

Zur Entwicklung  
eines Beschreibungsprofils  
für eine nationale  
Langzeit-Archivierungs-Strategie  
(National long-term Preservation Policy)  
- ein Beitrag aus der Sicht der Wissenschaften

Thomas Severiens und Eberhard R. Hilf

Institute for Science Networking Oldenburg  
an der Carl von Ossietzky Universität Oldenburg





Zur Entwicklung  
eines Beschreibungsprofils  
für eine nationale  
Langzeit-Archivierungs-Strategie

- ein Beitrag aus der Sicht  
der Wissenschaften

Thomas Severiens  
Eberhard R. Hilf

Institute for Science Networking Oldenburg

Herausgegeben von

nestor - Kompetenznetzwerk Langzeitarchivierung und  
Langzeitverfügbarkeit Digitaler Ressourcen für Deutschland

nestor - Network of Expertise in Long-Term Storage of Digital Resources

<http://www.langzeitarchivierung.de>

Projektpartner

Bayerische Staatsbibliothek, München

Bundesarchiv

Computer- und Medienservice / Universitätsbibliothek der Humboldt-Universität zu Berlin

Die Deutsche Bibliothek, Leipzig, Frankfurt am Main, Berlin (Projektleitung)

Generaldirektion der Staatlichen Archive Bayerns, München

Institut für Museumskunde, Berlin

Niedersächsische Staats- und Universitätsbibliothek, Göttingen

© 2006

nestor - Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit  
Digitaler Ressourcen für Deutschland

Der Inhalt dieser Veröffentlichung darf vervielfältigt und verbreitet werden, sofern der Name des Rechteinhabers „nestor - Kompetenznetzwerk Langzeitarchivierung“ genannt wird. Eine kommerzielle Nutzung ist nur mit Zustimmung des Rechteinhabers zulässig.

Betreuer dieser Veröffentlichung:

Niedersächsische Staats- und Universitätsbibliothek Göttingen  
(Heike Neuroth und Stefan Strathmann)

URN: <urn:nbn:de:0008-20051114021>  
<http://nbn-resolving.de/urn:nbn:de:0008-20051114021>

---

Die vorliegende Arbeit wurde von Prof. Dr. Eberhard R. Hilf und Thomas Severiens vom "Institute for Science Networking" in Oldenburg im Auftrag von nestor – Kompetenznetzwerk Langzeitarchivierung erstellt. Betreut wurde die Expertise von der Niedersächsischen Staats- und Universitätsbibliothek Göttingen. Der Inhalt gibt ausschließlich die Meinung der Autoren wieder.

Die Expertise konzentriert sich in erster Linie auf die Bedürfnisse der (natur-)wissenschaftlichen Forschung und ihre Anforderungen an die wissenschaftliche Informationsversorgung. Sie ist im Wesentlichen aus der Perspektive von Autoren und Nutzern von Bibliotheken geschrieben. Die Sicht der Institutionen, die für Langzeitarchivierung verantwortlich sind oder sein werden, ist partiell einbezogen. Die Anforderungen die sich aus der digitalen Langzeitarchivierung an Archive, Museen und weitere Gedächtnisinstitutionen ergeben, werden ebenfalls teilweise ergänzend herangezogen.

Die vorliegende Expertise ist ein wichtiger Schritt zur Anregung und Diskussion einer nationalen deutschen Langzeitarchivierungs-Policy im Rahmen von nestor.

nestor, das Kompetenznetzwerk Langzeitarchivierung will die unterschiedlichen Interessengruppen an der digitalen Langzeitarchivierung berücksichtigen und in Bezug auf eine nationale Langzeitarchivierungs-Policy, in gemeinsame, kooperative Entwicklungen einbeziehen.

Auf der Basis der vorliegenden Expertise und aus den Ergebnissen aus Umfragen und Arbeitsgruppendifkussionen wird nestor Empfehlungen zu einer nationalen Langzeitarchivierungs-Policy entwickeln und diese öffentlich zur Diskussion stellen. Dabei werden über die naturwissenschaftliche Sicht hinaus auch andere Sichtweisen und Interessengruppen berücksichtigt.

Die digitale Langzeitarchivierung ist ein umfangreiches neues Aufgabenfeld im digitalen Informations- und Dienstleistungszeitalter. Manche Lösungen, die für die Bewahrung des analogen kulturellen und wissenschaftlichen Erbes entwickelt wurden, lassen sich auf diesen digitalen Bereich übertragen; in anderen Fällen sind neue Lösungen gefragt. Eine nationale Langzeitarchivierungs-Policy hilft, den (gesetzlichen) Rahmen für die neuen Rollen, Zuständigkeiten und Verantwortlichkeiten zu entwickeln.

Für die Partner des Projekts nestor - Kompetenznetzwerk Langzeitarchivierung  
Heike Neuroth und Stefan Strathmann  
Niedersächsische Staats- und Universitätsbibliothek Göttingen



Zur Entwicklung eines  
Beschreibungsprofils für eine nationale  
Langzeit-Archivierungs-Strategie  
(National long-term Preservation Policy)  
- ein Beitrag aus der Sicht der Wissenschaften -

Thomas Severiens und Eberhard R. Hilf  
Institute for Science Networking Oldenburg

Dezember 2005





# Inhaltsverzeichnis

<b>1</b>	<b>Vorwort</b>	<b>3</b>
<b>2</b>	<b>Empfehlungen</b>	<b>5</b>
2.1	Handlungsfelder	6
2.2	Vorschläge für strukturelle Maßnahmen	17
<b>3</b>	<b>Materialien</b>	<b>23</b>
3.1	Einführung	23
3.1.1	Strategische Ziele einer nationalen LZA-Policy	23
3.1.2	Vielfalt der Entwicklungen (Länder, Inhalte, Autoren)	24
3.1.3	Geschichte	27
3.1.4	Ära ohne Langzeit-Gedächtnis	31
3.2	Museen und Archive	32
3.3	Begriffe und deren Eingrenzungen	35
3.4	Die Fragebogenaktion	38
3.4.1	Idee, Strategie, Durchführung, Rücklauf	38
3.4.2	Der Fragebogen	40
3.4.3	Archivierung und Extrakt von Fragebogen-Antworten	40
3.5	Aktivitäten im Ausland	43
3.5.1	Australien, Beispiel eines seit langem aktiven Industriestaates	44
3.5.1.1	PADI: Preserving Access to Digital Information	46
3.5.1.2	PANDORA Archive	47
3.5.1.3	The International Information Consortium Plan	49
3.5.2	Sehr kleine Staaten	50

3.5.3	Finnland als Beispiel eines mittelkleinen Industriestaates . . . . .	51
3.5.4	Frankreich . . . . .	51
3.5.5	Großbritannien . . . . .	53
3.5.5.1	Das Joint Information Systems Committee . . . . .	53
3.5.5.2	Die National Archives . . . . .	54
3.5.6	Niederlande . . . . .	54
3.5.7	Norwegen . . . . .	55
3.5.8	Die Schweiz . . . . .	56
3.5.9	Tschechische Republik . . . . .	59
3.5.10	Die Vereinigten Staaten von Amerika . . . . .	59
3.5.10.1	Das PREMIS-Projekt . . . . .	59
3.5.10.2	Das Digital Information Infrastructure and Preservation Program . . . . .	60
3.5.10.3	Der e-Government Act . . . . .	61
3.5.11	Europäische Union . . . . .	62
3.5.11.1	Langfristige Pläne . . . . .	62
3.5.11.2	Aktuelle Projekte . . . . .	64
3.5.11.3	Das Forschungsprogramm FP6 . . . . .	66
3.5.12	Die UNESCO . . . . .	66
3.5.13	Aktivitäten von Nationalbibliotheken weiterer Staaten . . . . .	67
3.6	Information in Wissenschaft und Lehre . . . . .	69
3.6.1	Dokument-Typen und Proprietät . . . . .	69
3.6.2	Redundanz . . . . .	74
3.6.3	Metadaten . . . . .	75
3.6.4	Vernetzte Inhalte, lebende Dokumente . . . . .	76
3.6.5	Internationale Langzeit-Archivierungskonferenz der Physik . . . . .	77
3.6.6	Zur Langzeit-Archivierung benachbarte Informationen . . . . .	81
3.7	Verschiedene Sichten . . . . .	82
3.7.1	Sicht des Staates . . . . .	82
3.7.2	Sicht der Erzeuger von Information . . . . .	85
3.7.3	Sicht der Nutzer . . . . .	86
3.7.4	Sicht der archivierenden Institutionen . . . . .	89

3.8	Szenarien . . . . .	93
3.8.1	Zentralistisches Szenarium . . . . .	93
3.8.2	Dezentrales Szenarium . . . . .	93
3.8.3	Dezentrales Szenarium mit Leitungsstrukturen . . . . .	95
3.9	Fragen- und Planungslisten . . . . .	102
3.9.1	Strategische Fragen an eine LZA-Policy . . . . .	102
3.9.2	Planungs-Checkliste . . . . .	102
3.9.3	Vorbereitende Schritte zu einer LZA-Policy . . . . .	103
3.9.4	Themen und Beschränkung einer nationalen LZA-Policy . . . . .	104
3.10	Erläuterungen zu Einzelthemen . . . . .	106
3.10.1	Sicherheit: Datenschutz, Geheimhaltung . . . . .	106
3.10.2	Redundanz . . . . .	107
3.10.3	Internationale Vernetzung . . . . .	107
3.10.4	Finanzierung . . . . .	108
3.10.5	Ausbildungsfragen . . . . .	109
3.10.6	Mögliche Marginalität von Publikationen . . . . .	109
3.10.7	Zugriffskonditionen . . . . .	110
3.10.8	Sicherung der Authentizität . . . . .	111
3.10.9	Datenschutz und Sicherheit . . . . .	112
3.10.10	Organisationsform . . . . .	113
3.10.11	Metadaten . . . . .	113
3.10.12	Migration von Formaten . . . . .	114
3.10.13	Gesetzliche Rahmenbedingungen . . . . .	114
3.10.14	Experten . . . . .	116
<b>4</b>	<b>Schlusswort</b>	<b>117</b>
<b>A</b>	<b>Anhang: Nationalbibliotheken weltweit</b>	<b>119</b>



# 1

## Vorwort

Ziel dieser Studie ist es, zur Entwicklung einer *nationalen Policy für die Langzeit-Archivierung digitaler Objekte aus Kultur, Wissenschaft und akademischer Lehre* (National long-term preservation policy), kurz LZA-Policy, beizutragen.

Sie wurde 2005 vom *Institute for Science Networking* in Oldenburg unter aktiver Unterstützung durch die *Niedersächsische Staats- und Universitätsbibliothek Göttingen SUB* für das BMBF-Projekt *nestor Kompetenznetzwerk Langzeit-Archivierung* erarbeitet.

Die Langzeit-Archivierung digitaler Dokumente wird von allen Gebieten der Wissenschaft<sup>1</sup>, der Kultur, der Bildung, der Museen und der klassischen Archive verlangt. Entsprechend offen und breit muss eine nationale LZA-Policy angelegt sein.

Eine Langzeit-Archivierung will Materialien der Gegenwart und Vergangenheit für die fernere Zukunft speichern. Die Realisierung bewegt sich daher in mehrfacher Hinsicht auf einem „moving ground“: Umgestalten werden sich das Management wissenschaftlicher, kultureller und von Bildungs-Informationen, die Art, die Wege und Gewohnheiten, fachlich zu kommunizieren und sich auf Werke des kulturellen und wissenschaftlichen Erbes zu beziehen, und die technischen Möglichkeiten für digitale innovative Dienste; für eine nationale LZA-Policy, die ja langfristige sichere und stabile Strukturen, Finanzierung und Einbettung vorausschauend garantieren soll – eine ungeheure Herausforderung.

Schon 1991 wurde in den USA bereits das erste große, bis heute funktionierende digitale Langzeit-Archiv, das *ePrint Archive*<sup>2</sup> von Los Alamos, jetzt Cornell University, gegründet. Eine Langzeit-Archivierung ist daher weniger eine neue technische, sondern vor allem eine nationale, teilweise sogar internationale organisatorische und politische Herausforderung. Als Leitlinie für die Beteiligten und zur Selbstbindung der Institutionen bedarf es einer nationalen LZA-Policy. Diese Expertise will daher vor allem zur Identifizierung und Beleuchtung von Handlungsfeldern beitragen und so die

---

<sup>1</sup>Einen wirklich großen Bereich der Wissenschaft überdeckte zuletzt: Alexander von Humboldt; *Kosmos; Entwurf einer physischen Weltbeschreibung*; Stuttgart; Band 1. 1845, Bd. 2. 1847, Bd. 3. 1850, Bd. 4. 1858.

<sup>2</sup><http://www.arxiv.org>

Erarbeitung einer nationalen Policy durch die Verantwortlichen unterstützen. Sie bildet keine Stellungnahme *des nestor*-Projektes sondern ist als Expertise und Essay aus der Sicht der Wissenschaften als Materialie *für* das Projekt angelegt.

Die organisatorischen und politischen Perspektiven werden daher in den Vordergrund gestellt; sie sollen eine breite Diskussion befördern, Kenntnisse verbreiten helfen und durch Beides die zukünftige Akzeptanz mit vorbereiten und so den Weg ebnen, eine technische und stabile Lösung für die Langzeit-Archivierung durch die beteiligten Institutionen kooperativ zu etablieren. Diese Expertise trägt einzelne Facetten zur vielschichtigen Diskussion im Vorfeld aus der Sicht der Autoren – mit ihrem vielleicht aus teils vierzig Jahren Berufserfahrung in der Naturwissenschaft verengtem Blick – zusammen. Sie soll den Leser, die Leserin anderen fachlichen und beruflichen Hintergrundes ermuntern, das gesamte Gebiet vom eigenen Standpunkt aus zu beleuchten und Fehlendes oder zu kurz Gekommenes beizutragen. Als selbst nicht an den Entscheidungsebenen der Archivierung Beteiligte haben die Autoren dieser Expertise den Vorteil des Blickes von außen, von den Anforderungen der Autoren und Nutzer her – die Entscheidungen zu politischen Weichenstellungen und erst recht die Realisierung und Detaillierung der Policy liegen beim Bund, den Ländern, den Archivierungsinstitutionen und den Fachorganisationen.

Es werden strategische Ausgangsüberlegungen zusammengestellt und die Ansätze in einigen anderen Ländern beleuchtet. Hierzu wurde eine international gestreute Umfrage durchgeführt und es wurden Informationen zu einschlägigen Aktivitäten im Ausland zusammengestellt. Wir hoffen dadurch, den Horizont der Diskussion zu weiten, und zur Akzeptanz einer zukünftigen nationalen LZA-Policy beizutragen.

Als Form der Expertise wurde ein induktiver, synoptischer und argumentativer Ansatz als in die Zukunft orientierte Hilfe gewählt, anstelle einer deduktiven Grundlagenarbeit, deren eher vergangenheitsbezogene Sicht den Blick voraus leicht verstellen könnte.

Für die Frage, was bei einer Archivierung digitaler Objekte unter den *gegenwärtigen* Gesetzen beachtet werden muss, wird auf die entsprechende *nestor*-Rechts-Expertise<sup>3</sup> verwiesen.

Die Autoren danken Stefan Strathmann für die kompetente und kritische Durchsicht und ein professionelles Lektorat, und Heike Neuroth für wertvolle Hinweise, beide an der *Niedersächsischen Staats- und Universitätsbibliothek Göttingen SUB*, sowie Kerstin Blöching für die sorgfältige Durchsicht. Über weiterführende oder korrigierende Hinweise der Leser und Leserinnen würden sich die Autoren freuen.

31. 12. 2005

Thomas Severiens

Eberhard R. Hilf

---

<sup>3</sup>Rechtsanwälte Goebel und Scheller; *Digitale Langzeit-Archivierung und Recht*; nestor Materialien 1, 2004; <http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:0008-20040916022>

## Teil 2

# Empfehlungen für eine nationale Langzeit-Archivierungs-Policy

Eine staatliche digitale Langzeit-Archivierung (LZA) dient der Bewahrung der wissenschaftlichen Erkenntnisse und der kulturellen Leistungen der Gegenwart und der Vergangenheit für zukünftige Generationen.

Wir werden hier auf die *Langzeit*-Archivierung fokussieren, also auf die Bewahrung über die nächste technische Umwälzung, über den derzeitigen Umbruch der Kommunikationskultur, der Geschäftsmodelle von Verlagen, der Bewertungsverfahren hinaus – jedenfalls soweit dies heute erkennbar scheint. Viele Anforderungen lassen sich aber schon aus den Erfahrungen der institutionellen Archive mit ihrem täglichen Umgang bei der digitalen Archivierung unter den üblicherweise sehr pragmatisch gewählten Anfangsbedingungen gewinnen.

Diese Expertise fußt auf einer Umfrage unter nationalen Bibliotheken, auf einer Vielzahl vertiefender Gespräche und Telefonate sowie auf einer Webrecherche.

Der erste Teil ist für eilige bzw. für professionell mit Langzeit-Archivierung befasste Leser gedacht. Es werden hier aus Teil II Handlungsfelder extrahiert und daraus synoptisch und kondensiert Vorschläge zu ersten Maßnahmen vorangestellt.

Im zweiten Teil<sup>1</sup> werden dann Teilergebnisse der Recherchen wiedergegeben, soweit sie weiterführend erscheinen, es werden Materialien analysiert und einschlägige Projekte und Aktivitäten in anderen Ländern vorgestellt.

---

<sup>1</sup>So wie beim Teekochen die Genießbarkeit des Tees (in Analogie zu Teil I dieser Expertise) von den Eigenschaften der verwendeten Teeblätter und der Zubereitung (Teil II) abhängen.

## 2.1 Handlungsfelder

Die notwendige *Spannweite einer nationalen LZA-Policy* lässt sich nur schwer unterschätzen. Hier werden jedenfalls einige Handlungsfelder abgesteckt, die gesichert dazugehören sollten.

Die nationale Langzeit-Archivierung *digitaler* Dokumente ist technisch und organisatorisch eine relativ homogene Herausforderung verglichen mit der breiten Vielfalt *physischer* Werke aus Wissenschaft, Kultur und Bildung, die für die Nachwelt aufbewahrt werden sollen und der für diese damit verbundenen notwendigen Vielfalt der technischen und organisatorischen Vorkehrungen.

Die *Vielfalt der Bereiche*, die einer Richtlinie und Regelung durch eine nationale LZA-Policy digitaler Dokumente bedürfen, bezieht sich vor allem auf Anforderungen, die sich aus dem Inhalt der Dokumente, der Akteure/Erzeuger, und den vermuteten Absichten und der Wünsche zukünftiger Leser und Nutzer ergeben. Sie spiegeln insgesamt die sehr unterschiedlichen Kulturen, Traditionen und Gewohnheiten wider, die die Kommunikation – hier mit den erwarteten zukünftigen Generationen – und damit die Anforderungen an die Langzeit-Archivierung bestimmen.

Es ist daher eine große Herausforderung, die entsprechend breit und vielstimmig anzulegende Diskussion mit Beiträgen aus allen Bereichen der Wissenschaft, Kultur und Bildung sowie der professionellen LZA-Institutionen bei der Entwicklung einer nationalen LZA-Policy zu organisieren und dann zu nutzen.

*Die zukünftigen Leser/Nutzer* werden z. B. Wissenschaftler sein, die Arbeiten ihrer Vorgänger für ihre berufliche Arbeit brauchen, Lehrende, die aufbereitete Inhalte aufbauend auf vorhandenem gesammeltem Wissen benötigen, Behörden und staatliche Organe, die Entscheidungen mit Rückgriff auf die Vergangenheit realisieren oder belegen müssen, Kunstbeflissene, die Informationen über Werke vergangener Epochen aufnehmen wollen, sowie Bürger, die sich über die Vergangenheit in allen ihren Facetten informieren wollen.

*Die mit der Langzeit-Archivierung zukünftig Beauftragten* werden wie bisher Bibliotheken, Museen und Archive, möglicherweise aber auch neue Institutionen sein. Denn das Anforderungsprofil wird sich ausweiten wie z. B. die Inhaltserschließung durch Fachexperten, die Unterstützung von Materialaufbereitungen durch Dritte, die Organisation der Kommunikation mit den zukünftigen Lesern.

Zu diesem sehr breiten Feld wollen wir aus der Sicht der Wissenschaften und beschränkt auf digitale Quellen – seien sie originär digital oder durch Digitalisierung gewonnen – beitragen. Zunächst benennen wir nun in kompakter Form einige *technische Handlungsfelder*, um dann Vorschläge für erste Maßnahmen vorzustellen.

1. Als Erstes müssen natürlich digitale Dokumente, die in der Zukunft erwartet werden sollen, dann noch vorhanden sein. Die **Sicherheit gegen Verlust** von LZA-Materialien im digitalen Zeitalter wird daher ein zentrales Thema einer natio-



nalen LZA-Policy sein. Sie sollte weitestmöglich auf *Redundanz* bauen. Dies ist digital ohne besonders großen technischen Aufwand umzusetzen, z. B. durch Backup-Kopien im Hause sowie zusätzlich durch Kopien in *anderen* verwandten Institutionen, mit denen dazu Vereinbarungen getroffen werden müssten.

Eine brisante Teilherausforderung ist die *politische Sicherheit* gegen Verlust, Verschließen oder Verfälschung. Hier wird in der internationalen Diskussion die Realisierung durch **Spiegelungen im Ausland** als richtige Lösung angesehen.

2. Aufbewahrte Dokumente können ihre Wirkung in der Zukunft nur insoweit entfalten, als sie dann auch eingesehen werden dürfen. Daher ist es eine zentrale Aufgabe für die nationale LZA-Policy, allgemeine Regeln für die **Zugriffsrechte** zu setzen. Aus Sicht der Nutzer heißt dies, die Quellen bis auf besonders sensitive Daten, die auf Grund rechtlicher Auflagen sicher nach außen geschützt werden müssen, langfristig, bequem und digital offen zugänglich zu halten.
3. Die nationale LZA-Policy wird generelle Regeln zur technischen **Lesbarkeit** gespeicherter digitaler Dokumente auch in ferner Zukunft setzen müssen. Prinzipiell kann dies nur durch technische Lösungen, die unabhängig vom Markt und der jeweiligen Wirtschaftskraft kommerzieller Firmen sind, geschehen. Technisch bedeutet dies nichtproprietäre offene Formate sowie eine international zu vereinbarenden regelmäßige<sup>2</sup> **Migration** auf aktuelle Formate. Verantwortlich müssen dazu staatliche Institutionen sein mit dem dazu professionell geschulten Personal.
4. Um langfristig die möglichst vollständige **Erhaltung von Funktionalität und Inhalt** zu erleichtern, die Nachnutzung technisch und inhaltlich auch in der fernen Zukunft zu ermöglichen, sind technische Vorkehrungen erforderlich. Die entsprechenden Anforderungen an die Dokumente sind am einfachsten bereits bei der Anlieferung zu stellen.

Für eine Langzeit-Archivierung sollte möglichst feingliedrig sowohl der **formale wie der semantische Gehalt** von Dokumenten maschinenlesbar eingearbeitet und festgehalten werden, soweit hierzu internationale Standards (Metadaten, fachspezifische Markup-Sprachen) verfügbar sind.

5. Eine **Kontext-Erhaltung** (z. B. Bezüge und Verlinkungen zu anderen Dokumenten) erleichtert in der Zukunft das Einordnen und Einarbeiten in das aktuelle Arbeitsfeld. Sie sollte, soweit möglich, bereits bei der Eingabe der Dokumente eingearbeitet werden und anschließend fortlaufend (soweit zukünftig intelligent möglich auch automatisiert) aktualisiert werden. Der Kontext und damit der intellektuelle Inhalt könnten sonst unwiederbringlich verloren gehen – insbesondere bei verteilt gespeicherten Teilen von Dokumenten (sogenannten „Kollagen“ aus verteilten Fragmenten), bei Dokumenten, die Teile einer thematischen Sammlung waren, oder bei Dokumenten, die kommunikative Elemente enthalten oder aus einem Online-kommunikativen Umfeld stammen.

---

<sup>2</sup>Erfahrungsgemäß z. B. einmal jährlich.

6. Große Besorgnis herrscht in der aktuellen Diskussion in der Wissenschaft, inwieweit digitale Dokumente in der fernen Zukunft intellektuell verstehbar gehalten werden können trotz der Weiterentwicklung der Fachsprachen. Durch begleitende Zusatzdienste wie der **Übertragung in die jeweils aktuell verstehbare Fachsprache** ließe sich dem entgegenwirken. Natürlich muss das Original dabei erhalten bleiben. Die Extraktion der Inhalte und deren Abbildung auf eine für die Wissenschaft und Kultur notwendige Langzeit-Archivierung der Inhalte wird einer ganz neuen Qualität und Quantität inhaltlicher Nachbearbeitung und Kondensation bedürfen, um in ferner Zukunft wirklich nachnutzbares Material zu erhalten. Eine solche fortlaufende Übertragung in weiter nutzbares Wissen, in Erkenntnis und Einsicht, wird in vielen Gebieten auch einen entscheidenden Vorteil für die nationale wissenschaftliche und wirtschaftliche Leistungsfähigkeit bedeuten und entsprechend einen Schwerpunkt der aktuellen Diskussion zur nationalen LZA-Policy bilden.

Konkret sieht dies in den nach Themen unterschiedenen Bereichen der Langzeit-Archivierung sehr verschieden aus. In den exakten Naturwissenschaften sind die Erkenntnisse und Aussagen über die Natur nicht natürlich-sprachlich, sondern durch mathematische Formeln, experimentelle Aufbauten, physische Sammlungen etc. gegeben. Die Originaldokumente aus der Vergangenheit sind oft nur noch für Geschichtswissenschaftler lesbar und von Interesse. In anderen Bereichen der Wissenschaften, Kultur, Bildung und bei Behörden sind aber natürlich-sprachliche Originaltexte, ja sogar die physischen Originaldokumente vor ihrer Digitalisierung primäre Erkenntnisquellen. Hier bezieht sich die Übertragung dann auf die digitalen erläuternden Zusatztexte (Interpretationen, Analysen, Quellenstudien etc.).

7. Angesichts der schier unendlichen Zahl aller denkbar archivierbaren digitalen Quellen insgesamt<sup>3</sup> und deren immer rascherer Steigerung und Komplexität insbesondere der originär digitalen Quellen wird eine vollständige Langzeit-Archivierung unmöglich bleiben.

Eine nationale LZA-Policy hat daher als eine ihrer Kernaufgaben, Regeln für die faire und sachkompetente **Auswahl der Quellen** zur Aufnahme in eine Langzeit-Archivierung zu setzen. Diese sollte dabei allein nach ihrem Inhalt und ihrem vermuteten Impact (intellektuellen und wissenschaftlichen Einfluss) auf zukünftige Nutzer geschehen, unabhängig vom Objekttyp, von Vertriebs- und Publikationswegen, Verlagen, Autorenkategorien oder Dokumententypen.

Entsprechend sind mittels einer nationalen LZA-Policy Organisations- und Expertengremien zu schaffen, die die Auswahl-Prinzipien (Definition von Klassen von Inhalten) bzw. -Kriterien (Qualitätsprüfungen der Inhalte) bei der Auswahl der Dokumente für eine Langzeit-Archivierung festlegen. Für einen gewissen Bereich der klassischen Archive, insbesondere solcher, die Dokumente von Behörden, Ämtern und staatlichen Instanzen archivieren, ist die Auswahl der zu archivie-

<sup>3</sup>Allein die Zahl der als digitalisierungswürdig von der EU angesehenen gedruckten Werke in Bibliotheken in Europa wird von ihr auf ca. 2 Milliarden geschätzt.

renden Dokumente bereits durch die Archivgesetze bzw. behördliche Regelungen festgelegt. Diese sind dann lediglich auf digitale Dokumente und Digitalisate geeignet zu übertragen und zu erweitern.

8. Die **offenen Quellen** des weltweiten Internet (Open Access) und die individuellen, institutionellen oder zentralen *Open Archive Data Provider*<sup>4</sup> enthalten einen ständig steigenden Anteil an den für die Zukunft wesentlichen Informationen<sup>5</sup>. Gemäß Punkt 7 sind sie daher „aus gleichem Recht“ wie traditionelle Verlags-Quellen für die Langzeit-Archivierung abzuschöpfen (harvesting), d. h. allein nach gemeinsamen Auswahl-Prinzipien und -Kriterien der Policy für die LZA auszuwählen. Dies wird in einigen anderen Ländern bereits so angegangen.
9. Die besondere föderale Struktur der Bundesrepublik sowie das breite vielfältige Netz der beteiligten Archivierungsinstitutionen verlangt nach einem umso strafferen zielgenauen eigenen **LZA-Gesetz**, um einen stabilen, langfristig sicheren, von aktuellen politischen und wirtschaftlichen Kräften möglichst abgekoppelten Rahmen zu setzen.

Außerdem muss das bestehende Urheberrechts-Gesetz so weiterentwickelt werden, dass es den Anforderungen der Langzeit-Archivierung gerecht wird. Hierbei wird sich das demokratische und liberale nationale Selbstverständnis in entsprechenden offenen Regelungen zum Kopierschutz, zum Ausmaß des zukünftigen Zugangs und zum Plagiatschutz äußern müssen. Weltweit wird in vielen Industrieländern die Langzeit-Archivierung als staatliche Aufgabe begriffen, werden offene Formate und freier Zugang gefordert.

10. Die zunehmende Umstellung der (akademischen) Lehre auf digitale multimediale Dokumente eröffnet ein weiteres Handlungsfeld für eine nationale LZA-Policy. **Lehrmaterialien** werden durch die Digitalisierung und ihre Verbreitung über das Netz leichter einer Begutachtung und Bewertung durch die Fachöffentlichkeit zugänglich. Ihr Stellenwert beim Wettbewerb unter den Hochschulen und bei der Karriereplanung der Wissenschaftler wird entsprechend zunehmen<sup>6</sup>. Der schiere Umfang und der multimediale Komplexitätsgrad der digitalen Lehrmaterialien werden nach Schätzungen<sup>7</sup> den der wissenschaftlichen Dokumente binnen kurzem um ein Vielfaches übertreffen.

<sup>4</sup>OA*i Data Provider* sind Online-Archive frei zugänglicher Quellen, deren Nachweis über eine international verabredete Schnittstelle erfolgt; <http://www.openarchives.org>

<sup>5</sup>Hierzu ist auch die ständig steigende Zahl der OA Zeitschriften zu rechnen; [http://rzblx1.uni-regensburg.de/ezeit/fl.phtml?colors=7&lang=de&selected\\_colors%5B%5D=1&bibid=BISOL](http://rzblx1.uni-regensburg.de/ezeit/fl.phtml?colors=7&lang=de&selected_colors%5B%5D=1&bibid=BISOL). Hinzu kommt, dass die kommerziellen Verlage derzeit ihre Geschäftsmodelle von dem auf die Papier-Ära ausgerichteten Schema auf neue die Möglichkeiten der digitalen Welt nutzende Dienste umstellen (siehe Teil II).

<sup>6</sup>Eberhard R. Hilf; *Offene Kommunikation für die Leistungsfähigkeit der Wissenschaft und Hochschulen*; DINI Symposium Wissenschaftliches Publizieren der Zukunft - Open Access; <http://www.isn-oldenburg.de/~hilf/vortraege/dini-oa05/> und [http://www.dini.de/veranstaltung/workshop/goettingen\\_2005-05-23/](http://www.dini.de/veranstaltung/workshop/goettingen_2005-05-23/)

<sup>7</sup>Auf dem DINI-Symposium, a. a. O., wurde der Faktor 1.000 geschätzt. Es gibt viel mehr Lehrende als aktive Forscher, es gibt viel mehr Lehrveranstaltungen als Arbeitsgruppen, es gibt viel mehr Lehreinheiten als Veranstaltungen.

Neben den Kriterien und Anforderungen an eine Langzeit-Archivierung für Lehrmaterialien (Module), wie sie sich aus denen der wissenschaftlichen Dokumente nach entsprechender Anpassung ableiten, gibt es spezifische Besonderheiten: Die große Zahl von weltweit möglicherweise bald vorhandenen Lehrmaterialien macht eine sehr viel spezifischere Auswahl für eine nationale Langzeit-Archivierung erforderlich.

Kriterien könnten sein:

- Module, die eigenständige fachspezifische herausragende Leistungen sind oder als solche eingeschätzt werden, die also nicht leicht durch andere Module surrogierbar sind;
- Module, die als „best practice“ für Standard-Lehrstoffe gelten<sup>8</sup>.

Beide Kriterien verlangen strenge fach- und anwendungsspezifische Auswahlverfahren, um den zukünftig erwarteten Nutzen ihrer Langzeit-Archivierung abzuschätzen.

Abgesehen von diesen Sonderfällen einer nationalen Langzeit-Archivierung von Lehr-Modulen ist die Langzeit-Archivierung von Lernmaterialien aber eine Aufgabe der Universitäten als Teil ihrer Wettbewerbsstrategie (sogenannte *institutionelle Langzeit-Archivierung*).

Eine weitere Besonderheit sind die sehr komplexen Rechtsfragen bei Lernmaterialien, die einer Antwort in einer nationalen LZA-Policy bedürfen, allein schon als Rechtsrahmen für den universitären Wettbewerb. So gilt es, die Rechte der beteiligten Gruppen (Autoren, Erzeuger, Regisseure, Betreiber, Dozenten, Tutoren, Studenten) unter dem Primat einer langfristigen Unterstützung und Qualität der akademischen Lehre einzugrenzen. Im Übrigen gelten die Anforderungen an eine nationale LZA-Policy von Dokumenten sinngemäß auch für Lehrmaterialien.

## Vorschläge für erste Maßnahmen

Zu den ersten Maßnahmen für die benannten Handlungsfelder einer nationalen LZA-Policy könnten gehören:

1. **Diskussion:** Um die volle Komplexität, Vielfalt und Breite der Anforderungen von Erzeugern digitaler Dokumente, von Archivierungsinstitutionen und von Lesergruppen an eine nationale LZA-Policy zu erfassen, bedarf es der **Diskussion mit Experten aus allen Gebieten der Wissenschaft, Bildung, Kultur und von Behörden**. So ist der Stand der möglichen semantischen Kodierung wissenschaftlicher Inhalte in verschiedenen Fächern sehr verschieden weit gediehen, muss aber in seiner weiteren Entwicklung frühzeitig antizipiert werden.

<sup>8</sup>Als Beispiel seien die Feynman Lectures genannt, Lehrbücher des Nobelpreisträgers Feynman zur Physik für die Anfangssemester, die weltweit als absoluter Kanon gelten: R. P. Feynman; *Feynman Lectures on Physics*; <http://www.amazon.com/gp/product/0201021153/104-0764664-9706313?v=glance&n=283155&n=507846&s=books&v=glance>

2. **Sicherheit:** Mit den mit Langzeit-Archivierung befassten Institutionen anderer Länder, insbesondere in USA, Australien, Frankreich, den Niederlanden und Großbritannien, sollten Verhandlungen aufgenommen werden über gemeinsame rechtliche, organisatorische und technische Rahmenbedingungen für eine gegenseitige Spiegelung vorhandenen LZA-Materials<sup>9</sup>.
3. **Zugriff:** Um die sich aus der nationalen LZA-Policy ergebenden Mindestanforderungen an digitale Dienste für die Langzeit-Archivierung festzulegen, wird es nationaler fach- und themenspezifisch besetzter Gremien bedürfen. Technisch notwendige Mehrwertleistungen sind z. B. Retrieval-Dienste, Dienste-Vernetzung, Mehrwertdienste wie Indexierung, dynamische Fachthesauri etc. Sie sollen den bequemen und inhaltlichen zukünftigen Zugriff auf langzeit-archivierte Inhalte sichern helfen.
4. **Lesbarkeit:** Um den Inhalt des für lange Zeit zu speichernden Wissens der Gegenwart und Vergangenheit in zukünftigen Zeiten sicher lesbar zu erhalten, haben sich technische Eingangs-Filter bewährt, die bereits bei der Aufnahme von Dokumenten in die Langzeit-Archivierung nur nichtproprietäre Formate zulassen bzw. nur solche, welche die vollständige Software zur Dekodierung mit archivieren<sup>10</sup>. Bei den für lange Zeit zu archivierenden Dokumenten, die diese Kriterien nicht erfüllen, und bei denen die Autoren bzw. Ersteller diese nicht mehr nachliefern können, wird schon bei der Aufnahme in das Archiv eine Konvertierung (Normalisierung) stattfinden müssen unter Benennung und Inkaufnahme der dabei unvermeidlichen Informationsverluste.

Hilfreich wäre dazu, wenn die nationale LZA-Policy die Verwendung offener Formate bereits bei der Erstellung von Dokumenten effektiv und organisatorisch abgesichert fördern und bekannt machen würde. Für die derzeit gängigen digitalen Formate sollten umfangreiche Tests der Konvertierung vorgenommen werden, deren Ergebnis auch der Bewertung durch die Autoren bedarf.

5. **Volle wissenschaftliche Information:** Die nationale LZA-Policy sollte von Autoren und Dokumenten-Erzeugern, die mit öffentlichen Mitteln finanziert werden, die Einhaltung gewisser Mindeststandards der Vollständigkeit der Information zum Zwecke der Nachprüfbarkeit, des Verstehen-Könnens, der Nachnutzung der Dokumentation und der Einbettung in vorhandenes Wissen (Verlinkung, Zitate) einfordern. Dies wird dann die Wissenschaftsorganisationen und Förderinstitutionen sowie die entsprechenden Landes- und Bundesministerien, die Universitäten und Archivierungsinstitutionen darin stärken, diese durchzusetzen. Die „LZA-Konditionierung an der Quelle/beim Erzeuger“ soll eine leis-

<sup>9</sup>Innerhalb Europas bietet sich hier das neue europäische digitale Archiv CCSDS (*Centre pour la Communication Scientifique Directe*) der nationalen französischen Forschungsorganisation CNRS als Mittler an; <http://ccsd.cnrs.fr/accueil.php3?lang=de>

<sup>10</sup>Zum Beispiel akzeptiert das 1991 gegründete eArchiv *ArXiv* <http://www.arxiv.org> nur nichtproprietäre Formate wie  $\text{\LaTeX}$ , *html* etc. und auch nur, wenn alle verwendeten Dienstprogramme ('Macros', Style sheets, Editierpakete etc.), die zum Lesen erforderlich sind, vom Autor mitgeliefert werden.

tungsfähige und sichere Langzeit-Archivierung dauerhaft sichern helfen bzw. wesentlich erleichtern<sup>11</sup>.

6. **Internationale Standards:** Eine entsprechend formulierte nationale LZA-Policy wird die archivierenden Institutionen darin stärken, bereits bei der Aufnahme von Dokumenten international vereinbarte Metadaten einzusetzen bzw. sie mit den einliefernden Institutionen zu vereinbaren und von den Autoren zu verlangen<sup>12</sup>.

Die zusätzlichen und qualitativ neuartigen und höherwertigen Anforderungen an die LZA-Institutionen müssten durch dafür zweckbestimmte Mittel zunächst durch das *Bundesministerium für Forschung und Technologie* BMBF, sodann – den gesetzlich zu treffenden Regelungen des Bundes und der Länder zur Langzeit-Archivierung entsprechend – durch verlässlich festzuschreibende regelmäßige Bundes- bzw. Landesmittel finanziert werden.

Die LZA-Institutionen und die entsprechend zu bildenden Gremien werden sich aktiv in den *internationalen Prozess der Verabredung von Standards* einbringen. Hierzu muss die nationale LZA-Policy einen verlässlichen Rahmen<sup>13</sup> setzen mit den Schwerpunkten:

- (a) Bildung eines nationalen Gremiums zur Koordinierung der Aktivitäten deutscher Experten in internationalen Gremien;
- (b) Förderung der Ausbildung und Schulung international tätiger Experten;
- (c) Koordinierte Information und Schulung der LZA-Institutionen;
- (d) Förderung der aktiven Teilnahme deutscher Experten in internationalen Gremien.

7. **Kontexterhaltung:** Verstehbarkeit und Nachnutzbarkeit setzt die Erhaltung des semantisch kontextuellen Rahmens voraus. Insbesondere implizite Bezüge sind sonst nicht mehr verstehbar oder könnten sogar falsch interpretiert werden<sup>14</sup>. Entsprechend ist der assoziierte Kontext zu modularisieren und den jeweiligen

<sup>11</sup>Beispiele sind die Beigabe von Computer-Codes bei theoretisch-physikalischen Arbeiten oder genaue Angaben über ein Experiment-Setup bei physikalisch-experimentellen Arbeiten.

<sup>12</sup>Als Vorbild mögen der Workflow und die Vereinbarungen zu Metadaten zwischen den Institutionen des Projektes *Dissertationen Online* <http://www.dissonline.de> gelten, die zwischen der *DublinCore Initiative*, *Die Deutsche Bibliothek*, Universitätsbibliotheken und vier Instituten, jeweils als Vertretern ihrer großen Fachgesellschaften abgeschlossen wurden. Das DissOnline-Projekt entstand bereits 1996 auf Initiative des entsprechenden Arbeitskreises der *Initiative Information und Kommunikation der wissenschaftlichen Fachgesellschaften in Deutschland* (IuK) <http://www.iuk-initiative.org> als ein früher und erfolgreicher Schritt zur Langzeit-Archivierung wissenschaftlicher Materialien; <http://elfikom.physik.uni-oldenburg.de/dissonline/PhysDis/>

<sup>13</sup>Leider hat 2003 der Projektträger des BMBF einen entsprechenden bundesweit koordinierten Antrag auf Förderung nicht einmal entgegengenommen.

<sup>14</sup>Beispielsweise werden mathematische Formeln ohne den Kontext der jeweils geltenden Axiome und Darstellungskonventionen zu einer unverständlichen eihung von Glyphen.

Informationen fest zuzuordnen. Die entsprechende technische Entwicklung maschinenlesbarer Kodierungen mit der dazu notwendigen internationalen Einbettung sowie die Verbreitung der Lösungen bei den Autoren, Nutzern und den Retrieval-Diensten sind enorme Herausforderungen. Dies erfordert in Deutschland eine neue energische und auf viele Fächer verteilte koordinierte Bundesförderung und ein entsprechendes Engagement der Wissenschaften. Es droht sonst die Abkoppelung von der derzeit hochdynamischen internationalen Forschung und Entwicklung.

8. **Verstehbarkeit wissenschaftlicher Dokumente:** Der stete Transfer des Wissens über die Originale in aktuellere Darstellungen ist ebenso notwendig wie die 'Kondensation' von Wissen (Reviewing) zu synoptischen Zusammenstellungen, um zukünftig Arbeiten verstehen und nutzen zu können. Dies gilt vor allem für Fachwissenschaften, in denen sich die Wissenschaftssprache und der Kontext oft besonders rasch ändern. In den Naturwissenschaften, aber auch in anderen Wissenschaftsgebieten, der Kunst bzw. bei Archivgut ist die Übersetzung der Kenntnisse über das Werk in aktuelle Sprache und Denkkontexte zum zukünftigen Verstehen oft bereits gängige Praxis (z. B. synoptische Darstellungen, Review-Artikel, Bücher). Nur so bleiben die langzeit-archivierten Originale (etwa mittelalterliche Handschriften, Kunstwerke, prähistorische Funde) durch das Wissen über sie lebendig und „verstehbar“.

Die *Organisation* kann fach- bzw. inhaltsspezifisch mittels Einbindung und Vernetzung der einschlägig arbeitenden wissenschaftlichen Institute realisiert werden durch

- (a) vertragliche Bindung geeigneter Autoren;
- (b) Beantragung von gemeinsamen Projekten mit Wissenschaftlergruppen zur experimentellen modellhaften Transformation/Kondensation von Wissen.

Die nationale LZA-Policy muss hier auch Überzeugungsarbeit anstoßen: Im Papierzeitalter lag der Schwerpunkt der Archivierung auf dem Sammeln der überschaubaren Zahl zugänglicher Dokumente meist sehr begrenzten Inhalts (z. B. bedingt durch die damals technisch notwendigen engen Umfangsbeschränkungen der wissenschaftlichen Zeitschriften). Daher enthalten die bisher noch sehr Papierorientierten nationalen Policies zur Archivierung Vorgaben wie: „Sammele im Sinne einer einfachen Ablage alles, was die vorgegebenen Kriterien erfüllt“. Im digitalen Zeitalter ist die Stoffmenge jedoch praktisch unbeschränkt. Eine reine Sammlung würde dem eigentlichen Auftrag zur Bewahrung des wissenschaftlichen und kulturellen Erbes für eine zukünftige Verstehbarkeit und Nutzung nicht gerecht, sie durch Mittel- und Kräfteentzug sogar behindern. Das zukünftig relevante Material würde im anschwellenden Meer des Irrelevanten zunehmend schwerer wieder auffindbar, weniger verstehbar, wäre schwieriger einzuordnen und unvollständiger erschließbar.

9. Die **Bewertung, Kondensierung, Einbettung, Filterung und Sprachmigration** bildet bei digitalen langzeit-archivierten Dokumenten einen neuen, zentralen und

wichtigen Fokus einer nationalen LZA-Policy, um die Verstehbarkeit für sehr lange Zeit zu sichern. Ein *nestor*-Arbeitskreis aus Experten verschiedener Fächer könnte an konkreten ausgewählten Beispielen aus der Vielfalt der Dokumente und Bereiche inklusive der Museen und klassischen Archive aufzeigen, welche Anforderungen sich aus dem jeweiligen Fachgebiet ergeben, welche Formen sinnvoll sind, und wie dann mit diesen Bündeln digitaler Dokumente (das Original mit seinen ergänzenden Dokumenten) in der Langzeit-Archivierung verfahren werden sollte (Speicherung und Retrieval). Der dann besser abschätzbare Aufwand findet über Willensaussagen in der nationalen LZA-Policy seine Entsprechung in staatlich einzusetzenden Mitteln. In diesem Arbeitskreis werden sich den Fachexperten z. B. aus der Wissenschaft ihnen ungewohnte Fragen stellen: welche der aktuell wichtigen Dokumente könnten etwa im Jahre 2030 als wesentlich für die aktuelle Arbeit angesehen werden? Ein erster Hinweis ergibt sich aus dem Blick von heute in die Vergangenheit: so werden wissenschaftliche Publikationen in vielen Gebieten schon nach wenigen Jahren zum Zwecke aktueller Forschung nicht mehr angefordert oder gelesen – wohl aber werden ihre intellektuellen Inhalte genutzt.

**Dokumentenformen:** Die nationale LZA-Policy sollte eine Erklärung enthalten, dass die zentralen staatlich finanzierten LZA-Institutionen jeweils die technischen Mittel bereitstellen, um möglichst alle aktuell in Bildung, Kultur und Wissenschaften eingesetzten Dokumentenformen und -typen lesen zu können, so dass sie gegebenenfalls bereits bei der Annahme in LZA-fähige Formate und Dokumenttypen projiziert werden können.

10. **Die Auswahl der Quellen** sollte in der Wissenschaft, Kultur und Bildung allein nach ihrem Gehalt für eine Nutzung in der fernen Zukunft geschehen, nicht nach der Dokumentenform, nach Veröffentlichungsweg oder Referierungsgrad. Die nationale LZA-Policy wird daher die Kriterien der Auswahl von Dokumenten für die Langzeit-Archivierung zusammen mit den anliefernden Institutionen und den Fach-Beratern als den Vertretern der Autoren regeln. Die anliefernden Institutionen müssen bei der Auswahl, Konditionierung und Archivierung aktiv mitarbeiten. Hierzu könnte ein kleiner Arbeitskreis konkrete Filter-Regeln ausarbeiten, die dann die entsprechenden Regelungen in einer nationalen LZA-Policy konkreter zu fassen gestatten würden.
11. Bei den **klassischen Archiven** liegt die Hoheit der Auswahl bei den die Archivierung beauftragenden Behörden, Institutionen und Ämtern und ist teilweise durch Gesetze und Regelungen festgelegt. Einer nationalen Policy bedarf es hier nur als übergeordnetem Rahmen, um den klassischen Archiven bezüglich ihres digitalen Materials bzw. ihrer Digitalisate, die ja u. A. völlig neue Möglichkeiten der vielfach redundanten Lagerung, Migration und Bearbeitung eröffnen, national einheitliche Rahmen-Regeln an die Hand zu geben, die technisch solche Möglichkeiten auch erst realisierbar machen (gemeinsame digitale Standards, Formate, Spiegelungen). Zugleich muss jeweils geregelt werden, in welcher Reihenfolge und Auswahl vorhandene Dokumente digitalisiert werden sollen.



Eine nationale Policy, die eine leistungsfähige Langzeit-Archivierung als Staatsinteresse formuliert, setzt dadurch den Exekutiven die Vorgabe, für die Bereitstellung ausreichend starken technischen Wissens und entsprechend ausgebildeten Personals in den LZA-Institutionen zu sorgen.

