

Zur Entwicklung eines Beschreibungsprofils für eine nationale Langzeit-Archivierungs-Strategie *digitaler* Dokumente - ein Beitrag aus der Sicht der Wissenschaften

Thomas Severiens und Eberhard R. Hilf [hilf@isn-oldenburg.de]
[ISN](#) Institute for Science Networking Oldenburg GmbH an der Carl von Ossietzky Universität

Auftrag: Entwicklungen im Ausland abfragen und auswerten als
 Expertise als [[nestor Materialie No. 7](#)] (zum downloaden [pdf] sowie als Print on demand)

Diese Seite trägt als *shadow file* die Metadaten und den Verweis auf den Quellort:
www.isn-oldenburg.de/~hilf/vortraege/nestor06/

- Dieses Dokument zum [Browsen](#) [zugleich der Quelltext zur Langzeitarchivierung]
- Dieses Dokument als [Slide-show](#)
- Dieses Dokument zum [Ausdrucken](#)

Urheberrechtsverfügung: *Jedes Slide darf von Ihnen auch einzeln bei Verweis auf die Quelle in eigenen Vorträgen verwendet werden. Ansonsten gilt creative commons (+ - -).*

Die ISN GmbH erarbeitet Expertisen für den Bereich der wissenschaftlichen Information

1

nestor Langzeitarchivierungs-Policy 2006, Koblenz

[zu slide 2](#)

I. Was ist die Herausforderung?

Gespräch unserer Generation mit den zukünftigen hat natürliche Schwierigkeiten..

Wir kennen ihre Kommunikations-Gewohnheiten, Erwartungen, technischen Möglichkeiten, Fragestellungen und Kenntnisse nicht.

Beispiel: die *Hohmann-Karte*: die 'BILD-Zeitung' von 1717.

Zitat:

Geographische Vorstellung der jämmerlichen Wasser-Flutt in Nieder-Teutschland, welche den 25.Dec. A.171, in der heiligen Christnacht mit unzähligen Schaden, und Verlust vieler tausend Menschen einen großen theil derer herzogth. Holstein und Bremen, die Grafsch. Oldenburg, Frislandt, Gröningen und Nort-Holland überschwemet hat, ediert von Joh. Bapt. Homann Der Röm. Keis. Mai. Geogr. In Nürnbergi
 siehe *Altkartendatenbank IKAR des Verbundes der Staatsbibliothek zu Berlin, ...*
 Drucke ab 1718;

Wie man heute weiß, gibt die Karten nicht unbedingt Fakten, wie wir Heutigen das erwarten, sondern will einen Eindruck verschaffen, wie schlimm die Flut damals war, die Schraffuren der Überflutungsgebiete enthalten dann schon mal wissentlich die eingezeichnete höchste Erhebung Norddeutschlands....

- Heute emailen wir ein Digitalisat;
- heute wollen wir, dass die Vorväter mehr Fakten aufbewahrt hätten;
- heute haben wir GPS und Autos, um die Küste abzufahren;
- heute wollen wir wissen, wo es überschwemmt war anno 1717;
- heute wissen wir mehr als die Archivare damals über sich...

Zum heutigen Verständnis eines Werkes aus der fernen Vergangenheit braucht es die Summe der seither gewonnenen Erkenntnisse über es.

2

nestor Langzeitarchivierungs-Policy 2006, Koblenz

[zu slide 3](#)

Den Generationenvertrag kann nur der Staat erfüllen

Die Verlage von 1717 gibt es lange nicht mehr.

Ihr Gewinnstreben (und damit ihr Interesse an einer Vorratshaltung) erlosch mit ihrem Ende.

Aber wie soll der Staat vorgehen?

Welche Vorgaben muss er machen?

3 nestor Langzeitarchivierungs-Policy 2006, Koblenz

[zu slide 3a](#)

II. Was können wir von anderen Staaten lernen?

Verfahren: Web-Umfrage

Fragestellungen erarbeitet mit H. Neuroth, R.Sietmann

1. Aktivitäten von wem? Stand? Gibt es eine nationale LZA-Policy?
2. Gibt es eine institutionelle LZA-Policy? Erfahrungen?
3. Gibt es eine zuständige Institution? Vereinbarungen?
4. Auswahlkriterien: wer wählt was aus? Formate?
5. Zugang? Finanzierung?
6. Nutzung? Erwartungen?

