffice.org bzw. LibreOffice genannt). Damit lässt sich das Programm prinzipiell anpassen und neu kompilieren, z. B. um es auf neueren Betriebssystemversionen oder anderen Rechnerarchitekturen ausführen zu können. Damit geht allerdings die Gefahr einher, dass das so veränderte Programm auch in der Verarbeitung (und Anzeige) nicht mehr der ursprünglichen Version entspricht. Außerdem muss sichergestellt werden, dass auch aktuellere Compiler den Quelltext kompilieren können, was bei nicht-rückwärts-kompatiblen Änderungen innerhalb einer Programmiersprache problematisch sein kann. Vermeiden lassen sich solche Herausforderungen ebenfalls nur durch zusätzliche Archivierung entsprechender Compilerversionen in einer entsprechenden Umgebung (was ebenfalls auf eine Emulation hinauslaufen kann).

Ein weiteres Problem bei Software, das allerdings für die Langzeitarchivierung nur eine sekundäre Rolle spielt, ist das der Softwareerosion (engl. code rot). Mit diesem Begriff wird im Allgemeinen die Zustandsveränderung von Programmen beschrieben, die aufgrund häufiger Überarbeitungen im Zuge von Weiterentwicklungen weniger performant und deutlich schlechter zu warten werden. Da in der Langzeitarchivierung eine Modifikation der ursprünglich archivierten Daten ausgeschlossen wird, tritt diese Problematik nur im Zusammenhang mit geänderten Kontexten (Betriebssystem o. ä.) auf. Auch hier lässt sich durch die Emulation einer als lauffähig bekannten Umgebung das Risiko minimieren.

### *Archivierung*

Für die Speicherung archivierter Daten gelten in weitaus höherem Maße die bereits im Abschnitt über den „Import vorhandener Daten“ genannten Anforderungen.

Von Zeit zu Zeit müssen die Daten umkopiert werden, um der unausweichlichen Alterung der Trägermedien entgegenzuwirken, teilweise können dabei auch neue Trägermedien zum Einsatz kommen, sofern die ursprünglich verwendeten sich als obsolet herausstellen (die Ausführungen zur Migration im Abschnitt „Datenformate“). Dabei muss – wie beim erstmaligen Import – sichergestellt werden, dass die ursprünglichen Informationen konsistent bleiben, d. h., es muss gewährleistet sein, dass weder beim Kopiervorgang noch bei der späteren Archivierung Änderungen an den Dateien (Bitfehler o. ä.) auftreten, weshalb auch hierbei Prüfsummen und weitere Datenintegritätsverfahren zum Einsatz kommen (regelmäßige Bitstream-Kontrolle).<sup>20</sup>

### *Metadaten*

Sind Daten in digitaler Form gespeichert, muss der Zugriff auf diese sichergestellt werden. Dazu ist es notwendig, die eigentlichen Informationen mit weiteren Daten anzureichern, den Metadaten. Diese geben Auskunft über das eigentliche Datenobjekt und ermöglichen den einfachen Zugriff sowie die Gruppierung von Datenobjekten in größeren Zusammenhängen. Neben der Tatsache, dass für Metadaten die gleichen Prinzipien der Sorgfalt bei der Speicherung zu gelten haben wie für die Primärdaten, ist für die Langzeitarchivierung relevant, dass sie einem standardisierten Schema folgen. Auf XML basierende Beispiele wie der Metadata Encoding and Transmission Standard (METS)<sup>21</sup> sind weitverbreitet, sodass dieser Aspekt aus technischer Sicht unproblematisch erscheint.

Neuere Entwicklungen in diesem Bereich adressieren vorrangig Konzeptregistrierungen auf Basis von ISO 12620:2009 (hier sei ISOcat genannt)<sup>22</sup> sowie komponentenbasierende und hierarchisch organisierte Metadaten, die auch über verteilte Repositories hinweg zugänglich gemacht werden können

20 Weitere Angaben zur Technik der Langzeitarchivierung finden sich in dem Beitrag von Giebel in diesem Band.

21 Vgl. McDonough, Jerome: METS: standardized encoding for digital library objects. In: International Journal on Digital Libraries 6/2 (2006), S. 148–158. doi: 10.1007/s00799-005-0132-1.

22 ISO/TC 37/SC 3. Terminology and other language and content resources — Specification of data categories and management of a Data Category Registry for language resources. International Standard ISO 12620:2009.

(aus dem linguistischen Bereich z. B. die Component Metadata Infrastructure, CMDI).<sup>23</sup>

- 23 Broeder, Daan/Schonefeld, Oliver/Trippel, Thorsten/van Uytvanck, Dieter/Witt, Andreas: A pragmatic approach to XML interoperability – the Component Metadata Infrastructure (CMDI). In: Proceedings of Balisage: The Markup Conference. Bd. 7. Balisage Series on Markup Technologies. Montréal 2011. doi: 10.4242/BalisageVol7.Broeder01; Broeder, Daan/Windhouwer, Menzo/van Uytvanck, Dieter/Trippel, Thorsten/Goosen, Twan: CMDI: a Component Metadata Infrastructure. In: Arranz, Victoria/Broeder, Daan/Gaiffe, Bertrand/Gavrillidou, Maria/Monachini, Monica/Trippel, Thorsten (Hrsg.): Proceedings of the LREC 2012 Workshop on Describing LRs with Metadata: Towards Flexibility and Interoperability in the Documentation of LR. 2012, S. 1–4; Beitrag von Georg Rehm in diesem Band.



# Speichertechnologie und Nachhaltigkeit

RALPH GIEBEL

## *Ausgangspunkte*

Obwohl Mitarbeiter und Mitarbeiterinnen aus Archiven, Museen und Bibliotheken, Geisteswissenschaftler, Historiker in Zusammenarbeit mit IT-Firmen wichtige Standards definiert haben, sind zentrale Fragen nach wie vor offen. An erster Stelle ist hier die geeignete Speichertechnologie zu nennen. Die Erwartung, ein für alle Zeit geeignetes Medium werde gefunden, hat sich aber bis heute nicht erfüllt. Nach wie vor sind Technologiewechsel ein akutes Problem. Haben die Technologiehersteller dieses Problem nicht erkannt? Sind es technische Zwänge? Oder müssen wir gar unsere Erwartungshaltungen neu definieren? Diesen Fragen will ich in diesem Artikel nachgehen.

Mein Ansatz führt mich in die Geschichte der Computertechnologie. Ich möchte zunächst den Stellenwert und die Bedeutung der Speicher im Zusammenwirken mit anderen Komponenten der IT-Technologie deutlich machen. Danach werde ich Marktmechanismen untersuchen sowie auf das Konzept und den Wert von Information im Allgemeinen eingehen. Schließlich möchte ich die Frage vertiefen, welchen Beitrag die Speichertechnologie zum gesteckten Ziel liefern kann, digitale Informationen, die heute und morgen unser kulturelles Erbe beschreiben, nachhaltig zu bewahren.

## *Begriffe und Definitionen*

In Wikipedia, unserer digitalen Enzyklopädie, ist über den traditionsreichen Begriff Archiv zu lesen:



„Ein Archiv (lateinisch archium aus griechisch ἀρχείο(ν) archeío(n) ‚Regierungs-‘ oder ‚Amtsgebäude‘) ist eine Institution oder Organisationseinheit, in der Archivgut zeitlich unbegrenzt aufbewahrt, benutzbar gemacht und erhalten wird (Archivierung).“<sup>1</sup>

Das Archiv ist also der organisatorische Überbau zum Zwecke der Aufbewahrung historischer Informationen. Dem Archiv sind keine Verfahren zur Archivierung eigen. Die verwendeten Technologien sind für das Ziel frei wählbar. Vorgabe ist allerdings eine unbegrenzte Aufbewahrungszeit.

Die Informationstechnologie kann zur Archivierung benutzt werden, allerdings sind ihre Apparate und Methoden in beständigem Wandel. Daher ist die Frage berechtigt, ob die Informationstechnologie nachhaltige Lösungen in der digitalen Welt bereitstellen kann. Aus der Definition heraus können wir die Frage nicht mit Ja beantworten. Ganz im Gegenteil lässt die Definition eher den Schluss zu: Sie ist gänzlich ungeeignet. Es geht um Wandlung, um ständige Dynamik und permanente Veränderung.

Rechenanlagen wurden im Ursprung als Hilfsmittel von Mathematikern erfunden, um für Mathematiker komplexe Fragen durch Algorithmen schneller beantworten zu können. Im Zyklus der Abarbeitung der Algorithmen werden Informationen erfasst, verarbeitet und Ergebnisse ausgegeben. Nicht mehr und nicht weniger. Dies geschieht heute schneller als noch zu Zeiten des Z22 oder ENIAC, doch immer noch nach den gleichen Regeln. Jedoch hat die Informationstechnologie das Anwendungsfeld der Mathematik längst verlassen und widmet sich heute unzähligen Aufgaben des alltäglichen Lebens. Sie ist das Pendant zur immer größer werdenden Bedeutung von Wissen und Information in unserer Gesellschaft.

### *Marktmechanismen oder der Wert von Informationen*

Im Jahre 2015 werden nach einer Studie im IT-Markt Waren und Dienstleistungen im Werte von 1,274 Billionen Euro weltweit umgesetzt.<sup>2</sup> Wer wen in diesem Markt antreibt, ist nicht immer klar auszumachen. Seit 1964 der erste Personal Computer (PC) und 1990 das Internet mit großer Geschwindigkeit die Massen

1 Vgl. <http://de.wikipedia.org/wiki/Archiv> (Letzter Aufruf: 27.07.2013).

2 Vgl. IDATE: Umsatz im ITK-Bereich weltweit im Zeitraum von 2005 bis 2014. [www.statista.de](http://www.statista.de) (Letzter Aufruf: 15.08.2013).

erreichten, war es neben der Industrie auch die Gesellschaft selbst, die die Entwicklung vorantrieb. Neue Anwendungsfelder werden fast täglich für Produkte der IT gefunden. Mit Beginn des 21. Jahrhunderts fing die Gesellschaft an, sich ihr digitales Universum zu schaffen. Damit wuchs auch der Berg an digital gespeicherten Informationen, und ein Ende der Entwicklung scheint nicht in Sicht. Im Gegenteil: Eine IDC-Studie „The Digital Universe“ aus dem Jahre 2012 prognostiziert das Anwachsen des Informationsberges bis 2020 auf 40 Zettabyte! Das sind 57-mal mehr elektronische Bytes, als es Sandkörner auf der Erde gibt. 2020 wird jeder Mensch pro Jahr 8 Gigabyte an Daten erzeugen.<sup>3</sup> Dabei werden 75 % aller Informationen von Privatpersonen erzeugt und konsumiert.<sup>4</sup> Das Bemerkenswerte daran ist aber, dass ein großer Teil davon durch kommerzielle Unternehmen bewegt wird (Netzwerk, Webseiten, Provider). Und schon lange fußen auf dieser Tatsache stabile Geschäftsmodelle. Noch interessanter aber ist die folgende Berechnung: 2020 werden 33 % der 40 Zettabyte an Daten für Analysezwecke Verwendung finden können und bekommen damit einen wirtschaftlich relevanten Wert. Der Wert steckt in Social-Media-Daten, in Auswertungen wissenschaftlicher Studien, medizinischen Informationen in Korrelation mit soziologischen Daten usw. Die Grundlage dieses Informationswachstums bildet die Allgegenwart des Computers. Sie stecken in der Tasche jedes Menschen und im Gehäuse vieler Maschinen. Dies wird auch als Ubiquitous Computing bezeichnet. Mark Weiser hat diesen Begriff schon 1991 in seinem Aufsatz in der „Scientific American“ geprägt!<sup>5</sup> Zur zeitlichen Einordnung will ich daran erinnern, dass Weiser diesen Gedanken ein Jahr vor der Erfindung des World Wide Web durch Tim Berners-Lee schrieb. Berners-Lee veröffentlichte 1992 die erste Webpräsenz, nachdem er bereits in den Jahren zuvor die Grundlagen dafür entwickelt hatte. Ich gehe noch 30 Jahre zurück. Im Jahre 1960 formulierte der Psychologe J. C. R. Licklider in seinem Aufsatz „Man-Computer-Symbiosis“

- 3 Gantz, John/Reinsel, David: The Digital Universe in 2020. Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. IDC Digital Universe Study. 2012. [www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf](http://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf) (Letzter Aufruf: 15.08.2013).
- 4 Gantz, John F.: The Expanding Digital Universe. IDC White Paper, sponsored by EMC. 2007. [www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf](http://www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf) (Letzter Aufruf: 27.07.2013).
- 5 Weiser, Mark: The Computer for the 21st Century. In: Scientific American 265/3 (1991), S. 94-104.

eine Zukunft der engen Verbindung zwischen Mensch und Computer.<sup>6</sup> Zum damaligen Zeitpunkt stand die Stapelverarbeitung auf der Basis von Lochkarten am Großrechner im Zenit ihrer Entwicklung. Was wir also heute als selbstverständlich betrachten, ist schon vor über 50 Jahren gedacht worden.

Information ist der Rohstoff des 21. Jahrhunderts, und dieser Rohstoff liegt zunehmend nur noch digital vor. Hieraus resultiert der wirtschaftliche Wert, welcher die Entwicklung der IT-Industrie in den nächsten Jahren weiter prägen wird. Information behält aber nur dann ihren Wert, wenn sie vorhanden (gespeichert), interpretierbar (vollständig) und verfügbar (zugänglich) sein wird. All dies sind Kriterien zur Bewertung von Nachhaltigkeit im Informationssektor.

### *Eine sehr kurze Geschichte der Speichertechnologie*

Wie kann die IT-Industrie zur Nachhaltigkeit des Erhalts digitaler Informationen beitragen, wo sie doch beständig neue Technologien hervorbringt, die die „alte“ Technik obsolet macht?

Ein Blick zurück: Elektronische Informationen wurden schon immer gespeichert. Die immer schnellere Verarbeitungsgeschwindigkeit der Rechner verlangt nach immer schnelleren Speichermedien. Immer mehr Daten müssen in immer kürzeren Zeiteinheiten zur Ein- und Ausgabe bereitgestellt werden. So war 1956 die Zeit reif für das erste in Serie hergestellte elektronische Speichermedium, die IBM RAMAC 305. 500 Kilogramm brachte dieses Gerät auf die Waage. Sie speicherte auf 50 Magnetplatten mit 61 Zentimeter Durchmesser ganze 5 Megabyte an Daten.<sup>7</sup> Vier Jahre zuvor erblickte das erste Magnetband (IBM 726) das Licht der Welt. Beide Medien wurden aus der Recheneinheit direkt angesteuert. Es konnte in den damaligen Rechnern nur genau ein Programm abgearbeitet werden. So stand der Rechner nach dem Programmdurchlauf meistens längere Zeit still, da der neue Durchlauf erst neu codiert werden musste. Die neu entstandenen Informationen wurden zumeist ausgedruckt. Einen anderen Aspekt

6 Licklider, J.C.R.: Man-Computer-Symbiosis. In: IRE Transactions on Human Factors in Electronics HFE-1 (1960), S. 4-11. <http://groups.csail.mit.edu/medg/people/psz/Licklider.html> (Letzter Aufruf: 15.08.2013).

7 Gereke, Kurt: IBM System Storage-Kompendium. Die IBM Speichergeschichte von 1952 bis 2010. 2010. [www-03.ibm.com/systems/data/flash/de/resources/ibm\\_storage\\_compendium\\_de.pdf](http://www-03.ibm.com/systems/data/flash/de/resources/ibm_storage_compendium_de.pdf) (Letzter Aufruf: 15.08.2013), S. 14.

beschreibt Held: „Im Schnitt fielen diese Rechner alle paar Tage einmal aus, oft war eine Röhre defekt (im Jahresbericht der RWTH Aachen von 1974 wurde z. B. der mittlere Fehlerabstand sogar der 1966 beschafften modernen Anlage CD 6400 der Firma Control Data noch mit 74,2 Stunden angegeben).“<sup>8</sup>

Trotzdem sahen immer mehr Kunden aus Wissenschaft, Industrie und Verwaltung den großen Wert der schnellen Datenverarbeitung und forderten sichere Speichergeräte. So wurde in den 1960er Jahren für die Ausgabe der Informationen auf Magnetbändern die automatische Fehlerkorrektur (Automatic Error Capture & Correction) erfunden. In den späten 1980er und frühen 1990er Jahren wurde das Hierarchische Speichermanagement (HSM) eingeführt. Es realisierte die automatische Verlagerung von weniger häufig benutzten Daten auf langsamere und kostengünstigere Medien. Daraus entwickelte sich die heute übliche FAST-Technologie. Der Wettkampf der Speichertechnologien war in vollem Gange. Ende der 1980er Jahre fanden die ersten Backup-Programme Einzug in die Rechenzentren. Da für die Wirtschaft relevante Informationen erhoben wurden, forderten behördliche Regularien eine Speicherung der Informationen für mehrere Jahre. In der Gesundheitswirtschaft sind das bis zu 30 Jahre. Riesige Magnetbandbibliotheken füllten sich. Doch damit begann auch die Ära der aufwändigen Migrationsprojekte, denn der technische Fortschritt machte auch vor Magnetbandtechnologien nicht halt. Ich werde später noch näher darauf eingehen.

1974 war ein entscheidendes Jahr. Der ALTAIR 8080 kam für 400 Dollar auf den Markt.<sup>9</sup> Der erste Minicomputer wurde einer breiten Bevölkerung angeboten. „[...] die technischen und sozialen Komponenten der Personal Computerentwicklung kamen endlich zusammen.“<sup>10</sup> Der ALTAIR war mit einem Mikroprozessor 8080 der kleinen aufstrebenden Firma Intel bestückt und den Computersystemen von DEC, IBM und anderen etablierten Computerfirmen ebenbürtig. Nur eben erreichte er eine Bevölkerungsschicht, die bisher von der

8 Held, Wilhelm (Hrsg.): Vom Anfang des Informationszeitalter in Deutschland. Geschichte der Zusammenarbeit der Rechenzentren in Forschung und Lehre. Vom Betrieb der ersten Rechner bis zur heutigen Kommunikation und Informationsverarbeitung. Wissenschaftliche Schriften der WWU Münster, Reihe XIX, Band 1. [http://miami.uni-muenster.de/servlets/DerivateServlet/Derivate-5085/Buchblock\\_Held.pdf](http://miami.uni-muenster.de/servlets/DerivateServlet/Derivate-5085/Buchblock_Held.pdf) (Letzter Aufruf: 15.08.2013), S. 9.

9 Ceruzzi, Paul E.: Eine kleine Geschichte der EDV. Bonn 2003, S. 272.

10 Ebd., S. 273.

Entwicklung ausgeschlossen war. Und sie nahm dieses Angebot gierig an. Von nun an wurde die Innovation auch aus privaten Haushalten getrieben.

Wir wissen heute, was dieser Meilenstein bedeutete. Innovation, Massenfertigung, Preisdruck und wieder neue Innovation setzten die Entwicklung im Massenmarkt in Gang. Und dieser Massenmarkt erzeugte digital gespeicherte Informationen in ungeahntem Ausmaß.

Die Speichertechnologie folgte natürlich dieser Entwicklung. Die Winchester-Baureihe der IBM wurde die Vorlage für viele Erfinder, die sie verbessern wollten. Es dauerte aber noch sechs Jahre, bis 1980 Seagate die erste 5 ¼“-Festplatte (ST506) herausbrachte. Die ST506 wurde zum legendären Quasi-Standard und kostete einen Bruchteil der bisherigen Wechselp Plattenspeicher. Professionellen Rechenzentren war diese Technologie der billigen Massenproduktion nicht geheuer. So dauerte es noch einmal fast zehn Jahre, bis Seagates Technologie die Rechenzentren erreichte. Sie galt als wenig stabil, und Experten trauten ihr den harten Einsatz im Rechenzentrum nicht zu. Plattenausfälle hätten verheerende Wirkung auf den Rechenzentrumsbetrieb bedeutet. Der Kostendruck aber erzwang deren Einsatz. Um diese Diskrepanz aufzulösen, beschrieben 1984 Petterson, Gibson und Katz in einem Aufsatz fünf verschiedene RAID-Verfahren, um die Standfestigkeit (MTTF – Mean Time To Failure) der billigen Festplatten zu erhöhen. Im RAID bieten sich die Festplatten selbst Schutz vor Datenverlust.<sup>11</sup> Damit war eine neue Generation von ausfalltoleranten Festplattensystemen eingeläutet. Neue Marktteilnehmer verdrängten langsam die Marktmacht der IBM in diesem Teilgebiet der IT. Die Firma EMC<sup>2</sup> war erst zehn Jahre im Geschäft, als sie 1989 ihr erstes IBM-kompatibles Speichergerät mit völlig neuer Architektur vorstellte – die Symmetrix ICDA. Als leistungsfähige Multiprozessormaschine mit integriertem Cache (ICDA – integrated cached disc array) betrieb die Symmetrix fünfzig 5 ¼“-Festplatten von Seagate oder anderer Hersteller und sicherte diesen Verbund über RAID-Verfahren und Fehlerkorrektursoftware. Diese Multiprozessor-Arrays werden nun ausschließlich zum Zwecke der sicheren und schnellen Datenverarbeitung in Rechenzentren entwickelt. Bis heute haben sich viele Marktbegleiter dieser Entwicklung angeschlossen.

11 Petterson, David A./Gibson, Garth/Katz, Randy H.: A Case for Redundant Arrays of Inexpensive Disks RAID. In: In Proceedings of the 1988 ACM SIGMOD international conference on Management of data. 1988, S. 109-116. [www.eecs.berkeley.edu/Pubs/TechRpts/1987/CSD-87-391.pdf](http://www.eecs.berkeley.edu/Pubs/TechRpts/1987/CSD-87-391.pdf) (Letzter Aufruf: 15.08.2013).

EMC<sup>2</sup> läutete damit eine neue Generation von Festplattencomputern ein. An diese Arrays konnten eine Vielzahl von Computern unterschiedlicher Hersteller parallel betrieben werden. Verschiedene Zugriffsprotokolle (FICON für Großrechner, FiberChannel, SCSI und NFS) ermöglichten einen kostensparenden und performanten Betrieb. Möglich machte das die hohe Kompatibilität der Zugriffsprotokolle. Der gleichzeitige Betrieb von Großrechnern und Mikrocomputern war neu und innovativ, und es wurden damals sehr heftige philosophische Debatten geführt. Um den Rechenzentrumsbetrieb über mehrere Standorte ausfallsicher zu gestalten, wurden geräteinterne Übertragungsprotokolle entwickelt. Diese ermöglichten die Synchronisation der Datenbestände in Echtzeit über mehrere Kilometer Entfernung.

Spätestens ab 1990 entwickelte sich der Markt für externe Speichertechnologien parallel zum Großrechner-, Server- und PC-Markt als eigenständige Sparte der Computertechnologie weiter.

Dieser kurze historische Abriss erläutert, dass Innovation und Modernisierung in der Informationstechnologie nicht trotz, sondern wegen des Nachhaltigkeitsgebotes umgesetzt werden. Migration und Systemwechsel sind nicht Risiken, sondern Voraussetzung für eine Nachhaltigkeit.

### *Das Netzwerk ist der Computer*

Die Zeit zwischen 1970 und 1990 ist geprägt von einer unglaublichen Dynamik an Erfindungen im Personal-Computer-Bereich. Der Personal Computer auf der Basis von Mikroprozessoren setzte zum Siegeszug an. Gleichzeitig wurde am Xerox-PARC das Ethernet-Protokoll für Netzwerke erfunden. Mit VisiCalc kam der erste Tabellenkalkulator für Minicomputer (neben Textprogrammen) auf den Markt und machte den Büroarbeitsplatz attraktiver. Nicht zuletzt waren es aber Computerspiele, die die weite Verbreitung der PCs beflügelten. Sie waren sehr anspruchsvoll in der Programmierung, zeigten die technologischen Grenzen der Rechner sehr schnell auf und beflügelten deren Weiterentwicklung. Zwei weitere Erfindungen des Xerox-PARC sollten zum Erfolg anderer Firmen beitragen: die Computermaus als Eingabegerät und die Windows-Oberfläche. Beide wurden von Apple zur Perfektion weiterentwickelt. Sehr spät bemerkte die IBM diese große Dynamik der Entwicklung. Trotzdem schaffte sie es durch ihre Marktmacht, den IBM-PC auf den Markt zu bringen. Der IBM-PC wurde zum Standard erklärt und bald von vielen Firmen in kompatiblen Varianten

nachgebaut. Das ist ein großer Verdienst der IBM. So kam wieder Orientierung in den Markt. Letztlich profitierten aber andere davon. Compaq schaffte es in den Verkaufszahlen noch vor die IBM.

Sun – ein weiteres junges Unternehmen – etablierte sich mit seinen Rechnern im universitären Bereich und baute neben DEC die leistungsfähigsten Workstations. Die Rechner von SUN und DEC wurden vom neuen Betriebssystem Unix angetrieben. Beide Firmen (DEC und Sun) sind heute schon wieder vom Markt verschwunden. Großrechner muteten in dieser Zeit wie Dinosaurier aus einer längst vergangenen Zeit an. Doch an diesem Beispiel zeigt sich die Durchsetzungskraft standardisierter Ingenieurskunst. Der Mainframe hat viele der damals dynamischen Firmen überlebt. PCs und Workstations (leistungsfähige Maschinen mit UNIX-Betriebssystem) wurden über Ethernet vernetzt und waren in der Summe leistungsfähiger als ein Zentralcomputer mit dummen Terminals. Größter Gewinner in dieser Zeit aber war Microsoft mit seinem Betriebssystem MS-DOS, welches in riesigen Stückzahlen mit allen IBM-kompatiblen PCs vermarktet wurde.

PCs und Workstations arbeiteten zunächst mit Floppy-Disks. Das waren dünne magnetische Scheiben im 8“- oder 5 ¼“-Format. Später fanden die ersten Festplatten im 5 ¼“-Format im PC-Gehäuse Platz. Das Betriebssystem wurde aber immer noch auf Band oder Floppy mitgeliefert und musste zum Start des Rechners jedes Mal neu eingelesen werden. Sehr zukunftsweisende Strategien, Information für längere Zeit abzuspeichern, gab es zur damaligen Zeit nicht. Es scheint sehr wahrscheinlich, dass es in dieser dynamischen Phase der Entwicklung oft zu Datenverlust gekommen ist.

### *Das Zeitalter des Internets*

Mit der Veröffentlichung des ersten Webbrowsers mit Namen Mosaik im Jahr 1993 wurde nun das digitale Zeitalter eingeläutet. Der Webbrowser war für alle nicht so geübten Nutzer des PCs leicht zu bedienen. Er war das Tor ins World Wide Web. Das World Wide Web wurde 1990 von Tim Berners-Lee am CERN beschrieben und aufgebaut, um die Rechner unterschiedlicher Plattformen aller am Projekt arbeitenden Wissenschaftler zum Zwecke des Informationsaustausches zu koppeln. Sehr entscheidend war, dass er das zugrundeliegende Protokoll TCP/IP, die Seitenbeschreibung HTML, das Transferprotokoll http und vor allem die URL-Beschreibung zur freien Nutzung zur Verfügung stellte. Das

schaftte die Akzeptanz vieler Entwickler weltweit. Die Domainnamen der Nutzer wurden dann von einem Gremium der W3C vergeben. Freie Nutzung von Protokollen und gute Organisation waren die Schlüssel für rasche Verbreitung des World Wide Web.

Die weitere rasante Entwicklung des Internets kennen wir. Jeder verkaufte Rechner konnte sofort an das Internet angeschlossen werden. TCP/IP wurde zum Standardprotokoll in jedem Betriebssystem. Sehr wahrscheinlich spielt dabei auch der freie Zugang zum Quellcode eine Rolle.

### *Standardisierung und Nachhaltigkeit*

Mit dem World Wide Web ist plötzlich ein Massenmarkt entstanden, der mit herkömmlichen, teuren Computern nicht zu bedienen war. Die vielfältigen leistungsstarken Spezialrechner waren ungeeignet für das Wohnzimmer der privaten Nutzer. Doch um preiswertere Technik anzubieten, war eine Standardisierung der Technologie notwendig. Schnittstellen, Protokolle, Datenformate und letztlich Betriebssysteme wurden in immer engeren Grenzen beschrieben. So konnte im Wettbewerb um die Vorherrschaft in der Versorgung des Marktes nur der bestehen, dessen Produkte in einem möglichst breiten Anwendungsspektrum reibungslos funktionierten.

Diese Entwicklung hatte selbstverständlich auch eine große Wirkung auf die Speicherindustrie. Es wurden immer neue Geräteklassen entworfen, die einerseits mit preiswerter Hardware ausgerüstet waren, doch andererseits die hohen Sicherheitsanforderungen der professionellen Anwender bedienten. Diese Eigenschaften brachte die HDD (Seagate ST506) mit. Sie setzte sich mit ihrem Formfaktor 5 ¼“ durch und schuf auch hier einen Markt für Nachahmer. Ständig wurde diese Technologie weiterentwickelt und ist bis heute nach ähnlichem Design das meist eingesetzte Speichermedium. Neue Übertragungsprotokolle wurden für die HDD entworfen (SCSI, FC, ATA, SATA, USB), Modelle in kleinerem Formfaktor (5 ¼“, 3 ½“, 1 ½“) wurden gebaut, durch feinere Abtastverfahren konnte die Speicherdichte erhöht werden (1986: 5 ¼“ mit 5 MB, 2013: 3 ½“ mit 4 TB). Diese Technologie war also eine geniale Erfindung, deren Basistechnologie noch immer nicht ausgeschöpft scheint. Mit kleinen Modifikationen kann sie immer noch den Markterfordernissen angepasst werden. Durch die Langlebigkeit dieser gleichen Technologie war es nun auch für viele Firmen wirtschaftlich, auf deren Basis Geräte zu entwickeln, die den speziellen Bedarf



der Kunden decken konnte. Bis heute ist die HDD (sie feierte 2006 ihr 50-jähriges Jubiläum) die Standardtechnologie sowohl im Home-PC als auch im Hochleistungsrechenzentrum.

Trotzdem, konkurrenzlos war diese Technologie nicht. Widerstand regte sich vor allem in Spezialthemen. Staatliche Regularien zwangen die Kunden zu Maßnahmen, die es ihnen abverlangten, den Nachweis der Unveränderlichkeit von elektronisch gespeicherten Daten zu erbringen. So wurden besondere 5 ¼“-Scheiben mit einer Spezialbeschichtung entwickelt, die ein Überschreiben oder gar Löschen der Informationen verhindern sollten. Die WORM-Technologie (Write Once Read Many) kam in Gestalt der ersten Optical Disc auf den Markt. Mit Hilfe von speziellen Laserstrahlen wurden die Bits und Bytes auf die Scheibe gebrannt. Somit waren diese nicht mehr zu verändern oder löschar. Große Roboter wurden entwickelt, die mehrere hundert WORM-Disks organisieren konnten. Sie prägten die ersten elektronischen Archive. Bald jedoch entdeckte die Unterhaltungsindustrie die Vorzüge der metallisch glänzenden Plastescheibe. Musik und Filme konnten in bekannten Einheiten auf diesen Scheiben verkauft werden. Sie fand Einzug in die Heimelektronik. Der Consumer-Markt ist aber anderen Regeln unterworfen als der IT-Markt. Im Kampf um Marktherrschaft spielt Nachhaltigkeit keine Rolle. Wichtig ist ein schickes Produkt, was die Massen anspricht. Auch sind Technologiewechsel sehr willkommen. Der Kunde soll in ein paar Jahren neue Geräte kaufen wollen. Das führte dazu, dass die WORM-Technik in sehr schneller Folge nicht-kompatible Datenformate für unterschiedlichste Einsatzzwecke hervorbrachte (CD, DVD, Blue-Ray). Professionelle Rechenzentren sahen sich somit in kürzester Zeit gezwungen, teure Migrationen von alte auf neue Geräte durchzuführen. Auch fiel die Forschung nach größerer Speicherdichte hinter der Magnetplatte weit zurück. Das war nicht akzeptabel. Somit kann diese Technik für unseren Zweck der Nachhaltigkeit als gescheitert angesehen werden.

Dessenungeachtet führte das Magnetband ein sehr stilles, aber doch sehr erfolgreiches Dasein im Rechenzentrum. Die Entwicklung von Kapazität, Lese- und Schreibgeschwindigkeit und Handhabung des Magnetbandes ist ausgesprochen gut geeignet, um Sicherungskopien der Festplatten auf dieses Medium zu schreiben. Vor allem aber war es möglich, das Band aus dem direkten Zugriff der Rechenanlage zu entfernen und in einem sicheren Raum zu verschließen. Dieser Vorgang wurde Auslagern genannt. Das mentale Gefühl der Sicherheit war dadurch leicht zu vermitteln. Das Magnetband dient bis heute als Sicherungsmedium in fast allen Rechenzentren. Leider ist es aber gerade das Entfernen aus

dem direkten Zugriff, was das Band für den Einsatz langfristiger Archive anfällig machte. Wenn die Magnetisierung gestört wird, sind Daten nicht mehr lesbar. Noch tragischer ist, dass das keiner merkt. Erst nach Jahren ist der Verlust spürbar. Abhilfe schaffen große Robotersysteme, die beständig mit der Bewegung, Umwälzung und Migration der Datenbestände auf den Magnetbändern beschäftigt sind. Das treibt die operativen Kosten in die Höhe, und die finanziellen Mehrwerte der Magnetbandtechnologie fallen nicht mehr ganz so günstig aus. Bei Technologiewechsel sind die Migrationskosten zumeist sehr hoch. Der Vorteil des geringen Preises wird inzwischen durch die massenhafte Fertigung von der HDD eingeholt.

In beide Einsatzbereiche, Archiv und Backup, dringt nun auch die HDD vor. Ihre Vorteile – Sicherheit, Geschwindigkeit, leichte Handhabbarkeit und vor allem Kompatibilität zu älteren Modellen – kommen hier zum Tragen. Gerade Letzteres führt zu signifikanten Kostenersparnissen im operativen Rechenzentrumsbetrieb. Migrationsaufwendungen fallen praktisch nicht mehr an. Im Rückblick zur Betrachtung der Werte, die in den Informationen stecken, ist eines in HDD-basierten Technologien von unschlagbarer Bedeutung: Gemeint ist der permanente Zugriff auf die Daten.

Die Geburtsstunde des Objektspeichers auf der Basis der HDD kann in den Monat April 2002 datiert werden. Die EMC CENTERA war das erste auf HDD-Technologie aufgebaute, objektorientierte Archivspeichergerät auf dem Markt.<sup>12</sup> Dieses Speichergerät konnte nur durch das Standardprotokoll IP angesprochen werden. Das war ein großer Vorzug für die Implementierung in bestehende Archivierungssoftware. Für die Sicherheit der als Objekt (Bild, Film, Dokument) behandelten Daten sorgt eine API (Programmierschnittstelle). Diese war und ist für alle Entwickler weltweit frei zugänglich. Hunderte Softwarefirmen nutzen diese Schnittstelle in ihren Produkten. Über Hash-Algorithmen werden den digitalen Objekten IDs (Fingerprint oder auch Signatur) mitgegeben. Über diese Signatur können die Objekte im Archiv wieder aufgefunden werden. Datenduplizierung und Migrationsfreiheit sind weitere Vorzüge dieser Technologie. Veränderungen an den Objekten lässt diese Technologie nicht zu. Damit kam man Kunden entgegen, die einen sehr hohen Schutzbedarf für ihre Daten hatten.

12 EMC Presse Release: EMC Launches Centera, Ushers in New Era of Content-Addressed Storage. 2002. [www.emc.com/about/news/press/us/2002/20020429-1254.htm](http://www.emc.com/about/news/press/us/2002/20020429-1254.htm) (Letzter Aufruf: 15.08.2013).

Unter anderem sind das auch Kultureinrichtungen, die darauf ihre einmaligen digitalisierten Kulturschätze abspeichern. Zahlreiche Nachahmungen der Objektspeicher entfachten einen großen Wettbewerb, der bis heute anhält.

Das Datenwachstum der Kunden führte zu weiteren Nutzungsszenarien der HDD.

Firmen wie Google, eBay, Amazon, Apple, Rackspace und andere entwickelten sich zu riesigen Informationsquellen und bieten ihren Kunden (vor allem im privaten Bereich) den Service an, Daten in ihrem Speichernetzwerk zu speichern. Informationsbestände, mehrere Petabyte groß, lagern sie in ihrem Speichernetzwerk. Neue Organisationskonzepte sind entwickelt wurden: Hadoop, MapReduce und andere. Es ist undenkbar, dass bei Technologiewechsel eine Migration dieser Datenbestände vorgenommen werden kann. Diese würde viele Jahre in Anspruch nehmen. Die Lösung sind neue migrationsfreie Scale-Out-Technologien. Auf Scale-Out-Technologie aufgebaute Geräte gibt es nun auch für traditionelle Rechenzentren. Sie können linear im laufenden Betrieb erweitert werden und kennen keinen Verwaltungsaufwand mehr. Migration findet nicht mehr statt. Schrittweise kann alte Technik durch neue ersetzt werden. Bisherige Verwaltungsgrenzen (Größe von Filesystemen, Zugriffsbandbreite) sind praktisch aufgehoben. Ihr Betriebssystem (zum Beispiel OneFS, ATMOS) sorgt für eine höhere Abstraktion zwischen den Betriebssystemen und dem Speichermedium. Somit werden die Verwaltungstools zur Organisation der Daten von lästigen Aufgaben befreit. Diese Speichergeräte sind selbst zum Hochleistungsrechner mutiert. Sie sind nun Teil eines Rechnernetzwerkes mit speziellen Aufgaben – der Datenorganisation.

Alle heute auf dem Markt befindlichen Speicherklassen (SAN, NAS, CAS, COS) können als Archiv durch bekannte Funktionen (WORM oder Clip-ID) eingesetzt werden.

Der Wert der Information für Wirtschaft und Gesellschaft ist nur gegeben, wenn auch der Zugang zu den Informationen gewährleistet ist. Die beschriebenen neuen Geräteklassen ermöglichen sehr einfach und leicht den Zugang zu den digitalen Objekten. Praktisch können diese Speicherklassen überall im Netzwerk platziert werden. Der Zugang kann dann über das Web oder spezielle Anwendungsserver ermöglicht werden.

Nachhaltig an diesen neuen Speicherkonzepten auf HDD-Basis sind folgende wesentliche Faktoren:

- Unbedingte Standardisierung der eingesetzten Grundlagentechnik (Schnittstellen)
- Großvolumige Produktion in hoher Qualität (Kosten)
- Hohe Verarbeitungsgeschwindigkeit (Performance)
- Breite Integrationsfähigkeit (Standardisierung)

*Noch ein paar Worte zur Software*

Ich sparte Betrachtungen zur Software nicht zufällig aus. Software unterliegt eigenen Entwicklungsregeln. Es ist mir dennoch wichtig, einen Aspekt in der jüngsten Entwicklung der Software besonders hervorzuheben.

Durch die Software werden die nachhaltig gespeicherten Daten interpretiert. Nur die Software bringt einen Sinn in die Abfolge der gespeicherten Nullen und Einsen.

Heute ist die Software neben Netzwerk, Computer und Speicher die vierte selbstständige Sparte der IT. Das ist sie erst seit 1968/69, als die US-Regierung die IBM zwang, auf ihren Rechnungen Hardware und Software getrennt auszuweisen. In den darauffolgenden Jahren gründeten sich Firmen, die einzig damit beschäftigt waren, Softwareprogramme zu schreiben. Wiederum die erfolgreichsten (z. B. Microsoft, Oracle oder SAP) waren Unternehmen, die es schafften, ihre Produkte als Quasi-Standard zu etablieren. Sie schufen das Fundament für eine Vielzahl an Individualsoftware. Der wachsende Druck zur Zusammenarbeit zwingt Unternehmen, Daten auszutauschen oder zusammenzuführen. Das ist mit Individualsoftware oft sehr problematisch. Standardsoftware hingegen bringt die Migrationstools meistens gleich mit.

Junge Unternehmen wie Google oder Amazon fußen auf hochstandardisierten Softwarekonzepten. Gegen wenig Geld und den freiwilligen Verzicht auf Individualität kann der Kunde Software zeitweise mieten. Dies stellt im Wettbewerb die etablierten, sehr komplexen Rechenzentren vor eine große Herausforderung. Das Problem sind deren hohe operationale Kosten, die durch das Management ihrer komplexen Systemlandschaften entstehen. IDC ermittelte, dass bis zu 75 % der im Rechenzentrum eingesetzten Mittel in das Management der komplexen Systemlandschaften fließen. Ein großer Teil davon entfällt auf die Wartung komplexer Softwarelandschaften.

Einen Ausweg bieten hier vor allem Virtualisierungstechnologien. Diese können heute schon als etabliert bezeichnet werden. Marktteilnehmer sind vor

allem VMWare, Microsoft mit Hiper-V und OpenStack. Eine weitere Stufe der Abstraktion der Systemlandschaft wird durch die Entwicklung zum Software-defined data center (SDDC) erreicht. Das SDDC wird in den nächsten Jahren die Rechenzentren revolutionieren. Die Abhängigkeit zwischen der Individualsoftware und der eingesetzten IT-Infrastruktur wird noch weiter aufgehoben. Die Infrastruktur muss nur leistungsfähig genug sein, die geforderten Transaktionen und Datenbewegungen in größtmöglicher Geschwindigkeit abzuwickeln. Eine direkte Abhängigkeit existiert nicht mehr.

Für den Erhalt der digitalen Schätze, die es zu bewahren gilt, ist das eine sehr hilfreiche Entwicklung. Nun können auch die unzähligen Interpretationsmöglichkeiten der Formate auf ein möglichst geringes Maß standardisiert werden. Für den Zugang ist es wichtig, dass alle notwendigen Informationen vorhanden sind: Metadaten, Formatinformationen u. a. Für die Entwicklung geeigneter Archivierungsstrategien bedeutet dies mehr Freiheit in der Entwicklung notwendiger Konzepte.

