nestor
Kompetenznetzwerk Langzeitarchivierung

nestor Handbuch:

Eine kleine Enzyklopädie der digitalen Langzeitarchivierung

8 Vertrauenswürdigkeit von digitalen Langzeitarchiven

Herausgeber:

Heike Neuroth
Hans Liegmann
Achim Oßwald
Regine Scheffel
Mathias Jehn

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

Dieser Artikel ist ein Auszug aus dem nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung

Dieser Artikel ist verfügbar unter der URL:
http://nestor.sub.uni-goettingen.de/handbuch/artikel/text_84.pdf

Die Online Version des Handbuches unter der URL:
http://nestor.sub.uni-goettingen.de/handbuch/

Kontakt:
Niedersächsische Staats- und Universitätsbibliothek Göttingen
Dr. Heike Neuroth
Forschung und Entwicklung
Papendiek 14
37073 Göttingen
neuroth@sub.uni-goettingen.de
Tel. +49 (0) 55 1 39 38 66

8 Vertrauenswürdigkeit von digitalen Langzeitarchiven

von Susanne Dobratz und Astrid Schoger

Grundkonzepte der Vertrauenswürdigkeit

Bedrohungen für den Informationserhalt
Informationen, die durch digitale Objekte repräsentiert sind, sind bedroht durch Einbußen in ihrer Integrität, Authentizität und Vertraulichkeit sowie den gänzlichen Verlust der Verfügbarkeit und Nutzbarkeit. Besondere Herausforderung für die Langzeitarchivierung stellen die physische Alterung der Datenträger, die Trennung der Informationen von ihren originären Datenträgern sowie die rapiden Veränderungen der für die Interpretation der digitalen Objekte erforderlichen technischen Infrastruktur dar.
Digitale Langzeitarchive haben den Erhalt der Informationen über lange Zeiträume hinweg zum Ziel. Deshalb ergreifen sie sowohl organisatorische als auch technische Maßnahmen, um diesen Bedrohungen entgegenzuwirken. Vertrauenswürdige digitale Langzeitarchive operieren nach ihren Zielen und Spezifikationen.
Folgende Begriffe orientieren sich am OAIS-Modell (siehe entsprechendes Kapitel im Handbuch).

Digitales Objekt, Metadaten
Ein digitales Objekt ist eine logisch abgegrenzte Informationseinheit in der Form digitaler Daten. Daten sind maschinenlesbare und –bearbeitbare Repräsentationen von Information, in digitaler Form (eine Bitfolge, also eine Folge von Nullen und Einsen). Zur Nutzung der Informationen müssen die digitalen Daten interpretiert (dekodiert) werden.
Der Informationsbegriff umfasst hier jeden Typ von Wissen, der ausgetauscht werden kann; zum Beispiel aus inhaltlicher Sicht etwa Werke geistiger Schöpfung, Ergebnisse der Forschung und Entwicklung, Dokumentationen des politischen, sozialen und wirtschaftlichen Handelns.
Zu den Daten, die die Inhaltsinformation repräsentieren (Inhaltsdaten), können weitere Daten hinzukommen, die z.B. der Identifizierung, der Auffindbarkeit, der Rekonstruktion und Interpretation oder dem Nachweis der Integrität und Authentizität sowie der Kontrolle der Nutzungsrechte dienen (Metadaten). Metadaten können zu unterschiedlichen Zeiten im Lebenszyklus digitaler Objekte entstehen (z.B. bei der Produktion, bei der Archivierung, bei der Bereitstellung für die Nutzung). Sie werden als Teile der logischen Einheit „digitales Objekt“ aufgefasst und können sowohl getrennt als auch gemeinsam mit den Inhaltsdaten verwaltetet werden.

Digitales Langzeitarchiv, Vertrauenswürdigkeit
Unter einem digitalen Langzeitarchiv wird eine Organisation (bestehend aus Personen und technischen Systemen) verstanden, die die Verantwortung für den Langzeiterhalt und die Langzeitverfügbarkeit digitaler Objekte sowie für ihre Interpretierbarkeit zum Zwecke der Nutzung durch eine bestimmte Zielgruppe übernommen hat. Dabei bedeutet „Langzeit“ über Veränderungen in der Technik (Soft- und Hardware) hinweg und auch unter Berücksichtigung möglicher Änderungen der Zielgruppe. Vertrauenswürdigkeit (engl. trustworthiness) wird als Eigenschaft eines Systems angesehen, gemäß seinen Zielen und Spezifikationen zu operieren (d.h. es tut genau das, was es zu tun vorgibt). Aus Sicht der IT-Sicherheit stellen Integrität, Authentizität, Vertraulichkeit und Verfügbarkeit Grundwerte dar. IT-Sicherheit ist somit ein wichtiger Baustein für vertrauenswürdige digitale Langzeitarchive.
Anhand eines Kriterienkatalogs kann die Vertrauenswürdigkeit digitaler Langzeitarchive geprüft und bewertet werden.