12. **Offene Quellen:** Offene digitale Quellen (*Open Access* OA), die also über das Web ohne finanzielle, rechtliche oder technische Barrieren vollständig zugänglich sind, nehmen rasch zu, weil sie die aktuelle Arbeit erleichtern. Sie sind instantan, weltweit, und für jeden zugänglich.

Konkrete Erfahrungen (siehe Teil II) wie etwa in Australien oder in Frankreich könnten gewonnen werden, wenn mit dem Aufbau eines zentralen Langzeit-Archivs begonnen würde, das die bereits vorhandenen OA-Archive in wissenschaftlichen Institutionen und Bibliotheken kopiert und zentral nachweist. Zugleich könnten dabei die in einer zukünftigen nationalen LZA-Policy geforderten Standards für Kennzeichnung und Retrieval bereits praktisch erprobt und den anliefernden Archiven nahegebracht werden<sup>15</sup>.

13. Die „Weiten des Web“ werden aber eine **aktive Suche nach relevanten Quellen** verlangen, die zweckmäßig durch Fachexperten ausgerichtet erfolgen müsste<sup>16</sup>.

Unterstützt werden kann dies durch fachverständlich automatisiertes *Harvesting* über das Internet. Die Fundstellen müssen dann aber den o. g. LZA-Filtern und Bewertungen zur Auswahl unterworfen werden<sup>17</sup>.

14. **Interaktion und LZA:** In der Zukunft wird digitale Information entsprechend den dann angepassten Arbeitsgewohnheiten immer stärker interaktiv genutzt werden. Dies hat zur Folge, dass ursprünglich statisch angelegte Langzeit-Archive als „Informations-Vorrathshäuser“ in der Zukunft zunehmend interaktiv genutzt und in den zukünftigen Kontext eingebunden werden wollen, was bedingt, dass die Erkenntnisse aus diesen Interaktionen fortlaufend ebenfalls mit in das gleiche Archiv gespeichert werden müssen, um professionell genutzt werden zu können und akzeptiert zu bleiben.

15. **Rechtsfragen zu verteilten Dokumenten:** Es sollte gezielt eine nichtstaatliche Arbeitsgruppe zu Rechtsfragen der Langzeit-Archivierung in Deutschland gegründet werden, der neben einschlägig ausgewiesenen Juristen auch Vertreter der LZA-Institutionen und der Fachwissenschaften angehören. Deren Aufgabe wird

---

<sup>15</sup>Als Beispiele mögen die Überlegungen im BMBF-geförderten Vorhaben *eSciDoc* für die Max-Planck Institute <http://www.escidoc-project.de/de/> sowie der Vorschlag eines zentralen *Open Access Initiative*-konformen Service-Providers von DINI (*Deutsche Initiative für NetzInformation*) <http://www.dini.de> gelten.

<sup>16</sup>Erste Tests ergaben, dass z.B. auf Arbeitsgruppen-Servern der Hochschulen mindestens zehnmals mehr wesentliches wissenschaftliches Material als OA angeboten wird, als bereits Eingang in das sich bildende weltweit verteilte Netz der nach einem technischen Standard der OAi (*Open Access Initiative*) <http://www.openarchives.org> arbeitenden registrierten *Data Provider* gefunden hat.

<sup>17</sup>Als Beispiele seien das Crawler-Filter-System <http://www.ecs.soton.ac.uk/~harnad/Temp/impact.html> der Gruppe um Stevan Harnad oder die neuen intelligenten Dokumenten-Crawler in <http://www.physnet.net> genannt.

es sein, die Anforderungen an ein *nationales LZA-Gesetz* auszuarbeiten und den Gesetzgebungsprozess aktiv zu begleiten<sup>18</sup>.

Als Kernpunkte eines nationalen LZA-Gesetzes bieten sich an:

- (a) Forderung an die Autoren/Erzeuger/Anlieferer von Dokumenten, gewisse Mindeststandards für eine leichtere spätere Langzeit-Archivierung einzuhalten, wie offene Formate, Vollständigkeit, Dokumentation, Verlinkung und internationale Standards (so genannte „LZA-Konditionierung an der Quelle“);
- (b) Schranken für die Rechte der Urheber und der (oft kommerziellen) Rechteinhaber sowie der staatlichen Vertraulichkeitsauflagen, so dass die Anforderungen künftiger Generationen an Zugang, Lesbarkeit und Verstehbarkeit des kulturellen und wissenschaftlichen Erbes erfüllbar werden;
- (c) Definition der Langzeit-Archivierung als staatliche Aufgabe;
- (d) Regelung der Teilung von Arbeit und Zuständigkeiten mit den Ländern;
- (e) Einhaltung der Mitwirkung an internationalen Standards;
- (f) Langfristige Sicherung der Langzeit-Archivierung durch Nachweis der Redundanz und der internationalen Spiegelung;
- (g) Regelung des offenen Zugangs, der Bewahrung des Urheberrechts und der Regelungen bei Nutzergruppen mit unterschiedlichen Rechten;
- (h) Regelung der besonderen rechtlichen Fragen bei multimedialen und verteilten Dokumenten mit Beiträgen mehrerer Autoren und bei der Beteiligung verschiedenartiger Personengruppen (Autoren, Erzeuger, Dozenten, Hersteller, Drehbuchschreiber, Produzenten, Darsteller etc.).

Die aktive Einflussnahme der von einem Gesetz Betroffenen auf den Gesetzgebungsprozess wird in Deutschland vom Bundesministerium der Justiz durch einen regelten formal installierten Prozess von Anhörungen aller Beteiligten gefördert<sup>19</sup>.

16. **Die Europäische Union will eine eigene europäische Langzeitarchivierungs-Policy entwickeln, die dann nachfolgend in nationale Policies eingearbeitet werden muss.** Es gilt daher, hier von Anfang an sehr intensiv und aktiv mitzuarbeiten und auf allen Ebenen Experten zu entsenden, um nationale Erfahrungen und Expertise einzubringen und um die nationalen Interessen und besonderen Anforderungen zu vertreten und durchzusetzen.

<sup>18</sup>Als Vorbild mag die Formulierung eines eigenen §52a des Urheberrechtsgesetzes für den Zweck der Nutzung von Dokumenten in Bildung und Wissenschaft gelten, das dem Urheberrecht Schranken nach dem Zweck auferlegt, wie sie 2004 die *Arbeitsgruppe Urheberrecht der Deutschen Initiative für NetzwerkInformation e.V.* <http://www.dini.de/documents/DINI-UrhG-K2-121104-CC.pdf> vorgelegt hat.

<sup>19</sup> Als Beispiel mag der derzeitige Prozess der Novellierung des Urheberrechtsgesetzes gelten, siehe die Dokumente des *Aktionsbündnis Reform des Urheberrechts für Bildung und Wissenschaft*; <http://www.urheberrechtsbuendnis.de>

Die **Formulierung der Anforderungen** an die nationale LZA-Policy könnte durch einen *Nationalen LZA-Rat* geschehen, dem Vertreter der LZA-Institutionen und aus Wissenschaft, Kultur und Bildung angehören. Dieser sollte von der Regierung (BMBF) einberufen und benannt werden<sup>20</sup>.

## 2.2 Vorschläge für strukturelle Maßnahmen

Es werden hier einige Vorschläge für strukturelle Maßnahmen aus den Materialien des Teils II dieser Expertise extrahiert und vorgezogen skizziert:

1. Wie bereits in anderen Ländern mit föderativer Struktur (z. B. USA, Australien) bewährt, wird sich die nationale Langzeit-Archivierung am besten durch eine neue *verteilte, aber integrative Organisations-Architektur* realisieren lassen: *verteilt*, um die Langzeit-Archivierung den rasch anschwellenden Dokumentenmengen gegenüber skalierbar und akzeptierbar machen; *integrativ*, um den steigenden professionellen Anforderungen durch Bündelung der Expertisen und Interaktion der Experten gerecht werden zu können. Entsprechend wird sich dies in einem **verteilten, arbeitsteiligen aber integrativen Zusammenwirken** von Landes- und Bundesinstitutionen mit integrativen Leitungsstrukturen widerspiegeln.

2. Der für eine breite **Akzeptanz** notwendige langfristig tragfähige und belastbare nationale Konsens erfordert eine entsprechend angelegte Diskussion aller zu beteiligenden Experten und Fachleute.

Deren *Organisation* wird zweckmäßig von *nestor* geführt, organisiert und von Arbeitsgruppen zu den Einzelfragen, zu denen jeweils externe entsprechende Fachexperten hinzugezogen werden, ausgearbeitet werden. So könnte die enorme Breite und Vielfalt der Thematik, der Fächer, der Anwendungen abgedeckt werden. Zu ergänzen wäre dies durch direkte Kontakte zu den entsprechenden Fachgesellschaften und Verbänden, um Filterregelungen zu konzipieren und zu erproben.

Die genauen technischen, finanziellen und organisatorischen Strukturen werden dann danach von den zu beteiligenden Institutionen **gemeinsam erarbeitet** und politisch durchgesetzt werden, z. B. durch *Die Deutsche Bibliothek*, die Sondersammelgebiet-Bibliotheken, Landesarchive und -bibliotheken, Museen, Rechenzentren, Multimediazentren, Messdaten-Archive, Rundfunkanstalten, Informationswissenschaftler, Wissenschaftsvertretungen etc.

3. Ergebnis einer nationalen LZA-Policy werden schließlich Verträge zwischen Partnern sein, welche die **föderale Struktur der Bundesrepublik** und die Forderungen der zahlreichen beteiligten und zu beteiligenden Ministerien des Bundes und der Länder abbilden. Die zukünftigen Nutzer einer Langzeit-Archivierung

---

<sup>20</sup>Für das Teilgebiet der Digitalisate gibt es etwas entsprechendes bereits: EUBAM (*Portal zu europäischen Angelegenheiten für Bibliotheken, Archive, Museen und Denkmalpflege*); <http://www.eubam.de>

werden dabei durch die Wissenschafts- und Kultus-Ministerien vertreten. Die rechtlichen Aspekte des Zugriffs, des Urheberrechts und von Verwertungsrechten wird von den Justizministerien wahrgenommen. Die Auflagen zum Schutz staatlicher Daten erfordern die Beteiligung der Innenministerien und bei anderen fachlichen Aspekten die weiterer Fachministerien.

Wegen dieser komplexen und sich möglicherweise auch zukünftig noch verschiebenden Zuständigkeitsverteilungen, für eine größere Akzeptanz und verbesserte Durchsetzbarkeit sowie auch zur dauerhaften Finanzierung wird daher als stabile Richtschnur für alle Beteiligten für ein **durchsetzungsstarkes LZA-Gesetz** des Bundes plädiert.

4. Eine **enge Ankopplung an die nationale und internationale Forschung** und Entwicklung neuer professioneller Dienste für die Langzeit-Archivierung wird in der sich auch technisch rasch weiterentwickelnden Welt mit ihren laufend neu aufkommenden Formaten, Dokumententypen sowie an Informations- und Kommunikationsgewohnheiten der Nutzer das notwendige Innovationstempo sicherstellen.
5. Eine weitere Herausforderung ist die Einbindung in die (zukünftig zunehmend integrierten) **internationalen vernetzten LZA-Organisationsstrukturen**.  
Nicht nur die UNESCO sondern vor allem die Europäische Union planen LZA-Policies, denen die nationale Policy und die nationalen Dienste entsprechen müssen. Es wäre ein entscheidender Vorteil, wenn in die dortigen Vorbereitungen, Projekte und Gremien Vertreter aus Deutschland entsandt und diese für ihre Arbeit energisch unterstützt würden.
6. Die Langzeit-Archivierung digitaler Inhalte eröffnet für das Netz der archivierenden Institutionen **neue Aufgabenfelder** und verlangt Expertise aus weiteren Bereichen. Dies muss sich in entsprechend ergänzender Finanzierung der Institutionen und in einer breit angelegten Förderung der entsprechenden Ausbildungslinien niederschlagen.
7. Für **die Finanzierung** der Aufgaben zur Langzeit-Archivierung werden Bund und Länder zweckmäßig gemeinsam einen Finanzplan aufstellen, der den Anforderungen der nationalen LZA-Policy gerecht wird. Es wird sich erweisen, dass es auch wesentlicher zusätzlicher finanzieller Anstrengungen bedürfen wird, damit Deutschland zukünftig auch in dem sich bildenden internationalen LZA-Netzwerk aktiv und angemessen beteiligt sein kann.
8. Die für eine nationale Langzeit-Archivierung notwendigen **technischen Dienste** verlangen eine professionelle Realisierung. Unter dem Rahmen einer nationalen LZA-Policy und entsprechender institutioneller Policies und der Anforderungen der Archivierungsinstitutionen, der Erzeuger- und Nutzer-Vertretungen können die notwendigen Dienstleistungen und technischen Dienste an kommerzielle Firmen vergeben werden. Insoweit die Rahmenbedingungen Wettbewerb

ermöglichen, wird dieser für eine rasche Anpassung an die jeweils aktuellen technischen Standards zu fairen Preisen sorgen. Hier könnten Erfahrungen gewonnen werden durch eine Kommission, die für eine Ausschreibung geeignet erscheinende begrenzte Aufgabenfelder definiert, Angebote kommerzieller Firmen prüft, den entsprechenden Dienst befristet akzeptiert, und die Qualität der Leistung fortlaufend evaluiert.

9. Publierte Dokumente (z. B. Artikel in Journalen) kommerzieller Verlage sollten den gleichen LZA-Aufnahmeregeln unterworfen werden wie alle anderen Dokumente. Im Allgemeinen haben die Verlage bereits Verträge mit nationalen Bibliotheken zur Archivierung ihrer gesamten Publikationen. Für die Aufnahme in eine nationale Langzeit-Archivierung müssen diese Verträge entsprechend erweitert werden (Auswahlfilter, offene Formate, Schlüsselfreigabe, Sperrfrist).
10. Die politische Durchsetzung einer koordinierten, zukunftssicheren und für die Bewahrung des kulturellen und wissenschaftlichen Erbes angemessenen nationalen Langzeit-Archivierung durch die politischen Entscheidungsträger in Bund und Ländern verlangt auch eine entsprechende **Neuordnung der nationalen Wissenschaftspolitik**: hin zu einer vorausschauenden, langfristigen Planung, der breiten und qualitätsfordernden Förderung der beteiligten Experten und der kulturellen und wissenschaftlichen Institutionen, hin zu einer engagierten Einflussnahme auf den politischen Prozess der Gesetzgebung und hin zu einer straffen Durchsetzung und Kontrolle der in einer nationalen Policy festgeschriebenen Rahmenbedingungen.
11. In die nationale LZA-Policy gehört ein Passus, der das **Recht der Autoren** in Bildung, Wissenschaft und Kultur auf eine staatliche Langzeit-Archivierung ihrer digitalen Werke festhält, sofern sie die allgemeinen LZA-Filter passieren (Gleichheit der Autoren vor dem Recht, Unabhängigkeit der Langzeit-Archivierung von kommerziellen Interessen, Priorität des Staatsinteresses).

### Vorschläge für erste organisatorische Maßnahmen

1. **Vorbereitung des politischen Prozesses hin zu einer nationalen LZA-Policy:** Die Initiative *nestor*<sup>21</sup> sollte einen Facharbeitskreis *Nationale LZA-Policy* bilden, dem Vertreter der LZA-Institutionen, der Wissenschaftsvertretungen, der Wissenschaft, der KMK (*Kultusminister-Konferenz*) sowie einschlägig kompetente Rechtsvertreter angehören. Dieser sollte den politischen Prozess zu einer nationalen LZA-Policy und einem diesbezüglichen nationalen LZA-Gesetz durch geeignete Vorlagen und Strukturvorschläge, durch Pressearbeit und durch die Initiierung von Vorlagen zu Teilthemen voranbringen und sich hierzu Arbeitsgruppen bedienen, die mit den entsprechenden externen Fachleuten besetzt werden könnten.

---

<sup>21</sup><http://www.langzeitarchivierung.de>

2. **Kommunikationsprozess zwischen den mit der Durchführung von Langzeit-Archivierung Befassten:** Die Kommunikation unter den mit LZA zukünftig befassten Institutionen könnte entscheidend vorangetrieben werden durch eine neue Stufe der Intensität (fach-)öffentlicher Aktivitäten, durch themenspezifische Arbeitskreise, durch die Erarbeitung von Empfehlungen, technischen Leitfäden sowie durch Trainingscamps und Schulungen. Notwendig ist dabei die energische Verbreiterung der Basis, d. h. die aktive Einbeziehung aller an Langzeit-Archivierung interessierten und mit ihr befassten Institutionen und Personen. Nur auf der so anzustrebenden breiten Basis von Kenntnis und Beteiligung lässt sich die notwendige Akzeptanz der Beteiligten erreichen.
3. **Ankopplung an den Wissenschafts- Kultur- und Bildungsbereich:** Die enge Verzahnung mit diesen Bereichen ermöglicht den frühen Blick auf zukünftige technische Entwicklungen auch für die Realisierung einer Langzeit-Archivierung. Sie erleichtert, Anforderungen aufzunehmen, sie bereitet die Akzeptanz zukünftiger Dienste vor. Dies könnte durch eine Vielzahl von Einzelaktivitäten unterstützt werden, wie:
  - Start einer Workshop-Reihe;
  - Einberufung einer Expertengruppe;
  - Förderung der Aktivitäten der Experten auf dem Gebiet der LZA;
  - Exemplarische Erprobung mit Experten jedes Faches zu Verfahren zum fachkompetenten Auswählen, Filtern, Konditionieren, Migrieren, Archivieren, Kondensieren, Verlinken, Auffinden und Einsammeln von Dokumenten für eine LZA.
4. Die **Einbettung in das Netz der Arbeiten zur nationalen Langzeit-Archivierung anderer Länder** kann ausgehen von der Koordinierung und dem Austausch von Erfahrungen der Experten aus Deutschland, die jetzt schon mit LZA-Themen befasst und in entsprechenden internationalen Gremien tätig sind. Praktischer Ausgangspunkt werden die Kontakte zu den jeweiligen nationalen Bibliotheken (siehe Anhang) und ihren Experten und Projekten sein. Die Projekte weiterer internationaler Gremien und Projekte wie der EU und der UNESCO sollten (z. B. die Verweise in Teil II nutzend) systematisch analysiert werden und der Kontakt zu ihnen zum Nutzen für die Entwicklung einer nationalen Langzeit-Archivierung aufgenommen werden. Aktuell sind hier die Projekte des FP 6 aber auch die Planungen zum FP 7 zu nennen. International kommt darüberhinaus den Industriestaaten eine Aufgabe als 'Paten-Nation' für kleine und für Entwicklungsländer zu, Expertenhilfe zu leisten, LZA-Funktionen zu übernehmen, Spiegelung von Materialien anzubieten und technische Hilfen anzubieten.
5. **Die neuen Aufgabenfelder für LZA-Institutionen** und die Berufsqualifikationen für in der Langzeit-Archivierung Beschäftigte erfordern eine Definition der notwendigen Berufsqualifikationen für das professionelle Arbeiten in der Langzeit-Archivierung. Darauf folgen Überlegungen und Anstrengungen, die vorhande-

nen Curricula, z. B. der Ausbildungen z. B. zum Archiv-Bibliothekar, entsprechend zu erweitern bzw. entsprechende Ausbildungslinien (etwa zum „Informations-Ingenieur“) zu schaffen, vielleicht als einem standardisierten Studiengang der Informations- und Kommunikationswissenschaft; dieser sollte dann Pflichtanteile zu den besonderen Themen der Langzeit-Archivierung enthalten. Die wissenschaftliche und kulturelle Information und Kommunikation mit den zukünftigen Generationen verlangt zu dessen professioneller Realisierung beruflich qualifizierte.

6. **Die Finanzierung** der Langzeit-Archivierung erfordert zunächst eine umfassende Analyse der jetzigen Ausgaben für Bibliotheks- und Archivierungsdienste. Aus dem Vergleich mit z. T. andersartigen und neuen Anforderungen einer digitalen Bibliothek werden Vorschläge für notwendige Umschichtungen folgen<sup>22</sup>. So könnten Maßzahlen gewonnen werden. Für die exakten Wissenschaften mit ihren Anforderungen an eine möglichst umfassenden Literaturversorgung aus allen existierenden relevanten Zeitschriften könnten diese z. B. sein: Aufgewendete Mittel im Verhältnis zur Bereitstellung bzw. Beschaffung von angeforderten Artikeln und Dokumenten aus dem gesamten weltweiten Dokumentenraum (nicht nur aus dem der lokal vorgehaltenen wissenschaftlichen Zeitschriften).

Für diese Fächer ist das weltweit „Gelesen werden können“ ein Teil des Berufszieles. Entsprechend kommt in der digitalen Welt den Bibliotheken als neue Rolle die Unterstützung der internationalen Verbreitung lokal erzeugter Informationen zu. Über die Festlegung geeigneter Maßzahlen für den entsprechenden Impact als Leistungsbewertung wird zwar noch geforscht<sup>23</sup>, aber die Zahl der Zitierungen lokal erzeugter Arbeiten je Wissenschaftler bezogen auf den finanziellen Aufwand für unterstützende Leistungen der Bibliothek (Web-posting mit Metadaten, Aufnahme in das digitale Universitätsarchiv und die Langzeit-Archivierung) kann ein pragmatischer Ansatz sein.

Die dadurch an vielen Orten lokal freigesetzten Mittel<sup>24</sup> könnten die Finanzierung für die Teilnahme der lokalen Institutionen an der Wertschöpfungskette bis hin zur Langzeit-Archivierung wesentlich unterstützen.

Schließlich müssen zukünftige finanzielle Bedürfnisse eines nationalen LZA-Netzwerkes geschätzt werden und diese dann durch Bund-Länder-Vereinbarungen für lange Zeit (über eine Legislaturperiode hinaus) gesichert und gesetzlich verankert werden.

---

<sup>22</sup>Robert Terry; *Publication costs are research costs*; Followup Conference on OpenAccess; CERN 2004; <http://www.zim.mpg.de/openaccess-cern/presentation-terry.ppt>

<sup>23</sup>Das Institut für Forschungsinformation und Qualitätssicherung (IFQ) der DFG (Deutsche Forschungsgemeinschaft); <http://www.dfg.de/ranking/> Siehe auch den Vortrag von J. Güdler; [http://www.iuk-initiative.org/index.php?option=com\\_content&task=view&id=37&Itemid=72](http://www.iuk-initiative.org/index.php?option=com_content&task=view&id=37&Itemid=72) auf der Jahrestagung 2005 der IuK Initiative; <http://www.iuk-initiative.org/pdf/IuK-2005-Guedler.pdf>

<sup>24</sup>Der Fachbereich Physik der Carl von Ossietzky Universität Oldenburg hat am 1.1.2003 einen wesentlichen Teil der aus Abbestellungen kommerzieller Zeitschriften gewonnenen Mittel für die Verbesserung der Online-Informationssuche und -versorgung sowie der besseren Web-Sichtbarkeit der lokal erzeugten wissenschaftlichen Dokumente umgeschichtet.

7. In einer kleinen technischen Arbeitsgruppe von *nestor* könnten Anforderungen an eine nationale LZA-Policy erarbeitet werden, die den **Markt der Dienstleistungen** zur Realisierung der verteilten Langzeit-Archivierung definieren und so das zeitbefristete Outsourcen von Leistungen an kommerzielle Firmen im gesicherten Wettbewerb erlauben und die entsprechende Leistungskontrolle sicherstellen.
8. **Zur politischen Durchsetzung** einer nationalen LZA-Policy könnte in einem ersten Schritt eine kleine „politische“ Task-Force von Experten aus den LZA-Institutionen, den Wissenschaftsorganisationen (HRK, WR), den Wissenschaften und dem Rechtswesen gebildet werden, die einen Entwurf einer LZA-Policy und ihre Projektion auf ein LZA-Gesetz in den politischen Prozess einbringt, diesen vorantreibt und die fachöffentliche Diskussion initiiert und begleitet.

Die Bildung eines staatlichen Komplementes hierzu wäre anzustoßen: Die Einrichtung eines LZA-Arbeitskreises, der die BLK (*Bund-Länder-Kommission*), die KMK (*Kultusminister-Konferenz*) und Vertreter aller einschlägigen Ministerien angehören.

Erfolgreiche Beispiele solcher Prozesse, wenn auch in kleinerem Maßstab, sind *Dissertationen online*, die Open-Access-Verbreitung durch die *Deutsche Initiative für NetzwerkInformation (DINI)*, das *Aktionsbündnis Urheberrechtsreform für Bildung und Wissenschaft*, die *Initiative Information und Kommunikation der wissenschaftlichen Fachgesellschaften (IuK)*.

Parallel sollte das BMBF eine fachlich abgesicherte eigene interne zukunftsfähige Policy für die breite, aber qualitätssichernde Förderung der für eine Langzeit-Archivierung notwendigen Entwicklungen, Erprobungen und Installationen entwickeln lassen und sodann die für die Durchsetzung notwendigen organisatorischen und finanziellen Umschichtungen der Förderung veranlassen. Dabei sollten der Breite und der rein auf die Sache bezogenen Förderung und der externen Qualitätskontrolle der Fördermittel-Ausgaben großer Wert zugemessen werden.

9. Die digitale Information zukünftiger Generationen verlangt eine radikale Umgestaltung des jetzigen Urheberrechts-Gesetzes. Es wird ein Recht der jetzigen Autoren auf Integrität des Dokumentes, auf eine Speicherung in einem nationalen Langzeit-Archiv und auf offenen Zugang durch die von ihm gewünschte Community enthalten. Es muss aber auch das **Leserecht zukünftiger Generationen** sichern, jedenfalls soweit es für die wissenschaftliche Arbeit, für kulturelles Schaffen und in der Bildung relevant ist. Diese LZA-Rechtfestsetzung muss den zukünftigen Leser vor Einschränkungen durch 'Besitzrechte' kommerzieller Firmen durch Verschlüsselungen etc. schützen.

Der Weg zu einem auch in ferner Zukunft nutzbaren Hort des Wissens verlangt gegenwärtig eine „Gedächtnisorganisation“, welche die Ressourcen an fachkompetenten Experten, an Maschinen, Programmen, Datenbanken und an internationalen Vereinbarungen stärkt und ausschöpft. Ihre Organisation verlangt eine langfristig angelegte nationale LZA-Policy und deren konsequente Durchsetzung.



# Teil 3

## Materialien zur Entwicklung einer nationalen Langzeit-Archivierungspolicy

### 3.1 Einführung

#### 3.1.1 Strategische Ziele einer nationalen LZA-Policy

Industrie- und Kulturnationen brauchen ein kollektives Langzeit-Gedächtnis ihrer Erungenschaften, Erkenntnisse, ihrer Kulturleistungen, ihres Bildungsgutes.

Im Papierzeitalter gründete sich dies zu einem guten Teil auf gedruckte und geschriebene Werke – mit seiner technisch notwendigen Kopplung von Wort und Werk an die physische Speicherung auf Papier.

Die Umstellung auf *digitale* Dokumente befreit von technischen Einschränkungen wie dem begrenzten Textumfang je Dokument, der Mühsal, Dauer und den Kosten der technischen Erstellung, Verbreitung, Lagerung, Besorgung, Kopie, Auswertung. Auf die frühe Grundsatzstudie<sup>1</sup> der IuK sei hier verwiesen.

Zugleich aber erweist sich die Umstellung als Herkules-Aufgabe: wegen ihres gewaltigen Umfangs, wegen des technischen Neulands, weil die Autoren allmählich auch komplexere Dokumenten- und Kommunikationsstrukturen erproben, wegen der ungeheuren Vielfalt der Fachgebiete, Nutzungsarten und -gruppen.

Die Realisierung der Langzeit-Archivierung von digitalen Materialien (nur diese behandeln wir in dieser Expertise) der Vergangenheit und der Gegenwart wird – wie in

---

<sup>1</sup>*Digitale Bibliotheken; – Rahmenbedingungen, Perspektiven, Anforderungen und Empfehlungen – ; zur Neuordnung von Strukturen der Information und Kommunikation in den Wissenschaften; Positionspapier der Initiative Information und Kommunikation der wissenschaftlichen Fachgesellschaften in Deutschland (IuK Initiative); Executive Board der IuK Initiative; März 2002; [http://www.iuk-initiative.org/index.php?option=com\\_content\&task=view\&id=40\&Itemid=0](http://www.iuk-initiative.org/index.php?option=com_content\&task=view\&id=40\&Itemid=0)*

der Papier-Ära – professioneller Institutionen und Strukturen bedürfen. Nur sie können die notwendige internationale Vernetzung, die Entwicklung und Adaption intelligenter professioneller Dienste gewährleisten.

Nur der Staat kann die Verantwortung übernehmen, das Konzert der Dienstleister, Bibliotheken, Archive, Museen, Institute, Universitäten so zu leiten, dass eine synergetische, professionelle und effiziente langfristig stabile Gesamtstruktur der verteilten heterogenen an der Langzeit-Archivierung beteiligten Institutionen entsteht.

Dazu bedarf er einer *nationalen LZA-Policy* – als Vorgabe für die Exekutive, als Leitlinie für die LZA-Institutionen und als Willenserklärung seinen Bürgern gegenüber, für die Zukunft des kollektiven Gedächtnisses zu sorgen.

Die Langzeit-Archivierung wird technisch Dienste brauchen, die professionell angeboten werden. Es braucht Verabredungen – lokal unter Kollegen, fachspezifisch unter Instituten, international unter Ländern – damit diese Dienste auch greifen können, Texte zukünftig von externen Dritten gelesen werden (Formate) und gefunden werden können (Metadaten).

Eine zusätzliche Herausforderung ist das unterschiedliche Entwicklungstempo verschiedener Fächer, Länder, Autorengewohnheiten und Nutzungsarten.

### 3.1.2 Vielfalt der Entwicklungen (Länder, Inhalte, Autoren)

**Entwicklung in verschiedenen Ländern:** Die Möglichkeiten des World Wide Web wurden besonders früh entdeckt in Ländern in Randlage (z. B. Australien, Finnland, Hawaii) mit ihren besonderen Bedürfnissen nach Anbindung an die internationale Vernetzung. Dies hat im Gefolge zu einem gewissen Startvorsprung auch auf dem Gebiet der digitalen Langzeit-Archivierung geführt<sup>2</sup>.

Die Länder können jedoch je nach ihrer Leistungsfähigkeit und Finanzkraft nur in sehr verschiedenem Umfang zum zukünftigen weltweiten Netz der Langzeit-Archivierung beitragen: Sehr kleine Länder werden auf Patenschaften und externe Archivierung ihrer digitalen Materialien im Ausland angewiesen sein<sup>3</sup>. Entsprechend werden die Industriestaaten solche Leistungen für Drittländer – koordiniert über UNESCO bzw. EU – vorhalten müssen. Eine nationale LZA-Policy muss also die Rolle des Staates im internationalen Kontext definieren.

**Entwicklung in verschiedenen Wissenschafts-Fachgebieten:** Die Anforderungen an eine nationale Langzeit-Archivierung, und damit auch an eine nationale LZA-Policy sind für verschiedene Gebiete und Archivgut-Arten in Bezug auf das allgemeine Ziel ziemlich gleichartig: „Materialien, die für relevant gehalten werden, nutzbar bis in ferne Zukunft zu erhalten“.

<sup>2</sup>Analoges mag man auch für Städte in Randlage erkennen wie Arizona, Konstanz, Oldenburg, Tucson (Arizona, USA).

<sup>3</sup>In Abschnitt 3.4 werden die Antworten aus einigen Ländern analysiert. In 3.5 werden die Entwicklungen, Projekte und Planungen verschiedener ausgewählter Staaten dargestellt.

Die Unterschiede manifestieren sich vor Allem in den unterschiedlichen historisch gewachsenen Kommunikationskulturen, aber auch in den fachspezifischen Anforderungen. Es gibt Fachgebiete, in denen

- die genaue graphische Form der Dokumente auch in der digitalen Fassung bewahrt werden muss;
- die genaue sprachliche Formulierung selbst Teil der zu archivierenden Aussage ist;
- der wissenschaftliche Inhalt nicht von ihrer Fassung in einer natürlichen Sprache oder auch nur der Wahl der Sprache abhängig ist;
- eine Langzeit-Archivierung der Werke selbst nicht Teil der Anforderungen ist;
- weitgehend lokal, regional, national gearbeitet<sup>4</sup> wird.

Die genaue und fachgerechte Erarbeitung der Anforderungen aus der Vielzahl der unterschiedlichen Fachgebiete für eine Langzeit-Archivierung und daraus folgend für eine nationale LZA-Policy verlangt den Kontakt und die Auseinandersetzung mit Experten des jeweiligen Fachgebietes und entsprechend anzulegender Analysen und Expertisen. Dabei wäre der Schwerpunkt der Autoren auf den Blickwinkel zu lenken, welche Fachinhalte in welcher Form etwa im Jahre 2030 aus den Arbeiten der Jetztzeit man erwarten sollte, nicht auf die technische Realisierung, da sich die allermeisten Wissenschaftler mit der Langzeit-Archivierung selbst ihrer eigenen Werke bisher nicht ernsthaft auseinandergesetzt haben.

Die im Rahmen dieser Expertise (siehe 3.4) durchgeführte Umfrage beschränkte sich nur auf nationale Bibliotheken, nicht auf klassische Archive, und nicht auf Autoren und Nutzer. Es wird daher weiterer entsprechender, auch methodisch anders angelegter Expertisen und Umfragen bedürfen.

**Digitale Inhaltsbeschreibungen:** Digitale Werke erlauben das (auch nachträgliche) Hinzufügen von inhaltsbeschreibenden Zusatzinformationen. Zugleich erleichtern diese dann aber das Auffinden der Dokumente im Web, das Auffinden gesuchter Dokumententeile, und die digitale Weiterverarbeitung der Inhalte durch den Nutzer.

Die entsprechende Inhaltsanalyse muss durch Experten des jeweiligen Fachgebietes erfolgen. Hier können wir nur aus den exakten Naturwissenschaften, in denen die Autoren dieser Expertise lange gearbeitet haben, ein Schlaglicht setzen.

Die Ergebnisse dieser Wissenschaften sind nichtsprachlicher Natur. Chemische Reaktionen, physikalische Gesetze, biologische Entwicklungen haben als semantischen Kern einen Prozess in der Natur bzw. deren abstrahiertes mathematisches Abbild. Die Sprache dient nur der Vermittlung der Ergebnisse. Nicht nur sind daher Texte in diesen

---

<sup>4</sup>Erkennbar etwa an einer rein nationalsprachlichen Kommunikation.

Gebieten oft sprachlich einfach gehalten ('basic broken english') sondern ihre Formulierung ist nicht selbst Teil der Wissenschaft, sondern allenfalls des Faches Geschichte der Wissenschaft (wer hat wann was zuerst gemacht).

Daher haben Autoren dieser Gebiete schon immer einen besonderen Wunsch nach direkter Kennzeichnung semantischen Inhalts, nach Unterscheidung z. B. zwischen mathematischer Struktur und ihren unendlich vielen wesentlich verschiedenen mathematischen Darstellungen und erst recht Formulierungen. Beispiel: Der 'Satz von Columbus'<sup>5</sup>: „Die Erde kann umrundet werden, ohne dass der Seefahrer von der Oberfläche 'herunter fällt'“ ist eine erfahrbare Aussage über die reale Erdoberfläche. Dessen mathematische (und erst recht sprachliche) Darstellungen können sich aber durchaus sehr verschiedener mathematischer Gebiete bedienen (z. B. integrale oder differenzielle Aussagen) und erst recht verschiedener Formulierungen (Wahl der Symbole, Schreibweisen).

Daher haben Versuche abstrahierender Fassungen wissenschaftlicher Inhalte, und semantischer Beschreibungen schon immer größtes Interesse in den exakten Naturwissenschaften gehabt. Von daher keine Überraschung wurde die erste Markup-Sprache für digitale Dokumente,  $\LaTeX$ , nach ihrer Erfindung<sup>6</sup> durch Donald Knuth sehr rasch, vor allem aber in der Mathematik und theoretischen Physik weltweit sich verbreitend eingesetzt.  $\LaTeX$  erlaubt elementare international verabredete semantische Kodierungen der groben Textstruktur, von mathematischen Formeln und chemischen Strukturen und Reaktionen. Es erlaubt die Trennung von Kodierung und Erscheinungsbild.

Naturwissenschaftler haben als Aufgabe, neue messbare Aussagen über die Natur zu finden. Es macht hier keinen Sinn, regional oder lokal erneut anderwärts bereits bekannte Naturgesetze zu entdecken. Daher ist der effektivere Weg die möglichst instantane internationale Verbreitung neuer gefundener Erkenntnisse, um Doppelarbeit zu vermeiden. Digitale Dokumente wurden sehr rasch und international genutzt zur Kodierung von Texten und Formeln in offenen Formaten (ASCII) zur instantanen barrierefreien Verbreitung. Daher wurde *Email* direkt nach seiner Erfindung 1976 weltweit zum internationalen Austausch digital gefasster ASCII-Texte genutzt<sup>7</sup>, daher wurde  $\LaTeX$  schon früh von Großforschungsanlagen (CERN, GSI etc.) ebenso wie in Großverlagen (etwa Elsevier) genutzt. Eine Schlüsselfunktion wird hier das neue *Tex Document Center* TeXDocC<sup>8</sup> der SUB Göttingen einnehmen: Es bietet als Dienst eine Online-Evaluierung an, ob  $\LaTeX$ -Texte korrekt langzeitarchivierbar sind.

Die Kodierung semantischer Fachinhalte wird vor allem in der Mathematik vorangetrieben. Ausgehend von *browsing-MathML*, zu *content-MathML* gibt es jetzt erstmals konkrete Werkzeuge<sup>9</sup>, die auch die **semantische Kodierung nichtmathematischer In-**

<sup>5</sup>Nach seinen zahlreichen unabhängigen Entdeckern je nach Anwendung nach Green, Stokes, Gauß etc. benannt.

<sup>6</sup>Donald Knuth; TeX 1977;  $\LaTeX$  1984.

<sup>7</sup>Arpanet 1962; <http://de.wikipedia.org/wiki/ARPANET> und Milnet 1983; <http://de.wikipedia.org/wiki/MILNET>

<sup>8</sup><http://www.texdocc.de>

<sup>9</sup>OMDoc (*Open Mathematical Documents*); <http://www.mathweb.org/omdoc/>

**halte** erlaubt. Anwendungen in anderen (zunächst benachbarten) Fächern wie den Geowissenschaften, der Informatik und der Physik sind in Arbeit.

Das *World Wide Web* wurde am CERN 1993 ersonnen<sup>10</sup> als unmittelbare Notwendigkeit zum Austausch zwischen Archiven<sup>11</sup>.

Eine maschinenlesbare, international vereinbarte Kodierung von Werk-Inhalten (semantische Beschreibung) steht in den meisten anderen Gebieten und Bereichen noch sehr am Anfang.

**Autoren und Leser:** Die notwendigen Kenntnisse – und erst recht die Akzeptanz – der geeigneten Vorbereitung für die Erstellung von digitalen Dokumenten für Langzeit-Archive fehlen bei den allermeisten Autoren und Lesern/Nutzern noch. Daher sind auch sehr viele lokale neu angelegte digitale Archive noch ziemlich leer, obgleich die Autoren in sehr vielen Fällen ihre digitalen Dokumente über ihr Online-Schriftenverzeichnis auf ihrem eigenen Server nachweisen – allerdings zumeist in einer für eine Langzeit-Archivierung ungeeigneten Form. Die LZA-Anforderungen an ein Dokument durch den Autor erfüllen zu lassen, wird man wohl nur durch eine Verpflichtung erreichen, oder durch Karriereanreize (z. B. öffentlich sichtbare Messergebnisse der Zahl der Zitierungen seiner Werke). Die zugehörige Schulung der Autoren aller Gebiete ist eine gigantische Aufgabe, die nur verteilt angegangen werden kann.

### 3.1.3 Zur Geschichte der Langzeit-Archivierung

Den Übergang von der 'Papier-Ära' der Information in gedruckten Dokumenten zum digitalen Zeitalter der Kommunikation von Wissen, technisch angestoßen durch den flächendeckenden Einzug der EDV mit dem PC in den 1980-er Jahren und beschleunigt durch die Entwicklung des World Wide Web seit 1993, stellten sich Experten damals als einfach und zeitlich unmittelbar realisierbar vor. Tatsächlich hat sich der Übergang aber als sehr allmählich und mühsam herausgestellt.

Eigentlich wäre dies zu erwarten gewesen: Schon die Theorie des Phasenübergangs in *endlichen* Systemen<sup>12</sup> sagt bei einem Umbruch eines Systems zu einem neuen Gleich-

<sup>10</sup> Tim Berners-Lee; <http://www.w3.org/People/Berners-Lee/>, Robert Cailliau; <http://robert.cailliau.free.fr/ByLetter/M/Me/CV.html>

<sup>11</sup> Robert Cailliau erläuterte einem der Autoren am Tisch vor der CERN-Cafeteria die Wahl des Namens „Wir Wissenschaftler sind ein World Wide Web und brauchen eine Technik, die die direkte barrierefreie Vermittlung von Inhalten stabil für alle zukünftigen Generationen ermöglicht.“. Er macht sich aber Sorgen um die Folgen der Internationalisierung für andere Bereiche.

<sup>12</sup> a) Eberhard R. Hilf et al.: *Die Zukunft des wissenschaftlichen Publizierens – Vom Publizieren zum Austausch von Informationen in der Wissenschaft*; Kolloquium am Fachbereich Physik; Carl von Ossietzky Universität Oldenburg; 13. Januar 2003; <http://physnet.uni-oldenburg.de/~hilf/vortraege/oldenburg03/oldenburg03.html#I-5>

b) E. F. Schumacher; *Small is beautiful*; Rowohlt; Reinbek bei Hamburg 1977 (Rowohlt); 316 Seiten; ISBN 3-498-06121-6; <http://www.langelieder.de/lit-schumacher.html>

c) H.-P. Baltes und Eberhard R. Hilf; *Spectra of Finite Systems*; Wissenschaftsverlag; 1971; ISBN 3-411-01491-1; <http://smallsystems.isn-oldenburg.de/publications/metadocs/ebs.spectra.finite.systems.prep.html>

gewichtszustand (für den Laien überraschenderweise) einen zeitlich um so allmählicheren Übergang voraus, je kleiner das System ist. Auch der Übergang vom Holztafel/Pergament-Zeitalter zum Papier-Zeitalter dauerte fast 100 Jahre - mit jahrzehntelangem Widerstand ausgerechnet der Bibliotheken gegen den Papierdruck<sup>13</sup>. Die Welt des Wissens und der Wissenden ist also beschränkt, wenn auch für den Einzelnen nicht überschaubar<sup>14</sup>.

Das in fast 600 Jahren gewachsene System der wissenschaftlichen Papier-Bibliotheken hat erst sehr allmählich zu schließlich festen, professionalisierten, etablierten und durchstrukturierten Geschäftsketten sowie zu entsprechenden Gewohnheiten geführt. Zugleich sind die an diesen traditionellen Strukturen Beteiligten in den Prozess selbst seit langem fest eingebunden, gedanklich, zeitlich und finanziell, sodass ihnen wenig Ressourcen bleiben, Neues zu probieren.

Erschwert wird dieser Übergang auch durch die hohe professionelle Arbeitsteilung bei gleichzeitigem Auseinanderdriften der Sektoren: Autoren, Bibliotheken und Verlage und Geldgeber „wissen“, was für sie zu tun ist. Einer Absprache oder gar Abstimmung der Sektoren untereinander schien es kaum zu bedürfen.

Es ist daher verständlich, dass der Umbruch zur digitalen Ära sich nur zögerlich und mit vielen Rückschlägen sowie weltweit in verschiedenen Ländern bzw. in verschiedenen Disziplinen mit sehr ungleichen Geschwindigkeiten vollzieht. So hat z. B. Australien bereits 1996 ein nationales E-Government-System etabliert, das die volle vertikale Tiefe von der Bundes- über die Landes- bis zu den Regional- und Orts-Ebenen miteinander verknüpft, inklusive einer verteilten integrierten E-Aktenhaltung mit gemeinschaftlich eingeführten Metadaten. Durch den frühen Beginn des Systems konnten inzwischen wertvolle Erfahrungen gesammelt werden, die schon entsprechend frühzeitig zu Anpassungen und Revidierungen geführt haben.

Anderenorts gab es bereits 1994 so verschiedenartige Vorreiter beim Einsatz digitaler Dokumente und des World Wide Web, wie z. B. Finnland, Kanada, und sogar die Kleinhändler von Oahu in Hawaii oder die Stadtverwaltung von Tucson, Arizona, USA.

P. Koerbin<sup>15</sup> sieht als derzeitige Herausforderungen des LZA-Vorhabens in Australien (und diese gelten natürlich analog für andere Länder) u. a.:

- Es können nicht alle E-Dokumente archiviert werden, welche die Wissenschaftler in der Zukunft möglicherweise vorzufinden wünschen könnten, erst recht nicht die riesige Menge der zur Zeit erzeugten Dokumente;
- Langzeit-Archivieren ist sehr arbeitsintensiv und damit teuer;

<sup>13</sup>Chiara Frugoni: *Das Mittelalter auf der Nase*; C.H.Beck Verlag; 2003; ISBN 3 406 50911 8.

<sup>14</sup>Astrophysiker würden formulieren: *Der eigene Ereignishorizont ist kleiner als der endliche Horizont des Universums.*

<sup>15</sup>P. Koerbin; *PANDORA Australia's Web Archive; Library Science Talks; SNL/CERN September 2004.*

- Es lässt sich nicht die volle Komplexität der Verlinkung des Dokuments im Netz abbilden und sichern.

Der Unterschied von Papier- zu digitalen Werken in Bezug auf die Langzeit-Archivierung wird auch von Wendy Smith et. al.<sup>16</sup>, Stephen Pinfield<sup>17</sup> und in verschiedenen Publikationen von J. Rothenberg<sup>18</sup> sowie in zahlreichen Artikeln, z. B. im Archiv von CLIR<sup>19</sup> ausführlich analysiert.

Insgesamt ist die Diskussion, was *technisch* zu tun wäre, was erforderlich ist, welche Anforderungen zu stellen sind, weitgehend und im Konsens abgeschlossen.

Was fehlt sind *politischer* Gestaltungswille, Durchsetzungs- und Finanzierungskraft, die Lösung der immensen organisatorischen, strukturellen und Management-Herausforderungen sowie die flächendeckende Implementation.

Ein Weg hierzu wird die bundesweite Schulung der Autoren als Erzeuger von wissenschaftlicher und kultureller Information sein, damit sie die Anforderungen einer Langzeit-Archivierung von Anfang an schon bei der Erstellung eines Werkes kennen und berücksichtigen, sowie eine breite öffentliche Diskussion.

In Deutschland verlief die Entwicklung sehr ungleich: Einerseits wurden z. B. bereits im Frühjahr 1994 an allen Fachbereichen der Physik Web-Server eingerichtet; 1995 wurden die je weltweit über alle Wissenschaftsinstitutionen verteilten offenen wissenschaftlichen Web-Dokumentensysteme Math-Net<sup>20</sup> und PhysNet<sup>21</sup> initiiert, über die jeweils der Nachweis der lokal gespeicherten digitalen wissenschaftlichen Dokumente erfolgt<sup>22</sup>.

Andererseits wurde bis heute keine breite organisierte nationale Vertretung in den internationalen Standardisierungsgremien erreicht und finanziert<sup>23</sup>. Seit 1995 gab es mehrere Anläufe in Deutschland, zusammen mit den großen kommerziellen und den Fachgesellschafts-Verlagen ein nationales bzw. internationales verteiltes E-Dokumentenarchiv zu entwickeln. Diese scheiterten<sup>24</sup> aber, weil der technologische Umbruch die staatlichen bzw. überstaatlichen Entscheidungsträger in ihrem eigenen Bereich noch

---

<sup>16</sup>Wendy Smith, W. E. Phillips, D. Campbell; *Preserving Access to Web Publications*; <http://ausweb.scu.edu.au/proceedings/smith/>

<sup>17</sup>Stephen Pinfield, Hamish James; *The Digital Preservation of e-Prints*; D-Lib Magazine, September 2003 Vol.9 Number 9 ISSN1082-9873; <http://www.dlib.org/dlib/september03/pinfield/09pinfield.html>

<sup>18</sup>Jeff Rothenberg; *Ensuring the Longevity of Digital Documents*; Scientific American 272(1) 1995.

<sup>19</sup><http://www.clir.org/pubs/issues/issues.html>

<sup>20</sup><http://www.math-net.org>

<sup>21</sup><http://www.physnet.net>

<sup>22</sup>So weist *PhysNet* <http://de.physnet.net/PhysNet/physdoc.html> mehrere Millionen wissenschaftliche Dokumente der Arbeitsgruppen von etwa 6.000 weltweit verteilten Physik-Institutionen nach.