### *Schlusswort*

Der kurze Streifzug durch einige Aspekte der jungen Informationstechnologie zeigt, wie eng diese mit wirtschaftlichen und gesellschaftlichen Entwicklungen verflochten ist. Es ist deutlich geworden, dass gerade Innovation eine Voraussetzung für Nachhaltigkeit und damit langfristig verfügbare Informationen ist. Nur der permanente Wechsel von Technologie sichert die erarbeiteten Inhalte. Diese Paradoxie ist grundlegend für die Informationstechnologie. Innovative Ideen können sich aber nur entfalten, wenn die gesellschaftliche Akzeptanz und der wirtschaftliche Nutzen eine bestimmte Reife erreicht haben. Es war nicht alles immer so geplant gewesen. Manchmal waren es auch Zufälle, die eine Idee in ebendiesem Nährboden aufgehen ließ. Doch braucht es auch die sehr planmäßige Ingenieurskunst, um eine Idee langfristig in der weiteren Entwicklung zu verankern. Heute haben historische Informationen einen wirtschaftlichen Wert erlangt, der die Entwicklungen von Verfahren und Lösungen für deren Erhalt attraktiv werden lassen. Große Technologiefirmen wiederum werden es schaffen, daraus nachhaltige Produkte zu formen.

So bin ich der Überzeugung, dass es ein guter Zeitpunkt ist, über nachhaltigen Schutz von digitalen Informationen zu sprechen. Aber zu erwarten, Speichertechnologie allein würde die Lösung bringen, ist falsch. Alle Komponenten der IT sind maßgeblich an Lösungen beteiligt.

# Die Gaming-Community als Pionier der digitalen Bewahrung

ANDREAS LANGE

Games sind die ersten digital geborenen Kulturgüter, die massenhaft und global produziert wurden. Wissen und Praktiken, die sich in der Games-Kultur herausgebildet haben, sind auch darüber hinaus für die Bewahrung digitaler Kulturgüter relevant. Als komplexe digitale Artefakte stellen Computerspiele höchste Anforderungen an Archivare, sodass ich die These aufstelle: Wenn es uns gelingt, Computerspiele zu bewahren, können wir alles Digitale bewahren, da alle digitalen Werke die gleiche technische Grundlage besitzen.

Digitale Werke wie Games sind von der Kombination aus Betriebssystem und Hardware in hohem Maße abhängig. Ohne deren Hilfe wären wir nicht in der Lage, den Nullen und Einsen im Bitstrom einen Sinn zu geben. Weil diese Nullen und Einsen nur eine begrenzte Lebensdauer haben, ist es eine der wichtigsten Aufgaben, sie unabhängig von der ursprünglichen physischen Systemplattform zu bewahren. Dies leisten im Kern Emulatoren, indem sie die ursprüngliche physische Hardware auf anderen Systemen imitieren und durch Virtualisierung auch dann noch anwendbar halten, wenn kein funktionsfähiger Computer eines bestimmten Systems mehr existiert.

Die Strategie der Migration – der Umarbeitung der Software auf ein neues System – ist im Bereich komplexer digitaler Artefakte nicht praktikabel. Zwar dient auch Migration dazu, ein Werk vom Ursprungssystem unabhängig zu machen, indem es so verändert wird, dass es auf aktuellen Systemen zugänglich ist. Allerdings ist diese Methode lediglich bei technisch einfachen Inhalten wie Texten oder Grafiken praktikabel, da der Migrationsaufwand mit zunehmender Komplexität der Werke steigt. Auch muss die Migration stets neu durchgeführt werden, wenn sich die aktuelle Systemplattform ändert. Da bei der Emulation jedoch das Werk selbst

unverändert bleibt und die ständige Anpassung der Werke an neue Plattformen wegfällt, bietet sie unbestreitbare Vorteile für die Langzeitbewahrung. Die Komplexität der Computerspiele als interaktive, zeit-, bild- und tonkritische Anwendungen macht ihre Bewahrung zur Nagelprobe für das Gelingen digitaler Bewahrung.

### *Die Anfänge der digitalen Bewahrung*

Es ist keine Besonderheit der Games-Kultur, dass die erste Initiative zur Bewahrung von Kulturgütern von den Fans ausging: So wie die Cinémathèque in Paris als eines der ersten Archive für die noch junge Filmkultur 1935 von Film-enthusiasten um Henri Langlois gegründet wurde, waren es auch die Fans früherer Computerspiele, die den kulturellen Wert der Spiele erkannten und erste Ansätze entwickelten, um diese Kultur zu bewahren. Gaming-Fans waren allerdings die ersten Liebhaber, die sich an diese Aufgabe unter den Bedingungen des Internets machten. Mehr noch: Sie gehörten zu den ersten Internet-Communities überhaupt, die das World Wide Web Mitte der 1990er Jahre mit ihrem Hobby besetzten und damit begannen, die Geschichte ihres geliebten Kulturguts kollaborativ aufzuschreiben. Als Anfangspunkt dieser Bewegung lässt sich der 07. Dezember 1992 festhalten: Mit `alt.games.video.classic` wurde die erste Newsgroup gegründet, die sich dem Austausch von Nachrichten und Diskussionen ausschließlich über klassische Videospiele widmete.

Diese erste Phase lässt sich mit einer mündlichen Geschichtsüberlieferung vergleichen: Vieles wurde eher dem Hörensagen nach festgehalten – Dinge, die man auf anderen Seiten gelesen hatte, wurden zusammenkopiert oder verlinkt. Genauigkeit wurde zwar angestrebt, systematisch und wissenschaftlich fundiert waren die Aktivitäten aber nicht. Dennoch: Parallel zu dieser ersten Geschichtsschreibung entstanden bereits umfangreiche Archive und Metadatenansammlungen, die ebenfalls von Fans zusammengetragen wurden.<sup>1</sup>

### *Bewahrung im Wettlauf mit der Technik*

Bedingt durch die schnelle technische Entwicklung der Abspieplattformen sind Spiele schon nach kurzer Zeit obsolet und nur noch mit Glück gebraucht zu be-

1 Siehe etwa das Projekt Home of the Underdogs (online 1998-2009, [http://de.wikipedia.org/wiki/Home\\_of\\_the\\_Underdogs](http://de.wikipedia.org/wiki/Home_of_the_Underdogs), letzter Aufruf: 15.08.2013) oder [www.mobygames.com](http://www.mobygames.com) (ab 1999).

kommen. Insofern waren die Fans der alten Spiele darauf angewiesen, eigene Bewahrungsansätze zu realisieren. Zwei Aspekte dieses Prozesses verdienen es, hervorgehoben zu werden:

- Erstens handelte es sich bei diesen Fan-Communitys von Anfang an um globale Communitys. Dies nicht nur in Bezug auf die technische Kommunikationsgrundlage Internet, die keine nationalen Grenzen kannte. Die globale Ausdehnung der Fangemeinschaft war dem globalen Charakter der Games selbst geschuldet, die von Anfang an für einen möglichst weltweiten Markt produziert und verkauft wurden. Wie sich dieser Unterschied auf die zukünftigen organisatorischen Strukturen der Computerspielebahrung auswirken wird, muss sich noch zeigen, sind doch Bewahrungsinitiativen traditionell oft in einen nationalen Kontext eingebunden.
- Zweitens war es die Community der Gamer selbst, die als Erste praktikable Lösungen entwickelt hat, um diese zu bewahren. Über das Internet bildeten sich offene, kollaborative, globale und üblicherweise nicht-kommerzielle Projekte, die heute auch von etablierten Gedächtnisorganisationen zur Kenntnis genommen werden. Für die Community lag es aufgrund der Komplexität der digitalen Spiele nahe, die Spiele nicht einzeln auf neue Plattformen zu adaptieren, sondern die obsoleten Plattformen selbst in Form von Emulatoren zu bewahren. Die Retro-Gamer waren insofern die ersten Anwender, die das bereits bekannte Modell Emulator nicht als Programmierumgebung für neue Hardware einsetzten, sondern als Werkzeug der Bewahrung.

Seitdem sind Millionen von Stunden an hochspezialisiertem Wissen in die Entwicklung der Emulatoren eingeflossen, die unter institutionellen Rahmenbedingungen womöglich hunderte Millionen gekostet hätten. Eines der ältesten und bis heute wichtigsten Emulatorenprojekte ist M.A.M.E., der Multiple Arcade Machine Emulator. Seinen Ursprung hat er in der Emulation historischer Spieleautomaten. Seit der Gründung 1996 sind mindestens 7304 Spieleautomaten (ROM sets) und 3910 einzelne Games für den Emulator adaptiert worden.<sup>2</sup> Auf das Projekt setzen weitere Entwicklungen auf, darunter MESS (Multi Emulator Super System), das seit 1998 entwickelt wird und 632 unterschiedliche Systeme

2 MAME wird auch heute noch weiterentwickelt, doch die letzten Zahlen über emulierte Automatenhardware stammen aus dem Jahr 2008.



emuliert, von denen das älteste aus dem Jahr 1948 stammt.<sup>3</sup> M.A.M.E. war von Anfang an ein offenes, nicht-kommerzielles Projekt, das ausschließlich von der Community getragen wurde.

*Probleme der digitalen Bewahrung: bitrot und Kopierschutz*

Unabhängig von der jeweils eingeschlagenen Bewahrungsstrategie ist die Frage relevant, wie lange ein Medium Informationen speichern kann. Der je nach Speichermedium einsetzende Datenverlust wird bitrot genannt. Für das Computerspielmuseum bedeutet dies einen regelmäßigen Verlust an Bits – speziell bei den magnetischen Datenträgern wie Compact Cassetten, Floppy-Discs oder Disketten, auf denen viele der mehr als 20 000 Spiele seiner Sammlung gespeichert sind.

Eine Stichprobe bei Spielen für den C64-Computer ergab, dass bereits ca. 10 % aller Spiele aufgrund nachlassender magnetischer Ladung so weit verrotten sind, dass sie nicht mehr gespielt werden können. Um digitale Sammlungen vor dem Verschwinden zu retten, müssen die Werke von ihren originalen auf stabilere Datenträger überführt werden. Erst mit dem virtuellen Abbild des Werkes, dem Datenträger-Image, kann das Werk von einem Emulator verarbeitet werden.

Da Computerspiele wie viele andere digitale Werke jedoch mit Kopierschutz (DRM) ausgestattet sind, der genau diesen Transfer verhindern soll, sind wir bei diesem Prozess mit weiteren technischen und rechtlichen Herausforderungen konfrontiert. Rein technisch stehen uns zwei alternative Strategien zur Verfügung:

- Erstens können die DRM-Maßnahmen gewaltsam außer Funktion gesetzt werden, gemeinhin auch als Cracken bezeichnet. Da dabei der Programmcode des Werkes verändert wird, ist diese Methode – unabhängig von der rechtlichen Problematik – keine befriedigende Option.
- Zweitens kann der Kopierschutzmechanismus intakt gelassen werden und wird einfach mitkopiert. In Anbetracht hunderter verschiedener DRM-Mechanismen stellt dieser Ansatz eine immense technische Herausforderung dar, die nur von wenigen Werkzeugen wenigstens teilweise beherrscht wird, etwa dem Kryoflux Controller.<sup>4</sup>

3 Stand vom 21.02.2012.

4 Der Kryoflux Controller ist eine Entwicklung, die ihre Wurzeln in der Software Preservation Society hat, die ursprünglich als Zusammenschluss von Gamern zum Erhalt von Spielen für die Commodore-Amiga-Plattform gegründet wurde.

### *Rechtliche Hürden bei Games*

Unabhängig von der Transfermethode sind rechtliche Fragen zu bedenken: Zwar gestattet uns das Urheberrecht in Deutschland, eine Sicherungskopie anzufertigen, doch dürfen wir technische Schutzmaßnahmen dabei nicht umgehen. Das Gesetz bietet nur den Ausweg, die Rechteinhaber um Erlaubnis zu fragen. Im Bereich der Computerspiele bleibt das größtenteils nur eine theoretische Option. Wir haben es mit einem großen Anteil an verwaisten Werken zu tun, bei denen es entweder nicht möglich ist, den Rechteinhaber herauszufinden, oder die Rechteinhaber selbst gar nicht wissen, welche Rechte sie tatsächlich besitzen. Zentrale Register, über die man Auskunft einholen könnte, haben sich im jungen Games-Markt noch nicht gebildet. Viele Firmen sind nach kurzer Lebensdauer wieder vom Markt verschwunden, die von ihnen gehaltenen Rechte wurden aufgesplittet und weiterverkauft.

Erschwerend kommt hinzu, dass Games ihrer komplexen Natur gemäß aus Werkteilen unterschiedlicher Rechteherkunft bestehen. Es ist üblich, dass man es bei einem Spiel mit einer Vielzahl an Rechteinhabern zu tun hat. Die zweite Möglichkeit, zu warten, bis das Urheberrecht abgelaufen ist, ist bei der gegenwärtigen Frist von 70 Jahren keine Lösung, da sich nach diesem Zeitraum kein funktionierendes Werk mehr auf einem alten Datenträger befinden würde.

### *Probleme und Grenzen der Emulatorenstrategie*

Jedes Werk, auch das digitale, benötigt Schnittstellen, über das es von Menschen wahrgenommen und im Fall von interaktiven Werken gesteuert werden kann. Diese Ein- und Ausgabeschnittstellen sind kein Bestandteil des Codes an sich, sondern sind seine Peripherie. Ein klassischer Videospieleautomat wie etwa Tempest (Atari 1980) besitzt einen Trackball als Eingabe- und einen Vektormonitor als Ausgabeschnittstelle. Führt man den originalen Programmcode über einen Emulator auf einem aktuellen Desktoprechner aus, fehlen beide historische Schnittstellen. Da beide jedoch wesentlich zum originalen Spielerlebnis beitragen, muss diese „historische Aufführungspraxis“ bedacht werden. Eine vom Computerspielmuseum praktizierte Lösung für historische Videospieleautomaten besteht darin, einen aktuellen Computer mitsamt Emulator in die alten Gehäuse mit den originalen Schnittstellen einzubauen, sodass das Look-and-feel des ursprünglichen Werkes weitgehend beibehalten wird.

Ein weiteres Problem besteht in der Notwendigkeit, die Emulatoren selbst immer wieder an die aktuellen Plattformen anzupassen. Immer, wenn ein neues Betriebssystem oder eine neue Hardwarearchitektur etabliert wird, müssen auch die Emulatoren auf diese angepasst werden, was aufwändige Neuprogrammierungen notwendig machen kann. Bisher haben die Programmierer von Emulatoren freiwillig und ohne Budget solche Portierungen – etwa von DOS auf Windows 95, von Windows 95 auf Windows XP – vorgenommen. Sie haben eine mehr oder weniger persönliche Beziehung zu den Spielen und Originalplattformen, jedoch können wir nicht davon ausgehen, dass die immer wieder notwendig werdenden Anpassungen auch mittel- bis langfristig auf diese Weise geleistet werden können.

Abhilfe könnten virtuelle Maschinen bringen, die speziell zu Bewahrungszwecken geschaffen werden. Bereits im Jahr 2000 hat Jeff Rothenberg eine solche Schnittstelle für Emulatoren vorgeschlagen. Sie würde die notwendigen Anpassungsprozesse auf ein Minimum reduzieren.<sup>5</sup> Seine Idee einer Emulation Virtual Machine (EVM) führt eine Zwischenebene ein: eine Schnittstelle, die als virtuelle Maschine zwischen dem aktuellen Betriebssystem und den Emulatoren implementiert wird und auf die die Emulatoren einmalig portiert werden müssen. Der Vorteil besteht darin, dass bei einem Wechsel der aktuellen Kombination aus Hardware und Betriebssystem lediglich die virtuelle Maschine an das neue System angepasst werden muss, nicht der jeweilige Emulator. Da wir es mittlerweile mit dutzenden von obsoleten Hardwareplattformen und mindestens ebenso vielen verschiedenen Emulatoren zu tun haben, ist der Nutzen leicht erkennbar. Ohne sie müssten alle Emulatoren einzeln angepasst werden, sobald das aktuelle Computersystem wechselt. Allerdings stellt die Programmierung einer solchen virtuellen Maschine technisch eine so große Herausforderung dar, dass mit einsatzfähigen Lösungen zumindest zeitnah nicht zu rechnen ist.

### *Gaming-Community und traditionelle Gedächtnisorganisationen*

Während die Gamer-Community von Anfang an auf Emulation statt auf Migration setzte, erwies sich die Migration für traditionelle Gedächtnisorganisationen zunächst als eine praktikable Methode. Doch wurden einerseits die zu bewah-

<sup>5</sup> Rothenberg, Jeff: An Experiment in Using Emulation to Preserve Digital Publications. Den Haag 2000.

renden Artefakte immer komplexer, andererseits wuchs mit jedem abermaligen Migrationsschritt die Bereitschaft, sich alternativen Methoden wie der Emulation zu öffnen.

Das EU-Forschungsprojekt KEEP (Keeping Emulation Environments Portable)<sup>6</sup> steht für diesen Wendepunkt, zusammen mit einer Reihe weiterer Initiativen – wie dem EU-Projekt PLANETS und der Open Planets Foundation, auf nationaler Ebene mit dem baden-württembergischen Langzeitarchivierungsprojekt bwFLA.<sup>7</sup> Alle diese Initiativen und Projekte greifen Werkzeuge und Ansätze auf, die in der Retro-Gamer-Community entstanden sind, und passen sie institutionellen Bedürfnissen an. Etwa beim Emulation Framework, das im Rahmen von KEEP entstanden ist. Es bietet die Möglichkeit, existierende Emulatoren mit bestehenden digitalen Sammlungen zu verbinden. Der Archivar muss sich hier nicht mehr mit den verschiedenen Einstellmöglichkeiten der unterschiedlichen Emulatoren auseinandersetzen, sondern findet die einheitliche Benutzeroberfläche eines Archivsystems vor.

Im Ergebnis all dieser Tendenzen erleben wir, wie sich die Verantwortung, das digitale Spieleerbe zu bewahren, von den Fans hin zum institutionellen Sektor verlagert. Auch wenn damit ein höherer materieller Ressourceneinsatz einhergeht, ist es ein notwendiger und folgerichtiger Schritt. Die Motivation der bisherigen Emulatoren-Programmierer und ihr emotionaler Bezug zu den Spielen und Plattformen werden zukünftigen Generation wohl fehlen. Einzelne Personen können nicht ersetzen, was Institutionen für die Bewahrung des Spieleerbes leisten können. Es ist dennoch bemerkenswert, wie viele Ressourcen bereits in Werkzeuge geflossen sind, die jetzt übergeben, weitergenutzt und -entwickelt werden, ohne dass dafür öffentliche Mittel geflossen sind.

### *Das Gaming-Museum der Zukunft*

Um großangelegte Archivsysteme auf Emulatorenbasis technisch zu realisieren, sind heute alle benötigten Komponenten verfügbar. Kombiniert man diese mit den ebenfalls existierenden Streamingtechnologien, ergibt sich daraus ein Bild, wie zukünftige Archive eingerichtet sein können, die den Zugang zu ihren Sammlungen über das Internet organisieren. In diesem Zusammenhang sind

6 Vgl. [www.keep-project.eu](http://www.keep-project.eu) (Letzter Aufruf: 15.08.2013).

7 Vgl. <http://bw-fla.uni-freiburg.de> (Letzter Aufruf: 15.08.2013).

Emulatoren nicht nur notwendiger Bestandteil einer Bewahrungsstrategie, sondern eröffnen auch breite Möglichkeiten des Zugangs.

Sicher trifft das hier Gesagte vor allem auf diejenigen digitalen Werke zu, die auf physikalischen Datenträgern vorliegen. Heute tritt der internetbasierte Zugang zu digitalen Werken als Service zunehmend an die Stelle des auf Speicherkarten, DVDs oder anderen Datenträgern angebotenen Werkexemplars. Hier sind Spiele ebenfalls Vorreiter der Entwicklung, die auch ihre Bewahrung vor neue Herausforderungen stellt. Für diese Spiele müssen nicht nur lokale Systeme, sondern ganze Serverstrukturen inklusive ihrer Online-Bezahlsysteme mit Hilfe von Emulatoren bewahrt werden. Prinzipiell ist das kein anderer Ansatz, doch kann etwa die soziale Interaktion – die für viele dieser Services wesentlich ist – nicht mit Hilfe von Emulation abgebildet werden. Hier wird man die digitale Bewahrung mit den traditionellen Mitteln und Fähigkeiten der Archivare ergänzen und Dokumentationen parallel zur Codebasis des Werkes erstellen müssen.

Akuter Handlungsbedarf besteht vor allem im juristischen Bereich. Solange die Rechtslage keinen praktikablen Transfer der Werke auf stabilere Datenträger zulässt, wird es nicht möglich sein, die Bewahrungsaktivitäten der Community auf eine tragfähige institutionelle Basis zu stellen. In Anbetracht der sich aufgrund von bitrot auflösenden Spielesammlungen ist schnelles Handeln gefordert. Ein möglicher Weg könnte eine neue Regelung im Urheberrecht sein, die es Archiven erlaubt, Kopierschutzmechanismen zu umgehen. In den USA wurde dies mit einer Ergänzung des Digital Millennium Copyright Act bereits 2003 vollzogen.

### *Mehr Einsatz für Games als Kulturgut nötig*

Ein Bewusstsein des Gesetzgebers, dass Games ebenso wie Bücher, Musikstücke und Filme digitale Werke darstellen und zum kulturellen Erbe gehören, ist jedoch noch kaum ausgeprägt. Dies zeigte sich auch bei der europäischen Richtlinie zu verwaisten Werken, die eine Möglichkeit geboten hätte, eine entsprechende Regelung zu verankern. Hier bleibt nur, langwierige Erklärungs- und Überzeugungsarbeit zu leisten. Europäische Einrichtungen haben erst kürzlich mit EFGAMP, der European Federation of Game Archives, Museums and Preservation Projects, einen eigenen Verband gegründet.<sup>8</sup> Neben Lobbyarbeit

8 Vgl. [www.computerspielemuseum.de/1243\\_European\\_Computer\\_and\\_Video\\_Game\\_Archives\\_and\\_Museums\\_Joining\\_Forces\\_to\\_Preserve\\_Gaming\\_Legacy.htm?nId=24](http://www.computerspielemuseum.de/1243_European_Computer_and_Video_Game_Archives_and_Museums_Joining_Forces_to_Preserve_Gaming_Legacy.htm?nId=24) (Letzter Aufruf: 15.08.2013).

und der Verbreitung von Wissen soll er auch Bewahrungs- und Beschreibungsstandards etablieren. Im Initiatorenkreis sind Organisationen aus der Gaming-Community ebenso vertreten wie klassische Gedächtniseinrichtungen, etwa die Dänische Nationalbibliothek und das Nationale Medienmuseum Englands.

Mit dieser Herangehensweise lassen sich die Qualitäten und Möglichkeiten beider Bereiche kombinieren. Neben den in der Retro-Gamer-Community entwickelten Werkzeugen können Nutzer zum Beispiel bei der Beschreibung von Werken wichtige Aufgaben übernehmen. In keinem anderen kulturellen Bereich sind die Möglichkeiten, große Metadatensammlungen mit Hilfe der Nutzer zu erzeugen, so erfolgreich erprobt worden wie bei Computerspielen. Mit den Jahren sind große und umfangreiche Sammlungen entstanden, die permanent von den Nutzern erweitert werden, ohne dass große Mittel zur Verfügung standen. Ein gutes Beispiel ist [mobygames.com](http://mobygames.com), dessen Inhalte ausschließlich von Nutzern erstellt wurden und das im Juli 2013 über detaillierte Beschreibungen von 73 844 Spielen auf über 138 verschiedenen Plattformen verfügt. Der Computerspiele-Leihservice Gamefly erkannte den Wert solcher Sammlungen und kaufte die Plattform 2010 auf.

All das zeigt, welche Chancen in der Kooperation und dem Zusammenwachsen der Retro-Gamer-Community mit den traditionellen Gedächtnisorganisationen liegen – auch für die Bewahrung digitaler Kulturgüter jenseits des Spielbereichs. Die neuen Möglichkeiten, den Zugang dauerhaft und niederschwellig zu organisieren, übertreffen die Schwierigkeiten, denen wir hier gegenüberstehen, bei weitem.



# Farbraum und Bildzustand im Kontext der Langzeitarchivierung

JAN FRÖHLICH

Wie wir Filme wahrnehmen, ist ein höchst subjektiver Akt. Doch der Wahrnehmung eines Films korrespondiert eine souveräne Beherrschung seiner Stilmittel: Bildgestaltung, Dramaturgie und Montage haben sich über Jahrzehnte hinweg entwickelt, wurden spielerisch variiert, verworfen und fortgeschrieben. Eines der subtilsten Stilmittel des Film – und des Kinos – ist die Farbe. Doch für sie gilt, mehr noch als für das reine Filmmaterial, eine sehr eingeschränkte Beständigkeit. Keine andere Eigenschaft des Films ist so gefährdet verlorenzugehen. Die zahlreichen rot- und grünstichigen Surrogate sind ein trauriger Beweis hierfür.

Mit der Einführung der digitalen Technik, welche sich praktisch flächendeckend vollzogen hat, sind diese Probleme in keinsten Weise gelöst, sondern bekommen eine neue Dimension.

Der Aspekt der Farbigkeit, insbesondere des Farbraums, hinsichtlich einer digitalen Langzeitarchivierung ist ein wenig beachtetes Thema innerhalb der Archivierungsbestrebungen.

## *Grundlagen*

Wie werden Bilder im Rechner gespeichert? Ein digitales Foto oder Film-Einzelbild ist aus zahlreichen Pixeln (Picture Elements) zusammengesetzt. Jedes Pixel besteht aus drei Werten: Rot, Grün und Blau. Als Modell für die weitere Betrachtung stellen wir uns eine Video-Leuchtreklame vor, wie wir sie aus Großstädten kennen. In einem gleichmäßigen rechteckigen Raster sind Elemente aus jeweils einer roten, grünen und blauen Glühbirne angeordnet. Jede Glühbirne wird



durch einen Dimmer angesteuert. Die Werte in der digitalen Repräsentation steuern, wie hell der Dimmer die jeweilige Glühbirne brennen lässt.

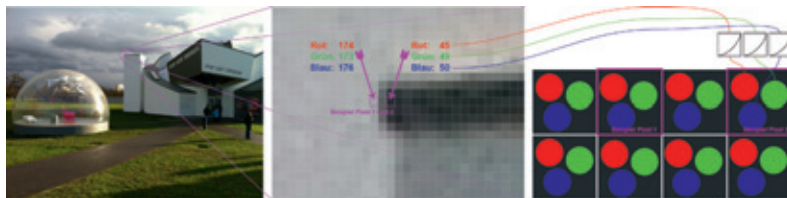


Abbildung 1: Foto und digitale Repräsentation

Codieren wir unsere Datei mit 8 Bit haben wir 256 Zustände pro Farbe zur Verfügung, also Werte von 0 bis 255. In dem Modell schaltet der Dimmer die jeweilige Glühbirne bei Wert 0 komplett ab, bei Wert 255 leuchtet sie mit voller Intensität. Wenn wir nun Daten für lange Zeit archivieren und in ferner Zukunft diese Daten wieder korrekt angezeigt werden sollen, stellt sich die Frage: Wie hell soll eine Glühbirne bei voller Intensität leuchten?

Im Falle von sRGB<sup>1</sup> ist festgelegt, dass beim maximalen Wert von 255 alle drei Farben gemeinsam ein Weiß mit einer Helligkeit von 80 Candela pro Quadratmeter (cd/m<sup>2</sup>) erzeugen sollen.

Welche Helligkeit müssen die drei Lampen nun erzeugen, wenn sie den mittleren Wert zwischen 0 und 255, also 127, darstellen sollen?

Die einfachste, aber falsche Antwort wäre, dass der halbe Wert in der digitalen Repräsentation zur halben Helligkeit am Monitor führt, also 40 Candela. Da wir Menschen dunkle Bildelemente aber feiner abgestuft wahrnehmen als helle Bildelemente, ergibt es Sinn, Bilder nicht linear zu codieren. Bei sRGB codiert der Schritt vom Wert 0,0,0 zu 1,1,1 einen Helligkeitsunterschied von  $\sim 0,0243$  cd/m<sup>2</sup>. Der Schritt vom Wert 254,254,254 zu 255,255,255 codiert einen Helligkeitsunterschied von  $\sim 0,7118$  cd/m<sup>2</sup>.

Die Vorschrift, nach der den digitalen Werten Display-Helligkeiten zugeordnet werden, wird oft als Gamma oder allgemein als Elektro-Optische-Transfer-

1 Stokes, Michael/Anderson, Matthew/Chandrasekar, Srinivasan/Motta, Ricardo: A standard default color space for the internet - srgb. Version 1.10. 1996.

Funktion (EOTF) bezeichnet. Wenn wir zukünftigen Generationen die originalgetreue Anzeige von digitalen Bildern ermöglichen wollen, müssen wir also nicht nur die maximale Helligkeit für die Anzeige dieser Bilder hinterlegen, sondern zusätzlich eine Vorschrift, die es erlaubt, für jeden digitalen Wert die Helligkeit der Darstellung zu berechnen. Abbildung 2 zeigt die EOTF von heutigen Monitoren und digitalem Kino:

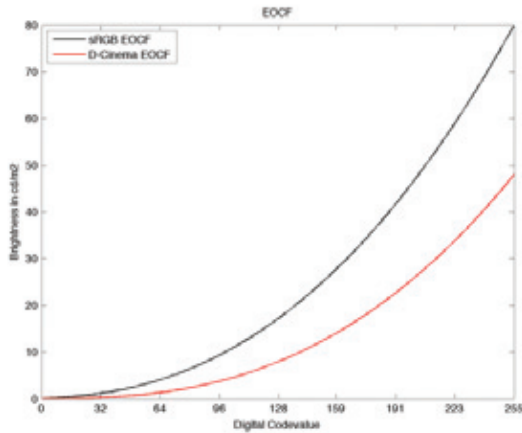


Abbildung 2: EOTF von Computermonitoren (sRGB) und digitalem Kino (SMPTE 428-1<sup>2</sup>)

Eine letzte Information fehlt nun noch, um die digitale Bildrepräsentation korrekt darzustellen. Im Beispiel der Bildeinwand gesprochen, ist dies der exakte Farbton der verwendeten roten, grünen und blauen Lampen sowie deren Mischungsverhältnis. Bei sRGB sind die Farborte von Rot, Grün und Blau als  $(0,64; 0,33)$ ,  $(0,30; 0,60)$  und  $(0,15; 0,06)$  in CIE  $xy$ -Koordinaten festgelegt. Das Helligkeitsverhältnis der drei Farben ist so zu wählen, dass sich bei Wert 255,255,255 ein Weiß mit dem Farbort  $(0,3127; 0,3290)$  ergibt.

Zum Abschluss des Grundlagenteils zwei Warnhinweise bezüglich der oft für die Darstellung des Farbumfangs verwendeten CIE-Normfarbtafel:

- 2 Society of Motion Picture and Television Engineers: SMPTE 428-1-2006, D-Cinema Distribution Master - Image Characteristics.

1. Das CIE xy-Diagramm ist nicht visuell gleichabständig. In den Blautönen entspricht eine geringe Ortsänderung im Diagramm einer deutlich größeren visuellen Farbänderung als im grünen Bereich.<sup>3</sup>
2. Ein Monitor kann nicht automatisch alle Farben darstellen, die innerhalb des von den Primärfarben aufgespannten Dreiecks liegen. Der in Abbildung 3 durch einen weißen Punkt angezeigte Farbton liegt in der 2-D-Darstellung zwar innerhalb des sRGB-Dreiecks, kann aber, wie in der 3-D-Ansicht leicht zu sehen ist, aufgrund seiner Helligkeit nicht von einem sRGB-Monitor dargestellt werden.

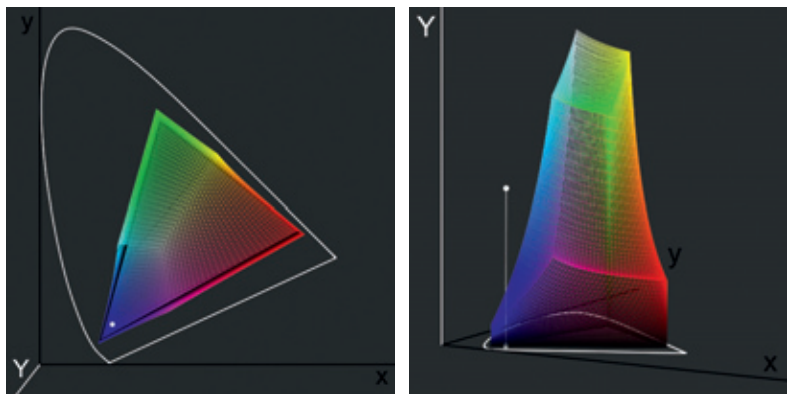


Abbildung 3: CIE-Normfarbtafel in der klassischen 2-D-Ansicht und in 3-D.

Das farbige Gitter beschreibt die Grenzen des sRGB-Farbraums. Der weiße Punkt, welcher einen hellblauen Farbton repräsentiert, scheint in der 2-D-Ansicht innerhalb des Farbraums zu liegen. In der 3-D-Ansicht wird sichtbar, dass dieser Farbton nicht von einem sRGB-Monitor dargestellt werden kann.

### *Der Bildzustand*

Ein Archivar wird sich vielleicht über diese Definition eines Farbraums wundern. Warum definieren wir die Farbigkeit eines Bildes über das Anzeigegerät?

3 MacAdam, David L.: Visual sensitivities to color differences in daylight. In: JOSA 32.5 (1942), S. 247-273.

Wäre es nicht sinnvoller, unsere Farben auf die Farbtöne der fotografierten Szene zu beziehen?

Dies liegt daran, dass die Bilder in den meisten Bildwiedergabesystemen zwischen Aufnahme und Anzeige modifiziert werden. Wird mit einer Kamera eine Szene aufgenommen und diese dann auf einem Monitor wiedergegeben, kann normalerweise keine Eins-zu-eins-Wiedergabe der Helligkeiten und Farbigkeiten der Szene erzielt werden. In vielen Fällen ist die Szene deutlich heller, als der Monitor im Stande ist wiederzugeben. Wenn wir die Szene nun dunkler darstellen wollen, müssen wir die Kontrastwiedergabe anpassen, um den gleichen Kontrasteindruck zu erzeugen.<sup>4</sup> Auch ästhetische Beweggründe spielen in der Bildverarbeitung einer Kamera eine Rolle. Viele Kamerasysteme bevorzugen Hauttöne und reduzieren dagegen die Intensität stark gesättigter Farben. Dies ist kein neues digitales Phänomen, sondern hat eine lange Tradition aus der analogen Fotografie. Neben dem Farbraum selber ist für die korrekte Interpretation also zusätzlich die Information notwendig, ob der Farbraum szenen- oder displayreferenziert ist. Displayreferenzierte Daten können direkt wiedergegeben werden, szenenreferenzierte Daten müssen noch für die jeweilige Displaytechnologie gerendert werden und sind damit flexibler bezüglich zukünftiger Präsentationsformen.

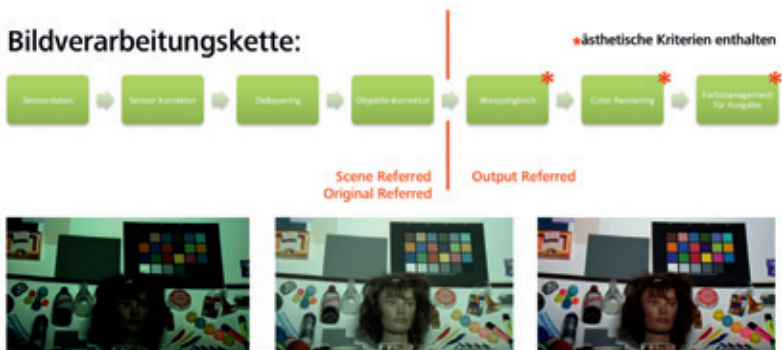


Abbildung 4. Bildverarbeitungskette.

4 Fairchild, Mark D.: Color appearance models. Third Edition, 6.7. Chichester 2013, S. 122.

*Anwendung für die digitale Langzeitarchivierung*

Welche Gründe gibt es, in der digitalen Langzeitarchivierung ausgabereferenzierte Farbräume zu verwenden?

*Beispiel 1: Fotografie von Gemälden*

Wenn ein Gemälde durch Fotografie digitalisiert wird, soll keine neue ästhetische Interpretation dieses Gemäldes erstellt werden, sondern es sollen möglichst die originalen Farbwerte reproduziert werden. Hier kann für die jeweilige Kamera eine Kameracharakterisierung erstellt werden, sodass sich die zu archivierenden Werte auf die Helligkeit und Farbigkeit des Gemäldes hinsichtlich einer speziellen Beleuchtung beziehen und nicht auf dessen Reproduktion auf einem Monitor. In der Folge können unabhängig von Displaytechnologien für jeden neuen Farbraum Renderings erstellt werden, die das Gemälde auf diesem Display möglichst originalgetreu wiedergeben.

*Beispiel 2: Retrodigitalisierung von Filmen*

Wenn Filme als Negativ oder Positiv vorliegen, ist es sinnvoll, die Negativ- beziehungsweise Positividichten zu scannen und in dieser Form zu archivieren. Es versteht sich von selbst, dass als Benutzungsstück zusätzlich ein Rendering für den Farbraum aktueller Projektoren erstellt werden muss. Dies kann z. B. ein DCP sein. Das DCP oder auch das visuell identische DCDM sind zur Archivierung jedoch nicht optimal, da hier das Rendering für den aktuellen Zielfarbraum des digitalen Kinos schon enthalten ist.

*Beispiel 3: Aktuelle Kinoproduktionen*

Die Situation bei aktuellen Kinoproduktionen ist analog zu der Situation in der Retrodigitalisierung zu sehen. Ein schönes Beispiel sind die aktuellen Renderings für die Darstellung auf Monitoren. Hier werden bei der Erstellung der displayreferenzierten Derivate oft gegen den Protest der Bildgestalter Korn- und Rauschreduktionsalgorithmen eingesetzt. Dies dient dazu, die Übertragung per DVB-T mittels niedrigbitratiger H.264-Kompression zu ermöglichen. Die hierdurch entstehenden Bilder widersprechen jedoch zum Teil dem ästhetischen Konzept des Films. Besonders die Anmutung der Materialität von Objekten verändert sich durch die Kornreduzierung. Für die Langzeitarchivierung dürften sowohl diese auf eine spezielle Übertragungstechnik optimierte

Version als auch die unbehandelte Version, die der originalen künstlerischen Intention entspricht, interessant sein. Es erscheint daher sinnvoll, für aktuelle Produktionen sowohl die Rohdaten als auch die gerenderten Derivate als Referenz zu archivieren.

### *Zusammenfassung*

Anhand dieser drei Beispiele wird gezeigt, dass der Postproduktionsprozess immer auf das Endformat hin konzipiert ist. Abhängig davon, ob eine Kino-, Fernseh- oder eine bestimmte digitale Distribution gewünscht wird, werden entsprechende Schritte in der Postproduktion umgesetzt. Allerdings sind diese Zielformate temporär. Aufgrund des Technologiewandels erwarten wir größere Farbräume seitens der Endgeräte.

Für eine Archivierung, welche nicht nur die physische Sicherung, sondern auch die originalgetreue „Aufführung“ in 100 Jahren zum Ziel hat, ist daher neben den Zielformaten auch die Sicherung szenenreferenzierter Daten notwendig.

Die Definition des Bildzustandes nach ISO 22028-1<sup>5</sup> ermöglicht es, in der digitalen Archivierung genau diese Auswahl zu treffen: objektreferenziert, szenenreferenziert oder displayreferenziert. Die Wahl des richtigen Bildzustands für die Archivierung wird wie die Dokumentation dieser Metadaten über die Nachhaltigkeit dieser digitalen Langzeitarchivierung entscheiden.

5 ISO 22028-1. Photography and graphic technology - Extended colour encodings for digital image storage, manipulation and interchange. International Standard ISO 22028-1:2004.



Lastenaufzug, Konrad-Zuse-Zentrum für Informationstechnik Berlin

# Zwischenrufe







## Facebook, Froschlaich, Folianten

KATHRIN PASSIG

300 Jahre alte Bücher können wir problemlos lesen, wenige Jahre alte Disketten nicht. Wenn alles ins Internet wandert, ist es dann womöglich eines Tages einfach weg? So oder ähnlich beschäftigt die Frage seit Jahren Podiumsdiskussionsteilnehmer, Journalisten und Privatleute. Die Sorge, ein neues Speichermedium werde dazu führen, dass demnächst alles verschwunden ist, ist aber schon etwas älter. Prominent vorgetragen wurde sie im Jahr 1492 von Johannes Trithemius in „De laude scriptorum manualium“ („Zum Lob der Schreiber“):

„Gedrucktes aber, da es auf Papier steht, wie lange wird es halten? Geschriebenes, wenn man es auf Pergament bringt, wird an die tausend Jahre Bestand haben [...] wenn Gedrucktes in einem Band aus Papier an die zweihundert Jahre Bestand haben wird, wird es hoch kommen.“

Trithemius wird gern mit spöttischem Unterton zitiert, aber seine Zweifel an der Haltbarkeit des Papiers waren nicht unberechtigt.

Trotzdem hat sich das Papier seither ganz gut bewährt. Dass Gedrucktes billiger und leichter zu vervielfältigen war, ermöglichte höhere Auflagen und damit einen Ausgleich der geringeren Überlebenswahrscheinlichkeit des einzelnen Textes. Auf dieselbe Strategie setzt der Frosch, wenn er große Mengen Laich unbeaufsichtigt in die Landschaft entlässt – der Laich ist weder robust noch gut bewacht, aber ein paar Fröschelein werden am Ende schon durchkommen. Passenderweise ist „De laude scriptorum manualium“ nicht als Handschrift erhalten, sondern nur in der auf Papier gedruckten Version.

Auch in der Architektur geht es nicht unbedingt um die Robustheit des Materials. 2012 zitierte ich bei Twitter aus einem Vortrag des Architekten Friedrich von Borries: Nachhaltigkeit sei kein Wert an sich, die nachhaltigste Architektur seien Bunker.<sup>1</sup> Daraufhin meldete sich Sascha Brossmann<sup>2</sup> zu Wort und gab zu bedenken, die nachhaltigste Architektur seien bisher immer noch Sakralbauten.<sup>3</sup> Die Kirche setzt nicht auf meterdicken Stahlbeton, sondern auf die Kraft einer Idee. Durch diese Idee bewegt sie die Menschen über viele Generationen hinweg zu regelmäßigen Reparaturen. Falls der Bau doch zerstört wird, errichtet man häufig an derselben Stelle einen neuen.