Ein digitales Langzeitarchiv entsteht als komplexer Gesamtzusammenhang. Die Umsetzung der einzelnen Kriterien muss stets vor dem Hintergrund der Ziele des Gesamtsystems gesehen werden. Sowohl die Realisierung des digitalen Langzeitarchivs als Ganzes als auch die Erfüllung der einzelnen Kriterien läuft als Prozess in mehreren Stufen ab:

1. Konzeption
2. Planung und Spezifikation
3. Umsetzung und Implementation
4. Evaluierung

Im Zuge der ständigen Verbesserung sind diese Stufen nicht als starres Phasenmodell zu betrachten sondern zu wiederholen.

Kriterien zur Evaluierung der Vertrauenswürdigkeit

Aktuelle internationale Entwicklungen
International beschäftigen sich mehrere Organisationen und Projekte mit der Vertrauenswürdigkeit digitaler Langzeitarchive. Ausgangspunkt aller Bemühungen war der RLG-OCLC-Bericht „Trusted Digital Repositories: Attributes and Responsibilities" (Mai 2002)[1]. Darauf aufbauend hat die RLG/NARA-Task Force 2006 den Entwurf „Audit Checklist for Certifying Digital Repositories"[2] veröffentlicht, der in den USA vom Center of Research Libraries sowie in Europa vom Digital Curation Centre und dem EU-Projekt „Digital Preservation Europe“ als Werkzeug zum Audit verschiedenster digitaler Langzeitarchive verwendet wurde und aufgrund dieser Erfahrungen nun überarbeitet wird. nestor hat unter Berücksichtigung nationaler Ansätze und Arbeitsergebnisse (z.B. „DINI-Zertifikat für Dokumenten- und Publikationsserver“[3]) den Kriterienkatalog nationalen Rahmenbedingungen und Bedürfnissen der deutschen Gedächtnisorganisationen angepasst und im Sommer 2006 als Entwurf zur öffentlichen Kommentierung veröffentlicht.[4]
Kriterien für die Vertrauenswürdigkeit digitaler Langzeitarchive befinden sich zurzeit im Prozess internationaler Abstimmung und Standardisierung im Rahmen der ISO. Kürzlich haben CRL, DCC, DPE und nestor zehn gemeinsame Prinzipien herausgearbeitet, die den oben genannten Kriterienkatalogen und Audit-Checklisten zu Grunde liegen.
Als Beispiel eines Kritereinkataloges zur Evaluierung der Vertrauenswüdigkeit digitaler Langzeitarchive wird im Folgenden der nestor-Kriterienkatalog vorgestellt.

Der nestor-Kriterienkatalog „Vertrauenswürdige digitale Langzeitarchive“
Der nestor-Kriterienkatalog richtet sich in erster Linie an Gedächtnisorganisationen (Archive, Bibliotheken, Museen) und dient als Leitfaden, um ein vertrauenswürdiges digitales Langzeitarchiv zu konzipieren, zu planen und umzusetzen. Ferner kann er auf allen Stufen der Entwicklung zur Selbstkontrolle eingesetzt werden.
Darüber hinaus dient der Katalog allen Institutionen, die selbst archivieren, sowie Dienstleistern aus dem kommerziellen und nichtkommerziellen Bereich und Drittanbietern von Produkten als Orientierungshilfe.