<sup>23</sup>Es gab aber bereits in 1995 und dann wieder 2003 hierzu Anträge an das BMBF.

<sup>24</sup>Verbund-Anträge an den Bund (*global-info* 1996); <http://web.archive.org/web/19990125095628/http://www.global-info.org/> bzw. an die EU (*DDD Multimedia scientific Distributed Document Database in Physics* 1995); <http://www.physik.uni-oldenburg.de/ddd-physics/>

nicht erreicht hatte und so Antragsteller und Entscheider nicht die gleiche Sicht der zukünftigen Entwicklung hatten.

Die nationale Langzeit-Archivierung von Dokumenten aus Wissenschaft, Kultur und akademischer Lehre wurde aber inzwischen von den großen nationalen Archivierungsinstitutionen und den großen Förderinstitutionen (DFG, BMBF) als Notwendigkeit erkannt und als Herausforderung angenommen. Beispiele der aktuellen Aktivitäten sind das Projekt *nestor* selbst sowie die Vorbereitung der Schaffung einer Infrastruktur zur Langzeit-Archivierung digitaler Publikationen im Rahmen von NEDLIB<sup>25</sup>.

Im letzten Jahrzehnt hatte sich die jahrhundertlange Symbiose einer Vielzahl von professionellen wissenschaftlichen Verlagen mit den Wissenschaftlern als Autoren und Lesern und den Bibliotheken als Archiven in eine ungewollte Abhängigkeit verwandelt:

- Zwei eng zusammenhängende kommerzielle Verlage<sup>26</sup> haben die meisten anderen Verlage aufgekauft, haben so fast ein Monopol im Markt (über 50% Marktanteil) und können daher die Preise beliebig<sup>27</sup> und kostenunabhängig festsetzen und die Bedingungen so gestalten, dass ein Ausstieg einer einzelnen Universitätsbibliothek schwierig bis unmöglich<sup>28</sup> ist.
- Einige Vertreter kommerzieller Wissenschafts-Verlage sehen ihre eigene Klientel inzwischen sehr kritisch<sup>29</sup>.

Die Bibliotheken können die wissenschaftlichen Zeitschriften nicht mehr bezahlen: Von den derzeit ca. 200.000 referierten Zeitschriften hat die Harvard University 106.000, die Stanford University 16.000, eine normale deutsche Universität dagegen vielleicht nur noch einige Hundert.

Folglich werden immer weniger Zeitschriftenartikel gelesen. Dies mag man am Trend ablesen, älteres Wissen nicht mehr durch Zitierung des Originalartikels der Erstentdeckung zu zitieren, sondern zunehmend durch Nennung von Review-Artikeln<sup>30</sup>. Es ist eben dem normalen Wissenschaftler meist nicht mehr möglich, an die Originalartikel

<sup>25</sup>NEDLIB (*Networked European Deposit Library*); DFG-Projekt 1998 - 2000; <http://www.kb.nl/coop/nedlib/>

<sup>26</sup>Elsevier bzw. Springer/Kluwer. Letzterer hat den bisherigen CEO von Elsevier als eigenen CEO eingestellt.

<sup>27</sup>Erkennbar z. B. an den ca. 40 % Rendite von ES im Jahr 2003.

<sup>28</sup>Als Beispiel sei das 'Bremer Rätsel' erwähnt; Wer, etwa bei einer Haushaltsnotlage, abbestellt, muss mehr zahlen, als wenn er es lässt; <http://www.isn-oldenburg.de/~hilf/vortraege/dini-0a05/0a05-11.html>

<sup>29</sup>So sagte der Vertreter des Deutschen Börsenvereins, M. Sprang, der dort als Vertreter der kommerziellen Verleger auftrat, am 11. 11. 2004 in der öffentlichen Anhörung im Bundespresseamt zum Referentenentwurf des BMJ zum Urheberrechtsgesetz, Korb 2: „Die Wissenschaftler sind unsere natürlichen Feinde...“.

<sup>30</sup>Analyse, vorgestellt auf der internationalen IUPAP-Konferenz zu *Workshop on Scientific Misconduct and the Role of Physics Journals in its Investigation and Prevention*; London 2003; <http://www.iupap.org/wg/communications/ethics/program.html>



mit vertretbarem Aufwand heranzukommen, auch unabhängig von den Kosten. So erklärt sich auch die Vielzahl der fehlerhaften und/oder ungenauen Zitierungen bahnbrechender alter Artikel<sup>31</sup>.

### 3.1.4 Ära ohne Langzeit-Gedächtnis

Der Aufbau eines dem digitalen Zeitalter angepassten digitalen LZA-Systems erfordert viel Zeit, Geduld, Geld und viele engagierte Institutionen und Personen, um die ungeheuren Möglichkeiten zum Nutzen der Wissenschaft, der Kultur und der akademischen Lehre auszuschöpfen.

Diese unsere Zwischenzeit, in der wir leben, also sagen wir 1993 - 2013, wird manchmal bereits als die „Zeit ohne Gedächtnis“ angesehen. Zeitschriften sind nicht mehr erreichbar, weil sie nicht mehr an einer Vielzahl von Bibliotheken gehalten werden. Die in dieser Zeit aufgebauten riesigen digitalen Archive von Dissertationen und wissenschaftlichen Dokumenten im proprietären *pdf*-Format von ADOBE<sup>32</sup> sind zum Teil bereits jetzt nicht mehr einwandfrei lesbar, weil die Übertragung auf neuere Versionen nicht mehr eindeutig bzw. fehlerfrei gelingt. Daher werden bei dem international ältesten und bereits seit 1991 erprobt stabilen und sicheren digitalen Langzeit-Archiv *ArXiv* (früher am LANL<sup>33</sup>, heute an der Cornell University<sup>34</sup>) Dokumente nur angenommen, wenn sie in einem nichtproprietären Format *und* mit allen erforderlichen *Macros* (ergänzender Programm-Code, um den Text verarbeiten zu können) eingereicht werden. *ArXiv* erzeugt aus diesen Dokumenten mit der vollen Information zweimal im Jahr dynamisch Textansichten<sup>35</sup> sowohl im jeweils aktuellen *pdf*-Format wie in anderen derzeit gängigen Darstellungsformaten (*html*, *postscript*). Da die Quellen mit ihrer vollständigen Information archiviert bleiben, ist dieses Archiv jederzeit intakt und aktuell lesbar.

Ein Grundproblem bei der Migration von Dokumenten zu einem informationsärmeren Format ist die Vernichtung von Information (genannt *Projektion*), also z. B. von *XML* zu *L<sup>A</sup>T<sub>E</sub>X* und von dort zu *pdf*. Ihr Gegenstück, genannt *Inflation*, bei der das Dokument zu einer informationsreicheren Struktur migriert wird (etwa von *L<sup>A</sup>T<sub>E</sub>X* zu *XML* oder *MathML*, *CML* etc.), erfordert, sich für den möglichen Informationszuwachs maschinell Vorschläge machen zu lassen, um durch Interaktion mit dem Autor (Rückfragen, versuchsweises Rückprojizieren und Nachfragen) diesem die Möglichkeit zu geben, nachzubessern und zu korrigieren<sup>36</sup>. Teilweise werden auch neuartige Werkzeuge ein-

<sup>31</sup>M. Cardona, W. Marx; *Verwechselt, vergessen, wieder gefunden*; Physik Journal 11/2004 S.27.

<sup>32</sup>ADOBE hat dies erkannt und bietet nun *pdf/A*, ein speziell für das Langzeit-Archivieren entwickeltes Dokumentenformat, dessen Programm-Code von ADOBE offengelegt wurde.

<sup>33</sup>LANL (*Los Alamos National Laboratory*); <http://xxx.lanl.gov/>

<sup>34</sup><http://arxiv.org/>

<sup>35</sup>Thorsten Schwander; Internationaler Workshop Mathematik Dissertationen; Göttingen, 2002; <http://www.ub.uni-duisburg.de/mathdiss/Schwander/vortrag.html>

<sup>36</sup>Besonders beliebt sind solche technischen Unterstützungen bei der Migration von mathematischen Formeln. Es werde z. B. automatisiert der Ausdruck *sin(x)* aus *L<sup>A</sup>T<sub>E</sub>X* zur semantisch reicheren Sprache *MathML* migriert. Die Rückprojektion ergibt dann möglicherweise den ungewollten Ausdruck

gesetzt<sup>37</sup>, die den Autor durch semantische Informationen kodierende Editoren unterstützen sollen. Das Ziel ist daher, dass der Autor den gesamten semantischen Inhalt seines Dokumentes gleich möglichst vollständig mitteilt, maschinenlesbar kodiert und sich hierbei weitestgehend durch intelligente Editorprogramme und Experten unterstützen lässt.

Schlimmer als die vollkommene Inkompatibilität verschiedener Fassungen, die meist zeitnah bemerkt wird, sind jedoch schleichende Konvertierungsfehler. Wer mag sich die Folgen ausdenken, wenn bei einer Konvertierung z. B. jedes millionste Minuszeichen fehlt<sup>38</sup>. Hier kann nur die Beschränkung auf offene Formate und Software und die Speicherung des vollen Quellcodes eine gewisse Sicherheit bringen, weil hier Fehler im Programm von Dritten selbst dann aufspürbar sind, wenn es die Erzeugerfirma eines proprietären Editor-Programms gar nicht mehr gibt.

Es bedarf größter Anstrengungen des Staates, gegründet auf einer nationalen LZA-Policy, die Autoren anzuleiten und von ihnen zu fordern, ihr Material LZA-fähig zu erzeugen. Wenn doch Dokumente in anderen als offenen Formaten archiviert werden sollen, sollten diese bereits *vor der Einlagerung* in ein LZ-Archiv auf offene nichtproprietäre Formate migriert werden, um die Lesbarkeit auch in ferner Zukunft zu ermöglichen.

Diese Expertise will zu einer zur nationalen LZA-Policy führenden Diskussion beitragen. Die LZA-Policy soll dann zukünftig als eine Art Leitfaden dienen, so dass die Vielzahl der Beteiligten koordiniert auf die nationale Realisierung hinarbeiten und der nationale Beitrag international eingebettet werden kann. So könnte die wissenschaftliche und technische Entwicklung in Deutschland nachhaltig unterstützt – und das Zeitfenster der nationalen Langzeit-Gedächtnislosigkeit verkleinert werden.

Wir werden besonders auf die Entwicklung im Ausland eingehen, weil hier zum Teil bereits wesentlich längere Erfahrungen vorliegen als in Deutschland.

Die derzeitigen Initiativen in Deutschland<sup>39</sup> auf dem Gebiet der Langzeit-Archivierung werden nun im Leitprojekt *nestor* des BMBF zusammengeführt. Zu den jeweiligen aktuellen Entwicklungen gibt *nestor* einen Newsletter<sup>40</sup> heraus.

## 3.2 Museen und Archive

Diese Expertise hat als Schwerpunkt die derzeit am intensivsten diskutierte *Langzeit-Archivierung* wissenschaftlicher Dokumente.

---

$s * \sqrt{-1} * n(x)$ . Erst der Autor kann bei einem Print-Ausdruck seinen Kodierungsfehler erkennen und ihn dann im  $\text{\LaTeX}$  Code nachbessern.

<sup>37</sup>OMDoC *Open Mathematical Documents*; <http://www.mathweb.org/omdoc/>

<sup>38</sup> Dies kam bei einem der Routine-Migrierungen im *ePrint ArXiv* zutage, als man neuere mit älteren, aus demselben Original erzeugten *pdf*-Ansichten verglich.

<sup>39</sup><http://www.LangzeitArchivierung.de>

<sup>40</sup><http://nestor.sub.uni-goettingen.de/newsletter/>

Für die Besonderheiten der Langzeit-Archivierung in *Museen* wird auf die *nestor*-Materialie von Dirk Witthaut et al.<sup>41</sup> sowie den Ratgeber für die Bewahrung digitaler Daten in Museen<sup>42</sup> verwiesen. Eine nationale LZA-Policy zielt bei Archiven und Museen insbesondere darauf ab, die für das Verständnis und die kulturelle Einordnung wichtigen *digitalen Sekundärinformationen* über physische Museums-Unikate, digitale Primärarchive und Digitalisate, analogen Regelungen bezüglich der Harmonisierung, der Sicherung und dem Zugang zu unterwerfen wie denen für digitale Materialien aus Kultur und Wissenschaft.

*Die klassischen Archive*, die relevante Unikate aus Kultur (z. B. Geschichte, Religion, Musik, Kunst), öffentlichem Leben, Bildung und Wissenschaft bzw. von Behörden und staatlichen Institutionen sammeln, sind zumeist in öffentlicher Trägerschaft, also in staatlicher, kommunaler bzw. kirchlicher Hand.

Im Archivwesen (hier meist 'klassische Archive' genannt) selbst wird traditionell<sup>43</sup> nach den Rechtsträgern unterschieden:

1. Staatliche Archive wie das Bundesarchiv und die Staats- bzw. Landesarchive;
2. Kommunalarchive der Städte, Kreise und Gemeinden;
3. Kirchliche Archive;
4. Haus- und Familien-Archive;
5. Wirtschaftsarchive (z. B. Industrie- und Handelskammern);
6. Politische Archive (Parlament, Parteien, Gewerkschaften, Verbände);
7. Medienarchive (Film, Fernsehen, Funk, Bild);
8. Hochschularchive;
9. 'Alternative' Archive (z. B. sozialer Bewegungen).

Den ersten acht Sparten entsprechen Fachgruppen des VdA. Die Auswahlprinzipien sind spezifisch durch den jeweiligen Träger gegeben. In dieser Expertise beschränken wir uns auf gemeinsame Anforderungen an eine nationale LZA-Policy, soweit sie *digitale* Werke betreffen. Spartenspezifische Ergänzungen müssen durch diese erarbeitet werden.

---

<sup>41</sup>Dirk Witthaut unter Mitarbeit von Andrea Zierer, Arno Dettmers, Stefan Rohde-Enslin; *Digitalisierung und Erhalt von Digitalisaten in deutschen Museen*; Frankfurt am Main; nestor c/o Die Deutsche Bibliothek, 2005. IV; [http://www.langzeitarchivierung.de/downloads/mat/nestor\\_mat\\_02.pdf](http://www.langzeitarchivierung.de/downloads/mat/nestor_mat_02.pdf)

<sup>42</sup>Stefan Rohde-Enslin; *Nicht von Dauer: Kleiner Ratgeber für die Bewahrung digitaler Daten in Museen*; nestor-ratgeber 2; BMBF-Projekt *nestor* Kompetenznetzwerk Langzeit-Archivierung; Herausgeber: Das Institut für Museumskunde der Staatlichen Museen zu Berlin und die Stiftung Preußischer Kulturbesitz; 22. Juli 2005; <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0008-20041103017>

<sup>43</sup>VdA (Verband deutscher Archivarinnen und Archivare e.V.); <http://www.vda.archiv.net/index.htm?fachgruppen.htm>

Archive und Museen haben vom Auftrag her schon immer Langzeit-Archivierung betrieben. Daher wurden und werden hier Kriterien und Verfahren angewandt, die bei wissenschaftlichen Dokumenten bisher noch nicht üblich sind und nun erst in der aktuellen Diskussion für eine digitale Langzeit-Archivierung gefordert werden müssen.

In diesen „klassischen“ Archiven<sup>44</sup> sind die primären Archiv-Objekte oft Bündel von Dokumenten. Um ein Werk authentisch, vollständig und verstehbar für die Nachwelt zu erhalten, werden neben dem Original auch sein Kontext, sein Entstehungszusammenhang (Akte, Vorgang) sowie von und auf das Objekt zielende Verweise mit gesammelt, mit Metadaten und Protokollinformationen versehen und zusammen archiviert.

Museen und Archive haben schon immer nicht alles gesammelt, was aktuell erzeugt wird, sondern eine strenge, oft auch formal geregelte und einer eigenen institutionellen Policy bzw. dem Auftraggeber unterworfenen Auswahl getroffen.

Archivierungsregeln wurden zum Teil von Bund- und Länderverwaltungen, bzw. von den kommunalen Auftraggebern festgelegt<sup>45</sup>. Hierzu zählen auch die Vorgaben zu Bildformaten, Metadaten oder z. B. die Beschränkung auf ASCII-Dateien bei Texten.

Eine nationale LZA-Policy muss die Bereiche benennen, in denen eine Langzeit-Archivierung als Bundes- oder Landesaufgabe wahrzunehmen ist. Spiegelungen zu externen Einrichtungen, erst recht nicht zu solchen in anderen Ländern, sind hier bisher noch nicht erprobt.

Der Zugang zu solchem Archivgut selbst wird bisher meist nur auf begründeten Antrag gewährt, bei physischen Materialien schon durch den dazu notwendigen Aufwand bzw. die Fragilität des primären Objektes eingeschränkt. Weitere Einschränkungen ergeben sich aus Aspekten des Geheimsschutzes, vor allem aber des Daten- und Persönlichkeitsschutzes (Sperrfristen für Behörden-Material zumeist 30 Jahre).

Für die digitalen Objekte, Akten und Vorgänge ist dagegen prinzipiell ein für das Objekt selbst schadensfreier, bequemer und instantaner Zugang aus aller Welt möglich. Daher ist in einer nationalen LZA-Policy der Zugang zu den digitalen Werken neu festzulegen. Der nun mögliche, i. A. gewünschte breite offene Zugang (auch über externe Kopien des Archivs) muss mit Restriktionen des Daten- und Persönlichkeitsschutzes, sowie nationaler Interessen neu austariert werden.

Die Kosten einer Langzeit-Archivierung digitalen Materials in Museen und Archiven, insbesondere die neu durch Forderungen aus einer nationalen LZA-Policy zu digitalem Material entstehenden, müssen ja vom Staat bzw. den Ländern oder Kommunen aufgebracht werden. Sie können dann bei einmal bestehenden gesetzlichen Regelungen aber auch eingefordert werden.

Die nationale LZA-Policy sollte den notwendigen gemeinsamen Rahmen setzen für alle Institutionen, die Langzeit-Archivierung betreiben, Museen und Archive einge-

---

<sup>44</sup>Wertvolle Fach-Hinweise zu diesem Abschnitt verdanken wir Karl-Ernst Lupprian.

<sup>45</sup>DOMEA-Konzept für Dokumenten-Management und elektronische Archivierung in der öffentlichen Verwaltung; <http://www.kbst.bund.de/- , 413/DOMEA-Konzept.htm>

schlossen. Nur so lassen sich nationale Standards setzen, diese in die internationalen Standards einbetten und die Staatsaufgabe der Sicherung der Rechte aller Beteiligten realisieren und mit den nachgeordneten Behörden abstimmen.

Die sichere Langzeit-Archivierung digitaler Materialien und Informationen in Museen und Archiven bedarf steter Forschung in dazu spezialisierten Einrichtungen zu seiner auch zukünftig sicheren Realisierung. Die zugehörige Forschung geschieht dabei sowohl lokal wie durch Anknüpfung an die allgemeine Entwicklung.

Aktuell hatte auf europäischer Ebene das DLM (*Document Lifecycle Management*) Forum<sup>46</sup> im Oktober 2005 seine vierte Tagung organisiert. National gibt es den *Arbeitskreis für die Archivierung von Unterlagen aus digitalen Systemen*<sup>47</sup>.

Die Ausbildung zum 'Archivar'<sup>48</sup> in seinen verschiedenen Ebenen wird zunehmend durch Fachkenntnisse zur Langzeit-Archivierung *digitaler* Dokumente ergänzt werden müssen. Dies kann dann durch Anpassung existierender Curricula oder durch eine eigene Ausbildungslinien z. B. zu einem „Verwaltungsinformatiker“ geschehen.

Insgesamt verlangt die Erarbeitung der Anforderungen der klassischen Archive eine eigene Expertise, zu der Experten aus diesem Bereich sowie die auftraggebenden Behörden und staatlichen Dienststellen hinzuzuziehen wären. Auch kam zu der Fragebogenaktion dieser Expertise (Abschnitt 3.4) kein relevanter Rücklauf.

### 3.3 Begriffe und deren Eingrenzungen

Nicht nur im angelsächsischen Umfeld sind inzwischen Begriffe, wie „Policy“, „Strategy“, „Guideline“ und „Best Practice“, verbreitet, die sich jedoch nur auf nichttriviale Weise ins Deutsche übersetzen lassen, so dass hier zunächst eine Erläuterung den weiteren Sprachgebrauch in diesem Text klären soll. Denn es gibt klare Unterschiede zwischen diesen oft leichthin als Synonym verwendeten Worten:

Das Wort „Policy“ ist nur schwer zu übersetzen, es gibt eine „National Policy“ und eine davon zu unterscheidende „Institutional Policy“.

„National Policy“ beschreibt die notwendigen langfristigen politischen Rahmenbedingungen, die auf regionaler, nationaler oder internationaler Ebene staatlich durch Vorgaben auferlegt werden und nicht durch die archivierenden Institutionen im Nachhinein beeinflussbar sind.

„Institutional Policy“ lässt sich wohl am einfachsten in Anlehnung an „Corporate Policy“ als Firmenpolitik übersetzen. Sie ist eine kurze und von der Leitung der Institution gegebene Richtlinie als Orientierung zu einem Sachthema. Sie hat keine Halbwertszeit, sondern ist die allgemeine Einstellung zu einem Themenkomplex und gilt

<sup>46</sup>[http://europa.eu.int/comm/secretariat\\_general/edoc\\_management/dlm\\_forum/](http://europa.eu.int/comm/secretariat_general/edoc_management/dlm_forum/)

<sup>47</sup><http://www.archive.nrw.de/archivar/2000-04/A17.htm>

<sup>48</sup>Genauer: Archivassistent/in bzw. Fachangestellte/r für Medien- und Informationsdienste mit verschiedenen Fachrichtungen; Archivar/Archivarin im höheren Dienst; Diplomarchivar (FH).

solange, bis äußere Umstände eine Änderung erzwingen oder eine andere Politik von der Institution eingeschlagen wird.

Eine „Strategy“ hingegen hat eine Halbwertszeit und wird regelmäßig, beispielsweise alle fünf Jahre, überarbeitet. Hier werden zeitlich erreichbare Ziele definiert und auch die Methoden, Mittel und Ressourcen zur Erreichung der Ziele spezifiziert. Eine „Strategy“ ist also wesentlich konkreter, möglicherweise aber kurzlebiger als eine „Policy“.

„Guidelines“ und „Best Practices“ wiederum beziehen sich auf einen spezifischen Teil eines Workflows, eine spezifische Aktivität, eine spezifische Person in ihrer Rolle und Funktion. Hier wird auf die maximale Effizienz und Qualität im Sinne eines „Business Process“ gezielt.

Diese Definitionen entsprechen in etwa den im europäischen ERPANET-Projekt (*Electronic Resource Preservation and Access Network*<sup>49</sup>) verwendeten und etablierten<sup>50</sup>.

In diesem Sinne beschränkt sich das Folgende auf eine „National Policy“, wobei sich hieraus implizit immer auch Aussagen über jenen Teil von „Institutional Policies“ ergeben, die in den staatlichen Rahmen passen müssen oder sogar Teil von diesen sind, also beispielsweise den in ein LZA-Netzwerk integrierten Institutionen.

Kernpunkte einer nationalen Policy sind jedoch die Architektur, Organisation und die Rahmensetzung für die archivierenden Institutionen mit gesetzlichem oder vertraglich geregeltem Auftrag, die Verabredungen mit anderen LZA-Institutionen im In- und Ausland sowie die Anforderungen an die LZA-Institutionen selbst (Ausbildung ihrer Mitarbeiter, Finanzierung, Ausstattung, Vertrauenswürdigkeit).

In dieser Expertise befassen wir uns für die angestrebte Langzeit-Archivierung digitaler Werke aus Kultur, Wissenschaft und akademischer Lehre mit einer *Nationalen Langzeit-Archivierungspolicy* (*National long-term Preservation Policy*) digitaler Dokumente, für die es auch im Ausland noch keine Vorbilder gibt. Belastbare Erfahrungen Dritter über viele Jahrzehnte als wirklichen Test kann es ja schon naturgemäß nicht geben.

Unter Langzeit-Archivierung verstehen wir das „Zur Verfügung halten“ digital gespeicherter Informationen für zukünftige Generationen - und als Voraussetzung das der dabei verwendeten Techniken.

Dagegen gibt es bereits bei vielen Nationen institutionelle Archivierungspolicies ihrer nationalen Bibliotheken für digitale Dokumente (*Preservation Policy*), oft für ganze Dokumentenklassen<sup>51</sup> und oft national bereits mit entsprechenden rechtlichen Rahmenbedingungen versehen.

Der Vielfalt der Fächer, Publikations-Kulturen und Aufgaben entsprechend werden Begriffe zum Teil mit sehr unterschiedlicher Bedeutung benutzt oder es werden Begriffe neu geschöpft in Unkenntnis ihrer andersartigen Nutzung in anderen Gebieten.

<sup>49</sup><http://www.erpanet.org>

<sup>50</sup>Andreas Aschenbrenner; private Email 2004.

<sup>51</sup>Die DDB archiviert alle digitalen Dissertationen aus Hochschulen in Deutschland; <http://www.dissonline.de>

So werden wir den Begriff digitale *Bibliothek* und *Archiv* synonym für alle Institutionen verwenden, die digitale Werke langzeitarchivieren, also sowohl für Online-Archive, für Bibliotheken wie für Archive. Tatsächlich bilden aber historisch die Bibliotheken und die Archive auf Grund ihres unterschiedlichen Auftrages und ihrer Auftraggeber getrennte Welten – jeweils weltweit in sich vernetzt aber voneinander getrennt<sup>52</sup>.

Wir beschränken uns hier auf *digitale* Dokumente und verwenden den Begriff „Dokument“ für jede Art digitalen Objektes, das langzeitarchiviert werden soll, den Begriff *Objekt* zumeist für dingliche Objekte (gedruckte Werke, physische Gegenstände, ...), *Werk* für ein Dokument zusammen mit seiner Einbettung (Übertragungen, Fassungen, Erläuterungen, Erkenntnisse über das Dokument etc.), soweit sie für das Verständnis und die Lesbarkeit in der Zukunft wesentlich erscheinen.

Wir unterscheiden zwischen 'Werken' für die Langzeit-Archivierung, die also auf Grund ihrer Aufbereitung bei der Speicherung und der anschließenden andauernden Einbettung bereit gehalten werden zum Gelesen und Verstanden werden, und 'Rohdaten', die in großen Mengen z. B. bei Forschungsexperimenten, als Wetterdaten etc. anfallen, und möglicherweise für lange Zeit gespeichert werden müssen, damit sie bei späteren Auswertungen, auch z. B. wegen neuartiger Fragestellungen, zur Verfügung stehen. Rohdaten sind also ohne Aufbereitung gemäß einer erst später aufkommenden Fragestellung und dann entsprechende Auswertung noch keine von Menschen nutzbare Erkenntnis. Sie müssen natürlich *mit* allen Schlüsseln und Codes und Experimentbeschreibungen abgespeichert werden, damit sie später entschlüsselt werden können. Über die spezifischen und besonderen Probleme der Langzeit-Speicherung von Rohdaten verweisen wir auf die parallele *nestor*-Expertise<sup>53</sup>.

---

<sup>52</sup>Im englischen Sprachraum wird klarer zwischen *archives* für ein klassisches Archiv und *archive* für einen Bestandteil als Teil eines Archivs, oder eine Registratur (Schriftgutablage) einer Einrichtung unterschieden.

<sup>53</sup>Thomas Severiens, Eberhard R. Hilf; *Langzeitarchivierung von Rohdaten*; nestor-Materiale 6; November 2005; <http://nbn-resolving.de/urn:nbn:de:0008-20051114018>

## 3.4 Vorgehen und Ergebnisse der Fragebogenaktion

### 3.4.1 Idee, Strategie, Durchführung, Rücklauf

Kern der Arbeiten zu dieser Expertise sollte ursprünglich eine Fragebogenaktion sein.

Die Hoffnung war, durch genaue Kenntnis der Planungen in anderen Ländern Anregungen für die zu entwickelnde nationale LZA-Policy Deutschlands zu bekommen.

Der Fragebogen wurde zusammen mit Experten der *Niedersächsischen Staats- und Universitätsbibliothek SUB Göttingen* zusammengestellt. Als außenstehender Experte wurde hierzu der Wissenschaftsjournalist Richard Sietmann hinzugezogen. Zur Abgrenzung von der PREMIS-Fragebogenaktion<sup>54</sup> wurden hier keine technischen Fragen zu *institutionellen* Policies in den Vordergrund gestellt, sondern es wurde nach dem Stand auf dem Weg zu einer *nationalen LZA-Policy* gefragt.

Hierzu wurde systematisch, Land für Land, teilweise ausgehend von den Rednerlisten einschlägiger Tagungen, nach Informationen über entsprechende Projekte und Vorhaben in anderen Ländern gesucht, um diese anschließend zu kontaktieren. Es wurde dann ein Fragebogen entwickelt, der nach dem Stand der Entwicklung und, sofern vorhanden, nach Erfahrungen mit der Langzeit-Archivierung fragte. Da dieser Fragebogen nicht nur an Projekte geschickt wurde, die aus den Recherchen hervorgingen, sondern auch weltweit an nationale Bibliotheken, Archive und National-Museen wurde dort auch nach weiteren Aktivitäten gefragt, um somit den Kreis der bekannten Projekte und Vorhaben zu erweitern.

Inhaltlich hatte der Fragebogen<sup>55</sup> sechs Kernkomplexe:

1. Gibt es eine *ationale LZA-Policy*? Ist sie abrufbar, wer hat sie geschrieben, welche Verbesserungen werden gesehen? Wenn noch keine nationale LZA-Policy existiert, welche Aktivitäten gibt es dazu, welches sind die Beteiligten, wie ist der Stand?
2. Gibt es im Institut des/der Befragten selbst eine LZA-Policy? Hat diese von den Erfahrungen Anderer profitiert? Wie sind die Erfahrungen?
3. Gibt es eine *Institution*, die für die nationale LZA-Policy zuständig ist? Welches sind deren Aufgaben, wer zahlt, gibt es formale Vereinbarungen, welche Metadaten werden benutzt?
4. Welches sind die Dokumenten-Auswahl-Kriterien? Wer wählt aus, welche Kriterien wurden vereinbart, welche Formate werden erlaubt?
5. Wie ist der Umfang der bisher auf Grund einer nationalen LZA-Policy archivierten Materialien? Wer darf darauf zugreifen, wer zahlt, welche Zeitspannen bis zur Lieferung werden akzeptiert?

<sup>54</sup><http://www.oclc.org/research/projects/pmwg/>

<sup>55</sup><http://www2.hu-berlin.de/nestor/questionnaire/q1.php>



6. Wie groß ist die Nutzung der Langzeit-Archive, seit wann gibt es das Archiv, wie sind die Nutzungserwartungen?

Es wurden gezielt 320 Fragebögen ausgesandt, sowie zusätzlich eine offene Verbreitung über einige große Email-Verteiler, die mindestens 41 Länder erreicht haben, vorgenommen.

Der Fragebogen wurde zusätzlich von der DDB über ihren eigenen Verteiler an Nationalbibliotheken und Nationalarchive<sup>56</sup> versandt.

Der *Rücklauf* war enttäuschend gering und die Antworten waren oft nur kurssrisch. Es wurden nur 17 auswertbare Fragebögen zurückgesandt. Damit verbot sich jede ernsthafte statistische Auswertung.

Alle zurückgesandten Fragebögen wurden vereinbarungsgemäß dem *nestor*-Partner Humboldt-Universität zur Ablage in einer nur projektintern zugänglichen Datenbank zugesandt.

Die Hintergründe der geringen offiziellen Reaktionen erschlossen sich aus daraufhin durchgeführten vertiefenden Telefonaten: Die angesprochenen national *Verantwortlichen* wollten nur ungern möglicherweise öffentlich werdende Statements über noch im Fluss befindliche bzw. noch nicht beschlossene eigene nationale Policy-Planungen abgeben, um den lokalen politischen Prozess nicht zu stören, obwohl den Institutionen im Fragebogen die vertrauliche Behandlung und nur anonymisierte Verwendung in dieser Expertise zugesichert worden waren.

Andere befragte *Experten* nationaler Bibliotheken durften keine Statements zu noch nicht endgültig beschlossenen nationalen Policies herausgeben, ohne diese vorher einem längeren (und ungewissen) hausinternen Freigabe-Prozess unterworfen zu haben. Teilweise wurde der Fragebogen auch als „Spam“ verworfen.

Die Rückläufe wurden Ende Mai 2004 zu einem ersten Resultat zusammengestellt und auf einem Workshop des Projektes *nestor* am 1. und 2. Juni in Göttingen intensiv diskutiert<sup>57</sup>. An ihm nahmen Experten der Informationsproduzenten, der Archivnutzer, der Staatsarchive, der wissenschaftlichen Museen, der Bibliotheken sowie aus der Wissenschaft teil. Der speziell bundesdeutsche föderale Aspekt wurde hierbei besonders beleuchtet. Die Ergebnisse dieser Diskussion sowie zahlreicher Gespräche mit Projekten und Institutionen weltweit, die auf dem Feld der Entwicklung von Strategien und Konzepten einer nationalen Langzeit-Archivierung des kulturellen und wissenschaftlichen digitalen Gutes aktiv arbeiten, haben zu der vorliegenden Expertise beigetragen und diese wesentlich geprägt.

Um dennoch eine sinnvolle Erfassung und tiefere Auswertung der international laufenden und abgeschlossenen Projekte umzusetzen, wurden einige Leiter oder von ihnen benannte Ansprechpartner der einschlägigen Projekte kontaktiert und, soweit

---

<sup>56</sup>Die Subskribenten sind den Autoren dieser Expertise nicht bekannt.

<sup>57</sup>Eberhard R. Hilf; *Anforderungen aus der Sicht eines Wissenschaftlers*; *nestor*-Workshop Juni 2004; Göttingen; <http://www.isn-oldenburg.de/~hilf/vortraege/nestor/nestor04-hilf.html>

möglich, im direkten Gespräch oder telefonisch befragt, wobei sich der ursprüngliche Fragebogen als guter Leitfaden herausstellte. Die Telefonate und Gespräche, oft auch am Rande internationaler Konferenzen und Workshops, erwiesen sich als besonders wertvoll: In einigen Fällen wurden wesentliche Diskrepanzen zwischen dem Online und dem in Publikationen dargestellten Stand einerseits sowie dem tatsächlichen fortgeschritteneren Stand und den entwickelten Strategien andererseits offenbar, die man nicht schriftlich in einem Fragebogen fixieren wollte.

Im folgenden Kapitel wird ein Überblick über den aktuellen Stand der Entwicklung nationaler Policies zur Langzeit-Archivierung in anderen Ländern gegeben.

Dabei heben wir insbesondere hervorstechende Ideen und Umsetzungen heraus, die für eine bundesdeutsche LZA-Policy von Interesse sein könnten. Die Auswahl der Länder resultiert aus der Quellenlage, dem Rücklauf der Fragebögen und den Gesprächspartnern.

### 3.4.2 Der Fragebogen

Die Rückantworten aus der Fragebogenaktion aus dem Ausland und die Hintergrundtelefonate werden entsprechend dem Wunsch der meisten Einsender und der Zusicherung auf dem Fragebogen selbst jeweils beispielhaft für einige Länder zusammengefasst und eingearbeitet, dabei aber anonymisiert.

Die Auswertung wird ergänzt durch Einzelanalysen von nationalen Institutionen und den Ansätzen zu nationalen Policies, sofern vorhanden.

Der Fragebogen, und auch eine Analyse der vorhandenen spärlichen offiziell herausgegebenen Quellen, war ein ungeeigneter Ansatz, wenn man für die Vorgehensweise in Deutschland etwas lernen will.

Die vorliegende Expertise ist daher keine deduktiv orientierte Arbeit, sondern fasst die Ergebnisse der Fragebogenaktion, zahlreicher Gespräche mit Experten, die in anderen Ländern auf diesem und verwandten Gebieten arbeiten, sowie die eigene Betrachtung von Projekten und Implementationen im Ausland zusammen, um den nationalen Weg zu einer LZA-Policy aktuell und ohne zu großen Zeitverzug zu unterstützen.

Die Einschätzung einer langfristigen Entwicklung in die ferne Zukunft ist natürlich eine schwierige Herausforderung: Ein einfaches Subsummieren und Mitteln der Aussagen vieler Experten der Gegenwart und das Studium von Quellen der Vergangenheit verstellt eher den Blick. Besser wäre es, die Vielzahl der Einschätzungen unabhängig zu erhalten, um einen Sinn für die Bandbreite zu bekommen.

### 3.4.3 Archivierung und Extrakt von Fragebogen-Antworten

Auf Grund technischer Anlaufschwierigkeiten des externen Fragebogen-Archivs wurden nur neun der 26 zurückgesandten und abgespeicherten Fragebogen in der externen

nestor Arbeitsbereiche - Mozilla

File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop <http://www2.hu-berlin.de/nestor/questionnaire/q1.php> Search Print

Start | Login | Logout | Kontakt | English

**nestor** Kompetenznetzwerk  
Langzeitarchivierung  
-Kommunikationsplattform-

Kommunikationsplattform

- Mailinglisten
- Arbeitsbereiche
  - Intern
  - Beirat
  - AG Repositories
  - AG Multimedia
- Fragebögen
- Statistiken

Kompetenznetzwerk

- Home

Informationsplattform

- Informationsdatenbank
- Wer Wo Was
- Reviews
- Projekte
- Kalender
- Aktuelles
- Newsletter

Links

- Padi
- Forum Bestandserhaltung
- Netzwerk Mediatheken

gefördert vom  
Bundesministerium  
für Bildung  
und Forschung

### Questionnaire to develop a National Long-Term Preservation Policy

This survey is part of NESTOR, a project aimed at developing criteria and a workflow for the long-term preservation of scientific digital data in Germany. It is funded by the German Ministry for Education and Research. We would appreciate it very much if you could spare a few minutes to respond to the questionnaire below. The outcome from this survey will be documented and published online. We will notify you by email as soon as the report becomes available.

The fields marked with an asterisk (\*) are mandatory.

\* Your Name:

\* Your Country:

\* Name of your institution:

URL of your institution (<http://...>):

URL of document archive of your institution (<http://...>):

\* Whom do you recommend to contact for further information?

1. Does a **national** long-term preservation strategy exist in your country? (as far as you know or are responsible)      yes  no

1.1. If your country agreed on a national long-term preservation strategy:

1.1.1. Is this strategy written down and available online? URL (<http://...>):

1.1.2. Who developed and wrote down the strategy? (Governmental Commission, National Library, National Archive, ...):

1.1.3. Did the commission use experiences from other countries?

1.1.4. Which advantages and problems do you see with this national strategy?

1.1.5. What would you prefer to be changed or could be implemented in a better way?

Abbildung 3.1: Auszug aus dem Fragebogen zur LZA-Policy

Datenbank langzeit-archiviert.

Sie kamen aus fünf verschiedenen Ländern. In keinem dieser Länder existiert eine nationale LZA-Policy, alle Befragten wünschen aber eine solche.

Die Aktivitäten sind überwiegend nicht auf Regierungsebene angesiedelt, sondern gehen von den nationalen Bibliotheken und einschlägigen wissenschaftlichen Projekten aus.

Als Status des Fortschrittes auf dem Weg zu einer nationalen LZA-Policy wird stets „early discussion“ vermerkt, was für einige Länder sicher stark untertrieben ist, aber die Befürchtung widerspiegelt, vorzeitig Informationen herauszugeben.

Als ein Hauptproblem wird die internationale Vereinbarung von spezifischen LZA-Metadaten genannt.

Die Dokumentenauswahl für die Langzeit-Archivierung wird zumeist von Bibliothekaren bzw. den archivierenden Institutionen durchgeführt.

Als Formate werden *XML*, *LaTeX*, aber auch *Postscript*, *Powerpoint*, *WORD* und *RTF* genannt.

Als derzeitiger Umfang der Archive werden jeweils einige Millionen Dateien genannt, die meist durch Web-Harvesting gewonnen wurden. Diese sind zu 98% rein digital archiviert.

Der Zugriff von außen ist bisher zumeist nur Experten erlaubt, solange rechtliche Regelungen fehlen, die der Offenlegung vorausgehen müssten.

**Die Schaffung international kompatibler nationaler LZA-Policies und Gesetzgebungen ist die vorrangige Herausforderung, um die Langzeit-Archivierung von Werken der Wissenschaft, der Kultur und der akademischen Lehre im weltweiten Verbund zu planen und dann zu realisieren.**

### 3.5 Aktivitäten zur Langzeit-Archivierung im Ausland

Die Entwicklung zu nationalen LZA-Lösungen ist unter den Experten in allen Staaten in der aktuellen Diskussion, besonders in deren jeweiligen nationalen Archivierungs-Institutionen.

Über die technischen Anforderungen ist eigentlich alles vielfach gesagt. Insbesondere sei das DFG-Projekt 1999-2000 zur Langzeit-Archivierung digitaler Publikationen mit der Zielsetzung, ein „organisatorisches und technisches Konzept für die Langzeit-Archivierung und die Bereithaltung digitaler Publikationen in einer Universalbibliothek“ zu erstellen, genannt<sup>58</sup>.

International gibt es eine gute vergleichende Übersicht über das Management vernetzter digitaler Publikationen<sup>59</sup> von Elizabeth Martin (*National Library of Canada*).

Die organisatorischen und politischen Wege zu einer nationalen LZA-Policy sind z. T. landesspezifisch und daher nur bedingt auf Deutschland übertragbar, so dass sich nur mit einem gewissen Abstand aus den Schwierigkeiten anderer Länder lernen lässt. Der erfolgsversprechendere Weg ist die extensive und intensive Kontaktaufnahme mit den Aktiven in anderen Ländern, um aktuelle Erfahrungen auszutauschen, um gemeinsam in internationalen Gremien zu arbeiten, um die internationale Vernetzung der Langzeit-Archivierung (Redundanz, Sicherheit) zu realisieren und so die nationalen Policy-Anforderungen zu extrahieren.

Im Folgenden werden daher Einblicke in den aktuellen Stand der Diskussion zur Langzeit-Archivierung einiger Länder gegeben, soweit sie für die anstehende Diskussion zu einer LZA-Policy in Deutschland interessant erscheinen.

In allen Ländern wird die digitale Archivierung (*preservation*) mehr oder weniger energisch angegangen, es gibt dazu Gesetze und (zumeist aus den eigenen Mitteln der beteiligten Bibliotheken durch Umschichtung gewonnene) Finanzen. Übersichten gibt es durch eine ganze Reihe von Studien<sup>60</sup>.

Dabei haben verschiedene Länder deutlich unterschiedliche Vor- und Nachteile, sehen sich auch je nach Größe des Landes und seiner industriellen Stärke unterschiedlichen Herausforderungen gegenüber: Sehr kleine Staaten sind auf einen engen Kontakt zu einem größeren Land angewiesen, um an der technischen Entwicklung zu partizipieren. Mittelkleine aber industrielle Staaten haben den großen Vorteil überschaubarer politischer Strukturen, falls sich diese nutzen lassen. Große Industriestaaten sind in der technologischen Entwicklung und Forschung meist am weitesten, haben aber das Problem starker, komplexer und gewachsener Strukturen, einer sehr breiten Palette professioneller Institutionen, die sich nur schwer neu ordnen und bündeln lassen. Hinzu kommen die Herausforderungen durch die sehr unübersichtliche politische Aufgaben-

---

<sup>58</sup>Abschlussbericht: Peter Rödiger, Uwe Borghoff; *Langzeit-Archivierung digitaler Medien*; <http://www.informatik.unibw-muenchen.de/reports/reports/2002-02.pdf>

<sup>59</sup>Elizabeth Martin; *Management of networked electronic publications*; <http://www.collectionscanada.ca/obj/r7/f6/r7-100-e.rtf>

<sup>60</sup>z. B. CLIR 2003, a.a.O.

und Zuständigkeitsverteilung (z. B. Bund/Länder, Staatenbund, Commonwealth).

Eine solide, den bekannten Anforderungen wirklich gerecht werdende national organisierte Langzeit-Archivierung, die entsprechend nationale staatliche Mittel, eine nationale Organisation und eine nationale LZA-Policy voraussetzt, fehlt aber bisher, selbst in Australien oder in den USA.

Insofern würde ein energischer nationaler Einstieg in Deutschland nicht nur den unausweichlichen nationalen „digitalen Gedächtnisschwund“ der Jetzt-Zeit verkürzen, sondern auch eine Spitzenposition in der internationalen Diskussion erreichbar erscheinen lassen.

### 3.5.1 Australien, Beispiel eines seit langem aktiven Industriestaates

Australien hat durch seine großen nationalen Bibliotheken und Archive eine langjährige umfangreiche Erfahrung auf allen Feldern der digitalen Langzeit-Archivierung.

Es gibt bisher keine nationale *Preservation Policy*, geschweige denn eine nationale *long-term Preservation Policy*. Dies hat politische Gründe: Australien ist politisch als *Commonwealth of Australia* organisiert. Es ist ein föderaler Bundesstaat mit weitgehend eigenständigen Staaten.

Aus bundesdeutscher Sicht lohnt sich daher besonders ein Blick auf die australischen Projekte, da die föderalen Strukturen den deutschen verwandt sind.

Eine gute synoptische Zusammenstellung von Links zu für die Langzeit-Archivierung relevanten institutionellen Policies und entsprechenden Berichten findet sich bei dem australischen *E-Government Resource Centre*<sup>61</sup>.

Die australische Nationalbibliothek NLA (*National Library of Australia*<sup>62</sup>), und das NAA<sup>63</sup>, die *National Archives of Australia*, sehen sich besonders verantwortlich für die professionelle Entwicklung und Vorbereitung der digitalen Archivierung. Sie haben hier bereits Wesentliches durch Analysen, Reports und durch pragmatische Teilprojekte geleistet.

In Australien gibt es bisher keine Pflichtexemplarregelung für digitale Publikationen, auch wenn dies derzeit intensiv diskutiert wird.

Australien ist weltweit eines der Länder mit der höchsten Quote von Internetnutzern in der Bevölkerung, nach den Vereinigten Staaten und Singapur. Die großen Entfernungen zwischen den Städten haben von Anfang an zu einem entsprechenden Druck geführt, so viele staatliche Dienste wie möglich Online anzubieten, digitale Medien adäquat langfristig zu archivieren und Online zugreifbar zu erhalten.

Australien hat bereits seit 1996 vertikal für alle staatlichen Dienste (E-Government)

---

<sup>61</sup><http://www.egov.vic.gov.au/Australia/Aust-Egov-Reports.htm>

<sup>62</sup><http://www.nla.gov.au>

<sup>63</sup><http://www.naa.gov.au>

Metadaten-Profilen definiert und durchgesetzt, also für das Commonwealth insgesamt, für seine Staaten und deren Städte.

Die NLA hat seit 1996 mit dem PANDORA<sup>64</sup> Projekt *Preservation*<sup>65</sup> eine Sammlung aufgebaut, um daran die Techniken und Organisationsformen der Langzeit-Archivierung zu entwickeln und zu implementieren. So wurden bereits Tests der Dateiformat-Migration, der Hardware-Emulation sowie Techniken zur Archivierung des Web-Content erprobt und weiterentwickelt. Es wurde mit der TRIM-Software ein Softwaresystem etabliert, das veraltete Versionen von Textdarstellungsdateien weitestmöglich anzeigen können soll.

Die institutionelle *Preservation Policy*<sup>66</sup> der Nationalen Bibliothek Australiens (NLA) fasst den Erkenntnisstand zusammen.

Die Auswahlkriterien<sup>67</sup> der NLA enthalten sowohl thematisch- wie qualitätsorientierte, aber vom Quellentyp unabhängige Aufnahmekriterien sowie eine sorgfältige Liste von Ausschlusskriterien<sup>68</sup>.

Die NLA sieht die Notwendigkeit, international zu kooperieren, um nachhaltig Methoden der Archivierung digitaler Objekte voranzutreiben.

Die Langzeit-Archivierung wird *durch öffentliche Mittel finanziert*, diese werden jedoch bisher durch interne Umschichtung innerhalb der Nationalen Bibliothek bzw. der beteiligten Bibliotheken aufgebracht.

Der *Workflow* wird durch die Bibliothek selbst entwickelt und durchgeführt.

In PANDORA wird die *Auswahl der Dokumente* durch Experten in- und außerhalb der NLA vorgenommen.

Die Kriterien sind durch das PANDORA-Projekt<sup>69</sup> dokumentiert. Es wird im Wesentlichen „roharchiviert“, d. h. alle, auch proprietäre Formate sind noch zugelassen, obgleich man sich der Problematik bewusst ist. Der *Umfang* ist bereits auf 3,3 Millionen Dokumente angewachsen, mit ca. 170 GByte Zuwachs pro Jahr. Der *Zugang ist generell frei* im Rahmen des gesetzlich jeweils Zulässigen. Die Archivierung startete 1996, mit zur Zeit etwa einer Million *Anfragen* je Jahr.