Wenn Informationen verlorengehen, dann liegt das in aller Regel nicht an Materialversagen, sondern an nachlassendem Interesse. Brennende Bibliotheken sind ein spektakuläres Bild, spielen aber als Ursache für Verluste keine wesentliche Rolle. Die meisten Inhalte hat einfach irgendwann niemand mehr für aufbewahrenswert, also weiterkopierenswert, gehalten. Alte Texte sind uns in aller Regel nicht deshalb überliefert, weil ihr Material so außerordentlich haltbar wäre, sondern weil sie immer wieder abgeschrieben oder nachgedruckt wurden.

Dass Inhalte verschwinden, weil es zu wenig Interesse an ihrer Erhaltung gibt, ist kein Problem der Vergangenheit. In der Frühzeit von Radio- und Fernsehsendungen klaffen große Lücken,<sup>4</sup> ebenso wie in der Frühgeschichte des Internets. Als der kostenlose Webhostingdienst Geocities 2009 nach 15 Jahren eingestellt wurde, waren es nur diverse private Initiativen, die noch schnell wenigstens einen Teil der Daten sichern konnten,<sup>5</sup> ähnlich wie bei der Ende April 2013 geschlossenen Blogplattform Posterous.

Das Problem hat in erster Linie damit zu tun, dass diese Daten selbst den für Archivierung von Kulturgut zuständigen Fachleuten der jeweiligen Periode als nicht besonders erhaltenswert gelten. Die Einsicht, dass sich schon wenige Jahre später Forscher für Moritatentafeln, Ansichtskarten, Comichefte und Fernseh-

1 Vgl. <http://twitter.com/kathrinpassig/status/252023189423669249> (Letzter Aufruf: 15.08.2013).

2 Vgl. <https://twitter.com/brsma> (Letzter Aufruf: 15.08.2013).

3 Vgl. <https://twitter.com/brsma/statuses/252053688791740416> (Letzter Aufruf: 15.08.2013).

4 Vgl. [http://en.wikipedia.org/wiki/Doctor\\_who\\_missing\\_episodes](http://en.wikipedia.org/wiki/Doctor_who_missing_episodes) (Letzter Aufruf: 15.08.2013).

5 Vgl. <http://archiveteam.org/index.php?title=GeoCities> (Letzter Aufruf: 15.08.2013).

serien interessieren werden, muss offenbar für jede kulturelle Ausdrucksform wieder separat erarbeitet werden.

Die Geocities-Seiten waren sehr unansehnlich – nicht ganz so hässlich wie MySpace, aber hässlich genug, um vor ähnlichen Problemen zu stehen wie eine bedrohte Tierart, die nicht gerade mit Knopfüchtigkeit und Flauschigkeit gesegnet ist. Aber Geocities enthielt einen großen Teil der Frühgeschichte des World Wide Web, und seine zweifelhafte Ästhetik ist Bestandteil dieser Geschichte.

Das Hässlichkeitsproblem ist mittlerweile etwas in den Hintergrund getreten, dafür beleidigt das Internet das Auge des Archivars durch fehlenden inhaltlichen Adel der Alltagskommunikation bei Twitter und Facebook, in Blogs, Chats und Kommentaren.

Zum fehlenden Erhaltungswillen kommen rechtliche Probleme: Archivierungswillige in den USA können immerhin versuchen, auf Fair Use zu plädieren,<sup>6</sup> in Deutschland ist die Lage noch schwieriger. Die Froschlaich-Strategie ist grundsätzlich problematisch, wenn Urheber oder Verwerter gar nicht wollen, dass Inhalte vervielfältigt werden, also bei privaten oder geschützten Daten. Alles offen Zugängliche hat deutlich bessere Chancen auf Erhaltung.<sup>7</sup> Kopiergeschützte Daten, also zum Beispiel die meisten E-Books, existieren zwar in großer Zahl auf verschiedenen Geräten, aber sie sind verschlüsselt, und wenn die zentrale Komponente wegfällt, weil etwa der Anbieter seine Tätigkeit einstellt,<sup>8</sup> bleibt nur das Schlechteste aus zwei Welten übrig: ein fragiles Speicherformat gepaart mit fehlender Quantität.

Wenn die Erhaltung digitaler Inhalte unter dem Aspekt der technischen Haltbarkeit der Speichermaterialien diskutiert wird, steckt darin ein Teil Wunschdenken: Die technischen Probleme sind ein angenehmeres Gesprächsthema, weil sie leichter zu bewältigen sind als die eigentlichen Archivierungshindernisse. Und die bleiben über die Jahrhunderte weitgehend gleich: Es fehlt an der Überzeugung, dass die jeweiligen Inhalte aufbewahrensenswert sind, und es fehlen die rechtlichen und institutionellen Rahmenbedingungen, die die Aufbewahrung möglich machen.

6 Vgl. [http://de.wikipedia.org/wiki/Fair\\_use](http://de.wikipedia.org/wiki/Fair_use) (Letzter Aufruf: 15.08.2013).

7 Vgl. [http://de.wikipedia.org/wiki/Langzeitarchivierung#Weitere\\_Beschr.C3.A4nkungen](http://de.wikipedia.org/wiki/Langzeitarchivierung#Weitere_Beschr.C3.A4nkungen) (Letzter Aufruf: 15.08.2013).

8 Vgl. [www.spiegel.de/netzwelt/web/netzwelt-ticker-yahoo-dreht-kunden-die-musik-ab-a-568086.html](http://www.spiegel.de/netzwelt/web/netzwelt-ticker-yahoo-dreht-kunden-die-musik-ab-a-568086.html) (Letzter Aufruf: 15.08.2013).

Dass es nicht auf das Material ankommt, belegt das Beispiel des Wawilow-Instituts in Sankt Petersburg. Die 1921 gegründete Einrichtung beherbergt heute noch die drittgrößte Saatgutsammlung der Welt. Während der Blockade der Stadt von 1941 bis 1944 verhungerten über 1 Millionen Menschen, darunter 6 Mitarbeiter des Instituts. Das Archivgut ist nicht nur essbar – es besteht unter anderem aus Kartoffeln und Reis –, es muss auch regelmäßig ausgepflanzt und neu vermehrt werden, weil es nach ein paar Jahren seine Keimfähigkeit verliert. Ein vergänglicheres, gefährdeteres Material ist kaum denkbar, und doch hat das Archiv die Blockade, den Krieg und den Rest des 20. Jahrhunderts überstanden. Seine Erhaltung hing nicht von der Robustheit des Archivguts ab, sondern von Überzeugung, Infrastruktur und politischer Unterstützung. Erst das Ende der Sowjetunion bedroht heute seinen Fortbestand.

Genauso wenig, wie man Kartoffelsorten in eine leichter archivierbare Form bringen kann, lassen sich digitale Inhalte einfach in robustere Formen überführen, ohne dass wesentliche Eigenschaften verlorengehen. (Auch wenn die Universitätsbibliothek der Helmut-Schmidt-Universität genau darauf in ihrem Aprilscherz 2012 setzt: Zunächst sei eine Pergamentisierung des Bestandes geplant, später dann eine Sicherheitskopie auf Granit.)<sup>9</sup> Die russischen Pflanzen genetiker haben nicht nach anderen Kartoffeln verlangt, und wir sollten damit aufhören, nach anderen, irgendwie weniger digitalen Dateien zu rufen. Was fehlt, sind den Entwicklungen angepasste Rechtsgrundlagen und Strategien – um das Kopieren zu erleichtern.

Die Haltbarkeit vieler Produkte lässt sich nur unter Verzicht auf andere Vorteile erkaufen: Flugzeuge werden nicht aus Hartgummi hergestellt, Fensterscheiben nicht aus Stahl, und die meisten Menschen ernähren sich nicht ausschließlich von Produkten, deren Verfallsdatum mehrere Jahre in der Zukunft liegt. In der Diskussion um die Haltbarkeit digitaler Inhalte aber sieht es oft so aus, als sei die Dauerhaftigkeit des Analogen uneingeschränkt von Vorteil.

„Das wohl wichtigste Versprechen des Digitalzeitalters wurde gebrochen“, hieß es vor einigen Jahren in der FAZ. „Stand am Anfang der Gedanke, dass Digitales aller Art durch seine entmaterialisierten binären Inhalte der herkömmlichen analogen Speichertechnik in jeder Hinsicht überlegen ist, stellt sich jetzt Ernüchterung ein. Bibliotheken, Archive, Unternehmen, Behörden, Medien und

9 Vgl. <http://wp.ub.hsu-hh.de/2241/einzigartiges-projekt-zur-langzeitarchivierung-an-der-ub-der-hsu/> (Letzter Aufruf: 15.08.2013).

Universitäten müssen sich schon lange mit dem Problem auseinandersetzen, dass nicht einmal 30 Jahre alte digitale Informationen für immer verloren sind.“<sup>10</sup>

Dabei haben auch haltbare Medien ganz konkrete Nachteile: Am gedruckten Text wird häufig gelobt, dass er nicht einfach über Nacht wieder geändert werden kann. Aber dieselbe Eigenschaft bringt es mit sich, dass Fehler, die sich einmal in den Text eingeschlichen haben, dort lange stehen bleiben. Haltbarkeit des Mediums und einfache Fehlerkorrektur schließen sich gegenseitig aus. Man kann Text in einen Marmorblock meißeln, dann sind auch die Fehler dauerhaft. Oder man schreibt ihn ins Netz, dann sind die Fehler leicht zu beheben, aber auch der übrige Text ist nicht ganz so dauerhaft wie ein Marmorblock.

Haltbare Medien erschweren nicht nur die Fehlerkorrektur, sondern eignen sich auch weniger für Informationen, die sich schnell verändern. Und da die haltbaren Medien zwar nicht immer Marmorblöcke, aber doch etwas größer und schwerer sind als die weniger dauerhaften, gilt dasselbe für sehr umfangreiche Werke. Nachschlagewerke vereinen häufigen Änderungsbedarf mit großem Umfang und wurden daher als Erste von der Digitalisierung verzehrt. Wäre die englischsprachige Wikipedia ein Lexikon auf Papier, dann hätte sie ohne Bilder knapp 1800 dicke Bände. Selbst die deutschsprachige Ausgabe wäre mit über 800 Bänden immer noch etwas unhandlich beim Umzug. Setzt man statt auf Papier auf ein wirklich haltbares Medium, wird die Lage nicht besser. Der österreichische Anbieter Memory of Mankind brennt für Privatpersonen und Institutionen Aufbewahrenswertes auf Keramiktafeln, die in einem ehemaligen Salzbergwerksstollen einige hunderttausend Jahre überdauern sollen. Für die 1,5 Millionen Einträge der deutschen Wikipedia wären mindestens eine halbe Millionen Tontafeln nötig. Und es ist deutlich schwerer, auf diese Tontafeln zuzugreifen, wenn man beispielsweise auf einer Zugfahrt dringend alles über Kuhmagnete oder Schmuckeremiten herausfinden muss.

Neben diesen konkreten Problemen bringt Haltbarkeit auch einige abstraktere Nachteile mit sich, zum Beispiel das Gegenstandsdenken. Als im letzten Winter der Verkauf der Stralsunder Archivbibliothek durch die Presse ging, drehte sich die gesamte Diskussion um die physischen Bücher: Die Gegenstände waren durch unsachgemäße Lagerung von Schimmel befallen, die Gegenstände waren verkauft worden. Dass diese Gegenstände auch Text enthalten

10 Vgl. [www.faz.net/aktuell/technik-motor/computer-internet/langzeitarchivierung-das-digitale-daten-desaster-1539385.html](http://www.faz.net/aktuell/technik-motor/computer-internet/langzeitarchivierung-das-digitale-daten-desaster-1539385.html) (Letzter Aufruf: 15.08.2013).

und dass man diese Inhalte eventuell besser zugänglich machen könnte als in einem feuchten und offenbar nicht besonders seriös verwalteten Archiv, davon war nicht die Rede. Wenn die Stralsunder Bibliothek den Bestand vor dem Verkauf digitalisiert hätte, hätte man in der Diskussion immerhin trennen können zwischen dem Wunsch nach Zugänglichkeit der Buchinhalte und dem Wunsch, die Bücher aus historischen Gründen in genau dieser Zusammenstellung an diesem Ort zu behalten. Der gedruckte Text lenkt die Aufmerksamkeit stark auf den physischen Gegenstand Buch, das beeinflusst das Nachdenken über Bücher und die Prioritäten beim Geldausgeben.

Weitere abstrakte Nachteile des Haltbaren hat der kanadische Kommunikationstheoretiker Harold Innis in den 1940er Jahren beschrieben.<sup>11</sup> Innis unterschied zwei Arten von Medien, die er als „space-binding“ und „time-binding“ bezeichnete. Zeitbindende Medien sind beispielsweise Steintafeln oder Pergament. Inhalte, die man solchen zeitbindenden Medien anvertraut, halten lange, erreichen aber nur ein begrenztes Publikum. Bei den raumbindenden ist es umgekehrt. Sie sind flüchtiger, haben aber eine größere Reichweite. Zu diesen raumbindenden Medien gehörten für Innis Radio, Fernsehen und die überregionale Tageszeitung, heute würde er vermutlich auch Internet und E-Books dazuzählen.

Aus diesen unterschiedlichen Eigenschaften der Medien leitet Innis Folgen für die jeweilige Gesellschaft ab. Die zeitbindenden Medien fördern Stabilität, Gemeinschaft, Tradition und Religion. Die raumbindenden fördern schnelle Veränderungen, Materialismus, Säkularismus und die Ausdehnung von Machtbereichen. Die Kirche trägt also nicht nur zur Beliebtheit des Pergaments bei, sondern das Pergament stärkt seinerseits die Dominanz der Kirche mit ihren Ideen, die für die Ewigkeit niedergelegt, aber nur für wenige zugänglich sind. Die Vorherrschaft der Tageszeitung hingegen führt zu Kommunikationsmonopolen und zur Vernachlässigung langfristiger Entwicklungen und dauerhafter Ideen.

Es ist schade, dass Harold Innis das Internet nicht mehr miterlebt hat. Das gesprochene Wort gehört für Innis zu den zeitbindenden Medien, und der Idealzustand einer Zivilisation war für ihn einer, in dem ein Gleichgewicht zwischen Gesprochenem und Geschriebenem herrscht. Das Internet ist zwar eindeutig ein raumbindendes Medium, aber wie viel von dem, was im Netz geschieht, hätte Innis wohl als mündliche Kommunikationsform betrachtet?

11 Innis, Harold: *Empire and Communications*. Toronto 1950.

Vieles Aufgeschriebene ist in Stil und Zweck dem mündlichen Austausch sehr nahe: schriftliche Diskussionen, Chats und Kurznachrichten. Sind Podcasts, Videos, MOOC-Onlinekurse<sup>12</sup> (die überwiegend auf Videos basieren) schriftliche oder mündliche Formen des Austauschs?

Ein dritter Nachteil herkömmlicher Medien ist eigentlich vielmehr ein selten erwähnter Vorteil der neuen: Wenn heute beklagt wird, dass digitale Inhalte wegen ihres Speichermediums vom Verschwinden bedroht sind, dann geht es dabei oft um Inhalte, die noch vor wenigen Jahren überhaupt nicht aufgezeichnet, geschweige denn archiviert oder öffentlich zugänglich gemacht wurden. Dazu gehört die gesamte Alltagskommunikation und damit ein wesentlicher Teil des Geschehens im Internet.

Dieser Artikel basiert auf einem Vortrag auf der Abschlusskonferenz der 8. Initiative „Nachhaltigkeit in der Digitalen Welt“ des Internet und Gesellschaft Collaboratory e. V. am 31. Mai 2013 im Jüdischen Museum Berlin, erstmals veröffentlicht auf ZEIT online am 3. Juni 2013.<sup>13</sup>

12 Vgl. <http://de.wikipedia.org/wiki/MOOC> (Letzter Aufruf: 15.08.2013).

13 Vgl. [www.zeit.de/digital/internet/2013-06/passig-kopieren-digitales-erinnern](http://www.zeit.de/digital/internet/2013-06/passig-kopieren-digitales-erinnern) (Letzter Aufruf: 01.08.2013).





*„Kultur und Wissen online, ein funktionierendes kulturelles Gedächtnis im Internet, setzen die digitale Langzeitarchivierung zwingend voraus.“*

– ELLEN EULER FÜR DIE DEUTSCHE  
DIGITALE BIBLIOTHEK



# Kulturelles Erbe und Nachhaltigkeit

STEFAN WILLER

## *Zur Begrifflichkeit von Erbe und Nachhaltigkeit*

In der 1972 verabschiedeten World Heritage Convention der UNESCO geloben die Unterzeichnerstaaten „Erfassung, Schutz und Erhaltung in Bestand und Wertigkeit des in seinem Hoheitsgebiet befindlichen [...] Kultur- und Naturerbes sowie seine Weitergabe an künftige Generationen“.<sup>1</sup> So wie in dieser Formel ist das Konzept des Erbes auch sonst begriffs- und diskursgeschichtlich mit dem der intergenerationellen Übertragung verknüpft. Das gilt für alle drei Aspekte des Erbe-Begriffs, wie er sich – in eben dieser Dreigliedrigkeit – seit dem 19. Jahrhundert herausgebildet hat: für die zivilrechtlich kodifizierte Eigentumsübertragung, für die biologische Weitergabe von Eigenschaften (bzw. deren Anlagen) und für die kulturelle Traditionsbildung. Was aber heißt es eigentlich, Praktiken der Weitergabe von Kultur mit der Bezeichnung Erbe zu belegen?

Sehr oft denkt man beim Kulturerbe an einen aus der Vergangenheit stammenden Schatz, den man aufzubewahren hat. Wie an der Festlegung der UNESCO-Konvention auf „Erfassung, Schutz und Erhaltung“ zu sehen ist, impliziert ein solcher Begriff des kulturellen Erbes ein konservatorisches Programm

1 UNESCO: Übereinkommen zum Schutz des Kultur- und Naturerbes der Welt (offizielle deutsche Übersetzung aus dem Bundesgesetzblatt), Art. 4. [www.unesco.de/welterbekonvention.html](http://www.unesco.de/welterbekonvention.html) (Letzter Aufruf: 13.07.2013). In der englischen Version lautet die Wendung „identification, protection, conservation, presentation and transmission to future generations of the cultural and natural heritage“. UNESCO: Convention Concerning the Protection of the World Cultural and Natural Heritage. [whc.unesco.org/archive/convention-en.pdf](http://whc.unesco.org/archive/convention-en.pdf) (Letzter Aufruf: 13.07.2013).

– zugleich aber auch, im Abzielen auf die „Weitergabe an künftige Generationen“, ein Programm, das in die Zukunft gerichtet ist. Genau hier kommt der Begriff der Nachhaltigkeit ins Spiel. Er spielt heute in der Kulturpolitik eine zentrale Rolle, vor allem in jener globalisierten Ausprägung, die maßgeblich von der UNESCO verkörpert wird. Francesco Bandarin, der frühere Direktor des World Heritage Center der UNESCO, nannte Nachhaltigkeit vor ein paar Jahren einmal den „Schlüssel sowohl zum Überleben des Welterbes als auch zu seiner Glaubwürdigkeit“ („sustainability remains the key to both the survival of World Heritage and its credibility“).<sup>2</sup>

Aus der Forstwirtschaft stammend und lange als ökologischer Terminus festgelegt, wird der Begriff Nachhaltigkeit seit den 1980er Jahren auf das Konzept der Generationen bezogen, um in Zukunftsprogrammen Ressourcen für kommende Generationen zu berücksichtigen.<sup>3</sup> Mittlerweile ist Nachhaltigkeit zu einem überaus anschlussfähigen politischen Begriff geworden, der bisweilen geradezu als Synonym für Generationengerechtigkeit verwendet wird.<sup>4</sup> Dabei entsteht in Bezug auf das Konzept der Generation das Versprechen einer Sinnstiftung durch zweifellose Evidenz, da sich im Muster einander ablösender Generationen historisch kontingente Veränderungsprozesse als gleichsam natürlicher Wandel, als Rhythmus eines natürlichen Reproduktionsgeschehens verstehen lassen.<sup>5</sup> Aus dieser vermeintlichen Natürlichkeit speist sich Nachhaltigkeit als kulturpolitische Formel für Sicherstellung, Bewahrung und Weitergabe an künftige Generationen. Nachhaltigkeit ist also, wie kulturelles Erbe, einerseits explizit vergangenheitsbezogen (in Form einer Aufforderung zur Bewahrung), andererseits in die Zukunft gerichtet – und zwar in eine maximal weite Zukunft. So heißt es weiter in dem erwähnten Artikel von Francesco Bandarin, wenn man von der Konservierung des Erbes spreche, dann denke man „per definitionem

2 Bandarin, Francesco: Protecting Heritage. In: Our Planet 14/2 (2003), S. 11-12, S. 12.

3 Vgl. Grunwald, Armin/Kopf Müller, Jürgen: Nachhaltigkeit. Eine Einführung. Frankfurt/M. 2006; Grober, Ulrich: Die Entdeckung der Nachhaltigkeit. Kulturgeschichte eines Begriffs. München 2010.

4 Vgl. Ekardt, Felix: Das Prinzip Nachhaltigkeit. Generationengerechtigkeit und soziale Gerechtigkeit. München 2005; Mathis, Klaus (Hrsg.): Efficiency, Sustainability, and Justice to Future Generations. Dordrecht 2011.

5 Vgl. Parnes, Ohad/Vedder, Ulrike/Willer, Stefan: Das Konzept der Generation. Eine Wissenschafts- und Kulturgeschichte. Frankfurt/M. 2008.

langfristig – nicht für ein oder zwei Jahre, sondern für immer“ („conservation is by definition long term – not for a year or two, but for ever“).<sup>6</sup>

Dieser zeitliche Imperativ ist keineswegs unproblematisch. Schutz „für immer“ heißt, dass dem zu schützenden und zu bewahrenden Erbe ein sehr weitgehendes Recht an der Gegenwart und der Zukunft eingeräumt wird. Genau in dieser begrifflichen Allianz von Erbe, Nachhaltigkeit, Überleben und Konservieren offenbart sich das eigentümliche Zeitregime einer konservatorischen – oder auch konservativen – Futurisierung. Das Überleben, von dem bei Bandarin die Rede ist („survival of World Heritage“), nähert sich dabei in auffallender Weise dem ewigen Leben. In diesem maximalen Sinn verstanden, muss das kulturelle Erbe in eine grenzenlose Zukunft hinein bewahrt werden.

Der Sinn des Konzepts Erbe liegt aber gerade darin, dass dieser Befehl einer grenzenlosen Konservierung nicht zweifelsfrei gilt, sondern umstritten ist. Gemeinsam ist allen als Erbe klassifizierten Übertragungen, dass sie ein komplexes Verhältnis zwischen Vererbendem, Vererbtem und Erbendem erzeugen; gemeinsam ist ihnen nicht zuletzt, dass sie eine Zäsur voraussetzen: Erst durch eine Unterbrechung in der Kette der Wesen, Dinge oder Ereignisse kommt es überhaupt zur Übertragung. Am deutlichsten ist dies in der juristischen Definition des Erbes als einer Übertragung von Todes wegen, aber auch andere Formen der Vererbung – ob kulturell oder biologisch – bedürfen des Momentes der Zäsur.<sup>7</sup> Speziell die kulturelle Überlieferung ist kein kontinuierlicher Vorgang, sondern geprägt von Umbrüchen, Konflikten und Widersprüchen. Der Begriff des kulturellen Erbes bringt eben diesen grundsätzlich strittigen und diskontinuierlichen Status der Traditionsbildung auf den Punkt – so wie auch die Vererbung von Eigentum oft genug zu erheblichen Konflikten führt und von vornherein niemals ohne die Zäsur des Todes zu denken ist.

### *Konservierung als Problem*

In der Tat ist der Befund der Unterbrechung, der Zäsur, historisch grundlegend für das UNESCO-Konzept des kulturellen Erbes. Seine Entstehung ist nicht

6 Bandarin, Francesco: Protecting Heritage. In: Our Planet 14/2 (2003), S. 11-12, S. 12.

7 Vgl. dazu ausführlich Willer, Stefan/Weigel, Sigrid/Jussen, Bernhard (Hrsg.): Erbe. Übertragungskonzepte zwischen Natur und Kultur. Berlin 2013. Darin zum Folgenden v. a. Willer, Stefan: Kulturelles Erbe. Tradieren und Konservieren in der Moderne, S. 160-201.

zu trennen von der faktischen Zerstörung materieller wie immaterieller kultureller Werte durch jenen Weltkrieg, der die internationalen Organisationen der Vereinten Nationen erst hervorgebracht hat.<sup>8</sup> Auch für zahlreiche Stätten, denen seit der Konvention von 1972 der Status des Welterbes zugesprochen wurde, gilt das konstitutive Moment der Bedrohung. Besonders prominent wurde dies im Fall der Buddha-Statuen im afghanischen Bamiyan, einem Fall, der zugleich eine interessante Paradoxie des konservatorischen Denkens liefert. Die Paradoxie zeigt sich darin, dass die Statuen, die erst zwei Jahre nach ihrer Zerstörung durch die Taliban (2001) in die berühmte UNESCO-Liste aufgenommen wurden, sogleich den Status des „gefährdeten Erbes“ erhielten.<sup>9</sup> Der Begriff der Gefährdung wurde somit auf den Zustand *vor* der Zerstörung zurückprojiziert. Demnach kann der Imperativ von Schutz und Erhaltung in diesem Fall nicht dem Zustand gelten, in dem das Monument sich zum Zeitpunkt seiner Aufnahme befand, sondern dem Zustand, der der Zerstörung vorausging – wobei hinzuzufügen ist, dass eben diese Zerstörung die Aufnahme des Monuments in die Liste politisch entscheidend mitbegründete.

Nimmt man den Begriff der Gefährdung ernst, dann müsste die Erhaltung der Stätte in der – sobald möglich – ins Werk zu setzenden Restauration der Statuen bestehen. Das wäre insofern bemerkenswert, als eine solche Restauration gegen die zentrale Idee des modernen Denkmalschutzes verstieße. Denn seit der um 1900 zuerst von deutschen und österreichischen Kunsthistorikern wie Georg Dehio und Alois Riegl ausgegebenen Devise „konservieren, nicht restaurieren“<sup>10</sup> gelten historische Rückbauten als ebenso verpönt wie modernisierende Eingriffe. Geschützt wird vielmehr der Status quo im Moment der Deklaration eines Bauwerks zum Denkmal, mit dem Ziel, die Vorgeschichte inklusive historischer Eingriffe, ja selbst Zerstörungen, möglichst vollständig offenzulegen. Internationale Geltung erhielt diese Lehrmeinung in der 1964 verabschiedeten Charta von Venedig, die in vielerlei Hinsicht der UNESCO-Welterbe-Konvention von 1972 vorarbeitete, auch in den einzelnen Formulierungen

8 Darauf hat dankenswerterweise Verena Metzke-Mangold (Deutsche UNESCO-Kommission) in der Diskussion zur Vortragsfassung dieses Beitrags am 31. Mai 2013 hingewiesen.

9 Vgl. UNESCO / World Heritage Center: World Heritage in Danger. <http://whc.unesco.org/en/158/> (Letzter Aufruf: 13.07.2013).

10 Vgl. Dehio, Georg/Riegl, Alois: Konservieren, nicht restaurieren. Streitschriften zur Denkmalpflege um 1900. Braunschweig 1988.

wie „lebendige Zeugnisse jahrhundertealter Traditionen“, „gemeinsames Erbe“, „Verantwortung gegenüber kommenden Generationen“. <sup>11</sup> In der Charta heißt es, weder bei der Konservierung noch bei der Restaurierung eines Bau- denkmals komme es auf stilistische Einheit an; vielmehr seien die „Beiträge aller Epochen zu einem Denkmal“ zu würdigen. <sup>12</sup> Überbauungen und „sich überlagernde Zustände“ <sup>13</sup> sind nicht zu entfernen, sondern ausdrücklich zu erhalten.

Der in der Charta von Venedig kodifizierte Vorrang des Konservierens gegenüber dem Restaurieren geht bis zur apodiktischen Forderung nach der Erhaltung von Ruinen. <sup>14</sup> Es ließe sich also argumentieren, dass die zerstörten Buddha-Statuen von Bamiyan gerade als Ruinen ein Monument des Welterbes darstellen: ein Trümmerdenkmal, in dem kein einziger Stein verrückt werden dürfte. <sup>15</sup> Um aus derartigen Ausweglosigkeiten des Erhaltungsimperativs herauszukommen, bietet allerdings bereits die Charta von 1964 einen kasuistischen Ausweg an. Demnach sind Zerlegungen und Neuzusammensetzungen von zerstörten oder ausgegrabenen Monumenten erlaubt, sofern man sie als sogenannte Anastylose klassifizieren kann, das heißt als Wiedervereinen vorliegender, aber zerstreuter Bestandteile eines Ganzen. <sup>16</sup>

Solche Anastylosen sind heute das typische Kennzeichen technisch fortschrittlicher Restaurierungen. So entstehen Bauten, die in zunehmend perfekter Weise den Zustand vor einer Zerstörung nachzuahmen vermögen, ja selbst die zwischenzeitliche Nichtexistenz des Originalgebäudes fast ungeschehen

11 Vgl. International Charter for the Conservation and Restoration of Monuments and Sites (The Venice Charter) mitsamt den offiziellen Übersetzungen in zahlreichen Sprachen.

[www.international.icomos.org/venicecharter2004/index.html](http://www.international.icomos.org/venicecharter2004/index.html) (Letzter Aufruf: 13.07.2013).

Im Englischen lauten die zitierten Formulierungen „living witnesses of [...] age-old traditions“, „common heritage“, „safeguarding for future generations“.

12 Ebd., Art. 11: „The valid contribution of all periods to the building of a monument must be respected, since unity of style is not the aim of a restoration.“

13 Ebd.: „the superimposed work of different periods“.

14 Ebd., Art. 15: „Ruins must be maintained [...]“.

15 Vgl. ebd., Art. 7: „The moving of all or part of a monument cannot be allowed except where the safeguarding of that monument demands it [...]“.

16 Vgl. ebd., Art. 15: „Only anastylosis, that is to say, the reassembling of existing but dismembered parts can be permitted.“



machen können. Zum Beispiel wird seit der Rekonstruktion der im Zweiten Weltkrieg zerstörten Dresdner Frauenkirche immer wieder die Heilung des Stadtbilds, vor allem der Elbsilhouette, betont. Bei den Wiederaufbauarbeiten, die zwischen 1994 und 2005 große mediale Aufmerksamkeit erzielten, entstand trotz oder gerade wegen der avancierten Rekonstruktionstechnik der Eindruck eines gleichsam selbsttätigen Auferstehens aus Ruinen, umso mehr, als es sich um einen sogenannten archäologischen Wiederaufbau handelte, der so weit möglich dem Prinzip der Anastylose folgte, also die noch vorhandenen Originalsteine verwendete. Für viele dieser Steine ließ sich aufgrund ihrer Lage in dem jahrzehntelang an Ort und Stelle gebliebenen Trümmerhaufen sowie durch den Einsatz avancierter Geoinformationssysteme der ursprüngliche Platz im Gebäude errechnen.<sup>17</sup>

So gesehen, müssen Restaurierungen heute nicht mehr als das genaue Gegenteil der Konservierung gelten, sofern ihre Initiatoren im genauen Wissen um die Logik und Praxis des Denkmalschutzes vorgehen. Umso ausgreifender, ja totaler wird somit aber der Imperativ der Erhaltung des in Monumenten gebundenen Kulturerbes, ob es sich nun um die archäologische Offenlegung von Überbauungen handelt oder um die Restaurierung kompletter Gebäude. In der Engführung von Erbe und Erhaltung wird der je gegenwärtige Umgang mit dem Ererbten oft geradezu zu einer Unmöglichkeit. Kulturdenkmäler sollen zwar materielle Zeugnisse historischer Prozesse sein – in einem durchaus differenzierten Geschichtsverständnis, das die Komplexität und Heterogenität jener Prozesse berücksichtigt –, wenn aber einmal die Kennzeichnung einer Stätte als Denkmal wirksam geworden ist, dann soll damit auch die Geschichte aufgehört haben: Jede weitere Veränderung wird als unangemessen, zerstörerisch und historisch falsch bewertet.

Für das Konzept des kulturellen Erbes ist das ein nicht zu unterschätzendes Problem. Seine Tragweite wird deutlich, wenn man sich klarmacht, welche große Rolle für die Entstehung dieses Konzepts der Begriff der Aneignung gespielt hat – also die komplementäre Vorstellung zum rein passivischen Empfang dessen, was einem von den Vorfahren vermacht wurde. Aneignung, Erwerb, Arbeit – das waren seit dem frühen 19. Jahrhundert die entscheidenden Kategorien, um einem bloßen Determinismus oder Fatalismus des Erbes entgegen-

17 Vgl. Wenzel, Fritz (Hrsg.): Berichte vom Wiederaufbau der Frauenkirche zu Dresden. Konstruktion des Steinbaus und Integration der Ruine. Karlsruhe 2007.

zuwirken und den aktiven Anteil vom Erblasser auf die Erben zu verschieben. Diese Überlegung findet sich sowohl in bürgerlich-liberalen als auch in kommunistisch-sozialistischen Reflexionen über das kulturelle Erbe.<sup>18</sup> Ihre einschlägig geschichtsphilosophische Interpretation liest man bereits in Hegels Heidelberger Antrittsvorlesung von 1817:

„Der Besitz an selbstbewußter Vernünftigkeit, welcher uns, der jetzigen Welt angehört, ist nicht unmittelbar entstanden und nur aus dem Boden der Gegenwart gewachsen, sondern es ist dies wesentlich in ihm, eine Erbschaft und näher das *Resultat* der Arbeit, und zwar der Arbeit aller vorhergegangenen Generationen des Menschengeschlechts zu sein.“<sup>19</sup>

Mit dem Konzept der Arbeit soll also von vornherein der Vorstellung vorgebeugt werden, bei jenem ererbten Besitz handle es sich um einen gesicherten Vorrat, den es nur zu verwalten gelte. Wenn weiter vom „mächtige[n] Strom“ der Tradition und vom tätigen Charakter des Weltgeistes die Rede ist, dann bleibt auch hier die Vorstellung einer Arbeit am Erbe leitend:

„Die Tat hat einen vorhandenen Stoff zu ihrer Voraussetzung, auf welchen sie gerichtet ist und den sie nicht etwa bloß vermehrt, durch hinzugefügtes Material verbreitert, sondern wesentlich *bearbeitet* und *umbildet*. Dies Erben ist zugleich Empfangen und Antreten der Erbschaft; und zugleich wird sie zu einem Stoffe herabgesetzt, der vom Geiste metamorphosiert wird. Das Empfangene ist auf diese Weise verändert und bereichert worden und zugleich erhalten.“<sup>20</sup>

Der bloße Erhalt von Monumenten – Inbegriff dessen, was heute als Kulturerbe fungiert – wäre demnach bloß die passive Seite des Vorgangs „Empfangen und Antreten der Erbschaft“. Damit hätte man aber den Anteil der Arbeit noch gar nicht in den Blick genommen. Das Erbe als Resultat von Arbeit zu verstehen

18 Vgl. dazu Willer, Stefan: *Erbfälle. Theorie und Praxis kultureller Übertragung in der Moderne*. München 2013 (in Vorbereitung).

19 Hegel, Georg Wilhelm Friedrich: *Vorlesungen über die Geschichte der Philosophie I, Einleitung* (Heidelberger Niederschrift). In: ders.: *Werke*, Bd. 18, hrsg. von Eva Moldenhauer und Karl Markus Michel. Frankfurt/M. 1986, S. 21.

20 Ebd., S. 21f.

heißt nach Hegel, es weiter zu bearbeiten, zu „metamorphosieren“. Zweifellos stellt sich die von Hegel für den philosophischen Geist formulierte Aufgabe anders, wenn von materiellen Überlieferungen die Rede ist; zu nivellieren oder ignorieren ist sie aber nicht.

Die Administratoren des UNESCO-Welterbezentrums in Paris wussten offenkundig um das Problem der Monumentalisierung und Stillstellung des Erbes, als sie im Jahr 2003 mit der Konvention über Intangible Cultural Heritage einen zweiten Teil des Welterbeprogramms auflegten. Seitdem werden – in Analogie zur Liste des Natur- und Kultur-Welterbes – mündliche, theatralische und musikalische Überlieferungen, Rituale, Feste, traditionelles Wissen und Handwerk aus der ganzen Welt aufgelistet. Dieses nicht in einzelnen Stätten verortbare Erbe wird als „Haupttriebfeder kultureller Vielfalt“ und als „Garantie für nachhaltige Entwicklung“ verstanden.<sup>21</sup>

Wie soll ein solches Erbe gepflegt, wie überhaupt definiert werden? Die Frage lässt sich nicht leicht beantworten, weil schon die Semantik des Schlüsselworts *intangible* schwierig zu fassen ist. Mit Blick auf das zugrundeliegende lateinische Verb *tangere*, berühren, ergibt sich eine interessante Doppeldeutigkeit. Spricht man in wörtlicher Übersetzung von einem unberührbaren Erbe, so kann das zum einen heißen, dass man es nicht berühren *kann*, weil es immateriell und daher ungreifbar ist. Immaterielles Kulturerbe ist denn auch die gängige deutsche Version von *Intangible Cultural Heritage*. Zum anderen schwingt hier aber die Bedeutung mit, dass man dieses Erbe nicht berühren *darf*, weil es tabu ist. Der auch im Englischen vorhandene Doppelsinn der Benennung ist wohl kaum zufällig unterlaufen. Vielmehr liegt den Proklamationen der UNESCO zu diesem Thema eine wesentliche Ambivalenz zugrunde. Sie bleibt allerdings fast gänzlich unausgesprochen, wie der definitorische Artikel der *Convention for the Safeguarding of Intangible Cultural Heritage* zeigt: „Dieses immaterielle / unberührbare, von Generation zu Generation überlieferte Kulturerbe wird fortwährend von

21 *Convention for the Safeguarding of Intangible Cultural Heritage*, Präambel: „the importance of the intangible cultural heritage as a mainspring of cultural diversity and a guarantee of sustainable development“. Zu Praktiken und Problemen im Umfeld dieses UNESCO-Programms vgl. Bernecker, Roland: Vom immateriellen Kulturerbe zur kulturellen Vielfalt. Genese eines neuen UNESCO-Übereinkommens. In: Albert, Marie-Theres (Hrsg.): *Perspektiven des Welterbes*. Frankfurt/M. 2006, S. 98-108; Jadé, Mariannick: *Le patrimoine immatériel. Perspectives d'interprétation du concept de patrimoine*, Paris 2006.

Gesellschaften und Gruppen neugeschaffen.“<sup>22</sup> Dass es zwischen Überlieferung und Neuschöpfung zu Konflikten kommen könnte, wird gar nicht erst ausformuliert; eher scheint man stillschweigend vorauszusetzen, dass Innovation mit der Pflege des Bestehenden in Übereinstimmung zu bringen ist – oder gar: darin aufgeht.

Auch in diesem Programm liegt die selbstgestellte Aufgabe der Unesco im safeguarding, im Sicherstellen des Erbes. Die Flüchtigkeit und die Variabilität des Immateriellen haben deshalb gegenüber dem Imperativ der unverfälschten Erhaltung zurückzustehen. Dazu passt die Proklamation der „Meisterwerke des mündlichen und immateriellen Erbes der Menschheit“, einer vorbereitenden und flankierenden Maßnahme zur Konvention von 2003.<sup>23</sup> Mit der Idee der Meisterwerke – noch dazu der kanonisch aufgelisteten – wird Immaterialität von vornherein auf das Abgeschlossene und Vorbildliche festgelegt und somit eigentlich auch schon hintertrieben. In jedem Fall verstärkt sich der Aspekt des Unberührbaren: Ein Meisterwerk ist nichts, was verändert oder dem etwas hinzugefügt werden sollte, und es verlangt auch nicht nach fortgesetzter Innovation.

Auch hier stellt sich Konservierung als Problem dar, mit interessanten, teils sogar kuriosen Auswirkungen auf lokaler Ebene. Wenige Jahre nach Inkrafttreten des Programms veröffentlichte der Ethnologe Markus Tauschek einen Aufsatz, in dem er am Beispiel des Karnevals in der belgischen Kleinstadt Binche – seit 2003 zum immateriellen Meisterwerk deklariert – einen „Unesco-Effekt“ beschrieb.<sup>24</sup> Teil dieses Effekts ist eine bezeichnende Unsicherheit der Akteure über den neuartigen Status ihrer kulturellen Praxis. Konkret geht es dabei um die Nationalität der Mitglieder in den Karnevalsgesellschaften. Traditionelle Aufnahmebedingung ist die belgische Staatsbürgerschaft, wobei diese Regel fak-

22 Convention for the Safeguarding of Intangible Cultural Heritage, I. General provisions, Art.

2 Definitions: „This intangible cultural heritage, transmitted from generation to generation, is constantly recreated by communities and groups.“

23 Vgl. Kuutma, Kristin: The politics of contested representation. UNESCO and the masterpieces of intangible cultural heritage. In: Hemme, Dorothee/Tauschek, Markus/Bendix, Regina (Hrsg.): Prädikat „Heritage“. Wertschöpfungen aus kulturellen Ressourcen. Berlin, Münster 2007, S. 177-195.