[1] “Trusted Digital Repositories: Attributes and Responsibilities: An RLG-OCLC Report”, RLG/OCLC Working Group on Digital Archive Attributes, 2002, http://www.rlg.org/en/pdfs/repositories.pdf
[2] “Audit Checklist for Certifying Digital Repositories”, RLG-NARA Task Force on Digital Repository Certification, 2006, http://www.rlg.org/en/pdfs/rlgnara-repositorieschecklist.pdf
[3] „DINI-Zertifikat für Dokumenten- und Publikationsserver“, Deutsche Initiative für Netzwerkinformation (DINI)/AG Elektronisches Publizieren 2003, http://www.dini.de/documents/Zertifikat.pdf
[4] „Kriterienkatalog vertrauenswürdige digitale Langzeitarchive Version 1 (Entwurf zur öffentlichen Kommentierung)“ herausgegeben von der nestor-Arbeitsgruppe Vertrauenswürdige Archive - Zertifizierung. (Frankfurt am Main: nestor-Materialien 8), 2006, http://nbn-resolving.de/urn:nbn:de:0008-2006060710

Grundprinzipien bei der Herleitung und Anwendung der Kriterien

Abstraktion
Ziel ist es, Kriterien zu formulieren, die für ein breites Spektrum digitaler Langzeitarchive angewendet werden können und über längere Zeit Gültigkeit behalten sollen. Deshalb wird von relativ abstrakten Kriterien ausgegangen. Den Kriterien werden jeweils ausführliche Erläuterungen und konkrete Beispiele aus verschiedenen Bereichen mitgegeben. Die Beispiele entsprechen dem heutigen Stand der Technik und Organisation und sind unter Umständen nur im Kontext einer spezifischen Archivierungsaufgabe sinnvoll. Sie haben keinen Anspruch auf Vollständigkeit.

Dokumentation
Die Ziele, die Konzeption und Spezifikation sowie die Implementierung des digitalen Langzeitarchivs sind angemessen zu dokumentieren. Anhand der Dokumentation kann der Entwicklungsstand intern und extern bewertet werden. Eine frühzeitige Bewertung kann auch dazu dienen, Fehler durch eine ungeeignete Implementierung zu vermeiden. Insbesondere erlaubt eine angemessene Dokumentation auf allen Stufen, die Schlüssigkeit eines digitalen Langzeitarchivs umfassend zu bewerten. Auch alle Qualitäts- und Sicherheitsnormen fordern eine angemessene Dokumentation.

Transparenz
Die Transparenz wird durch die Veröffentlichung geeigneter Teile der Dokumentation realisiert.
Transparenz nach *außen* gegenüber Nutzern und Partnern ermöglicht diesen, selbst den Grad an Vertrauenswürdigkeit festzustellen. Transparenz gegenüber Produzenten und Lieferanten bietet diesen die Möglichkeit, zu bewerten, wem sie ihre digitalen Objekte anvertrauen.
Die Transparenz nach *innen* dokumentiert gegenüber den Betreibern, den Trägern, dem Management sowie den Mitarbeitern die angemessene Qualität des digitalen Langzeitarchivs und sichert die Nachvollziehbarkeit der Maßnahmen.
Bei denjenigen Teilen der Dokumentation, die für die breite Öffentlichkeit nicht geeignet sind (z.B. Firmengeheimnisse, Informationen mit Sicherheitsbezug), kann die Transparenz auf einen ausgewählten Kreis (z.B. die zertifizierende Stelle) beschränkt werden.
Durch das Prinzip der Transparenz wird Vertrauen aufgebaut, da es die unmittelbare Bewertung der Qualität eines digitalen Langzeitarchivs durch Interessierte zulässt.

Angemessenheit
Das Prinzip der Angemessenheit berücksichtigt die Tatsache, dass keine absoluten Maßstäbe möglich sind, sondern dass sich die Bewertung immer an den Zielen und Aufgaben des jeweiligen digitalen Langzeitarchivs ausrichtet. Die Kriterien müssen im Kontext der jeweiligen Archivierungsaufgabe gesehen werden. Deshalb können ggf. einzelne Kriterien irrelevant sein. Auch der notwendige Erfüllungsgrad eines Kriteriums kann – je nach den Zielen und Aufgaben des digitalen Langzeitarchivs – unterschiedlich ausfallen.

Bewertbarkeit
Für die Vertrauenswürdigkeit existieren zum Teil - insbesondere unter Langzeitaspekten - keine objektiv bewertbaren (messbaren) Merkmale. In diesen Fällen ist man auf Indikatoren angewiesen, die den Grad der Vertrauenswürdigkeit repräsentieren. Transparenz macht auch die Indikatoren für eine Bewertung zugänglich.

