nestor
Kompetenznetzwerk Langzeitarchivierung

nestor Handbuch:

Eine kleine Enzyklopädie der digitalen Langzeitarchivierung

3 State of the Art

Einführung / Überblick

Herausgeber

Heike Neuroth
Hans Liegmann (†)
Achim Oßwald
Regine Scheffel
Mathias Jehn
Stefan Strathmann

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Im Auftrag 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

Kontakt

editors@langzeitarchivierung.de
c/o
Niedersächsische Staats- und Universitätsbibliothek Göttingen
Heike Neuroth
Forschung und Entwicklung
Papendiek 14
37073 Göttingen
Tel. +49 (0) 55 1 39 38 66

Der Inhalt steht unter folgender Creative Commons Lizenz:
http://creativecommons.org/licenses/by-nc-sa/2.0/de/

3 State of the Art

LZA-Aktivitäten in Deutschland aus dem Blickwinkel von nestor

Mathias Jehn, Sabine Schrimpf

Die Situation in Deutschland

Bibliotheken, Archive und Museen sind das wissenschaftliche, juristisch-administrative und kulturelle Gedächtnis einer Stadt, eines Landes, einer Nation. Sie sind Orte der Forschung und Wissensvermittlung, des Lernens und der Anschauung. Sie tragen die Verantwortung für die Erhaltung physisch vorhandener Originale ebenso wie für die langfristige Nutzbarkeit digitaler Informationen bzw. nachträglich angefertigter Digitalisate von anderen Kulturmedien. Gerade elektronische Publikationen oder, weiter gefasst, digitale Ressourcen nehmen in den meisten deutschen Einrichtungen einen stetig wachsenden Stellenwert ein und beeinflussen nachhaltig den Auftrag von Gedächtnisorganisationen. Die rasante Entwicklung auf diesem Gebiet stellt neue Anforderungen

hinsichtlich der dauerhaften Bewahrung und Zugänglichkeit dieser digitalen Objekte: So muss das digital publizierte Wissen auch unter den Bedingungen eines ständig stattfindenden Technologiewandels langfristig verfügbar gehalten werden, da der wissenschaftliche und technische Fortschritt eine regelmäßige Neubewertung älterer Wissensstände erfordert. Der digitalen Langzeitarchivierung kommt hierbei eine Schlüsselrolle zu. Letztlich stellt sie eine wesentliche Bedingung für die Konkurrenzfähigkeit des Bildungs- und Wissenschaftssystems und damit mittelbar auch für die wirtschaftliche Leistungsfähigkeit eines jeweiligen Landes dar.

Die digitale Langzeitsicherung erweitert das Aufgabenspektrum der archivierenden Institutionen, sodass neue organisatorische und technische Anstrengungen zur Sicherung und langfristigen Nutzbarkeit digitaler Objekte erforderlich sind. Ein Archiv, das sich erst bei anstehenden Lieferungen des elektronischen Schriftguts Gedanken über dessen Übernahme, Erschließung und die dauerhafte Speicherung macht, wird an der Komplexität der Aufgabe scheitern. Die dauerhafte Lesbarkeit von elektronischen Medien ist insbesondere durch den schnellen technischen Wandel von Datenträgern und -formaten sowie durch die permanente Veränderung und Weiterentwicklung der für die Nutzung notwendigen Anwendungsprogramme gefährdet. Die Arbeit, die im Bereich der physischen Datenträger geleistet wurde, vorgegeben durch gesetzliche Sammelaufträge oder Archivgesetze, hat deutlich werden lassen, dass sowohl für solch große Bereiche der Netzpublikationen wie ebooks, e-Journals, elektronische Hochschulschriften oder thematische Websites (bzw. Online-Ressourcen) gemeinsame und tragfähige Langzeitarchivierungsstrategien bislang noch fehlten. Dazu kommt, dass die Aufgaben sich in eine Vielzahl von Teilaspekten gliedern und daraus resultierenden Teilaufgaben von einer Institution allein nicht zu leisten sind. Neben den Bibliotheken werden auch die Archive in Zukunft mit einer wachsenden Zahl von Abgaben elektronischen Schriftguts rechnen müssen. Dieses Schriftgut aus den Behörden wird von Anfang an elektronisch („digital born“) erstellt und voraussichtlich die volle Bandbreite an Formen digitaler Unterlagen umfassen.