24 Tauschek, Markus: „Plus oultre“ – Welterbe und kein Ende? In: Hemme, Dorothee/Tauschek, Markus/Bendix, Regina (Hrsg.): Prädikat „Heritage“. Wertschöpfungen aus kulturellen Ressourcen. Berlin, Münster 2007, S. 197-224, S. 214.

tisch vielfach unterlaufen wird. Nach der Aufnahme in die UNESCO-Liste begannen sich aber die Karnevalisten über die Frage zu entzweien, ob die Reinheit und Authentizität der Festkultur durch eine verschärfte Einhaltung der Aufnahmekriterien wiederhergestellt werden solle oder ob nicht gerade die Kür durch die UNESCO die Aufforderung zu mehr Offenheit und Integration bedeute. Das Fazit des Ethnologen lautet:

„Wenngleich die Argumente konträr sind [...] so liegt doch beiden ein und dieselbe Wahrnehmung zugrunde: Die Entscheidungen dürfen nach der Vorstellung der lokalen Akteure nicht mehr allein nur vor Ort getroffen werden. Sie befinden sich nun in einem globalen Bezugsrahmen.“<sup>25</sup>

Angesichts dieses maximal weiten Bezugsrahmens ist die Frage nach dem kulturellen Erbe auch immer eine Frage nach der Beziehung zwischen der Grenzenlosigkeit des programmatischen Anspruchs und der konkreten Begrenzung im Einzelnen. Dabei geht es darum, das politische Ziel nach dem Einschluss möglichst vielfältiger, vielgestaltiger Zeugnisse menschlicher Kultur mit konkreten Praktiken des Ausschlusses ins Verhältnis zu setzen. Denn das Konzept World Heritage beruht seiner Logik und seiner Funktion nach darauf, dass die meisten Orte, Riten, Praktiken, Monumente und Dokumente der Welt *nicht* zum Erbe gehören, dass es vielmehr äußerst strikte Wertmaßstäbe für die Verleihung des Erbe-Status gibt. Das entscheidende Kriterium der UNESCO und der von ihr beauftragten Evaluierungsorganisationen ist das des Herausragens: Die einzelnen Bestandteile des Welterbes sind „outstanding“.<sup>26</sup> Als solche werden sie aus den normalen Verhältnissen, den ansonsten üblichen Ökonomien und Lebensbedingungen, hervorgehoben; sie werden im Wortsinne exemplarisch (von lat. *ex-imere*, herausnehmen).

Auf diese Weise entsteht ein Gefälle, ein Wertekontrast, zwischen dem Herausgenommenen und seiner Umgebung, die, wenn man so will, auf Normalhöhe verbleibt. Offenkundig steht dabei die numerische Begrenztheit der auf der Welterbeliste eingetragenen Stätten in enger Beziehung zu ihrer jeweiligen territorialen Eingrenzung. Wie sehr beides zusammenhängt – die Frage nach Einschluss und Ausschluss vonseiten der UNESCO und die Territorialität der

25 Ebd., S. 219.

26 Convention Concerning the Protection of the World Cultural and Natural Heritage, Art. 1 und 2.

gleichsam eingerahmten Stätten –, zeigt sich bei jeder landschafts-, stadt- und verkehrsplanerischen Maßnahme, die die konkreten territorialen Grenzen zwischen Erbe und Nicht-Erbe berührt.<sup>27</sup>

Daraus entstehen geradezu regelförmig Konflikte. Sie entzündeten sich etwa an tatsächlich vorgenommenen Veränderungen einer Kulturlandschaft, wie im Fall der Waldschlösschenbrücke in Dresden: Das 2004 in die Liste aufgenommene Dresdner Elbtal stand wegen des geplanten Brückenbaus durch die Elbauen seit der Jahressitzung des Welterbe-Komitees 2006 unter dem offiziellen Vorbehalt des World Heritage in Danger; im Sommer 2009 wurde die Stätte definitiv von der Liste gestrichen, als der durch Bürgerentscheid befürwortete (allerdings höchst umstrittene) Bau nicht gestoppt wurde. Mitunter betreffen solche Konflikte aber auch die bloße Beeinträchtigung eines bestimmten Blickwinkels auf ein bestimmtes Gebäude: So wurde der Kölner Dom, Weltkulturerbe seit 1996, im Jahr 2004 auf die Gefährdungsliste gesetzt, weil am gegenüberliegenden Rheinufer in Köln-Deutz mehrere Hochhäuser gebaut werden sollten; jedoch gilt diese Gefahr seit 2006 als offiziell gebannt, nachdem die Höhe der Häuser begrenzt worden war.

Ein weiteres Beispiel dafür, wie komplex die Bestimmung des Herausragenden sein kann, ist die Welterbestätte *Frontiers of the Roman Empire*, die derzeit aus drei Grenzgebilden von jeweils ziemlicher Ausdehnung besteht: dem Hadrianwall in Nordengland, dem Antoninuswall in Schottland und dem obergermanisch-raetischen Limes in Südwestdeutschland. Letzterer wurde im Juli 2005 als 31. deutsche Stätte in die UNESCO-Welterbeliste aufgenommen. Während die beiden Wallanlagen in Großbritannien in weiten Teilen erhalten sind, handelt es sich beim Limes nur mehr um die Überreste der römischen Grenzbefestigungen, wie sie sich auf einer Strecke von über 500 Kilometer zwischen Koblenz und Regensburg finden. Auffallendes Charakteristikum dieser Stätte ist ihre weitgehende Unsichtbarkeit, denn der Grenzwall ist ein sogenanntes Bodendenkmal, während es sich bei den sichtbaren Befestigungsanlagen zum Großteil um Nachbauten aus dem 19. Jahrhundert handelt.

Nun könnte man erwägen, den deutschen Limes eindrucksvoller zu gestalten, sodass er mit den britischen *Frontiers*-Stätten mithalten könnte. Doch die Vorstellung einer womöglich vollständigen Rekonstruktion des gesamten Limes würde nicht nur den heutigen Stand archäologischer und althistorischer Forschungen

27 Vgl. Pfeifle, Florian: UNESCO-Weltkulturerbe. Vom globalen Völkerrecht zur lokalen Infrastrukturplanung. Köln 2010.

unterbieten, in der prinzipiell die Einschätzung in Frage gestellt wird, das Römische Reich habe über klar definierte Grenzen verfügt (stattdessen versteht man den Limes heute eher als „Kontaktzone“<sup>28</sup>). Auch ohne solche historischen Einschränkungen widerspräche eine Rekonstruktion der bereits zitierten modernen Denkmalschutz-Prämisse der Erhaltung von Ruinen – die durchaus auf unsichtbar gewordene, mit dem Boden verwachsene Ruinen auszudehnen ist. Wie aber steht es mit den bereits existierenden Rekonstruktionen, etwa dem aus wilhelminischer Zeit stammenden Nachbau des Kastells Saalburg? Sie sind, in Ergänzung zum eigentlichen Bodendenkmal, integraler Bestandteil der Welterbestätte, obwohl sie nach heutigem Kenntnisstand unzureichende, historistische Rückprojektionen sind. Aber gerade deshalb lassen sie sich ihrerseits historisieren und genau aus diesem Grund ins Erbe eingemeinden, wie der International Council on Monuments and Sites (ICOMOS), der im Auftrag der UNESCO die Aufnahme des Limes überprüfte, in seinem Bericht festhielt:

„ICOMOS considers that the Roman remains need to be differentiated from reconstructions. Whereas reconstructions carried out in the 19th century can be said to now have a certain historical interest, it does not consider that reconstructions carried out since the inception of the Venice Charter can be considered authentic or of sufficient value as to be included in the nomination.“<sup>29</sup>

Das heutige historische Interesse schließt also fraglos das 19. Jahrhundert mit ein; somit können die während dieser Zeit vorgenommenen Eingriffe in die antike Stätte ihrerseits ins Erbe eingeschlossen werden. Zugleich wird eine klare zeitliche Grenze gesetzt, jenseits derer die Erbwürdigkeit von Rekonstruktionen endet. Diese Grenze ist – bemerkenswert genug – die Implementierung der Charta von Venedig. Was seitdem, also seit der Kodifizierung des heute herrschenden Regimes von Denkmalschutz, ein-, an- oder nachgebaut wurde, ist gemäß der Diktion des Prüfberichts unauthentisch, besitzt keinen hinreichenden Wert und muss aus dem Erbe ausgeschlossen werden. Dies ist eine für die UNESCO und die von ihr beauftragten Institutionen höchst charakteristische

28 So der programmatische Titel einer aktuellen Einführung: Waldherr, Gerhard: Der Limes. Kontaktzone zwischen den Kulturen. Stuttgart 2009.

29 Vgl. [http://whc.unesco.org/archive/advisory\\_body\\_evaluation/430ter.pdf](http://whc.unesco.org/archive/advisory_body_evaluation/430ter.pdf) (Letzter Aufruf: 13.07.2013), S. 167.

Art der Differenzierung. Eine gewisse Authentizität von Rekonstruktionen wird bis zu einem bestimmten Grad zugestanden, doch alle neueren Veränderungen werden als Bedrohung, ja Zerstörung gewertet: „any further reconstructions [ ... ] could risk putting the site under threat.“<sup>30</sup>

Der Befund der Bedrohung, der einmal am Ausgangspunkt des Konzepts Welterbe stand – als kulturell produktiver Wunsch, das Überlieferte vor konkret materieller Zerstörung zu retten –, erscheint heute oft genug als administrativer, ja bürokratischer Schutzmechanismus. Dieses mechanische Verständnis von maximal langfristiger, potenziell endloser Erhaltung des einmal als erhaltenswert Deklarierten stellt schließlich das Konzept Erbe als solches in Frage, weil somit das konstitutive Spannungsverhältnis von Erbe als Nachlass und Erban-eignung bis zur Wirkungslosigkeit reduziert wird.

### *Kritik der Nachhaltigkeit*

Die derzeit wirksame Engführung, wenn nicht sogar Identifizierung von Erhaltung, Nachhaltigkeit und Erbe ist zugleich eine eigentümliche Art der Zukunftsverwaltung. Am besten wäre sie wohl mit dem grammatischen Modell des Futur II, der vollendeten Zukunft, zu beschreiben: Die jeweils gegenwärtige Zukunftsplanung soll vorwegnehmen, wie die kommenden Generationen mit ihrer Vergangenheit umgehen werden – wie sie also das aufnehmen, was ihnen einst hinterlassen wird. Das Rezeptionsverhalten zukünftiger Populationen wird damit nicht nur antizipiert, sondern festgelegt, in Form von konkreten Handlungsanweisungen, wie sie mit den ihnen übertragenen kulturellen Werten umzugehen haben. Damit wird eine künftige kulturelle Weltgesellschaft schon jetzt zur Erbgemeinschaft, geradezu zur Gemeinschaft der Epigonen gemacht. Das ist kein bloßes Problem von Denkmal- und Brauchtumspflege und lässt sich auch nicht auf Kulturpolitik im engeren Verständnis beschränken. Vielmehr offenbart sich Nachhaltigkeit als genauer Gegenbegriff zu dem der Prognose, also des Vorauswissens. Im grammatisch-semantischen Gegensatz von „nach“ und „voraus“ artikuliert sich eine gedankliche Spannung im Inneren des Programms einer zukunftsbezüglichen Politik.

Diese Spannung ist sowohl erkenntnistheoretischer als auch zeittheoretischer Natur. Was die Erkennbarkeit der Zukunft betrifft, so lässt sich die Idee der

30 Ebd.



Nachhaltigkeit als Versuch verstehen, der essenziellen Ungewissheit entgegenzuarbeiten, die doch an der Basis allen Zukunftswissens und Zukunftsdenkens steht. Während dieses nicht nur ergebnisoffen sein, sondern auch zukünftige Kontingenzen mit in Rechnung stellen muss, versucht man sich heute im Interesse der Generationengerechtigkeit und Nachhaltigkeit an Festlegungen zukünftiger Umgangsweisen mit dem, was, von heute aus betrachtet, einmal Hinterlassenschaft und Erbe sein wird. Was die Zeitlichkeit der Zukunft betrifft, geht es in den Verfügungen über das kulturelle Erbe, die die UNESCO-Konventionen im Namen „der Welt“ treffen, um eine Art der Vergangenheitsrepräsentation, die geradezu auf eine Stillstellung der historischen Zeit hinauszulaufen scheint. Aufgabe der Kulturerbepflege ist es, die Monumente und auch die immateriellen Erbstücke in dem Zustand zu erhalten, in den sie die Geschichte versetzt hat. Wenn aber einmal diese Erbpflege begonnen hat, dann soll das historische Vergehen der Zeit aufhören. Jede weitere Veränderung wird als unangemessen, zerstörerisch und historisch falsch bewertet.

Dieses auf Dauer gestellte Moratorium hat – nach den Vorgaben der UNESCO, aber vielfach auch in der gängigen Praxis regionaler und lokaler, öffentlicher und privater Kulturinstitutionen – testamentarischen Charakter: Es betrifft Veränderungen, die möglicherweise einmal von zukünftigen Generationen vorgenommen werden könnten; es will jene Generationen also dem Zeitregime der nachhaltigen Zukunft unterwerfen, ohne zu bedenken, dass dieses Zeitregime eben ein *heutiges* ist, über dessen zukünftige Wirksamkeit keine verlässlichen Aussagen gemacht werden können. Angesichts des enormen Erfolgs und der großen sozio-kulturellen Wirksamkeit von World Heritage ist es an der Zeit, wieder stärker an einer Dialektik des Erbes zu arbeiten und diese Arbeit als eine Kritik der Nachhaltigkeit zu betreiben.

Dieser Beitrag ist die geringfügig überarbeitete Niederschrift des Vortrags, den ich am 31. Mai 2013 bei der Abschlussveranstaltung der 8. Initiative „Nachhaltigkeit in der Digitalen Welt“ des Internet und Gesellschaft Collaboratory e. V. im Jüdischen Museum Berlin gehalten habe.





Hochleistungsrechner, Konrad-Zuse-Zentrum für Informationstechnik Berlin

Der Streit um die Zukunft:  
Auswahl, Beständigkeit und  
Vernetzung



# Nachhaltigkeit und Semantic Web

FELIX SASAKI

## *Einführung*

Dieser Aufsatz betrachtet das Semantic Web aus der Sicht von Nachhaltigkeit. Zunächst werden die Grundlagen des Semantic Web behandelt. Anschließend werden verschiedene Aspekte von Nachhaltigkeit diskutiert: Mehrsprachigkeit, Adäquatheit sogenannter Aussagen im Semantic Web und Identifikatoren. Anschließend wird ein Kernbereich gegenwärtiger Aktivitäten im Semantic Web vorgestellt: die Entwicklung und Verknüpfung von Vokabularen. Diese Aktivität geht über das Semantic Web aus technischer Sicht hinaus und stellt die eigentliche Herausforderung von Nachhaltigkeit maschinenlesbarer Informationen im Web dar.

## *Was ist das Semantic Web?*

Der Begriff Semantic Web wurde bereits vor über zwölf Jahren von Tim Berners-Lee geprägt.<sup>1</sup> Er ist ein schillernder Begriff, der mit der Zeit immer wieder neu ausgelegt wurde. Ursprünglich wurden darunter technologische Grundlagen für eine weltumspannende, intelligente Wissensbasis verstanden. Die Tradition der künstlichen Intelligenz und ihre technischen Grundlagen spielen dementsprechend eine maßgebliche Rolle: formale Logik und Inferenzmechanismen, um aus bestehenden Informationen neues Wissen zu generieren. Die Vision ist ein

1 Berners-Lee, Tim/Hendler, James/Lassila, Ora: The Semantic Web. In: Scientific American 284/5 (2001), S. 34.

globales dezentrales Frage-Antwort-System. Der Mensch soll Informationen im Web über reine Volltextsuche hinaus und letztlich in natürlicher Sprache erfragen können.

Diese Vision wurde zunächst nicht Realität. Die Sicht auf das Semantic Web hat sich geändert und dem Web für Menschen angenähert. Was aber ist das Web für Menschen? Es besteht aus zwei grundlegenden Bausteinen:

- Repräsentationen von Inhalten. In den Anfangstagen des Web waren Inhalte zumeist statische HTML-Textseiten. Inzwischen sind Inhalte zunehmend multimedial, interaktiv, dynamisch und von Nutzern generiert.
- Verlinkungen zwischen Inhalten. In den Anfangstagen des Web dienten Links dazu, statistische Inhalte zwischen Webseiten zu verbinden. Inzwischen ist es üblich, Mailkonversationen, Tweets und andere Interaktionen im sozialen Web zu referenzieren.<sup>2</sup>

Obwohl sich die Natur des Web in der beschriebenen Weise geändert hat, sind die beiden Hauptbestandteile Inhalte + Verlinkungen zentral geblieben. Generalisiert kann man über Links auch als Identifikatoren oder URIs (Universal Resource Identifier) sprechen.<sup>3</sup> Die folgende Abbildung demonstriert die Rolle von Identifikatoren im Web für Menschen.

```
<p>All content on this site is licensed under  
a  
<a href="http://creativecommons.org/licenses/by-sa/3.0/"  
>Creative Commons License</a>.</p>
```

- 2 Die W3C-Mailinglisten sind ein Bereich, in dem ausführlich von derartigen Referenzen Gebrauch gemacht wird. Ein Beispiel ist <http://lists.w3.org/Archives/Public/public-multilingualweb-lt/2013Jul/0005.html> (Letzter Aufruf: 15.08.2013). Hier werden unter anderem Mailthreads, Transkriptionen von Online-Meetings und Wikiinhalte via URIs referenziert.
- 3 Jacobs, Ian/Walsh, Norman: Architecture of the World Wide Web, Volume One. 2004. [www.w3.org/TR/2004/REC-webarch-20041215/](http://www.w3.org/TR/2004/REC-webarch-20041215/) (Letzter Aufruf: 15.08.2013).

Der Paragraph dient dazu, die Lizenz zu verdeutlichen, unter der eine Webseite genutzt werden kann. Der Link führt den menschlichen Nutzer auf die Dokumentation der Creative Commons CC By 3.0 Lizenz. Zugleich identifiziert der Link die Lizenz: Ein menschlicher Nutzer erkennt anhand des Links die Lizenz, ohne die Inhalte der Webseite <http://creativecommons.org/licenses/by/3.0/> aufrufen zu müssen.

Mit dieser Form von Inhalten lässt sich das Web durchsuchen. Eine Anfrage wie „Finde alle Inhalte, welche auf die Webadresse der Creative Commons Lizenz verweisen“ kann von einer textbasierten Suchmaschine beantwortet werden. Die Suchmaschine muss nur nach dem Creative Commons Link in Webseiten suchen. Nicht möglich ist aber eine Anfrage wie die folgende: „Finde alle Inhalte, welche unter der Creative Commons Lizenz stehen.“ Dazu müsste die Intention eindeutig sein, mit welcher der Verweis auf die URI gesetzt wurde: Dient der Link dazu, Lizenzinformation über den Ausgangspunkt des Verweises zu bieten, oder ist er beispielsweise nur ein Verweis, um die Creative Commons Lizenz an sich zu dokumentieren?

Die Beantwortung dieser Frage lässt sich mit den Mechanismen des Web für Menschen – Inhalte + Links / Identifikatoren – nicht automatisieren. Das Semantic Web, wie es heute hauptsächlich verstanden wird, füllt diese Lücke. Es gibt eine Antwort in Form *maschinenlesbarer* Identifikatoren. Dies verdeutlicht die folgende Abbildung.

```
<p>All content on this site is licensed under
a
<a property="http://creativecommons.org/ns#license"
href="http://creativecommons.org/licenses/by-sa/3.0/"
>Creative Commons License</a>.</p>
```

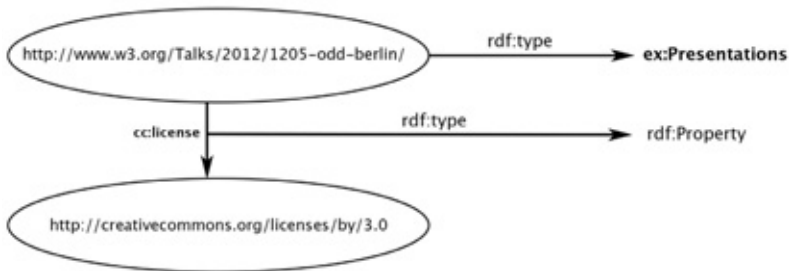
Im Unterschied zum vorherigen Beispiel wird hier die Intention oder Art des Links auf maschinenlesbare Weise ausgedrückt. Die URI <http://creativecommons.org/ns#license> drückt aus, dass das Linkziel Lizenzinformationen über die Webseite bereithält. Die Suchanfrage „Finde alle Inhalte, welche unter der Creative Commons Lizenz stehen“ ist nun möglich: Es müssen alle Links gefunden werden, welche auf die Creative Commons Lizenz verweisen und zugleich als Lizenzlink typisiert sind.



*Semantic Web schreiben: Aussagen, zugrundeliegende Vokabulare und Schreibweisen*

Die dem Semantic Web zugrundeliegende Technologie heißt RDF (Resource Description Framework).<sup>4</sup> Mit RDF lässt sich die Verlinkungsstruktur in maschinenlesbarer Weise wie beschrieben ausdrücken. Mit RDF werden maschinenlesbare Aussagen gemacht. Gegenstand bzw. Subjekt der Aussage ist z. B. die URI der entsprechenden Webseite. Thema bzw. Prädikat der Aussage ist z. B. die URI <http://creativecommons.org/ns#license>. Das Objekt der Aussage ist im Beispiel die URI für die Creative Commons CC BY Lizenz. Da RDF-Aussagen immer die drei Bestandteile Subjekt (= Ausgangspunkt), Prädikat (= typisierter Link) und Objekt (= Linkziel oder Wert) haben, spricht man auch von RDF-Tripeln.

RDF-Aussagen wären wenig nützlich, wenn es beliebige Tripel gäbe. Im Beispiel ist es sowohl für Autoren von Tripeln als auch für Suchanfragen hilfreich, dass die Lizenz-URIs, d. h. Prädikat und Objekt, wohl bekannt sind. Deshalb bietet das Semantic Web die Möglichkeit, Vokabulare zu definieren, dazu die folgende Abbildung:



Hier wird angenommen, dass es ein Vokabular für Präsentationen gibt. Die Subjekt-URI entstammt diesem Vokabular; man sagt in RDF-Terminologie: Sie ist eine Instanz der Klasse von Präsentationen. Dies drückt das Tripel mit dem in

4 Vgl. Manola, Frank/Miller, Eric: RDF Primer. 2004. [www.w3.org/TR/2004/REC-rdf-primer-20040210/](http://www.w3.org/TR/2004/REC-rdf-primer-20040210/) (Letzter Aufruf: 15.08.2013).

RDF vordefinierten Prädikat `rdf:type` im oberen Bereich der Abbildung aus.

`rdf:type` wird auch in einem Tripel genutzt, bei dem das Prädikat `cc:license` (die Kurzform der Prädikat-URI aus dem ersten Beispiel) das Subjekt ist. Dieses Tripel drückt aus, dass `cc:license` eine RDF-Eigenschaft (property) ist, die in Aussagen verwendet werden kann.

Klassen und Eigenschaften sind das Rückgrat, auf dem Vokabulare im Semantic Web beruhen. In den RDF-Technologien werden sie in RDF-Schema und OWL (Web Ontology Language) definiert. OWL erlaubt es, im Vergleich zu RDF-Schema, komplexere Inferenzen zu ziehen. So lässt sich mit OWL eine Eigenschaft `istKindVon` als symmetrisch zu einer Eigenschaft `istVaterVon` definieren. Ein Tripel wie `a istKindVon b.` führt dann zur Inferenz `b istVaterVon a.`

RDF-Aussagen und auch Vokabulardefinitionen lassen sich in Webseiten einbetten.<sup>5</sup> Das erste Beispiel hat diese Art der Schreibung von Semantic Web demonstriert. Die Einbettung in Webseiten hat den Vorteil, dass Informationen des Web für Menschen und des Semantic Web eng miteinander verknüpft sind. In manchen Anwendungsfällen bietet es sich jedoch an, RDF-Aussagen und Vokabulare unabhängig von Webseiten zu schreiben. Hierfür werden verschiedene Schreibweisen bereitgestellt, u. a. RDF/XML oder Turtle.<sup>6</sup>

### *Herausforderungen zur Nachhaltigkeit des Semantic Web*

Im Folgenden werden Herausforderungen diskutiert, welche für eine nachhaltige Nutzung des Semantic Web gelöst werden müssen. Teilweise sind diese Herausforderungen aus dem Web für Menschen wohl bekannt. Andere Aspekte sind spezifisch für das Semantic Web.

### *Semantic Web und Mehrsprachigkeit*

Die URI für das Prädikat `cc:license` ist auch für den menschlichen Leser verständlich – vorausgesetzt er versteht Englisch. Man könnte annehmen, dass

5 Adida, Ben/Herman, Ivan/Sporny, Manu/Birbeck, Mark: RDFa 1.1 Primer. 2012. [www.w3.org/TR/2012/NOTE-rdfa-primer-20120607/](http://www.w3.org/TR/2012/NOTE-rdfa-primer-20120607/) (Letzter Aufruf: 15.08.2013).

6 Prud'hommeaux, Eric/Carothers, Gavin: Turtle – Terse RDF Triple Language. 2013. [www.w3.org/TR/2013/CR-turtle-20130219/](http://www.w3.org/TR/2013/CR-turtle-20130219/) (Letzter Aufruf: 15.08.2013).

im Semantic Web eine Lingua franca ausreichen würde. Wenn man sich jedoch die Anforderung einer maschinenlesbaren Repräsentation von menschlichem Wissen im Web vor Augen führt, werden die Grenzen einer Lingua franca schnell sichtbar. Man denke beispielsweise an Semantic-Web-Vokabulare, um Terme aus einer Fachdomäne wie dem juristischen Bereich zu definieren. Selbst dem juristischen Laien wird klar, dass eine sprach- und länderunabhängige Definition unmöglich ist.

Das Semantic Web bietet technische Mittel um URIs zu erstellen, die für Menschen lesbar sind. Sprachtags identifizieren die verwendete Sprache. Eine umfassende Analyse von Ell verdeutlicht, dass Mehrsprachigkeit im Semantic Web nicht verbreitet ist.<sup>7</sup> Nur 4,78 % der Beschreibungen haben einen Sprachtag. Nur 0,7 % besitzen mehr als einen Sprachtag. Die am häufigsten verwendeten Sprachen sind Englisch, Deutsch, Französisch und Italienisch.

Was hat nun die Problematik der Mehrsprachigkeit mit Nachhaltigkeit zu tun? Sprache ist ein Teil unseres kulturellen Erbes. Dies trifft auch auf den digitalen Bereich zu und hier sowohl auf das Semantic Web als auch auf das Web für Menschen. Studien belegen, dass im Web unterrepräsentierte Sprachen sich in einem Teufelskreis befinden:<sup>8</sup> Eine geringe Menge an Inhalten bedeutet wenig Anreiz, weitere Inhalte zu erstellen. Letztlich droht das „digitale Aussterben“ der Sprache. Ausführlich werden dieses Thema und der gegenwärtige Stand europäischer Sprachen im Web und bezüglich sprachtechnologischer Unterstützung in den META-NET White Paper diskutiert.<sup>9</sup>

Automatisierte Ansätze können helfen, diese Situation zu verbessern. Ein berühmtes Beispiel ist die thailändische Wikipedia. Sie wurde zunächst mittels maschineller Übersetzung erstellt und von regionalen Wikipedianern verbessert.

7. Ell, Basil/Vrandečić, Denny/Simperl, Elena: Labels in the Web of Data. In: Aroyo, Lora/Welty, Chris/Alani, Harith/Taylor, Jamie/Bernstein, Abraham/Kagal, Lalana/Fridman Noy, Natasha/Blomqvist, Eva (Hrsg.): *The Semantic Web – ISWC 2011 – 10th International Semantic Web Conference*, Bonn, Germany, October 23-27, 2011. Proceedings, Part I. 2011, S. 162-176.
8. Kornai, András: *Language Death in the Digital Age*. Präsentation beim META-FORUM 2012. 2012. [www.meta-net.eu/events/meta-forum-2012/report](http://www.meta-net.eu/events/meta-forum-2012/report) (Letzter Aufruf: 15.08.2013).
9. Rehm, Georg/Uszkoreit, Hans: *META-NET White Paper Series — Europe’s Languages in the Digital Age*. 31 Bände. Heidelberg, New York, Dordrecht, London 2012. Vgl. auch die Pressemitteilung zur Veröffentlichung der White Paper unter [www.meta-net.eu/whitepapers/press-release](http://www.meta-net.eu/whitepapers/press-release) (Letzter Aufruf: 15.08.2013).

Die gleiche Möglichkeit bietet sich für das Semantic Web: Maschinelle Übersetzung kann mehrsprachige Beschreibungen im großen Umfang erzeugen. Wie auch bei der thailändischen Wikipedia ist aber ein Verständnis um die Grenzen der automatischen Verarbeitung und die Rolle von Nachbearbeitungsschritten wichtig.

*(Nachhaltige) Adäquatheit von Aussagen im Semantic Web*

Semantic Web umfasst Aussagen über Ressourcen. In zunehmendem Maße sind Ressourcen Webseiten selbst. So ist die Semantic-Web-Wissensquelle dbpedia eine umfangreiche Sammlung von Aussagen, die aus Wikipedia generiert wird.<sup>10</sup> Dies geschieht hauptsächlich auf der Basis von sogenannten Wikipedia-Info-boxen. Sie enthalten strukturierte Informationen, aus denen sich recht einfach RDF-Tripel generieren lassen. Die jeweilige Wikipedia-Seite dient als Subjekt. Listenartige Einträge in den Infoboxen (z. B. Geburtsdaten in Seiten über Personen) sind die Basis von Prädikaten und Objekten.

Die Nachhaltigkeit derart generierter Aussagen beruht hauptsächlich auf der Wikipedia-Community. Letztlich sind es Wikipedianer, welche die Beschreibungen in den Infoboxen und somit die Basis der Tripel erzeugen. Aus dieser Situation ergeben sich verschiedene Fragestellungen:

- Sind die Wikipedia-URIs für die jeweiligen Artikel stabil?
- Ist der Inhalt stabil?
- Sind die gemachten Aussagen adäquat?

Alle Fragen lassen sich nur mit einer automatischen Verarbeitung von Wikipedia beantworten. Die Stabilität von URIs als Subjekte (Wikipedia-Seiten-URIs) bzw. Prädikate oder Objekte (Teil der Infoboxen) ist dabei noch die kleinste Herausforderung. Zentral ist die Beziehung zwischen den URIs der Wikipedia-Seiten, den Infoboxen und dem eigentlichen Inhalt, dem Text des Wikipedia-Artikels. Was für Tripel drücken die in einem Artikel behandelten Entitäten (Personen, Orte, Ereignisse etc.), die Beziehungen zwischen Entitäten (z. B. „A ist mit B

<sup>10</sup> Bizer, Christian/Lehmann, Jens/Kobilarov, Georgi/Auer, Sören/Becker, Christian/Cyganiak, Richard/Hellmann, Sebastian: DBpedia – A Crystallization Point for the Web of Data. In: Journal of Web Semantics. Science, Services and Agents on the World Wide Web 7 (2009), S. 154–165.

verheiratet“) oder ein übergreifendes Thema des Wikipedia-Artikels aus? Sind diese Tripel kompatibel zu den Informationen, welche sich aus den Infoboxen ergeben?

Die Extraktion von Wissen aus den eigentlichen Wikipedia-Inhalten steckt noch in den Kinderschuhen.<sup>11</sup> Sie ist deshalb von großer Bedeutung, weil die Problematik in Wikipedia nur exemplarisch ist. Für die nachhaltige Nutzbarkeit von Informationen im Semantic Web gilt es, die Beziehung von den Semantic-Web-Aussagen zu den eigentlichen Inhalten skalierbar, d. h. weitestgehend automatisch, zu analysieren und zu verifizieren.

### *Systeme für Identifikatoren*

Bisher war in diesem Artikel durchweg vom Identifikatorensystem URI die Rede. Es gibt jedoch eine Vielzahl von Mechanismen, um Ressourcen zu identifizieren. Der Begriff Ressource hat dabei eine sehr weite Bedeutung. Die Identifikation kann sich auf physische Objekte beziehen (z. B. natürliche Personen), abstrakte Konzepte (z. B. politische Ideen) oder auch Informationsquellen im Web (z. B. eine Webseite zu Creative Commons).

Besonders im Bibliotheksbereich gibt es hohe Anforderungen an die Nachhaltigkeit von Identifikatoren. Nur so können kulturelle und wissenschaftliche Güter nachhaltig bewahrt werden. Das Web konnte diesen Anforderungen lange nicht genügen, man denke an das berüchtigte „404 not found“-Erlebnis, wenn eine eventuell langfristig genutzte Webseite plötzlich nicht mehr gepflegt wird. Die Bibliotheken haben ihre eigenen Systeme für Identifikatoren entwickelt.<sup>12</sup> Prominente Beispiele sind bibliotheksspezifische URN oder DOIs. Dies führte zeitweise zu einer Trennung von Communitys: auf der einen Seite Identifikatorensysteme mit dem Anspruch auf Nachhaltigkeit, auf der anderen Seite das Web mit einer Unsicherheit bezüglich der langfristigen Verfügbarkeit von Informationen.

Gerade das Semantic Web beruht auf dieser Verfügbarkeit. Und nicht nur dieser Teil der Web-Community legt inzwischen großen Wert auf nachhaltig

11 Vgl. [http://en.wikipedia.org/wiki/Knowledge\\_extraction](http://en.wikipedia.org/wiki/Knowledge_extraction) (Letzter Aufruf: 15.08.2013).

12 Vgl. Abschnitt 9.4 in Neuroth, Heike/Oßwald, Achim/Scheffel, Regine/Strathmann, Stefan/Huth, Karsten (Hrsg.): *nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*. Version 2.3. Göttingen 2010. <http://nestor.sub.uni-goettingen.de/handbuch/> (Letzter Aufruf: 15.08.2013).

nutzbare URIs. Dazu gehört auch das soziale Web: URIs auf Tweets werden inzwischen wie selbstverständlich in Nachrichtenmedien verwendet, sowohl im Digitalen als auch im Analogen, d. h. auf Papier.

Besonders für den Bereich wissenschaftlicher Publikationen hat es sich gezeigt, dass es kein Entweder-oder zwischen Identifikatorensystemen geben muss. So gibt die DOI Foundation auf ihrer Webseite folgende Empfehlung: „DOI names used through a HTTP proxy server (in the <http://dx.doi.org> representation as a URL) will continue to be persistent.“<sup>13</sup> Dies bedeutet, dass eine DOI wie 10.1000/182 auch als Webadresse verwendet werden kann. Der Domänenname <http://dx.doi.org> dient dabei als persistentes, also nachhaltig verfügbares, Präfix – so kann die URI <http://dx.doi.org/10.1000/182> gebildet werden. Dieses Verfahren erlaubt zugleich die Weiterverwendung von DOIs außerhalb des Webbereiches als auch die Nutzung des http-Präfixes im (Semantic) Web.

*Vokabulare – ist mein Autor Dein Autor – und wie soll man das aufschreiben?*

Mehrsprachigkeit, die Adäquatheit von Aussagen im Semantic Web und Identifikatorensysteme – das alles sind Herausforderungen, die in hohem Maße mit technischen Aspekten zu tun haben. Es gibt jedoch eine Herausforderung, bei der die Beziehung zwischen Communitys zentral ist.

Zugespitzt kann diese Herausforderung auf folgende Frage gebracht werden: Ist mein Autor Dein Autor – und wie soll man das aufschreiben? Hinter dieser Frage stehen zum einen die Vokabulare für maschinenlesbare Informationen selbst. Wie in diesen Vokabularen ein Autor oder Verfasser definiert wird – auf diese Frage gibt es schon innerhalb einer Community wie dem Bibliotheksbereich verschiedene Antworten. Die Formulierung einer Antwort in Form von generellen Vokabularen außerhalb der Community ist hingegen schwer zu finden.

Zum anderen zeichnet sich in den letzten Jahren eine Entwicklung bezüglich maschinenlesbarer Vokabulare ab, welche die Communitys polarisiert: Es wird zunehmend von semantischen Technologien gesprochen.<sup>14</sup> Diese sind vom

13 Vgl. [www.doi.org/doi\\_handbook/Glossary.html](http://www.doi.org/doi_handbook/Glossary.html) (Letzter Aufruf: 15.08.2013).

14 Eine Unterscheidung von Semantic Web vs. Semantic Technologies bietet [www.cambridgesemantics.com/semantic-university/semantic-web-vs-semantic-technologies](http://www.cambridgesemantics.com/semantic-university/semantic-web-vs-semantic-technologies) (Letzter Aufruf: 15.08.2013).

Semantic Web im engeren Sinne, d. h. der hier vorgestellten Technologiebasis RDF, abgegrenzt. Semantische Technologien zeichnen sich vor allem dadurch aus, dass sie eine Vielzahl von Möglichkeiten zur Repräsentation und Verarbeitung von semantischen Informationen im Web bieten. Der Google Knowledge Graph oder der Facebook Open Graph sind prominente Beispiele hierzu. Sie werden von einzelnen Firmen vorangetrieben. Schema.org wiederum ist eine Initiative mehrerer großer Suchmaschinenhersteller.<sup>15</sup> Alle diese Anstrengungen haben das gleiche Ziel wie das Semantic Web. Sie nutzen jedoch unterschiedliche Schreibweisen für semantische Informationen. Teilweise sind diese 1:1 kompatibel zu RDF. Schema.org etwa lässt sich ohne Verlust der semantischen Modellierung in RDF umwandeln.

Man könnte diese Entwicklung als Siegeszug der Kernidee des Semantic Web auffassen. Kritiker monieren, dass es sich um proprietäre Anstrengungen handelt und deshalb eine Monopolisierung der Semantik im Web droht. Schema.org ist dieser Kritik besonders ausgesetzt. Das hat mit dem Erfolg von schema.org zu tun, aber auch damit, dass schema.org daran arbeitet, existierende semantische Modellierungen in das schema.org-Vokabular zu integrieren. Ein prominentes, ehemals externes Vokabular, bei dem dies gelungen ist, heißt GoodRelations. Es dient der Auszeichnung von Produkten im Web und ist ein wichtiger Beitrag zur Entwicklung des E-Business. GoodRelations wurde als RDF-Vokabular entwickelt und ist nun in schema.org aufgegangen.

Diese Rolle von schema.org als Integrationspunkt für existierende Vokabulare spaltet: Sie wird als Chance oder als Gefahr angesehen. Erreicht hier Monopolisierung eine neue Qualität, welche communityspezifische Vokabulare erdrückt?

*Schlussbemerkung: Nachhaltigkeit maschinenlesbarer Informationen  
braucht Verständigung zwischen Menschen*

Die Diskussion um Vokabulare im Web hat gezeigt, dass die eigentliche Herausforderung für Nachhaltigkeit im Semantic Web sozialer Natur ist. Um beispielsweise die Frage „Ist mein Autor dein Autor?“ aus zwei Perspektiven (Bibliothek vs. schema.org) zu beantworten, werden zwei Plattformen benötigt.

Das W3C (World Wide Web Consortium) hat bei der Diskussion um schema.org die Rolle einer solchen Plattform eingenommen. Eine Mailingliste dient

15 Vgl. <http://schema.org/> (Letzter Aufruf: 15.08.2013).

der Diskussion von Vokabularaspekten im Allgemeinen; komplementiert wird sie von speziellen Listen, z. B. für den Bibliotheksbereich.<sup>16</sup>

Ein offenes Forum allein kann natürlich nicht sicherstellen, dass die Interessen kleinerer Communitys genauso berücksichtigt werden wie globale Unternehmen. Aber es kann helfen, Zusammenarbeit zwischen Communitys auszubauen und aufrechtzuerhalten. Die Nachhaltigkeit von Ergebnissen kann so auch auf einer nicht-technischen Ebene befördert werden.

<sup>16</sup> Vgl. die Mailinglisten <http://lists.w3.org/Archives/Public/public-vocabs/> sowie <http://lists.w3.org/Archives/Public/public-schemabibex/> (Letzter Aufruf: 15.08.2013).





*„Die Ellipse ist das elementare Bild der Neuzeit: Sie macht uns klar, wie entscheidend bei der Wahrnehmung der Welt unsere eigene Perspektive ist. Auf der documenta 13 machte die Uhr in der kurfürstlichen Aue von Kassel am Ende des Küchengrabens das schlagartig deutlich.*

*Mit seiner Theorie der elliptischen Funktionen hat im frühen 19. Jahrhundert der Potsdamer Mathematiker Carl Gustav Jacobi der deutschen Mathematik Weltgeltung verschafft: ein höchst aktuelles Wissen. Jacobis elliptische Funktionen wurden bei einem UNESCO Future Forum zitiert: Bei gesellschaftlichen Umbrüchen müssten wir zurückschwingen und uns zunächst auf jene Prinzipien und Messlatten verständigen, die im Übergang von der alten in die neue Welt als maßgebende Funktionen zu gelten haben.*

*Was aber ist maßgebend? Wovon leiten wir heute bindende Anforderungen ab? Die Austauschprozesse, seien sie öffentlich, privat oder marktmäßig, zeitigen verschiedene, mitunter einander ausschließende Anforderungen an Rechtssicherheit und Freiheit. Und die kulturellen Traditionen zeitigen eine wachsende kulturelle Vielfalt:*

*Kulturen haben verschiedene Wertesysteme mit je ganz eigenen ethischen und rechtlichen Anforderungen an ein Netz der Netze, dessen „Code“ (Lawrence Lessig) – das kommt erschwerend hinzu – keineswegs in allen Fällen offengelegt und transparent ist. Gemeinsam geteilte Prinzipien können wir nur auf der Basis der Zusammenarbeit und des gemeinsamen Lernens erreichen. Danke, CoLab!“*

– VERENA METZE-MANGOLD



# Semantische Nachhaltigkeit und Kontrolle: Gedanken zu schema.org und Linked Open Data

STEFAN GRADMANN

## *Linked Open Data: ein fröhliches Chaos*

„It’s massively successful. It is trivially simple. Massively successful like karaoke – anybody can do it.“<sup>1</sup>

„Anyone can say anything about anything.“ (Tim Berners-Lee)<sup>2</sup>

Die Entwicklung des World Wide Web von einem ersten Entwurf als Web der Dokumente über die zu Anfang jedenfalls von den Allmachtsphantasien der künstlichen Intelligenz geprägte Phase des Semantic Web bis zu seiner heutigen Ausprägung als Web der Dinge, als Linked Data Web, ist auf den ersten Blick eine unglaubliche Erfolgsgeschichte: Hier ist in wenig mehr als 20 Jahren eine gigantische, weltumspannende und komplex vernetzte Maschine für die Wissensgenerierung entstanden. Vielleicht am besten charakterisiert ist das entstandene Resultat durch den Titel von Heath und Bizer:<sup>3</sup> Das World Wide Web ist inzwischen in der Tat zu einem Global Data Space geworden.