Insgesamt werden durch die NLA die organisatorisch-politischen Herausforderungen einer nationalen LZA-Policy, denen wir uns in Deutschland stellen sollten, derzeit noch umgangen. Sie ist zwar die nationale Bibliothek. Es fehlt jedoch eine nationale *long-term Preservation Policy* der Regierung des Commonwealth selbst; diese wird jedoch in der nationalen Diskussion gefordert.

---

<sup>64</sup><http://www.nla.gov.au/policy/plan/pandora.html>

<sup>65</sup><http://pandora.nla.gov.au/activities.html#preservationmet>

<sup>66</sup><http://www.nla.gov.au/policy/digpres.html>

<sup>67</sup>*Selection Guidelines for Archiving and Preservation*; National Library of Australia; <http://pandora.nla.gov.au/selectionguidelines.html>

<sup>68</sup>a. a. O.: 3.7.

<sup>69</sup><http://www.pandora.nla.gov.au>

Obgleich die Beschäftigung mit den einzelnen LZA-Herausforderungen die Qualität des jetzigen „institutionellen Archivs mit nationalem Blick“ der NLA in der Zukunft weiter erhöhen wird, liegt der Schwerpunkt noch nicht auf der echten Langzeit-Verfügbarkeit, sondern auf einem pragmatischen Beginn und dem Sammeln von Erfahrungen. Aus der Sicht der fernen Zukunft könnten also beachtliche Teile des Archivs nicht mehr lesbar werden und damit für immer verloren gehen.

### 3.5.1.1 PADI: Preserving Access to Digital Information

Im Kooperationsprojekt PADI<sup>70</sup>, *Preserving Access to Digital Information*, wird ein Portal zu allen Aspekten der Langzeit-Archivierung aufgebaut und betrieben.

PADI ist im wesentlichen ein auch aktuell gepflegtes Portal (*Subject Gateway*). Es enthält Sammlungen von Links und Informationen über alle Aspekte der Langzeit-Archivierung, wie z. B. *Preservation Metadata*, *Identifiers* und *Standards*.

PADI wird von der *National Library of Australia* (NLA) betrieben und ist als internationales Kooperationsprojekt angelegt.

Die NLA hat ein *Memorandum of Understanding* mit der *Digital Preservation Coalition*<sup>71</sup> (DPC) vereinbart zu den Themen: Sammlung aktueller Links; Einspeisung der Informationen aus PADI zu Aktivitäten in UK; gemeinsame Erstellung von Review-Texten; Mitwirkung an Auswahl und Archivierung (*Safekeeping Project*<sup>72</sup>). In ihm werden durch eine internationale Gruppe von LZA-Experten für die Langzeit-Archivierung wichtige Dokumente gesammelt.

Das Projekt sammelt herausragende Links zum Thema *digitale Langzeit-Archivierung*. Sein Start war 2001. Es soll ein verteiltes Netzwerk von Institutionen im In- und Ausland, die sich mit Langzeit-Archivierung befassen, gebildet werden. Dabei sollen die an den verschiedenen Orten/Staaten eingesetzten technischen und organisatorischen Strategien erfasst und verglichen werden. Auf formale Verträge wird verzichtet.

Der Kontakt zum *Safekeeping Project* geschieht über die *National & International Preservation Activities* (NIPA<sup>73</sup>) der *National Library of Australia*.

Safekeeping wird über das CLIR<sup>74</sup> *Council on Library and Information Resources*, USA, mitfinanziert.

**Empfehlung:** Die aktive Beteiligung der deutschen LZA-Gruppe an den Projekten *Safekeeping* und *PADI* sollte weiter unterstützt und verbreitert werden.

Partner von PADI sind u. a. DPC (*Digital Preservation Coalition*) und das EU-Projekt Erpanet<sup>75</sup> (*Electronic Resource Preservation and Access Network*).

<sup>70</sup><http://www.nla.gov.au/padi/>

<sup>71</sup><http://www.dpconline.org/graphics/index.html>

<sup>72</sup><http://www.dlib.org/dlib/january02/berthon/01berthon.html>

<sup>73</sup><http://www.nla.gov.au/preserve/pres.html>

<sup>74</sup><http://www.clir.org>

<sup>75</sup><http://www.erpanet.org/about.php>



PADI hat ein *Memorandum of Understanding* mit der SUB Göttingen unterzeichnet, das den Metadaten-Austausch zwischen dem *PADI Subject Gateway* und der *nestor Informationsdatenbank* ermöglicht. Der Datenaustausch zwischen den beiden Subject Gateways wird zur Zeit vorbereitet, und wird später dazu führen, das in beiden Datenbanken ein umfassender Datenbestand nachgewiesen werden kann.

PADI hat ein *International Advisory Board* mit Vertretern der wesentlichen nationalen Bibliotheken und LZA-Institutionen berufen. Aus Deutschland ist Hans Liegmann von der DDB, der Projektleiter von *nestor*, vertreten.

### 3.5.1.2 PANDORA Archive

Die NLA betreibt mit dem PANDORA Archive, *Australia's Web Archive*, bereits ein Langzeit-Archiv, in das zur Zeit jährlich etwa 21 Millionen Dateien von zusammen etwa 680 GB Volumen aufgenommen werden. Bisher hat PANDORA etwa eine Million registrierte Nutzer. Alle LZA-Aktivitäten werden nur indirekt über verschiedene öffentliche Haushalte finanziert.

Eine für die Überlegungen zu einer nationalen LZA-Policy in Deutschland wesentliche Leitidee ist in PANDORA formuliert:

**Parallel zum pragmatischen Beginn müssen sowohl das System wie das Konzept der Langzeit-Archivierung weiterentwickelt werden.**

Die NAA und NLA planen derzeit, ihre Aktivitäten zu bündeln und zukünftig gemeinsam ein verteiltes Netzwerk von Archiven zu betreiben.

In PANDORA wurde Mitte 1996 mit einer ersten Auswahl von Dokumenten begonnen. Die Archivierung selbst begann dann Ende 1996.

PANDORA hat zehn Partner-Institutionen, primär das *Australian Institute of Aboriginal and Torres Strait Islander Studies*, das *Australian War Memorial* und die *National Library of Australia*, aber auch die Staatsbibliotheken der australischen Bundesstaaten.

Als nationales Gesetz gab *The National Library Act* 1960 den Auftrag, eine nationale Kollektion von „Bibliotheksmaterial“ sowie eine Sammlung aller Quellen im In- und Ausland, die von Australien oder den Australiern handeln, für die Zukunft zu archivieren.

Als Auswahlkriterium ist das Sammeln von Australica für die australische Regierung natürlich naheliegend, während die Motivation für ausländische National-Bibliotheken, ausgerechnet Australica zu sammeln, natürlich nicht gleichrangig gegeben ist. Der Sammelauftrag „Kollektionen von vorhandenem Bibliotheksmaterial“ ist auf die Bewahrung des Vorhandenen ausgerichtet. Es fehlt eine klare Formulierung der für die langfristige Planung der Langzeit-Archivierung in der Zukunft notwendigen Kriterien. Insofern ist der *National Library Act* kein Ersatz für eine *National long-term Preservation Policy* digitaler Werke.

Die NLA hat als Titel und ist die Führungs-Bibliothek (*Leading Library*) für die anderen Bibliotheken in Australien.

Die gesetzlich geregelte Hinterlegung von zu archivierenden Dokumenten enthält nicht die digitale Fassung. Daher ist die NLA frei, was sie digital archiviert, sofern es sich nicht um Australica handelt. Ihr eigener pragmatischer Ansatz ist dabei, mit der Langzeit-Archivierung früh begonnen zu haben, um im realen Betrieb Erfahrungen zu sammeln.

Die Auswahl der digitalen Dokumente wird also pragmatisch vorgenommen (proportional zur technischen Leistungsfähigkeit). Die Aufnahme eines Werkes geschieht auf Antrag. Die E-Dokumente werden manuell auf ihre Archivierbarkeit geprüft. Die Dokumente werden öffentlich zugänglich gehalten.

Die Arbeit an notwendigen Positionen für eine nationale australische LZA-Policy<sup>76</sup> wurde schon frühzeitig begonnen.

Die NIAC (*National Initiatives and Collaboration*) der NLA hat bereits seit 1995 an einem Statement zu Prinzipien der Langzeit-Archivierung und des langfristig gesicherten Zugangs zu digitalen Objekten gearbeitet<sup>77</sup>. Schon darin wurde eine das Commonwealth und die Staaten einbeziehende, nationale LZA-Policy mit geeigneten Regelungen und Gesetzen gefordert<sup>78</sup>.

Das Statement nennt sieben Prinzipien, die hier zusammengefasst werden:

1. Eine stabile Langzeit-Archivierung ist nur realisierbar durch eine Kooperation und Interaktion aller Beteiligten (Autoren, Verlage, Designer, Archivare).
2. Schon die Autoren müssen durch geeignete Rechtesetzungen dafür sorgen, dass der spätere Zugang erlaubt bleibt<sup>79</sup>.
3. Ein nationales Netzwerk der Archive, Sammelbibliotheken und anderer Informationsprovider muss mit klaren Verantwortlichkeiten über formelle Vereinbarungen gebildet werden.
4. Digitale Dokumente sollten nur solange gespeichert werden, wie geeignete Fach-Gutachter der Meinung sind, dass sie für zukünftige Generationen relevant sein sollten.
5. Es sollte keine Zensur geben.
6. Die LZA-Strategien sollten kontinuierlich an den Stand der Forschung angepasst werden.

---

<sup>76</sup>National Strategy for Provision of Access to Australian Electronic Publications; A National Library of Australia Position Paper; <http://www.nla.gov.au/policy/paep.html>

<sup>77</sup>Statement of Principles for the Preservation of and Long-Term Access to Australian Digital Objects; <http://www.nla.gov.au/preserve/digital/princ.html>

<sup>78</sup>a. a. O.: Principle 7: Role of Government.

<sup>79</sup>Dies mag als ein früher Ruf nach *Open Access* angesehen werden.

7. Es sollte eine nationale LZA-Policy entwickelt und durchgesetzt werden.

Ebenfalls bereits 1995 hat das *Australian Cultural Ministers Council* in seiner *National Conservation and Preservation Policy for movable cultural heritage*<sup>80</sup> nationale LZA-Policy Statements gegeben.

**Ein Beispiel: Der Staat New South Wales:** Der Bundesstaat *New South Wales* ist Mitglied des *Commonwealth of Australia*, des Bundes der Provinzstaaten (Bundesländer). Er möge hier als ein Beispiel eines Bundeslandes von Australien gelten, das sich an den Plänen für eine LZA beteiligt.

Bereits in den Jahren vor 1998 wurde in der Landesregierung die Behörde OICT<sup>81</sup>, *Office of Information and Communication Technology*, aufgebaut.

Gesprächspartner eines der Autoren dieser Expertise<sup>82</sup> waren 1998 Dr. Dita Bartels, Deputy Director, *International Development, University of New South Wales* und Ken Bullock, Experte für Metadaten.

Das OICT verfolgt neun Strategie-Linien<sup>83</sup>. Wichtig ist hier die Strategie 3: *Strategic Value of Information* mit den auch für die Entwicklung einer nationalen LZA-Policy in Deutschland wesentlichen Unterthemen: Sicherheit der Informationssysteme; Organisation; Authentizität; Zugangskontrolle; Integrität der Dokumente; Vertraulichkeit; Verfügbarkeit; Audit.

Dies entspricht den inzwischen allgemein diskutierten Themen. Gesprächspartner von uns waren Ian Brandt, *Government Online Project*, und Jo Hein, *Fedlink (Government Intranet)* sowie der Kanzler der ANU (*The Australian National University*<sup>84</sup>) Satis Arnold. Ziel war, den Stand und die Planungen der Archivierung der Webpräsenz der Universität kennenzulernen.

Eine *Audit Checklist*<sup>85</sup> zum Information Management wurde vom Staat New South Wales entwickelt und im praktischen Einsatz angepasst.

### 3.5.1.3 The International Information Consortium Plan

Um das Jahr 1998 herum waren die australischen Behörden, die Regierung (genauer: ihre Behörde OGIT, später OICT genannt), ANU und einige australische Staaten, vor allem New South Wales, weit offen für eine enge Kooperation mit Deutschland auf dem Gebiet der Metadaten, des E-Government und der Langzeit-Archivierung.

<sup>80</sup><http://www.nla.gov.au/preserve/natpol.html>

<sup>81</sup><http://www.oict.nsw.gov.au/content/2.0.Pol-Gui.asp>

<sup>82</sup>Eberhard R. Hilf; Reise 1998 auf Einladung der australischen Regierung.

<sup>83</sup><http://www.oict.nsw.gov.au/content/1.2.2.Blueprint-Map.asp>

<sup>84</sup><http://www.anu.edu.au>

<sup>85</sup><http://www.oict.nsw.gov.au/Guidelines/4.3.12.g-IM-Audit.asp#app1>

Der *Director of Information Management of OGIT* der australischen Regierung, C. Palmer, hatte Deutschland mehrfach besucht und u. a. auf dem jährlichen Workshop der IuK 1998 in Hamburg<sup>86</sup> vorgetragen. Das Angebot an das anwesende BMBF zu einem Gespräch wurde aber von diesem negativ beschieden. Anschließende Briefe ihrer Behörde an diverse Ministerien (BMA, BMI, BMBF) zu einer Kooperation auf den Gebieten E-Government und Langzeit-Archivierung wurden schließlich sogar in der Inter-Staatssekretärs-Kommission der Bundesministerien behandelt, blieben aber letztlich ohne Ergebnis für Aktivitäten in Deutschland<sup>87</sup>. Daher blieb auch der Vorschlag eines *Global International Information Consortium*<sup>88</sup> folgenlos.

Vergleichen wir nun die vielfältigen tatsächlichen vergangenen und gegenwärtigen Aktivitäten in Australien mit den offiziellen Antworten der Experten im aktuellen Fragebogen:

- Gibt es eine nationale LZA-Policy?  
Antwort: „No“.
- Gibt es Aktivitäten, eine nationale LZA-Policy zu entwickeln?  
Antwort: „No activities in place“.
- Wie ist der Status der Diskussion?  
Antwort: „Some early discussions“ ,
- usf.

Der Grund für diese Diskrepanzen ist natürlich, dass die NLA und die NAA zwar all die genannten Aktivitäten betreiben, es aber eben keine offiziellen Aktivitäten des Bundes sind.

### 3.5.2 Sehr kleine Staaten

Kleine Länder nehmen - und das wurde auch aus der Umfrage deutlich - eine Sonderrolle im sich bildenden internationalen Netzwerk zur Langzeit-Archivierung ein. Ihre nationale LZA-Policy muss sicherstellen, dass sie in die internationale Entwicklung eingebunden werden und dass bei den bilateralen Verträgen mit anderen Ländern zur exterritorialen Langzeit-Archivierung der Dokumente aus dem eigenen Lande die spezifischen Anforderungen des eigenen Landes berücksichtigt werden. Auch kann

<sup>86</sup><http://web.archive.org/web/20040315031807/http://www.iuk-initiative.org/iuk-hh/foils-palmer.pdf>

<sup>87</sup>Die damalige Leiterin der Kommission, Frau Brigitte Zypries, wurde zur Bundes-Justizministerin ernannt.

<sup>88</sup>E. R. Hilf; *Future Professional Services in Scientific Information by joint R and D of Universities, Libraries and Industrial Providers with report on the German Government Digital Library programme Global-Info*; Keynote International Address at AIC Worldwide Conference, Sydney; 1998; <http://www.isn-oldenburg.de/~hilf/vortraege/australia/AiC-1998/>

es besondere Dokumententypen und -Klassen geben. Als praktische Hilfe zur Kontaktaufnahme ist im Anhang eine Liste der Web-links zu möglichst vielen Nationalbibliotheken auch sehr kleiner Staaten angegeben.

### 3.5.3 Finnland als Beispiel eines mittelkleinen Industriestaates

In Finnland gibt es seit 2004 Bemühungen der Nationalbibliothek, zusammen mit dem Nationalarchiv, dem *Center for Scientific Computing* und dem Finnischen Filmarchiv ein Langzeit-Archiv aufzubauen.

Stand der Entwicklung ist nach den Aussagen in zwei Fragebögen und nach einem persönlichen Gespräch die frühe Diskussionsphase. Die Nationalbibliothek hat jedoch<sup>89</sup> die Aufgabe, bis 2006 eine nationale Policy für die Langzeit-Archivierung digitaler Inhalte zu entwickeln. Der Schwerpunkt wird hierbei jedoch zunächst auf den staatlichen Sammlungen liegen. Man plant, sich an „internationalen Standards“ zu orientieren. Finanziert werden soll das Vorhaben aus Steuermitteln. Erwartet werden etwa 500 GB Daten jährlich.

Das Projekt EVA<sup>90</sup> bearbeitet die Aufgabe, Dokumente aus dem finnischen Teil des Internets zu archivieren und entsprechende eine Policy zu definieren.

Koordiniert wird EVA von der *Helsinki University Library - The National Library*. EVA ist ein Förderprojekt des finnischen Ministeriums für Erziehung. Es beschränkt sich auf Dokumente von Verlagen<sup>91</sup>.

### 3.5.4 Frankreich

In Frankreich gibt es eine gesetzliche Pflicht zur Ablieferung gedruckter und elektronischer Werke. Die Ablieferung elektronischer Werke kann entweder Datenträgergebunden oder Online erfolgen, wobei für die Online-Ablieferung entweder ein *HTTP*- bzw. *FTP*-Upload (Push) oder ein Harvesting der Inhalte (Pull) genutzt werden kann. Die Workflows für gedruckte und elektronische Objekte sind in der *Bibliothèque nationale de France* (BnF) voneinander getrennt. Seit 1992 digitalisiert die BnF auch selbst Text- und Bilddokumente und seit 2001 auch Filme.

Erste Erfahrungen mit der Langzeit-Archivierung konnten bei einer Migration der Inhalte des Archives bereits gesammelt werden. In 2000 wurden Abbilder eines kleinen Teils des Web erstellt, um einen Überblick über die entstehenden Kosten und die Wirkung des Workflows zu erhalten und um verschiedene Software-Systeme zu testen.

<sup>89</sup>*Aspects juridiques du depot de documents en archive ouverte „HAL“*; [http://www.lib.helsinki.fi/english/infoe/NationalLibraryStrategy\\_2004.htm](http://www.lib.helsinki.fi/english/infoe/NationalLibraryStrategy_2004.htm)

<sup>90</sup><http://www.lib.helsinki.fi/eva/english.html> und: Kirsti Lounamaa and Inkeri Salonharju; *EVA - The Acquisition and Archiving of Electronic Network Publications In Finland*; <http://www.lib.helsinki.fi/tietolinja/0199/evaart.html>

<sup>91</sup>Vergleichbares gilt für andere Länder, z. B. China.

Das Archivierungssystem ist zentralisiert angelegt. Es gibt eine nationale Archivierungs-Policy in Form eines Gesetzes. Die in Archiven und Bibliotheken beschäftigten Mitarbeiter werden zumeist direkt aus staatlichen Mitteln finanziert. Es gibt nur eine einzige Ausbildungsstätte für Bibliothekare und Archivare. Insgesamt gibt es viele Aktivitäten, die über die Grenzen Frankreichs hinaus relevant und anwendbar sein könnten.

Das *Centre National d'Etudes Spatiales* CNES<sup>92</sup> hat wesentlich zur Entwicklung des OAIS-Modells<sup>93</sup> für digitale Langzeit-Archive beigetragen.

Das *Centre National de la Recherche Scientifique* (CNRS)<sup>94</sup>, die nationale Forschungsorganisation Frankreichs, entwickelt derzeit zusammen mit der Universität Lyon das Projekt ETOL<sup>95</sup> (European Thesis on line), das u. a. plant, die Langzeit-Archivierung für nur zunächst Dissertationen auf europäischer Ebene im Verbund zu realisieren. Träger ist das neue nationale eprint-Archiv CCSD (*Centre pour la Communication Scientifique Directe*<sup>96</sup>), das eine europäische Rolle anstrebt. Eine Kooperation mit der DDB wird aktuell etabliert.

Ein weiteres Projekt, HAL (*Hyper Article en Ligne*<sup>97</sup>), befindet sich in der Entwicklung. Hier wird insbesondere die Verlinkung aus einem Langzeit-Archiv heraus hin zu aktueller Literatur in Sammlungen Dritter erprobt. In HAL sollen alle naturwissenschaftlichen Publikationen gesammelt und *langfristig* archiviert werden. HAL soll gleichzeitig den dort abliefernden Wissenschaftlern die Pflicht zur eigenen Ablieferung in der BnF abnehmen. Eine europaweite Kooperation ist geplant und es wird wenigstens für die Physik eine weltweite Kooperation unter Einbeziehung (und Spiegelung) des zentralen Archivs *ArXiv* angestrebt.

Neuerdings hat das CNRS eine *institutionelle Policy* für die Archivierung wissenschaftlicher Dokumente herausgegeben<sup>98</sup>. Am 6. September 2005 hat nun die nationale französische Forschungsorganisation CNRS beschlossen, ein nationales Archiv<sup>99</sup> aller digitalen wissenschaftlichen Werke, die in wissenschaftlichen Institutionen in Frankreich entstanden sind zu errichten. Die Teilnahme der Autoren soll verpflichtend sein. Das Archiv wird technisch auf der Software des CCSD aufbauen, an der auch das ISN mitgewirkt hat.

---

<sup>92</sup>[http://www.cnes.fr/html/\\_ .php](http://www.cnes.fr/html/_ .php)

<sup>93</sup>OAIS Reference Model for an Open Archival Information System; <http://www.ccsds.org/documents/650x0b1.pdf>

<sup>94</sup><http://www.cnrs.fr>

<sup>95</sup>*European Thesis on line*; CCSD, Frankreich; <http://www.ccsd.cnrs.fr/accueil.php3>

<sup>96</sup>Wissenschaftlicher Leiter des CCSD ist Franck Laloë; <http://ccsd.cnrs.fr>

<sup>97</sup><http://hal.ccsd.cnrs.fr/>

<sup>98</sup>Durch das CCSD der CNRS; Oktober 2005; [http://publiccnrs.inist.fr/IMG/pdf/guide\\_bis\\_juridique\\_publiCNRS.pdf](http://publiccnrs.inist.fr/IMG/pdf/guide_bis_juridique_publiCNRS.pdf)

<sup>99</sup>*Shared scientific communications portal*; CNRS; 6. September 2005; <http://www.cnrs.fr/singleportal.html>

### 3.5.5 Großbritannien

Großbritannien hat ein Pflichtexemplar-Gesetz. Als Pflichtexemplar-Bibliothek agiert ein Verbund der *British Library*, der *National Library of Wales*, der *National Library of Scotland*, der *Bodleian Library* in Oxford, der *Cambridge University Library* und der *Trinity College Library* in Dublin. Das Pflichtexemplar-Gesetz umfasst keine elektronischen Publikationen.

Es gibt zahlreiche nationale und lokale Projekte zur Langzeit-Archivierung, jedoch wurde auch hier der wirkliche Durchbruch vom Projekt zur Produktion noch nicht vollzogen.

#### 3.5.5.1 Das Joint Information Systems Committee

JISC<sup>100</sup> wurde 1993 gegründet als gemeinsame Anstrengung der nationalen Förderorganisationen, soweit sie für die akademische Lehre zuständig sind. Es hat sich zu einer nationalen zentralen Drehscheibe für den Aufbruch aller Zweige der Weiterbildung nach der Schule ('post-16') auf dem Gebiet der digitalen Dienste für Lehre *und* Forschung entwickelt<sup>101</sup>.

JISC hat ein Programm zur Langzeit-Archivierung digitaler Dokumente und des Dokumenten-Managements<sup>102</sup> aufgelegt.

Es bietet sich als der UK-Ansprechpartner für die Zusammenarbeit mit anderen Ländern sowie internationalen Projekten an.

Zu seinen Zielen gehört die Unterstützung der Archivierung digitaler Werke der Wissenschaft und akademischen Lehre, sowie die Beratung auf dem Weg zu institutionellen Policies bei der Langzeit-Archivierung<sup>103</sup>.

Eine einschlägige Teilaktivität zur Langzeit-Archivierung von JISC ist die in 2001 gegründete DPC (*The Digital Preservation Coalition*<sup>104</sup>). Hier werden Strategien, nicht die Policies im Sinne der Definitionen, zur Umsetzung der Langzeit-Archivierung entwickelt. Wichtig sind hierbei insbesondere die Handbücher, die mit Leitlinien für einzelne Aspekte der Langzeit-Archivierung erscheinen. Die geförderten Projekte zur Archivierung digitaler Ressourcen sind breit gestreut<sup>105</sup>. Zusammen mit PADI und der *National*

<sup>100</sup>JISC (The Joint Information Systems Committee); <http://www.jisc.ac.uk>

<sup>101</sup>JISC Fünffjahresplan 2001-2006; [http://www.jisc.ac.uk/index.cfm?name=strategy0105\\_summary](http://www.jisc.ac.uk/index.cfm?name=strategy0105_summary)

<sup>102</sup>[http://www.jisc.ac.uk/index.cfm?name=programme\\_preservation](http://www.jisc.ac.uk/index.cfm?name=programme_preservation)

<sup>103</sup>Aus dem Fünffjahresplan von JISC:

Recommendation 18 : *The JISC will work with others to help address the challenges of digital preservation and seek to foster collaboration through the establishment of a Digital Preservation Coalition with external partners;*  
Recommendation 19: *The JISC Digital Preservation Focus will oversee further development of policy, guidance and training to support institutions in managing and preserving for the long term their intellectual assets and resources in digital form.*

<sup>104</sup><http://www.dpconline.org/graphics/index.html>

<sup>105</sup><http://www.dpconline.org/graphics/join/projects.html>

*Library of Australia* wird ein Handbuch<sup>106</sup> aktuell gehalten mit den hier einschlägigen Kapiteln zur digitalen Langzeit-Archivierung und zu entsprechenden institutionellen Policies.

Das DCC (*Digital Curation Center*<sup>107</sup>) will als Forschungs- und Entwicklungsinstitut zur digitalen Langzeit-Archivierung in Großbritannien beitragen und dazu einschlägige Dienste entwickeln und anbieten. Aktuelle Forschungsthemen sind Annotationen in Datenbasen, Datenarchivierung, Metadaten, Datentransformation, Sicherheit, und unterstützende Technologien.

Ein zentrales Leitthema des DCC ist die Notwendigkeit steter Arbeit an der Verbesserung der Qualität von langzeitarchivierten Werken, um sie für die ferne Zukunft nutzbar zu erhalten. Das DCC wird von JISC finanziert.

Jüngst wurde auf dem von der *British Library*, CCLRC, DCC und JISC finanzierten *Warwick Workshop on Digital Curation and Preservation: Defining the research agenda for the next decade* eine Empfehlung<sup>108</sup> gegeben. Das *Science and Innovation Investment Framework: 2004 - 2014* der Regierung identifiziert die Langzeit-Archivierung als wichtige Aufgabe. Diese wird der OST (*Office of Science Technology working group on e-Infrastructure*) zugewiesen.

### 3.5.5.2 Die National Archives

Dokumente der englischen Regierung werden vom *The National Archives*<sup>109</sup> gesammelt. Es hat kürzlich eine erste Arbeitsgruppe<sup>110</sup> durch das NCA (*National Council on Archives*<sup>111</sup>) eingesetzt, um Leitlinien und praktische Informationen zu erarbeiten.

Das NCA hat mit an einem Aktionsplan zu einer nationalen Policy für Archive<sup>112</sup> gearbeitet. Dieser enthält eine Fülle von einschlägigen Einzelbeiträgen. Eine umfassende nationale LZA-Policy digitaler Dokumente steht aber noch aus.

**Eine breit angelegte Kooperation der deutschen Experten mit den hier genannten LZA-Aktivitäten Großbritanniens wird die eigenen Vorhaben befördern und die internationale Einbettung vorantreiben.**

### 3.5.6 Niederlande

Die Niederlande unterscheiden sich von den anderen betrachteten Nationen insbeson-

<sup>106</sup>*The Preservation Management of Digital Material Handbook*; <http://www.dpconline.org/graphics/handbook/>

<sup>107</sup><http://www.dcc.ac.uk/about/>

<sup>108</sup>[http://www.dcc.ac.uk/training/warwick\\_2005/Warwick\\_Workshop\\_report.pdf](http://www.dcc.ac.uk/training/warwick_2005/Warwick_Workshop_report.pdf)

<sup>109</sup><http://www.nationalarchives.gov.uk>

<sup>110</sup><http://www.nationalarchives.gov.uk/policy/idac/meetings/4may2004.htm>

<sup>111</sup>[http://www.ncaonline.org.uk/about\\_nca/achievements/](http://www.ncaonline.org.uk/about_nca/achievements/)

<sup>112</sup>*Government Policy on Archives: Action Plan*; InterDepartmental Archives Committee; [http://www.nationalarchives.gov.uk/policy/idac/pdf/action\\_plan.pdf](http://www.nationalarchives.gov.uk/policy/idac/pdf/action_plan.pdf)



dere durch das Fehlen jeder Form von Pflichtexemplar-Regelung und durch die allgegenwärtige Präsenz großer STM-Verlage (Springer mit Kluwer und Elsevier), die weltweit etwa 50% der STM-Journale verlegen. Die niederländische Nationalbibliothek, die *Koninklijke Bibliotheek* (KB), hat entsprechende Verträge mit diesen Verlagen abgeschlossen, um die Archivierung zu ermöglichen. Das *National Preservation Office* (NPO) finanziert die allermeisten Aktivitäten auf diesem Gebiet und ist in Räumlichkeiten der KB untergebracht.

Die in der Zukunft notwendige Schulung der Autoren, ihre Informationen in ihrem semantischen Gehalt bei der Niederschrift zu dokumentieren (z. B. durch die Nutzung von fachspezifischen Markup-Sprachen (wie z. B. *MathML*, *CML*, *PhysML*<sup>113</sup>) bzw. semantischen Beschreibungssprachen (wie *OMDoC*), wird sich in den Niederlanden schwieriger gestalten. Alle Universitäten des Landes haben einen gemeinsamen Vertrag mit *Microsoft* geschlossen, flächendeckend das Betriebssystem WINDOWS und damit als Editor WORD einzusetzen, so dass nichtproprietäre semantische Primärsprachen wie  $\text{\LaTeX}$  sehr viel seltener eingesetzt werden.

Es gibt zahlreiche Projekte zur Entwicklung von Techniken und Strategien zur Langzeit-Archivierung, darunter u. a. ein „digital preservation research testbed“ (*Testbed Digitale Bewaring*<sup>114</sup>).

Seit 1991 gibt es ein Programm zur Archivierung des gedruckten Kulturgutes, den *Delta Plan*. 1997 wurde dann ein Programm zur Archivierung von Bibliotheksbeständen, der *Metamorfoze Plan*<sup>115</sup> beschlossen.

Die KB betreibt u. A. ein Archivsystem für digitale Publikationen, das stetig in diversen Projekten weiterentwickelt wird. Die KB hat u. A. das europäische Projekt NEDLIB<sup>116</sup> geleitet, in dem zahlreiche „Guidelines“ für digitale Archive entwickelt wurden. Seit 2000 kooperiert die KB eng mit IBM zur Entwicklung eines neuen Archivsystems, dem *Deposit of Netherlands Electronic Publications-implementation* (DNEP-i). Dies ist eine Implementation des OAIS-Modells.

**Die KB bietet sich im Rahmen einer deutschen nationalen Langzeit-Archivierung als technisch und organisatorisch besonders kompetenter Partner an. Dies sollte intensiv genutzt werden.**

### 3.5.7 Norwegen

Die Dokumentenauswahl in Norwegen zur Archivierung folgt noch eng den Papier-

---

<sup>113</sup>Romeo Anghelache; *Metadata and XML in physics related scientific documents*; <http://hermes.aei.mpg.de/meta.html>

<sup>114</sup><http://www.digitaleduurzaamheid.nl/home.cfm>

<sup>115</sup>Metamorfoze (*Nationaal Programma voor het behoud van het papier erfgoed*); (2005-2008); <http://www.metamorfoze.nl/programma/beleidsplan2005-2008.pdf>

<sup>116</sup><http://www.kb.nl/coop/nedlib/index.html?related.html>

Archivierungsregeln. Das Gesetz<sup>117</sup> mit späteren ergänzenden Regulierungen<sup>118</sup> stammt noch aus dem Jahre 1989.

Die archivierende Bibliothek ist die NB (*Nasjonallbibliothek*). Standorte sind in Oslo, sowie speziell für staatliche Dokumente in Rana.

Offener Zugang wird derzeit nur für die norwegischen Newsgroups (Kopie zur Archivierung) gewährt. Periodika und Monografien sowie einige Server werden gespiegelt und archiviert (bisher auf CD-ROM). Elektronische Dokumente, wie z. B. DVD, CD, Disketten, magnetische Tapes, die unter die Pflichtexemplar-Regelung fallen, werden in zwei Exemplaren eingelagert. Das Gleiche gilt für Dokumente aus Fernsehen, Rechner-Netzen und aus dem Bereich der Telekommunikation.

Indirekte gesetzliche Grundlage der Langzeit-Archivierung ist der *Freedom of Information Act*<sup>119</sup>, der festschreibt, dass der Bürger ein Recht auf Einsicht in staatliche Dokumente hat. Die Einschränkungen nach Inhalt, Form und Zweck sind im *Information Act* festgelegt. Der König selbst kann hierzu eventuell vorhandene Rechte von Personen, Firmen oder Behörden einschränken.

Norwegen nahm am europäischen Projekt NEDLIB teil<sup>120</sup>.

Es besteht eine Ablieferungspflicht für alle Dokumente, die originär digital sind und die der Öffentlichkeit präsentiert wurden, und zwar unabhängig von der Quelle.

Die Auswahl geschieht durch die *Nasjonallbibliothek* entsprechend dem vermuteten kulturellen Wert für nachfolgende Generationen.

Die Nationalbibliothek hat ein Recht darauf, öffentlich zugängliche Dokumente Dritter für ihr Archiv zu kopieren.

### 3.5.8 Die Schweiz

Nationale Bibliothek ist die *Schweizerische Landesbibliothek* (SLB) (*Swiss National Library SNL*)<sup>121</sup>.

Auf dem Weg zu einer LZA-Policy gibt es noch keine Vereinbarungen. Es gibt aber an der SLB ein entsprechendes Projekt: *E-Helvetica*<sup>122</sup>.

Kooperationen bestehen innerhalb des Projektes mit den Schweizerischen Bundes-

<sup>117</sup>Act of Legal Deposit No. 32 of 9 June 1989 relating to the legal deposit of generally available documents; [http://www.nb.no/fag/for\\_utgjevarar\\_og\\_trykkeri/pliktavleveringhtml/index.php](http://www.nb.no/fag/for_utgjevarar_og_trykkeri/pliktavleveringhtml/index.php)

<sup>118</sup>The statutory regulations, §30.

<sup>119</sup>Act of 19 June 1970 No. 69 relating to public access to documents in the public administration as subsequently amended, most recently by Act of 10 January 1997 No. 7; (short title: Freedom of Information Act); <http://www.ub.uio.no/ujur/ulovdata/lov-19700619-069-eng.pdf>

<sup>120</sup>NEDLIB (*Networked European Deposit Library*), „a collaborative project of European national libraries to develop infrastructure for a European deposit library“.

<sup>121</sup><http://www.sn1.admin.ch>

<sup>122</sup><http://www.e-helvetica.ch>

Kantons- und Universitäts-Bibliotheken sowie mit einzelnen Verlagen. Sie sind in einem frühen Stadium.

Es gibt jedoch bereits ein Gesetz, das *Bundesgesetz über die Schweizerische Landesbibliothek*<sup>123</sup>, welches die SLB als in der Schweiz für die Langzeit-Archivierung verantwortlich und zuständig erklärt.

Das Gesetz beschränkt den Auftrag darauf, „Helvetica“ zu sammeln, der Nachwelt zu erhalten und zugänglich zu machen.

Dieser Auftrag bezieht sich auf alle Dokumentenarten und -typen, inklusive Adresswerken, Telefonbüchern, Dissertationen, Zeitungen etc., sofern sie von Schweizer Autoren (auch im Ausland) verfasst wurden, sowie auf Werke aller Art, die sich auf die Schweiz beziehen oder diese zum Thema haben.

Finanziert wird die Langzeit-Archivierung ausschließlich von der Schweizer Regierung.

Es werden Dokumente sowohl über interaktive Webseiten wie durch *ftp*, *Email*, bzw. das *OAI-PMH-Protocol* importiert.

Metadaten nach den *DublinCore*-Vereinbarungen werden bisher nur für Dissertationen eingesetzt.

Es gibt noch keinen organisierten Workflow vom Autor bis in die Langzeit-Archivierung, wohl aber einen internen Workflow innerhalb des Projektes in der SLB. Es existieren hier formulierte Kriterien für die Auswahl eines Werkes zur Aufnahme in die Langzeit-Archivierung (sogenannte Sammelrichtlinien<sup>124</sup>).

Es müssen alle Formate angenommen werden. Der Workflow ist möglichst genau analog zu dem von gedruckten Materialien gehalten.

Ziel ist, auch die Originale (Druckwerke, Autoren-Fassungen (z. B.  $\LaTeX$ -Dateien) mit aufzubewahren.

Der Umfang ist bisher gering: Etwa 100 Dokumente pro Jahr, meist Dissertationen, alle mit parallel gedruckter Fassung.

Bisher sind die technischen Realisierungen zu Verfügbarkeit, Integrität und Lesbarkeit noch wenig ausgearbeitet. Entsprechend ist das Archiv bisher **nicht** öffentlich zugänglich. Dies soll frühestens 2006 geschehen.

Das Projekt *E-Helvetica* läuft seit 2001 und hat letztlich die Umsetzung des §2 des Bundesgesetzes über die SLB zum Ziel, laut dessen die SLB die Aufgabe hat, „gedruckte oder auf anderen Informationsträgern gespeicherte Informationen, die einen Bezug zur Schweiz haben, zu sammeln, zu erschließen, zu erhalten und zu vermitteln“.

Technisch und konzeptionell basiert die E-Helvetica-Lösung auf den Ergebnissen des europäischen Projektes NEDLIB (1998 - 2001). Das Projekt soll bis 2006 laufen und

<sup>123</sup>[http://www.admin.ch/ch/d/sr/432\\_21/](http://www.admin.ch/ch/d/sr/432_21/)

<sup>124</sup>[http://www.e-helvetica.admin.ch/pdf/ger/tp-org/sammelrichtlinien\\_online\\_helvetica.pdf](http://www.e-helvetica.admin.ch/pdf/ger/tp-org/sammelrichtlinien_online_helvetica.pdf)

dann in die Betriebsphase eines Langzeit-Archives übergehen. Ab dann soll auch der Online-Zugriff auf das Archiv möglich sein.

Interessant ist das Projekt insbesondere, weil hier im deutschsprachigen Ausland bereits praktische Erfahrungen gesammelt wurden, insbesondere bezüglich der Mehrsprachigkeit der Lösung sowie der strikten Trennung der Organisation der Langzeit-Archivierung und der aktuellen Archivierung. Auch dürfte eine Analyse der sich in der Entwicklung befindlichen Geschäftsgänge von Interesse sein.

Neben dem zentralen nationalen Ansatz zur Langzeit-Archivierung der SLB gibt es einen verteilten Ansatz, die „E-Archiving“-Aktivitäten des Konsortiums der *Schweizer Hochschulbibliotheken*. Diese wollen nicht vorrangig selbst ein Archiv aufbauen, sondern insbesondere den Zugriff auf kommerzielle und bezahlte Online-Inhalte von Zeitschriften auch nach Ablauf eines Abonnements oder nach Auflösung des Anbieters sicherstellen. Dieser Aspekt ist eng an die Langzeit-Archivierung gekoppelt<sup>125</sup>. Eine nationale Policy ist, soweit bekannt, nicht vorgesehen.

Ideen zu einer Policy wurden von *Price Waterhouse Coopers* für den Bereich der Archive mit der *Gesamtschweizerischen Strategie zur dauerhaften Archivierung von Unterlagen aus elektronischen Systemen* entwickelt und zusammengetragen<sup>126</sup>. Zwar auf den archivarischen Bereich beschränkt, gibt diese Studie doch einen Eindruck von der Problemvielfalt, so in deren Kapitel 3.1.2 eine Auflistung der Problemkreise und Kernfragen. Die Studie stellt u.a. heraus, dass langfristig Kooperationen zwischen den Archiven notwendig sein werden, um die digitalen Informationen auch langfristig stabil und verstehbar vorzuhalten und den Nutzern eine zeitgemäße Zugriffsmöglichkeit zu bieten. Die Studie mahnt einen Paradigmenwechsel in den Archiven an, nicht mehr einfach nur Vergangenes zu archivieren, sondern innovativ neue Techniken zu integrieren, bevor die mit diesen erstellten Inhalte in die Archive aufgenommen werden. Politisch mahnt die Studie neben zusätzlicher finanzieller und personeller Ausstattung der Archive insbesondere eine partnerschaftliche Kooperation der Archive an.

Die Studie entwickelt mit dem fiktiven Szenario des Archivwesens im Jahre 2015 eine Idee von den Chancen und Schwierigkeiten auf dem Weg zu einer LZA-Policy.

Die Autoren vermuten, dass es problematisch sein wird, die notwendigen politischen Entscheidungen zu initiieren, weil das Fachwissen in den Parlamenten fehle. Als weitere Hürden werden die notwendige verstärkte Weiterbildung des Personals bei gleichzeitig erhöhtem Personalbedarf durch die digitale Archivierung und parallel die Anschaffung der notwendigen technischen Infrastruktur eingestuft. Auf diese thematisch relevanten Mahnungen der PWC-Studie für die Schweiz weisen wir gerne hin, widerstehen jedoch der Versuchung ihre Übertragbarkeit auf die deutschen Anstrengungen zu analysieren.

---

<sup>125</sup>Siehe auch *Koninklijke Bibliotheek – Nationale Bibliotheek van Nederland* (KB); <http://www.kb.nl> in den Niederlanden.

<sup>126</sup>[http://www.vsa-aas.org/uploads/media/d\\_strategie.pdf](http://www.vsa-aas.org/uploads/media/d_strategie.pdf)

### 3.5.9 Tschechische Republik

Tschechien, als kleines Industrieland, nutzt seinen Vorteil der kurzen Kommunikationswege. Die *National Library of the CR* (NLC)<sup>127</sup> hat ein Langzeit-Archiv, *Archive of Czech web resources*<sup>128</sup>, von tschechischen Webseiten aufgesetzt. Es werden für Autoren als Dienste u.a. ein Metadaten- und ein URN-Erzeuger angeboten.

Die Auswahl der Dokumente<sup>129</sup> für die LZA geschieht unabhängig von der Quelleninstitution (z. B. Verlag oder Einzelperson), verlangt jedoch, dass das Dokument originär digital Online ist (also keine Kopie eines Druckerzeugnis). Die Dokumente müssen *Open Access* sein. Der Dokumentenraum beschränkt sich auf Dokumente, die in Tschechien bzw. von Tschechen (auch im Ausland) erzeugt wurden. Entscheidungen zur Auswahl trifft die NLC.

Es existiert bisher keine institutionelle oder nationale LZA-Policy; diese wird aber gewünscht und im Rahmen eines Forschungsprojektes erarbeitet, an dem Wissenschaftsinstitutionen, Bibliotheken und das Ministerium für Kultur beteiligt sind.

Der Umfang des Archivs betrug 2004 etwa 10 Millionen Datensätze (500GB/Jahr).

Das Archiv selbst ist bisher noch nicht öffentlich.

### 3.5.10 Die Vereinigten Staaten von Amerika

#### 3.5.10.1 Das PREMIS-Projekt

Die Aktivitäten in den Vereinigten Staaten werden von vielen Playern vorangetrieben, wobei insbesondere die *Library of Congress* (LOC<sup>130</sup>) und das *Online Computer Library Center* (OCLC) als Knoten öffentlich sichtbar sind.

Im PREMIS-Projekt (*PREservation Metadata Implementation Strategies*<sup>131</sup>) werden u. A. schwerpunktmäßig Metadaten-Konzepte für die Langzeit-Archivierung entwickelt.

Hierzu wurden zunächst Anwendungsszenarien (*Use Cases*) entwickelt, für die dann ein Core Set der Metadaten entwickelt wird. Es werden im PREMIS-Projekt Strategien für die Beschreibung und Speicherung digitaler Objekte in einem Archiv sowie für den Austausch von Objekten und ihren Metadaten zwischen Archiven bzw. Archivsystemen entwickelt und in Pilotanwendungen getestet.

PREMIS ist ein internationales Projekt, in das mehrere Staaten und viele Institutionen eingebunden sind. An der Finanzierung ist auch die RLG<sup>132</sup> beteiligt.

<sup>127</sup><http://www.nkp.cz>

<sup>128</sup><http://www.webarchiv.cz/index-e.html>

<sup>129</sup><http://www.webarchiv.cz/criteria.html>

<sup>130</sup><http://www.loc.gov>

<sup>131</sup><http://www.oclc.org/research/projects/pmwg/>

<sup>132</sup>RLG ist eine internationale gemeinnützige Organisation von Bibliotheken, Museen und Archiven mit

Aus Beispielen Dritter wurde versucht, einen Kernset of *Metadata* zu entwickeln und zu erproben. Ausgangspunkt waren bestehende Preservation-Metadaten Dritter.

**Diese Aktivitäten könnten im Rahmen der Entwicklung einer bundesdeutschen LZA-Policy beobachtet und möglichst aktiv unterstützt werden.**

Die hier entwickelten Lösungen könnten anschließend in das Konzept eines bundesdeutschen Langzeit-Archivs (Netzwerk) einfließen. Die aktive Mitarbeit im PREMIS-Projekt wird aufwendige und teure Eigenentwicklung von Lösungen verringern und die notwendige Expertise in Deutschland stärken. Das bei OCLC entstehende Core Set für *Metadaten der Langzeit-Archivierung* könnte eine entsprechende Relevanz entwickeln, die das ebenfalls zunächst bei OCLC für einfache zentrale inhaltsbeschreibende Metadaten initiierte *DublinCore* bereits erreicht hat.

Die Niedersächsische Staats- und Universitätsbibliothek SUB Göttingen lässt sich im Rahmen des BMBF-Projektes *Cashmere-int*<sup>133</sup> in den PREMIS-Arbeitsgruppen vertreten.

Im Mai 2005 hat PREMIS seinen Abschlussbericht<sup>134</sup> vorgelegt, der ein *Metadaten-Lexikon*<sup>135</sup> mit einer Anleitung<sup>136</sup> enthält.

### 3.5.10.2 Das Digital Information Infrastructure and Preservation Program

Bausteine des NDIIP<sup>137</sup>, sind z. B. der METASCHOLAR Workshop<sup>138</sup> zu *Open Access and Digital Preservation* sowie Vorträge von Thomas Robertson über das Projekt LOCKSS (*Lot Of Copies Keep Stuff Save*)<sup>139</sup>. LOCKSS entwickelt ein Peer-to-Peer Proxy-Netzwerk, das Copyright- und Urheberrechtsregeln beachtet und den persistenten Zugang zu elektronischen Zeitschriften-Artikeln ermöglicht, insbesondere für den Fall, dass der Verlag nicht mehr existiert. Weiterführend sind auch Beiträge von J. Willinski über OJS<sup>140</sup> und die Technik digitaler Langzeit-Archivierung<sup>141</sup> sowie von Ted Harrison<sup>142</sup> vom Oktober 2004 zum NDIIP.