Adressaten: Nationale Bibliotheken (über 500)

Eine Webumfrage zu staatlichen Aufgaben ist ungeeignet

3 II. Was können wir von anderen Staaten lernen?

[zu slide 4](#)

Typische Antworten

1. Aktivitäten von wem? Stand? *Some early discussion*
Gibt es eine nationale LZA-Policy? *Nein*
2. Gibt es eine institutionelle LZA-Policy? *nur teilweise*
Erfahrungen? *Ja (aber bekannt)*
3. Gibt es zuständige Institution? *Vereinbarungen? Nein*
4. Auswahlkriterien: wer wählt was aus? *Selbst*
Formate? *Alles*
5. Zugang? *Beschränkt*
Finanzierung? *Selbst*
6. Nutzung? *Mengenproportional*
Erwartungen? *Hoch*

Statt einer Webumfrage daher eigene Recherchen, Gespräche

4 nestor Langzeitarchivierungs-Policy 2006, Koblenz

[zu slide 5](#)

Was lernen wir von der Vielfalt anderer Staaten?

- *Große, mittlere und kleine Staaten*
[Von der Vielfalt gewachsener Institutionen bis zu 'Einer macht es!;]
- *Industriestaaten und Entwicklungsstaaten*
[Von technologisch die Entwicklung treibend bis zu 'niemand da!;]
- *Interne Organisationsformen*
Von Commonwealth bis zu Diktaturen

Mitverantwortung Deutschlands für andere Staaten; [über Unesco, EU]
Lernen von und kooperieren mit vergleichbaren Staaten;

5 nestor Langzeitarchivierungs-Policy 2006, Koblenz

[zu slide 6](#)

Die Expertise gibt eine Übersicht über aktuelle Aktivitäten

1. *Es gibt bisher keine nationalen LZA-Policies,-- aber alle wollen sie*
2. *Pragmatische Vorerfahrungen mit der Archivierung von digitalen browsing/printing Images*
(insbesondere im kommerziellen pdf-Format)
3. *Es gibt gute institutionelle Policies*

4. *Es werden internationale Vereinbarungen gefordert:*
- *zu LZA-Metadaten (Nachweis)*
 - *zur Spiegelung von digitalen Materialien*
 - *zur gemeinsamen Erarbeitung der notwendigen Techniken*
 - *zur Auswahl und Arbeitsteilung*

Ansprechpartner sind die nationalen Bibliotheken und Archive von 195 Ländern ...

6 *nestor* Langzeitarchivierungs-Policy 2006, Koblenz

[zu slide 7](#)

III. Vielfalt der nationalen Herausforderungen

1. Vielfalt der Fachgebiets-Kommunikationskulturen

- *Wissenschaft: Exakte Wissenschaften vs. Geisteswissenschaften [Fakten, Formen, Phrasen als Mittler und Content]*
- *Originale aus Kultur, Bildung, Museen*
- *Kommunikationswege ändern sich [Kollaborationen (CREW); vs. Mathematik (Beweise)]*
- *Vielfalt der Fragen: Ist es richtig? Ist es original?, ...*

Mitwirkung der Fachgebiets-Experten bei der Definition, was wie langzeit-archiviert werden soll.
(ungewohnt ...)

7 III. Vielfalt der nationalen Herausforderungen

[zu slide 8](#)

2. Vielfalt der LZA-Anforderer

1. *Wissenschaftler*
2. *Industrie und Wirtschaft*
3. *Museen*
4. *Behörden: National-, Landes-, Lokal-; Vereine;*
5. *Privatpersonen.*

Mitwirkung der Anforderer und der Nutzer ist das Tor zur Akzeptanz

8 III. Vielfalt der nationalen Herausforderungen

[zu slide 9](#)

2. Vielfalt der Dokumenten-Arten

1. *Text-Dokumente;*
2. *Graphiken;*
3. *Ton/Bild-Mitschnitte;*
4. *Digitalisate;*
5. *Konglomerate, Patchwork, Kondensate.*

Technische Herausforderung:

Konditionierung vor dem Eingangstor zur LZA als nationale Policy.

Mitwirkung der Autoren und der Erzeuger

9 III. Vielfalt der nationalen Herausforderungen

[zu slide 10](#)

2. Vielfalt der Nutzer, Nutzung, Autoren

1. *Lesen (Entziffern, verstehen, einordnen können)*
2. *Weiterverarbeiten, z. B.:*
 - *Chemie (Reaktionen);*
 - *Mathematik (Beweise);*
 - *Physik (Numerische Beispiele)]*

3. kopieren, downloaden
4. Integrieren in eigene Werke (Kondensate; Reviews; ...)
5. Verwenden in Veranstaltungen (eLearning)

Kondensieren, Einbetten, Aktualisieren, Quelltexte aufbewahren

10

[zu slide 11](#)

IV. Realisierungs-Herausforderungen

1. Organisationsmodell

Verschiedene Staaten diskutieren dies:

zentral, verteilt, oder verteilt mit starker Leitungsstruktur?