- 1 Nelson, Ted: Visionary lays into the web. 2001. <http://news.bbc.co.uk/2/hi/science/nature/1581891.stm> (Letzter Aufruf: 15.08.2013).
- 2 Klyne, Graham/Caroll, Jeremy: Resource Description Framework (RDF). Concepts and Abstract Data Model. W3C. 2002. [www.w3.org/TR/2002/WD-rdf-concepts-20020829/](http://www.w3.org/TR/2002/WD-rdf-concepts-20020829/) (Letzter Aufruf: 15.08.2013).
- 3 Heath, Tom/Bizer, Christian: Linked Data: Evolving the Web into a Global Data Space. 2011. <http://linkeddatatoolkit.com/editions/1.0/> (Letzter Aufruf: 15.08.2013).

Dabei waren in der Entwicklung vom Document Web der ersten Generation zum heutigen World Wide Web zwei elementare Erweiterungen maßgeblich: Die syntaktische Erweiterung der WWW-Protokolle um das Resource Description Framework (RDF) erlaubte die semantische Typisierung von Verbindungen zwischen WWW-Entitäten und damit – zusammen mit der Schemasprache RDFS – eine maschinelle Verarbeitung von WWW-Inhalten im Sinne einfacher deterministischer Schlussoperationen, und eine Erweiterung im Repräsentationsraum des World Wide Web erlaubt nunmehr die Repräsentation im Prinzip beliebiger Entitäten der „wirklichen“ und der gedachten Welt im World Wide Web.

Dabei sind drei Charakteristika bemerkenswert, die im Folgenden polemisch betrachtet werden sollen:

- Es gibt keine Regierung des World Wide Web, welche die Macht hätte, über opportune bzw. nicht genehme Äußerungen zu bestimmen. Das als Motto vorangestellte Zitat von Berners-Lee sagt es sehr treffend: Jede Äußerung zu jedem Thema ist im World Wide Web für jedermann möglich, eine Zensur findet nicht statt und ist auch so gut wie nicht implementierbar. Staaten wie China, denen dies ein Dorn im Auge ist, haben keine andere Wahl, als ihre gesamte Population hinter einer riesigen Proxyarchitektur zu verstecken und sie damit aus dem World Wide Web zu verbannen: Im World Wide Web wäre sie nicht effektiv kontrollierbar.
- Das massive Wachstum des Linked Data Web seit 2007 wurde technisch ermöglicht vor allem durch eine massive Vereinfachung des komplexen Schichtenmodells des Semantic Web und eine Rückbesinnung auf RDF als zentralen technischen Ankerpunkt sowie darauf basierende robuste technische Standards des W3C wie etwa SPARQL.
- Technische Vereinfachung und die urdemokratische Entwicklung des Web als Aussagensystem mitsamt der damit verbundenen Freiheit, Ontologieressourcen frei und unkontrolliert zu kreieren und zu verwenden, haben ihren Preis: Der Blick auf eine URL wie <http://ws.nju.edu.cn/falcons/ontologysearch/result.jsp?query=person> offenbart ein massives Qualitätsproblem des Linked Data Web, denn die Ontologiesuchmaschine Falcons macht dort sichtbar, dass es an Ontologieressourcen für die Modellierung etwa von Personen ganz sicher nicht mangelt: Es sind deren im Gegenteil deutlich zu viele.

Das W3C ist sich gerade des letzten Problems durchaus bewusst und hat es lange Zeit auch bewusst toleriert, wie das folgende Zitat von Klyne und Caroll belegt:

„To facilitate operation at Internet scale, RDF is an open-world framework that allows anyone to say anything about anything. In general, it is not assumed that all information about any topic is available. A consequence of this is that RDF cannot prevent anyone from making nonsensical or inconsistent assertions, and applications that build upon RDF must find ways to deal with conflicting sources of information. (This is where RDF departs from the XML approach to data representation, which is generally quite prescriptive and aims to present an application with information that is well-formed and complete for the application’s needs.)“<sup>4</sup>

Neben den anderen Großbaustellen des World Wide Web (fehlende Versionierungsmethoden, mangelnde Expressivität in den Bereichen Provenienz und Autorisierungsmerkmale) hat das Linked Data Web also ein massives Qualitätsproblem vor allem in Gestalt der vielen koexistierenden und unkontrollierten Ontologien, die einerseits zum Teil enorme Überschneidungsbereiche haben, andererseits aber doch immer nur partiell semantisch redundant sind, sodass ein Konstrukt wie owl:sameAs streng genommen so gut wie nie korrekt anwendbar ist.

Damit ähnelt das fröhliche Chaos des Linked Data Web – wollte man eine architektonische Metapher gebrauchen – am ehesten einer dynamisch wuchernen brasilianischen Favela, und es beginnt in der Community das Nachdenken darüber, wie dieses Qualitätsproblem anzugehen wäre: Ontology-Mapping sowie die Bewertung und Modellierung semantischer Nähe und Überschneidung sind zwei der momentan aktivsten Forschungsfelder in diesem Umfeld.

### *Schema.org: Ordnung, Disziplin und Kontrolle*

Ganz anders als die favelaartige Herangehensweise mutet die Strategie der Schema.org-Initiative an, die – wieder in einer architektonischen Metapher – weit eher dem wohlorganisierten, geordneten Bau einer Kathedrale ähnelt.

4 Klyne, Graham/Caroll, Jeremy: Resource Description Framework (RDF). Concepts and Abstract Data Model. World Wide Web Consortium (W3C). 2002. [www.w3.org/TR/2002/WD-rdf-concepts-20020829/](http://www.w3.org/TR/2002/WD-rdf-concepts-20020829/) (Letzter Aufruf: 15.08.2013).

Lanciert wurde diese Initiative am 02. Juni 2011 von den Suchmaschinen Google, Bing (Microsoft) und Yahoo, einige Monate später kam die führende russische Suchmaschine Yandex dazu. Das Ziel von schema.org wird im Blog der Initiative wie folgt benannt: „create and support a standard set of schemas for structured data markup on web pages.“<sup>5</sup>

Es geht also um den Aufbau eines kontrollierten Kernvokabulars, der zwar Erweiterungen durchaus zulässt, jedoch immerhin eine zentrale Redaktionsinstanz kennt. Neben diesem Charakteristikum unterschied schema.org sich zumindest im ersten Angang noch in einem zweiten Punkt entscheidend vom Linked Data Web: Die Aussagensyntax von schema.org war ursprünglich auf microdata beschränkt, RDFa kam im September 2011 dazu; im ersten Ansatz war schema.org also bewusst nicht als Teil des offenen World Wide Web entworfen.

Dabei unterliegt der Weltsicht von schema.org, wie sie unter <http://schema.org/docs/full.html> dokumentiert ist, einer eigenartigen Verzerrung, vergleicht man sie mit derjenigen von Top-Level-Ontologien wie etwa SUMO<sup>6</sup> und der darunterliegenden Mid-Level Ontology MILO. Deutlich überrepräsentiert sind hier – wenig verwunderlich! – die Personengruppen und Institutionen / Firmen, die das klassische, auf Werbeeinnahmen basierende Geschäftsmodell der großen Suchmaschinen befeuern. Und kaum prominent sind Entitäten, mit denen sich im World Wide Web kein Geld verdienen lässt. So nehmen etwa Diätpläne (diet) in der Klasse Creative Work als direkte Subklasse eine deutlich prominentere Rolle ein als die zwei Hierarchiestufen tiefer angesiedelten wissenschaftlichen Aufsätze (scholarly article), die dann außerdem kurioserweise noch als eigene Unterklasse medical articles gesondert behandeln – auch hier ist wahrscheinlich wieder eine monetär geprägte Perspektive im Spiel. Oder – um ein anderes Beispiel zu nennen – es sind innerhalb von Creative Work gleich drei eigene Unterklassen für TVEpisode, TVSeason und TVSeries vorgesehen.

Dennoch muss zugunsten von schema.org festgehalten werden, dass die Ontologie das selbstgesetzte Ziel eines einheitlichen Standardvokabulars ohne die massiven Redundanzen des Linked Data Web zumindest ein Stück weit erreicht: Das Schema <http://schema.org/Person> jedenfalls wirkt deutlich aufgeräumter als das oben angeführte Beispiel der Personenontologien im Linked Data Web –

5 O'Connor, Michael: schema blog. 2011. <http://blog.schema.org/search?updated-max=2011-12-12T12:10:00-08:00> (Letzter Aufruf: 15.08.2013).

6 Vgl. [www.ontologyportal.org/](http://www.ontologyportal.org/) (Letzter Aufruf: 15.08.2013).

wenngleich auch hier wieder die kommerzielle Verzerrung des Suchmaschinen-Geschäftsmodells durchschlägt, denn unter den ansonsten eher generischen Attributen von Person finden sich völlig isoliert und überraschen die zwei sehr spezifischen Attribute `hasPOS` und `brand`.

Ähnlich steht es mit `schema.org`-basierten Anwendungen: Das hierfür generell als Quelle interessante Verzeichnis <http://linter.structured-data.org/examples/> führt ganz überwiegend Verzeichnis- und Vertriebsdienste aus dem Bereich der Kreativindustrie auf.

Vor allem aber fällt auf, wie wenig Verbindung die Vokabularschemata von `schema.org` mit den Ontologien des Linked Data Web haben, modellieren sie doch durchweg Entitäten, die selbst schon eine Repräsentation im Linked Data Web haben. `Schema.org` fügt also den Vokabularressourcen im World Wide Web eine weitere, nahezu 100 % semantisch redundante Vokabularschicht hinzu, was unmittelbar zur Frage nach der Motivation der Initiatoren von `schema.org` führt: Warum diese großangelegte, hochgradig redundante Initiative?

### *Eine versteckte Agenda?*

Von Beginn an war die Diskussion über `schema.org` im World Wide Web von heftigen Spekulationen und mitunter geradezu paranoiden Untertönen geprägt, genährt allerdings nicht zuletzt durch die notorische Intransparenz von Google: Die Firma ist bis heute Antworten auf eine ganze Reihe sich unmittelbar aufdrängender Fragen schuldig geblieben.

So ist etwa bis heute unklar, was die harten Konkurrenten Google, Microsoft und Yahoo dazu bewogen hat, ausgerechnet im Bereich ihres Suchmaschinen-Kerngeschäfts zu kooperieren. Handelt es sich womöglich um eine Reaktion auf das als bedrohlich wahrgenommene Funktionsmodell des Linked Data Web, in der Absicht, Kunden möglichst nachhaltig an das Suchmaschinen-Geschäftsmodell zu binden?

Und auch die Motivation der ursprünglichen Einengung auf die Microdata-Syntax und der damit verbundenen Ablehnung von RDF ist niemals wirklich deutlich geworden, zumal diese Linie dann schon nach wenigen Monaten verlassen wurde: Warum hat man sich nicht von vornherein auf das RDF-Paradigma eingelassen?

Und schließlich: Gibt es einen Zusammenhang zwischen `schema.org` und dem wenig später lancierten Knowledge Graph, der wie eine googlespezifische



Antwort auf das Linked Data Web daherkommt, und wenn ja, welcher Art ist dieser Zusammenhang? Manu Sporny hat zu diesem Thema interessante Details zusammengetragen.<sup>7</sup>

Genereller gesprochen: Ist das Verhältnis der Paradigmen schema.org und Linked Open Data in Analogie zu begreifen zu der metaphorischen Opposition The Cathedral vs. The Bazaar, wie sie in einem Essay von Raymond für die Opposition Open Source vs. Closed Source eingeführt worden war?<sup>8</sup> Ist schema.org also nützlich oder zumindest unschuldig oder womöglich gar böse?

Letzterer Position scheint Stewart zuzuneigen, wenn er in seinem Blog schreibt:

„Schema.org appears to be Linked Data Lite with extremely limited support for vocabularies outside of the service. [...] There is a subtle air of intimidation throughout the schema.org announcements and documentation. [...] Again, I could just be paranoid, but this is Microsoft and Google we're talking about. Whatever happened to ‚do no evil?‘“<sup>9</sup>

Und selbst wenn man so weit nicht gehen möchte und jeden paranoiden Denkreiflex vermeidet, bleibt die Eingangsfrage unbeantwortet, was Google, Microsoft und Yahoo dazu bewogen hat, eine eigene, semantisch redundante Initiative zu lancieren, die das Linked Data Web im Grunde noch einmal zu erfinden versucht.

### *Semantischer Darwinismus - oder doch Zensur?*

Die Antwort könnte – vordergründig betrachtet – in zwei alternativen Annahmen bestehen. Entweder handelt es sich bei schema.org um den zutiefst unschuldigen, raubtierkapitalistischen Versuch, eine große, aber diffuse semantische

7 Vgl. Sporny, Manu: Google Indexing RDFa 1.0 + schema.org Markup. In: The Beautiful, Tormented Machine. 2012. <http://manu.sporny.org/2012/google-indexing-schema-rdfa/> (Letzter Aufruf: 15.08.2013).

8 Raymond, Eric S.: The Cathedral & the Bazaar. Beijing u. a. 2001.

9 Stewart, Darin: Schema.org: Webmaster One-Stop or Linked Data Land Grab? Gartner Blog. 2011. <http://blogs.gartner.com/darin-stewart/2011/06/04/schema-org-webmaster-one-stop-or-linked-data-land-grab/> (Letzter Aufruf: 15.08.2013).

Ressource (das Linked Data Web) durch eine andere, bessere Lösung zu ersetzen, und damit nur um eine radikale Variante dessen, was wir möglicherweise bald ontologischen Darwinismus nennen könnten. Oder aber es handelt sich um eine Art von Outsourcing-Modell für Zensur, das letztendlich darauf abzielt, nur noch diejenigen Inhalte im World Wide Web effektiv präsent werden zu lassen, die das „richtige“ Vokabular verwenden – nur dass dann eben die Zensurhoheit nicht mehr bei politisch-administrativen Instanzen liegt, wie in der Vergangenheit, sondern bei Google & Co.

Vielleicht jedoch besteht die richtige Antwort in einer sehr viel weiter gehenden Annahme, die in gewisser Hinsicht substantiell über den traditionellen Zensurbegriff hinausgeht. Denn die traditionelle Zensur war eine Methode, Menschen an der Veröffentlichung ihres Schaffens zu hindern, nicht jedoch an der geistigen Produktion selbst. Insbesondere waren die Sprache und die sonstigen künstlerischen Hilfsmittel dabei niemandes partikulares Eigentum – dies könnte sich, zumindest bezogen auf die Wahrnehmbarkeit von Aussagen im World Wide Web, mit schema.org in dem Sinne ändern, dass damit die Ausdrucksmittel selbst oligopolistisches Eigentum des Suchmaschinenkonsortiums werden. Wie weit die Zusammenarbeit des Schema.org-Konsortiums mit dem W3C diese Annahme relativieren kann, ist aus meiner Sicht momentan nicht einschätzbar.

In einer sehr pessimistischen Lesart ginge es dann also bei schema.org nicht nur darum, was gesagt werden kann (bzw. nicht gesagt werden kann) und wer das zugrundeliegende Vokabular kontrolliert, sondern vor allem darum, welche Aussagen überhaupt effektiv wahrgenommen werden (und welche einfach nicht registriert werden). Und damit ginge es dann bei schema.org gar nicht primär um Geld, sondern um die Leitwährung des World Wide Web: um Aufmerksamkeit!

Anmerkung des Autors: Ich habe an anderer Stelle (so zuletzt in Gradmann, Stefan: *Semantic Web und Linked Open Data*. In: *Grundlagen der praktischen Information und Dokumentation*. Berlin 2013, S. 219-228) die Entwicklung des Semantic Web und des Linked-Data-Paradigmas überblicksartig nachgezeichnet, sodass an dieser Stelle ein erneutes Referieren dieser Entwicklung verzichtbar ist: Die nachstehenden Ausführungen mögen vielmehr als Vertiefung von Aspekten gelesen werden, die dort ebenfalls schon angerissen sind. Weiter verweise ich in diesem Zusammenhang auf die drei grundlegenden, lehrbuchartigen Einführungen: Antoniou, Grigoris/Van Harmelen, Frank: *A Semantic*

Web Primer, 2nd Edition. Cambridge 2008. <http://dl.acm.org/citation.cfm?id=1373341> (Letzter Aufruf: 15.08.2013); Hitzler, Pascal/Krötzsch, Markus/Rudolph, Sebastian/Sure, York: Semantic-Web-Grundlagen. Heidelberg 2008. [www.semantic-web-grundlagen.de/](http://www.semantic-web-grundlagen.de/) (Letzter Aufruf: 15.08.2013); Heath, Tom/Bizer, Christian: Linked Data: Evolving the Web into a Global Data Space. 2011. <http://linkeddatabook.com/editions/1.0/> (Letzter Aufruf: 15.08.2013). Alle drei sind auch für eine vertiefende Einführung empfehlenswert.

# Wikipedia und Nachhaltigkeit

INTERVIEW MIT JENS BEST, PRÄSIDIUMSMITGLIED VON  
WIKIMEDIA DEUTSCHLAND

Enzyklopädien haben in der Vergangenheit den jeweiligen Wissensstand ihrer Zeit dokumentiert. Heute hat die Wikipedia als kollaborativ erstelltes Online-Angebot diese Funktion. Während bei gedruckten Enzyklopädiendruckwerken der physische Erhalt der Bücher im Vordergrund stand, ergeben sich bei der Wikipedia in Hinblick auf die Nachhaltigkeit ganz neue Fragen.

*Was verstehen Sie unter Nachhaltigkeit in Hinblick auf die Wikimedia-Projekte?*

*Jens Best:* Freiwillige Kollaboration in einem offenen Netzwerk ist nur dann nachhaltig, wenn sie lebendig gehalten werden kann. Man kann sich gerne an den beeindruckenden Zahlen der Wikimedia-Projekte berauschen. Fünftgrößte Website der Welt, 1,5 Millionen Artikel in der deutschsprachigen Wikipedia, 17 Millionen Dateien in der Wikimedia Commons und 500 Millionen Nutzer jeden Monat sind Zahlen, die beeindruckend sind. Sie zeigen, dass die Produktion und Rezeption von freiem Wissen erfolgreich ihren Weg in die globale Gesellschaft finden.

Dennoch sind die Herausforderungen enorm. Die wichtigste Frage bezüglich Nachhaltigkeit ist nicht, ob wir die nächsten 100 Jahre unsere Server und unseren Traffic zahlen können, sondern, ob es uns gelingt die aktiven Gemeinschaften, die freies Wissen und freie Bildung konkret erstellen, zu fördern und zu vergrößern. Dies setzt an einer der wichtigsten Wurzeln einer modernen Gesellschaft an. Es geht um die Frage, ob eine Gesellschaft erkennt, dass Wissensgewinnung und Wissensvermittlung eine ihrer zentralen Aufgaben ist, und dies gemäß den technischen Möglichkeiten gerecht und frei für ihre Bürger umsetzt.

Die Wikipedia ist für viele Menschen in vielen Bereichen ihres alltäglichen und professionellen Lebens eine der ersten Quellen, um ihr Wissen zu erweitern. Sie hilft, besser zu verstehen, wie Dinge funktionieren. Nachhaltigkeit bedeutet nicht nur die ständige Verbesserung der Artikel, sondern auch die Arbeit an den Prozessen der individuellen und gemeinschaftlichen Erkenntnisvermittlung. Wikipedia hat gezeigt, dass Wissen in seiner enzyklopädischen Erfassung kein Projekt von Wenigen sein muss, sondern qualitativ auf breiter und vollkommen offener Basis möglich ist. Dass dies funktioniert, haben viele immer noch nicht wirklich verstanden. Es ist auch keine Selbstverständlichkeit. Nur weil das Internet die Möglichkeit geschaffen hat, in offen kollaborativer Arbeit unglaubliche Ergebnisse zu produzieren, ist damit kein neues Paradigma des Wissens vom Himmel gefallen. Wir stehen im Gegenteil meiner Meinung nach immer noch sehr am Anfang dieses kulturellen Prozesses, bei dem die Wikipedia eines der ersten beeindruckenden Zwischenergebnisse ist.

Insofern verstehe ich unter Nachhaltigkeit nicht unbedingt, dass es auch in 100 Jahren noch eine Wikipedia geben muss (wobei mich das als Präsidiumsmitglied natürlich freuen würde). Aber die digitalen Netzwerke, die neue Formen der Offenheit und Kollaboration fördern, und die Prinzipien, die eine demokratischere Form der Erstellung, Vermittlung und Bewahrung von Wissen ermöglichen, gilt es, kulturell verantwortlich weiterzuentwickeln. Eine Gesellschaft, deren primärer Antrieb nicht die Kommerzialisierung, sondern die Demokratisierung von Information ist, muss diese nutzbar machen.

*Welche Strategie verfolgt die Wikimedia, um die Verfügbarkeit ihrer Inhalte langfristig sicherzustellen?*

Die Wikimedia-Projekte, allen voran die Wikipedia, sind eher ein Prozess als ein fertiges Produkt, das es zu konservieren gilt. Insofern geht es in erster Linie um die Sicherstellung und Verbesserung des Betriebes dieses globalen Masterprojektes kollaborativer Zusammenarbeit.

Um den Betrieb der Website sicherzustellen, hat die Wikimedia Foundation Rücklagen für mehrere Jahre. Es wird jedes Jahr neu über den Spendenanteil entschieden, der in diese Rücklagen geht. Darüber hinaus gibt es eine aktive Diskussion über die längerfristige Sicherstellung unter dem englischen Stichwort *endowment*. Die Wikimedia Foundation ist im Kern dafür verantwortlich, den technischen Betrieb sicherzustellen. Sie wird dabei, insbesondere finanziell, von

den einzelnen nationalen Chapters unterstützt. Sprich, es wird ein wesentlicher Teil der in Deutschland eingesammelten Spenden an die US-Stiftung weitergeleitet, damit diese Serverfarmen und andere technische Einrichtungen, Dienste und Entwicklungen bezahlen kann.

Schaut man sich die juristischen und gesellschaftlichen Rahmenbedingungen an, so ist der Betrieb der Wikimedia-Projekte einer ständigen Entwicklung unterworfen, nicht immer zum Vorteil für freies Wissen. Wenn sich in den USA die Laufzeiten von urheberrechtlichen Werken verlängern, kann das dann auch Fotos und Digitalisate betreffen, die auf Wikimedia Commons gespeichert wurden. In solchen Momenten wird sowohl in der Wikimedia Foundation, den nationalen Chapters, aber auch insbesondere in der Community der Freiwilligen viel über alternative Wege diskutiert. Die Abwägungen, die hier getroffen werden müssen, sind nicht immer einfach. Andererseits aber kann nicht bei jeder Verschlechterung nationaler rechtlicher Rahmenbedingungen die globale Struktur der Wikimedia angepasst werden. Wir arbeiten vielmehr auf politischer und gesellschaftlicher Ebene dafür, dass es zu besseren Bedingungen für die Grundlagen von freiem Wissen und freier Bildung kommt – in allen Ländern, in denen die Wikipedia und ihre Schwester-Projekte eine Basis haben.

Eine der wichtigsten Aufgaben der Wikimedia Deutschland sowie ihrer Schwester-Organisationen in anderen Ländern ist es, die Teilnahme von Freiwilligen zu fördern und auf eine gesunde Basis zu stellen. Wissen ist ein ständiger Erkenntnisprozess, der nur dann zu informierten Menschen führt, wenn er lebendig gehalten wird. Und genau darin liegt die wichtigste langfristige Aufgabe der Wikimedia-Bewegung.

Um Wissen mit den Werkzeugen des freien Web aufzuarbeiten, muss die Wikimedia-Bewegung auch vermehrt mit anderen Wissensinstitutionen zusammenarbeiten. Diesen Institutionen beim digitalen Wandel zu helfen, ihnen in Kooperationen zu zeigen, wie der Übergang zu mehr freiem Wissen und seiner offenen Vermittlung möglich ist, wird mehr und mehr zu einer der zentralen Aufgaben der Wikimedia Deutschland und ihren Schwester-Organisationen in anderen Ländern. Dies geschieht durch gemeinsames Lernen und das Ausprobieren der neuen Möglichkeiten, die sich für Wissensarbeiter jeglicher Art durch diese webbasierten offenen Strukturen eröffnen. Wir stehen da, trotz vieler beeindruckender Zahlen, immer noch am Anfang eines kulturellen Wandels von historischer Dimension. Und dieser will Schritt für Schritt, von einem Artefakt zum nächsten Artikel, ergründet werden.

Wikimedia stellt die Inhalte ihrer Projekte, die ja ständig erweitert werden, zur Verfügung, indem wir sie mit anderen teilen und mit anderen Institutionen und Organisationen in Austausch treten. Ähnlich der Grundidee des Internets liegt das Wesen der Wikimedia-Projekte nicht darin, Inhalte an einer zentralen Stelle zu sammeln, sondern die Teilnehmer umfassend zu vernetzen.

*Im Internet werden dauerhafte Strukturen immer wichtiger. Ist die Wikimedia darauf vorbereitet?*

Die Bausteine für Wissen, also Daten, Informationen und Zusammenhänge, werden nicht nur im enzyklopädischen Projekt Wikipedia zusammengetragen. Es gibt zum Beispiel für Fotografien, Grafiken und Digitalisate das Projekt Wikimedia Commons. Sie können dort eine weitere Heimat unter freien Lizenzen finden. Bei Wikimedia Commons können einzelne Informationsobjekte mit ihren Metadaten hochgeladen werden. Die Objekte können dann in die Wikipedia eingebunden werden. Es sind aber auch viele andere Einbindungen und Nutzungen innerhalb und außerhalb des Wikimedia-Universums möglich. Dafür braucht es Standards, damit dies einfach, rechtssicher und pluralistisch geschehen kann.

Dennoch ist keines der Wikimedia-Projekte die zukünftige Bibliothek von Alexandria. Wer im Web dauerhaft und nachhaltig handeln will, muss offen denken. Dafür sind die freien Lizenzen, unter denen alle Inhalte der Wikimedia-Projekte stehen, eine wichtige Grundlage. Ich denke, man kann sich – als nicht-staatliche Organisation sowieso – auf nachhaltige Strukturen nur vorbereiten, indem man die offene Kollaboration mit allen Beteiligten in den Vordergrund stellt. Wikimedia hat die Aufgabe, diese offene Zusammenarbeit voranzutreiben.

In den letzten Jahren ist in einigen wichtigen Bereichen bereits einiges besser geworden. Wo man früher gegen Wände gelaufen ist, entstehen heute Fenster und Türen, durch die ein neuer Wind weht, der die gesamtgesellschaftliche Aufgabenteilung der Schaffung, der Vermittlung und der Bewahrung von Wissen voranbringt. Wikimedia ist bereit, ihren Platz als zivilgesellschaftliche Teilnehmerin in diesem faszinierenden Aufbruch einzunehmen. Wikimedia agiert hier allerdings in einem Spannungsfeld. Einerseits ist die Pflege und Förderung der Freiwilligenprojekte, allen voran die Wikipedia, unsere Kernaufgabe, andererseits stellt uns gerade diese Verpflichtung mehr und mehr vor Herausforderungen, die weit über die Sicherstellung des Wachstums der eigenen Projekte hinausgehen. Es sind spannende Zeiten.

Um also in aller Kürze auf die Frage zu antworten: Die Wikimedia ist darauf vorbereitet, im Wandel zu einem offenen, dauerhaften Wissens-Web eine aktive Rolle zu spielen. Eine Nichtregierungsorganisation allein kann diese historische Aufgabe am Beginn des digitalen Informationszeitalters allerdings nicht bewältigen. Wikimedia ist bereit, ihrer Pionierrolle gerecht zu werden und an vorderster Stelle die mannigfaltigen Details dieses Wandel gemäß unseren Vorstellungen von freiem Wissen und freier Bildung zu erforschen, mitzugestalten und weiterzuentwickeln – aber die wichtigste Aufgabe wird sein, an einer Gesellschaft zu arbeiten, die dieses als eine ihrer dringlichsten nachhaltigen Aufgaben erkennt.

*Wie sehen Sie das Verhältnis von Objekten und Kontextualisierung?*

Wenn man den Blick vom einzelnen Objekt auf seine Zusammenhänge und Netzungen lenkt, scheint die Welt im ersten Moment komplizierter zu werden. Die Muster der kulturellen, sozialen, wissenschaftlichen und zwischenmenschlichen Verbindungen können verwirrend sein, wenn man nicht in der Lage ist, die Werkzeuge zu benutzen, die der Mensch für das Leben auf dieser neuen Komplexitätsstufe gerade entwickelt. Dies bewirkt Phasen der Irritation, die aber unverzichtbar sind, damit eine Gesellschaft sich entwickelt und sich schrittweise an die neuen Bedingungen anpassen kann. Es ist eine Aufgabe der Wikimedia, diese Entwicklung so mitzuformen, dass möglichst viele Menschen die Gelegenheit haben, mit den neuen Werkzeugen Wissen zu nutzen und zu gestalten.

Ein wichtiger Schritt ist ein junges Projekt im Wikimedia-Universum: Wikidata. Wikidata ist im Kern eine Datenbank, die wie die anderen Wikimedia-Projekte frei bearbeitet werden kann. Wikidata ermöglicht den Einsatz von strukturierten Datensätzen unter freier Lizenz. Dies stellt einen wichtigen Schritt sowohl für den Betrieb der Wikipedia als auch mittelfristig für die breite offene Verwendung von Big Data dar. Für den Betrieb der Wikipedia wird der globale Abgleich von Datensätzen quer durch alle Sprachversionen erleichtert. Aber mittelfristig können auf Wikidata Werkzeuge und Anwendungen aufsetzen, die die Vermittlung von komplexen Datensätzen in vielen Bereichen vereinfachen. Wenn zum Beispiel ein Lehrer in hoffentlich nicht zu vielen Jahren das Auf und Ab der Bevölkerungsstrukturen verschiedener Städte verdeutlichen will, kann er auf die unter freier Lizenz verfügbaren Datensätze von Wikidata zurückgreifen. Mittelfristig wird es also möglich sein, dass ein Individuum jenseits bestehender Zugehörigkeit zu Institutionen auf geprüfte Objekte und Objektstrukturen



zugreifen kann, um diese in neue Zusammenhänge zu stellen. Wikidata kann hier ein wichtiger Knotenpunkt in einem Netzwerk freier Datenbanken werden, wie sie heute schon in vielen wissenschaftlichen Bereichen existieren. Die intelligente Nutzbarmachung dieser Ressourcen stellt gleichzeitig einen wichtigen Wendepunkt in der gesellschaftlichen Wissenvermittlung und eine große Herausforderung für unsere bestehenden Lernsysteme dar.

*Welche Bedeutung haben Standards für die Wikimedia?*

Die Findungsprozesse von Standards in Bezug auf Wissensbewahrung und Wissensvermittlung sind so mannigfaltig, dass eine eindeutige Antwort nicht möglich ist. Sicher ist ein wichtiger Aspekt der politischen und gesellschaftlichen Arbeit der Wikimedia Deutschland die Förderung von rechtlichen Standards, die freies Wissen und freie Bildung ermöglichen. Creative Commons ist für die Wikimedia ein wichtiger globaler Standard, um die Nutzung freier Inhalte rechtssicher zu ermöglichen.

*Wie sehen Sie die Rolle der Wikimedia im Verhältnis zu öffentlichen Gedächtnisinstitutionen einerseits und Firmen andererseits?*

Wikimedia will nicht das Apple des freien Wissens werden – dies würde die gesamte Grundüberzeugung auf den Kopf stellen. Wikimedias wichtigste Aufgabe neben der Förderung der Freiwilligen-Projekte ist die wachsende Zusammenarbeit mit anderen Trägern, Vermittlern und Bewahrern von Wissen in der Gesellschaft. Dies umfasst den Bereich Kultur, gerne kurz umschrieben mit dem Stichwort GLAM (Galleries, Libraries, Archives, Museums), sowie die Bereiche Bildung und Politik.

Wikimedia hat global in all diesen Feldern in den letzten Jahren seine Aktivitäten verstärkt. Viele Leuchtturmprojekte, mit denen wir in unterschiedlichster Intensität zusammenarbeiten, zeigen, welche Chancen und Herausforderungen auf uns zukommen, wenn Wissen breiter und demokratischer verfügbar wird. Wenn Wissen in einen gelebten sozialen Kontext gesetzt wird, geht es zwar auch immer um Macht, aber eben auch um Verstehen. Das Internet stellt insgesamt eine bedeutende soziokulturelle Herausforderung dar und das auf globaler und lokaler Ebene. Viele Kooperationen der letzten Jahre – es sei nur das Google-Books-Projekt genannt – haben gezeigt, wie sensibel dieser Prozess anzugehen

ist, obwohl es technisch möglich wäre, viel schneller vorzugehen. Wikimedia spielt hier mehr und mehr die Rolle einer Nichtregierungsorganisation, die im Zentrum des Spannungsfeldes von techno-sozialer Innovation, neuer demokratischer Wissensvermittlung und kultureller Bewahrung steht. Diese Rolle ist in dieser Kombination noch wenig erforscht, und somit ist jede Kooperation, jeder Diskurs darüber, wie rechtlich und kulturell am besten vorzugehen ist, ein neuer Balanceakt.



# Was wollen wir archivieren?

MICHAEL HOLLMANN

Archivare bewerten! Im Gegensatz zu der landläufig verbreiteten Meinung, Archivarinnen und Archivare würden in ihren Archiven beinahe alles horten, was in ihre Reichweite gelangt, gehört es zu ihren wichtigsten und komplexesten Aufgaben, aus der großen Masse der ihnen angebotenen Unterlagen diejenigen auszuwählen, die es „wert“ sind, dauerhaft, d. h. zeitlich unbegrenzt, überliefert zu werden. Wichtig ist in diesem Zusammenhang, dass der Begriff Unterlage im archivischen Sinne umfassend verstanden werden will und alle Formen der Informationsaufzeichnung von der traditionellen Akte bis hin zu audiovisuellen und genuin digitalen Aufzeichnungen einbegreift.<sup>1</sup> Soweit es sich um den staatlichen Bereich handelt, enthalten alle (deutschen) Archivgesetze des Bundes und der Länder spezielle Vorschriften, mit denen die Archive ausdrücklich zur Bewertung ermächtigt werden, in aller Regel sogar in Form eines nur durch vergleichsweise schwache Mitwirkungsrechte der abgebenden Stellen eingeschränkten Monopols.

Warum ist das so, und wie rechtfertigt sich eine – zumindest für nach 1945 entstandene Unterlagen geltende – durchschnittliche Kassationsquote von ca. 75 %? Für die vergleichsweise geringe und damit überschaubare Menge überlieferter Quellen für die Epochen des Mittelalters und der frühen Neuzeit und selbst für das 19. Jahrhundert verbietet sich eine Kassation grundsätzlich. Auch für die

1 Inhaltlich stützt sich die nachfolgende thesenartige Ausarbeitung zu meinem am 31. Mai 2013 gehaltenen Referat im Wesentlichen auf meinen Beitrag zur „Bestandspolitik“, der voraussichtlich noch 2013 im Metzler Handbuch Archiv veröffentlicht werden wird. Lepper, Marcel/Raulff, Ulrich (Hrsg.): Handbuch Archiv. Geschichte, Aufgaben, Perspektiven. Stuttgart 2013.

Unterlagen des frühen 20. Jahrhunderts gilt, dass angesichts der enormen Kriegsverluste die auf uns gekommenen archivischen Quellen beherrschbar sind und die wissenschaftliche Forschung für diese zentrale Epoche der deutschen und europäischen Geschichte auf jede noch verfügbare Information angewiesen ist.

Nach 1945 wuchs jedoch das Verwaltungsschriftgut in einem Maße an, dass es unerlässlich ist, die exponentiell steigende Masse von Unterlagen einer intensiven Sichtung und Bewertung zu unterziehen. Ausschlaggebend ist dabei, dass nach menschlichem Ermessen kein Archiv jemals über die personellen und finanziellen Ressourcen verfügen wird, um die Gesamtheit aller in seinem Zuständigkeitsbereich entstandenen Unterlagen übernehmen, konservieren, erschließen und auf unbegrenzte Dauer für die Benutzung bereitstellen zu können. Daher sind die Archive – auf eine kurze Formel gebracht – bemüht, mit einem Minimum an Unterlagen ein Maximum an Informationen dauerhaft zu sichern.

Eine so motivierte, erhebliche Reduzierung der Unterlagen liegt auch – so paradox das klingen mag – nicht zuletzt im Interesse der Benutzerinnen und Benutzer. Diese sind darauf angewiesen, dass die Archivarinnen und Archivare im Zuge der Bewertung die ursprüngliche, durch vielfache Redundanz gekennzeichnete Informationsmenge gleichzeitig reduzieren und verdichten. Nur dann besteht überhaupt die Möglichkeit, das für die jeweilige Fragestellung verfügbare Archivgut überschauen und umfassend auswerten zu können.

Damit die im Ergebnis der Bewertung zu Archivgut umgewidmeten Unterlagen ein immer noch umfassendes Bild ihrer Entstehungszeit vermitteln, bedarf es eines methodisch durchdachten und Willkür nach Möglichkeit ausschließenden Vorgehens. Die Regeln und Prinzipien der archivischen Bewertung müssen späteren Archivbenutzern nachvollziehbar sein und ihnen eine angemessene quellenkritische Beurteilung des Archivguts ermöglichen. Es kann daher nicht verwundern, dass die Bewertung einen zentralen Platz in der archivischen Fachdiskussion einnimmt, wobei die Frage nach dem Vorrang formaler oder inhaltlicher Kriterien je nach historischem Kontext und Struktur der Unterlagen durchaus unterschiedlich beantwortet werden kann. Als Folge der Diskussion der letzten 20 Jahre haben viele Archive Überlieferungsprofile und Sammlungskonzepte formuliert und publiziert, deren Rahmen im Wesentlichen jeweils durch den institutionellen Ort des Archivträgers und die von diesem formulierte Aufgabe des Archivs abgesteckt wird.

Unterhalb der Ebene der Gesamtprofile und -konzepte führt in der Praxis eine Analyse der Aufgaben und Tätigkeiten einer Provenienzstelle sowie der

Struktur ihrer Unterlagen dazu, dass für diese Provenienzstelle ein Bewertungskatalog erstellt wird. Dieser unterscheidet die Unterlagen einer anbietenden Stelle grundsätzlich danach,

- ob sie – da offensichtlich ohne bleibenden Wert – dem Archiv gar nicht erst angeboten werden müssen und gleich durch die anbietende Stelle vernichtet werden können,
- ob sie den Kernbereich der Aufgabenwahrnehmung einer Behörde etc. betreffen und deshalb a priori archivwürdig sind oder
- ob zur Bewertung eine direkte Autopsie notwendig ist.

Das inhaltliche Ziel der Bewertung ist, für den Zuständigkeits- und Arbeitsbereich des Archivträgers ein möglichst umfassendes Bild der Gesamtgesellschaft zu dokumentieren. Es geht nicht vorrangig darum zu zeigen, dass ein Ministerium, eine Behörde etc. existiert und gearbeitet hat, obwohl auch dies durchaus ein Zweck archivischer Dokumentation ist. In der Hauptsache soll aber dokumentiert werden, welche Aufgaben und wie sie diese Aufgaben wahrgenommen hat. Damit wird quasi aus der Perspektive einer Institution heraus ein Blick auf deren zeitgenössische Umwelt ermöglicht.

Es liegt auf der Hand, dass sich zwangsläufig größere Unterschiede in der Herangehensweise ergeben, je nachdem, ob es sich um eine politische Institution wie ein Bundesministerium oder eine exekutive Behörde etwa der Finanzverwaltung handelt. Eine besondere Herausforderung stellen die sogenannten Massenakten dar, also in großer Menge anfallende, gleichförmige Akten, die insbesondere die Leistungsverwaltung produziert. Die Ansätze für eine angemessene Reduzierung der Unterlagen reichen hier von statistischen Verfahren bis hin zu gezielten, oft das Wissen der anbietenden Stelle einbeziehenden Auswahl des Typischen und des Besonderen.

Entwickelt wurden Theorie und Praxis der archivischen Bewertung, um eine immer größer werdende Masse von Akten, Karten, Fotografien, Tönen, Filmen u. a. m. zu bewältigen und diese Masse methodisch geregelt in einen Überlieferungskörper zu transformieren. Gelten diese in der analogen Welt anwendbaren Prinzipien und Regeln auch noch im digitalen Zeitalter? Kann man jetzt nicht einfach alles digitalisieren und über immer leistungsfähiger werdende Suchmaschinen recherchierbar machen?

Sicher hat der Einsatz von IT die archivistische Arbeit von Grund auf revolutioniert. An die Stelle von Erschließungskarteien und Findbüchern sind längst moderne Formen getreten, die es den Benutzern in immer umfassenderem Maße ermöglichen, archivistische Erschließungsinformationen über das Internet zu nutzen. Und die Online-Präsentation von digitalisiertem oder auch genuin digitalem Archivgut schreitet zügig voran. Derzeit ist aber kaum vorstellbar, dass alles bereits existierende Archivgut digitalisiert und online gestellt werden könnte. Das liegt nur zum Teil an der schiereren Menge des Archivguts oder daran, dass viele Archivalien nur unter besonderen rechtlichen Bedingungen zugänglich gemacht werden können, die das Internet nicht garantieren kann. Digitalisierung und Online-Präsentation sind teuer und auch im dauerhaften Erhalt sehr kostenintensiv. Angesichts der vergleichsweise geringen Nutzungsfrequenz – um der Wahrheit willen muss das hier bekannt werden – erzwingen die bestehenden Ressourcenbeschränkungen auch hier ein strategisch durchdachtes und nachhaltiges Vorgehen.