In Deutschland wurde das Thema zum ersten Mal 1995 in einem Positionspapier „Elektronische Publikationen“ der Deutschen Forschungsgemeinschaft (DFG) aufgegriffen und als Aufgabenbereich der Virtuellen Fachbibliotheken benannt. In Anbetracht sowohl des Umfangs der Aufgabe als auch der föderalen Struktur Deutschlands mit der Verantwortlichkeit seiner Bundesländer für Wissenschaft und Kultur, war es folgerichtig, dass der Ansatz zu einer erfolgreichen Lösung dieser Probleme nur ein kooperativer sein konnte. Aus

der gemeinsamen Arbeit an konzeptionellen Fragen der künftigen Entwicklung digitaler Bibliotheken im Rahmen des vom Bundesministeriums für Wissenschaft und Forschung (BMBF) getragenen Projektes „digital library konzepte“ ist eine Initiativgruppe Langzeitarchivierung hervorgegangen, deren Arbeitsplan im Rahmen einer 6-monatigen Folgeprojekts im Jahre 2002 auf zwei Workshops ausgewählten Experten des Informationswesens zur Diskussion gestellt wurden. Diese „Initialzündung“ für eine kooperative Lösung der Langzeitarchivierung digitaler Ressourcen resultierte in einem Papier mit Abschlussempfehlungen für zentrale Komponenten einer kooperativen digitalen Langzeiterhaltungsstrategie für Deutschland. Seit dem Jahr 2003 besteht mit dem BMBF-geförderten Projekt *nestor* ein nationales Kompetenznetzwerk zur Langzeitarchivierung und Langzeitverfügbarkeit digitaler Objekte, das als einziges seiner Art die in Deutschland identifizierbaren Kompetenzen bündelt und die Kontakte zu entsprechenden Initiativen und Fachgruppen koordiniert.[1] Mit der Einrichtung von *nestor* sollte gemeinsam den Defiziten bei der Langzeitarchivierung – unter Einbeziehung der „Produzenten“ digitaler Ressourcen, d. h. Verlage, Universitäten, Forschungseinrichtungen, Behörden, Wissenschaftler sowie technischer Dienstleister wie Rechen-, Daten- und Medienzentren und Großdatenbankbetreiber – begegnet werden. Die gemeinsame Fragestellung betrifft die dauerhafte Erhaltung sowohl genuin digitaler Objekte als auch retrodigitalisierter Ressourcen sowie die nachhaltige Verfügbarkeit dieser Informationen für spätere Generationen.

Mittlerweile verteilen sich in *nestor* die notwendigen Fachkompetenzen für den Aufgabenkomplex „Langzeitarchivierung digitaler Ressourcen“ über ein breites Spektrum von Personen, die in vielen Institutionen, Organisationen und Wirtschaftsunternehmen tätig sind. *nestor* bringt so die Experten der Langzeitarchivierung und aktive Projektnehmer zusammen und fördert den Austausch von Informationen, die Entwicklung von Standards sowie die Nutzung von Synergieeffekten. Alle Sparten der Gedächtnisinstitutionen werden bei der Herausforderung unterstützt, die Bewahrung und Verfügbarkeit aller digitalen Ressourcen selbst zu gewährleisten, die Bewahrung und Verfügbarkeit der wichtigsten Objekte jedes Fachgebiets zu organisieren sowie schließlich die Bewahrung und Verfügbarkeit digitaler Archivalien garantieren zu können.

Für Bibliotheken, Archive und Museen ist mit der Einrichtung von *nestor* ein wichtiger Schritt für die verteilte Übernahme konkreter Aufgaben und Absprа-

1 nestor ist das Akronym der englischen Übersetzung des Projekttitels: „Network of Expertise in long-term storage and availability of digital Resources in Germany“. Siehe: http://www.langzeitarchivierung.de.

chen in Deutschland getan. Ein zentrales Aufgabenfeld des Netwerks ist beispielsweise die Sicherung der Authentizität (im Sinne der Vertrauenswürdigkeit) des archivierten Dokuments. Im Prozess der Planung von Erhaltungsstrategien sind u.a. drei wichtige Arbeitsschritte zu vollziehen:

1. Da ein nationaler Alleingang in der globalen Informationsgesellschaft ein sicherer Misserfolgsfaktor wäre, ist es erstens wichtig, eine Bestandsaufnahme, Analyse und Auswertung der internationalen Entwicklungen vorzunehmen und zu prüfen, welche der bereits existierenden Lösungsvorschläge der deutschen Situation angemessen sein könnten.
2. Die Entwicklung von Norm-Standards ist unbedingt erforderlich. Diese sollten in Übereinstimmung mit den sich aktuell im internationalen Rahmen abzeichnenden Standardisierungsinitiativen erarbeitet werden.
3. Der Aufbau einer dezentralen und kooperativen Infrastruktur für die Archivierung digitaler Dokumente in Deutschland, die nicht nur Zuständigkeiten klar definiert sondern auch effektive und effiziente Kooperationsstrukturen etabliert, ist notwendig.

Zur Umsetzung dieser Ziele müssen weitere finanzielle Mittel zur Verfügung gestellt werden, weil mit der Langzeitarchivierung und -verfügbarkeit digitaler Objekte völlig unterschiedliche Bereiche betroffen sind.[2] Sobald einmal mit der Langzeitarchivierung begonnen wird, muss die langfristige Finanzierung gewährleistet sein. Zwar ist heute immer noch unklar, wie sich die Kosten in der Zukunft entwickeln werden, jedoch ist es sicher, dass einerseits große Geldsummen für den Aufbau und Betrieb von Langzeitarchivierungssystemen benötigt werden, andererseits der finanzielle Spielraum für den öffentlich-rechtlichen Bereich begrenzt sein wird. Es sind daher Strategien nötig, wie Gedächtnisorganisationen mit den begrenzten Mitteln die besten Ergebnisse erzielen können.

Kurzer Überblick über die Langzeitarchivierungssysteme und -projekte

In Deutschland gibt es schon einige Institutionen, die mit der digitalen Langzeitarchivierung begonnen haben. Auf Grund der komplexen und innovativen

2 Ein wichtiges Ergebnis der ersten Projektphase von 2003 bis 2006 war die Verabschiedung gemeinsamer Richtlinien: nestor hat in einem „Memorandum zur Langzeitverfügbarkeit digitaler Informationen in Deutschland“ die notwendigen Anstrengungen von politischen Entscheidungsträgern, Urhebern, Verlegern, Hard- und Softwareherstellern sowie kulturellen und wissenschaftlichen Gedächtnisorganisationen zusammengestellt, um die Rahmenbedingungen einer nationalen Langzeitarchivierungs-Policy abzustecken. Siehe: http://www.langzeitarchivierung.de/downloads/memo2006.pdf.

Herausforderungen, die mit dem Thema digitale Langzeitarchivierung verbunden sind, geschieht dies meist im Projektverbund.

Mit kopal („kooperativer Aufbau eines Langzeitarchivs digitaler Informationen") haben die Deutsche Nationalbibliothek in Kooperation mit der Niedersächsischen Staats- und Universitätsbibliothek (SUB) Göttingen, der Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) und IBM Deutschland kooperativ eine technische Lösung für die Bewahrung und langfristige Verfügbarkeit digitaler Dokumente erarbeitet.[3] Seit der Aufnahme des Produktivbetriebs im August 2006 hat kopal umfangreiche digitale Bestände von der Deutschen Nationalbibliothek und der SUB Göttingen in das Archivsystem eingespielt. Die beteiligten Institutionen konnten dabei wertvolle Erfahrungen für den Betrieb und die weitere Entwicklung des Archivsystems sammeln. Das kopal-Langzeitarchiv ist nach dem Projektende im Juni 2007 bei der Deutschen Nationalbibliothek und der (SUB) Göttingen, in den Routinebetrieb gegangen. Anlässlich des Abschlussworkshops „kopal goes live" am 13. Juni 2007 wurde ein Memorandum unterzeichnet, in dem sich die Partner langfristig den Aufgaben der Langzeitarchivierung verpflichten und einen Rahmen für die weitere Zusammenarbeit gesetzt haben.