Deutschland ist eine Bundes-Republik, daher

- durchsetzungsstarke nationale LZA-Policy
- Viel engere Kooperation der verteilten Kompetenz- und LZA-Institutionen

Organisations-Netz ist notwendig

11

IV. Realisierungsherausforderungen

[zu slide 12](#)

2. Technische Besonderheiten

digital weniger heterogen und problematisch als bei analogen Quellen

- Annahme nur von korrekten Codes [TeXDocCenter; W3C-Checker]
- Refreshing (halbjährlich)
- Migration, dynamische Erzeugung von Ansichten
- Redundanz [Calliau: digitale Kopien sind Klone]
- Spiegelungen im In- **und** Ausland
- Instanz für das Reagieren bei Ausfall von LZA-Archiven

Best Practise Beispiele:

Alle .tex Dokumente seit 1987 weltweit;

Alle .html seit 1993; [z.B. alle meine Vorträge [[nestor Workshop 2004](#); oder [Kooperation in der Fachinformation](#), Darmstadt Januar 1995 oder [Integrated Information Management in PhysicsAPS E-Print workshop](#); Los Alamos, Oktober 1994...]

Alle Dokumente im [ArXiv](#) seit 1991.

Respekt vor dem Werk: Quelldokument mit archivieren.

12

IV. Realisierungsherausforderungen

[zu slide 13](#)

3. Langzeit-Archivierung des Inhalts

- Metadaten: bibliographisch, LZA, Struktur, Inhalt
[Beispiel: OmDoc als neue Sprache für Semantik]
[Beispiel: Mathematik braucht inhaltliche Prüfbarkeit: contentMML]
- Vollständigkeit zur Rekonstruktion des Inhalts
[numerische Programme, Daten, Rohdaten, Beschreibungen]
- Auswahl nach der Relevanz unabhängig von der Quelle.

Archivieren mit Einbettung; Erläuterungsmaterial; Übertragung in Verstehbares

13 IV. Realisierungsherausforderungen [zu slide 14](#)

5. Nutzbarkeit und Retrieval

- *Intelligente Add-On-Dienste (werden vom Leser erwartet)*
- *Vernetzung der LZA-Institutionen zu weltweitem Nachweissystem*
- *Mehrwertdienste über dem Archiv.*

Vom Vorhalten auf Verdacht (Bibliothek) zum Anliefern on demand

14 IV. Realisierungsherausforderungen [zu slide 15](#)

6. Der Umgang mit Innovation durch Menschen

Das Tempo der technologischen Innovation wird sich weiter erhöhen (Vergleich nestor 2004 zu nestor 2006)

- *Bundesweite Standards für Anforderungen an Ausbildungslinien, sodass LZA professionell betrieben werden kann.*
- *Engere Kopplung an die technologische und wissenschaftliche Entwicklung*

Von LZA-Bibliotheken zu LZA-Institutionen digitaler Materialien durch Verzahnung mit der Wissenschaft (Wie schwer taten Sie sich schon mit einem Wissenschaftler)

15 [zu slide 16](#)

V. Zu einer Nationalen LZA-Policy

Was ist als Nächstes zu tun?

- *Den Prozess in Gang bringen zu einem straffen Gesetz [sehr viele politische Partner machen es zu einem zähen Geschäft.]*
- *spezifische LZA-Metadaten festlegen und verbreiten*
- *Internationale Beteiligung und Vernetzung professionalisieren*
- *Best Practise Beispiele von professioneller LZA unter Beteiligung der Erzeuger/Nutzer Community.*
- *Wege zur Nutzerakzeptanz mit den Erzeugern/Nutzern erproben*
- *Vernetzung der LZA-Beauftragten und Vernäherung mit der Wissenschaft*
- *Autoren-Incentives schaffen (Zertifikate, Zitierungen)*

Wo liegen die Chancen, im internationalen Wettbewerb voranzukommen?

- *Straffes Gesetz, dass die entsprechende Förderung langfristig sichert*
- *Öffnung der LZA-Institutionen zur Wissenschaft und zum Ausland*
- *Best Practise mit der Community eines kleinen Bereichs statt 'pragmatischem Archivieren'.*

15 [zum Beginn](#)

Schluss: Was muss auf der Wissenschaftlerseite getan werden:

- *Exemplarisch für einzelne Wissenschaftscommunities die Nützlichkeit, die Anforderungen an den Autor demonstrieren und die Kommunikation mit der gewählten Community zu Auswahlkriterien, Metadaten, Formaten, Erwartungen an die LZA in Gang setzen.*

Sie stehen vor einer Langzeit-Herausforderung: die Aufgaben und die Komplexität werden mit der Erfahrung wachsen ...