Die LOC hat im September 2004 Verträge mit acht strategischen Partnern abgeschlossen, die zusammen mit jeweils eigenen Partner-Institutionen Teilaspekte behandeln:

- dem Ziel, Lösungen für den Informationszugang und das -Management zu erarbeiten. Aus Deutschland ist die Bayerische Staatsbibliothek Mitglied; <http://www.rlg.org>

<sup>133</sup><http://www.iwi-iuk.org/cashmere/>

<sup>134</sup>*Data Dictionary for Preservation Metadata*; Final Report of the PREMIS Working Group (May 2005); <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>

<sup>135</sup><http://www.oclc.org/research/projects/pmwg/premis-dd.pdf>

<sup>136</sup><http://www.oclc.org/research/projects/pmwg/premis-examples.pdf>

<sup>137</sup><http://www.digitalpreservation.gov>

<sup>138</sup>*Symposium on Open Access and Digital Preservation*; METASCHOLAR; Atlanta, USA; 2004; <http://www.metascholar.org/OADP-Symposium.html>

<sup>139</sup>Thomas Robertson; LOCKSS; <http://lockss.stanford.edu>

<sup>140</sup>*Open Journal Systems*; <http://pkp.sfu.ca/ojs/>

<sup>141</sup><http://dewey.library.nd.edu/morgan/travel/open-access-symposium/#shulenburger>

<sup>142</sup>Ted Harrison; *The NDIIP: The National Digital Information Infrastructure and Preservation Program plan*; <http://www.cs.odu.edu/~mln/teaching/cs791-s04/?method=getElement&element=~week15~harrison.ppt>

die CDL (*California Digital Library at the University of California*<sup>143</sup>) erarbeitet Web-Archivierungswerkzeuge für die Langzeit-Archivierung von Regierungsdokumenten und Dokumenten aus dem politischen Raum;

- die *University of California at Santa Barbara* entwickelt mit Partner-Institutionen ein *National Geospatial Federated Digital Repository* für die Langzeit-Archivierung geographischer Informationen der USA;
- die *Educational Broadcasting Corporation* (EBC) (Thirteen/WNET New York) wird erste Prozeduren, Strukturen und nationale Standards für die Langzeit-Archivierung der öffentlichen Fernsehprogramme erarbeiten.

In den Vereinigten Staaten selbst scheint dieses Vorhaben vielen Wissenschaftlern und Bibliothekaren der lokalen Bibliotheken zwar bekannt zu sein, wird jedoch als „Regierungsprojekt“ gesehen und entsprechend distanziert von den inhaltskompetenten Playern kommentiert. LOC selbst nimmt zu den Erfolgsaussichten des Vorhabens diplomatisch wie folgt Stellung: „One problem is the large scope of the US, so a national strategy is difficult to manage.“

Eingebunden in dieses Vorhaben sind zahlreiche Regierungsinstitutionen, Bibliotheken, Archive, wissenschaftliche Institutionen, Universitäten und einige kommerzielle Unternehmen. Die *Johns Hopkins University* als eine der vier im *Archive Ingest Handling Test* evaluierenden Universitäten berichtet: „We're still testing the ingestion, validation and transfer of digital content“. LOC beschreibt den Stand des Projektes als „Discussion, some testing and some evaluation“, was aber in Anbetracht des Vorhabens deutlich untertrieben sein dürfte.

Interessant wird es sein, welchen Ansatz einer Archivierung man in den Vereinigten Staaten wählen wird, ob die Langzeit-Archivierung eine Aufgabe der LOC werden wird, eine andere Regierungsinstitution damit beauftragt oder doch ein dezentraler Ansatz gewählt werden wird. Die Argumente für und wider die einzelnen denkbaren Modelle könnten für eine bundesdeutsche Initiative zumindest interessant, wenn auch nicht unbedingt übertragbar sein.

Die Entwicklungen in den Vereinigten Staaten können dabei auf eine verhältnismäßig lange Historie und die entsprechende Menge praktischer Erfahrungen bauen: Überlegungen in der LOC für eine nationale Policy begannen 1998. Ein Beispiel einer institutionellen Policy als Vorlage für bundesdeutsche Institutionen könnte die 2000 entwickelte *Policy for Preservation of Digital Resources*<sup>144</sup> der *Columbia University Library* sein.

### 3.5.10.3 Der e-Government Act

Im Jahre 2002 wurde der *E-Government Act* (Gesetz zu einem durch alle Behörden

<sup>143</sup>[http://www.cdlib.org/programs/digital\\_preservation.html](http://www.cdlib.org/programs/digital_preservation.html)

<sup>144</sup><http://www.columbia.edu/cu/libraries/services/preservation/dlpolicy.html>

durchgängigen E-Informationssystem) von 2001 neu gefasst und mit Implementierungsanleitungen<sup>145</sup> versehen. Sie enthält eine Langzeit-Archivierung für die Erhaltung von Regierungsdokumenten.

Einen umfangreichen Report zur Langzeit-Archivierung digitaler Werke hat das *National Science Board* der NSF (*National Science Foundation* herausgebracht<sup>146</sup>. Einschlägig ist hier das vierte Kapitel (*Perspectives on Digital Data Collections Policy*), das eine nationale NSF Policy einklagt, und sich mit den dazu notwendigen Teilthemen wie Zugang und Management befasst. Zwar sind Daten-Sammlungen ein relativ spezieller Sektor der digitalen Daten für eine Langzeit-Archivierung, doch sind viele der Argumente des NSB auch auf digitale Dokumente übertragbar.

Ein Vergleich zu anderen Ländern ist auch über die Linksammlung des *E-Government Resource Centre*<sup>147</sup> möglich.

**Deutschland sollte auf Expertenebene wie auf institutioneller und Regierungsebene enge Kontakte zu den entsprechenden Partnern und Projekten der USA aufbauen, um zeitnah an den technischen Fortschritten, neuen Diensten und Standards zu partizipieren, dies umso mehr, als die USA ebenso wie Deutschland ein moderner Industriestaat mit einer vergleichbaren bundesstaatlichen politischen Gliederung in Bundesstaaten.**

### 3.5.11 Europäische Union

#### 3.5.11.1 Langfristige Pläne

Die Europäische Union plant eine Intensivierung der Aktivitäten und der Zuwendung zur Bewahrung des wissenschaftlichen und kulturellen Erbes. Sie will im Sommer 2005 hierzu eine Erklärung herausgeben. Außerdem soll die Förderung im 7. Rahmenprogramm (ab 2007) intensiviert werden<sup>148</sup>. Die als Schrift, Bild und Ton archivierten Dokumente seien entsprechend zu bewerten, sowohl als Kulturgut als auch im Hinblick auf ihre wirtschaftliche Bedeutung. In der angekündigten Studie sollen die derzeitigen Hindernisse zu einer Langzeit-Verfügbarkeit analysiert werden und gleichzeitig ein Vorschlag für eine Empfehlung vorgelegt werden, „mit dem alle betroffenen öffentlichen Akteure mobilisiert und öffentlich-private Partnerschaften im Bereich der Digitalisierung unseres Kulturerbes erleichtert werden sollen“.

Die *European Commission on Preservation and Access (ECPA)*<sup>149</sup> will die Anstrengun-

<sup>145</sup><http://www.egov.vic.gov.au/International/TheAmericas/UnitedStates/E-government-Act-2002/e-government-act-2002.htm#implementationguidance>

<sup>146</sup>*Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century*; Report of the National Science Board; Pre-publication draft; NSB-05-40 vom 23. Mai 2005; [http://www.nsf.gov/nsb/documents/2005/LLDDC\\_report.pdf](http://www.nsf.gov/nsb/documents/2005/LLDDC_report.pdf)

<sup>147</sup><http://www.egov.vic.gov.au/index.php?env=-2570:m1699-1-1-7:10-0-1-n0-0-0&reset=1>

<sup>148</sup>Presseerklärung der EU: „Die IKT müssen in Europa besser und häufiger zur Evaluierung des kulturellen und audiovisuellen Erbes eingesetzt werden“; IP/05/528 vom 3. Mai 2005.

<sup>149</sup><http://www.knaw.nl/ecpa/>



gen zu institutionellen und nationalen LZA-Policies bündeln. Sie hat hierzu z. B. eine umfangreiche Umfrage über Absichten und den Stand zur Archivierung durchgeführt und die Ergebnisse länderweise zusammengestellt<sup>150</sup>.

Im 7. Rahmenprogramm sollen verstärkt Mittel eingesetzt werden „zur Entwicklung von Suchmaschinen für die breite Öffentlichkeit“. Für den Zeitraum 2005 - 2008 soll „die Digitalisierung von Inhalten und der Zugang sowie verstärkt Maßnahmen zur Vernetzung von Archiven in Europa eingesetzt werden“.

Eine Zuarbeit liefert das neue Programm *eContentplus*<sup>151</sup>. Es soll die Fragmentierung auf dem Markt der europäischen digitalen Inhalte verringern und die Zugänglichkeit und Nutzbarkeit von geografischen Informationen, kulturellen Inhalten und Lehrmaterialien verbessern.

Im Mittelpunkt des neuen Programms stehen jene Segmente des Inhaltmarktes, die in Europa stark fragmentiert sind und in denen die Marktkräfte derzeit noch nicht in der Lage sind, Wachstum zu schaffen.

Besonders gefördert werden die drei Bereiche Raum- (oder geografische) Daten, Lehrmaterialien und kulturelle Inhalte. Das Programm *eContentplus* ist Teil eines umfangreichen Maßnahmen-Pakets der EU zur Verbesserung der Voraussetzungen für Innovation und Kreativität auf dem konvergierenden Inhaltmarkt.

Dazu gehören das Programm *MEDIAplus*<sup>152</sup> und die in diesem Jahr anstehende Modernisierung der Richtlinie *Fernsehen ohne Grenzen*<sup>153</sup>.

In einer Initiative *i2010 Digital Libraries*<sup>154</sup> soll der Zugang zu digitalen Informationen (originär digital oder Digitalisate gedruckter Werke) des kulturellen Erbes einfacher und attraktiver gestaltet werden.

Kernpunkte des Programms sind:

- Freier Zugang für jedermann: nur so könnten die Bürger, Wissenschaftler und Firmen maximalen Nutzen aus den gespeicherten Informationen ziehen;
- Digitalisierung von analog vorliegenden Werken<sup>155</sup> (wie Druckwerken), um sie einer breiteren Leserschaft zugänglich zu machen;
- Langzeit-Archivierung, um die Dokumente zukünftigen Generationen zu erhalten.

<sup>150</sup><http://www.knaw.nl/ecpa/map/index.html>

<sup>151</sup>*eContentplus*; EU-Programm; <http://www.econtentplus.net/>

<sup>152</sup>*Mediaplus*; EU-Programm; [http://europa.eu.int/comm/avpolicy/media/index\\_fr.html](http://europa.eu.int/comm/avpolicy/media/index_fr.html)

<sup>153</sup>[http://europa.eu.int/comm/avpolicy/regul/regul\\_de.htm](http://europa.eu.int/comm/avpolicy/regul/regul_de.htm)

<sup>154</sup>[http://europa.eu.int/information\\_society/activities/digital\\_libraries/index\\_en.htm](http://europa.eu.int/information_society/activities/digital_libraries/index_en.htm)

<sup>155</sup>Diese Herkulesaufgabe ist schwer vorstellbar: Allein für die Bibliotheken des (größeren) Europas wurden 2.5 Mrd. gedruckte Werke geschätzt. Siehe: D. Fuegi, M. Jennings; *International Library Statistics: Trends and Commentary based on the Libecon data*; 2004; <http://www.libecon.org/pdf/InternationalLibraryStatistic.pdf>

In *i2010 Digital Libraries* sind mehrere Stränge geplant:

Am 30. 09. 2005 wurde der Strang *Communication* von der Kommission beschlossen. Er zielt schwerpunktmäßig auf ein Konzept zur Langzeit-Verfügbarkeit *kulturellen Erbes*<sup>156</sup>.

Ein zweiter Strang *online Consultation* hat als Kern einen Fragebogen<sup>157</sup>, der aktuell verschickt wurde. Daraus soll dann eine Empfehlung zur Digitalisierung und digitalen Langzeit-Archivierung, vorgesehen für 2006, erarbeitet werden. Weiterhin soll bis dann eine Überarbeitung des *Copyright framework* vorliegen.

Ebenfalls für 2006 ist ein Strang *Scientific Information* vorgesehen, der naturgemäß vor allem auf die Langzeit-Archivierung aktueller originär digitaler Dokumente zielen soll. Komplementiert wird dieses Konzept durch das FP 7 (2007), das Forschungsaktivitäten zur Digitalisierung, und zur Langzeit-Archivierung kulturellen Erbes sowie *Centres of Competence on digitization preservation* fördern will, sowie das Projekt *eContentplus*.

Das von Frankreich initiierte Programm *i2010* entstand ursprünglich als Reaktion auf das Google-Projekt *Google Book Library Project*<sup>158</sup>, bei dem von einer sehr großen Zahl von digitalisierten Büchern dem Anfragenden über *Google Book Search*<sup>159</sup> nur 'snippets' (kurze Textzitate), bzw. wenn Google das Einverständnis des entsprechenden Verlages hat, von einzelnen Seiten gegeben werden. Nur für Bücher von vor 1864 kann der Volltext eingesehen werden. Das Projekt *dient also nur der Suche nach Büchern*, und kann die Aufgaben staatlicher Bibliotheken und Archive in der Langzeit-Archivierung nicht ersetzen. Zur Langfrist-Verfügbarkeit dieser Digitalisate im Besitz einer kommerziellen Firma können naturgemäß keine Garantien gegeben werden.

### 3.5.11.2 Aktuelle Projekte

Zu den bereits durchgeführten, laufenden und angekündigten Maßnahmen der Europäischen Union gehören die Entwicklung von Suchmaschinen, gemeinsamen Standards sowie die Unterstützung der Digitalisierung von Inhalten.

Fortgeführt und erweitert werden soll die Entwicklung von Standardisierungs- und Suchverfahren für die Nutzung von Archiven schriftlicher und audiovisueller Dokumente in einem digitalen Umfeld.

Bisherige Förderprogramme waren der Austausch bewährter Praktiken, die Festlegung von Benchmarks und der Dialog der betroffenen Akteure.

<sup>156</sup>*Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: i2010: Digital Libraries*; SEC (Commission of the European Communities); 1194, 2005; [http://europa.eu.int/information\\_society/activities/digital\\_libraries/doc/en\\_comm\\_digital\\_libraries.pdf](http://europa.eu.int/information_society/activities/digital_libraries/doc/en_comm_digital_libraries.pdf)

<sup>157</sup>[http://europa.eu.int/information\\_society/activities/digital\\_libraries/consultation/index\\_en.htm](http://europa.eu.int/information_society/activities/digital_libraries/consultation/index_en.htm)

<sup>158</sup><http://books.google.com/intl/en/googlebooks/library.html>

<sup>159</sup><http://books.google.com/intl/en/googlebooks/about.html>

Projekte sind MINERVA<sup>160</sup>, MINERVA Plus und DIGICULT-Forum<sup>161</sup>. Außerdem wurde ein Prozess der Standardisierung und der Tests für einen wechselseitigen Zugang bei öffentlichen Bibliotheken eingeleitet<sup>162</sup>.

Bei den audiovisuellen Archiven leisten die Maßnahmen PRESTO und PRESTO-SPACE<sup>163</sup> einen Beitrag zur Entwicklung von schnellen und qualitativ hochwertigen Digitalisierungsverfahren für Fernseharchive. Teilprojekte sind Bricks<sup>164</sup>; *Calimera*<sup>165</sup>, ein Koordinierungsprojekt zur Zusammenstellung von Informationen und Adressen mit einer Analyse der Aktivitäten in den EU-Ländern mit seinen kulturellen Institutionen<sup>166</sup>; *Delos*<sup>167</sup>, das im FP6 technische Lösungen erarbeiten will; *Epoch*, in dem sich über 100 Partner aus allen Bereichen der Technologien zur Bewahrung des kulturellen Erbes widmen wollen, und *MinervaPlus*. Das EU-Programm MEDIA<sup>168</sup> unterstützt die Katalogisierung von Filmmaterial in digitaler Form. Eine der Prioritäten des Programms MEDIA 2007 wird die Nutzung digitaler Technologien durch die audiovisuelle Industrie in Europa sein.

Demnächst soll eine Empfehlung zum Filmerbe angenommen werden, mit der insbesondere die Standardisierung von Filmkatalogen und die Interoperabilität der Datenbanken der audiovisuellen Archive gefördert werden sollen.

Im bisherigen IST FP6<sup>169</sup> gab es noch einmal eine Ausschreibung zu *LT Preservation and Access of Digital Cultural Heritage Sources* mit verschiedenen *Integrating Projects* IP und STREP (*Specific Targeted Research Projects*).

<sup>160</sup> *Minerva* EU Programm; [http://europa.eu.int/comm/education/programmes/socrates/minerva/index\\_en.html](http://europa.eu.int/comm/education/programmes/socrates/minerva/index_en.html) und *Education and Training*; [http://europa.eu.int/comm/education/index\\_en.html](http://europa.eu.int/comm/education/index_en.html)

<sup>161</sup> *Digicult Forum Technology Challenges for Digital Culture*; <http://www.digicult.info/pages/index.php>

<sup>162</sup> TEL Projekt (*The European Library*), das heute von der niederländischen Nationalbibliothek verwaltet wird und Zugang zu 90 Sammlungen gewährt.

<sup>163</sup> <http://www.prestospace.org/>

<sup>164</sup> BRICKS *Building Resources for Integrated Cultural Knowledge Services*; <http://www.bricksfactory.com/pls/bricks/ecolnet.home>

<sup>165</sup> <http://www.calimera.org/default.aspx>

<sup>166</sup> Das (offensichtlich längst nicht mehr aktuell gehaltene) offizielle öffentliche Statement des Projekts *Calimera* zu *Policies and Strategic Activities* in Deutschland soll hier nicht vorenthalten werden: *There is no officially formulated vision or policy for library, museum or archival services. These areas have never been subject to federal legislation. National strategies in the information society/network society are formulated on the professional rather than the political level. Depending on the organization of cultural services within the individual States, there are varying degrees of state-determined policy. Recent attempts to move more responsibility for cultural affairs to the federal level have been rejected. Consequently it is doubtful whether extensive, co-ordinated activities will emerge in the near future. Individual initiatives in the various institutions will provide the incentive for new developments and will be copied or modified by other institutions. Professional organizations generally provide the strongest influence on formulating and pushing developments. They lack, however, the ability to directly create policies and activities to which all organizations are committed to follow.*

<sup>167</sup> <http://www.delos.info>

<sup>168</sup> [http://europa.eu.int/comm/avpolicy/media/index\\_en.html](http://europa.eu.int/comm/avpolicy/media/index_en.html)

<sup>169</sup> *DigiCult* <http://www.cordis.lu/ist/digicult/index.html>

Im EU-Projekt SEPIA<sup>170</sup> wird die Langzeit-Archivierung von fotografischen Materialien bearbeitet. Die Partner<sup>171</sup>, ein Zusammenschluss aus Nationalbibliotheken und Wissenschaftsinstituten, wenden sich in SEPIA II dem Austausch von Erfahrungen, der Öffentlichkeitsarbeit und der semantischen maschinenlesbaren Beschreibung zu.

Der Beitrag aus Deutschland zielt auf die Langzeit-Archivierung der ca. 10 Millionen Fotografien des Bundesarchivs<sup>172</sup> ab. Koordinierungsstelle ist die ECPA<sup>173</sup>. Ein wichtiges Ergebnis von SEPIA ist SEPIADES<sup>174</sup>.

Das EU-Projekt Erpanet<sup>175</sup> (*Electronic Resource Preservation and Access Network*) soll als europäische Plattform des Informationsaustauschs zur Langzeit-Archivierung dienen. Es hat die Knoten HATII (*The Humanities Advanced Technology and Information Institute*, Glasgow), das *Nationaal Archief van Nederland*, das *Istituto di Studi per la Tutela dei Beni Archivistici e Librari* der *Universita degli Studi di Urbino Carlo Bo*, das *Schweizerische Bundesarchiv*, und PADI. Standort ist Glasgow, Schottland.

### 3.5.11.3 Das Forschungsprogramm FP6

Im IST *Sixth Framework Programme FP6*<sup>176</sup> gab es im *Call 5* eine *Strategic Objective* zu *Access to and preservation of cultural and scientific resources*. Die Mitarbeit in den *Specific Programmes*<sup>177</sup> und den *Advisory Groups*<sup>178</sup> durch kompetente Vertreter aus Deutschland sollte auch zu einer stärkeren Beteiligung in den aktuellen Ausschreibungen führen.

**Deutschland als einer der großen Industriestaaten in der EU sollte eine starke und aktive Rolle in den EU-Entwicklungslinien übernehmen und sich hierzu vielfältig als Projektnehmer an den aktuellen Calls sowie engagiert an den weiteren Planungen und Richtungsentscheidungen beteiligen.**

## 3.5.12 Die UNESCO

In der UNESCO wird diskutiert, ob sie sich eine Rolle bei der Entwicklung und Festlegung von allgemeinen übernationalen Regeln für nationale LZA-Policies zu Eigen machen soll.

<sup>170</sup>SEPIA (*Safeguarding European Photographic Images for Access*); <http://www.knaw.nl/ecpa/sepia/home.html>

<sup>171</sup><http://www.knaw.nl/ecpa/sepia/home/partners.html>

<sup>172</sup><http://www.bundesarchiv.de/>

<sup>173</sup>ECPA (*European Commission on Preservation and Access*); Amsterdam, The Netherlands; <http://www.knaw.nl>

<sup>174</sup>Aasbø, K., Ortega García, I., Isomursu, A., Johansson, T., Klijn, E., Klijn, E.(ed); *SEPIADES: Recommendations for cataloguing photographic collections*; 2003, 248 pp. 90-6984-397-8; <http://www.knaw.nl/ecpa/sepia/workinggroups/wp5/sepiadestool/sepiadesdef.pdf>

<sup>175</sup><http://www.erpanet.org/about.php>

<sup>176</sup><http://www.cordis.lu/fp6/find-doc.htm#wp>

<sup>177</sup><http://www.cordis.lu/fp6/find-doc-general.htm#wps>

<sup>178</sup><http://www.cordis.lu/fp6/eags.htm>

Zentrale Anlaufstelle ist das *UNESCO Archives Portal*<sup>179</sup>. Zentrale Dokumente sind die *Charter on the Preservation of the Digital Heritage*<sup>180</sup>, die am 17. Oktober 2003 von der Generalversammlung der UNESCO verabschiedet wurde. Ebenfalls 2003 wurden *Guidelines for the Preservation of Information Digital Heritage*<sup>181</sup> herausgegeben. Eine *UNESCO-Perspektive zur Archivierung des digitalen Erbes* wurde von Abdelaziz Abid geschrieben<sup>182</sup>.

Nationale Kommissionen für die Mitarbeit in der UNESCO bereiten die Zuarbeit vor und organisieren hierzu z. B. internationale Konferenzen, so auch zur Frage der nationalen LZA-Policies und deren internationaler Harmonisierung<sup>183</sup>. Erstes Ziel sind Anforderungen an LZA-Policies zusammenzustellen, so die Auswahl von Kriterien für Dokumente und die Aufteilung der Aufgaben und Verantwortlichkeiten.

### 3.5.13 Aktivitäten von Nationalbibliotheken weiterer Staaten

Um die Suche nach Aktivitäten zur Entwicklung einer nationalen LZA-Policy weiterer Staaten zu erleichtern, wird hier zunächst auf Verzeichnisse Dritter verwiesen. Der Pflegezustand dieser Listen nationaler Bibliotheken ist jedoch sehr unterschiedlich. Eine Liste von nationalen Bibliotheksgesellschaften in der Welt geben z. B. *UNECA*<sup>184</sup>, die *Lidicity Company*<sup>185</sup>, die Firma *theemedia*<sup>186</sup>. Das IPL (*The Internet Public Library*) hat eine Liste von größeren Bibliotheken in der Welt<sup>187</sup>.

Eine Liste des IFLANET (*International Federation of Libraries and Institutions*) zu nationalen Bibliotheken ist ebenfalls nicht auf dem neuesten Stand oder weitestmöglich vollständig<sup>188</sup>. Eine Liste der Links zu den Nationalbibliotheken von 134 Ländern findet sich im Anhang. Sie verwendet u. A. auch die Listen der *University of Queensland*<sup>189</sup> und von *WikiPedia*<sup>190</sup>. Am sinnvollsten wäre die *gemeinsame* Pflege einer solchen Linksammlung über WikiPedia.

<sup>179</sup>[http://www.unesco.org/webworld/portal\\_archives/pages/Resources/Internet\\_Resources/Archival\\_Legislations/Arab\\_States/index.shtml](http://www.unesco.org/webworld/portal_archives/pages/Resources/Internet_Resources/Archival_Legislations/Arab_States/index.shtml)

<sup>180</sup>[http://www.unesco.nl/images/09f4ffcbdaddeeb0eccdd3ac1a0662398charter\\_en.pdf](http://www.unesco.nl/images/09f4ffcbdaddeeb0eccdd3ac1a0662398charter_en.pdf)

<sup>181</sup>Sie wurden von der *National Library of Australia* für die *Information Society Division* der *United Nations Educational, Scientific and Cultural Organization* erarbeitet; <http://www.unesco.nl/images/guidelines.pdf>

<sup>182</sup>Abdelaziz Abid; UNESCO document *Preserving Our Digital Heritage; Principles and Policies*; [http://www.unesco.nl/main\\_6-3.php](http://www.unesco.nl/main_6-3.php)

<sup>183</sup>Z. B. die *Koninklijke Bibliotheek* zusammen mit der niederländischen nationalen Kommission für die UNESCO die *International Conference Preserving the Digital Heritage: Principles and Policies*; Den Haag, 4.11.2005; [http://www.unesco.nl/main\\_6-3.php](http://www.unesco.nl/main_6-3.php)

<sup>184</sup>[http://www.uneca.org/eca\\_programmes/it\\_for\\_development/library/Library\\_associations.htm](http://www.uneca.org/eca_programmes/it_for_development/library/Library_associations.htm)

<sup>185</sup><http://www.lidicity.com/lieguo/lgsk.html>

<sup>186</sup>[http://www.theemedia.com/libraryonline/national\\_libraries\\_address\\_m\\_2.htm](http://www.theemedia.com/libraryonline/national_libraries_address_m_2.htm)

<sup>187</sup><http://www.ipl.org/div/subject/browse/hum45.70.00/>

<sup>188</sup><http://www.ifla.org/VI/2/p2/national-libraries.htm>

<sup>189</sup><http://www.library.uq.edu.au/ssah/jeast/websites.html>

<sup>190</sup>[http://en.wikipedia.org/wiki/List\\_of\\_national\\_libraries](http://en.wikipedia.org/wiki/List_of_national_libraries)



## 3.6 Information in Wissenschaft und Lehre

Hier werden – notwendigerweise unvollkommen und kursorisch – weitere Materialien zu den relevanten Einzelthemen zusammengetragen, welche die Empfehlungen dieser Expertise unterstützen sollen.

### 3.6.1 Dokument-Typen und Proprietät

Eine nationale LZA-Policy will einen langfristig verlässlichen Rahmen für die Erhaltung und Nutzbarkeit kulturellen und wissenschaftlichen Erbes setzen. Dies ist eine besondere Herausforderung angesichts der raschen Entwicklung in den Formen und Wegen wissenschaftlicher Information und Kommunikation.

Nach den technischen Innovationen zu digitalen Dokumenten (Texte, Bilder, Filme etc.) und ihren Verknüpfungen (*Internet, World Wide Web, Email, XML*) werden neuerdings auch Instrumente entwickelt, die die inhaltliche Vernetzung ermöglichen wie Import/Export von Funktionalitäten (*webservices, SOAP*), die Kopplung verteilter Rechner, interaktive Programme und verteiltes Editieren (*Wiki, SVN, VRVS*).

Zunehmend werden alle diese Möglichkeiten in der Wissenschaft und Bildung auch genutzt werden, z. B. bei Blended learning, Lernmaterial-Kollagen, lebenden Dokumenten mit verteilten Autoren, und interaktiven Auswertungen auf den Rechnern der Nutzer bzw. der Erzeuger von Information.

Einerseits wird gegenüber dem Papierzeitalter durch die Archivierung *digitaler* Dokumente viel Geld gespart: die Vertriebskette und die Geschäftsmodelle vereinfachen sich, die gesamte Speicherung wird bequem, einfach und platzsparend, die mühsame Hand-Katalogisierung wird durch automatisch indexierende Suchmaschinen abgelöst und die Verbreitung von Dokumenten lässt sich problemlos realisieren.

Andererseits kommen aber möglicherweise auch sehr hohe Kosten neu hinzu: Die Zahl und der Umfang zu speichernder Dokumente wächst in der digitalen Welt - technisch nicht mehr eingeschränkt - sehr stark. Zunehmend kommen Dokumente hinzu, die technisch anspruchsvoller, komplexer, interaktiver und fragmentierter sind.

Eine geeignete Prozesskette der Dokumente für die Langzeit-Archivierung auszuwählen, zusammenzustellen sowie fachverständlich in den Kontext anderer Dokumente einzubetten, verlangt die direkte Einbeziehung von Fachwissenschaftlern.

Die Prozesskette Annehmen, Einlagern, Speichern, Migrieren und Spiegeln verlangt höherqualifiziertes Personal sowie im digitalen Zeitalter einen höheren Aufwand in den LZA-Institutionen.

Eine LZA-Policy, die ja Kultur- und Wissenserbe für eine zukünftige Nutzung bewahren soll, muss daher diesen Entwicklungen Rechnung tragen. Informationen aus den Archiven müssen so bereitgehalten werden, dass sie in zukünftige Szenarien importierbar und dort nutzbar sind.

Oft wurde die Archivierung in verschiedenen Stufen gesehen:

1. „Archiv anlegen“: Die Dokumente werden abgespeichert, „as is“. Digital wäre das also „Bitstream“. Im Papierzeitalter und davor wurde die Information stets im Klartext gespeichert, konnte also immer gelesen, wenn auch nicht immer verstanden werden (z. B. die Hieroglyphen). Das Verstehen, die Entzifferung ohne Informationsverlust konnte so auf eine spätere Zeit verschoben werden. Bei proprietär verschlüsselten Bitstreams hingegen sind die Daten unrettbar verloren, wenn der digitale Schlüssel für den Editor nicht mehr funktioniert. Dies würde daher eine neue Abhängigkeit von dem Fortbestand einzelner Firmen der Industrie, die diese Schlüssel als Geschäftsgrundlage besitzen, erzeugen - ein für eine *nationale* Langzeit-Archivierung nicht akzeptabler Zustand.

2. Im Papierzeitalter war die formale Erschließung auf der Ebene von „Sammlungen“ bzw. die rudimentäre inhaltliche Erschließung einzelner Objekte noch relativ kostengünstig.

Digital muss dagegen schon vorher ein gewisser Aufwand getrieben werden: Fachverständige Auswahl der Dokumente, technische Prüfung auf digitale Lesbarkeit und Darstellbarkeit in offenen Formaten, Hinzufügen von Metadaten, Annahme nur bei Beifügung der Quelltexte, Verlinkung durch die Bibliothek mittels Suchmaschinen, sowie die Prüfung der rechtlichen Zulässigkeiten (UrhG; Autorenverfügung). Wie bei E-Dissertationen bundesweit bereits erprobt, kann ein Teil dieser Arbeiten von anderen Institutionen (z. B. Universitätsbibliotheken) erledigt werden, die als Zwischenglied in der Prozesskette vom Autor zur LZA-Institution dienen. Im Papierzeitalter war das Sammeln das Problem. Im E-Zeitalter ist es die zukünftige Rückholbarkeit und Lesbarkeit.

3. „In Bibliothek einstellen“: Die papierenen Dokumente wurden lesbar und einsehbar gehalten und strukturiert beschrieben. Dem entsprach eine feine formale Erschließung auf der Ebene der einzelnen Objekte, bei wenig ausgeprägter inhaltlicher Erschließung und daher gemäßigtem Kostenaufwand je Dokument.

Bei digitalen Dokumenten entfällt dieser Schritt. Er ist durch den ersten Schritt erledigt. Auf dieser Ebene kommen aber neue verwaltungstechnische und kostenträchtige Aufgaben hinzu, wie die Migrierung, Spiegelung und Einbettung in Portale.

4. „Dokumentation“: Im Papierzeitalter wurde unter diesem Begriff verstanden, dass das Material aufbereitet und im Kontext dargestellt wird, z. B. über eine feine formale Erschließung durch Experten auf der Ebene auch von Teilen von Objekten.



Dies war in der Regel sehr kostenintensiv und war daher auf wenige als besonders wichtig angesehene Gebiete beschränkt.

Papierdokumente, einmal archiviert, bleiben jedenfalls lesbar (solange das Papier erhalten bleibt). Man kann daher die semantische Analyse auf später verschieben.

Digital sieht das ganz anders aus: E-Dokumente, einmal abgespeichert ohne die Programme zur sicheren Lesbarkeit mit abzuspeichern, sind unrettbar und ein für alle Mal verloren<sup>191</sup>. Als ersten Schritt haben daher am 18. Januar 2005 Musik- und Buchwirtschaft eine Vereinbarung mit Der Deutschen Bibliothek geschlossen. Sie darf künftig Musikaufnahmen, E-Books sowie Büchern beiliegende CD-ROM trotz Kopierschutz für wissenschaftliche und kulturelle Zwecke kopieren (siehe Pressenotiz<sup>192</sup>).

Digitale Dokumente werden in der Zukunft häufiger interaktiv in den aktuellen Wissenschaftsprozess eingebunden werden. Dies geschieht i. A. räumlich und personell verteilt in Wissenschaftsinstituten. Hierfür das kulturelle Erbe sinnvoll vorzuhalten, heißt, es auffindbar, rückholbar, lesbar, verstehbar für lange Zeit zu archivieren, und als Herausforderung den aktuellen Diskussionsprozess für die Langzeit-Archivierung zu kondensieren.

Eine nationale LZA-Policy wird daher die neuen Kernaufgaben einfordern: Die Dokumente für die Langzeit-Archivierung müssen geeignet ausgewählt werden; sie sollten möglichst bereits für die Aufnahme in das Langzeit-Archiv mit Metadaten versehen, in ein offenes Format gebracht, auf Vollständigkeit geprüft und in ihren Kontext eingebettet werden können. Die Organisation der verteilten Langzeit-Archivierung wird dem durch einen höheren Grad an gemeinsamen Standardisierungsanforderungen Rechnung tragen. Das Finanzierungs-Gesamtkonzept muss dem entsprechen. Die Kosten der Speicherung der Dokumente und Lagerung verschwinden dagegen hinter dem im E-Zeitalter notwendigen umfangreichen finanziellen Aufwand für eine sichere und rückholfähige (d. h. in der Zukunft Lesbarkeit und Verstehbarkeit sicherstellende) Langzeit-Archivierung.

Quantitativ ist daher nicht mehr die Zahl der archivierten Dokumente die höchste Hürde (sie wird in ungeahnte Höhen steigen), sondern die der geschilderten Begleittätigkeiten (technisch Sichern, Migrieren, Refreshen, Einbetten, semantisch Annotieren, d. h. suchfähig machen, für Diskussionen in der Zukunft importierbar machen usw.).

Eine ständige Konditionierung und Format-Migration der digitalen Archivierungsmaterialien, ihre fragmentierte Abrufbarkeit und ihre Integrationsfähigkeit in zukünftige Dokumente werden daher für eine Langzeit-Archivierung notwendig.

Daher ist z. B. die Verwendung von dem zur Zeit verbreitetsten proprietären Format *ACROBAT-pdf* oder gar von *WORD*-Texten für LZA-Textdokumente auszuschließen:

---

<sup>191</sup>Bei proprietären Formaten ist das immer so, da der Editor-Code aus geschäftlichen Überlegungen bisher weder der Bibliothek, ja nicht einmal dem Staat verfügbar gemacht wird.

<sup>192</sup><http://www.isn-oldenburg.de/~hilf/urhg/urhg-presse-ddb.pdf>

- Wissenschaftliche Informationen eines Dokumentes (z. B. von mathematischen Formeln) werden auf ihre graphischen Abbilder reduziert und sind so nicht mehr in neuem Kontext nachnutzbar.
- Laut *Adobe ACROBAT Professional* können Textpassagen als Zitate nur mittels eines weiteren proprietären Editorsystems (*WORD*) in neue Dokumente digital exportiert werden.
- Schon jetzt sind ältere *pdf*- Fassungen nicht mehr den Inhalt korrekt erhaltend migrierbar<sup>193</sup>.
- Die Langfristpflege und Verwendbarkeit des *pdf*-Formats durch die Industrie ist nicht gewährleistet.
- Eine Migrierung von *pdf* zu einem ASCII Text ist nicht mehr verlustfrei möglich (es kann kein Konvertierungsprogramm etwa von *pdf* zu  $\text{\LaTeX}$  geben) und ist von *ACROBAT* verständlicherweise nicht gewünscht. Damit aber würde die zukünftige Migrierbarkeit staatlich langzeit-archivierter Dokumente abhängig von einer jetzt florierenden kommerziellen Firma<sup>194</sup>.

Problematisch wären bei der Langzeit-Archivierung insbesondere proprietäre Formate, die nur von einzelnen wissenschaftlichen Disziplinen genutzt werden. Würde eine solche später als eigenständige Disziplin verschwinden, wäre das Wissen über die Lesbarkeit ihrer Daten möglicherweise verloren. Im Falle, dass diese Disziplin zu einem späteren Zeitpunkt wieder als wichtig betrachtet würde, wären dann Teile des Wissens zwar als Datenfluss vorhanden, aber inhaltlich unerschließbar.

Dass dies nicht abseitig ist, kann man sogar am *Browsing MathML*<sup>195</sup> sehen: Hier werden mathematische Formeln „optisch mnemotechnisch“ notiert und vereinbart. Die Verstehbarkeit durch den Leser steht und fällt mit der übereinstimmenden Fachausrichtung des Lesers *und* des Autors. Geht auch nur eine verloren, ist der Inhalt nicht mehr rekonstruierbar.

Es muss also verlangt werden, dass Dokumente, die in einem proprietären Format (etwa *Microsoft-WORD*) zur Langzeit-Archivierung angeliefert werden, bereits am „Eingang“ vom Autor/Erzeuger selbst oder seinem Institut, der Bibliothek seiner Universität oder der LZA-Institution in ein LZA-fähiges Format gewandelt werden. Zumindest sollte dann das Ergebnis der Migration durch Dritte (etwa seinem Institut, seiner Bibliothek oder einer LZA-Institution) vom Autor noch selbst autorisiert werden.

---

<sup>193</sup>Thorten Schwander; *LaTeX als Archivierungssprache*; SUB Göttingen 2002; Workshop MathDiss International; Ergebnisse und Visionen; <http://www.ub.uni-duisburg.de/mathdiss/Schwander/vortrag.html>

<sup>194</sup>PDF wurde einst zur Organisation der Druckvorstufe entwickelt, entsprechend bestand keinerlei Bedarf, Semantik zu integrieren.

<sup>195</sup>*Putting mathematics on the Web with MathML*; W3C; *Presentation MathML* ist eine Variante eines W3C-Standards, die die graphischen Notierungen von mathematischen Inhalten verabredet, also z. B. 'schreibe  $C^*$  und meine eine  $C$ -Stern-Algebra'; Man braucht also die professionelle Hintergrundkenntnis von Leser und Autor zum Verständnis; <http://www.w3.org/Math/XSL/>

Allerdings hat die Industrie den Markt der Langzeit-Archivierung erkannt: *Adobe Systems* plant<sup>196</sup> ein Format, das dem Archivierer mehr Kontrolle darüber geben soll, wie ein Dokument für eine Langzeit-Archivierung geeignet gespeichert werden kann. *PDF/A* tritt also als die proprietäre Konkurrenz zum *Extensible Markup Language Format XML* auf.

*PDF/A* könnte sich besonders für juristische Dokumente und Vertragstexte eignen, soweit es bei diesen auf die korrekte Erhaltung der äußeren Form ankommt.

*Adobe* hat den Code für *PDF/A*<sup>197</sup> ab Version 1.4 „für alle Zeiten“ offen gelegt, um mit dem offenen XML hier gleichzuziehen. Trotzdem ist bisher nicht sicher, ob sich wirklich eine stabile weltweit verteilte und von *Adobe* unabhängige Entwickler-Community bilden wird bzw. ob sie *PDF-A* als Standard akzeptiert und ob es möglich ist, dieses Format auch zukünftig weiterzuentwickeln, wenn es *Adobe* einmal nicht mehr geben sollte. *PDF/A* ist als *ISO Standard*<sup>198</sup> registriert worden.

Entsprechend werden den Autoren *Editoren* an die Hand gegeben, in dem diese ihre Texte erstellen und abspeichern können. Bei  $\text{\LaTeX}$  wird der Text mit allen Layout-Angaben in *ASCII*, also im Klartext abgespeichert und ist damit auf jedem Editor weltweit ohne Probleme weltweit und mit jedem Betriebssystem lesbar.

In neuerer Zeit versuchen kommerzielle Firmen, ihre eigenen Standards nach Anpassungen als „offen“ zu vermarkten. Ein Format wird aber nicht allein dadurch „offen“, dass man es z. B. „Office Open XML“ nennt. Dass MS das Format 'offenlegt', heißt nicht automatisch, dass es alle für ein offenes Format im Sinne internationaler Vereinbarung notwendigen Rechte darüber an die Nutzer- und Entwicklergemeinschaft abgibt. Insbesondere folgt daraus nicht, dass man in andere offene Formate konvertieren kann, oder dass man selbst Anpassungen vornehmen kann. Davon unterschieden ist *Open Office*<sup>199</sup>. Es wurde von SUN mit Unterstützung von Google entwickelt; Es kann der FSF<sup>200</sup> entsprechend 'auf ewig' von jedem benutzt, übertragen und weitergegeben werden. Das Format wird vom *OASIS Standard*<sup>201</sup> unterstützt. An der Entwicklung beteiligten sich u. A. das *National Archive of Australia*. Weiterführende links finden sich bei *WikiPedia* und bei D. Wheeler<sup>202</sup>.

<sup>196</sup>siehe White Paper of Adobe Inc.: *PDF as a Standard for Archiving*; <http://www.adobe.com/products/acrobat/pdfs/pdfarchiving.pdf>

<sup>197</sup>Markenname auch *X-pdf*.

<sup>198</sup> *Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1)*; Published Standard vom 28.9.2005; ISO (*International Organization for Standardization*); <http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=38920&ICS1=37&ICS2=100&ICS3\unhbox\voidb@x\bgroup\let\unhbox\voidb@x\setbox\@tempboxa\hbox{9\global\mathchardef\accent@spacefactor\spacefactor}\accent229\egroup\spacefactor\accent@spacefactor9>

<sup>199</sup>OO (*Open Office*); <http://www.openoffice.org>

<sup>200</sup>FSF (*Free Software Foundation*); <http://www.gnu.org/fsf/fsf.html>

<sup>201</sup><http://www.oasis-open.org/>

<sup>202</sup><http://de.wikipedia.org/wiki/OpenDocument> <http://de.wikipedia.org/wiki/OpenOffice.org> sowie <http://www.dwheeler.com/essays/why-opendocument-won.html>

### 3.6.2 Redundanz

Im Papierzeitalter wurde die Redundanz der Materialien durch Stationierung der entsprechenden physischen Zeitschriftenbände in möglichst vielen verteilten Bibliotheken realisiert. Dies war teuer, aufwendig und arbeitsintensiv (man muss ja die Vertriebswege etc. dazuzählen). Inzwischen sind aber viele Zeitschriften nur noch an wenigen Orten vorhanden.

Daher liest ein Autor meist nicht mehr die Originalquellen, die er gleichwohl zitiert; er müsste einen unverhältnismäßigen Aufwand treiben, sie sich tatsächlich zur Einsicht zu besorgen. Die Folge: Zitierfehler pflanzen sich oft über viele Autoren-Generationen fort<sup>203</sup>.

Digital ist die Redundanz im Prinzip einfach zu bewerkstelligen. Calliau, der Mit-erfinder des *World Wide Web*, bezeichnet die digitalen Kopien genau wie das Original sogar als „das Dokument“, weil das Wort Kopie nicht trifft: Jede E-Kopie eines E-Dokuments *ist* das Original. Insofern ist 'klonen' die treffendere Bezeichnung zur Abgrenzung vom physischen 'kopieren' bei Papier-gebundenen Dokumenten.

Um aber die Dokumente zu sichern, müssen sie gespiegelt (zu anderen Orten kopiert) werden. Das Spiegeln selbst ist preiswert. Die Vereinbarungen zwischen den Spiegel-Partnern müssen aber solide getroffen werden und durch den Rahmen einer nationalen Policy langfristig gesichert werden.

**File-refreshing** ist bei digitalen Dokumenten stetig begleitend notwendig. Das Material muss in neue 'state of the art' Techniken und Verfahren überführt werden. Denn Träger altern und die Anzahl statistischer Fehler wächst; Format-Leseprogramme wandeln sich.

Dies ist alles Routine, allerdings schreckt hier der schiere Umfang der notwendigen Arbeiten.

Um später effektiv mit den für die Langzeit-Archivierung vorgesehenen originalen Quellen arbeiten zu können, benötigt man aber, wie geschildert, zusätzlich:

- **Verknüpfung** durch *Links* mit anderen und insbesondere aktuelleren Dokumenten (die z. B. von semantisch arbeitenden Suchmaschinen gefunden werden können), die sogenannte Seitwärts- und Vorwärtsvernetzung;
- **Transfer** des *Inhaltes* in die jeweils aktuelle Wissenschaftssprache. Dies ist jedenfalls in der Physik und verwandten Gebieten wichtig, aber versteckter auch in anderen Wissensgebieten (Fachausdrücke oder Methoden geraten in Vergessenheit);
- **Intellektuelle Einbeziehung** in Reviews, die durch ihre geprüfte Qualität und ihre kompaktere Darstellung im Kontext die Nutzbarkeit des Wissens enorm er-

<sup>203</sup>Manuel Cardona und Werner Marx; *Verwechselt, vergessen, wieder gefunden*; Physik Journal 11/2004; <http://www.pro-physik.de/Phy/pdfs/ISSART21255DE.PDF>

höhen und für einen Leser in ferner Zukunft und anderem Umfeld das Lesen und die Arbeit erleichtern.

### 3.6.3 Metadaten

Zusätzliche LZA-Metadaten werden - neben den für Dokumente üblichen Metadaten (etwa den im *OAI registry*<sup>204</sup> registrierten Metadatensätzen) - benötigt, um die *Langzeit-Archivierung* sicherzustellen. Weitere Metadaten sind erforderlich und müssen speziell für die Langzeit-Archivierung international vereinbart und dann kodiert werden. Insbesondere benannt werden müssen so die Redundanz-Instanzen, die Zugriffserlaubnisse als Funktion der Zeit, die Software, die notwendig ist, um das Dokument auszu-lesen, die Versionsnummer und die Links zu weiterführendem Material.

Eine nationale LZA-Policy muss verlangen, ein als geeignet vereinbartes Minimal-Metadaten-Profil verbindlich für alle Mitglieder des nationalen Archivierungsnetzes festzuschreiben, um das Gesamtsystem funktionieren zu lassen. Das Niveau des nationalen Langzeit-Metadaten-Profiles sollte natürlich nicht hinter den bereits etablierten institutionellen der nationalen Bibliotheken zurückfallen (z. B. National Library of Australia, CEDARS<sup>205</sup>). Indem die nationale LZA-Policy dies als staatliche Aufgabe festlegt, erlegt sie dem Staat auf, dies zu finanzieren.

Schon 1998 hat Michael Day<sup>206</sup> auf der *Joint RLG and NPO Conference on Guidelines for Digital Imaging* einen Vortrag zu *Issues and Approaches to Preservation Metadata*<sup>207</sup> gehalten, der für die Definition von LZA-Metadaten richtungweisend ist und z. B. in der Preservation Policy von Australien als Richtschnur für die Festlegung der LZA-Metadaten verwendet wurde.

Nach M. Day sind LZA-Metadaten spezielle *administrative Metadaten*, um technische Informationen abzuspeichern, die der Langzeit-Archivierung digitaler Objekte dienen, um Migrations- und Emulations-Strategien zu bezeichnen, die Authentizität zu sichern, Rechte-Bemerkungen unterzubringen und zu vermerken, zu welcher Sammlung es im Archiv gehört<sup>208</sup>.

PREMIS<sup>209</sup> (*PREservation Metadata: Implementation Strategies*) hat im Mai 2005 einen *Final Report*<sup>210</sup> zu gegenwärtigen Praktiken und Trends bei der Langzeit-Archivierung kulturellen Erbes herausgegeben. Im Wesentlichen ist dies die Auswertung einer Fra-

<sup>204</sup><http://www.openarchives.org>

<sup>205</sup>*Cedars Guide to Preservation Metadata*; <http://www.leeds.ac.uk/cedars/guideto/metadata/guidetometadata.pdf>

<sup>206</sup>Michael Day ist Metadata Officer des UKOLN: (*UK Office for Library and Information Networking*); <http://www.ukoln.ac.uk>

<sup>207</sup><http://www.rlg.org/preserv/joint/day.html>

<sup>208</sup>*Selection Criteria, Guidelines, Decision-making Aids*; Joint RLG and NPO Preservation Conference; 1998; <http://www.rlg.org/preserv/joint/selection.html>

<sup>209</sup><http://www.oclc.org/research/projects/pmwg/>

<sup>210</sup>*Implementing Preservation Repositories for digital Materials: Current Practise and Emerging Trends in the Cultural Heritage Community*; <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>

gebogenaktion von etwa zwanzig Bibliotheken in aller Welt. Gefragt wurde nicht nach der *nationalen* Policy, sondern nach den technischen Strukturen der *institutionellen* E-Archive. Für die spätere Arbeit einer LZA-Policy-Expertengruppe ist es aber sicher eine Fundgrube. Neben dem vollständigen Bericht inklusive „Data Dictionary“ werden die Bestandteile und ergänzenden Beispiele auch einzeln zum Download angeboten. Die künftige Weiterentwicklung und Pflege des „Data Dictionary“ und der zugehörigen XML-Schemas wird durch die *PREMIS maintenance activity*<sup>211</sup> koordiniert.