Aber auch in der digitalen Welt liegt eine methodische Verdichtung der existierenden Information im Interesse des Benutzers. Mehr noch als in der konventionellen Welt der Papierakten ergibt sich die Qualität einer archivalischen Information im Falle genuin digitaler und damit prinzipiell flüchtiger Aufzeichnungen aus ihrem Provenienzkontext, der durch technische und inhaltliche Metadaten dauerhaft evident gehalten werden muss. Und die Geschwindigkeit von Recherche und Informationspräsentation ändert gar nichts an dem Umstand, dass ein Benutzer in der Fülle der Informationen diejenigen finden und erkennen muss, die für seine spezielle Fragestellung wichtig sind. Die Vorstellung gewaltiger unstrukturierter Datenmengen mit unklaren Kontexten ist aus archivistischer Sicht eine Horrorvorstellung. Daher gewinnen in der Welt flüchtiger digitaler Informationen die archivistischen Kernaufgaben der Übernahme, Bewertung und kontextbezogenen Erschließung eine nochmals gesteigerte Bedeutung.

# Die Dark Ages des Internet?

JENS CRUEGER

Wenn künftige Historiker eines Tages beginnen werden, die Geschichte des Internets intensiv zu erforschen, so werden sie vermutlich darüber klagen, wie quellenarm die Frühphase des Internets sei, wie wenige Quellendokumente aus dieser Zeit die Archive bewahrt hätten. Wichtige Belege über die Aktivitäten in den ersten Jahren des World Wide Web werden für diese Historiker schlicht nicht mehr auffindbar sein. Das mag zunächst verwundern, denn als quellenarm gelten in der Geschichtsforschung üblicherweise vor allem lange vergangene Epochen wie die Antike, wo Zeiten ganz ohne überlieferte Schriftlichkeit denn auch als „Dark Ages“ die Historiker an die Grenzen ihrer wissenschaftlichen Möglichkeiten führen. Die Zeitgeschichte hingegen, also die jüngst vergangene Gegenwart, bietet dem Historiker üblicherweise eher zu viel als zu wenig Quellenmaterial. Jedoch drohen der Internetgeschichtsschreibung ihre digitalen Dark Ages, denn ein Großteil der Webseiten aus den Anfangsjahren des World Wide Web ist bereits heute unwiederbringlich verloren.<sup>1</sup> Dies ist ein ernstes Problem für die Geschichtswissenschaft, die von der historisch-kritischen Analyse ihrer Quellen lebt, gar ein „Vetorecht der Quellen“ (niemals gegen die Aussage der Quellen!) kennt und bei fehlender Quellengrundlage gerne schweigen möchte. Somit drohen Teile der frühen Internetkultur, im Dunkel der Geschichte zu versinken.

Doch die Überlieferungssituation der frühen Internetquellen ist nicht das einzige Problem für eine Geschichtsschreibung des Digitalen: De facto gibt es diesen Zweig der Geschichtswissenschaft nämlich im deutschsprachigen

1 Haber, Peter: Digital Past. Geschichtswissenschaft im digitalen Zeitalter. München 2011, S. 61.



Raum noch gar nicht. Erst sehr wenige Historiker nehmen sich der Geschichte des Digitalen überhaupt an, es fehlt eine an die Digitalität angepasste Methodik, und die Institutionen, die für die Fachkommunikation und die Qualifizierung wichtig wären, befinden sich noch im Aufbaustadium. Der vorliegende Text soll daher als Plädoyer für eine Geschichtsforschung und Geschichtsschreibung der digitalen Welt verstanden werden. Es geht dabei nicht nur darum, eine medien- oder technikgeschichtliche Perspektive einzunehmen, vielmehr muss die digitale Welt – allen voran das Internet – als ein Kultur- und Diskursraum geschichtswissenschaftlich erfasst werden. Es gilt, eine Kulturgeschichte des Internets zu schreiben, die sich nicht nur auf das Innenleben des World Wide Web beschränkt, sondern die historischen Wechselwirkungen und Austauschbeziehungen zwischen Online- und Offline-Welt gleichermaßen beschreibt. Ohne eine solche Geschichtsschreibung des Internet sind viele Gegenwartsdiskussionen kaum interpretierbar, denn bereits heute gehört die Online-Perspektive gewissermaßen zum historischen Kontext unserer Gegenwartsphänomene, auch wenn diese auf den ersten Blick nichts mit dem Internet zu tun haben. Gleichwohl werden Ereignisse aus der Offline-Welt im Internet nachvollzogen, dokumentiert, diskutiert, bewertet, Themen werden gesetzt und Stimmungen erzeugt. So spiegelt das Internet die Gegenwartsgeschichte und übt gleichzeitig einen aktiven Einfluss darauf aus. Es ist daher dringend geboten, den Internetdiskurs stärker in den Fokus von Geschichtsforschung zu nehmen.

Überdies stellt eine geschichtswissenschaftliche Einordnung dessen, was oft als digitale Revolution bezeichnet wird, eine wichtige Zukunftsaufgabe für die Geschichtswissenschaft dar. Die Digitalisierung unserer Gesellschaft kulturhistorisch einzuordnen und Deutungsangebote zu geben ist von offenkundiger Relevanz. Wann immer nach einem Narrativ der Digitalisierung gerufen wird, der einer kulturellen Spaltung unserer Gesellschaft in die „Onliner“ und die „Offliner“ entgegenwirken soll, kann dies als ein Aufruf an die Adresse der Geschichtswissenschaft verstanden werden. Die Geschichtsschreibung vermag, solche Narrative zu entwickeln, und hätte somit die Möglichkeit, die digitale Revolution all jenen zu erklären, die bislang nach einer solchen Erklärung suchen.

Doch leider ist der Status quo in der Geschichtswissenschaft derzeit sehr ernüchternd. Die von Christian Keitel und Peter Haber 2008 durchgeführten Befragungen unter Fachhistorikern lieferten eine erschreckende Bilanz, konn-

ten sie doch „keinen Historiker ermitteln, der sich bereits mit dem Thema [genuin digitaler Dokumente] beschäftigen würde“.<sup>2</sup> So haben zwar Medienhistoriker und jene, die sich mit der „Geschichte der technischen Entwicklung oder der Digitalisierung beschäftigen“, die Bedeutung genuin digitaler Quellen sehr wohl erkannt, von der „Mehrheit historischer Disziplinen“ indes werden „genuine Internetquellen gänzlich“ ignoriert.<sup>3</sup> Zwar lassen sich mittlerweile erste Institutionalisierungsschritte der sogenannten digitalen Geschichtswissenschaft verzeichnen, das Netzwerk digitale Geschichtswissenschaften (NDG)<sup>4</sup> hat mit der ersten Ausgabe der „Zeitschrift für digitale Geschichtswissenschaften“ (ZDG) ein Publikationsorgan für dieses Thema geschaffen, das sinnigerweise als Open Access publiziert wird.<sup>5</sup> Auf dem Mainzer Historikertag im September 2012 gründete sich überdies die Arbeitsgemeinschaft für „Digitale Geschichtswissenschaft“ im Verband der Historiker und Historikerinnen Deutschlands e. V., die sich zum Ziel gesetzt hat, „Potenziale der Digital Humanities bzw. eHumanities für die Geschichtswissenschaft zu vermessen, zu nutzen und sichtbar zu machen“.<sup>6</sup> Jedoch ist das Themenfeld der digitalen Geschichtswissenschaft(en) breit, reicht es doch von digitalen Werkzeugen und Methoden (u. a. Text Mining, Opinion Mining) über die digitale Aufbereitung von Quellen (Digitalisate, digitale Editionen), die digitale Vermittlung von Geschichte (Museen, Unterricht, E-Learning) bis hin zur digitalen Fachkommunikation.

Catherina Schreiber sieht die „Vorbehalte, auf die die Verwendung von Internetquellen bislang in der Geschichtswissenschaft zumeist stieß“, im „Fehlen einer

- 2 Keitel, Christian: Über den Zusammenhang zwischen Quellenkritik und Informationserhalt. Ergebnisse der Anfrage „Forschen mit ‚digitalen Quellen‘“. 2009. <http://hsozkult.geschichte.hu-berlin.de/forum/id=1173&type=diskussionen> (Letzter Aufruf: 15.08.2013).
- 3 Schreiber, Catherina: Genuine Internetdaten als historische Quellen – Entwurf einer korrelationalistischen Quellentheorie. In: Zeitschrift für digitale Geschichtswissenschaften 1 (2012). <http://universaar.uni-saarland.de/journals/index.php/zdg/article/view/292> (Letzter Aufruf: 25.05.2013), S. 1-15, S. 2.
- 4 Vgl. [www.ndg.uni-saarland.de/](http://www.ndg.uni-saarland.de/) (Letzter Aufruf: 25.05.2013).
- 5 Vgl. <http://universaar.uni-saarland.de/journals/index.php/zdg/index> (Letzter Aufruf: 25.05.2013).
- 6 Vgl. [www.historikerverband.de/arbeitsgruppen/ag-digitale-gw.html](http://www.historikerverband.de/arbeitsgruppen/ag-digitale-gw.html) (Letzter Aufruf: 25.05.2013).

systematischen Quellentheorie für genuine Internetquellen“ begründet.<sup>7</sup> Die Diskussion über die Quellenkritik des Digitalen befindet sich immerhin in den Anfängen, wozu Artikel wie jene von Eva Pflanzelter, Catherina Schreiber und Peter Haber einen wichtigen Beitrag geleistet haben.<sup>8</sup> Im Raum steht dabei die von Peter Haber aufgeworfene Frage, ob es einer den „digitalen Medialitäten“ angepassten Hilfswissenschaft bedarf oder der digitale Wandel „nicht vielmehr den epistemologischen Kern der geschichtswissenschaftlichen Arbeitsweise“ berührt und daher „als ein Kernbestand des Faches konzeptionalisiert werden“ muss.<sup>9</sup>

Digitale Quellen, die als Grundlage historiographischer Forschung dienen können,<sup>10</sup> umfassen einerseits retrodigitalisierte Dokumente, die aus gedruckten Dokumenten nachträglich mittels Scan in ein digitales Format (als Bild- oder Volltextdatei) gebracht wurden. Die zweite Gruppe digitaler Quellen sind die sogenannten genuin digitalen Quellen (digital born). Einerseits sind dies hybrid hergestellte Materialien, die digital produziert wurden, um als Vorstufe für einen späteren Druck zu dienen. Dabei handelt es sich teilweise um reine Druckvorlagen, oftmals werden die Dokumente jedoch auch mit Metadaten und Schutzmechanismen für eine weitere hybride Nutzung versehen. Andererseits existieren im Internet und anderen digitalen Kontexten gigantische Mengen an Quelldokumenten, die ausschließlich in digitaler Form existieren und in keiner Absicht stehen, aus dem digitalen Zustand jemals in eine gedruckte Form gebracht zu werden. Dazu gehören „neben den gängigen in Textform vorliegenden Quellen aus Internetseiten, Chat- und Forenbeiträgen, Wiki- und Blogartikeln natürlich

7 Schreiber, Catherina: Genuine Internetdaten als historische Quellen – Entwurf einer korrealistischen Quellentheorie. In: Zeitschrift für digitale Geschichtswissenschaften 1 (2012). <http://universaar.uni-saarland.de/journals/index.php/zdg/article/view/292> (Letzter Aufruf: 25.05.2013), S. 1-15, S. 4.

8 Pflanzelter, Eva: Von der Quellenkritik zum kritischen Umgang mit digitalen Ressourcen. In: Gasteiner, Martin/Haber, Peter: Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften. Wien u. a. 2010, S. 39–49; Schreiber, Catherina: Genuine Internetdaten als historische Quellen – Entwurf einer korrealistischen Quellentheorie. In: Zeitschrift für digitale Geschichtswissenschaften 1 (2012). <http://universaar.uni-saarland.de/journals/index.php/zdg/article/view/292> (Letzter Aufruf: 25.05.2013), S. 1-15; Haber, Peter: Digital Past. Geschichtswissenschaft im digitalen Zeitalter. München 2011, S. 104-112.

9 Haber, Peter: Digital Past. Geschichtswissenschaft im digitalen Zeitalter. München 2011, S. 106.

10 Im Folgenden nach ebd., S. 103.

audiovisuelle Medien, aber auch Emails, Animationen, virtuelle Kontakte in virtuellen Gemeinschaften sowie graphisch umgesetzte virtuelle Realitäten von Online-Spielen bis hin zu Source-Codes und Computerviren“.<sup>11</sup>

Die Frage, welche Teile dieser digitalen Datenflut bewahrenswert und daher archivierungswürdig sind, lässt sich abstrakt kaum beantworten. Historiker sind es, die gegenüber den Archiven artikulieren müssten, welche digitalen Quellen ihnen in welcher Form erhaltenswürdig erscheinen. Allerdings stellt das Forschungsinteresse der Historiker stets eine Momentaufnahme dar, und künftige Historikergenerationen fragen möglicherweise nach ganz anderen digitalen Quellen. Zumal sich die Geschichtswissenschaft der digitalen Welt erst am Anfang ihrer Entwicklung befindet, sind derzeit kaum Aussagen darüber möglich, welche Arten von digitalen Quellen archiviert werden sollten. Außerdem bedarf es zunächst einer definierten quellenkritischen Methodik, aus der die Archive dann ableiten können, „wie digitale Objekte überhaupt erhalten werden sollen“.<sup>12</sup> Dies betrifft etwa Eigenschaften archivierter digitaler Quellen (significant properties) und die Frage, wie trotz Transformationen die Authentizität der Archivalien gewahrt werden kann. Da digitale Quellen „nicht unveränderlich gegeben“ sind, sondern „von den Archiven in Annahme von Benutzerinteressen konstituiert“ werden,<sup>13</sup> sind diese Fragen ebenso drängend wie fruchtbar.

Solange allerdings primär auf die äußere Quellenkritik und vor allem auf die Frage der Authentifizierung abgehoben wird, verstellt dies leider den Blick auf Probleme der inneren Quellenkritik. Dadurch werden „wichtige Potenziale, die durch das veränderte Kommunikationsschema [...] gegeben sind“, leider kaum thematisiert.<sup>14</sup>

11 Schreiber, Catherina: Genuine Internetdaten als historische Quellen – Entwurf einer korrealistischen Quellentheorie. In: Zeitschrift für digitale Geschichtswissenschaften 1 (2012). <http://universaar.uni-saarland.de/journals/index.php/zdg/article/view/292> (Letzter Aufruf: 25.05.2013), S. 1-15, S. 2.

12 Keitel, Christian: Über den Zusammenhang zwischen Quellenkritik und Informationserhalt. Ergebnisse der Anfrage „Forschen mit ‚digitalen Quellen‘“. 2009. <http://hsozkult.geschichte.hu-berlin.de/forum/id=1173&type=diskussionen> (Letzter Aufruf: 15.08.2013).

13 Ebd.

14 Schreiber, Catherina: Genuine Internetdaten als historische Quellen – Entwurf einer korrealistischen Quellentheorie. In: Zeitschrift für digitale Geschichtswissenschaften 1 (2012). <http://universaar.uni-saarland.de/journals/index.php/zdg/article/view/292> (Letzter Aufruf: 25.05.2013), S. 1-15, S. 6.

Schreiber führt beispielsweise das veränderte Konzept der Autorenschaft und „Auswirkungen der Digitalität auf die Quelleninhalte“ an.<sup>15</sup> Ihr Credo, wonach die „meisten Grundsätze einer digitalen Quellenkritik [...], wiewohl sie ihren Ansatz aus der spezifischen Medialität und Materialität genuiner Internetquellen nehmen, auch auf analoge oder hybride Medien angewandt werden“ könnten, stellt einen wertvollen Diskussionsimpuls dar.<sup>16</sup> Da Internetquellen keinem homogenen Quellentypus zuzuordnen sind, insbesondere Internettexpte in vieler Hinsicht in vor-digitalen Traditionen stehen und eine Wirkungsgeschichte digitaler Quellen oftmals auch offline verläuft, lassen sich digitale Quellen nicht losgelöst von der etablierten Quellenkritik materieller Artefakte interpretieren.<sup>17</sup>

Digitale Quellen stellen jedoch besondere Anforderungen an die historisch-kritische Methodik, von denen ich nur einige beispielhaft anführen möchte. So muss etwa die ästhetische Dimension stärker erfasst werden, als die sehr auf Textkritik konzentrierte Geschichtswissenschaft dies üblicherweise tut. Diese Ästhetik umfasst sowohl die graphische Dimension etwa von Webseitenlayouts oder digitalen Welten als auch die formale Ästhetik von Quellcodes. Ebenso muss der materielle Charakter der Gerätschaften, die zum Zugriff auf das Internet nötig sind (PCs, Smartphones, Tablets), als Teil des Aneignungsprozesses reflektiert werden. Somit ist eine Interpretation von Internetquellen „zwingend multimodal“.<sup>18</sup>

Die Findbarkeit digitaler Ressourcen im World Wide Web mittels direkter Verlinkung, aber ebenso auch mittels des Rankings in Volltextsuchmaschinen stellt einen für die Quellenkritik nicht unwesentlichen Faktor da. Dabei geht es nicht nur um die wirkungsgeschichtliche Frage nach jener „Ökonomie der Aufmerksamkeit, die über das Suchen respektive das Gefundenwerden entscheidet“.<sup>19</sup> Vielmehr lässt sich anhand des Quelltextes einer Ressource etwa ableiten, ob die betreffende Seite vom Autor des Quelltextes mit der Intention der guten Findbarkeit programmiert wurde oder nicht. Hieran knüpft sich die Frage nach

15 Ebd., S. 6.

16 Ebd., S. 7.

17 Ebd., S. 8.

18 Ebd., S. 12.

19 Haber, Peter: Digital Past. Geschichtswissenschaft im digitalen Zeitalter. München 2011, S. 87.

dem Adressatenkreis sowie nach der technischen und soziotechnischen Zugänglichkeit (Stichwort digital divide).<sup>20</sup>

Die hohe Frequenz an Versionierungen bei vielen digitalen Quellen (etwa bei Webseiten von Tageszeitungen) deutet an, wie wichtig eine bewusste Archivierungsstrategie aus Sicht der Geschichtswissenschaft ist. Dabei stellen nicht nur in kurzer Folge aktualisierte redaktionelle Inhalte ein historiographisches Quellenproblem dar. Denn genauso müssten die auf diesen Webseiten eingeblendeten kommerziellen Werbebanner eigentlich jeweils archiviert werden, um die Darstellungsvarianten der Webseite vollständig nachvollziehbar zu machen.

Es ist dringend an der Zeit, dass in der Geschichtswissenschaft ein stärkerer Fokus auf das Internet und die Digitalisierung gesetzt wird. Es geht dabei nicht nur um die Vermeidung von weiteren Quellenverlusten und drohenden Dark Ages. Vielmehr geht es um die kulturhistorische Einordnung und Interpretation jenes gewaltigen gesellschaftlichen Transformationsprozesses, der durch die Digitalisierung bedingt wurde und weiterhin bedingt wird. Die Geschichtswissenschaft wäre in der Lage, hierzu relevante Beiträge zu erarbeiten. Sie sollte dies auch tun.

20 Schreiber, Catherina: Genuine Internetdaten als historische Quellen – Entwurf einer korrelationalistischen Quellentheorie. In: Zeitschrift für digitale Geschichtswissenschaften 1 (2012). <http://universaar.uni-saarland.de/journals/index.php/zdg/article/view/292> (Letzter Aufruf: 25.05.2013), S. 1-15, S. 7f.



# Eine Strategie zur Förderung der digitalen Langzeitarchivierung

GEORG REHM

## *Einleitung und Überblick*

Das Thema digitale Langzeitarchivierung wird in Deutschland zwar seit einigen Jahren diskutiert, eine tatsächliche operative Umsetzung durch die Implementierung entsprechender Infrastrukturen oder das Anbieten von Dienstleistungen steckt allerdings noch in den Kinderschuhen.<sup>1</sup> Über die Gründe kann vorzüglich spekuliert werden – vermutlich liegt ihr Kern in dem Umstand, dass es sich um ein Thema handelt, das inhärent eine Vielzahl verschiedener Gebiete involviert und somit nur interdisziplinär bearbeitet und umgesetzt werden kann. Die digitale Langzeitarchivierung berührt u. a. die Informations- und Archivwissenschaft (Umfang, Struktur und Inhalte von Metadatenvokabularen, Prinzipien der Archivierung, Strukturierung, Annotation und Verschlagwortung von Objekten etc.), die Bibliothekswissenschaft (bestmögliche Distribuierung von Informationen, Katalogisierungsprinzipien, Informationsversorgung), die Rechtswissenschaft (Urheberrechts- und Copyrightfragen, z. B. im Hinblick auf das Kopieren und Archivieren digitaler Objekte aus dem Web), die Kulturwissenschaft (etwa bezüglich der Frage nach der Differenzierung zwischen zu archivierenden und nicht zu archivierenden Objekten; ab welcher Schöpfungshöhe soll ein digitales Objekt für die Nachwelt langfristig gespeichert werden?), die Geschichtswissenschaft und Gedächtnisinstitutionen (u. a. als Nutzer und Anbieter von Archiven digitaler Objekte), sogar die Architektur (wie sollte ein Rechenzentrum baulich beschaffen sein, um ein digitales Archiv für die nächsten 500 Jahre sicher beherbergen zu können, und wo sollte es im besten Fall aufgebaut werden, um nicht von Naturkatastrophen betroffen zu sein?) sowie



die Informatik und Informationstechnologie (u. a. schnelle, redundant ausgelegte Netzwerke, Massenspeicher, Backups, Zugriffssysteme, APIs, Nutzeroberflächen und Anfragesprachen, Import und Export von Daten, Suchalgorithmen, Datenintegrität, Revisionssicherheit etc.). Je nach Typ und Natur der zu archivierenden digitalen Objekte kommen gegebenenfalls noch die domänenspezifischen Fachwissenschaften hinzu. Dazu später mehr.

Die verschiedenen Beiträge in diesem Band verdeutlichen, dass die digitale Langzeitarchivierung eine wichtige gesellschaftliche Aufgabe ist. Ohne eine Strategie und die Initiierung konkreter Maßnahmen besteht die Gefahr des unwiederbringlichen Verlustes einer Vielzahl digitaler und analoger Kulturgüter. Der Erarbeitung und speziell auch der Umsetzung einer derartigen Strategie steht jedoch die Realität gegenüber. Diese ist insbesondere gekennzeichnet von einer globalen Wirtschaftskrise, die sich auch auf die Bereitstellung von Fördergeldern für Themengebiete auswirkt, die von Gesellschaft und Politik zwar als relevant und wichtig wahrgenommen werden, für die aber letztlich keine ausreichenden Mittel zur Implementierung zur Verfügung gestellt werden können. Die oben punktuell aufgeführten beteiligten Gebiete und die Bandbreite der zu bearbeitenden Fragestellungen illustrieren, wie umfangreich ein derartiges Mammutvorhaben tatsächlich wäre. Gerade in einer Zeit, in der über Konzepte wie das Recht auf Vergessenwerden<sup>1</sup> und den digitalen Radiergummi diskutiert wird, dürfte es eine Herausforderung sein, genügend Aufmerksamkeit zu bekommen und anschließend Ressourcen in Form von Fördermitteln für die digitale Langzeitarchivierung einzuwerben. Andererseits ist das Thema gerade jetzt mit den Spionageprogrammen Prism der National Security Agency (NSA) und Tempora des UK Government Communications Headquarters (GCHQ) wieder im Mittelpunkt einer global geführten öffentlichen Debatte angelangt. Dies stellt ein großes Risiko für die Fördersituation dar, kann aber auch mit einem positiven Effekt verbunden sein. Dieser Aspekt wird im letzten Abschnitt wieder aufgegriffen.

Aufgrund der Komplexität der Thematik, der immensen Höhe der zu erwartenden Kosten und des gewaltigen Umfangs der zu archivierenden Daten erscheint es zum derzeitigen Stand der Debatte unwahrscheinlich, dass in absehbarer Zeit eine nationale Strategie zur digitalen Langzeitarchivierung entwickelt und auch umgesetzt werden wird. Der drohende Verlust digitaler Kulturgüter ist

1 Vgl. [http://de.wikipedia.org/wiki/Recht\\_auf\\_Vergessenwerden](http://de.wikipedia.org/wiki/Recht_auf_Vergessenwerden) (Letzter Aufruf: 15.08.2013).

selbstverständlich nicht auf die Bundesrepublik beschränkt – ganz im Gegenteil. Die digitale Langzeitarchivierung ist keine nationale, sondern eine internationale Herausforderung und Aufgabe. Die Diskussion sollte zusätzlich zur nationalen auch auf die internationale Ebene im Europaparlament und in die Europäische Kommission verlagert und dort verstärkt werden – gleiches gilt für die UNESCO. Bis zur jüngsten Restrukturierung innerhalb der Europäischen Kommission wurden gelegentlich Projekte im Bereich Cultural Preservation and Digital Heritage gefördert, und zwar mit einem Gesamtbudget von ca. 15 bis 20 Millionen Euro pro Jahr.<sup>2</sup> Seit der erwähnten Umstrukturierung heißt die Abteilung nun jedoch Creativity.<sup>3</sup> Entsprechend dieser Neuausrichtung hin zu digitaler Kreativität wurden die Mittel für digitale Nachhaltigkeit drastisch gekürzt. Mit anderen Worten: Das Thema Langzeitarchivierung besitzt derzeit und in den kommenden Jahren auf der Ebene der Europäischen Kommission keine Sichtbarkeit, weshalb nur eingeschränkte Ressourcen zur Verfügung stehen. Ein möglicher Weg, dennoch zu einer Förderung dieses wichtigen Themas zu gelangen, könnte es also sein, die digitale Langzeitarchivierung mit verwandten Themen zu verbinden, die derzeit mehr Visibilität genießen, und auf diese Weise nach und nach auch die digitale Nachhaltigkeit wieder in den Fokus der Aufmerksamkeit zu rücken.

#### *Digitale Langzeitarchivierung: Status quo*

Die digitale Langzeitarchivierung ist ein internationales Thema und eine internationale Herausforderung. In Deutschland wird das Thema nicht auf Bundes-, sondern in erster Linie auf Landesebene bearbeitet, sodass Expertise, Infrastrukturen, Prozesse, Lösungen, Erfahrungen und Best-Practice-Ansätze von Bundesland zu Bundesland und von Institution zu Institution stark variieren.

Die zentrale Initiative der letzten Jahre ist nestor (Network of Expertise in long-term storage and availability of digital resources), das deutsche Kompetenznetzwerk zur digitalen Langzeitarchivierung.<sup>4</sup> In nestor arbeiten Bibliotheken, Museen und Archive gemeinsam zum Thema Langzeitarchivierung und

2 Vgl. [http://cordis.europa.eu/fp7/ict/telearn-digicult/digicult-projects\\_en.html](http://cordis.europa.eu/fp7/ict/telearn-digicult/digicult-projects_en.html) (Letzter Aufruf: 15.08.2013).

3 Vgl. [http://cordis.europa.eu/fp7/ict/creativity/creativity\\_en.html](http://cordis.europa.eu/fp7/ict/creativity/creativity_en.html) (Letzter Aufruf: 15.08.2013).

4 Vgl. [www.langzeitarchivierung.de](http://www.langzeitarchivierung.de) (Letzter Aufruf: 15.08.2013).

Langzeitverfügbarkeit digitaler Objekte. nestor wurde in zwei Projektphasen (2003–2006, 2006–2009) mit insgesamt ca. 1,2 Millionen Euro vom BMBF gefördert.<sup>5</sup> Auch nach dem Ablauf der Förderung ist nestor weiter aktiv und agiert als selbstständiger Kooperationsverbund mit Partnern wie z. B. der Bayerischen Staatsbibliothek, dem Bibliotheksservice-Zentrum Baden-Württemberg, der Deutschen Nationalbibliothek und dem Institut für Deutsche Sprache. Nach eigenen Angaben pflegt nestor enge Kontakte zu entsprechenden Initiativen anderer Länder und beteiligt sich an europäischen und internationalen Projekten.

nestor hat im Rahmen seiner Laufzeit und auch im Anschluss wertvolle Impulse für die Entwicklung eines allgemeinen Rahmens für das Thema sowie einer Strategie geliefert. Die zahlreichen Ergebnisse von nestor flossen ein in Handbücher zur Langzeitarchivierung,<sup>6</sup> drei DIN-Standards sowie einige Ratgeber und Broschüren.<sup>7</sup> Außerdem wurde ein Konzept für die Aus- und Weiterbildung entwickelt. Ebenfalls hat nestor einen nicht zu unterschätzenden ersten Beitrag zum Aufbau einer Community rund um das Thema Langzeitarchivierung geliefert.

Konkrete technische Lösungen für die Langzeitarchivierung hat nestor nicht hervorgebracht. Diese waren allerdings auch nicht das Ziel der Initiative, werden jedoch von den beteiligten Einrichtungen und Gedächtnisinstitutionen dringend und zwingend benötigt.

### *Gemeinsame Bedarfe – gemeinsame Lösungen*

Derzeit verwenden interessierte oder das entsprechende Mandat besitzende Institutionen für die digitale Langzeitarchivierung eine heterogene Mischung

5 Vgl. [www.langzeitarchivierung.de/Subsites/nestor/DE/Header/Ueberuns/Projektartikel.html](http://www.langzeitarchivierung.de/Subsites/nestor/DE/Header/Ueberuns/Projektartikel.html) (Letzter Aufruf: 15.08.2013).

6 Neuroth, Heike/Oßwald, Achim/Scheffel, Regine/Strathmann, Stefan/Huth, Karsten (Hrsg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Göttingen 2010. [www.langzeitarchivierung.de/Subsites/nestor/DE/Publikationen/Handbuch/handbuch\\_node.html](http://www.langzeitarchivierung.de/Subsites/nestor/DE/Publikationen/Handbuch/handbuch_node.html) (Letzter Aufruf: 15.08.2013); Neuroth, Heike/Strathmann, Stefan/Oßwald, Achim/Scheffel, Regine/Klump, Jens/Ludwig, Jens (Hrsg.): Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme. Version 1.0. Göttingen 2012. [www.langzeitarchivierung.de/Subsites/nestor/DE/Publikationen/Handbuch/handbuch\\_node.html](http://www.langzeitarchivierung.de/Subsites/nestor/DE/Publikationen/Handbuch/handbuch_node.html) (Letzter Aufruf: 15.08.2013).

7 Vgl. [www.langzeitarchivierung.de/Subsites/nestor/DE/Publikationen/publikationen\\_node.html](http://www.langzeitarchivierung.de/Subsites/nestor/DE/Publikationen/publikationen_node.html) (Letzter Aufruf: 15.08.2013).

verschiedenster Ansätze, Softwarewerkzeuge, Datenbanken und Repository- sowie Content-Management-Systeme. Ein wesentlicher Grund für den aktuellen, in jeder Hinsicht von Heterogenität geprägten Stand besteht in der Tatsache, dass die digitale Langzeitarchivierung in Deutschland aufgrund der Kulturhoheit Ländersache ist und somit alle Länder und sogar die einzelnen Organisationen innerhalb der Länder ihre eigenen Wege verfolgen und eigene Lösungen erarbeiten, d. h. entweder selber Software entwickeln oder mal diese und mal jene Software einsetzen und dabei immer wieder den gleichen Problemen begegnen.

Der vorliegende Beitrag ruft dazu auf, nicht nur eine nationale, sondern eine internationale Strategie für digitale Langzeitarchivierung zu erarbeiten und diese, möglichst europaweit, als digitale Infrastruktur technisch umzusetzen und interessierten Einrichtungen und gegebenenfalls auch Unternehmen anzubieten. Zahlreiche Gründe sprechen für einen derartigen europaweit koordinierten Ansatz.

Zunächst ist zwischen zentralen und dezentralen Aufgaben zu differenzieren. Die zentralen, immer wiederkehrenden Aufgaben beinhalten technische Aspekte der Speicherung und Verfügbarmachung von Daten, die Spezifikation abstrakter Metadatenschemata sowie die Bearbeitung rechtlicher Fragestellungen und Herausforderungen. Es ergibt keinerlei Sinn, dass z. B. dezentral arbeitende Gedächtnisinstitutionen das Rad immer wieder neu erfinden, Software für Repositorien evaluieren, Rechtsexperten beauftragen oder Untersuchungen verschiedener Query-Sprachen durchführen. Derartige abstrakte technische Aspekte sollten im besten Fall von einer gemeinsamen technischen Infrastruktur angeboten und bestmöglich unterstützt werden. Eine solche technische Infrastruktur kann beispielsweise konzeptualisiert werden als eine Art virtueller Cloud-Speicher, der über verabredete Programmierschnittstellen (APIs) zugreifbar ist. Aus Gründen der Ausfallsicherheit und Redundanz wäre es hilfreich, ein derartiges Cloud-Archiv als verteilten Speicher aufzubauen, an dessen Bereitstellung viele verschiedene Rechenzentren teilnehmen – auch über Ländergrenzen hinweg.

Der evidenteste Fall einer zentralen, maximal generischen Aufgabe betrifft die Archivierung von Webseiten: Hier reichte letztlich eine einzelne Einrichtung aus, die sämtliche Webseiten Deutschlands, Europas oder auch der Welt konstant traversiert, archiviert und langfristig verfügbar hält – im Idealfall auf mehrere, räumlich weit entfernte, redundant speichernde Datenzentren verteilt. An einer solchen Aufgabe könnte sich z. B. ein institutionalisiertes Konsortium aus mehreren europäischen Datenzentren beteiligen. Die Präzedenzfälle sind

alle weltweit operierenden Suchmaschinen (Google, Bing) sowie das Internet Archive.<sup>8</sup> Wenn es um die Archivierung typischer digitaler Objekte mit einer generischen, sehr weit verbreiteten Struktur und etablierten Formaten geht (Webseiten, Sounddateien, Filme etc.), können und müssen derartige zentrale Dienste Verwendung finden, und zwar vollautomatisch. Die selektive, punktuell, auf Anfrage durchgeführte Archivierung einzelner oder kleinerer Webseiten kann und darf keine dezentral durchzuführende Aufgabe kleinerer oder größerer Landesbibliotheken sein.

Zu den dezentralen Aufgaben gehören die regional-, institutions-, domänen- und objektspezifischen sowie alle fachwissenschaftlichen Aspekte; die Grenze zwischen zentralen und dezentralen sowie zwischen domänenunabhängigen und domänenabhängigen Aufgaben ist dabei nicht als trennscharf zu verstehen. Dezentrale Aufgaben sind etwa die Spezifizierung von Metadatenschemata für bestimmte, insgesamt eher seltene oder ungewöhnliche Typen analoger oder digitaler Objekte oder die Entwicklung hochgradig spezialisierter Scanner oder Archivierungsverfahren. Gedächtnisinstitutionen und vergleichbare Einrichtungen, die sich mit speziellen Domänen oder Fachwissenschaften beschäftigen, können wiederum im Hinblick auf ihre spezifischen Lösungen kooperieren und von Synergien profitieren. Je inhärent dezentraler und fachwissenschaftlicher die Archivierungsaufgabe, desto unwahrscheinlicher ist es, diese vollautomatisch durchführen zu können.

Die Planung und die Umsetzung eines derartigen Vorhabens verursachen immense Kosten. Daher ist es zwingend notwendig, Synergien mit ähnlichen Initiativen zu identifizieren und diese bestmöglich zu nutzen. Nur wenn eine derartige kontinentweite Allianz aus Archiven, Bibliotheken, Universitäten, Administrationen, Museen und anderen Gedächtnisinstitutionen usw. gebildet wird, kann die Umsetzung eines derartigen sehr ambitionierten Plans in den Bereich des Machbaren rücken.

### *Technische Herausforderungen*

Ein derartiges Vorhaben wird mit verschiedenen technischen Herausforderungen konfrontiert. Exemplarisch seien an dieser Stelle einige wenige genannt, die sich auf die Archivierung des World Wide Web beziehen: In den Anfangstagen

8 Vgl. [www.google.com](http://www.google.com); [www.bing.com](http://www.bing.com); [www.archive.org](http://www.archive.org) (Letzter Aufruf: 15.08.2013).

und -jahren des World Wide Web bestanden Webseiten aus statischen Dokumenten, die beim Aufruf aus dem Dateisystem des Webservers geladen und über das Netzwerk zur Darstellung an den Browser des Nutzers durchgereicht wurden. Derartige statische Dokumente sind heutzutage nur noch die Ausnahme, denn Dokumente bestehen typischerweise aus mehreren einzelnen Komponenten, Widgets, Mashups und Objekten, die aus verschiedenen Quellen stammen und rechtlich verschiedene Urheber besitzen. Nach dem Aufruf wird die Seite entweder vom Webserver oder vom Browser dynamisch zusammengesetzt; ein Aufruf aus Berlin kann zu einem anderen Dokument führen als ein Aufruf aus Köln, Hamburg, London, New York, Hong Kong oder Lima. Variationen können stattfinden in Bezug auf die Sprache, die grafische Gestaltung sowie Anzahl und Ausprägung dynamisch eingebetteter Objekte und auch auf die Frage, in welchen (anderen) Webservices der Nutzer gerade angemeldet ist (Stichworte: Personalisierung von Inhalten, personalisierte Werbung). Welche Variante, welche Sicht eines Dokuments kann oder soll archiviert werden? Ist diese Variante archivierungswürdiger als jene? Und: Ergibt das Archivieren eines derart emergenten Mediums überhaupt einen Sinn, wenn der individuelle Wahrnehmungs- und Erlebniskontext nicht zur Verfügung steht, wenn also z. B. nur einzelne Bestandteile einer Seite im Archiv vorliegen und andere nicht, wenn aus der dynamischen und interaktiven Webseite ein statistisches Dokument geworden ist, das zwar als solches betrachtet, mit dem aber nicht interagiert werden kann, weil Hyperlinks und eingebettete Objekte nicht funktionieren? Im Hinblick auf genuine Webapplikationen, die nicht etwa wie digitale Dokumente wirken, sondern tatsächlich die Anmutung typischer Softwareanwendungen besitzen (z. B. Google Mail, Apple iWorks, Bing Maps etc.), verkompliziert sich dieses Problem um ein Vielfaches.<sup>9</sup>

Doch nicht nur die Ebene der Inhalte, auch die Ebene von Software und Hardware spielen zentrale Rollen: Abhängig vom verwendeten Browser kann sich das Aussehen eines Dokuments verändern. Gerade die Anfangsjahre des World Wide Web waren gekennzeichnet von den Browser wars, in denen insbesondere Netscape und Microsoft versuchten, die Konkurrenz durch neue, allerdings proprietäre Funktionen zu übertrumpfen, um auf diese Weise mehr Marktanteile zu bekommen.<sup>10</sup> Während Standardisierungsbemühungen seitens des World Wide Web

9 Ausführlich hierzu Rehm, Georg: Hypertextsorten. Definition - Struktur - Klassifikation. Norderstedt 2007.

10 Vgl. [http://en.wikipedia.org/wiki/Browser\\_wars](http://en.wikipedia.org/wiki/Browser_wars) (Letzter Aufruf: 15.08.2013).

Consortium (W3C) und anderer Organisationen die Situation hier mittlerweile deutlich entschärft haben, stellt sich dennoch die Frage, ob man nicht zusätzlich auch die Ebene der Software und auch der Hardware in den Archivierungsprozess einbeziehen sollte. Ist es ein authentisches Erlebnis, ein HTML-Dokument von 1995 mit einem aktuellen Browser wie etwa Chrome oder Firefox darzustellen und zu lesen? Oder sollte das Dokument nicht vielmehr mit der Software und im besten Fall auch der Hardware (inklusive typischer Internet-Bandbreite) seiner Zeit dargestellt werden? Der erste grafische Webbrowser überhaupt wurde von Tim Berners-Lee auf einem Rechner der Firma NeXT entwickelt und trug den Namen WorldWideWeb. Ist es die Aufgabe von Gedächtnisinstitutionen, ein authentisches Gesamterlebnis des Bedienens dieser längst nicht mehr verfügbaren Software auf einer seit vielen Jahren nicht mehr verfügbaren Hardware zu ermöglichen? Eventuell können Emulatoren und entsprechende Softwarearchive einen entscheidenden Beitrag leisten, um derartige Zeitreisen zu realisieren. Gegebenenfalls könnte hier auch das Erhalten exemplarischer Hard- und Software sowie typischer Webseiten aus verschiedenen Epochen des World Wide Web und klar gekennzeichneten Designperioden bzw. mit unterschiedlichem Funktionsumfang ein möglicher Ansatz sein, wie er auch in der Welt der Dinge im Musealen verfolgt wird.

### *Synergien durch verwandte Initiativen*

Bei den verschiedenen Vorträgen und Diskussionen im Rahmen der Initiative „Nachhaltigkeit in der Digitalen Welt“ des Internet und Gesellschaft Collaboratory e. V.<sup>11</sup> stellte sich deutlich heraus, dass verschiedene Entwicklungen der letzten Jahre, insbesondere auf der europäischen Ebene, nicht oder nur marginal bekannt waren. Dies betrifft nicht nur den aktuellen Stand z. B. frei verfügbarer Crawler-Technologien, sondern etwa auch die Bereiche Forschungsinfrastrukturen und Metadatenschemata zur Beschreibung digitaler Objekte.

Wie bereits im Abschnitt „Gemeinsame Bedarfe – gemeinsame Lösungen“ erläutert wurde, ruft dieser Artikel dazu auf, eine internationale Strategie für digitale Langzeitarchivierung zu entwickeln, diese als europaweite Infrastruktur zu implementieren und interessierten Einrichtungen und Unternehmen als Service

11 Vgl. [www.collaboratory.de/w/Initiative\\_Nachhaltigkeit\\_in\\_der\\_Digitalen\\_Welt](http://www.collaboratory.de/w/Initiative_Nachhaltigkeit_in_der_Digitalen_Welt) (Letzter Aufruf: 15.08.2013).

anzubieten. Umgesetzt werden kann dieser Vorschlag nur in enger Kooperation mit bestehenden Initiativen, die sich mit digitalen Forschungs-, Metadaten- und Kommunikationsinfrastrukturen beschäftigen. Hierzu zählen insbesondere langfristig angelegte Vorhaben wie Europeana, CLARIN und META-SHARE des europäischen Exzellenznetzwerks META-NET.<sup>12</sup>

Diesen Infrastrukturen, Angeboten und Systemen ist gemein, dass sie in ihren teils sehr umfangreichen Katalogen diverse Metadaten über digitale Objekte unterschiedlicher Art beinhalten, die online nach verschiedenen Kriterien durchsucht und recherchiert werden können.<sup>13</sup> Ebendies ist eine der zentralen Kernfunktionen einer Plattform für digitale Nachhaltigkeit<sup>14</sup> und auch für die digitale Langzeitarchivierung: die Erfassung, Repräsentation und Speicherung von Metadaten über digitale Objekte und die Bereitstellung möglichst komfortabler Such- und Rechercheschnittstellen. Obwohl sich die oben genannten Plattformen in vielerlei Hinsicht unterscheiden, ist dies ihr gemeinsamer Kern.<sup>15</sup>

12 Vgl. [www.clarin.eu](http://www.clarin.eu); [www.europeana.eu](http://www.europeana.eu); [www.meta-share.eu](http://www.meta-share.eu); [www.meta-net.eu](http://www.meta-net.eu) (Letzter Aufruf: 15.08.2013).