Daneben wurde mit dem Pilotsystem „Bibliothekarisches Archivierungs- und Bereitstellungssystem" wurde ein weiteres Archivsystem an der Bayerischen Staatsbibliothek München in Zusammenarbeit mit dem Leibniz Rechenzentrum entwickelt.[4] Ziel des von der DFG geförderten Kooperationsprojektes war der Aufbau einer organisatorischen und technischen Infrastruktur für die Langzeitarchivierung und Bereitstellung von Netzpublikationen aus dem breiten Spektrum der Bayerischen Staatsbibliothek als Universal-, Landes- und SSG-Bibliothek sowie als Digitalisierungszentrum. Im Nachfolgeprojekt BABS II soll das Pilotsystem zu einem vertrauenswürdigen digitalen Langzeitarchivs als Teil kooperativer Strukturen und Evaluierung gemäß dem *nestor*-Kriterienkatalog ausgebaut werden. Evaluierbarkeit und Test der Skalierbarkeit des Gesamtsystems sollen einen langfristigen Betrieb mit Wachstumspotential gewährleisten.

Mit edoweb in Rheinland-Pfalz, BOA in Baden-Württemberg und Saardok im Saarland liegen kooperativ entwickelt und betriebene technische Plattformen für die Sammlung, Erschließung und langfristige Verfügbarkeit von regionalen elektronischen Pflichtexemplaren vor.[5]

---

3 http://kopal.langzeitarchivierung.de/

4 http://www.babs-muenchen.de/index.html?pcontent=startseite

5 http://www.lbz-rlp.de/cms/rheinische-landesbibliothek/digitale-angebote/edo-

Neben der Entwicklung kompletter Archivsystem-Lösungen befassen sich zahlreiche Institutionen in unterschiedlichen Projekten mit weiteren Aspekten der digitalen Langzeitarchivierung. *nestor* bündelt alle derartigen Projekte in Deutschland, im deutschsprachigen Raum sowie die mit Beteiligung deutscher Partner auf der *nestor*-Homepage. Das Themenspektrum der aufgeführten Projekte reicht von den hier beispielhaft vorgestellten Archivsystemen über die Strategiebildung hinsichtlich Langzeitarchivierung bis zur konkreten Entwicklung von Langzeitarchivierungswerkzeugen.

Neben diesen Beispielen aus Deutschland liegen auch einige gute Beispiele für erfolgreiche internationale Kooperationsprojekte im Bereich der Langzeitarchivierung vor. Im Bereich der technologischen Forschung sind die von der EU geförderten Forschungsprojekte PLANETS und CASPAR wichtige Einrichtungen,[6] etwa bei der Implementierung des Open Archival Information System, kurz OAIS-Modell.[7] Durch die Abgrenzung und eindeutige Benennung von Funktionsmodulen, Schnittstellen und Typen von Informationsobjekten ist es gelungen, eine einheitliche Sprache und eine über die Grenzen der Anwendergemeinschaften Archive, Datenzentren und Bibliotheken hinweg geltende allgemeine Sicht auf die Kernfunktionen eines digitalen Archivs zu schaffen. Gerade durch diese Allgemeingültigkeit ist der Abstraktionsgrad des Modells relativ hoch. Das Open Archival Information System beschreibt ein Informationsnetzwerk, das den Archivar und den Nutzer als Hauptkomponenten des digitalen Archivs versteht.

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

web/, http://www.boa-bw.de/, http://saardok.sulb.uni-saarland.de/

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

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

8 Die Kriterienkataloge sind hinterlegt in: http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91.

beispielhaft können hier DRIVER und DPE genannt werden.[9] Die Anbindung der eigenen Überlegungen an die Förderlinien der Europäischen Kommission ist wichtiger Bestandteil der Arbeit. Über Europa hinaus bestehen enge Bindungen z.B. an die frühere amerikanische Research Libraries Group und die australische Nationalbibliothek, gemeinsam mit außereuropäischen Partnern in den USA und Asien wird einmal jährlich eine internationale Konferenz organisiert (IPRES).[10]

9 Siehe: http://www.driver-repository.eu/ und http://www.digitalpreservationeurope.eu/.

10 Siehe: http://rdd.sub.uni-goettingen.de/conferences/ipres/ipres-en.html.