### 3.6.4 Vernetzte Inhalte, lebende Dokumente

Auch die *organisatorischen Strukturen der Informationsflüsse* (das Informationsmanagement) *in der Wissenschaft* sind im Umbruch, um den Anforderungen aus Wissenschaft, akademischer Lehre und Kultur die neuen technischen Möglichkeiten nutzend bestmöglich zu entsprechen.

So wird das neue Förderprogramm des BMBF, *d-grid*<sup>212</sup>, in einem ersten Schritt die in einigen Gebieten der Naturwissenschaften schon immer eingesetzte Verkopplung von verteilten Rechnern bei großen numerischen Programmen (Astrophysik, Theoretische Physik, Geophysik) auf weitere Gebiete der Wissenschaft ausdehnen. Es wird aber auch die inhaltliche und organisatorische Vernetzung durch die Bildung verteilter Wissenschaftscluster fördern.

Speziell soll das Management für verteilte interaktive lebende (also ständig von den Autoren aktualisierte) Dokumente mit möglicherweise einer Mehrzahl wechselnder, verteilter Autoren unterstützt werden. Diese neuen, nun technisch möglichen und vermutlich sich rasch verbreitenden Dokument-Typen werden eine ganz neue zusätzliche Herausforderung für die Langzeit-Archivierung bilden. Sie lassen sich nicht langzeitarchivieren. Auch der Vorschlag, sozusagen als „Protokoll“ den jeweiligen Stand des Dokumentes durch Harvesting und Speicherung der Web-Server der involvierten Institute jeweils in festen Zeitintervallen abzubilden, ist wenig befriedigend. Ein Großteil dieses Materials aus mehreren Jahren ist dann fast inhaltsgleich (riesige redundante Speicherung), aber der intellektuelle Nachvollzug artet möglicherweise trotzdem in „archäologische Forschung“ aus und schließlich kann der intellektuelle Kontext ganz verloren gehen.

Stattdessen haben Wissenschaftler schon früh neue Wege beschritten: Tom Finholt beschrieb sein System CREW<sup>213</sup> in seinem Vortrag<sup>214</sup> auf der *IuK 99 Dynamic Documents*, der Jahrestagung der IuK (*Initiative Information und Kommunikation der wissenschaftlichen Fachgesellschaften in Deutschland*<sup>215</sup>). Bei diesem weltweit verteilten geophysikalischen Collaboratory-Großforschungsprogramm wird die wissenschaftliche Infor-

<sup>211</sup><http://www.loc.gov/standards/premis/>

<sup>212</sup>D-GRID Initiative; Förderprogramm des BMBF; <http://www.d-grid.de/>

<sup>213</sup><http://www.crew.umich.edu>

<sup>214</sup><http://elfikom.physik.uni-oldenburg.de/IuK99/iuk99-proceedings.html>

<sup>215</sup><http://www.iuk-initiative.org>

mation zeitnah über kommunikative Kanäle verbreitet, bei denen Beiträge, Annotationen, Bemerkungen und Kommentare Dritter aus dem Programm, möglich sind. Externe dürfen nur mitlesen (sogenannte „Lurker“).

Aus dem steten Strom der Kommunikation im Projekt heraus geschieht die Extraktion von wesentlichen Inhalten und die Würdigung der Leistungen des einzelnen Wissenschaftlers durch spezielle Experten des Programms, welche die Kommunikation verfolgen sollen und dann jeweils schriftlich thematisch zusammenfassen. Nur diese „Kondensate“ sollen die Brücke zur Zukunft/Nachwelt bilden. Sie müssen daher in sich vollständig und verständlich sein und dann langzeitarchiviert werden.

Wir vermuten, dass die Anforderungen aus der Wissenschaft sich rasch ausweiten werden auf die Projektion von fragmentierten, lebenden, mit dem Leser interaktiven, kommunikativen, stetig weiter entwickelten Inhalten auf archivierbare Dokumente sowie deren Integration in denkbare aktuelle kommunikative Programme und Vorhaben – eine Herkules-Aufgabe.

**Eine nationale LZA-Policy sollte dieses Kondensieren und Zusammenfassen von LZA-Wissen für die Nachwelt als weiteres eigenständiges staatliches Ziel in Fürsorge für die Information zukünftiger Bürger festschreiben und so dessen Aufbau und Realisierung sowie dessen Finanzierung durch staatliche Mittel sichern.**

### 3.6.5 Internationale Langzeit-Archivierungskonferenz der Physik

Viele Anforderungen aus der Wissenschaft an eine Langzeit-Archivierung sind schon seit langem klar und wurden immer wieder formuliert.

So gab es im Jahre 2001 in Lyon eine internationale Konferenz, organisiert von der IUPAP (*International Union of Pure and Applied Physics*) zur Langzeit-Archivierung verteilter Dokumente in der Physik<sup>216</sup>. Auf einen Beitrag aus den Reihen der Wissenschaftler wird verwiesen<sup>217</sup>.

Die anwesenden Wissenschaftler waren sich (im Gegensatz zu den anwesenden Bibliothekaren) einig, dass bei weitem nicht alles, was (auch referiert) veröffentlicht wird, für eine Langzeit-Archivierung geeignet oder gar sinnvoll ist. Sie befürworteten einerseits eine Filterung, andererseits eine Hinzunahme von (auch unreferierten) Online-Quellen und schließlich eine aktive (d.h. von Experten als Dritten auszuführende) Kondensierung und Präparation von Inhalten, so dass sie für die Nachwelt als kulturelles und wissenschaftliches Erbe lesbar, bequem nutzbar und verständlich sind. Vor dieser langfristig sinnvollen Kraftanstrengung schreckten natürlich die Zuständigen zunächst zurück.

In Lyon wurde als „Langzeit“ die Zeit „nach mindestens zehn Jahren“ definiert, für

---

<sup>216</sup>Long Term Archiving of Digital Documents in Physics; International Workshop; IUPAP International Union of Pure and Applied Physics; Lyon 2001; <http://publish.aps.org/IUPAP/>

<sup>217</sup> E.R.Hilf; *Physics Archiving: Requirements, Perspectives, and some Approaches in Germany*; <http://physnet.uni-oldenburg.de/~hilf/vortraege/lyon01/talk.html>

Zeiten also, in denen die jetzigen technischen Realisierungen überholt und durch andere ersetzt sind, und in denen die derzeit kommerziell tätigen Verlage und Firmen vielleicht nicht mehr existieren. Dabei ahnte man damals (2001) noch nicht viel von dem erst danach richtig einsetzenden Konzentrationsprozess in der Verlagswirtschaft mit der Einstellung vieler traditioneller Zeitschriften (z. B. möglicherweise bis zu 50% bei der jetzigen Zusammenlegung von Kluwer und Springer).

Interessant an der Tagung war, dass sich die Teilnehmer aus aller Welt – Wissenschaftler, Vertreter der Verlage der weltweit führenden großen Fachgesellschaften<sup>218</sup> und der Nationalbibliotheks-Experten<sup>219</sup> (USA, FR, UK, ...) einig waren in den Grundzügen der Policy für eine Langzeit-Archivierung. Dies schlug sich in entsprechenden Empfehlungen der EPS<sup>220</sup> nieder. Diese Ergebnisse seien hier zusammengefasst:

- Langzeit-Archivierung in den exakten Wissenschaften *ohne* Erhaltung von Original-Form, -Format, -Sprache etc., sondern allein des Inhalts;
- Verzicht auf Vollständigkeit der Archivierung des aktuellen Dokumentenraumes (Publikationen mit nur marginalem Inhalt sollten also garnicht erst langzeitarchiviert werden) zugunsten der Langzeit-Archivierung des vollständigen Inhaltes relevanter Arbeiten. An die Stelle der zur Zeit aus Platzmangel in Papierzeitschriften eher kursorischen Art, Publikationen zu schreiben, tritt die Forderung nach vollständiger Dokumentation des Inhaltes relevanter Arbeiten unabhängig vom Typ der Quelle (Journal, Instituts-Server, Zentral-Archiv, referiert oder nicht). So sollten also etwa Messreihen, mathematische Formeln oder Beschreibungen von Messanordnungen den Originalartikeln beigelegt<sup>221</sup> werden;
- Forderung nach Langzeit-Archivierung des gesamten Inhaltes in offenen oder offen dokumentierten nichtproprietären Formaten;
- Qualitätsprüfung der für die Langzeit-Archivierung vorgesehenen Dokumente *vor* der Annahme zur Langzeit-Archivierung auf Inhalt durch die Beteiligung von Fachwissenschaftlern.
- Vollständigkeit, sowie offenes dokumentiertes Format (bzw. der Konditionierung (Migration zu offenen Formaten etc. ) *vor* der Einlagerung);
- Klare Trennung der LZA-Gründe: Soll sie der Bewahrung wissenschaftlich-kulturellen Erbes dienen oder nur der Berufskarriere, des Ansehens und der Erinnerung an Personen? Nur das Erstere sei Staatsaufgabe;
- International zu vereinbarende Metadaten und deren strikte Anwendung bereits bei der Einlagerung;

<sup>218</sup>IoPP (*Institute of Physics Publishing*), UK; APS (*American Physical Society*), USA.

<sup>219</sup> DDB oder andere führende Bibliotheken aus Deutschland waren nicht auf der Tagung vertreten.

<sup>220</sup> *Long Term Archiving of Distributed Documents*; Recommendations of the European Physical Society; <http://www.isn-oldenburg.de/~hilf/IuK/arve/eps-ltadp1.html>

<sup>221</sup> Dies soll auch die Aufdeckung von Fälschungen erleichtern



- Leistungsfähige Retrieval-Werkzeuge und freier Zugang künftiger Generationen zum kulturellen Erbe.

Die IUPAP-Kommission für Langzeit-Archivierung hat einen Report<sup>222</sup> (vor der Konferenz) herausgegeben sowie eine Conference Summary<sup>223</sup>.

Die Kernpunkte dieses Reports, soweit sie für eine LZA-Policy wichtig und auch für andere Fachdisziplinen und Informationskulturen anwendbar sind, seien hier zusammengefasst:

- Ein *Registry* mit Testfunktionalitäten für alle verteilten Langzeit-Archive sollte als Frühwarnung fungieren, ob irgendein Archiv in der Gefahr ist, nicht mehr einwandfrei lesbar zu sein oder dass dort immer noch veraltete Hard- und Software benutzt wird, die bald nicht mehr funktionieren könnte. Bei einer solchen Alarmmeldung sollten dann Vorschläge für eine Migration auf einen sicheren anderen Server gegeben werden;
- Die *digitale Fassung* eines auch gedruckt vorliegenden Dokumentes sollte für eine Langzeit-Archivierung als *das zu archivierende Original* behandelt werden. Dem stimmten auch alle anwesenden Verlage zu.
- Die *Redundanz* eines Langzeit-Archivs wird als entscheidend angesehen: Grundsätzlich sollte ein Dokument niemals nur in einer LZA-Institution gelagert werden.
- Die *Redundanz sollte auch für die Zuständigkeiten* gelten: Es sollte stets für jedes Material mindestens zwei auch politisch unabhängige LZA-Institutionen geben, um die Gefahr eines (z. B. politisch motivierten) Vernichtens von Information des wissenschaftlichen und kulturellen Erbes zu verhindern.
- Es wurde eine Art *Zertifizierung* mit Mindestanforderungen der *LZA-Institutionen* gefordert (entsprechend z. B. der DINI-Zertifizierung von *Open Access* Dokumenten-Sammlungen) an Universitäten: Als Möglichkeit wurde das OAIS-Modell der CCSDS<sup>224</sup> genannt. Reine Spiegel oder lokale Backups wurden als Realisierung von Redundanz ausgeschlossen.
- Der Zugang und die Zugangsmöglichkeiten müssen offengehalten werden („helle“ statt „dunkle“ Archive). Nur Langzeit-Archive mit Zugang würden zur Bewahrung und zur Nutzung kulturellen Erbes beitragen.
- Neben der Langzeit-Archivierung von Inhalten wurde vorgeschlagen, auch *Dienste zu archivieren*, welche die Inhalte erschließen wie *Linking* oder *intelligente Suche*, bis es auf Grund der fortschreitenden technologischen Entwicklung leistungsfähigere gibt.

<sup>222</sup><http://www.iupap.org/finalrep.html>

<sup>223</sup>[http://publish.aps.org/IUPAP/ltaddp\\_report.html](http://publish.aps.org/IUPAP/ltaddp_report.html)

<sup>224</sup>CCSDS (*The Consultative Committee for Space Data Systems*); <http://public.ccsds.org/default.aspx>

- Den nationalen Bibliotheken wird eine zentrale Rolle für die *Konzeption* und *Organisation des Gesamtkonzertes der Langzeit-Archivierung* zugewiesen. Dazu müssten aber nationale LZA-Gesetze und Regelungen in den Staaten überhaupt erst einmal gesetzt werden.
- Neben den referierten Journalen sollten unbedingt *die auf Webservern lagernden Informationen und Dokumente, die LZA-relevant sind, mit gleichem Recht langzeitarchiviert werden*<sup>225</sup>.
- Die Langzeit-Archivierung von *Material aus Entwicklungsländern* wird ebenso als Problem gesehen wie der Zugang von dort zu den weltweit verteilten Langzeit-Archiven. Daher sollten letztere auch einen Dienst zur Email-Zusendung von nachgefragten Dokumenten bieten.
- Die Langzeit-Archivierung von *Email-Diskussionsarchiven*, wo man i. A. nicht nachträglich die Einwilligung des jeweiligen Autors einholen kann, wurde teilweise als wichtig angesehen, aber kontrovers diskutiert.
- *Internationale Standards und offene Formate* sollen die Migration in zukünftig neu aufkommende Formate verlustfrei sicherstellen. Das PDF-Format wird als sehr problematisch bezeichnet. Idealerweise sollten Formate nicht nur offen sein, sondern auch ohne Format-Beschreibung dekodiert werden können.
- Automatisiertes *Refreshing* und Speicherfehler-Korrektur wird gefordert.
- Bei Dokumenten mit mehreren *Versionen* sollten diese ebenfalls archiviert werden und zugänglich sein<sup>226</sup>.
- *Der Zugang zur Langzeit-Archivierung in der fernen Zukunft sollte frei sein* (Zugang zum kulturellen Erbe sollte für alle Menschen möglich sein).
- Die *Finanzierung der Langzeit-Archivierung wurde weitgehend als Staatsaufgabe gesehen*, schon weil die Verlage gar nicht so lange existieren.
- Schließlich wurde als unerlässlich angesehen, dass *die Beteiligten* (Autoren, Nutzer, Bibliothekare, Verlage) sehr viel enger als bisher und organisierter in Bezug auf die Langzeit-Archivierung *zusammenarbeiten*.

Leitlinien dieser Art müssen in eine nationale LZA-Policy aufgenommen werden, um einen verlässlichen Rahmen für die organisatorische und finanzielle Realisierung zu setzen.

<sup>225</sup>Die Filterung und Auswahl solcher Dokumente aus den Physik-Servern weltweit wird in PhysNet <http://www.physnet.net> durch neue intelligente Software unterstützt.

<sup>226</sup>Wie dies bei ePrint ArXiv seit 1991 erfolgreich praktiziert wird; <http://www.arxiv.org>

### 3.6.6 Zur Langzeit-Archivierung benachbarte Informationen

Obgleich die folgenden Quellen sich mit der Archivierung von digitalen Dokumenten befassen, nicht aber unbedingt mit den besonderen Herausforderungen der *Langzeit-Archivierung* und erst recht nicht mit den speziellen Problemen einer nationalen LZA-Policy für Deutschland, seien sie hier doch kurz angegeben. Sie sind (zum Teil auch schon etwas ältere) Pflichtlektüre für die direkt mit der Langzeit-Archivierung Befassen.

Linda Cantara hat 2003 für das DLF (*Digital Library Federation*), einem Zusammenschluss amerikanischer Universitätsbibliotheken, ein Kompendium *Archiving Electronic Journals*<sup>227</sup> herausgegeben. Wichtig sind hier vor allem die *Minimum criteria for an archival repository of digital scholarly journals*<sup>228</sup> sowie der *LOCKSS Report*<sup>229</sup> und der Stand des Programmes *Preservation of electronic scholarly journals*<sup>230</sup>.

Don Waters von der Yale University hat bereits 1995 erste Überlegungen zur Archivierung digitaler Information<sup>231</sup> formuliert. Die Empfehlungen von D. Waters et al.<sup>232</sup> sind auch als Buch erhältlich.

Zusammen mit RLG hat OCLC Arbeitspapiere<sup>233</sup> zu Kriterien für ein digitales Archiv im Januar 2002 zur Diskussion gestellt<sup>234</sup>.

Das CCSDS (*Consultative Committee on Space Data Systems*) hat ein *Reference Model for an Open Archival Information System (OAIS)* herausgegeben, das auf die Herausforderungen der Aufnahme, Speicherung, auf Management und Retrieval eingeht<sup>235</sup>.

Diese und weitere Arbeitsmaterialien könnten für die zukünftigen Beteiligten einer *Kommission zur Ausarbeitung einer LZA-Policy* von Interesse sein.

---

<sup>227</sup><http://www.diglib.org/preserve/ejp.htm>

<sup>228</sup><http://www.diglib.org/preserve/criteria.pdf>

<sup>229</sup>LOCKSS: *A Distributed Digital Archiving System Progress Report; For The Digital Library Federation Preservation Web Site; Mellon Electronic Journal Archiving Program; Stanford University Libraries; <http://www.diglib.org/preserve/stanfordfinal.html>*

<sup>230</sup><http://www.diglib.org/preserve/presjour.htm>

<sup>231</sup>*Some Considerations on the Archiving of Digital Information; <http://www.ifla.org/documents/libraries/net/waters1.htm>*

<sup>232</sup>Donald Waters and John Garret; *Preserving Digital Information; Report of the Task Force on Archiving of Digital Information; 5; 1996; <http://www.rlg.org/legacy/ftpd/pub/archtf/final-report.pdf>*

<sup>233</sup><http://www.rlg.org/pr/pr2001-attributes.html>

<sup>234</sup>*Attributes of a Trusted Repository for Digital Materials; RLG and OCLC Release Paper for Comment.*

<sup>235</sup>*Reference Model for an Open Archival Information System; CCSD Blue Book; Januar 2002; <http://public.ccsds.org/publications/archive/650x0b1.pdf>*

## 3.7 Verschiedene Sichten

Die Anforderungen an ein LZA-System werden von den beteiligten Parteien sehr verschieden gesehen. Eine **nationale Policy** soll darauf hinwirken, all diese Anforderungen und Rahmenbedingungen synoptisch zu integrieren, langfristig stabil zu vereinen und einen verlässlichen, langfristig stabilen Rahmen zu setzen.

**Innerhalb der Geschäftskette der Langzeit-Archivierung vom heutigen Autor über die das Archiv betreibende Institution mit ihren staatlichen Auftraggebern zu den zukünftigen Lesern haben die jeweils Beteiligten durchaus verschiedene Anforderungen, Sichten und Erwartungen, deren Integration und Sicherung als staatliche Aufgabe durch eine langfristig angelegte, strikte nationale Policy abgedeckt werden muss.**

Ein Sonderfall sind die klassischen Archive (siehe 3.2). Bei ihnen sind die Auftraggeber für die Archivierung oft zugleich die Erzeuger der zu archivierenden Dokumente und erst recht zum Teil auch die Nutzer derselben (z. B. Parlamentsprotokolle). Sie bedürfen daher einer eigens auf sie abgestellten Expertise. Die folgenden Absätze beziehen sich daher vor allem auf Wissenschafts- und Kulturarchive.

### 3.7.1 Sicht des Staates

Generell ist die Langzeit-Archivierung eine staatliche Aufgabe der Wissenschafts- und Kulturpolitik. Nur der Staat kann unabhängig von den wirtschaftlichen Entwicklungen und Abhängigkeiten und dem Schicksal kommerzieller Firmen eine langfristig sichere Archivierung garantieren.

Eine der Staatsaufgaben ist die Sicherung der langfristigen industriellen und intellektuellen (Wissenschaft, Bildung und Kultur) Wettbewerbsfähigkeit und Leistungsfähigkeit im internationalen Wettbewerb. Der Auftrag dazu ist, jederzeit und für sehr lange Zeit den Zugriff und das Verständnis des wissenschaftlichen und kulturellen Erbes, die dazu notwendigen Informationen, Ergebnisse und Erkenntnisse aus der Vergangenheit sicherzustellen.

Politisch ist also der Staat (in Zusammenarbeit mit den Ländern) zum Handeln und Gestalten aufgefordert.

Die Langzeit-Archivierung darf nicht von wechselnden kommerziellen Interessen abhängen und muss den Generationenvertrag daher als staatliche Aufgabe erfüllen. Sie ist nachhaltig und effizient durchzuführen. Die Wahrung der Integrität und Vertraulichkeit (soweit gesetzlich vorgeschrieben) aller Informationen in einem Archiv hat dabei oberste Priorität. Auch diese können nur vom Staat akzeptabel und parteienunabhängig gewährleistet werden.

Gleichzeitig ist der öffentliche Zugang zu den archivierten Informationen aus Bildung, Wissenschaft und Kultur ein wesentliches Charakteristikum einer lebenden De-

mokratie (siehe US-Nationalarchiv) und für den raschestmöglichen wissenschaftlichen Fortschritt notwendig.

Rahmenbedingungen: Die föderale Struktur der Bundesrepublik mit der weitgehend an die Länder vergebenen Kulturhoheit ist für die Entwicklung einer *nationalen* Policy eine besondere Herausforderung. Das Grundgesetz bestimmt, dass die Kulturpolitik Sache der Länder ist und sich der Bund auf integrative nationale Aufgaben beschränken muss.

Die im Prinzip eigenständig möglichen Policies der Länder und des Bundes müssen, um eine einheitliche, über viele Jahrzehnte stabile professionelle Gesamtstruktur bilden zu können, mit gemeinsamen einheitlichen Standards, Archivierungskriterien und Retrieval-Systemen synoptisch zusammengeführt und entwickelt werden. Dies kann nur in einem vorgeschalteten gemeinsamen Diskussionsprozess der Länder, des Bundes, den die spätere Langzeit-Archivierung tragenden Institutionen sowie von Experten aus der Wissenschaft und deren Fachdisziplinen geschehen.

Weitere Rahmenbedingungen sind: Die Bundesrepublik ist Teil der Europäischen Union. Daher würde sie einer möglichen zukünftigen *europäischen* LZA-Policy unterworfen. Sie sollte daher jetzt aktiv diesen Prozess mit anstoßen und bei deren Entwicklung mitarbeiten, um bestmöglich die eigenen Interessen zu wahren, und die sinnvollste Übertragung auf ihre föderale Struktur nicht zu verstellen.

Die Bundesrepublik ist Teil der *internationalen* Staatengemeinschaft. Entsprechend sollte sie in den notwendigen internationalen Gremien (z. B. im Rahmen der UNESCO) mitarbeiten bzw. diese initiieren, um so die internationale Entwicklung mit zu gestalten, die Erfahrungen im Ausland zu nutzen, der Verantwortung gegenüber kleinen und -Entwicklungsländern gerecht zu werden, und auch die notwendigen internationalen Spiegelungen eigenen Materials bestmöglich zu sichern.

Für die Gültigkeit der eigenen nationalen LZA-Policy müssen der *Verpflichtungsgrad* für die beteiligten Institutionen festgelegt und die *Durchsetzungsfähigkeit* entsprechend verankert werden.

Es muss der *Geltungsbereich* bestimmt werden.

In einer nationalen LZA-Policy sollen die politischen Rahmenbedingungen und das organisatorische Zusammenwirken der LZA-Institutionen sowie die generelle Infrastruktur und Architektur des LZA-Gesamtsystems beschrieben und festgelegt werden. Daraus ergibt sich dann zwingend die Notwendigkeit, einheitliche Standards, Kriterien und Schnittstellen zu definieren, die jedoch selbst nicht in einer nationalen Policy enthalten sein sollten, um für innovative Techniken offen zu sein und flexibel auf neue Medientypen reagieren zu können.

Der Gestaltungsspielraum von Policies der Bundesländer ist dadurch eingeeengt auf Auswahlkriterien für die Aufnahme von Dokumenten, auf die Wahl der an der Langzeit-Archivierung teilnehmenden und staatlich finanzierten Archivierungs-Institutionen, auf den Umfang der Beteiligung an Spiegelungs-Aufgaben dritter Länder und des Bundes, auf die eigene Beteiligung und eigenen Beiträge zur wissenschaftlichen Ent-

wicklung auf diesem Gebiet, auf die Ausbildung von Personal sowie auf Zuarbeiten ihrer Institutionen zur nationalen Langzeit-Archivierung.

Der Gestaltungsspielraum der Policy des Bundes ist entsprechend ebenso durch die föderale Struktur eingeschränkt.

Dementsprechend müssen die angeforderten Leistungen, die Kriterien und die Prinzipien mit den Ländern ausgehandelt werden. Der Nachteil dieses komplizierteren Weges wird möglicherweise ausgeglichen durch die breitere Vielfalt an Experten in der Diskussion und die dadurch sicher weit höhere Akzeptanz.

Eine breite und gemeinsame Entwicklung der Policy ist wünschenswert; muss aber die „Breite“ entsprechend der der politischen Institutionen (Bund/Länder) realisiert werden oder wäre es sinnvoller, die Beteiligten nach inhaltlichen oder praktischen Gesichtspunkten zu benennen?

Einerseits ist die föderale Struktur der Bundesrepublik ein entscheidendes, prägendes Element der politischen Rahmenbedingungen - aber es stellt sich die Frage, inwieweit die Entwicklung einer „nationalen Policy“ diesem Umstand Rechnung tragen muss.

Die Ergebnisse der gegenwärtigen Reformkommission des Bundes und der Länder zur Neuordnung der Bund/Länder-Zuständigkeiten müssen abgewartet werden. Sie wird vermutlich noch klarer die Zuständigkeit der Länder für die Kulturpolitik herausstellen. Die nationale Langzeit-Archivierung wird damit noch mehr Aufgabe der Länder.

Andererseits gibt es spezifische Bundesaufgaben: Die Einhaltung einheitlicher bundesweiter Standards, das Sicherstellen der Vernetzung der Archive, die internationale Einbettung, die Harmonisierung der Auswahlkriterien usw.

Bei den klassischen Archiven des Bundes, der Länder, der Kommunalbehörden besteht das Interesse der Auftraggeber an der Langzeit-Archivierung darin, die eigene zukünftige Arbeitsfähigkeit zu erleichtern (Verfügbarkeit von Vorgängen und Akten) bzw. Entscheidungen der Vergangenheit belegen zu können.

Durch die zur Zeit politisch angestrebte stärkere Arbeitsteilung des Bundes und der Länder wird die Integration noch mehr zur Aufgabe des Bundes werden. Je loser aber die Kopplung, um so stärker müssen die zusammenbindenden Regeln gefasst werden.

Daher empfehlen wir eine durchsetzungsstarke Formulierung der nationalen LZA-Policy in einem Bundesgesetz und nicht nur eine Sammlung loser Empfehlungen.

Die Abstimmung unter den Beteiligten wird durch ein solches Gesetz sicher befördert, auch wenn sie den einzelnen Institutionen manchmal als zusätzliche Last erscheinen mag.

### 3.7.2 Sicht der Erzeuger von Information

Die Erzeuger wissenschaftlicher Information haben als professionelle Wissenschaftler die dienstliche Aufgabe zu verlangen, dass ihre jetzt erzeugten wesentlichen wissenschaftlichen Ergebnisse langfristig verfügbar, lesbar, verstehbar und archiviert gehalten werden<sup>236</sup>.

Dabei gibt es bisher einen Zielkonflikt: Einerseits das Interesse des Autors, gelesen und zitiert zu werden. Dies könnte er oder ein von ihm Beauftragter digital leicht durch das „Aufs-Netz-Stellen“ auf einem beliebigen Server realisieren, möglichst natürlich mit einfachen Metadaten und in ein *OAI-Dataprovider Repository* verlinkt. Andererseits möchte er auch langfristig mit seinem wissenschaftlichen Beitrag sichtbar bleiben; bisher fehlen aber weitgehend die hierzu notwendigen LZA-Dienste.

Die Informationserzeuger wollen dabei für die ersten Jahre nach der Publikation über deren Sichtbarkeit und Verfügbarkeit selbst entscheiden können. Dies sollte ein LZA-System erlauben, soweit dem nicht gesamtstaatliche Interessen entgegenstehen.

Die Dokumente, Texte und Daten, allgemein die Informationen, werden meist zunächst für eine Übergangsfrist - eventuell entsprechend vom Autor vorgesehener Einschränkungen der Nutzungsrechte durch Dritte - im Archiv vorgehalten, aber nur nach Maßgabe der Rechte freigegeben. Dies gilt insbesondere in dem Fall, dass Autoren die Nutzungsrechte an einen Verlag abgetreten haben. Nach Ablauf einer Schutzfrist wird das Langzeit-Archiv dann die Information der Allgemeinheit zugänglich machen. Allerdings: Lässt der Autor zu, dass der Zugang durch seine Rechteabtretungen an Verlage zeitweilig beschränkt wird, so schadet er der Sichtbarkeit seiner eigenen wissenschaftlichen Leistung (Impact) und verlangsamt so auch den Fortschritt der Wissenschaft.

Diese Regeln sind zur Zeit in der Diskussion (siehe *Korb 2 UrhG-Reform BMJ*). Gleichzeitig wandeln sich die Policies der meisten (über 90 %) kommerziellen Verlage<sup>237</sup> dahingehend, dem Autor zu erlauben, seine Arbeiten auch selbst auf dem Server seiner Institution zu speichern. Hinzu kommt der immer größere Sektor der frei zugänglichen Volltext-Dokumente (*Open Access*).

Die Erzeuger intellektueller und wissenschaftlicher Information haben ein großes Interesse, dass ihre Werke durch ein LZA-System möglichst breit und auch langfristig verfügbar und nutzbar gemacht werden.

Hierzu gehört neben der Möglichkeit, auf die vollständige und unverfälschte Information zuzugreifen, auch deren Erschließung durch Hinzufügen eventuell erklärender Zusatzinformationen und die Möglichkeit, die Information zu durchsuchen. Das LZA-

---

<sup>236</sup>Das bildet sich in der Wissenschaftstheorie ab, dass den Wissenschaftsprozess gliedert in „Vorwissenschaftliche Erfahrungen auf dem Gebiet, auf dem man arbeiten will, sammeln; Begriffe definieren; Hypothese aufstellen; wissenschaftlich arbeiten (also je nach Gebiet 'experimentieren, argumentieren, schaffen, ..'); Ergebnisse darstellen; *Ergebnisse mitteilen* (sic!). Siehe auch: *The logic of modern Physics*; P. W. Bridgeman; MacMillan; New York; 1927;

<sup>237</sup>Sherpa Project: *Publisher copyright policies and self-archiving*; <http://www.sherpa.ac.uk/romeo.php>

System soll aus Sicht der Erzeuger sicherstellen, dass die vollständige und intellektuell unveränderte Information für die zukünftigen Nutzer verfügbar und verstehbar gehalten wird.

Die Autorenwünsche sind durchaus vielfältig. In verschiedenen Fächern sind außerdem durchaus sehr verschiedene „Kulturen“ vorherrschend.

Sichtbarkeit in der eigenen Fachwelt ist das strategische Ziel bei einer Veröffentlichung als Voraussetzung, zum wissenschaftlichen Erkenntnis-Fortschritt beizutragen. Entsprechend bestimmt sich daraus das wissenschaftliche Renommee der Autoren. Es wird daher versucht, den dadurch erreichten wissenschaftlichen Impact über die Zahl der Leser und die Zahl der Zitate zu messen. Das Renommee der Autoren bestimmt sich hieraus nachfolgend.

Es werden in weiten Bereichen der Wissenschaften Manuskripte auch über kommerzielle Verlage vermarktet und über sie vertrieben, und es werden Autorenhonorare gezahlt. Auch gibt es einen Randbereich von der Lehrbuch- und Populärliteratur bis hin zur Belletristik mit einem privaten Gewinnstreben der Autoren und der Finanzierung durch die privaten Leser.

Primär aber streben die Autoren nach wissenschaftlicher Anerkennung, dem Bekanntwerden ihres Beitrages in ihrer Fachwelt, also im weiteren Sinne nach dem Erkenntnis-Fortschritt ihres Faches. Sie wollen, dass ihre Dokumente langfristig, im Inhalt vollständig, sicher, integer, auffindbar, verbreitbar und weiterverarbeitbar archiviert werden. Teilweise wird es abgelehnt, sogar auch nur die (bescheidenen) Pauschalvergütungen der Verwertungsgesellschaften (*VG Wort*) anzunehmen.

Der Staat vertritt das Allgemein-Interesse der maximalen Förderung von Forschung und Entwicklung, um seine internationale industrielle Wettbewerbsfähigkeit zu sichern. Dies wird am besten durch ein offenes Verbreiten (*Open Access*) aktueller Dokumente über die lokalen institutionellen Server sowie zum Anderen durch die offene Langzeit-Archivierung von staatlichen Institutionen gemäß einer nationalen LZA-Policy erreicht.

Im übrigen bleibt es durchaus fraglich, ob die „Autoren“ jedes ihrer Dokumente für lange Zeit bewahrt wissen möchten, aber sie werden nicht mehr gefragt, wenn sie ihre Rechte bei der Publikation an einen Verlag abgetreten haben. Nach dem *Korb 2 Referentenentwurf des BMJ von 2004 zum UrhG* erhielten die Verlage sogar das Recht auf alle zukünftig neuen Nutzungsarten.

### 3.7.3 Sicht der Nutzer

Hier können wir nur aus den Anforderungen heutiger Nutzer extrapolieren. Dieses stößt jedoch bei der Einbringung digitaler Informationen in ein Langzeit-Archiv an Grenzen, da bisher erst wenige Nutzer digital erzeugte („born digital“) Information aus einem Depotsystem regelmäßig beziehen oder bereits in einzelnen Zugriffen bezogen haben.



Gespräche mit Wissenschaftlern (im Rahmen von Tagungen in 2003 und 2004), also mit aktiven Erzeugern und Nutzern von Information, haben gezeigt, dass hier bisher meist wenig Bewusstsein der Problematik des Zugriffs auf digitale Inhalte durch zukünftige Nutzer gebildet wurde<sup>238</sup>. Erstaunlich ist dies insbesondere, weil umgekehrt praktisch jeder befragte Wissenschaftler zahlreiche Beispiele aus dem eigenen Umfeld berichten konnte, in denen der Zugriff auf nur wenige Jahre alte digitale Information Probleme bereitete, teilweise sogar gescheitert war.

Zukünftige Nutzer von langfristig archivierter Information sind darauf angewiesen, dass diese für sie auffindbar, beschaffbar, lesbar und verstehbar ist. Notwendig ist, dass ein Archiv sie automatisch auch auf aktuelle Texte an anderen Orten hinweist, die den Inhalt des damaligen Dokuments in der heutigen Sprache wiedergeben und aufarbeiten. Diese Links innerhalb des Archives auf extern liegende aktuelle(re) Quellen sind essentiell für eine spätere breite Nutzbarkeit eines Archives und die zukünftige wissenschaftliche Arbeit zu den Originalinformationen.

**Langzeit-archivierte Dokumente müssen in die jeweils aktuelle Forschung und ihre Dokumente eingebettet angeboten werden, um in aktuellem Kontext und in der jeweils aktuellen Wissenschaftssprache verstanden werden zu können.**

In abgewandelter Form gilt dies für alle Wissenschaftsgebiete und Kulturerzeugnisse: Ohne die Fähigkeit, in aktuelle Notenschriften, Sprachen, Schriften und Symbole übertragen werden zu können, werden die langzeitarchivierten Originale, wird ererbtes Wissen und Kultur allmählich unverständlich.

Eine Besonderheit der Museen und klassischen Archive ist, dass die primären Objekte Unikate i. A. nicht-digitaler Art sind (Kunstwerke, Moorleichen, mittelalterliche Handschriften), die es physisch zu bewahren gilt. Aber gerade in diesem Bereich waren schon immer die sekundären Informationen - Transkripte, Fotos, Erläuterungen, Erkenntnisse über das Original etc. - ein integraler Teil des Archivs/Museums. Diese müssen mitarchiviert werden, um das Original zu verstehen und einordnen zu können. Für diese sekundären Informationen bietet sich die digitale Langzeit-Archivierung an. Um so mehr muss natürlich Sorge getragen werden, dass die Originale ebenso lange erhalten werden.

Die Langzeit-Archivierung der Werke zusammen mit ihrer Einbettung mag als sehr weitgehende Anforderung und Anstrengung erscheinen. Sie ist aber nach allem Gesagten unverzichtbar, um das Wissen langfristig verfügbar und nutzbar zu halten.

Entsprechendes wird in der Mathematik mit der Diskussion zu *Content-MathML* oder *Presentation-MathML* seit Jahren verfolgt. Der Hermes-Workshop<sup>239</sup> im Oktober 2004 versuchte, die Relevanz der Semantik und der Nachführung der Wissenschafts-

---

<sup>238</sup>In 2003 und 2004 wurden außerdem zwei hausinterne Umfragen zu Aspekten der elektronischen Fachinformation von den Autoren am Fachbereich Physik der Universität Oldenburg im Auftrage des IBIT (*Informations-, Bibliotheks- und IT-Dienste; Universitätsbibliothek*); <http://www.bis.uni-oldenburg.de>) durchgeführt, in denen u.a. die Motivation zur Verwendung von Metadaten und Bestandserhaltung erfragt wurde.

<sup>239</sup>*Hermes and Semantic Authoring with TeX and MathML*; Workshop; Golm-Potsdam 2004; <http://hermes.aei.mpg.de/w20041026/>

prache auch für andere Fächer zu diskutieren.

Natürlich ist die inhaltliche Bearbeitung nicht primär die Aufgabe von Bibliothekaren in einer Archivierungsinstitution. Zukünftige Nutzer haben möglicherweise einen anderen Ausbildungsstand, eine andere Gesamtsicht auf eine wissenschaftliche Frage. Es werden weiterführende Fakten aufgedeckt werden, die Wissenschaftssprachen werden sich weiterentwickelt haben, die Gewohnheiten, Schreibweisen und Begriffe und sogar der Kontext werden sich gewandelt haben. Der zukünftige Nutzer wird aber trotzdem verlangen, Informationen zu finden und in einer dem Autor nachempfundenen Sicht aktuell lesen und verstehen zu können. Er/Sie will, soweit sinnvoll, Inhalte importieren und weiterbearbeiten können. Er/Sie will das wissenschaftliche Erbe aufnehmen und weiterverarbeiten.

Eine technische Herausforderung ist das enorme Anwachsen der Quellen und Dokumente im Internet. Sie verlangen nach intelligenten, fachverständigen Suchmaschinen, nach einer Auswahl für die Aufnahme in die Archivierung und nach synoptischer Aufbereitung.

Dies ist (und war) von Bibliotheken nicht zu leisten, ist aber im neuen Kontext unabdingbar. Der einzige Ausweg ist, die wissenschaftlichen Institute und Institutionen und ihre Wissenschaftler selbst mit in den Prozess einzubeziehen - das Ende der inhaltlichen, räumlichen und personellen Trennung von Bibliotheken und Wissenschaftlern?

Zugleich wird (und muss) der Staat in einer Art Kosten/Nutzenanalyse den optimalen Aufwand für die Langzeit-Archivierung – bisher eher ein Nebenzweig der Bibliotheken – neu bestimmen. Abzusehen ist, dass diese Kosten sicher um eine Größenordnung höher sein werden als die bisherigen, dass dieses Geld aber gut angelegt ist und dabei nur einen kleinen Teil<sup>240</sup> der Wissenschaftskosten insgesamt beträgt. Insgesamt gibt es einen Trend zur Zunahme des Anteils der Kosten des Informationssektors an den Gesamtkosten.

Einigkeit herrscht in der Wissenschaft, dass für eine zukünftige Nutzbarkeit der Informationsgehalt des Dokumentes erhalten bleiben muss *und* der Kontext so rekonstruierbar ist, dass das Dokument verstanden werden kann. Rothenberg schreibt<sup>241</sup>

Dies hat technische Konsequenzen für die Filter zur Auswahl bei der Aufnahme von Dokumenten. Als Beispiel möge ein Dokument aus dem Fachgebiet Mathematik gelten: Dessen mathematische Formeln können in der international vereinbarten Kodierungssprache *Content Mathematical Markup Language* MathML als langfristig stabiler, offen kodierter Sprache abgespeichert werden. Nur dann lassen sich die mathematischen Formeln in der Zukunft vom Nutzer digital weiterverarbeiten (z. B. durch automatisierte Umwandlung in ein numerisches Programm zur Berechnung von Anwendungen und Beispielen, ohne sie manuell abschreiben zu müssen). Ergänzt werden muss

<sup>240</sup>Die Größenordnung wird im Bereich von 1% liegen als Anteil an den Gesamtkosten eines Institutes oder einer Universität.

<sup>241</sup>Jeff Rothenberg; *Preserving Authentic Digital Information*; 2000, Council on Library and Information Resources; <http://www.clir.org/pubs/reports/pub92/rothenberg.html>: „...preservation implies the *usability* of that which is preserved“.

die Langzeit-Archivierung mathematischer Texte durch eine getrennt zu archivierende Speicherung der formelsprachlichen Notation mathematischer Inhalte, da sich diese im Laufe der Jahre den jeweiligen Forschungsschwerpunkten und -Interessen entsprechend wandeln kann. In diesem mathematischen Sprach-Archiv sind dann die synonymen Formulierungen desselben Inhaltes ineinander überführbar. Entsprechendes gilt für andere Fächer<sup>242</sup>.

Die Nutzung etablierter, internationaler Standards erhöht die Wahrscheinlichkeit, dass Informationen in ferner Zukunft noch verstanden werden können. Eine Garantie sind sie jedoch keinesfalls. Jeder verwendete Standard ist zukünftig mit zu migrieren und nachzupflegen. So werden zur Zeit die Standards für die Langzeit-Archivierung, und erst recht die semantische Beschreibung von Inhalten erst noch entwickelt.

Bei den klassischen Archiven gibt es neben den Erzeugern der digitalen Dokumente auch die Öffentlichkeit als zukünftigen Nutzer. So haben die Bürger eines demokratischen Staates Interesse an einem möglichst offenen Zugang zu allen politischen Akten. Bei anderen Archivsparten dienen die Materialien der allgemeinen Information der Bürger. Auch hier ist bei der Langzeit-Archivierung digitaler Dokumente die Anforderung, dass die Werke dann auch gelesen und verstanden werden können.

### 3.7.4 Sicht der archivierenden Institutionen

Die Umsetzung bundesweit einheitlicher Standards der digitalen Bestandserhaltung und Verfügbarhaltung digitaler Inhalte setzt Standards für die sich hierdurch verändernden und erweiternden Berufsfelder voraus. Der Beruf des Bibliothekars ist sehr alt und hat sich immer wieder an neue Anforderungen anpassen müssen. Im Papierzeitalter war dies ein professioneller hoch entwickelter Beruf, der die Ablage und Bereitstellung von Büchern und Zeitschriften in Bibliotheken kompetent sicherstellen und die entsprechenden Register pflegen konnte. Die Aufgaben, Arbeitsplatzbeschreibungen, Kompetenzen etc. sind gewachsen und fortlaufend professionalisiert worden.

Für den Langzeit-Archivar im digitalen Zeitalter müssen das Berufsbild und seine Aufgaben erst noch entwickelt und die entsprechende Ausbildung international standardisiert werden. Zu seinen Fähigkeiten gehören nicht nur eine gute Grundkenntnis digitaler Techniken, eine professionelle Kenntnis der Dokumentenformen und -Transfertechniken und der Datenbankverwaltung sondern auch eine aktive Teilnahme an der weltweiten Entwicklung des Faches und neuer innovativer Dienste, ein enger Kontakt zu Arbeitsgruppen, die neue Techniken entwickeln oder erproben, eine gute Kenntnis der internationalen Standards digitaler Information und der Retrieval-Techniken. Da dies auch vom Persönlichkeitsprofil her *andere* Qualifikationen anspricht

---

<sup>242</sup>Z. B. Chemie mit *CML*, Physik mit *PhysML* und erst recht für die geistes- und gesellschaftswissenschaftlichen Fächer, wobei dort die entsprechenden Werkzeuge noch entwickelt werden. Die phantastische Möglichkeit, semantische Inhalte digital gespeichert dem Dokument beizufügen, wird die Wissenschaften effektiv und professionell in ihrer Arbeit unterstützen (z. B. Transskripte von altjapanischen Lettern, althochdeutsche Bild-Ikonen-Interpretationen etc.)

(oft überspitzt formuliert durch „vom Sammler zum Informationsjäger“, „vom Katalogisierer zum Kodierer“ von Inhalt mittels internationaler Standards, zu einer professionellen Kommunikation mit Autoren, Gutachtern, Editoren, Referenten und Lesern), ist der Umbau auch personalpolitisch und von der Umstellung der Ausbildung her eine enorme Herausforderung.

Hieraus folgt, dass Langzeit-Archivierung nicht eine Aufgabe von Institutionen entsprechend heutiger Archive allein sein kann, sondern möglicherweise neuer Organisationsformen und jedenfalls der Kooperation *verschiedenartiger* Einrichtungen bedarf.

Aus der Sicht der jetzt mit der Langzeit-Archivierung beauftragten oder noch zu beauftragenden Einrichtungen stellen sich einfache Grundfragen, die aber Antworten erfordern, die der nationalen Policy und den gesetzlichen Rahmenbedingungen nicht widersprechen dürfen:

1. Welche Formate dürfen Dokumente haben, die zur Langzeit-Archivierung angenommen werden? Welche Anforderungen bestehen technisch, welche sind innerhalb des Archives zu lösen, welche durch die Autoren vor der Ablieferung?
2. Welche Inhalte, welche Dokumententypen können überhaupt archiviert werden? Wer soll die Auswahlkriterien festlegen dürfen?
3. Welche Zusatzinformationen über die Dokumente sollen gespeichert werden in Bezug auf Kontexteinbettung und Metadaten? Wer soll diese nach welchen Standards erfassen?
4. Welche und wie detaillierte maschinenlesbare Inhaltsbeschreibungen sollen eingesetzt werden?
5. Wie oft und wie sind von wem die Archiv-Inhalte zu pflegen?
6. Müssen bereits anderweitig archivierte Dokumente gespiegelt werden? Welche Import- und Exportschnittstellen muss ein Langzeit-Archiv bieten, um mit externen Archiven kooperieren zu können?
7. Wie organisiert man (technisch, finanziell, juristisch) die Langzeit-Archivierung?
8. Mit welchen internationalen Partnern soll kooperiert werden (z. B. für die Spiegelung oder die Standards-Entwicklung)?

Diese Liste ist keinesfalls abgeschlossen. Weitere Fragen werden im praktischen Prozess der Realisierung auftreten, wie auch ein australischer Telefon-Gesprächspartner eindrucksvoll schilderte.

Die angestrebte nationale LZA-Policy und ein zugehöriges LZA-Gesetz müssen hier den notwendigen langfristig verlässlichen Rahmen setzen, damit die technischen und organisatorischen Realisierungen geordnet und zielgerichtet erfolgen können.

Die vom Staat zu treffenden Regelungen einer nationalen LZA-Policy für die Aufgaben der eingebundenen Institutionen und die Kostenverteilung müssen aus den anzustrebenden zukünftigen Szenarien abgeleitet werden. Dies erfordert einen intensiven Diskussionsprozess.

Der Blick ist im digitalen Zeitalter auf die komplexe Aufgabe langfristig in aller Zukunft sicherer *Informationserhaltung* zu richten anstatt sich wie im Papierzeitalter auf die Archivierung von Dokumenten 'wie sie sind' auf Vorrat und vielleicht auch nur für eine spätere Auswertung zu beschränken.

Weiterhin muss eine nationale LZA-Policy die föderalen Strukturen der Bundesrepublik beachten. Die hier vorgestellten Extrem-Szenarien einer rein zentralen sowie einer rein dezentralen Langzeit-Archivierung sind daher mehr als Abgrenzung gemeint, nicht als realisierbare Alternativen.

Unabhängig von der konkreten Umsetzung und Etablierung muss jedes LZA-System zahlreiche weitere Kriterien erfüllen, zu denen seine Stabilität und Flexibilität gehören.