13 Rehm, Georg/Schonefeld, Oliver/Witt, Andreas/Chiaros, Christian/Lehberg, Timm: SPLICR. A Sustainability Platform for Linguistic Corpora and Resources. In: von Storrer, Angelika/Geyken, Alexander/Siebert, Alexander/Würzner, Kay-Michael (Hrsg.): KONVENS 2008 (Konferenz zur Verarbeitung natürlicher Sprache). Textressourcen und lexikalisches Wissen. Berlin 2008, S. 86-95; Rehm, Georg/Schonefeld, Oliver/Witt, Andreas/Lehberg, Timm/Chiaros, Christian/Bechara, Hanan/Eishold, Florian/Evang, Kilian/Leshtanska, Magdalena/Savkov, Aleksandar/Stark, Matthias: The Metadata-Database of a Next Generation Sustainability Web-Platform for Language Resources. In: Proceedings of LREC 2008, May 28–30, Marrakech, Morocco, 2008. 2008, S. 371-378.

14 Rehm, Georg/Schonefeld, Oliver/Trippel, Thorsten/Witt, Andreas: Sustainability of Linguistic Resources Revisited. In: Balisage 2010. International Symposium on XML for the Long Haul: Issues in the Long-term Preservation of XML, August 2, Montreal, Canada, 2010. 2010; Rehm, Georg/Schonefeld, Oliver/Witt, Andreas/Hinrichs, Erhard/Reis, Marga: Sustainability of Annotated Resources in Linguistics. A Web-Platform for Exploring, Querying and Distributing Linguistic Corpora and Other Resources. In: Literary and Linguistic Computing 24/2 (2009), S. 193-210.

15 Dieser gemeinsame Kern wiederum wurde von verschiedenen Initiativen zum Anlass genommen, entsprechende Open-Source-Technologien wie z. B. Fedora Commons zu entwickeln, die das Management, die Nachhaltigkeit und die Verlinkung beliebiger digitaler Inhalte erlauben. Vgl. [www.fedora-commons.org](http://www.fedora-commons.org) (Letzter Aufruf: 15.08.2013).



In allen genannten und verschiedenen weiteren Initiativen, die sich ebenfalls mit der Speicherung und Verfügbarmachung digitaler Objekte beschäftigen, stehen immer wieder die gleichen Fragestellungen im Mittelpunkt.<sup>16</sup> Hierzu zählen unter anderem rechtliche Aspekte (Urheberrechtsfragen in Bezug auf die ursprünglichen Objekte, in Bezug auf ihre digitalen Entsprechungen und auch in Bezug auf die zugehörigen Metadaten)<sup>17</sup>, bestmögliche Suche in den Katalogen (Information Retrieval, intelligentes Querying, Vorschläge, Nutzermodelle etc.), Erarbeitung und Spezifikation von Metadatenschemata zur Beschreibung der digitalen Objekte der jeweiligen Domäne, Speicherung der Metadaten sowie der digitalen Objekte (zentrale vs. dezentrale Repositorien, Hochverfügbarkeit etc.) und Import bestehender Datenbestände. Die Bearbeitung derartiger Fragen findet typischerweise in fachwissenschaftlichen EU-geförderten Projektkonsortien statt. Die Community um digitale Langzeitarchivierung sollte mit den oben genannten und verschiedenen weiteren Initiativen in Dialog treten,

- 16 Schmidt, Thomas/Chiarcos, Christian/Lehmburg, Timm/Rehm, Georg/Witt, Andreas/Hinrichs, Erhard: Avoiding Data Graveyards. From Heterogeneous Data Collected in Multiple Research Projects to Sustainable Linguistic Resources. In: Proceedings of the E-MELD 2006 Workshop on Digital Language Documentation. Tools and Standards – The State of the Art, East Lansing, Michigan. 2006; Rehm, Georg/Witt, Andreas/Hinrichs, Erhard/Reis, Marga: Sustainability of Annotated Resources in Linguistics. In: Proceedings of Digital Humanities 2008, June 25–29, Oulu, Finland, 2008. 2008; Gavrilidou, Maria/Labropoulou, Penny/Desipri, Elina/Piperidis, Stelios/Papegeorgiou, Haris/Monachini, Monica/Frontini, Francesca: The META-SHARE Metadata Schema for the Description of Language Resource. In: Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012). European Language Resources Association (ELRA), Istanbul, Turkey. 2012, S. 1090-1097.
- 17 Lehmburg, Timm/Chiarcos, Christian/Rehm, Georg/Witt, Andreas: Rechtsfragen bei der Nutzung und Weitergabe linguistischer Daten. In: Rehm, Georg/Witt, Andreas/Lemnitzer, Lothar (Hrsg.): Datenstrukturen für linguistische Ressourcen und ihre Anwendungen – Data Structures for Linguistic Resources and Applications. Proceedings of the Biennial GLDV Conference 2007. Tübingen 2007, S. 93-102; Weitzmann, John Hendrik/Rehm, Georg/Uszkoreit, Hans: Licensing and Sharing Language Resources: An Approach Inspired by Creative Commons and Open Science Data Movements. In: LREC 2010 Workshop Legal Issues for Sharing Language Resources: Constraints and Best Practices, May 17, Malta, 2010. 2010; der Beitrag von Eric W. Steinhauer, Wissen ohne Zukunft? Der Rechtsrahmen der digitalen Langzeitarchivierung von Netzpublikationen, in diesem Band.

um gemeinsame Themen und mögliche Synergien zu identifizieren und im Anschluss strategische Allianzen zu schmieden.

Die oben genannten Initiativen sind Beispiele einiger Fachwissenschaften, die innerhalb geförderter Projekte Plattformen und Infrastrukturen entwickeln, um ihre Datensammlungen (Forschungsdaten, Primärdaten, Analysedaten, Mess- und Sensordaten etc.) sowie zugehörige Werkzeuge und Technologien in zentralen One-Stop-Shop-Repositoryn zu sammeln und der jeweiligen Forschungs-Community anzubieten. CLARIN macht dies für die Computerlinguistik und verfolgt einen komplexen E-Science-Ansatz, wohingegen META-NET<sup>18</sup> und META-SHARE<sup>19</sup> die angewandt ausgerichtete und mit der Industrie eng kooperierende Sprachtechnologie adressieren und einen eher einfachen Peer-to-peer-Ansatz präferieren, der verteilte Knoten vorsieht, die jeweils organisations-, einrichtungs- oder auch landesspezifische Metadatenkataloge beinhalten, die in regelmäßigen Abständen in einem gemeinsamen Katalog aggregiert werden. Vergleichbare Initiativen existieren in diversen weiteren Fachwissenschaften. Europeana wiederum ist eine deutlich umfangreicher finanzierte Initiative, die die Metadatenkataloge zahlreicher Museen und Gedächtnisinstitutionen zusammenführt, dabei aber mit dem Problem hochgradig heterogener Metadaten konfrontiert wird. Eine weitere relevante, interdisziplinäre Community ist aus dem Einsatz von Auszeichnungssprachen wie SGML und XML für Aufgaben in den Geisteswissenschaften hervorgegangen; bei den jährlich stattfindenden internationalen Konferenzen der Reihe Digital Humanities werden mittlerweile auch neue Ansätze und Lösungen präsentiert, die für die digitale Nachhaltigkeit relevant sind.

Es zeigt sich also, dass derzeit die zentralen Impulse im mittelbaren Bereich der digitalen Langzeitarchivierung, die auch genügend Sichtbarkeit und Unterstützung und somit Förderung erhalten, primär aus den verschiedenen Fachwissenschaften mit ihren jeweils spezialisierten Initiativen wie z. B. CLARIN und META-SHARE stammen. Diese Wissenschaften erzeugen und

18 Rehm, Georg/Uszkoreit, Hans: Multilingual Europe. A challenge for lang tech. In: MultiLingual April/May (2011), S. 51-52; dies. (Hrsg.): The META-NET Strategic Research Agenda for Multilingual Europe 2020. Heidelberg, New York, Dordrecht, London 2013.

19 Piperidis, Stelios: The META-SHARE Language Resources Sharing Infrastructure. Principles, Challenges, Solutions. In: Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012). European Language Resources Association (ELRA), Istanbul, Turkey. 2012, S. 36-42.

analysieren digitale Objekte jeglicher Art, Komplexität und Größe. Die zentrale Sammlung und Bereitstellung der Daten hat unter anderem genuin wissenschaftliche Gründe, sodass interessierte Kollegen beispielsweise Experimente nachvollziehen können. Es geht dabei aber auch um die Beschleunigung des Forschungstransfers durch die Bereitstellung etwa von Sprachressourcen oder neuen Technologien, sodass interessierte Unternehmen von den neuesten Forschungsergebnissen profitieren können. Interessanterweise wird der Aspekt der Langzeitarchivierung in derartigen Initiativen zwar auch immer wieder diskutiert, er steht aber nicht im Mittelpunkt der Diskussionen, sodass im Falle einer langfristigen Kooperation zwischen den Experten für Langzeitarchivierung und den oben genannten fachwissenschaftlichen Initiativen beiderseitig mit gewinnbringenden Synergien gerechnet werden kann.

#### *Schlussfolgerungen und Empfehlungen*

Das Thema digitale Langzeitarchivierung wird mit sehr hohen Kosten, in verschiedenen Bereichen mit einer unklaren Rechtslage und riesigen Datenmengen konfrontiert. Außenstehenden erscheint das Anliegen von Fürsprechern des Themas oftmals zwar wichtig, es wird aber nicht als dringend wahrgenommen, was dazu führt, dass die digitale Nachhaltigkeit gerne aufgeschoben bzw. ihre Priorität herabgestuft wird. Eine Lösung für dieses Dilemma kann und muss es sein, die digitale Langzeitarchivierung bzw. zentrale Aspekte der digitalen Langzeitarchivierung unter einem anderen Oberthema zu bearbeiten. Ein mögliches Oberthema könnte die Entwicklung einer europäischen Suchmaschine bzw. einer digitalen Infrastruktur mit zahlreichen integrierten Services für Europa sein. Ein alternatives Thema könnte eine möglichst generische digitale Infrastruktur für E-Science und digitale Forschungsdaten sein. Solche EU-Infrastrukturen könnten zahlreiche positive Strahleffekte ausüben und Innovation sowie Forschungstransfer fördern. Unter diesem Mantel könnte auch das Thema der digitalen Langzeitarchivierung bearbeitet werden.

Dieser Beitrag strebt den Versuch eines Brückenschlags an, zwischen Experten für digitale Langzeitarchivierung und verschiedenen fachwissenschaftlich ausgerichteten Initiativen, die sich eher mit dem Aufbau domänenspezifischer Forschungsinfrastrukturen beschäftigen. Eine weitere relevante, bereits erfolgreich auf europäischer Ebene tätige Initiative in diesem Zusammenhang ist Europeana. Nicht nur Deutschland, sondern Europa benötigt eine gemeinsame

Position und Strategie für digitale Langzeitarchivierung, die als kontinentweite Infrastruktur implementiert und interessierten Einrichtungen und ggf. auch Unternehmen angeboten werden kann. Der Artikel appelliert daran, eine enge Kooperation mit den bestehenden Initiativen aufzubauen, die sich mit digitalen Forschungs-, Metadaten- und Kommunikationsinfrastrukturen beschäftigen.

Deutschland benötigt dringend eine Strategie zur digitalen Langzeitarchivierung. Sie sollte kleine, realistische, finanzierbare, umsetzbare Schritte enthalten. Die Planung der Strategie sollte aufgrund der zu erwartenden massiven Kosten für die Infrastruktur und Technologieentwicklung nicht isoliert erfolgen, sondern übergreifend von der Europäischen Kommission koordiniert werden, um z. B. eine gemeinsame technologische Infrastruktur zu entwickeln und bereitzustellen (etwa nach dem Vorbild von Europeana). Die digitale Langzeitarchivierung sollte ein integraler Bestandteil der nächsten Version der digitalen Agenda der Europäischen Kommission werden. In „A Digital Agenda for Europe“<sup>20</sup> tauchen die Begriffe *archive* oder auch *sustainability* in der hier gemeinten Lesart nicht ein einziges Mal auf. Die Entwicklung und Umsetzung einer nationalen bzw. internationalen Strategie zur digitalen Langzeitarchivierung birgt genügend Potenzial, um zahlreiche Strahleffekte in verschiedene Bereiche der Informationstechnologie zu generieren. Dabei ist es essentiell, eine Infrastruktur zu entwickeln oder die Entwicklung einer Infrastruktur fortzuführen und dabei vorhandene Technologielücken zu schließen, um auf diese Weise sowohl den europäischen Bürgerinnen und Bürgern als auch Kommunikation, Wirtschaft und Kultur wichtige und relevante neue Dienste anzubieten, die einen genuine Mehrwert schaffen. Die Anforderungen, Methoden und Prinzipien, die für die Langzeitarchivierung wissenschaftlicher Forschungsdaten gelten, können letztlich auch für alle anderen Daten angewendet werden, seien es private Daten, Webseiten oder auch öffentliche Verwaltungsdaten, weshalb wiederum ein gemeinsames technologisches Fundament eingesetzt werden kann, das auf konkreten Ebenen spezifische Services anbietet. Die Entwicklung von META-SHARE (2010–2012)<sup>21</sup>

20 European Commission: A Digital Agenda for Europe. 2010. [http://ec.europa.eu/information\\_society/digital-agenda/publications/](http://ec.europa.eu/information_society/digital-agenda/publications/) (Letzter Aufruf: 15.08.2013).

21 Vgl. Federmann, Christian/Giannopoulou, Ioanna/Girardi, Christian/Hamon, Olivier/Mavroidis, Dimitris/Minutoli, Salvatore/Schröder, Marc: META-SHARE V2. An Open Network of Repositories for Language Resources Including Data and Tools. In: Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012). European Language Resources Association (ELRA), Istanbul, Turkey. 2012, S. 3300-3303.

hätte beispielsweise durch eine verfügbare Basisinfrastruktur drastisch vereinfacht werden können, die Evaluation verschiedener verfügbarer Open-Source-Frameworks hätte entfallen können, ebenso große Teile der Entwicklung der Nutzerschnittstelle.

Kurzfristig ist der Einsatz bereits verfügbarer Infrastrukturen zu testen und für den Zweck der digitalen Langzeitarchivierung zu evaluieren. Falls dabei technologische oder konzeptuelle Lücken identifiziert werden, können diese gemeinsam mit den Fachwissenschaftlern geschlossen werden. Von Relevanz sind hierbei auch die jeweiligen Funktionsmerkmale der domänen- und fachwissenschaftsspezifischen Infrastrukturen: Wo liegen Gemeinsamkeiten vor, wo Unterschiede? Basierend auf derartigen Evaluationen, kann die Funktionalität einer abstrakten und generischen Basisinfrastruktur konzipiert werden.

Essentiell ist der Aufbau einer belastbaren europaweiten Community und Technologieallianz rund um das Thema digitale Langzeitarchivierung, die in der Lage ist, mit einer Stimme zu sprechen. Hierbei können insbesondere Verbindungen zu Unternehmen, denen entsprechende digitale Services kostenpflichtig angeboten werden, mittelfristig für eine substantielle Gegenfinanzierung und auch für Visibilität auf der Ebene nationaler und internationaler Entscheidungsträger sorgen. Ebenfalls sind gute Verbindungen zu den diversen Initiativen rund um Forschungsinfrastrukturen zu etablieren. Operationalisiert werden können solche Verbindungen über gemeinsame Forschungsthemen wie z. B. Metadatenschemata, Semantic Web<sup>22</sup>, Linked Open Data bis hin zu Technologiekomponenten, die von beiden Seiten benötigt werden. Eines dieser gemeinsamen Themen könnte eine vollautomatische und möglichst generische Beantwortung der kritischen Frage sein, welche digitalen Objekte archiviert und welche ignoriert werden sollen. Es erscheint hilfreich, diese Frage automatisch von einem transparenten und objektiven Algorithmus beantworten zu lassen, der jedoch auch Ausnahmen zulässt. Als Indizien hierfür könnten entsprechende positive Rückmeldungen (likes, favourites, bookmarks etc.) der Nutzer in den verschiedenen sozialen Medien und die digitale Sichtbarkeit der Objekte sowie ihrer Urheber in den großen Suchmaschinen sowie auch in traditionellen Archiven und Bibliotheken dienen. Zu den großen technischen Herausforderungen zählt dabei später, in der Phase des Betrachtens, allerdings die adäquate Rekonstruk-

22 Ausführlich hierzu der Beitrag von Felix Sasaki, Nachhaltigkeit und das Semantic Web, in diesem Band.

tion des kommunikativen und situativen Kontextes (vgl. Abschnitt „Technische Herausforderungen“).

Mittel- bis langfristig sollte gemeinsam mit den verschiedenen Initiativen der Fachwissenschaften und insbesondere auch mit Europeana eine generische, europaweite Cloud-Infrastruktur entwickelt und implementiert werden, die Archivierung und Retrieval beliebiger digitale Objekte ermöglicht, d. h. ebenfalls die digitale Langzeitarchivierung mit Services unterstützt, die sowohl von Gedächtnisinstitutionen als auch von anderen interessierten Organisationen und Unternehmen genutzt werden können. Eine derartige Infrastruktur sollte als inhärente und dringend benötigte Komponente für Nachhaltigkeit der digitalen Infrastruktur des digitalen Marktes konzeptualisiert und anschließend extern kommuniziert werden. Dieses Thema ist in jeder Hinsicht kompatibel mit den beiden neuen großen Förderprogrammen Connecting Europe Facility<sup>23</sup> und Horizon 2020<sup>24</sup>. Hierzu ist auf der europäischen Ebene hinreichend Sichtbarkeit für dieses Thema zu generieren, denn nur auf der Ebene der EU können mittel- bis langfristig ausreichende Mittel zur Verfügung gestellt werden, um zu einer tatsächlich benutzbaren Implementierung zu gelangen, die europaweit Verwendung finden kann. Aufgrund der Höhe der zu investierenden Kosten erscheint es nicht realistisch, dass eine rein nationale Anstrengung dies gewährleisten könnte, sodass eine europaweite digitale Infrastruktur anzustreben ist. Diese könnte bestehen aus mehreren verteilten Datenzentren, bei denen Langzeitarchivierung, Hochverfügbarkeit und Hochgeschwindigkeit zum Designprinzip erklärt werden. Im Zeitalter von Prism und Tempora kann bereits jetzt prognostiziert werden, dass die EU oder möglicherweise auch neutrale Länder wie z. B. die Schweiz Gegenmaßnahmen ergreifen werden, und zwar durch die Förderung und Realisierung einer Gruppe vertrauenswürdiger digitaler Dienste, die den Bürgerinnen und Bürgern zur Verfügung gestellt werden. Abhörsichere Kommunikationsverfahren mit vollständiger Endpunkt-zu-Endpunkt-Verschlüsselung und Kontrollmechanismen wie z. B. Trust-Centern sind notwendig, um für maximale Transparenz und Sicherheit zu sorgen, was

23 European Commission: Connecting Europe Facility. Commission adopts plan for 50 billion Euros boost to European networks. 2011. <http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/1200> (Letzter Aufruf: 15.08.2013).

24 European Commission: Horizon 2020. The Framework Programme for Research and Innovation. 2012. <http://ec.europa.eu/research/horizon2020/> (Letzter Aufruf: 15.08.2013).

ebenfalls für das Anbieten von Diensten für die digitale Langzeitarchivierung an Unternehmen ein missionskritischer Faktor sein wird. Optimalerweise ist es nur eine Frage der Zeit, bis die Europäische Union erkennt, dass derartige Dienste einen inhärenten Teil der europäischen Infrastruktur darstellen und ebenso wie Brücken, Straßen und Glasfasernetze gefördert werden müssen. Mit der Digital Component des 11 Milliarden Euro umfassenden Programms Connecting Europe Facility werden derzeit die ersten Schritte geplant. Falls also in einigen Jahren EU-geförderte Cloud-Speicher und sonstige Cloud-Dienste entstehen werden, sollten die Initiativen für digitale Langzeitspeicherung hinreichend gut aufgestellt und in der Lage sein, davon zu profitieren, indem alle generischen, domänen- und fachspezifischen Infrastrukturen daran angekoppelt werden.

Mittel- bis langfristig kann die Community um die digitale Langzeitarchivierung auch entscheidende Impulse liefern für die seit Jahren schwelende, aber auch stagnierende Debatte um eine nationale und internationale Reform des Urheberrechts, da sie verschiedene Aspekte dieser Debatte berührt und streng genommen erst durch eine vollständige Überarbeitung eine solide Rechtsgrundlage erhalten kann.<sup>25</sup> Geschieht diese Überarbeitung nicht bzw. zu spät, wird Deutschland digital vom Rest Europas bzw. vom Rest der Welt abgehängt. Die notwendige Überarbeitung des Urheberrechts könnte auch dafür sorgen, dass der Umgang mit digitalen Massendaten in der Forschung (etwa im Bereich Information Retrieval oder Sprachtechnologie) vereinfacht bzw. überhaupt erst ermöglicht wird. Durch entsprechende Forschungsprogramme können neue Innovationen, Verfahren, Algorithmen, Produkte, Dienstleistungen und Spin-offs entstehen.

25 Steinhauer, Eric W. (2013): "Wissen ohne Zukunft? Der Rechtsrahmen der digitalen Langzeitarchivierung von Netzpublikationen". In: Was bleibt? Nachhaltigkeit der Kultur in der digitalen Welt, herausgegeben von Klimpel, Paul und Keiper, Jürgen, iRights Media. In diesem Band, 61ff.

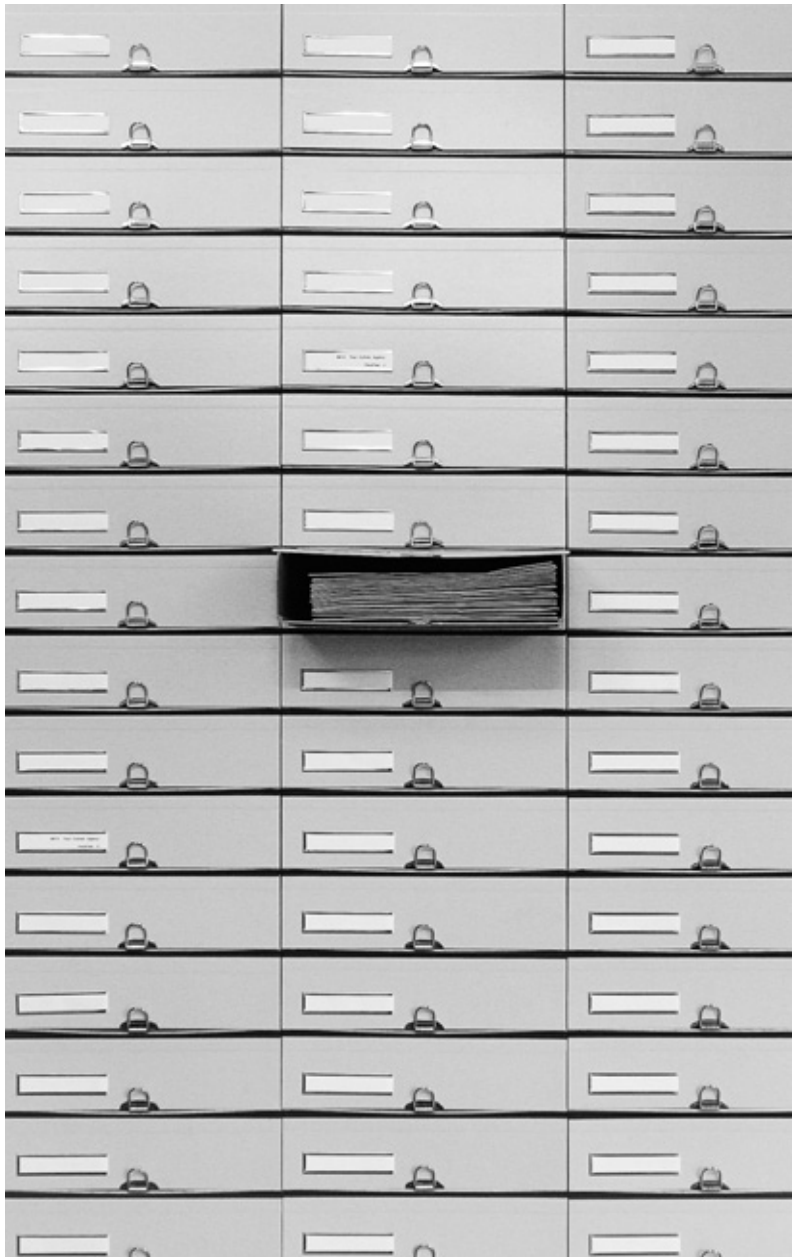






*„Eine fehlende systematische digitale Langzeitarchivierung für digitale Forschungsdaten und digitales Kulturerbe wäre für die Wissenschaft und die Kulturgeschichte gleichbedeutend mit einem kontinuierlichen Verschwinden von Beständen in Bibliotheken und Museen – Regalmeter um Regalmeter, Vitrine um Vitrine.“*

– RALF SCHNEIDER



# Texte, Quellen



## Textausschnitt aus dem Abschlussbericht der Internet Enquete Kommission des Deutschen Bundestags

Sechster Zwischenbericht der Enquete-Kommission zu Bildung und Forschung (BT-Drs. 17/12029), S. 44f.:

„Neben neu entstehenden Datensammlungen ist es eine große Herausforderung, analoge Archive zu digitalisieren, das heißt so – ebenfalls dank moderner Technik – neu zu erschließen und für die Fach-Community beziehungsweise die Öffentlichkeit leichter zugänglich sowie nachnutzbar zu machen. Projekte wie die Deutsche Digitale Bibliothek (DDB) und – auf europäischer Ebene – die Europeana sollen entsprechende Digitalisate zentral zur Verfügung stellen. Neben Fragen der Finanzierung ist hier oft die Rechtklärung bezüglich der Archivmaterialien von großer Bedeutung. Hierzu sei beispielhaft auf den Punkt ‚Umgang mit verwaisten Werken‘ im Bericht der Enquete-Projektgruppe Urheberrecht verwiesen. Gleichzeitig erfordern Langzeitarchivierungen besondere rechtliche Rahmenbedingungen. Die Rechtklärung und Finanzierung spielen gerade auch bei der Kooperation zwischen öffentlicher Hand und Privatwirtschaft, wie beispielsweise beim Digitalisieren der Bestände der Bayerischen Staatsbibliothek durch Google, eine wichtige Rolle. Dabei ist insbesondere zu fragen, welche Rechte an den Digitalisaten an welchen Kooperationspartner fallen und wie eine freie Nutzung öffentlicher Archive mit dem Geschäftsinteresse derer in Einklang gebracht werden kann, die diese Materialien auf ihre Kosten digitalisieren.“

Anmerkung E. Steinhauer: Wichtig scheint mir, dass wir beim Thema der digitalen Langzeitarchivierung unterscheiden zwischen der Befugnis, Kulturgüter zu digitalisieren, was in der Politik meist im Vordergrund steht, und der Sammlung und Erhaltung originär digitaler Objekte. Der zuletzt genannte Bereich wird meist leider übersehen.



## Vancouver-Deklaration der UNESCO

Auch in der UNESCO, der Organisation der Vereinten Nationen für Erziehung, Wissenschaft und Kultur, wird die Zukunft des kulturellen Erbes in der digitalen Welt diskutiert. Davon gibt die Vancouver-Deklaration Zeugnis, die erstmals in einer Übersetzung von Ina Goertz auf Deutsch veröffentlicht wird.

### *DIE VANCOUVER-ERKLÄRUNG*

*Das Gedächtnis der Menschheit im digitalen Zeitalter:  
Digitalisierung und Erhalt*

Die digitale Technologie bietet nie dagewesene Möglichkeiten der Weitergabe und Speicherung von Daten. Dokumente und Daten in digitaler Form sind für Wissenschaft, Bildung, Kultur sowie die wirtschaftliche und soziale Entwicklung von großer Bedeutung, aber die Sicherstellung ihrer langfristigen Haltbarkeit und Konservierung ist ein noch längst nicht gelöstes Problem. Die einzelnen Länder unterscheiden sich zwar erheblich in ihren Möglichkeiten zur Umsetzung von Maßnahmen, mit denen ein dauerhafter Zugang zu digitalen Ressourcen garantiert wird, aber die grundlegenden Herausforderungen sind universeller Art. Eine engere Zusammenarbeit im Umgang mit diesen Ressourcen wird für alle von Vorteil sein.

Gegenwärtig gehen digitale Informationen verloren, weil ihr Wert unterschätzt wird, weil es keinen rechtlichen und institutionellen Rahmen gibt oder weil das Wissen, die Fähigkeiten und die finanziellen Mittel für ihre Bewahrung fehlen. Um diesen Fragen auf den Grund zu gehen und Lösungen zu erarbeiten, berief der Generaldirektor der UNESCO eine internationale Konferenz ein,



die vom 26. bis 28. September 2012 in Vancouver, British Columbia, Kanada, stattfand: Das Gedächtnis der Menschheit im digitalen Zeitalter: Digitalisierung und Erhalt.

Über 500 Teilnehmer aus 110 Ländern diskutierten zwei zentrale Themenkomplexe, die für Dokumente, Unterlagen und Daten in einer digitalen Umgebung relevant sind:

- die Digitalisierung von analogem Material und
- die Haltbarkeit, Archivierung und Zugänglichkeit von authentischen, zuverlässigen und unverfälschten digitalen Materialien.

Als Diskussionsergebnis einigten sich die Konferenzteilnehmer auf folgende Punkte:

1. Wie in Artikel 19 der Allgemeinen Erklärung der Menschenrechte verankert, haben alle Menschen das Recht, „über Medien jeder Art und ohne Rücksicht auf Grenzen Informationen und Gedankengut zu suchen, zu empfangen und zu verbreiten“. Die Bürger üben dieses Recht aus, wenn sie auf Informationen in digitaler Form zugreifen. Glaubwürdigkeit und Unversehrtheit des dokumentarischen Erbes und der Dokumentationssysteme sind daher Voraussetzungen für die kontinuierliche Ausübung dieses Rechts.
2. Durch ihre Digitalisierung können wertvolle analoge Dokumente davor geschützt werden, dass sich ihr Zustand durch häufige Benutzung verschlechtert. Bei audiovisuellen Dokumenten ist die Digitalisierung das einzige Mittel, ihre Haltbarkeit auf Dauer zu sichern.
3. Viele Objekte werden originär in digitaler Form erzeugt, allerdings ohne die angemessene Berücksichtigung von Maßnahmen zur Sicherstellung ihrer dauerhaften Zugänglichkeit, ihrer langfristigen authentischen, zuverlässigen und unverfälschten Bewahrung, die mit dem technologischen Wandel Schritt halten kann. Die Probleme des Zugangs und der Erhaltung betreffen daher auch genuin digitale Medien.
4. Ein besseres Verständnis für die digitale Umgebung ist wesentlich für die Einführung digitaler Bewahungsverfahren, die grundlegenden, in institutionellen regulatorischen Rahmen verankerten Rechtsgrundsätzen folgen, eine Ausgewogenheit zwischen Zugänglichkeit und Datenschutz, zwischen Recht auf Wissen und wirtschaftlichen Rechten herstellen sowie das Eigentum an

und die Kontrolle über indigenes Kulturgut und überliefertes Wissen in digitalem Format respektieren.

5. Der digitalen Bewahrung sollte entwicklungspolitische Priorität eingeräumt werden; Investitionen in die Infrastruktur sind von zentraler Bedeutung, um die Zuverlässigkeit archivierter digitaler Dokumente sowie ihre langfristige Zugänglichkeit und Benutzbarkeit zu gewährleisten.
6. Es müssen Bildungs- und Schulungsprogramme für Informatiker entwickelt und zur Verfügung gestellt werden, um diese darauf vorzubereiten bzw. wieder in die Lage zu versetzen, sowohl Digitalisierungs- als auch Bewahrungspraktiken einzuführen, die den Anforderungen von Regierungen und Bürgern entsprechen.
7. Es besteht dringender Bedarf an einem Strategieplan für die Erarbeitung von Lösungen, Vereinbarungen und Maßnahmen, die einen langfristigen Zugang und zuverlässigen Erhalt von digitalen Medien garantieren. Dieser Strategieplan sollte Themen wie transparente Verwaltung (Open Government), offen zugängliche Daten (Open Data), freier Zugang (Open Access) und elektronische Regierung (E-Government) beinhalten. Er sollte nationale und internationale Prioritäten miteinander verzahnen und voll und ganz im Einklang mit den Menschenrechten stehen.

### *Empfehlungen*

Unter Berücksichtigung der aktuellen und sich herausbildenden Herausforderungen fordern die Teilnehmer

*das UNESCO-Sekretariat dazu auf,*

1. sich aktiv dafür einzusetzen, Rahmenbedingungen und Verfahren der digitalen Bewahrung Realität werden zu lassen, indem es die Verwaltung und den Erhalt digitaler Objekte in allen geeigneten Formen fördert und zu diesem Zweck mit anderen Organisationen, Fonds und Programmen der Vereinten Nationen zusammenarbeitet,
2. die Arbeit der internationalen Verbände im Bereich der Archive, Bibliotheken und Museen zu unterstützen, um internationale rechtliche Rahmenbedingungen für Ausnahmen und Einschränkungen des Urheberrechts zu sichern, um den Erhalt vom und Zugang zum kulturellen Erbe in digitalem

Format zu garantieren sowie den Erwerb von und Zugang zu diesem Erbe auf eine kulturell geeignete Art und Weise zu gewähren,

3. mit internationalen Fachverbänden und anderen internationalen Organen zusammenzuarbeiten, um wissenschaftliche Rahmenpläne für die Digitalisierung und den digitalen Erhalt zu entwickeln sowie Schulungsprogramme und globale Bildungsansätze zu erarbeiten, mit denen die Fähigkeiten der Mitarbeiter in Archiven, Bibliotheken und Museen erhöht werden, mit digitalen Daten umzugehen und sie zu erhalten,
4. ein Forum für alle Interessen- und Anspruchsgruppen einzurichten, um eine Diskussion anzustoßen über eine Standardisierung der Verfahren zur Digitalisierung und der digitalen Bewahrung, einschließlich der Einrichtung von digitalen Registern,
5. in Zusammenarbeit mit internationalen Fachverbänden und Forschungsprojekten Richtlinien, Maßnahmen, Verfahrensweisen und Best-Practice-Modelle für die Digitalisierung und den digitalen Erhalt zu entwerfen und zu veröffentlichen,
6. ein Notfallprogramm zu schaffen, das auf den Erhalt von dokumentarischen Materialien ausgerichtet ist, die durch Naturkatastrophen oder bewaffnete Auseinandersetzungen in Gefahr sind, und ein Programm für die Wiederherstellung von analogen und digitalen Medien zu entwickeln, die aufgrund von überalterter Hardware und Software bereits unzugänglich sind oder es demnächst zu werden drohen,
7. die Mitarbeit von denjenigen Fachleuten im Bereich des Kulturerbes anzuregen, die über Kenntnisse von Konzepten, Methoden und Instrumenten digitaler Forensik verfügen, um die Erfassung und verlässliche Archivierung von authentischen, kontextbezogenen und bedeutsamen Daten zu sichern sowie für eine geeignete Vermittlung des Zugangs zu den Daten zu sorgen,
8. die Richtlinien zur Umsetzung der UNESCO-Charta von 2003 zur Bewahrung des digitalen Kulturerbes zu aktualisieren und in Erwägung zu ziehen, die Bewahrung des und den Zugang zum digitalisierten Kulturerbe in die vorgeschlagene Empfehlung zum dokumentarischen Erbe aufzunehmen, über die auf der 190. Sitzung des UNESCO-Exekutivrats entschieden wird,
9. mit nationalen und internationalen Organisationen im Bereich Forschung und Kulturgut zusammenzuarbeiten, um Beurteilungskriterien zu entwickeln, mit denen Verwahrungsorte dahingehend überprüft werden können,

ob sie digitale Bestände verlässlich bewahren oder für diesen Zweck verbessert werden können,

10. die Zusammenarbeit mit internationalen Normungsgremien zu unterstützen, um die Einheitlichkeit verschiedener Referenzquellen zu digitalem Erhalt zu erhöhen und die Entwicklung von Standards zu fördern, die mit den von der UNESCO gebilligten Grundsätzen im Einklang stehen,
11. sich für die Überzeugung einzusetzen, dass eine gute Verwaltung von zuverlässigen digitalen Daten von zentraler Bedeutung für eine nachhaltige Entwicklung ist, und unter der Schirmherrschaft des Programms Gedächtnis der Menschheit einen globalen digitalen Strategieplan zu entwickeln und einzuführen, um alle betroffenen Stakeholder, insbesondere Regierungen und die Industrie, dazu aufzufordern, in eine verlässliche digitale Infrastruktur und den digitalen Erhalt zu investieren.

*Die Mitgliedstaaten fordern sie dazu auf,*

1. Gesetze zu erarbeiten und in Kraft zu setzen, mit denen die Rechte aller Bürger auf wesentliches Wissen garantiert wird,
2. staatliche Maßnahmen zu entwickeln, die den Erhalt des digitalen Erbes in einer sich rasant wandelnden technologischen Umgebung ermöglichen und unterstützen,
3. die Zusammenarbeit zwischen ihren gesetzgebenden Organen und Archiven, Bibliotheken und Museen sowie anderen einschlägigen Organisationen zu unterstützen, um rechtliche Rahmenbedingungen zu entwickeln, die den Erhalt vom und Zugang zum digitalen Kulturerbe fördern,
4. Strategien für eine transparente Verwaltung und frei zugängliche Daten zu entwickeln, um damit der Anforderung gerecht zu werden, Vertrauen in den Schutz von digitalen Regierungsdokumenten zu schaffen und zu bewahren,
5. rechtliche Garantien zu liefern, dass Informationen, die den Bürgern rechtlich zustehen, in einem frei zugänglichen Format zur Verfügung stehen,
6. Organisationen im privaten Sektor zu ermutigen, in vertrauenswürdige digitale Infrastruktur und den digitalen Erhalt zu investieren,
7. eine Empfehlung für die Förderung einer gesetzlichen Hinterlegungspflicht für digitale Formate zu entwickeln,
8. eine oder mehrere geeignete Aufsichtsbehörden einzurichten, z. B. einen Datenombudsmann, der das notwendige Maß an Unabhängigkeit überwacht

und schützt, das Archive, Bibliotheken, Museen und andere für das Kulturerbe relevanten Organisationen benötigen, um den Zugang zu digitalen Informationen in einer Art und Weise zu gewähren und zu erhalten, die das öffentliche Vertrauen darin stützt, dass Information als erhaltenswert erachtet werden und dass sie verlässlich erhalten werden,

9. digitales dokumentarisches Kulturerbe zu benennen und für die Aufnahme in das Register des Gedächtnisses der Menschheit vorzuschlagen,
10. sicherzustellen, dass analoge Inhalte in digitaler Form zur Verfügung gestellt werden, um zu verhindern, dass sie in Zukunft in einer Welt von überwiegend digitalem Informationsabruf nicht in Vergessenheit geraten,
11. das öffentliche Bewusstsein dafür zu schärfen, wie enorm wichtig die dauerhafte Erhaltung unsere Kulturgüter in digitaler Form ist,
12. die Anwendung von Standards sowie weithin anerkannten Richtlinien und besten Verfahrensweisen der Digitalisierung und der digitalen Bewahrung bei den zuständigen nationalen Organisationen und Gemeinschaften zu unterstützen.

*Die Fachverbände im Bereich des Kulturerbes fordern sie dazu auf,*

1. mit anderen Fachverbänden, internationalen und regionalen Organisationen sowie gewerblichen Unternehmen zusammenzuarbeiten, um sicherzustellen, dass wichtige digital erzeugte Materialien erhalten bleiben, indem sie die Einführung einer gesetzlichen Hinterlegungspflicht befürworten und unterstützen,
2. zur Entwicklung einer zusammenhängenden, konzeptuellen und praktischen Vision für eine digitale Strategie beizutragen, die den Umgang mit und den Erhalt von aufgezeichneten Informationen in all ihren Formen in der digitalen Umgebung beinhaltet,
3. ihre Mitglieder dazu anzuhalten, die Verlässlichkeit, die Authentizität, das Urheberrecht und die zukünftige Verwendung von digitaler Information zu berücksichtigen und Maßnahmen für alle Aspekte der Verwaltung und des Erhalts von digitalen Materialien zu entwickeln,
4. mit dem Privatsektor bei der Entwicklung von Produkten zusammenzuarbeiten, die eine langfristige Aufbewahrung und den Erhalt von im Digitalformat aufgezeichneten Informationen ermöglichen,
5. ihre Mitglieder zu ermutigen, die spezifischen Bedrohungen zu benennen und zu bewerten, denen ihre digitalen Informationen ausgesetzt sind, und

geeignete Verfahren und Maßnahmen zur Entschärfung dieser Gefahren einzuführen.