**Stabilität des LZA-Systems:** Langfristige Stabilität eines nationalen Archivierungs- und Verfügbarkeitssystems erfordert eine stabile organisatorische Gesamtstruktur, eine auf Langfristigkeit ausgelegte Aufgabenverteilung bei technisch und archivarisches professionell ausgebildetem Personal. Dies erfordert bereits eine organisatorische Mindestgröße, um sicherzustellen, dass natürliche personelle Fluktuationen keinen Einfluss auf die Qualität und Verfügbarkeit der angebotenen Dienstleistungen haben. Entsprechend sind die Finanzierungswege fest zu etablieren und als hoheitliche Aufgabe zu budgetieren. Eine nationale Archivierungs-Policy digitaler Dokumente muss entsprechend robust gegen Störungen angelegt sein.

**Akzeptanz:** Ein LZA-System muss von Informationserzeugern (Autoren) und Nutzern (Lesern) als professioneller, sicherer und vertrauenswürdiger Dienst bekannt und akzeptiert sein. Hierzu bedarf es einer deutlichen Außendarstellung, insbesondere in der Startphase, in der ein entsprechendes Archivierungs-System zwar schon digitale Objekte einsammelt, hingegen zuerst nur wenige nachweist oder ausliefert.

Ein Beispiel: In einem Interview zur Langzeit-Archivierung und Verfügbarhaltung von Primärdaten mit Prof. Zielsny, Fachhochschule Gelsenkirchen, betonte dieser mehrfach, wie wichtig es sei, dass ein LZA-System *vertrauenswürdig* sei: „Die chemische und pharmazeutische Industrie würde ein LZA-System gerne nutzen [...], natürlich nur, wenn es bezüglich Vertraulichkeit - ich erinnere daran, dass Teile der Daten für Patente relevant sind – klare Abmachungen gibt [...] schon ein ausländischer Betreiber wäre wohl nicht akzeptabel, ideal wäre es als staatliche Aufgabe [...]“.

**Flexibilität:** Die nationale LZA-Policy setzt den Rahmen und die Anforderungen, damit ein LZA-System trotz sich rasch entwickelnder Techniken und Dokumentenformate stabil und nachhaltig funktioniert kann. Neue Techniken werden beispielsweise neue Dokumentenformate und -typen hervorbringen, die dann ohne Änderung der Stabilität und Verfügbarkeit des Dienstes in das Archiv aufzunehmen sind. Darüber hinaus werden auch neue Strategien und Verhaltensweisen der Autoren und Nutzer

in ihrer Informationsbeschaffung und in ihrer wissenschaftlichen Kommunikation das LZA-System vor sich stets wandelnde Nutzeranforderungen stellen. Dies bedeutet, dass eine Archivierungsstrategie, resultierend aus den Rahmenbedingungen einer nationalen Policy, sich bestmöglich auf teils noch unbekannte Szenarien der wissenschaftlichen Kommunikation und Information frühzeitig einstellen können muss, ohne den festen nationalen Rahmen zu verlassen.

Einige Entwicklungen zeichnen sich bereits ab:

1. Die rasche Zunahme nicht-textueller Information;
2. Die rasche Zunahme kommunikativer Information;
3. Die rasche Zunahme von Dokumenten, die ihrerseits als Bausteine andere Dokumente Online aufgenommen und eingebaut haben (*Patchwork*-Dokumente, die vor allem in der Lehre („blended learning“), aber auch zunehmend bei der Mitteilung wissenschaftlicher Inhalte vorkommen);
4. Die stärkere Verwendung von Summaries<sup>243</sup>, Review-Darstellungen von länger zurückliegenden primären Erkenntnissen, Artikeln und Primär-Informationen<sup>244</sup>;
5. Die erwünschte stärkere direkte Einwebung von langzeit-archivierten Informationen in aktuelle Darstellungen („parat sein“);
6. Die zunehmende Vernetzung von Wissensfragmenten und deren synoptische Zusammenfügung zum Verständnis<sup>245</sup>.

---

<sup>243</sup>IUPAP Workshop on *Scientific Misconduct and the Role of Physics Journals in its Investigation and Prevention*; London 2003; <http://www.iupap.org/working/workshop.shtml>

<sup>244</sup>a.a.O.; IUPAP Conference on Misconduct; London 2003.

<sup>245</sup>Siehe das *Content in Context* example by OCLC; <http://www.oclc.org/reports/2004contextdemo.htm>

## 3.8 Szenarien für eine Langzeit-Archivierung

### 3.8.1 Zentralistisches Szenarium

Wie die Beleuchtung der Aktivitäten in anderen Ländern zeigte, ist die Einrichtung eines zentralen digitalen Archivs, beispielsweise beim Nationalarchiv oder bei der Nationalbibliothek, verbreitet. Hier lassen sich viele Techniken und Verfahren aus Implementierungen anderer Länder übernehmen und nachnutzen (Frankreich, Niederlande etc.).

Das Problem einer zentralen Datenhaltung, die damit gegenüber äußeren Katastrophen anfällig und gefährdet wäre, lässt sich leicht durch die Einrichtung eines oder zweier „Außenstandorte“ umgehen, welche die Daten gespiegelt vorhalten, jedoch als reine Spiegel kein eigene Datenannahme, Konvertierung etc. betreiben.

Bei dieser Architektur gibt es also eine einzelne Institution, die für die Langzeit-Archivierung und Verfügbarhaltung aller digitalen Objekte einer Nation zuständig ist. Diese Institution ist idealerweise, auf Deutschland übertragen, eine bundesunmittelbare Körperschaft welche mit ausreichend Budget versehen ist, um Personal kompetent auszubilden, die anschließend das Archiv betreibt, die Daten technisch erhält und dokumentiert, so dass diese langfristig nutzbar und verstehbar bleiben. Der Vorteil ist: Die Expertise der Datenerhaltung muss nur einmal aufgebaut werden. Im Gegenzug besteht aber die Gefahr, dass sich damit eine Behörde abseits der aktuellen Entwicklung herausbildet, die auf Grund ihrer Abkoppelung nur schleppend mit der aktuellen technischen Entwicklung Schritt halten kann.

Die Integration der Archivierungsaufgabe in die Forschung auf diesem sich rasch entwickelnden Umfeld der wissenschaftlichen Information und Kommunikation wäre jedoch bei diesem zentralen Ansatz relativ einfach und effizient durch eine große eigene leistungsfähige Forschungs- und Entwicklungsabteilung der nationalen Institution zu etablieren.

**Die föderale Struktur der Bundesrepublik ist jedoch nicht abbildbar** in eine zentrale LZA-Architektur. Die Kulturhoheit der Länder würde hierbei verletzt werden, so dass entsprechende Verträge zwischen dem Bund und allen einzelnen Ländern zu schließen wären. Ob eine Föderalismusklausele in Anlehnung an DBibIG §25 ausreichend wäre, vermögen die Verfasser dieses Textes nicht abschließend zu beurteilen. Zugleich wäre es schwierig, die ja vorhandene breit verteilte Fachkompetenz zu nutzen.

### 3.8.2 Dezentrales Szenarium

Die Archivierung des kulturellen und wissenschaftlichen Erbes im Papierzeitalter wurde (bisher) sehr dezentral vorgenommen. Der Bund betreibt zahlreiche Archive mit verschiedenen Datenformaten und -typen, wie *Die Deutsche Bibliothek*, das *Bundesarchiv*,

den *Deutschen Wetterdienst* etc. Die Länder betreiben Staatsarchive und die DFG fördert Sondersammelgebiet-Bibliotheken. Hinzu kommen zahlreiche Instituts-, Institutionen- und Privat-Archive und -Bibliotheken, die bestenfalls mit einem der öffentlichen Archive kooperieren.

Diese sehr dezentrale Architektur spiegelt die föderale Struktur der Bundesrepublik wider und hat sich im Papierzeitalter mit seinem Fokus auf eine reine Sammlung von Dokumenten bewährt. Die Expertise der Archivierung ist auf viele Institutionen verteilt. Es gibt eine gewisse Redundanz im Sammelgut, da datenträgergebundene Information (beispielsweise papierene Bücher) einerseits von der Deutschen Bibliothek als Pflichtexemplar-Bibliothek, weiterhin bei Erscheinen in Bundesländern mit eigenen Pflichtexemplar-Regelungen von der Pflichtexemplarbibliothek des Bundeslandes und bei Zuordnung zu einem wissenschaftlichen Fachgebiet durch eine der Sondersammelgebiet-Bibliotheken gesammelt werden.

Eine zentrale Recherchemöglichkeit über all diese technisch heterogenen, in sich geschlossenen Archive konnte bisher allerdings nicht eingerichtet werden. Auch sind die Archivsysteme und die verschiedenen Sammelkriterien und -schwerpunkte für die Mehrzahl der Erzeuger von Information ebenso wie für die Nutzer wenig durchschaubar, weil heterogen und unkoordiniert. Eine stabile professionelle Standardisierung ließe sich nicht durchsetzen. Der Ausbildungsstand ist sehr ungleich. So ist teilweise in wichtigen regionalen digitalen Archiven nicht einmal der Umgang mit den verbreitetsten, international abgesichert standardisierten offenen Formaten oder erst recht dem in der Physik, Mathematik und weiteren Gebieten sowie in den internationalen Verlagen (wie z. B. Elsevier) intern eingesetzten  $\text{\LaTeX}$  vorhanden.

**Eine solche dezentrale Struktur skaliert nicht** mit der stetig steigenden Anzahl von zu archivierenden Objekten und der zunehmenden Komplexität originärer Objekte. Der Aufwand für die höherwertige notwendige Schulung der Mitarbeiter und den Aufbau der notwendigen fachlichen und technischen Expertise in jeder einzelnen Archivierungsinstitution steigt hierbei überproportional an, weil nun an sehr vielen Orten höchstwertige Spezialkompetenz zu sehr vielen Themen und technischen Gebieten vorgehalten werden muss. Er stößt sehr schnell an die finanziellen Grenzen aller beteiligten Institutionen und übersteigt das Budget für die Langzeit-Archivierung insgesamt.

Die Gefahr des Verlustes des Sammlungsgutes einzelner Institutionen ist kontinuierlich gegeben, da langfristig nicht auszuschließen ist, dass einzelne Institutionen gerade wegen der steigenden Kosten durch die Digitalisierung oder aus anderen Gründen geschlossen werden, bevor ein Transfer des digitalen Sammlungsgutes in eine andere Institution durchgeführt werden kann.

Dem möglichen Datenverlust sollte daher durch die regelmäßige, vertraglich gesicherte, vollständige Spiegelung an in- und ausländischen LZA-Institutionen entgegen gewirkt werden. Die damit verbundenen Kosten digitaler Speicherung (im Vergleich zu den jetzigen Kosten des Vorhaltens von Zeitschriften an verschiedenen Bibliotheken) sind außerordentlich gering. Der Speicherbedarf für das gesamte Weltwissen zur



Mathematik etwa passt digital auf wenige Datenträger im Wert von einigen hundert Euro.

Bei traditionellen Archiven im Papierzeitalter war ein Informationsverlust selten (etwa bei Brand oder Wetterschaden in der Bibliothek). Rein digitale Information hingegen ist dagegen relativ leicht auch schon durch geringfügig erscheinende Nachlässigkeiten löscher, und die Gefahr ist wegen des ständigen Hantierens mit ihr (Refreshing, Migration) stets gegeben. Genau hier sollte eine nationale Policy ansetzen, so dass die Daten geordnet an andere Archive und archivierende Institutionen gespiegelt werden. Eine Redundanz der Datenhaltung für den Fall des Zusammenbruchs der regulierenden öffentlichen Ordnung verlangt zudem eine *internationale* Spiegelung aller Daten – in einem rein dezentralen heterogenen System nicht realisierbar.

**Beide Architekturen sind für den Aufbau eines LZA-Systems für das digitale kulturelle und wissenschaftliche Erbe der Bundesrepublik Deutschland ungeeignet.**

### 3.8.3 Dezentrales Szenarium mit Leitungsstrukturen

Basierend auf dem Vorbild Australiens bietet sich für die Bundesrepublik Deutschland die Entwicklung eines digitalen Langzeit-Archives als dezentrale Struktur mit einem koordinierten Netz von Leitungsinstitutionen an.

Dieses Modell sieht vor, dass der Bund mehrere seiner bereits bisher mit der papierenen Langzeit-Archivierung beauftragten Institutionen zu LZA-Institutionen für digitale Informationen ausbaut. Die Bundesländer können und sollen ihrerseits Institutionen wie die Staatsbibliotheken mit der Beteiligung an dem LZA-Netzwerk beauftragen. Die Einbindung der fachlichen Sichten, die wesentlich für die Wissenschaftlichkeit eines entsprechenden Netzwerkes sind, werden durch die Sondersammelgebiet-Bibliotheken und Archive eingebracht.

Das somit gebildete Gremium von Leitungsinstitutionen sollte die technischen und internen organisatorischen Anforderungen an die Informationsflüsse und Verarbeitungen im LZA-Netzwerk definieren und über ein kontinuierliches Auditing umsetzen.

Das Netzwerk sollte offen für die Einbindung weiterer Teilnehmer sein, insofern sie Inhalte entsprechend der im Netzwerk verbindlich durch die nationale Policy einzufordernden stabilen Standards einbringen wollen.

Entscheidend ist, dass das Netzwerk intern eine redundante Datenspeicherung gewährleistet und sicherstellt, dass bei Wegfall einzelner Institutionen ein „Fall-Back-Szenario“ greift, und die digitalen Daten bereitgestellt und in andere Archive überführt werden.

Bezüglich des Auditing erscheint eine hierarchische Struktur innerhalb des Netzwerkes notwendig. Dabei sollte entsprechend der Auditing-Standards die höherstehende Institution regelmäßig prüfen, dass die direkt untergeordneten Archive Schnittstellen so

anbieten, dass jederzeit ein „Fall-Back-Szenario“ technisch funktioniert und dass Spiegel der Daten vorhanden sind.

Die Expertise zum Betrieb eines Langzeit-Archives muss bei dieser Architektur nur in wenigen Institutionen vorhanden sein, die dann jederzeit über definierte und implementierte Schnittstellen jene Daten von anderen Archiven übernehmen können, für deren weitere Pflege dort die Expertise oder die finanzielle Ausstattung fehlt.

Der nur in einer digitalen Welt mögliche Vorteil dieses Szenariums ist, dass es den gesetzlichen Rahmenbedingungen und der Aufteilung in Bund-Länder-Hoheiten entspricht. Das Netzwerk ist hinreichend flexibel, um zeitnah auf aktuelle technische Entwicklungen reagieren zu können. Es gibt durch die hierarchische Struktur die Garantie einer ordentlichen Archivführung und langfristigen Stabilität. Umgekehrt wird durch die Offenheit für die Aufnahme neuer Partner gewährleistet, dass sich keine starre zentrale Behörde herausbildet. Dies ist eine allgemeine Gefahr, da die Flexibilität einer Institution i. A. mit seiner Größe abnimmt, ebenso wie seine Außenkommunikation im Verhältnis zur Binnenkommunikation. Ein Netzwerk dagegen skaliert mit einer beliebig wachsenden Anzahl digitaler Archivbestände und erleichtert die Interaktion aller Beteiligten.

Die mit der Langzeit-Archivierung beauftragten verteilten Institutionen sind bei diesem Organisationsmodell einfacher an die aktuelle, breit gestreute Forschung auf dem Gebiet der Archivierung und digitalen wissenschaftlichen Publikation ankoppelbar.

Die Abgrenzung zum rein dezentralen Szenario ergibt sich durch die Skalierung: Ohne zentrale Organisations-, Regelungs- und Standardisierungs-Vereinbarungen gäbe es einen mit der Zahl der beteiligten Institutionen quadratisch ansteigenden<sup>246</sup> Bedarf an bilateralen Verträgen, Abmachungen, Koordinierungen, Kommissionen etc.

Die Abgrenzung zum zentralen Szenario besteht insbesondere in der Einführung einer Hierarchie für Auditierungs-Kontrolle und eines „Fall Back“. Dieses Modell wurde von einem australischen Telefongesprächspartner basierend auf den Erkenntnissen und Erfahrungen des praktischen Betriebes eines LZA-Systems dringend empfohlen. Es vereinigt die Vorteile des dezentralen Szenarios mit seiner Skalierbarkeit und Offenheit mit den für einen langfristig stabilen Betrieb notwendigen Strukturen.

In diesem dezentralen aber mit straffen Leitungsstrukturen versehenen Szenario, das durch eine strikte nationale LZA-Policy und nachfolgend ein LZA-Gesetz abgesichert ist, lassen sich auch Grundfragen aus der Sicht der archivierenden Institution beantworten:

1. **Dokumentenformate und Aufgabenteilung mit den Autoren:** Jeder der Partner im Archivnetzwerk ist frei, eigene Dokumentenformat-Auswahlen für das jeweilige Facharchiv zuzulassen. Insbesondere fachspezifische Formate können in dafür professionellen speziellen Archiven verwaltet werden. Hier wird dann eine einem allgemein verarbeitbaren Schema entsprechende Version dieser speziellen

<sup>246</sup> Also  $(n^2 - n)/2$  bei  $n$  beteiligten Institutionen, also im Wesentlichen quadratisch ansteigend.

Dateien erzeugt (Beispiel: XML-Repräsentation von Messdaten), die der Sicherung der Transferierbarkeit auf ein anderes Archiv dient. Solange das Facharchiv mit der fachlichen Expertise existiert, können die Daten dort als lebendes Archiv mit all den zu ihrer Interpretation notwendigen Diensten angeboten werden. Die Aufgabenteilung zwischen Erzeugern und Archiv kann fachspezifisch gelöst werden.

2. **Auswahlkriterien:** Die Auswahlkriterien sollten in einem Netzwerk nach allgemeinen Kriterien festgelegt werden. Hierbei sollte für jeden Objekttyp die Frage nach der potentiellen Relevanz für die Nutzer in der Zukunft gestellt werden. Die Beurteilung, ob ein Objekt relevant ist, kann dabei unter fachlichen Aspekten und lokal beantwortet werden. Auf einschlägigen internationalen Konferenzen wiesen Wissenschaftler<sup>247</sup> immer wieder darauf hin, dass z. B. ein Großteil der aktuellen referierten Zeitschriften-Artikel zwar archiviert wird, aber wegen ihres nur marginalen Inhalts nicht langzeit-archiviert werden sollte.

Bisher besteht oft nur eine mehr formale Aussonderung/Kassation durch Archivare/Bibliothekare/Dokumentare mit Kriterien wie etwa: „Speichere alle Dissertationen der Universitäten“.

Bei der zur Zeit wieder viel diskutierten Frage der Archivierung von Diplomarbeiten und anderer Examensarbeiten ist eine inhaltliche Auswahl notwendig: Einerseits sind nicht alle Examensarbeiten publikationswürdig und -fähig. Außerdem sind die Zustimmungen des Autors *und* des Prüfers/Betreuers notwendig<sup>248</sup>.

Andererseits werden von der ständig steigenden Gesamtmenge an auch wesentlichen<sup>249</sup> wissenschaftlichen Publikationen, die auf den *Open-Access*-Servern liegen<sup>250</sup>, bisher kaum Dokumente in die Filterung zur Langzeit-Archivierung eingebracht und entsprechend konditioniert. Es droht daher ein wesentlicher Teil des aktuellen wissenschaftlichen Diskurses für die Nachwelt verloren zu gehen. Auch hier könnten dezentrale Akteure fachspezifische Crawler und Filter viel

<sup>247</sup>a) Claus Montonen; Zitat aus seinem Vortrag: „Entscheiden, ob und was gespeichert werden soll. Die Karrierewünsche sind kein kulturelles Motiv. Daher extensive und vielfältige Filterung.“; IUPAP-Conference *Long Term Archiving of Distributed Digital Documents*, Lyon, 2001; <http://www.aps.org/IUPAP>  
b) Jean Zinn-Justin; Vortrag zur Kompression der marginalen Publikationen; IuK-96 Workshop *Neue Medien in den Wissenschaften*; München 1996.

c) Jean Zinn-Justin; *Peer review and electronic publishing*; 1997; In : *The Impact of Electronic Publishing on the Academic Community*, Session 3; *The content and quality of academic communication, Peer review and electronic publishing*; International Workshop; Academia Europaea and the Wenner-Gren Foundation; Wenner-Gren Center, Stockholm April 1997; <http://www.portlandpress.com/pp/books/online/tiepac/session3/ch3.htm>

<sup>248</sup>Die Publikation des Inhaltes der Arbeit ohne den Label „Examensarbeit“ ist natürlich dem Autor freigestellt.

<sup>249</sup>So publiziert der zur Zeit berühmteste lebende theoretische Physiker t’Hooft fast nur noch auf nicht referierenden OA-Servern.

<sup>250</sup>Als *Open Access* gelten alle Dokumente, die in Volltext frei über das Internet lesbar sind - seien sie vom Erzeuger selbst auf seinen eigenen Web-Server gestellt (*individual archiving*) oder in den Server seines Instituts oder Hochschule eingebracht (*institutional archiving*) oder zu zentralen Archiven gesandt (*central archiving*), unabhängig davon, ob sie referieren oder nicht.

kompetenter einsetzen und so LZA-wertes Material aus dem Web in die nationale Langzeit-Archivierung einspeisen.

3. **Erfassung und Standards der Zusatzinformationen:** Beschreibende Zusatzinformationen (Metadaten) spalten sich in inhaltsbeschreibende, technische und kontextuelle. Die übergreifende Zusammenstellung einer *Metadata-Set-Definition* erlaubt es, eine Kompatibilitätsschnittstelle zwischen den in einem Netzwerk zusammengeschlossenen Fach- und Spezialarchiven zu entwickeln, ohne hierbei die einzelnen Archive in ihrer inhaltlichen Tiefe der Beschreibung zu begrenzen. Die frühzeitige Kooperation und das Verabreden von Standards zur Beschreibung des digitalen Archivgutes werden später die aufwendige oder gar unmögliche Entwicklung sogenannter „Crosswalks“ zur (immer mit Verlust behafteten) Überführung von Metadaten zwischen verschiedenen Standards vermeiden helfen. Noch gibt es in vielen Archiven nur wenig digitales Archivgut, oder zumindest nur wenige Typen und Genres digitalen Archivgutes, so dass sich die noch fehlenden Absprachen mit absehbarem Aufwand sicher besser früher als später treffen lassen.
4. **Maschinenlesbare Inhaltsbeschreibungen:** Über die Art und inhaltliche Tiefe sowie die Kodierung von Inhaltsbeschreibungen sollten die kooperierenden Institutionen einen Mindeststandard definieren können, ohne inhaltliche Brüche in den Dokumentenbeständen zu verursachen. Allerdings ist dies absehbar eine der konzeptionell schwierigeren Aufgaben.

Der Aufwand im Papierzeitalter ergab sich aus der möglichst feinen Katalogisierung und Einordnung (wie den RAK- und MAK-Indizierungen<sup>251</sup>).

Bei E-Dokumenten sind die Probleme andere: Die Katalogisierung kann weitgehend entfallen; die Erschließung wird heute besser durch Volltext-erschließende Suchmaschinen automatisiert vorgenommen, die einen sehr feinen Index von großen Dokumentenmengen extrem kostengünstig erstellen können.

Stattdessen ergeben sich neue Probleme: Die Lesbarkeit, Authentizität, Verstehbarkeit der Dokumente muss langfristig erhalten bleiben. Dies verlangt wiederum angesichts der viel größeren Dokumentenmengen und der stetig wachsenden Zahl unterschiedlicher oft proprietärer Formate und neuer Dokumententypen einen andersartigen, möglicherweise finanziell vergleichbaren oder höheren Aufwand als früher.

5. **Import- und Exportschnittstellen, Kooperation mit Institutionen außerhalb des Netzwerkes:** Die Offenheit eines LZA-Netzwerkes, jederzeit weitere Partner aufnehmen zu können, Daten von Archiven und Sammlungen aufnehmen zu können, die diese gespiegelt haben möchten und die den inhaltlich qualitativen Ansprüchen eines Langzeit-Archives entsprechen, bedarf der Definition von Schnittstellen für den Import von Archivgut. Diese Schnittstellen resultieren aus der im

<sup>251</sup>Eine Zusammenstellung der online zugänglichen Informationen; *Zugang zum Wissen*; News 24.1.2006; <http://www.zugang-zum-wissen.de/news-01-06.html>

LZA-Netzwerk eingesetzten internen Technik und Organisation, die beispielsweise auf OAIS (*Open Archival Information System*<sup>252</sup>) beruhen. OAIS definiert bereits Exportschnittstellen, die einerseits geschaffen wurden, um das Archiv für den Bürger über beispielsweise Web-Schnittstellen oder Retrieval-Funktionalitäten nutzbar zu machen, andererseits dem Export großer Datenmengen im Rahmen einer internationalen Zusammenarbeit dienen zu können.

Sollte jede Institution für sich alle Standards selbst lernen und einhalten sowie die Schnittstellen technisch implementieren, würde insgesamt jeder nationale Budgetrahmen gesprengt. Im Verbund jedoch ist dies einfach möglich, da die Datenhaltung transparent gespiegelt werden kann und die Expertise unter den Partnern weitergereicht aber lokal weiterentwickelt werden kann. Hierzu sind regelmäßige gemeinsame und offene Workshops und Schulungen sinnvoll, um die durch die Vernetzung entstehende Synergie optimal abzuschöpfen.

6. **Technische, finanzielle und rechtliche Organisation der Langzeit-Archivierung:** Technisch ist die Langzeit-Archivierung sehr anspruchsvoll und aufwendig, denn fortlaufend sind die Datenträger zu erneuern (Refreshing) und gegebenenfalls auf andere Formate zu migrieren, um die Unversehrtheit des konservierten Datenstroms zu gewährleisten.
7. Hierzu ist die aktuelle Entwicklung sorgsam zu beobachten. Diese Aufgabe kann in einem Netzwerk stellvertretend für alle beteiligten Archive von wenigen Partnern übernommen werden, so dass hier ein hoher synergetischer Gewinn entsteht. Als Beispiel mag das *TeXDoc Center*<sup>253</sup> dienen.

Finanziert werden muss die Langzeit-Archivierung digitaler Information analog der von papierener Information durch den Staat als hoheitliche Aufgabe. Entgelt für die Bereitstellung von Archivgut kann einen kleineren Teil der Kosten decken, analog zur jetzigen Regelung der DDB für Papier-Dokumente, wenn dies der Politik der Institution entspricht. Aus Sicht der Nutzer in der Wissenschaft ist ein freier Zugang (Open Access) für die wissenschaftliche Arbeit am effektivsten.

Mit wesentlichen Einnahmen wäre jedoch erst zu rechnen, wenn die Information im Archiv nur noch dort zu finden und zu nutzen ist, also erst nach einigen Jahren des Aufbaus. Diese Verknappung von Information entspricht aber gerade nicht der Intention des Staates auf möglichst ungehinderten Zugang der Wissenschaft zu den Quellen, entsprechend der bisherigen Einsicht in Bücher an den Bibliotheken.

Sinnvoll und realistisch wäre es also, nur Einnahmen für wesentliche Mehrleistungen (aufbereitete und bearbeitete Inhalte in ansprechender Präsentation) oder für den Verkauf von Informationen an kommerzielle Kunden zu fordern.

Klar ist: Der Zugang zu den Informationen muss für Privatpersonen und die

<sup>252</sup>NASA/Science Office of Standards and Technology (NOST): *ISO Archiving Standards - Reference Model Papers*; [http://ssdoo.gsfc.nasa.gov/nost/isoas/ref\\_model.html](http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html) und *Overview*; <http://ssdoo.gsfc.nasa.gov/nost/isoas/>

<sup>253</sup>TeX Document Center; SUB Göttingen; <http://www.texdocc.org/>

Wissenschaft frei bleiben, soweit es sich um „freies Material“ für den wissenschaftlichen Fortschritt, für die akademische Ausbildung und Kultur nützliche Dokumente handelt, um dem eigentlichen Auftrag einer Langzeit-Archivierung für zukünftige Generationen gerecht zu werden.

Der aktuelle Stand zu Open Access sowie Handlungsanleitungen für Autoren und Leser wurden von uns auf einer separaten Webseite<sup>254</sup> zusammengestellt.

Rechtlich ist für die Langzeit-Archivierung zu klären, ob es für das jeweilige Archivgut das Recht zur Archivierung, zu technisch notwendiger Migration und zur Zurverfügung-Stellung an die Nutzer gibt und wie dieses langfristig gesichert werden kann, insbesondere vor dem Hintergrund von DRM (Digital Rights Management)-Techniken. Diese sollen ja den Zugriff auf Information nach Ablauf des digitalen Schlüssels auf Wunsch kommerzieller Verleger technisch unmöglich machen. Hier ist zu klären, wie es Archiven im staatlichen Interesse zukünftig erlaubt sein soll, diese Schlüssel zu umgehen oder zu „knacken“. Der Umgang mit DRM-geschützten Objekten ist deshalb bereits in einer nationalen Policy zu klären. Ausführungen zur langfristigen Überprüfbarkeit finden sich z. B. in den zahlreichen Arbeiten von Jeff Rothenberg<sup>255</sup>.

Das Recht zur Archivierung und Zur-Verfügung-Stellung muss daher entsprechend durch gesetzgeberische Maßnahmen geregelt werden; ggf. sind Sperrfristen abzuschaffen oder enger zu fassen.

8. **Auswahl der internationalen Partner:** Die generellen politischen Auswahlkriterien für internationale Partner für eine LZA-Policy gehören zur politischen Leitfunktion des Staates. Internationale Partner sind notwendig, um die Grenzen übergreifende Redundanz des digitalen Archivgutes sicherzustellen und gemeinsame Regeln, Schnittstellen und Standards zu vereinbaren.

Die nationale LZA-Policy setzt dann die Regeln zwischen den nationalen LZA-Institutionen sowie für die internationale Einbindung. Sammlungen von Web-Adressen der nationalen Bibliotheksgesellschaften aller Staaten in der Welt sind im Anhang angegeben.

9. **Vertraulichkeit:** Die Materialien für die Langzeit-Archivierung werden auch vertrauliche Dokumente enthalten müssen (z. B. interne Regierungsdokumente). Die Regeln und Abmachungen mit den archivierenden Institutionen im Ausland (Spiegel) über den Zugriff durch definierte Benutzerkreise und die Befristung der Vertraulichkeit müssen vom Staat gesetzt werden. Nach ihrer Freigabe wandern sie dann in das öffentliche Archiv.
10. **Datenschutz:** Die Integrität der LZA-Materialien muss gegen unautorisierte Eingriffe durch Nutzer von außen gesichert werden. Die zugehörigen Mindeststandards sollten national einheitlich definiert werden.

<sup>254</sup><http://www.zugang-zum-wissen.de>

<sup>255</sup>u.a.: Jeff Rothenberg; *Preserving Authentic Digital Information*; 2000; Council on Library and Information Resources CLIR; <http://www.clir.org/pubs/reports/pub92/rothenberg.html>

Regeln für die Integrität der LZA-Materialien gegen Manipulationen (absichtlich oder unabsichtlich) ist ebenfalls Aufgabe einer nationalen LZA-Policy. Beispielsweise werden in Australien den Dokumenten je nach festgelegter Integritätsgarantie-Stufe entsprechende Labels gegeben (*Safelink*). Die Realisierung mit geeigneten technischen Mitteln ist dagegen in der Eigenverantwortung der LZA-Institution.

## 3.9 Fragen- und Planungslisten

Zur Auflockerung des eher apodiktischen Stils der bisherigen Kapitel dieser Expertise wenden wir uns unserem Thema in diesem Kapitel sokratisch zu.

### 3.9.1 Strategische Fragen an eine LZA-Policy

Als Grundfragen für eine nationale LZA-Policy stellen sich:

- Für *wie lange* sollen welche Dokumentmengen archiviert und verfügbar gehalten werden?
- *Wer archiviert?*
- Wie wird die *Qualität* der Langzeit-Archivierung gesichert?
- Wie werden die rechtlichen, finanziellen, organisatorischen *Strukturen* stabil gehalten?

### 3.9.2 Planungs-Checkliste

Enthält die Checkliste für eine nationale LZA-Policy

1. Definitionen zur Archivierung, Langzeit, Objekt, Information;
2. Beauftragung eines nationalen Archiv-Netzwerkes;
3. Grundsätze der nationalen LZA-Policy;
4. Gesellschaftsform und Organisationsstruktur;
5. Sicherungs- und Migrations-Strategie;
6. National einheitliche Anforderungen an Zugriffs- und Nutzungsregelung und den Umgang mit DRM;
7. Einheitliche Sammelkriterien unter Beteiligung der Autoren/Erzeuger;
8. Datenschutzregelungen;
9. Sicherheit;
10. Finanzierung und Budgetierung;
11. Aufsicht und Monitoring des nationalen LZA-Netzwerkes;



## 12. Regelung der Zuständigkeit zur Weiterentwicklung der nationalen LZA-Policy?

Sind als weitere Aspekte geregelt:

1. Bundesweite Rahmenbedingungen für die Ausbildung des Personals;
2. Die Einbindung in die aktuelle Forschung;
3. Das nationale Risikomanagement und die Risikokontrolle;
4. Ein nationaler Katastrophenplan;
5. Einrichtung eines nationalen *Registry* mit *Fall-Back*-Strategie;
6. Ein Plan, wie das Langzeit-Archiv bei Totalverlust wiederhergestellt werden kann (Recovery-Plan)?

Viele dieser Punkte gelten nicht nur spezifisch für die Langzeit-Archivierung *digitaler* Objekte.

### 3.9.3 Vorbereitende Schritte zu einer LZA-Policy

Die folgende Zusammenstellung soll es zukünftig erleichtern, durch die zu beteiligenden Institutionen die notwendigen Arbeitspapiere zur Vorbereitung politischer Entscheidungen für eine nationale Langzeit-Archivierung digitaler Objekte zu erstellen.

Ziel dieser Liste ist es, hierfür zu regelnde oder vorzugebende Aspekte zusammenzustellen, um späteren Nachregelbedarf zu minimieren.

Werden mit der nationalen Policy folgende Standardfragen<sup>256</sup> beantwortet:

- Wie wird national die Sicherheit gegen Verlust von an verteilten Institutionen archivierten Materialien gewährleistet? Gibt es Verträge zwischen inländischen bzw. ausländischen LZA-Institutionen dazu?
- Haben die an einer nationalen konzertierten Langzeit-Archivierung digitaler Dokumente beteiligten Institutionen Zugriffsregelungen getroffen, die den maximalen Zugriff aus Wissenschaft und Kultur frei oder staatlich finanziert erlauben? Wird der jeweils gültigen Fassung des UrhG durch geeignete Regelungen Rechnung getragen?
- Haben sich die beteiligten Institutionen auf die Speicherung langfristig lesbarer oder rekonstruierbarer, offener Formate geeinigt?

<sup>256</sup>Entsprechendes findet sich zur Papier-Archivierung z. B. bei: Mirijam M. Foot; *Building Blocks for a Preservation Policy*; National Preservation Office; 2001; <http://www.bl.uk/services/npo/npo8.pdf>

- Verlangen die LZA-Institutionen von den Erzeugern/ Autoren von digitalen LZA-Werken die volle wissenschaftliche Information zur Speicherung (Quelltexte)?
- Werden die bundesweit gemeinsam vereinbarten internationalen Standards lokal auch eingesetzt?
- Sind geregelte intensive Kooperationen mit den jeweils für das Material fachlich einschlägigen Instituten vereinbart und installiert, welche die Auswahl von für lange Zeit zu archivierenden Dokumenten regeln, die Integrität der Speicherungen überprüfen, Quellen bewerten und Quellen jeweils aktuell verstehbar halten (Transfer in die aktuellen Sprachen)?
- Ist die institutionelle LZA-Policy der beteiligten Institutionen beim Bund hinterlegt und mit den Partnern für eine sinnvolle Arbeitsteilung abgestimmt?

Die Checkliste für den Bund selbst ist einfacher:

- Wurde eine stabile, verteilte aber integrative Organisationsstruktur geschaffen?
- Wurde ein Gesetz zur Ausfüllung der nationalen LZA-Policy erlassen?
- Sind Vereinbarungen mit den Ländern zu Arbeitsverteilung, Arbeitsteilung und Finanzierung getroffen?
- Wurde die Einführung geeigneter bundesweiter Studieninhalte zur digitalen Langzeit-Archivierung unterstützt?
- Wurden Regelungen zur Verpflichtung der Langzeit-Aufbewahrung, der Zugänglichkeit bzw. Vertraulichkeit sensibler Daten erlassen?
- Wie wird die Einhaltung des nationalen LZA-Gesetzes erreicht?

### 3.9.4 Themen und Beschränkung einer nationalen LZA-Policy

Dokumente sollten nicht durch kommerzielle Verlage langzeit-archiviert werden<sup>257</sup>.

<sup>257</sup>Hierzu zwei Beispiele:

Die von der DPG (*Deutsche Physikalische Gesellschaft*) <http://www.dpg-physik.de> herausgegebene Zeitschrift *Physikalische Blätter* enthielt neben ihren gedruckten Artikeln auch bereits seit 1995 einige Artikel, deren Quelle als E-Dokument nachgewiesen wurde. Vor einigen Jahren hat nun der Verlag VCH Wiley dieses Journal übernommen und gibt es als *Physik Journal* heraus. Die Archive wurden ebenfalls übernommen und werden auf den Verlagsseiten nachgewiesen. Dies gilt jedoch nicht für das E-Archiv der frühen Jahre, das also für die Nachwelt verloren ist.

Die Verlage Springer und Kluwer haben fusioniert. Die neue Konzernstrategie soll in etwa auf das Einstellen eines größeren Teils der wissenschaftlichen Zeitschriften des entstandenen gemeinsamen Pools, die Zusammenlegung von verwandten Zeitschriften beider Häuser, Übernahme des Markennamens Springer aber der rigideren Urheberrechtsregeln von Kluwer für alle wissenschaftlichen Zeitschriften hinauslaufen. Bei der Langzeit-Archivierung hofft man dann auf *Die Deutsche Bibliothek* und die KB ...

Die Langzeit-*Verfügbarkeit* von Artikeln in wissenschaftlichen Zeitschriften ist zunehmend de facto nicht mehr gegeben, auch wenn „im Prinzip“ durch Verträge mit nationalen Bibliotheken, oder Institutionen wie OCLC versucht wird, sie abzudecken.

Verfügbar meint ja in der Wissenschaft, dass aktive Forscher eine Arbeit, auf der ihre Forschung aufbauen könnte, finden, einsehen und lesen können. Dies ist in der eigenen (Universitäts-)Bibliothek immer häufiger nicht mehr der Fall. Zusätzlich droht ja, dass die bewährte Online-Bestellung digitaler Files bei einschlägigen Bibliotheken verboten werden<sup>258</sup> soll.

Die Interessen des individuellen einzelnen Autors durch das Urheberrecht sind dem übergeordneten Ziel der Bewahrung von Erkenntnissen und des kulturellen Erbes für die Wissenschaft unterzuordnen. Es gibt zahlreiche denkbare Zwecke der Langzeit-Archivierung, die sehr von der involvierten Institution abhängen. So sammelt die australische Nationalbibliothek primär alle digitalen Materialien, die mit dem Land Australien zusammenhängen, unabhängig von Format, Sprache, Nationalität des Autors oder Verlagsart (Auswahl nach dem Themenkreis). Die DDB sammelt nur E-Dissertationen aus deutschen Universitäten (Auswahl nach dem Dokumententyp und Erstellungsland). Viele Bibliotheken sammeln nur pdf-Files aus der eigenen Universität (Auswahl nach dem File-Typ und Erstellungsort).

Die logische und regulierende Reichweite einer nationalen LZA-Policy sollte sich daher beschränken auf :

- Die Vernetzungsregeln der beteiligten verteilten LZA-Institutionen;
- Die Standards der Dokumentenbeschreibung;
- Die Mindest-Anforderungen an Sicherheit, Qualitätssicherung, Redundanz, Lesbarkeit, Migrierung;
- Grundsätze zur Ausbildung;
- Gesetzliche Absicherung der Archivierungs- und Leserechte;
- Organisation konzeptioneller Vorgaben;
- Förderung der einschlägigen Forschung und ihre Vernetzung mit den Aktivitäten der LZA-Institutionen.

Der Zweck der Archivierung ist die *zukünftige* Nutzung, die sich aber in Abhängigkeit vom Archivgut stark unterscheiden kann. Sie kann von der Nutzung ausschließlich durch Archivare und Historiker, über Wissenschaftler der Disziplinen, aus denen das Archivgut ursprünglich stammt, bis hin zur breiten Öffentlichkeit reichen.

---

<sup>258</sup>Referenten-Entwurf des Bundesministeriums für Justiz; November 2004; <http://www.urheberrechtsbuendnis.de/links.html>

Die Realisierung der Vorgaben einer gegebenen nationalen LZA-Policy geschieht durch individuelle LZA-Institutionen über ihre *institutionelle Policy*. Sie müssen Fragen beantworten wie: Soll ein Archiv von einer einzelnen Institution oder als Verbund betrieben werden? Wer soll die Sammelkriterien bestimmen? Welche existierenden Dienste sollen eingebunden werden? Welche sind neu zu schaffen? Welches Archivgut soll mit Institutionen im Ausland ausgetauscht und gespiegelt werden?

## 3.10 Erläuterungen zu Einzelthemen

### 3.10.1 Sicherheit: Datenschutz, Geheimhaltung

Jedes LZA-System sollte das Archivgut so aktuell wie möglich aufnehmen. Gerade digitale Medien erfordern es, dass diese in das Archiv integriert werden, solange noch das Wissen über die verwendeten Dateiformate, die zur Anzeige oder Ausführung notwendige Software und die für die Software notwendige Hardware vorhanden sind, um eine spätere Migration sicher ausführen zu können. Entsprechend verschärft gegenüber dem traditionellen Archivwesen stellt sich die Problematik der Sicherung des Archives vor unerlaubter Bearbeitung. Außerdem ist das Archivgut ständig weiter zu pflegen. Pintfield et. al.<sup>259</sup> schreiben hierzu: „Digital information is lost when it is left unattended while hardware, software and media continue to develop“.

Dennoch sind Versuche einer vorläufigen „Bitstream Conservation“ nicht überflüssig, und der Nutzen von Format-Registries nicht relevant eingeschränkt. Denn es geht eben auch darum, heute Daten zu speichern und möglicherweise erst morgen den Zugriff zu diesen Daten technisch zu ermöglichen (wir können nicht alle Probleme sofort lösen). Aber ein möglicher Verlust an Information durch Verlust der Kenntnisse über das Format, durch Projektion von Wissen bei der Archivierung z. B. nur noch auf Abbilder wird dabei dann in Kauf genommen werden müssen.

Digitale verteilte Archive werden immer auch personenbezogene, vertrauliche oder geheime Informationen enthalten, die erst nach Ablauf einer Schutzfrist externen Nutzern zugänglich gemacht werden sollen<sup>260</sup>. Dennoch sollen diese Inhalte auch während dieser Schutzfrist im Archiv gepflegt werden können, um sicherzustellen, dass nach Ablauf der Schutzfrist der inhaltliche Datenstrom unverändert ist, dass die Daten in einem Format verfügbar sind, das mit dann vorhandener Hardware lesbar ist.

Gleichzeitig ist es sinnvoll, diese Inhalte so vorzuhalten, dass sie auch bewusst nicht verfälschbar sind und auch im Rahmen einer internationalen Kooperation ausgetauscht werden können, ohne vertrauliche Inhalte preiszugeben. Entsprechend leistungsstarke Verschlüsselungs- und Signatur-Verfahren sind zu entwickeln und einzusetzen, um die Integrität der Quellen überprüfen zu können.

---

<sup>259</sup>Stephen Pinfield, Hamish James; *The Digital Preservation of e-Prints*; D-Lib Magazine; September 2003; Volume 9 Number 9; ISSN 1082-9873; <http://www.dlib.org/dlib/september03/pinfield/09pinfield.html>

<sup>260</sup>Als Beispiel seien Protokolle aus der Regierungsarbeit genannt.

Dabei kann es durchaus aus nationalen Interessen heraus notwendig sein, vertrauliche Daten verschlüsselt ins Ausland auszulagern, um diese im Bedarfsfall von dort - ohne gelesen/entschlüsselt worden zu sein - reimportieren zu können. Diese Vorkehrung sollte unabhängig von der Regierungs- oder Staatsform erfolgen, da die Geschichte zeigt, dass Staatsformen, in der Zeitskala einer Langzeit-Archivierung gesehen, plötzlich wechseln können.

### 3.10.2 Redundanz

Die Speicherung des wissenschaftlichen und kulturellen Erbes ist neben der Bereitstellung von Archivgut für öffentliche und wissenschaftliche Nutzer die Kernaufgabe eines LZA-Systems.

In der Vergangenheit bestand das Archivgut aus Informationen, die fest mit ihrem Träger verknüpft waren. So lässt sich die Tinte nicht ohne weiteres von dem Papier trennen und die in der Schrift oder Abbildung enthaltene Information nur schwer auf ein anderes Medium transferieren. Mikrofilme sind ein Beispiel für diesen Transfer. Eine Folge dieser engen Bindung der Information an ihr Trägermedium ist, dass die enthaltene Information bisher in Archiven nur selten redundant vorgehalten wird, außer bei Büchern und Zeitschriften, aber auch dort durch die Auflage im Voraus begrenzt und immer mit dem Risiko, die gesamte Information durch die Zerstörung des Informationsträgers zu verlieren.

Für digitale Informationen gelten diese Einschränkungen zunächst einmal nicht, denn sie lassen sich relativ einfach beliebig oft klonen<sup>261</sup>. Ohnehin ist hier der Transfer der Information auf immer wieder aktuelle Datenträger der technisch neuesten Generation notwendig, um deren Integrität und Lesbarkeit zu gewährleisten.

Damit eröffnet sich hier jetzt erstmals die Möglichkeit des beliebig redundanten Vorhaltens des Archivgutes.

Die Kooperation verschiedener Archiv-Institutionen ist für die Langzeit-Archivierung damit essentiell notwendig, um die redundante Haltung des Archivgutes zu gewährleisten. Dies ist eine wesentliche neue und zusätzliche Aufgabe. Archivierungsregeln, die bisher nur in einzelnen Archiven festgelegt wurden, müssen daher schon deshalb zukünftig national und international harmonisiert werden, um Spiegelung, Arbeitsteilung und gemeinsames Retrieval weltweit zu ermöglichen.

### 3.10.3 Internationale Vernetzung

Die Vorhaltung des Archivgutes sollte dabei keinesfalls nur innerhalb der eigenen Staatsgrenzen erfolgen, sondern international vernetzt sein. Nur so ist langfristig auch die

---

<sup>261</sup> Papier- oder Xerokopien eines gedruckten Originals sind dagegen natürlich nicht mit dem Original identisch.

bewusste Löschung oder Verfälschung von digitalem Archivgut auszuschließen. Wie groß der wissenschaftliche Wert von Archivgut ist, welches politisch turbulente Zeitphasen im Ausland unverfälscht übersteht, haben gerade deutsche Historiker in den vergangenen Jahrzehnten immer wieder erfahren.

Hierbei ist an Ereignisse wie Krieg, Naturkatastrophen oder wirtschaftliche Umbrüche etc. zu denken; jede digitale Kopie erhöht die Chance, dass das digitale Objekt die Zeiten übersteht (LOCKSS<sup>262</sup>).

Die Organisation geschieht über ein Netz von *Registries*, die automatisiert die an den verschiedenen Institutionen vorhandenen LZA-Dokumente registrieren und prüfen, ob diese auch an mindestens einem anderen unabhängigen Ort verfügbar und lesbar sind, und ob diese Dokumente identisch sind.

### 3.10.4 Finanzierung

Der finanzielle Aufwand bestimmt sich aus den Aufgaben. Er wird wesentlich höher sein als bisher. Viel größere Datenmengen sind zu bewältigen und mehr menschliche hoch qualifizierte Arbeit ist notwendig. Dies wird nicht ausbalanciert durch die Einsparungen bei einer digitalen Archivierung anstelle einer Papier-Archivierung.

Wer finanziert welche Aufgaben in der Langzeit-Archivierung? Hier ist insbesondere zu klären, wie die Aufteilung des notwendigen Budgets zwischen Bund und Ländern erfolgen soll. Es ist zu klären, wer das Budget legt. Diese Frage ist für die Durchsetzung einer nationalen Policy zentral, sie wird aber nicht in der Policy selbst geklärt werden können, sondern ist sozusagen eine „voraussetzende Folge“.

Die Finanzmittel für eine Realisierung der nationalen LZA-Strukturen müssen vom Bund und den Ländern aufgebracht werden. Dagegen bestimmt sich die Archivierungsarbeit selbst (auswählen, sammeln, speichern, spiegeln, migrieren, aufbereiten, anbieten) aus dem Aufgabenbereich und der institutionellen Policy der verteilten LZA-Institutionen.