*Den privaten Sektor fordern sie dazu auf,*

1. mit Archiven, Bibliotheken, Museen und anderen einschlägigen Organisationen zusammenzuarbeiten, um die langfristige Zugänglichkeit von digitalen Informationen sicherzustellen,
2. sich an anerkannte Metadatenstandards zu halten, die in Zusammenarbeit mit Informatikern für die Beschreibung und / oder Verwaltung von digitalen Ressourcen entworfen wurden, um die Interoperabilität von Quellen zu ermöglichen, die als authentisch, garantiert verlässlich und unverfälscht angesehen werden können,
3. bei der Mitwirkung an nationalen und internationalen Initiativen zur Standardisierung und bei grenzüberschreitenden und anderen Partnerschaftsinitiativen, bei denen es darum geht, in digitalen Formaten erzeugte Informationen langfristig zu erhalten, die Probleme der Bewahrung digitaler Medien nicht außer Acht zu lassen.



## Referenzen, Quellen

### *Wichtige Webseiten*

- Berliner Appell: [www.berliner-appell.org](http://www.berliner-appell.org)
- nestor, das deutsche Kompetenznetzwerk zur digitalen Langzeitarchivierung: [www.langzeitarchivierung.de](http://www.langzeitarchivierung.de)
- Schwerpunktinitiative „Digitale Information“: [www.allianzinitiative.de](http://www.allianzinitiative.de)
- Deutsche Digitale Bibliothek: [www.deutsche-digitale-bibliothek.de](http://www.deutsche-digitale-bibliothek.de)
- Europeana: [www.europeana.eu](http://www.europeana.eu)

### *Studien*

- Das Gesamtkonzept für die Informationsinfrastruktur in Deutschland von der Kommission Zukunft der Informationsinfrastruktur im Auftrag der Gemeinsamen Wissenschaftskonferenz des Bundes und der Länder: [www.allianzinitiative.de/fileadmin/user\\_upload/KII\\_Gesamtkonzept.pdf](http://www.allianzinitiative.de/fileadmin/user_upload/KII_Gesamtkonzept.pdf)
- Wissenschaftsrat: Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020, Berlin 2012 (Drs. 2359-12).



- Die Studie des Wissenschaftsrates baut auf die KII-Studie auf und ist im Umfang kürzer: [www.wissenschaftsrat.de/download/archiv/2359-12.pdf](http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf)
- Charles Beagrie Limited: Dauerhaften Zugriff sicherstellen: Auf dem Weg zu einer nationalen Strategie zu Perpetual Access und Hosting elektronischer Ressourcen in Deutschland: [www.allianzinitiative.de/fileadmin/hosting\\_studie\\_d.pdf](http://www.allianzinitiative.de/fileadmin/hosting_studie_d.pdf)

### *Materialien*

Zuvorderst sind natürlich hier alle nestor-Publikationen zu nennen:

- [www.langzeitarchivierung.de/Subsites/nestor/DE/Publikationen/publikationen\\_node.html](http://www.langzeitarchivierung.de/Subsites/nestor/DE/Publikationen/publikationen_node.html)

WICHTIG SIND DARUNTER INSBESONDERE:

- Heike Neuroth, Achim Oßwald, Regine Scheffel, Stefan Strathmann, Karsten Huth (Hrsg.): nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Göttingen 2010. <http://nestor.sub.uni-goettingen.de/handbuch/index.php> sowie
- nestor: Memorandum zur Langzeitverfügbarkeit digitaler Informationen in Deutschland. 2006. <http://files.dnb.de/nestor/memorandum/memo2006.pdf>

### *Weiterführende Links*

- Im Blog [memento-movie.de](http://memento-movie.de) findet sich eine Linkliste, insbesondere für den audiovisuellen Bereich: [www.memento-movie.de/links/](http://www.memento-movie.de/links/)

DOKUMENTATION DER ABSCHLUSSVERANSTALTUNG DER INITIATIVE „NACHHALTIGKEIT IN DER DIGITALEN WELT“

Fotos (Marco Rahn, CC-BY): <http://flic.kr/s/aHsjFKx4yd>

Videos: [www.youtube.com/playlist?list=PL2vfezQbomhoIgO4Wd8KECvpgIgYyNysD](http://www.youtube.com/playlist?list=PL2vfezQbomhoIgO4Wd8KECvpgIgYyNysD)

#### HEISE ONLINE

- [www.heise.de/newsticker/meldung/Kampf-um-die-Deutungshoheit-im-Web-Schema-org-vs-Linked-Data-1874722.html](http://www.heise.de/newsticker/meldung/Kampf-um-die-Deutungshoheit-im-Web-Schema-org-vs-Linked-Data-1874722.html)
- <http://m.heise.de/newsticker/meldung/Experten-gegen-Archivierung-von-Tweets-und-Chats-1874777.html?from-classic>

#### NIGHT OUT @ BERLIN

- <http://nightoutatberlin.jaxblog.de/post/Ums-Erbe-zerstritten-Nachhaltigkeit-in-der-Digitalen-Welt-und-Erbe-Übertragungskonzepte-zwischen-Natur-und-Kultur.aspx>

#### HÖRSAAL | DEUTSCHLANDRADIO WISSEN

- [http://wissen.dradio.de/digitale-nachhaltigkeit-was-es-nicht-lohnt-aufzuheben.88.de.html?dram:article\\_id=250223](http://wissen.dradio.de/digitale-nachhaltigkeit-was-es-nicht-lohnt-aufzuheben.88.de.html?dram:article_id=250223)
- [http://wissen.dradio.de/digitale-nachhaltigkeit-die-ordnungsmacht.88.de.html?dram:article\\_id=250354](http://wissen.dradio.de/digitale-nachhaltigkeit-die-ordnungsmacht.88.de.html?dram:article_id=250354)



Serverraum, Konrad-Zuse-Zentrum für Informationstechnik Berlin

# Biographien



THOMAS BÄHR

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Thomas Bähr studierte Ingenieurwesen und Organisationswissenschaften in Berlin und Hildesheim. Er arbeitete 15 Jahre in der Redaktionsleitung bei Bertelsmann und Springer STM. Sein Schwerpunkt war der Transferprozess von Printprodukten in die digitale Welt, insbesondere die Entwicklung von Konzepten zur medienneutralen Datenhaltung sowie die Durchführung der damit verbundenen Veränderungsprozesse in den betroffenen Organisationsstrukturen. Seit 2009 ist er Leiter der digitalen Langzeitarchivierung an der Technischen Informationsbibliothek. Neben dem Betrieb eines Langzeitarchivierungssystems und der Archivierung von Textdokumenten liegt der Fokus dort in der Entwicklung von Konzepten und praktischen Anwendungen zur Archivierung von audiovisuellen Materialien und 3-D-Objekten. Thomas Bähr engagiert sich u. a. in der Open Planets Foundation und in nestor, dem Kompetenznetzwerk zur digitalen Langzeitarchivierung.

TOBIAS BEINERT

*Autor (mit Armin Straube): Aktuelle Herausforderungen der digitalen Langzeitarchivierung*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Tobias Beinert ist seit 2006 in mehreren Projekten im Münchener Digitalisierungszentrum der Bayerischen Staatsbibliothek (BSB) aktiv. Seit Januar 2008 engagiert er sich im Bereich Langzeitarchivierung als Vertreter im Kompetenznetzwerk nestor (Co-Leiter AG Kooperation, AG Recht). Tobias Beinert wirkte an der DIN 31645 „Leitfaden zur Informationsübernahme in digitale Langzeitarchive“ mit. Wissenschaftliche Mitarbeit an Studien zur Entwicklung von Geschäfts- und Organisationsmodellen für die Archivierung von retrodigitalisierten Objekten sowie zur Vertrauenswürdigkeit und Skalierbarkeit der Infrastruktur für die Langzeitarchivierung an der Bayerischen Staatsbibliothek. Seine derzeitigen Arbeitsschwerpunkte sind unter anderem Webarchivierung und die Einführung eines neuen Langzeitarchivierungssystems an der BSB.



JENS BEST

*Autor: Wikipedia und Nachhaltigkeit – Interview mit Jens Best, Präsidiumsmitglied von Wikimedia Deutschland*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Jens Best ist IT-Berater und arbeitet als Senior Strategy Consultant, Redner und Autor in den Bereichen Digitaler Wandel, Social Media und E-Commerce. Best berät Unternehmen, Institutionen und zivilgesellschaftliche Gruppen beim Einsatz des Internets im wirtschaftlichen und gesellschaftlichen Kontext. Die Veränderungen in der Kommunikation, geschaffen durch

die Möglichkeiten der digitalen Medien, sind für Jens Best Grundlage neuer Kollaborations- und Partizipationsmöglichkeiten in allen gesellschaftlichen Feldern. Er ist u. a. Präsidiumsmitglied von Wikimedia Deutschland und Mitglied des ikosom. Jens Best wurde 1972 in Frankfurt/M. geboren.

Foto: Katja Ullrich

## DR. CHRISTOPH BRUCH

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Dr. Christoph Bruch studierte Politikwissenschaft in an der Goethe-Universität in Frankfurt/M. und an der Freien Universität Berlin. Seine Dissertation schrieb er über die Durchsetzung des Freedom of Information Act der USA. Nach beruflichen Stationen an der Freien Universität Berlin, dem Deutschen Institut für Urbanistik und als freiberuflicher Publizist leitet er seit April 2007 die Open-Access-Stelle der Max-Planck-Gesellschaft. Ehrenamtlich engagiert er sich bei einer Reihe von NGOs: Aktionsbündnis „Urheberrecht für Bildung und Wissenschaft“, EURALO-ICANN At Large Europe, Humanistischen Union, Mehr Demokratie und Stiftung Haus der Demokratie und Menschenrechte. Dabei nimmt sein Engagement für Informationsfreiheitsgesetze und direktdemokratische Beteiligungsformen eine herausragende Position ein.



## JENS CRUEGER

*Autor: Die Dark Ages des Internet?*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Jens Crueger war nach dem Abitur 2003 für eine Legislatur jüngster Abgeordneter der Bremischen Bürgerschaft, Landtag (u. a. Mitglied im Ausschuss für Medien und Kommunikationstechnologie). Danach hatte er diverse weitere politische Funktionen; er war z. B. Fraktionsmitarbeiter und Büroleiter eines Abgeordneten. Seit dem Ausstieg aus der Politik in 2012 ist er als wissenschaftlicher Politikberater, u. a. zu Nachhaltigkeit in der Politik, tätig.

Jens Crueger studierte Geschichte, Philosophie und Soziologie an der Universität Bremen (B.A. in 2011). Er wurde 2012 mit dem Ludwig-Windthorst-Preis der Stadt Meppen (für einen parlamentsgeschichtlichen Essay) und mit dem Publikumspreis „Zukunft des Wissenschaftsjournalismus“ auf dem Tag des Wissenschaftsjournalismus 2012 ausgezeichnet. 2013 wird er sein Masterstudium mit dem Schwerpunkt „Geschichte in der Öffentlichkeit“ an der Universität Bremen abschließen.





DR. ELLEN EULER LL.M.

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Dr. Ellen Euler, LL.M. (geb. 1977), Leiterin der Geschäftsstelle der Deutschen Digitalen Bibliothek, studierte Rechtswissenschaft in Halle/Saale, Leipzig, Münster/Westfalen sowie Siena und Bologna (Italien). Nach dem juristischen Staatsexamen promovierte sie am Lehrstuhl für Informationsrecht bei Prof. Dr. Thomas Dreier an der Albert-Ludwigs-Universität Freiburg. Ihre Dissertation befasst sich mit der Fragestellung, wie der rechtliche Rahmen (insbesondere das Urheberrecht) beschaffen sein muss, damit Gedächtnisinstitutionen ihren gesetzlichen Aufgaben des Bestandsaufbaus, der Bestandserhaltung und Vermittlung auch im digitalen Zeitalter adäquat nachkommen können. Zu spezifischen Rechtsfragen der digitalen Langzeitarchivierung hat Euler auf verschiedenen Fachtagungen referiert, sich als aktives Mitglied bei nestor eingebracht, war als Sachverständige zum Pflichtexemplarrecht in Nordrhein-Westfalen vor dem Landtag und veröffentlichte zahlreiche Fachbeiträge zur Thematik (bspw. Zur Langzeitarchivierung digital aufgezeichneter Werke und ihrer urheberrechtlichen Einordnung und Beurteilung, In: AfP (2008), 474ff.).



FREDERIK FISCHER

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Frederik Fischer ist Gründer und Geschäftsführer von Tame. Das News-Startup entwickelt eine Bewertungsmethode für journalistisch relevante Konversationen auf Twitter. Vor der Gründung hat Fischer als freier Journalist mit den Schwerpunkten Netzpolitik und Digitale Kultur für TV, Radio und Print gearbeitet. In seiner journalistischen Tätigkeit sowie als Geschäftsführer von Tame beschäftigt er sich mit der Vergänglichkeit digitaler Artefakte – von ganzen Web-Communities bis hin zu einzelnen Tweets. Frederik Fischer hat Medienwissenschaft und Volkswirtschaft in Aarhus, Amsterdam, Washington und London studiert. Seine Forschung an der City University London konzentrierte sich auf Technologie-Startups im Journalismus sowie deren Auswirkungen auf den Medienwandel.



JAN FRÖHLICH

*Autor: Farbraum und Bildzustand im Kontext der Langzeitarchivierung*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Jan Fröhlich (geb. 1979 in Freiburg) studierte Audiovisuelle Medien in Stuttgart. Als Diplomarbeit entwickelte er einen Filmscanner für historische Filme mit nicht standardgemäßer Perforation. Von 2004 bis 2012 gestaltete er als Technischer Leiter die Entwicklung der CinePostproduction GmbH von einem traditionellen Filmlabor zu einem führenden digitalen Postproduktions-

dienstleister. In dieser Zeit war er an einer Reihe von technisch zukunftsweisenden Projekten beteiligt. So begleitete er zuletzt den ersten abendfüllenden stereoskopischen Animationsfilm Europas, „Konferenz der Tiere“, ins Kino. Jan Fröhlich hat die Forschungsprojekte „Werkzeuge der Medienproduktion“ (Nondestruktive JPEG2000 basierte Workflows) und „CineSave“ (Langzeitarchivierung von digitalen Daten) mit initiiert. Seit April 2012 promoviert er an der Universität Stuttgart zum Thema Farbkorrektur und Farbworkflows. Er ist Lehrbeauftragter für den Bereich Postproduktion an zahlreichen Filmhochschulen und Mitglied der SMPTE, FKTG sowie dem Bundesverband Kamera (BVK).

#### RALPH GIEBEL

*Autor: Speichertechnologie und Nachhaltigkeit*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Ralph Giebel arbeitet als Head of Business Development seit 1995 für die Firma EMC in Deutschland. Er ist dort verantwortlich für die Entwicklung des Lösungsvertriebs in den Bereichen Öffentliche Auftraggeber und Gesundheitswesen.

Der Diplom-Ingenieur für angewandte Energietechnik sammelte nach seinem Studium an der Hochschule Zittau/Görlitz zunächst Erfahrungen als Systemprogrammierer und später als Abteilungsleiter Technologie. Danach startete er seine Beraterkarriere in der IT-Branche bei EMC. Seit 2001 ist er ohne Unterbrechung im Vertical Öffentliche Auftraggeber unterwegs.

Seit vielen Jahren beschäftigt sich Giebel schwerpunktmäßig mit dem Thema Langzeitarchivierung digitaler Objekte. So führte er die EMC Deutschland in das Kernteam, welches mit der Programmierung und der Organisation der Deutschen Digitalen Bibliothek befasst ist. Er vertritt die EMC Deutschland GmbH in der PPP Gesellschaft „Initiative D21“ und ist dort als Projektleiter für das Projekt „Digital Heritage“ tätig. 2011 war er maßgeblich an der Gründung des Bündnis für Bildung – BfB e.V. beteiligt. Zudem arbeitet er in verschiedenen Arbeitskreisen der BITKOM mit.



#### PROF. DR. STEFAN GRADMANN

*Autor: Semantische Nachhaltigkeit und Kontrolle: Gedanken zu schema.org und Linked Open Data*

Prof. Dr. Stefan Gradmann hat an der KU Leuven (Belgien) einen Lehrstuhl für Literaturwissenschaft inne und ist Direktor der dortigen Universitätsbibliothek. Daneben ist er Präsident der Deutschen Gesellschaft für Informationsmanagement und Informationspraxis. Vor seiner heutigen Position war er Lehrstuhlinhaber am Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu Berlin, stellvertretender Leiter des Rechenzentrums der Universität Hamburg und wissenschaftlicher Bibliothekar in unter-





schiedlichen Leitungspositionen. Er war von Beginn an intensiv beteiligt am Aufbau der Europeana. Seine Schwerpunkte in Forschung und Lehre sind die Digitalen Geisteswissenschaften und dabei insbesondere die Dekonstruktion des Dokument-Begriffs im digital vernetzten Kontext sowie die Frage, wie Bedeutung und Verstehen im Netz gedanklich zu fassen sind. Konkret manifestiert sich dies vor allem in einer intensiven Auseinandersetzung mit den Ansätzen des Semantic Web und der Linked Open Data und in der Frage, wie der Einsatz von Informationstechnologie in den hermeneutisch geprägten Disziplinen jenseits der Emulation traditioneller Arbeitsmethoden gestaltbar ist.



PROF. MONIKA HAGEDORN-SAUPE

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Prof. Monika Hagedorn-Saupe ist seit 1994 Leiterin des Referates „Besucherbezogene Museumsforschung und Kulturstatistik“ und zugleich stellvertretende Leiterin des Instituts für Museumsforschung in Berlin, an dem sie seit 1985 Mitarbeiterin ist. Sie ist Mitglied im Vorstand des Deutschen Museumsbundes (DMB) und Sprecherin der Fachgruppe „Dokumentation im DMB“ und leitet die Arbeitsgruppe „Museum Information Centres“ von ICOM CIDOC (International Council of Museums – Comité internationale pour la documentation). Hagedorn-Saupe ist zudem eine der drei deutschen VertreterInnen in der MSEG (früher NRG – National Representatives Group), der europäischen Expertengruppe zur Digitalisierung im Kulturbereich, und vertritt ICOM Europa im Vorstand der Europeana Foundation. Seit 2006 ist sie Honorarprofessorin an der Hochschule für Technik und Wirtschaft Berlin.



DR. MICHAEL HOLLMANN

*Autor: Was wollen wir archivieren?*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Michael Hollmann (geb. 1961) studierte Geschichte und Germanistik in Mainz und wurde dort mit einer Arbeit zur spätmittelalterlichen Landesgeschichte promoviert. Seit 1989 ist er im Bundesarchiv beschäftigt, zunächst als Referendar, von 1991 an als Leiter verschiedener Referate des Bundesarchivs in Koblenz und Berlin. Von 2006 bis 2011 leitete er die Abteilung „Bundesrepublik Deutschland“, die neben der konventionellen Überlieferung der Ministerien und Behörden des Bundes zuständig ist für den Aufbau eines digitalen Archivs. Seit Mai 2011 ist Michael Hollmann Präsident des Bundesarchivs.

## JÜRGEN KEIPER

*Autor: Überblick über die Expertendiskussion zur Nachhaltigkeit  
Leiter (mit Dr. Paul Klimpel) der Initiative „Nachhaltigkeit in der Di-  
gitalen Welt“*

Jürgen Keiper arbeitet seit 2006 bei der Deutschen Kinemathek – Museum für Film und Fernsehen und ist dort verantwortlich für IT und IT-Projekte wie LOST FILMS, die Internet-Präsentation zu „Wir waren so frei ... Momentaufnahmen 1989/90“, First We take Berlin (mit Peter Paul Kubitz) und die Entwicklung kollaborativer Systeme. Gemeinsam mit Chris Wahl Gründung des Blogs *memento-movie.de* ([www.memento-movie.de](http://www.memento-movie.de)) zum audiovisuellen Erbe. Zuvor Herausgeber der Zeitschrift „Film und Kritik“ (1992-1999), wissenschaftlicher Mitarbeiter des Instituts für Theater-, Film- und Medienwissenschaft der Goethe-Universität Frankfurt am Main (1993-1997), danach beim Deutschen Filminstitut – DIF (1997-2006). Dort u. a. verantwortlich für das IST-Projekt COLLATE – Collaboratory for Annotation, Indexing and Retrieval of Digitized Historical Archive Material (2000-2003), [filmportal.de](http://filmportal.de) (2003-2006), MIDAS – Moving Image Database for Access and Re-use of European film collections (2005-2006) und Convenor der „Standardisation working group for filmographic entries on an european level“ in Zusammenarbeit mit dem CEN – European Committee for Standardization (2005-2006). Zahlreiche Lehraufträge, Vorträge und Expertisen zu dem Thema Film und Neue Medien. Mitarbeiter der AG Media in nestor.



## DR. PAUL KLIMPEL

*Autor: Einführung – Was bleibt?*

*Leiter (mit Jürgen Keiper) der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Dr. Paul Klimpel (geb. 1970 in Minden) studierte Jura in Bonn, später in München, wo er sich auch für Philosophie, Psychologie und Sozialwissenschaften an der Jesuitischen Hochschule für Philosophie einschrieb. Nach Abschluss des Philosophiestudiums 1998 kam er zum Referendariat nach Berlin. Seine Dissertation an der Humboldt-Universität beschäftigt sich mit „Bevormundung und Freiheitsschutz“. 2002 kam er zur Stiftung Deutsche Kinemathek, deren Verwaltungsdirektor er von 2006 bis 2011 war. In dieser Funktion war er auch Geschäftsführer des Netzwerks Mediatheken. Er befasst sich mit den rechtlichen Rahmenbedingungen von Museen und Archiven, insbesondere mit dem Urheberrecht und mit den organisatorischen, technischen und rechtlichen Veränderungen in Gedächtnisorganisationen infolge der Digitalisierung. Seit 2011 koordiniert er den Bereich Kulturelles Erbe im Internet und Gesellschaft Collaboratory e. V., seit 2012 arbeitet er als Rechtsanwalt und ist Leiter des iRights.Lab Kultur.



Foto: Jürgen Keiper



ANDREAS LANGE

*Autor: Die Gaming-Community als Pionier der digitalen Bewahrung  
Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Andreas Lange ist Direktor des Computerspielmuseums in Berlin. Sein Studium der Religions- und Theaterwissenschaft an der Freien Universität Berlin schloss er 1994 mit der Abschlussarbeit „Die Geschichten der Computerspiele betrachtet unter mythentheoretischen Gesichtspunkten“ ab.

Lange ist als Experte vielfältig mit digitaler Unterhaltungskultur befasst. So baute er im Auftrag des fjs e. V. seit 1996 das Computerspielmuseum auf und war seitdem als Kurator für mehr als 30 Ausstellungen verantwortlich. Zudem ist Lange regelmäßig als Redner, Autor und Experte in akademischen und wissenschaftlichen Kontexten tätig.

Lange ist Mitglied der Akademie des Deutschen Entwicklerpreises, der Jury des Deutschen Games Awards LARA und der Beiräte der Deutschen Gamesstage sowie der Stiftung digitale Spielekultur. Lange war Projektleiter des Computerspielmuseums für das europäische Forschungsprojekt KEEP (Keeping Emulation Enviroments Portable) und ist Sprecher der AG Emulation des deutschen Kompetenznetzwerk zur digitalen Langzeitarchivierung, nestor, sowie des 2013 gegründeten Europäischen Verbandes der Computerspielerarchiv-, -museen und Bewahrungsprojekte (EFGAMP).

Foto: Joerg Metzner



DR. VERENA METZE-MANGOLD

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Dr. Verena Metze-Mangold (geb. 1946 in Deutschland) ist seit 1997 Vizepräsidentin der Deutschen UNESCO-Kommission. Von 2005 bis 2011 war sie Bereichsleiterin in der Intendanz des Hessischen Rundfunks. Nach ihrem Studium der Politikwissenschaft, Soziologie und Geschichte leitete sie von 1976 bis 1987 die Evangelische Medienakademie (cpa) im Gemeinschaftswerk der Evangelischen Publizistik in Frankfurt. Sie wechselte zum Hessischen Rundfunk und übernahm die Leitung der Presse- und Öffentlichkeitsarbeit, baute die Medienforschung auf und war 1989/90 Pressesprecherin des ARD-Vorsitzenden. Als Mitglied der Deutschen UNESCO-Kommission hatte sie seit 1982 diverse Funktionen in Kommissionen und Delegationen inne, u. a. nahm sie 2003 und 2005 am World Summit on Information Society der Vereinten Nationen in Genf teil. Seit 2001 vertritt sie Deutschland im zwischenstaatlichen Rat „Information for All“. Schwerpunkt ihrer Veröffentlichungen sind Themen wie Internationale Entwicklung, WTO und Völkerrecht, Presse- und Informationsfreiheit, Menschenrechte sowie interkulturelle Kommunikation.

## KATHRIN PASSIG

*Autor: Facebook, Froschlaich, Folianten*

Kathrin Passig hält sich derzeit als Fellow am Forschungskolleg Humanwissenschaften in Bad Homburg auf. Die Journalistin und Schriftstellerin lebt ansonsten in Berlin vom Schreiben von Sachbüchern. Ihre Beiträge zu den Veränderungen, die die Digitalisierung mit sich bringt, erschienen unter anderem in „Merkur – Deutsche Zeitschrift für europäisches Denken“, in der „Süddeutschen Zeitung“ und in der „ZEIT“. Zuletzt erschien „Internet - Segen oder Fluch“ (Berlin 2012, zusammen mit Sascha Lobo). Mehr unter: <http://kathrin.passig.de>.

Foto: Jan Bölsche



## DR. GEORG REHM

*Autor: Eine Strategie zur Förderung der digitalen Langzeitarchivierung  
Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Dr. Georg Rehm ist Wissenschaftler am Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) in Berlin und Projektmanager des EU-geförderten Exzellenznetzwerks META-NET ([www.meta-net.eu](http://www.meta-net.eu)), das aus 60 Forschungszentren in 34 Ländern besteht. Das Ziel von META-NET sind innovative Forschungsergebnisse im Bereich Sprachtechnologie (insbesondere Maschinelle Übersetzung) und ihr Transfer in die Industrie sowie die Einwerbung zusätzlicher Fördermittel auf europäischer und nationaler sowie regionaler Ebene (in enger Kooperation mit der Europäischen Kommission, Luxemburg). Mit dem Thema Digitale Nachhaltigkeit beschäftigt sich Georg Rehm seit vielen Jahren, insbesondere seit der Leitung zweier Projekte zum Thema „Nachhaltigkeit linguistischer Ressourcen“ am Sonderforschungsbereich 441 der Universität Tübingen (2006-2008). Georg Rehm hat einen Abschluss in Computerlinguistik und Künstlicher Intelligenz, Informatik und Sprachwissenschaft. Er promovierte in Computerlinguistik mit einer Arbeit über semantische Suchmaschinen. Weitere Informationen unter: <http://georg-re.hm>.



## BEATE RUSCH

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Beate Rusch studierte Sinologie und Germanistik in Berlin, Taipeh und Shanghai, bevor sie sich im Bibliothekswesen engagierte.

Seit 2010 ist sie stellvertretende Leiterin des Kooperativen Bibliotheksverbundes Berlin-Brandenburg, der im Konrad-Zuse-Institut zuhause ist. Seit 2012 baut sie zusätzlich die Servicestelle Digitalisierung Berlin (digiS) auf.





PROF. DR. FELIX SASAKI

*Autor: Nachhaltigkeit und Semantic Web*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Prof. Dr. Felix Sasaki ist Senior Researcher am DFKI (Deutsches Forschungszentrum für Künstliche Intelligenz), W3C Fellow und Professor an der Fachhochschule Potsdam im Bereich Informationswissenschaften. Er leitet am DFKI das deutsch-österreichische Büro des W3C (World Wide Web Consortium).

Von 1993 bis 1999 studierte Felix Sasaki Japanisch und Linguistik in Berlin, Nagoya (Japan) und Tokio. Ab 1999 arbeitete er bei dem Department of Computational Linguistics and Text Technology an der Universität Bielefeld, wo er 2004 einen PhD zum Thema „Integration of heterogeneous linguistic resources using XML-based and RDF-based representations“ absolvierte. Felix Sasaki hat langjährige Erfahrung in verschiedenen Standardisierungsbereichen wie Internationalisierung, Web Services und Multimedia Metadaten. Sein Hauptfokus liegt auf der Anwendung von Web-Technologien bei der Repräsentation und Bearbeitung von multilingualen Informationen.



DR. RALF SCHNEIDER

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Dr. Ralf Schneider (geb. 1971) ist seit 2012 Geschäftsführer des Center of Digital Tradition (CODIGT) am ZAK | Zentrum für Angewandte Kulturwissenschaft des Karlsruher Instituts für Technologie (KIT).

Er lehrt seit vielen Jahren in der Medien-, Kultur- und Literaturwissenschaft an der Karlsruher Fakultät für Geistes- und Sozialwissenschaften. Während seiner zwölfjährigen Funktion als Geschäftsführer des Karlsruher Studienzentrums Multimedia koordinierte er neben seiner Lehrtätigkeit Forschungs- und Lehrprojekte mit Universitäten in Deutschland, Frankreich und dem Maghreb. Die Vermittlung der theoretischen und medientechnischen Zusammenhänge in der akademischen Lehre sowie die Beschäftigung mit wissensbewahrenden Medien wie Enzyklopädien aus den Epochen vor dem Internet prägten sein Interesse für Konzepte und Praktiken der analogen und digitalen Wissensvermittlung. Am CODIGT arbeitet er mit seinen Kolleginnen und Kollegen an Themen wie der digitalen Langzeitarchivierung von Forschungsdaten und Daten des Kulturerbes. Zuletzt publizierte er 2013 im Thorbecke Verlag zu Erinnerungsorten und -brüchen.

THORSTEN SIEGMANN

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Thorsten Siegmann ist Kulturwissenschaftler und interessiert sich für die Auswirkungen der Digitalisierung, an der er in der Staatsbibliothek zu Berlin – Preussischer Kulturbesitz aktiv mitarbeitet. Er koordiniert aktuell das

Projekt Europeana Collections 1914–1918. Neben der Digitalisierung großer Bibliotheksbestände aus dem Zeitraum des Ersten Weltkrieges beschäftigt ihn die Frage, wie digitalisierte Kulturgüter aus Kultureinrichtungen nutzbar gemacht werden können. Zuvor hat Siegmann für das Institut für Museumsforschung der Staatlichen Museen – Preußischer Kulturbesitz gearbeitet.

DR. ERIC W. STEINHAUER

*Autor: Wissen ohne Zukunft? Der Rechtsrahmen der digitalen Langzeitarchivierung von Netzpublikationen*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Eric W. Steinhauer (geb. 1971 in Unna) studierte Rechtswissenschaft, katholische Theologie und Philosophie in Münster sowie Politik- und Erziehungswissenschaft in Hagen. Nach dem juristischen Staatsexamen war er Bibliotheksreferendar an der Universität Freiburg. Seine Dissertation befasst sich mit der Wissenschaftsfreiheit katholischer Theologen. Er arbeitet seit 2003 in unterschiedlichen Funktionen an Universitätsbibliotheken und ist seit 2009 Dezernent für Medienbearbeitung an der Universitätsbibliothek der FernUniversität Hagen. Er hat mehrere Lehraufträge u. a. an der Humboldt-Universität zu Berlin und der Bergischen Universität Wuppertal, ist Autor zahlreicher Fachveröffentlichungen und war als Sachverständiger an mehreren bibliotheksbezogenen Gesetzgebungsverfahren auf Länderebene beteiligt. Im Kompetenznetzwerk nestor hat er sich mit Fragen der Langzeitarchivierung beschäftigt. Fachlich interessiert er sich besonders für Bibliotheken als Gedächtnisinstitutionen und die damit zusammenhängenden rechtlichen Probleme und kulturwissenschaftlichen Fragestellungen vor allem vor dem Hintergrund der Herausforderungen der Digitalisierung.

Foto: Mathias Schindler



TOBIAS STEINKE

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Tobias Steinke (geb. 1971 in Berlin) studierte Informatik an der TU Darmstadt und arbeitete danach in einer Softwarefirma der Telekommunikationsbranche. Seit 2003 ist er als wissenschaftlicher Mitarbeiter in der IT-Abteilung der Deutschen Nationalbibliothek (DNB) in Frankfurt mit digitaler Langzeitarchivierung beschäftigt. Er verantwortete dort das Projekt kopal zum Aufbau eines digitalen Langzeitarchivs und vertritt die DNB in nationalen und internationalen Gremien (u. a. IIPC und METS Editorial Board) sowie in Standardisierungsaktivitäten im Bereich der Langzeitarchivierung (DIN NABD 15). Bei EU- und DFG-geförderten Projekten zur Langzeitarchivierung wirkte er aktiv mit (u. a. KEEP, SHAMAN, DP4Lib und LuKII). Seine Tätigkeit umfasst auch den Aufbau eines Workflows für Webarchivierung.





ARMIN STRAUBE

*Autor (mit Tobias Beinert): Aktuelle Herausforderungen der digitalen Langzeitarchivierung*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Armin Straube (geb. 1979 in Halle/Saale) ist seit September 2012 Geschäftsführer von nestor, dem deutschen Kompetenznetzwerk zur digitalen Langzeitarchivierung. Die Geschäftsstelle ist in der Deutschen Nationalbibliothek in Frankfurt/M. angesiedelt.

Nach dem Studium der Geographie (Universität Halle), Geschichte und VWL (FernUniversität Hagen) und des Archives and Records Managements (University College Dublin) war er in einem EU-Projekt zur Datenlieferung an die Europeana tätig.

Inhaltlich arbeitet er in den nestor AGs Kooperation und Vernetzung, Policy und Zertifizierung.



DR. MAIK STÜHRENBERG

*Autor: Digitale Langzeitarchivierung aus Sicht der IT*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Dr. Maik Stührenberg arbeitet als wissenschaftlicher Mitarbeiter an der Universität Bielefeld im Bereich Computerlinguistik / Texttechnologie und befasst sich seit seinem Studium der Computerlinguistik mit der nachhaltigen Strukturierung und Speicherung von multimodalen Informationen mit Hilfe standardisierter Auszeichnungs- und Metasprachen. Er gründete 2007 den Arbeitskreis „Sprach- und texttechnologische Methoden im eLearning“ der GSCL und ist aktuell E-Learning-Beauftragter der Fakultät für Linguistik und Literaturwissenschaft der Universität Bielefeld.

Als ein Vertreter der Universität Bielefeld arbeitet er beim DIN im NA 105 Normenausschuss Terminologie (NAT), NA 105-00-06 AA Arbeitsausschuss Sprachressourcen an der Entwicklung standardisierter Formate zur langfristigen Speicherung linguistischer Daten mit. Nachhaltigkeit war auch Gegenstand seiner im Jahr 2012 verfassten Dissertation mit dem Thema „Auszeichnungssprachen für linguistische Korpora. Theoretische Grundlagen, De-facto-Standards, Normen“.



DR. ANNE VAN RAAY

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Dr. Anne van Raay bearbeitet seit Anfang 2012 im IQF-Projekt zur Langzeitarchivierung komplexer digitaler Objekte am Zentrum für Angewandte Kulturwissenschaften (ZAK) des Karlsruher Instituts für Technologien (KIT) rechtliche Fragestellungen. Sie ist dort für den Aufbau der Informationsstelle Recht des in Gründung befindlichen Center of Digital Tradition (CODIGT)

zuständig. Nach Studium in Heidelberg und Referendariat am OLG Düsseldorf war sie zunächst für den Verbraucherzentrale Bundesverband e. V. (vzbv) in Berlin tätig, anschließend Doktorandin und wissenschaftliche Mitarbeiterin bei Prof. Dr. Thomas Dreier M.C.J. am Zentrum für Angewandte Rechtswissenschaft (ZAR) der Universität Karlsruhe (TH). Seit 2007 als Rechtsanwältin in Karlsruhe zugelassen, war sie 2008 bis 2010 Promotionsstipendiatin des Landes Baden-Württemberg (Schlieben-Lange-Programm); 2011 erfolgte die Promotion an der Universität Freiburg. Neben Rechtsfragen nachhaltiger Datenspeicherung beschäftigt sie sich aktuell etwa mit den rechtlichen Bedingungen des Zugangs und der Weiterverwertung von Daten der öffentlichen Hand.

Foto: FOTO DIGITAL heidi offerdingner

### PROF. DR. CHRIS WAHL

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Dr. Chris Wahl (geb. 1974 in Stuttgart) studierte Theater-, Film- und Fernsehwissenschaft, Romanische Philologie und Deutsch als Fremdsprache an der Ruhr-Universität Bochum sowie an der Escola Superior de Teatro e Cinema in Lissabon. 2003 promovierte er in Bochum über „Das Sprechen des Spielfilms“. Seitdem hat er mehrere DFG-Forschungsprojekte zu den Ufa-Sprachversionsfilmen der 1930er Jahre, zu den kinematographischen Stilmitteln Zeitlupe und Mehrfachbelichtung und zum Begriff Filmstil durchgeführt. 2009 erschien sein Buch über Sprachversionsfilme.

Seit Juni 2013 ist er Heisenberg-Proffessor für das Audiovisuelle Kulturerbe an der Hochschule für Film und Fernsehen „Konrad Wolf“ in Potsdam-Babelsberg, wo er bereits seit 2010 angestellt ist. Eines seiner Anliegen ist die Verbesserung der Kommunikation über Fragen des Filmberbes zwischen verschiedenen Interessensgruppen wie Wissenschaftlern, Archivaren, Produzenten usw. Dieser Aufgabe widmet er sich nicht nur als Sprecher der AG Filmwissenschaft in der Gesellschaft für Medienwissenschaft, sondern auch mittels des vom BKM-geförderten Blogs [memento-movie.de](http://memento-movie.de) ([www.memento-movie.de](http://www.memento-movie.de)), das er gemeinsam mit Jürgen Keiper betreibt.



### DR. KATRIN WELLER

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Katrin Weller befasst sich seit Jahren mit Forschungsfragen rund um Social-Media-Angebote und deren Nutzung. Seit Januar 2013 ist sie wissenschaftliche Mitarbeiterin bei GESIS – Leibniz-Institut für Sozialwissenschaften. Hier baut sie im Datenarchiv für Sozialwissenschaften einen neuen Forschungsbereich „Social Web Daten“ auf. Bis 2012 war sie wissenschaftliche Mitarbeiterin in der Abteilung für Informationswissenschaft an der Heinrich-Heine-Universität Düsseldorf und beschäftigte sich dort zuletzt insbesondere mit den Möglichkeiten zur Datenanalyse bei Twitter und mit Veränderungen der





wissenschaftlichen Arbeitswelt durch Online-Technologien. Ihre Promotion zum Thema „Knowledge Representation in the Social Semantic Web“ schloss sie 2010 ab. Sie hat Geschichte, Informationswissenschaft und Medienwissenschaft in Düsseldorf studiert.



PD DR. STEFAN WILLER

*Autor: Kulturelles Erbe und Nachhaltigkeit*

PD Dr. Stefan Willer (geb. 1970) ist stellvertretender Direktor des Zentrums für Literatur- und Kulturforschung, einer geisteswissenschaftlichen Forschungseinrichtung in Berlin. Er lehrt als Privatdozent an der Technischen Universität Berlin, außerdem hatte er Gastprofessuren an der Ludwig-Maximilians-Universität München, der Humboldt-Universität zu Berlin und der Stanford University inne. Zu seinen Arbeitsschwerpunkten zählen kulturelle Theorien und Imaginationen von Generation, Erbe und Zukünftigkeit. Buchveröffentlichungen u. a.: „Botho Strauß zur Einführung“ (2000), „Poetik der Etymologie. Texturen sprachlichen Wissens in der Romantik“ (2003), „Das Beispiel. Epistemologie des Exemplarischen“ (als Mitherausgeber, 2007), „Das Konzept der Generation. Eine Wissenschafts- und Kulturgeschichte“ (als Mitautor, 2008), „Erbe. Übertragungskonzepte zwischen Natur und Kultur“ (als Mitherausgeber, 2013), „Prophetie und Prognostik. Verfügungen über Zukunft in Wissenschaften, Religionen und Künsten“ (als Mitherausgeber, 2013), „Erbfälle. Theorie und Praxis kultureller Übertragung in der Moderne“ (in Vorbereitung für 2013).



STEFAN WOLF

*Autor: Informationen für die Zukunft – EINE Infrastruktur?*

*Mitglied der Initiative „Nachhaltigkeit in der Digitalen Welt“*

Stefan Wolf (geb. 1960 in Konstanz) ist nach Studium in Freiburg und Würzburg und der Ausbildung zum „Dipl.-Bibl. (FH) an wissenschaftlichen Bibliotheken“ in Stuttgart seit 1994 am Bibliothekservice-Zentrum Baden-Württemberg angestellt. Im Laufe der Zeit war und ist er für Sacherschließung, Sonderregelwerke der Formal- und Sacherschließung, bibliotheksfachliche Aspekte der digitalen Bibliothek sowie die digitale Langzeitarchivierung zuständig. Er arbeitete an Projekten wie dem Baden-Württembergischen Online-Archiv BOA, Literatur im Netz und dem Oberwolfach Digital Archive mit, koordinierte die Entwicklung der Software SWBcontent zur Speicherung und Präsentation von Webpublikationen und wirkte auch an der Entwicklung eines digitalen Bildarchivs der Staatsgalerie Stuttgart mit.

Im Zuge seiner beruflichen Tätigkeit ist er Mitglied in verschiedenen Gremien: u. a. nestor, AG Hosting und Langzeitarchivierung der Kommission Zukunft der Informationsinfrastruktur und die Gesprächsreihe zwischen Archiven (Bund / Länder / Kommunen) und Pflichtexemplarbibliotheken

(DNB, Staats- und Landesbibliotheken, Verbände). Und er redigierte das Schichtenmodell der digitalen Langzeitarchivierung in Baden-Württemberg.





Harddrives. Foto: Paul R. Potts



Internet & Gesellschaft  
**<Co:laboratory>**

Besuchen Sie das Internet & Gesellschaft Collaboratory e. V. auf  
[www.collaboratory.de](http://www.collaboratory.de)  
[www.youtube.com/collaboratoryvideo](http://www.youtube.com/collaboratoryvideo)

[www.berliner-appell.org](http://www.berliner-appell.org)

Eine Publikation des Internet & Gesellschaft Collaboratory e. V.  
bei iRights.Media



9 783944 362038

€ 8,00 (D)

**iRIGHTS**  
media