Die verschränkte Verantwortung und daraus notwendige Finanzierung von Bund und Ländern verlangt eine entsprechend gemeinsame *Kontrolle*. Sie verlangt weiterhin ein inhaltsbezogenes gemeinsam getragenes *Audit*. Die *Erfolgskriterien* können dabei zum Teil andere als im Papierzeitalter (etwa Zahl der archivierten, also vorgehaltenen Arbeiten) sein, nämlich u. A. die Zahl der Downloads, die Zahl der genutzten Dokumente und der zitierten Arbeiten - also genereller gesehen des „Impacts“.

Eine langfristig sichere Finanzierung bei wechselnden nationalen politischen Konstellationen kann nur durch ein Bundesgesetz für die nationale LZA-Policy sichergestellt werden.

Die Abfolge ist also: Erst die nationale LZA-Policy erstellen, dann den organisatorischen Rahmen festsetzen, dann die finanziellen Anforderungen definieren und diese

<sup>262</sup><http://lockss.stanford.edu>

dann in Budgets und Vereinbarungen von Bund und Ländern gießen.

Die Finanzierung der Langzeit-Archivierung spiegelt den politischen Willen wider. Nach allem ist deutlich geworden, dass ein wesentlicher Umbruch von der jetzigen Kurzzeit-Archivierung von Dokumenten, die auf das Fortbestehen der aktuellen proprietären Lese-Programme setzt, hin zu einer zukunftssicheren Langzeit-Archivierung notwendig ist. Sie wird neben den großen organisatorischen und strukturellen Gestaltungen auch sehr hohe finanzielle Anforderungen stellen.

Genau daran lässt sich dann die Ernsthaftigkeit des politischen Willens ablesen.

### 3.10.5 Ausbildungsfragen

Die neuartigen und höher qualifizierten Arbeiten in der gesamten LZA-Kette verlangen entsprechend ausgebildetes Personal.

Zur Sicherung eines bundesweit gleichmäßigen Qualitätsstandards sind daher einheitliche geeignete Ausbildungsinhalte zu definieren, die zu den als notwendig gesehenen Berufsqualifikationen hinführen. Ihre *Realisierung* kann durch eine Anpassung vorhandener oder aber auch durch eigene neue Studiengänge erfolgen.

Erste unkorrelierte Versuche, etwa die beiden Lehrstühle für Informationswissenschaft (Berlin, Konstanz), oder der Master-Studiengang *Information Engineering* (Osnabrück und Twente), reichen dazu schon allein wegen der großen benötigten Zahl von Studienabgängern nicht aus. Ab 2006 bietet aber die Staatliche Akademie der Bildenden Künste in Stuttgart einen neuen Master-Studiengang zum *Konservatoren für Neue Medien und Digitale Information*<sup>263</sup> an.

Die zu Beteiligten müssen für diesen neuen Teil der Ausbildungslinien daher der Bundesregierung ein Gesamtkonzept mit Studieninhalten, Dauer und Prüfungsanforderungen vorschlagen, die sich aus einer vertieften Berufsanforderungs-Analyse ergeben, unabhängig von der Realisierung durch Anpassung von vorhandenen oder durch neue Studiengänge.

### 3.10.6 Mögliche Marginalität von Publikationen

Publikationen und Dokumente, die nur wenig oder keinen relevanten Inhalt haben, sollten keinen Eingang in die Langzeit-Archivierung finden. Hierzu gehören viele Dokumente der täglichen Kommunikation und des Managements (Raumpläne an Universitäten etc.). Hierzu gehören aber auch sehr viele, wenn nicht die meisten Publikationen in referierten wie in nicht referierten wissenschaftlichen Zeitschriften und Sammlungen.

---

<sup>263</sup><http://www.abk-stuttgart.de>

Durch die zeitweise noch weit verbreitete Unsitte, Wissenschaftler für Forschungs- und Hochschulstellen nach der Zahl ihrer Publikationen, nicht nach deren Inhalt auszuwählen, hat sich der Anteil marginaler Publikationen stark erhöht.

Die Marginalität von Arbeiten lässt sich quantitativ stützen: Im *Citebase*-Server wird als Funktion der Zeit die Zahl der Zitierungen und die Zahl der Downloads einer Arbeit dargestellt<sup>264</sup>. Die in Citebase analysierten Arbeiten sind im wesentlichen solche des *ePrint ArXiv* der Cornell University, von denen bekannt ist, dass ca. 90% von ihnen nach der Erstverbreitung über das *ArXiv* in streng referierten Zeitschriften erscheinen. Die Fachwelt hat also - trotz freien Zugriffs auf die Arbeit - meist schon im ersten Jahr nach der Entstehung sehr oft sehr wenig Interesse an der Arbeit. Hierzu kommt die große Zahl nie zitierter Arbeiten.

Dass in der riesigen Menge Publikationen scheinbar marginalen Inhalts trotzdem Schätze verborgen liegen könnten, ist nicht auszuschließen, zumindest aber in der Naturwissenschaft mit ihrem engen Netz an Nachprüfungen eher unwahrscheinlich.

Zusätzlich sollten technisch nicht mehr lesbare Dokumente nicht langzeit-archiviert werden<sup>265</sup>.

### 3.10.7 Zugriffskonditionen

Es werden auch rechtliche Aspekte tangiert, die, wenn man die entsprechenden Abschnitte des Archivrechtes als Policy versteht, durch eine nationale Policy geregelt werden müssen.

Die Art und die Konditionen des Zugriffes durch die Nutzer sind in einer jeweils lokalen institutionellen Policy zu regeln, allerdings nach den Vorgaben der nationalen Policy. Wer darf auf das Archivgut zugreifen? Gibt es eine Sperrzeit für aktuelle Inhalte, vor deren Ablauf kein Zugriff erlaubt ist? Wie lang ist diese Sperrzeit? Von wo darf auf das Archivgut zugegriffen werden? Über das Internet oder nur innerhalb bestimmter Institutionen? Sollen Hilfen zur Hardware- und Software-unabhängigen Darstellung

<sup>264</sup><http://www.citebase.ac.uk>; Wegen eines Brandes im Institut war dieser Dienst in 2005 zeitweilig außer Betrieb; <http://citebase.eprints.org/offline.php>

<sup>265</sup>Das bekannteste Beispiel sind die Publikationen in den hochangesehenen, streng referierten Zeitschriften *Physical Review* (PR) und *Physical Review Letters* (PRL). Die Manuskripte werden zum weitaus überwiegenden Teil seit ca. 1983 (Erfindung der Text-Markup-Sprache TeX bzw. L<sup>A</sup>T<sub>E</sub>X) in dieser Sprache abgefasst und als reiner ASCII Code elektronisch an den Verlag übermittelt.

Der Verlag (APS American Physical Society) hat damals bis ca. 1994 aus den in langfristig lesbarer Form eingegangenen digitalen Files jeweils nach Eingang ein Druck-Image File erstellt (in *runoff*) und diese dann zum Drucken der Zeitschrift verwendet. Routinemäßig wurden dann die *runoff*-Files archiviert und die L<sup>A</sup>T<sub>E</sub>X-Files gelöscht.

Nachdem in 1994 beschlossen wurde, ein e-Archiv der PR- bzw. PRL-Artikel aufzubauen, musste man feststellen, dass die runoff-Files mit der neuen *runoff*-Version nicht mehr lesbar und druckbar waren, und dass ein Konverter nicht vorhanden war und nicht mehr erstellt werden konnte. Es blieb nichts anderes übrig, als die gedruckten Zeitschriften wieder einzuscannen, ein zehnjähriges Unterfangen, was jetzt aber einen wesentlichen und exzellent aufbereiteten Wissens-Schatz darstellt: PROLA *Physical Review Online Archive*; (Inhalt: 1875-heute; PR und PRL); <http://prola.aps.org/>



durch das Archiv angeboten werden?

Ob mit DRM-Verfahren (technisch realisiertes Digital Rights Management) geschützte Informationen in das Langzeit-Archiv aufgenommen werden sollen, ist eine Frage der institutionellen Policy, ebenso, ob das Archiv abgelaufene sowie nicht wiederbeschaffbare Schlüssel „knacken“ (technisch selbst umgehen) will.

In der nationalen Policy ist aber klarzustellen, dass der Gesetzgeber die Durchsetzungsfähigkeit der Schranken des UrhG für Wissenschaft und Bildung sichern muss: Dies ist insbesondere in der nationalen LZA-Policy zu regeln, da hierzu die entsprechenden rechtssichernden Rahmenbedingungen zu schaffen sind, damit nachgeordnete institutionelle Policies greifen können.

Die Zugriffsrechte unterliegen nationalem und internationalem Recht (z. B. der EU). Heikel sind hierbei heutige Verfügungen über Zugriffskonditionen zukünftiger Generationen, die sich möglicherweise ein Urheberrecht wählen werden, das sich vom heutigen wesentlich unterscheidet. Die zukünftigen Generationen werden sich nicht unbedingt den Wünschen heutiger Gesetzgeber, kommerzieller Firmen oder Autoren unterwerfen.

### 3.10.8 Sicherung der Authentizität

Welche Maßnahmen zur Sicherstellung der Authentizität des Archivgutes werden getroffen? *Hashes* und digitale Schlüssel? Spiegelung des Archivgutes in internationaler Kooperation, um gewollte Verfälschungen zu verhindern? Wie soll das Langzeit-Archiv vor einem Datenverlust bei Wegfall der Finanzierung von Teilen des Archives gesichert werden (Fall-Back-Strategie)?

Grundsätzlich ist die Sicherstellung der Authentizität ein wichtiger Punkt einer Policy. Eine nationale Policy wird aber nicht bestimmte technische Lösungen vorschreiben, zumal sich die technischen Randbedingungen auch recht kurzfristig ändern können. Wichtig ist aber die Festlegung auf offene Formate (Open Source), Quelltext-Verpflichtung und in ASCII kodierte Dateien bei Textdokumenten, um die langfristige Rückholbarkeit zu sichern.

Zur Langzeit-Sicherung der Authentizität gehört die möglichst unfälschbare *Signatur* der Autoren. Hier bemühen sich auch kommerzielle Dienstleister, technische Lösungen anzubieten. So hat am 25.11.2005 die Firma ADOBE einen Viewer vorgestellt, der dem deutschen Signaturgesetz entspricht und nach den „Common Criteria for Technology Security Evaluation“ zertifiziert ist<sup>266</sup>; Siehe auch das Faltblatt *Elektronische Signatur* des BSI (*Bundesamt für Sicherheit in der Informationstechnik*<sup>267</sup>, und international die entsprechenden Anforderungen des Europäischen Rates und des Parlaments sowie der Vereinten Nationen<sup>268</sup>.

<sup>266</sup><http://www.adobe.de/signatur>

<sup>267</sup><http://www.com2-net.de/public.files/tf/bsi-signatur.pdf>

<sup>268</sup>Zu Literatur siehe das Faltblatt des BSI.

### 3.10.9 Datenschutz und Sicherheit

In Deutschland ist die Speicherung personenbezogener Daten – und erst recht die Veröffentlichung von Zusammenstellungen solcher Daten ohne Einwilligung der Personen in wesentlichen Bereichen nicht zulässig. Jede LZA-Policy sollte daher eine Datenschutz-Regelung enthalten, denn es werden auch personenbezogene Daten gespeichert.

Es ist eine Regelung zu treffen, die die Interessen der inneren und äußeren Sicherheit angemessen berücksichtigt.

Es werden auch Daten und Dokumente langzeit-archiviert, die gesetzlichen Beschränkungen in der Verbreitung unterliegen und deren Kenntnis auch in der Zukunft nur einem eingeschränkten Personenkreis zugänglich gemacht werden kann. In diesem Bereich werden eine Mehrzahl von Ministerien und Ämtern Anforderungen an die Einschränkung der Zugänglichkeit anmelden.

Die ständig adjustierte Balance zwischen Datenschutz-Belangen und öffentlicher Zugänglichkeit ist bekannt und bildet sich auf die Langzeit-Archivierung ab, die dem Rechnung tragen muss.

Zugleich verlangen viele Verlage Auflagen wie „öffentlich zugänglich erst nach einer gewissen Anzahl von Jahren“, oder es stehen der öffentlichen Zugänglichkeit Rechte des Autors oder seiner Erben entsprechend dem Urheberrechtsgesetz entgegen.

Andererseits muss der Gesetzgeber an der Ausgestaltung einer nationalen Policy auch in der Weise mitwirken, dass die berechtigten und formulierten Anforderungen des Staates an die Zugänglichkeit von Wissen für die (zumindest wissenschaftliche) Öffentlichkeit, um die Forschung und die technologische Entwicklung zu fördern, nicht leiden. Dieses Allgemeininteresse setzt den Forderungen der Autoren und der Verlage Grenzen, sogenannte Schranken des Verwertungsrechts. Dies ist zur Zeit in der aktuellen Diskussion. Derzeit wird der sog. *Korb 2 der Urheberrechtsreform* für den Bundestag vorbereitet. Der zugehörige Referentenentwurf des Bundesministeriums der Justiz wird von den sechs großen Wissenschaftsorganisationen, derzeit 259 Institutionen und wissenschaftlichen Fachgesellschaften, sowie bereits 3.564 Einzelwissenschaftlern abgelehnt, denen, die die „Göttinger Erklärung“ des *Aktionsbündnis Urheberrecht für Bildung und Forschung*<sup>269</sup> unterzeichnet haben. Auf die Stellungnahmen des Bündnisses sowie die dort befindliche Sammlung von Stellungnahmen Dritter, wie z. B. KMK, DBV etc.<sup>270</sup> wird verwiesen.

Im Korb 2 des UrhG, und sicher erst recht in den bereits geplanten weiteren Körben 3 bis 8, für die jedesmal etwa ein Jahr für die politische Durchsetzung bis zum Erlassen des Gesetzes gerechnet wird, sollen wesentliche Änderungen der Zugänglichkeit von Wissen festgelegt werden, die sich direkt auf die LZA-Policy auswirken könnten. Die Gesetzgebung ist im Fluss und die parallele Entwicklung einer LZA-Policy und deren Realisierung muss die jeweils vorliegenden Anforderungen umsetzen können, was eine gewisse Flexibilität des Rahmens der Langzeit-Archivierung erfordert, bzw. besser

<sup>269</sup><http://www.urheberrechtsbuendnis.de>

<sup>270</sup><http://www.urheberrechtsbuendnis.de/links.html>

vorher aktiv mit auf den Gesetzgebungsprozess einwirken.

### 3.10.10 Organisationsform

Welche Gesellschafts- und Organisationsform soll ein Langzeit-Archiv oder ein die Langzeit-Archivierung betreibendes Netzwerk haben? D. h. soll es ein loses Netz archivierender Institutionen mit bilateralen Verträgen geben oder eine nationale LZA-Struktur mit eigener Organisationsform? Dies wirkt sich direkt auf die Verbindlichkeit, die Durchsetzungsfähigkeit von gemeinsamen Regeln und indirekt auf die Finanzierung aus. Wir plädieren in dieser Expertise für eine straffe Organisation von selbständigen nationalen und regionalen sowie von fachbezogenen LZA-Partnern mit breit verteilter dezentraler Langzeit-Archivierung.

### 3.10.11 Metadaten

Auf diesem Feld der spezifisch für die Langzeit-Archivierung zu bedenkenden Metadaten kann auf die Zusammenstellungen von PADI (*Preservation Access to Digital Information*) der *National Library of Australia* bzw. deren Literatursammlung zu *Preservation Metadata*<sup>271</sup> zurückgegriffen werden, ebenso zu deren sehr nützlicher Literatursammlung zu *Persistent Identifiers*<sup>272</sup>. Von PADI gibt es eine sehr gute Literatursammlung zu Standards<sup>273</sup>. Die SUB Göttingen arbeitet bereits über das Projekt *nestor* mit PADI zusammen.

Die Langzeit-Archivierung des kulturellen Erbes und der wissenschaftlichen Erkenntnisse und Ergebnisse aus Instituten und Institutionen in der Bundesrepublik ist eine gesamtstaatliche Aufgabe, die in enger Kooperation im internationalen Kontext erfolgen muss und daher international vereinbarter Metadaten bedarf.

Aus gesamtstaatlicher Sicht muss eine konsistente Policy entwickelt, festgelegt und verfolgt werden, welche die teilweisen Zuständigkeiten der Länder (Kulturhoheit) berücksichtigt und möglichst als Stärke, nicht als Schwäche eines Gesamtsystems sieht.

Dies setzt eine mit den Ländern vereinbarte Koordinierung der Zuständigkeiten voraus. Zugleich muss der Bund sicherstellen, dass national einheitliche Standards eingehalten werden, dass länderübergreifend das Gesamtsystem sicher, stabil und koordiniert langfristig funktioniert (Bundeszuständigkeit), unabhängig von möglichen politischen Wendungen der Kulturpolitik der Länder.

Der Bund muss die Einbindung in ein internationales, verteiltes LZA-System ebenso koordinieren und sicherstellen wie die Teilnahme an den fortschreitenden technischen und konzeptionellen Entwicklungen.

---

<sup>271</sup><http://www.nla.gov.au/padi/topics/30.html>

<sup>272</sup><http://www.nla.gov.au/padi/topics/36.html>

<sup>273</sup><http://www.nla.gov.au/padi/topics/43.html>

Daneben gibt es nationale Anforderungen an die Länder, die in deren Zuständigkeit fallen und daher mit ihnen auszuhandeln sind. Nach nationaler Festlegung der Anforderungen an das Gesamtsystem in Bezug auf Sicherheit und Redundanz kann die dann die internationale Einbindung vorangetrieben werden.

Die entsprechende Förderung der Forschung, Entwicklung und Anwendungsvorbereitung fällt ebenfalls größtenteils in Bundeskompetenz, um die Einhaltung gemeinsamer Standards, technischer und semantischer Schnittstellen langfristig zu garantieren.

### 3.10.12 Migration von Formaten

Das Migrieren von Dokumenten zu anderen Formaten ist inzwischen eine eigene Disziplin geworden. Grundsätzlich ist die *Projektion* von hochwertigen, offenen, international vereinbarten LZA-Formaten (wie z. B. SGML, XML, *content-MathML*, CML) auf andere offene Formate (wie  $\text{\LaTeX}$ ) möglich und es ist die weitere Projektion zu proprietären Darstellungsformaten (etwa *postscript* oder *pdf*) sichergestellt, solange diese Formate von den entsprechenden Firmen unterstützt werden. Projektion heißt aber immer auch Informationsverlust.

Daher ist der einzige Weg, um die vollständige Information langfristig erhalten zu können, dass die Autoren so bei der Erstellung ihrer Dokumente betreut werden, dass sie offene LZA-Formate nutzen.

Dabei ist die Projektion auf Darstellungsformate i. A. trotzdem eine trickreiche Angelegenheit. Hier sind Anleitungen der LZA-Bibliotheken erforderlich, als Beispiel seien die der *West Virginia University*<sup>274</sup> genannt.

Der umgekehrte Weg der *Entfaltung*, des Versuchs einer Migration zu höherwertigen LZA-Formaten ist immer mit Mehrdeutigkeiten der Interpretation verbunden, was denn nun gemeint sein könnte, sei es nun bei OCR von *Browsing Images* oder bei der Entfaltung von  $\text{\LaTeX}$ -Dateien zu *Content-MathML*.

### 3.10.13 Gesetzliche Rahmenbedingungen

Die gesetzlichen Grundlagen für die papier- und datenträger-gebundene Archivierung bilden dabei das *Gesetz über die Sicherung und Nutzung von Archivgut des Bundes* (Bundesarchivgesetz BArchG), das *Gesetz über die Deutsche Bibliothek (DBibLG)* und das *Gesetz über den Deutschen Wetterdienst (DWDG)* auf Bundesebene sowie die Landes-Archivgesetze der Bundesländer.

Das Bundesarchiv als obere Bundesbehörde steht unter unmittelbarer Kontrolle des Beauftragten der Bundesregierung für Kultur und Medien<sup>275</sup>. Es hat den Auftrag, Ar-

<sup>274</sup><http://www.libraries.wvu.edu/theses/electronic.htm>

<sup>275</sup> BArchG Gesetz über die Sicherung und Nutzung von Archivgut des Bundes; <http://bundesrecht.juris.de/barchg/BJNR000620988.html>

chivgut des Bundes zu archivieren, nutzbar zu machen und wissenschaftlich zu bewerten. Dieser über die reine Archivierung erweiterte Auftrag sichert die erst für eine Nutzung in ferner Zukunft sinnvolle Langzeit-Archivierung. Die Schutzfrist auf Zugriff ist i. A. auf 30 Jahre gesetzt. Im Gesetz sind die Zugriffsbeschränkungen genau geregelt, ebenso die Auswahlprinzipien (z. B. muss ein 'relevantes Interesse des Bundes' vorliegen) für die Archivierung.

Die Deutsche Bibliothek ist eine rechtsfähige bundesunmittelbare Anstalt des öffentlichen Rechts und als „zentrale Archivbibliothek“ (DBibIG §1) der Bundesrepublik angelegt<sup>276</sup>. Der Bereich des Archivgutes ist begrenzt auf seit 1913 in Deutschland verlegte oder hergestellte Druckwerke, deutschsprachige Werke im Ausland, fremdsprachige Übersetzungen bzw. Werke über Deutschland sowie von Emigranten (zwischen 1933-1945) verfasste Werke. Sie sind zu sammeln, zu inventarisieren und bibliographisch nachzuweisen. Die (unbefristete) Archivierungspflicht ergibt sich aus DBibIG Abschnitt 2, insbesondere §21<sup>277</sup>, wobei §25<sup>278</sup> die Abgrenzung zur Kulturhoheit der Bundesländer und deren teilweise vorhandene Ablieferungspflicht von Pflichtexemplaren regelt. Die DDB hat außerdem die gesetzliche Aufgabe, die Beziehungen zu entsprechenden ausländischen bzw. internationalen Institutionen zu pflegen<sup>279</sup> und hat von daher bei der Herausforderung, eine nationale LZA-Policy für digitale Werke international einzubetten, eine besondere Rolle.

Der Deutsche Wetterdienst wiederum ist nur eine teilrechtsfähige Anstalt (Teilkörperschaft) des öffentlichen Rechts im Geschäftsbereich des Bundesministeriums für Verkehr (derzeit angesiedelt beim Bundesminister für Verkehr, Bau- und Wohnungswesen). Die Archivierungspflicht ergibt sich aus DWDG, §4, Abs. 1.9<sup>280</sup>. Das Gesetz regelt in §3 die enge Kooperation mit dem Bundesministerium der Verteidigung und weiteren nicht genannten Ressorts.

Weitere relevante Rechtsvorschriften gibt es beispielsweise auch im Bereich des Datenschutzes, des Urheber- und Verwertungsrechtes etc. Was davon in eine nationale Policy einfließen sollte (oder umgekehrt, wo die nationale Policy die Gesetzgebung beeinflussen würde), bedarf einer genauen und gründlichen Aufarbeitung und insbesondere eines pragmatischen politischen Gestaltungsvorschlages. Dies könnte in Kooperation von *nestor* und dem *Arbeitskreis Urheberrecht* von DINI<sup>281</sup> und mit Experten des *Aktionsbündnis Urheberrecht für Bildung und Wissenschaft*<sup>282</sup> geschehen.

Im Übrigen wird auf die Rechts-Expertise<sup>283</sup> im *nestor*-Vorhaben verwiesen, die jedoch den hier mitentscheidenden §52a des Urheberrechtsgesetzes wegen der damals noch in der Novellierung geplanten Verfallsklausel §137k ausgeklammert hatte.

<sup>276</sup><http://bundesrecht.juris.de/dbiblg/BJNR002650969.html#BJNR002650969BJNG000100311>

<sup>277</sup> „Die Deutsche Bibliothek hat die Pflichtstücke aufzubewahren.“

<sup>278</sup> „Die landesrechtlichen Regelungen über die Ablieferung von Pflichtstücken bleiben unberührt.“

<sup>279</sup> DBibIG §2.4.

<sup>280</sup><http://bundesrecht.juris.de/dwdg/BJNR287100998BJNE001101308.html>

<sup>281</sup><http://www.dini.de>

<sup>282</sup><http://www.urheberrechtsbuendnis.de>

<sup>283</sup> Rechtsanwälte Göbel und Scheller; *Digitale Langzeitarchivierung und Recht*; *nestor*-Materialien 1; [http://www.Langzeitarchivierung.de/downloads/mat/nestor\\_mat\\_01.pdf](http://www.Langzeitarchivierung.de/downloads/mat/nestor_mat_01.pdf)

Inzwischen erscheint es jedoch vor allem durch die Aktivitäten des *Aktionsbündnis Urheberrecht für Bildung und Wissenschaft* sehr wahrscheinlich, dass die Verfallsklausel aufgehoben oder zumindest wesentlich verlängert wird und weitere Anforderungen der Wissenschaft berücksichtigt werden könnten.

Die aktive Einflussnahme der jeweils Betroffenen auf einen Gesetzgebungsprozess wird in Deutschland vom Bundesministerium der Justiz durch offene Anhörungen aller Beteiligten gefördert, wie wir bei der Novellierung des UrhG erfahren konnten.

Der Übergang von den Strukturen und Methoden der Langzeit-Archivierung im Papierzeitalter zu den angestrebten neuen langzeit-stabilen Strukturen und Methoden für digitale Informationen vollzieht sich derzeit außerordentlich rasch. Dabei wird zur Zeit in allen Industrieländern auf eine nationale Policy hingearbeitet und es werden tastende Versuche in Form von Projekten unternommen, die den Weg vorzeichnen sollen, um zu den noch unbekannt optimalen Strukturen und Methoden zu kommen.

#### **3.10.14 Experten**

Eine Teilaufgabe war die Zusammenstellung einer Liste von Experten an ausländischen Institutionen, die als Ansprechpartner für die Verfolgung einer nationalen LZA-Policy dienen könnten. Rechtlich ist eine solche publizierte Sammlung in Deutschland ohne die Einwilligung der genannten Personen nicht zulässig, sehr wohl aber der hier beschrittene Weg der im Text verteilten bzw. im Anhang genannten Quellen und Namensnennungen.

# 4

## Schlusswort

Eine Zusammenfassung der organisatorischen und strukturellen Identifizierung von Handlungsfeldern sowie von organisatorischen und strukturellen Vorschlägen für einen Weg zu einer nationalen Policy für die Langzeit-Archivierung wurde in dieser Expertise vorgezogen an den Anfang gestellt. In Teil II folgte eine Analyse des (wenig ergebnisreichen) Rücklaufs einer Umfrage unter nationalen Bibliotheken sowie von relevant erscheinenden Materialien und Informationen über Projekte in anderen Ländern.

Die technologische Entwicklung schreitet unvermindert rasch voran. Die nationalen Aktivitäten der professionellen Institutionen zur Archivierung digitaler Dokumente haben sich rasch verstärkt und verbreitert. Das BMBF-Verbundprojekt *nestor* hat sich des Themas der nationalen Langzeit-Archivierung angenommen. Den für die Erarbeitung einer Policy für die nationale Langzeit-Archivierung tätigen Experten – und denen aller Fächer, Gebiete und Bereiche, die noch hinzugezogen werden, war der zweite Teil mit seinen Materialien gewidmet.

Diese Expertise soll auch dazu dienen, das Interesse und die Aufmerksamkeit der Nutznießer einer Langzeit-Archivierung zu wecken – der Autoren und Leser, der Wissenschaftler und Künstler, der kleineren Archive, Bibliotheken und Museen – und die notwendige breite Diskussion auf dem Weg zu einer nationalen Langzeit-Archivierungspolicy zu befeuern und hierzu auf relevante Entwicklungen im Ausland hinzuweisen.

Wir hoffen, dass die Beschäftigung mit der Langzeit-Archivierung als Blick auf die Zukunft verstanden wird.





# Appendix A

## Nationalbibliotheken weltweit

Ein erster Schritt zur Einbettung der nationalen LZA-Policy in die internationalen Aktivitäten zur Langzeit-Archivierung ist die Kontaktaufnahme zu den entsprechenden Partnern anderer Länder. Für einen Industriestaat wie Deutschland entsteht eine zusätzliche Verantwortung, für kleine bzw. für Entwicklungsländer technische Hilfsdienste zu leisten, Expertise zu vermitteln und Aufgaben der internationalen Einbettung und der Spiegelung/Archivierung von Werken zu übernehmen.

Hier geben wir eine Liste der Web-Adressen von National-Bibliotheken anderer Staaten als einen möglichen Einstieg zur Kontaktaufnahme.

Erstaunlicherweise sind entsprechende offizielle Listen von IFLA etc. nicht sehr vollständig und auch nicht aktuell. Relativ gute Informationen finden sich in Wikipedia<sup>1</sup>.

Die alphabetische Einordnung ist von der gewählten Schreibweise abhängig. Die politische Zuordnung wird nicht immer den aktuellen Gegebenheiten entsprechen. Eine Gewähr wird also nicht übernommen.

Für Hinweise auf noch fehlende oder inkorrekte Links wären die Autoren dankbar (Mail an *hilf at isn-oldenburg.de*).

### Linkliste von National-Bibliotheken

#### 1. Ägypten

- (a) Egyptian National Library  
<http://www.darelkotob.org/ENGLISH/HTML/ABOUT.HTM>
- (b) Egyptian National Agricultural Library  
<http://nile.enal.sci.eg/>

#### 2. Albanien

- The National Library of the Republic of Albania  
<http://www.library.uiuc.edu/spx/class/nationalbib/natbibalbania.htm>

---

<sup>1</sup>[http://en.wikipedia.org/wiki/List\\_of\\_national\\_libraries](http://en.wikipedia.org/wiki/List_of_national_libraries)

3. Algerien  
Algeria Bibliotheque Nationale  
<http://www.biblionat.dz>
4. Andorra  
Biblioteca Nacional d'Andorra  
<http://bibnac.andorra.ad>  
<http://www.bibliotecanacional.ad/cast/web.htm>
5. Argentinien  
Argentina Biblioteca Nacional  
<http://www.bibnal.edu.ar>
6. Armenien  
National Library of Armenia  
<http://www.nla.am>
7. Aruba  
Aruba National Library  
<http://www.bibliotecanacional.aw>
8. Australien  
National Library of Australia  
<http://www.nla.gov.au>
9. Aserbeidjan  
National Library of Azerbaijan  
<http://www.anl.aznet.org>
10. Bahamas  
College of the Bahamas Library  
<http://www.cob.edu.bs/library/>
11. Bangladesh  
National Library of Bangladesh  
<http://www.nanl.gov.bd>
12. Barbados  
Barbados National Library Service  
natlib@caribsurf.com
13. Belgien  
Bibliotheque Royale de Belgique  
<http://www.kbr.be>
14. Belize  
National Library Service of Belize  
<http://www.nlsbze.bz/>
15. Belorussland  
National Library of Belarus  
<http://natlib.org.by/en/>

16. Benin  
Benin Bibliotheque Nationale  
[http://www.bj.refer.org/benin\\_ct/tur/bnb/Pagetitre.htm](http://www.bj.refer.org/benin_ct/tur/bnb/Pagetitre.htm)
17. Bermuda  
Bermuda Library  
<http://www.bermudanationallibrary.bm>
18. Bhutan  
National Library of Bhutan  
<http://www.ibrary.gov.bt>
19. Bosnien und Herzegovina  
Narodna i univerzitetska biblioteka Bosnia i Hercegovina Sarajevo  
<http://www.nub.ba/>
20. Botswana  
Botswana National Library Service  
[http://www.gov.bw/government/ministry\\_of\\_labour\\_and\\_home\\_affairs.html#national\\_library\\_service](http://www.gov.bw/government/ministry_of_labour_and_home_affairs.html#national_library_service)
21. Brasilien  
Brazil Biblioteca Nacional  
<http://www.bn.br>
22. Brunei  
Dewan Bahasa and Pustaka Brunei Library  
[http://www.dbp.gov.bn/library\\_Info/main.htm](http://www.dbp.gov.bn/library_Info/main.htm)
23. Bulgarien 'St Cyril and Methodius' National Library  
<http://www.nationallibrary.bg/>
24. Burkina Faso  
Bibliotheque Nationale du Burkina  
[http://www.culture.gov.bf/Site\\_Ministere/textes/ministere/ministere\\_sr\\_bn.htm](http://www.culture.gov.bf/Site_Ministere/textes/ministere/ministere_sr_bn.htm)
25. Chile  
Biblioteca Nacional de Chile  
[http://www.dibam.cl/biblioteca\\_nacional.htm](http://www.dibam.cl/biblioteca_nacional.htm)
26. China, Volksrepublik
  - (a) National Digital Library of China  
<http://www.nlc.gov.cn/>
  - (b) Hong Kong Public Library  
<http://www.hkpl.gov.hk/>
  - (c) Bibliotheka Central de Macau  
<http://www.library.gov.mo/>
27. Costa Rica  
Biblioteca Nacional  
<http://ns.binasss.sa.cr/>

## 28. Dänemark

- (a) Royal Library  
<http://www.kb.bib.dk> und <http://www.kb.dk>
- (b) Danish National Library of Science and Medicine  
<http://www.dnlb.dk/>
- (c) National Veterinary and Agricultural Library  
<http://www.dvjb.kvl.dk/>
- (d) National Library of Education  
<http://www.dpb.dpu.dk/>

## 29. Deutschland

Die Deutsche Bibliothek  
<http://www.ddb.de>

## 30. Dominikanische Republik

Biblioteca Nacional  
<http://www.bnrd.gov.do>

## 31. Ekuador

Biblioteca Nacional del Ecuador  
[http://cce.org.ec/index.php?action=seccion&id\\_c=49](http://cce.org.ec/index.php?action=seccion&id_c=49)

## 32. El Salvador

Biblioteca Nacional de El Salvador  
<http://www.binaes.gob.sv/>

## 33. Estland

National Library of Estonia  
<http://www.nlib.ee>

## 34. Äthiopien

National Library and Archives  
[http://www.nale.gov.et/national\\_library.html](http://www.nale.gov.et/national_library.html)

## 35. Faroer

Foroya Landsbokasavn  
<http://www.flb.fo>

## 36. Finnland

Helsinki University Library  
<http://www.lib-helsinki.fi/english>

## 37. Frankreich

Bibliothque nationale de France  
<http://www.bnf.fr>

## 38. Georgien

National Library of Georgia und National Parliamentary Library of Georgia  
<http://www.nplg.gov.ge>

39. Griechenland  
The National Library of Greece (Ethnike Bibliotheke tes Hellados)  
<http://www.nlg.gr>
40. Grönland  
Public and National Library of Greenland  
<http://www.katak.gl>
41. Großbritannien
  - (a) British Library  
<http://www.bl.uk>
  - (b) Scotland National Library of Scotland  
<http://www.nls.uk>
  - (c) Wales National Library of Wales  
<http://www.llgc.org.uk/>
  - (d) National Art Library  
<http://www.nal.vam.ac.uk/>
42. Guatemala  
Biblioteca Nacional de Guatemala  
<http://www.biblionet.edu.gt/>
43. Guyana  
National Library  
<http://www.natlib.gov.gy/>
44. Honduras  
Biblioteca Nacional de Honduras  
<http://www.binah.gob.hn>
45. Ungarn  
National Szechenyi Library  
<http://www.oszk.hu/eng/>
46. Island  
Landsbokasafn Islands National and University Library of Iceland  
<http://www.bok.hi.is>
47. Indien
  - (a) National Library Calcutta  
<http://www.nlindia.org/>
  - (b) National Science Library  
<http://www.niscair.res.in/InformationResources/info.asp?a=topframe.htm&b=leftcon.asp&c=ns1/ns1.htm&d=test>
  - (c) Indian National Scientific Documentation Centre (INSDOC)  
<http://www.nlindia.org>

48. Indonesien  
National Library of Indonesia  
<http://www.pnri.go.id>
49. Iran  
National Library of Iran  
<http://www.nli.ir>
50. Irland  
National Library of Ireland  
<http://www.nli.ie>
51. Israel  
Jewish National and University Library  
<http://jnu1.huji.ac.il>
52. Italien
  - (a) Biblioteca Nazionale Centrale "Vittorio Emanuele III"  
<http://www.bncrm.librari.beinculturali.it>
  - (b) Biblioteca Nazionale Centrale Firenze  
<http://www.bncf.firenze.sbn.it>
  - (c) National Central Library in Rome  
<http://www.bncrm.librari.beniculturali.it/>
53. Jamaika  
National Library of Jamaica  
<http://www.nlj.org.jm>
54. Japan  
National Diet Library  
<http://www.ndl.go.jp/en/>
55. Jordanien  
Department of Libraries, Documentation and Archives  
<http://www.nl.gov.jo/>
56. Kanada
  - (a) Libraries and Archives Canada  
<http://www.collectionscanada.ca>
  - (b) National Library of Canada  
<http://www.nlc-bnc.ca/>
  - (c) National Library of Québec  
<http://www.bibliat.gouv.qc.ca> und <http://www.bnquebec.ca/>
57. Kasachstan  
National Library of Kazakhstan  
<http://www.nlrk.kz>

58. Kenia  
National Library Services  
<http://www.knls.or.ke/>
59. Kiribati  
National Library and Archives Republic of Kiribati
60. Kolumbien  
Biblioteca Nacional de Colombia  
<http://www.bibliotecanacional.gov.co/>
61. Kroatien  
National and University Library  
<http://www.nsk.hr>
62. Kuba  
Biblioteca Nacional "Jose Marti"  
<http://www.bnjm.cu/>
63. Korea (South)  
National Library of Korea Republic of Korea (South)  
<http://www.nl.go.kr>
64. Kuwait  
Kuwait Central Library National Council for Culture, Art & Letters  
<http://www.nlp.gov.kw>
65. Kirgisien  
National Library of Kyrgyzstan  
<http://www.nlkr.gov.kg/>
66. Litauen  
National Library of Latvia  
<http://www.lnb.lv>
67. Libanon  
Bibliotheque Nationale du Liban  
<http://www.baakleennationallibrary.com/> und <http://www.bnlib.org/>
68. Libyen  
Libya (Socialist Peoples' Libyan Arab Jamahiriya) National Library of Libya  
<http://www.nll.8m.com/>
69. Liechtenstein  
Liechtensteinische Landesbibliothek  
<http://www.lbfl.li/>
70. Litauen  
Martynas Ma vydas National Library of Lithuania  
<http://www.lnb.lt>

71. Luxemburg  
Bibliothèque Nationale  
<http://www.bn1.lu/>
72. Mazedonien  
National and University Library Former Yugoslavian Republic of Macedonia  
<http://www.nubsk.edu.mk/>
73. Malaysia  
National Library of Malaysia  
<http://www.pnm.my>
74. Mali  
Bibliothèque Nationale  
<http://w3.culture.gov.ml/biblio/index.html>
75. Malta  
National Library of Malta  
<http://www.libraries-archives.gov.mt/nlm/>
76. Mauritius  
National Library of Mauritius  
<http://ncb.intnet.mu/mac/nlibrary/>
77. Mazedonien  
National and University Library  
<http://www.nubsk.edu.mk/EngNUB1.htm>
78. Mexiko
  - (a) Biblioteca Nacional de Mexico  
<http://biblional.bibliog.unam.mx/bib01.html>
  - (b) National Newspaper and Periodicals Library of Mexico  
<http://biblional.bibliog.unam.mx/hem/hemeroteca.html>
79. Moldovien  
National Library of Moldova  
<http://www.bnrm.md/>
80. Monaco  
Bibliothèque Louis Notari  
[http://www.monaco-mairie.mc/06\\_culture/mediat/mediat\\_bs.php](http://www.monaco-mairie.mc/06_culture/mediat/mediat_bs.php)
81. Montenegro  
Central National Library of the Republic of Montenegro  
[http://www.heritage.cg.yu/cnb\\_e.htm](http://www.heritage.cg.yu/cnb_e.htm)
82. Montenegro  
Central National Library of the Republic of Montenegro  
[http://www.heritage.cg.yu/cnb\\_e.htm](http://www.heritage.cg.yu/cnb_e.htm)



83. Marokko  
Bibliotheque Generale et Archives National Library and Archives of Morocco  
<http://www.minculture.gov.ma/fr/Biblioth%E8que%20G%E9n%E9rale%20et%20Archives%20de%20Rabat.htm>
84. Namibien  
National Library of Namibia  
<http://yaotto.natlib.mec.gov.na>
85. Nepal  
The National Library of Nepal  
<http://www.nnl.gov.np/nnl/>
86. Niederlande  
Koninklijke Bibliotheek  
<http://www.konbib.nl> und <http://www.kb.nl>
87. Neuseeland  
National Library of New Zealand  
<http://www.natlib.govt.nz>
88. Nicaragua  
Biblioteca Nacional 'Ruben Dario'  
National Health Library  
<http://www.minsa.gob.ni/bns/indexbns.htm>
89. Nigeria  
National Library of Nigeria  
<http://www.nlbn.org/>
90. Norwegen  
National Library of Norway  
<http://www.nb.no>
91. Oman  
Sultan Quaboos University Library  
<http://www.squ.edu.om/lib/>
92. Österreich  
Osterreichische Nationalbibliothek  
<http://www.onb.ac.at>
93. Pakistan  
National Library of Pakistan  
<http://www.nlp.gov.pk>
94. Panama  
Biblioteca Nacional  
<http://www.binal.ac.pa>
95. Papua New Guinea  
National Library of Papua New Guinea  
<http://www.dg.com.pg/~ola/>

96. Peru  
Biblioteca Nacional de Peru  
<http://www.binape.gob.pe>
97. Philippinen  
The National Library  
<http://www.nlp.gov.ph>
98. Polen  
Biblioteka Narodowa  
<http://www.bn.org.pl>
99. Portugal  
Biblioteca Nacional  
<http://www.bn.pt>
100. Qatar  
Public Libraries Department of National Council for Culture Arts and Heritage National Library of Qatar  
[http://www.cnc.com.qa/english/divisions.asp?content=public\\_libraries\\_department](http://www.cnc.com.qa/english/divisions.asp?content=public_libraries_department)
101. Rumänien  
Biblioteca Nationala  
<http://www.bibnat.ro>
102. Russland  
National Library of Russia  
<http://www.nlr.ru/eng/>  
Russian State Library  
<http://www.rsl.ru>
103. San Marino  
biblioteca@omniway.sm
104. Saudiarabien  
King Fahad National Library  
<http://www.kfnl.gov.sa>
105. Senegal  
Archives Nationales du Senegal  
<http://www.archivesdusenegal.gouv.sn/biblio.html>
106. Serbien  
National Library of Serbia  
<http://www.bms.ns.ac.yu/bmseng101.htm> und <http://www.nbs.bg.ac.yu>
107. Singapur  
National Library of Singapore  
<http://www.nlb.gov.sg>

108. Sloakei  
Slovak National Library  
<http://www.snk.sk/>
109. Slovenien  
National and University Library of Slovenia  
<http://www.nuk.uni-lj.si/vstop.cgi>
110. Südafrika  
National Library of South Africa  
<http://www.nlsa.ac.za>
111. Spanien
  - (a) Biblioteca Nacional  
<http://www.bne.es>
  - (b) Biblioteca de Catalunya  
<http://www.gencat.es/bc/> und <http://www.bnc.es/>
112. Sri Lanka  
National Library of Sri Lanka  
<http://www.natlib.lk>
113. Schweden  
The Royal Library Kungliga Biblioteket  
<http://www.kb.se>
114. Schweiz  
Bibliothèque Nationale Suisse  
<http://www.snl.ch>
115. Syrische arabische Republik  
Assad National Library  
<http://www.alassad-library.gov.sy/>
116. Taiwan  
National Central Library NCL Die nationale Bibliothek von Taiwan  
<http://www2.ncl.edu.tw/english/index.asp>
117. Thailand  
National Library  
<http://www.nlt.go.th/>
118. Tamil  
Nadu Agricultural University Agricultural College and Research India  
<http://md3.vsnl.net.in>
119. Trinidad and Tobago  
National Library and Information System  
<http://www.nalis.gov.tt/>

120. Tunesien  
Bibliothèque Nationale  
<http://www.bibliotheque.nat.tn>
121. Tschechische Republik  
National Library in Prague  
<http://www.nkp.cz>
122. Türkei  
Milli Kutuphane  
<http://web.mkutup.gov.tr>
123. Uganda  
National Library of Uganda  
<http://www.nlu.go.ug>
124. Ukraine
  - (a) National Parliamentary Library of Ukraine  
<http://www.nplu.kiev.ua>
  - (b) Kirovohrad Regional Library  
<http://www.library.kr.ua>
  - (c) National Parliamentary Library of Ukraine  
<http://www.rada.kiev.ua/LIBRARY/>
  - (d) Vernadsky National Library of Ukraine  
<http://www.nbu.gov.ua/>
125. Vereinigte Arabische Emirate  
Cultural Foundation National Library  
<http://www.cultural.org.ae/>
126. USA
  - (a) Library of Congress  
<http://lcweb2.loc.gov> und <http://lcweb.loc.gov/>
  - (b) National Library of Medicine  
<http://www.nlm.nih.gov>
  - (c) National Agricultural Library  
<http://www.nal.usda.gov>
  - (d) National Archives and Records Administration  
<http://www.nara.gov>
  - (e) National Library of Medicine  
<http://www.nlm.nih.gov/>
127. Uruguay  
Biblioteca Nacional del Uruguay  
<http://www.mec.gub.uy/>
128. Vatikan

- (a) Vatican Library  
[http://bav.vatican.va/en/v\\_home\\_bav/home\\_bav.shtml](http://bav.vatican.va/en/v_home_bav/home_bav.shtml)
  - (b) Biblioteca Apostolica Vaticana  
[http://www.vatican.va/library\\_archives/vat\\_library/index.htm](http://www.vatican.va/library_archives/vat_library/index.htm)
129. Venezuela  
Biblioteca Nacional de Venezuela  
<http://www.bnv.bib.ve>
130. Vietnam  
Bibliotheque Nationale  
<http://www.nlv.gov.vn/>
131. Zypern  
Municipal Library  
[http://libraries.theeuropeanlibrary.org/Cyprus/page\\_en.xml](http://libraries.theeuropeanlibrary.org/Cyprus/page_en.xml)



## Über die Autoren

*Thomas Severiens* ist Projektmanager im ISN. Er entwickelt und betreibt seit 1995 mit dem *PhysNet* einen internationalen weltweit verteilten Dienst für die Physik. Er hat Lehraufträge der Universitäten Enschede und Osnabrück im Fach Information Engineering. Er ist Chair der *Dublin-Core Arbeitsgruppe Tools* und Mitglied der *XQuery Arbeitsgruppe* des *World Wide Web Consortium W3C*. Forschungsschwerpunkt ist die Einbettung verteilter Digitaler Bibliotheken in eine Semantik-Web Infrastruktur. Erreichbar unter [thomas@severiens.de](mailto:thomas@severiens.de)

*Eberhard R. Hilf* ist Geschäftsführer des ISN. Er war von 1971 - 2000 Professor für Theoretische Physik an den Hochschulen Düsseldorf, Darmstadt, Oldenburg. Seit 1989 arbeitet er auf dem Gebiet des Managements wissenschaftlicher Information, jüngst zur semantischen Inhaltsbeschreibungssprache *Physics Markup Language*. Er ist Mitglied zahlreicher nationaler und internationaler wissenschaftlicher Organisationen und Boards. Erreichbar über <http://www.isn-oldenburg.de/~hilf>

Web-Adresse der Expertise:

<http://nbn-resolving.de/urn:nbn:de:0008-20051114021>

Ein *Forum* zu dieser Expertise soll über *Zugang zum Wissen*<sup>2</sup> eingerichtet werden. LeserInnen, die Annotationen, Kommentare oder Anregungen geben möchten, können diese dort nach Registrierung eingeben.

Alle (etwa 1.200) *Links* wurden im November 2005 zuletzt geprüft. Der Quelltext dieser Expertise ist in  $\text{\LaTeX}$  geschrieben und wird bei *Die Deutsche Bibliothek* zur Langzeit-Archivierung hinterlegt werden.

Rechtschreibung, Kommaregeln und Grammatik folgen einer gemäßigten Reformvariante, unterscheiden also z. B. zwischen abstrakter und konkreter Bedeutung eines Wortes<sup>3</sup> bzw. nutzen Kommata für die korrekte grammatische Zuordnung<sup>4</sup>.

---

<sup>2</sup><http://www.zugang-zum-wissen.de>

<sup>3</sup>„Sie hatten sich nicht auseinandergesetzt, um sich in Ruhe miteinander auseinanderzusetzen.“ Waren sie im Streit oder nicht?

<sup>4</sup>„Er riet der Stiftung Geld zu geben für die Bibliothek der Universität“. Zahlt oder erhält die Stiftung Geld?