Überblick über die nestor-Kriterien für die Vertrauenswürdigkeit digitaler

Langzeitarchive

Aus Gründen der Übersichtlichkeit wird im folgenden der Term „digitales Langzeitarchiv" mit „dLZA" abgekürzt.
Der Katalog gliedert sich in drei Kapitel: „Organisatorischer Rahmen“, „Umgang mit Objekten“ und „Infrastruktur und Sicherheit“.

A. Organisatorischer Rahmen

1. Das dLZA hat seine Ziele definiert.
1.1 Das dLZA hat Kriterien für die Auswahl seiner digitalen Objekte entwickelt.
1.2 Das dLZA übernimmt die Verantwortung für den dauerhaften Erhalt der durch die digitalen Objekte repräsentierten Informationen.
1.3 Das dLZA hat seine Zielgruppe(n) definiert.
2. Das dLZA ermöglicht seinen Zielgruppe(n) eine angemessene Nutzung der durch die digitalen Objekte repräsentierten Informationen.
2.1 Das dLZA ermöglicht seinen Zielgruppe(n) den Zugang zu den durch die digitalen Objekte repräsentierten Informationen.
2.2 Das dLZA stellt die Interpretierbarkeit der digitalen Objekte durch seine Zielgruppe(n) sicher.
3. Gesetzliche und vertragliche Regelungen werden eingehalten.
3.1 Es bestehen rechtliche Regelungen zwischen Produzenten und dem digitalen Langzeitarchiv.
3.2 Das dLZA handelt bei der Archivierung auf der Basis rechtlicher Regelungen.
3.3 Das dLZA handelt bei der Nutzung auf der Basis rechtlicher Regelungen.
4. Die Organisationsform ist für das dLZA angemessen.
4.1 Die Finanzierung des digitalen Langzeitarchivs ist sichergestellt.
4.2 Es steht Personal mit angemessener Qualifikation in ausreichendem Umfang zur Verfügung.
4.3 Es bestehen angemessene Organisationsstrukturen für das dLZA.
4.4 Das dLZA betreibt eine langfristige Planung.
4.5 Die Fortführung der festgelegten Aufgaben ist auch über das Bestehen des digitalen Langzeitarchivs hinaus sichergestellt.
5. Es wird ein angemessenes Qualitätsmanagement durchgeführt.
5.1 Alle Prozesse und Verantwortlichkeiten sind definiert.
5.2 Das dLZA dokumentiert alle seine Elemente nach einem definierten Verfahren.
5.3 Das dLZA reagiert auf substantielle Veränderungen.

B. Umgang mit Objekten

6. Das dLZA stellt die Integrität der digitalen Objekte auf allen Stufen der Verarbeitung sicher.
6.1 Aufnahme (Ingest): Das dLZA sichert die Integrität der digitalen Objekte.
6.2 Archivablage (Archival Storage): Das dLZA sichert die Integrität der digitalen Objekte .
6.3 Nutzung (Access): Das dLZA sichert die Integrität der digitalen Objekte.
7. Das dLZA stellt die Authentizität der digitalen Objekte und Metadaten auf allen Stufen der Verarbeitung sicher.
7.1 Aufnahme (Ingest): Das dLZA sichert die Authentizität der digitalen Objekte.
7.2 Archivablage (Archival Storage): Das dLZA sichert die Authentizität der digitalen Objekte.
7.3 Nutzung (Access): Das dLZA sichert die Authentizität der digitalen Objekte.
8. Das dLZA betreibt eine langfristige Planung seiner technischen Langzeiterhaltungsmaßnahmen.
9. Das dLZA übernimmt digitale Objekte von den Produzenten nach definierten Vorgaben.
9.1 Das dLZA spezifiziert seine Übergabeobjekte (Submission Information Packages, SIPs).
9.2 Das dLZA identifiziert, welche Eigenschaften der digitalen Objekte für den Erhalt von Informationen signifikant sind.
9.3 Das dLZA erhält die physische Kontrolle über die digitalen Objekte, um Langzeitarchivierungsmaßnahmen durchführen zu können.

10. Die Archivierung digitaler Objekte erfolgt nach definierten Vorgaben.
10.1 Das dLZA definiert seine Archivobjekte (Archival Information Packages, AIPs).
10.2 Das dLZA sorgt für eine Transformation der Übergabeobjekte in Archivobjekte.
10.3 Das dLZA gewährleistet die Speicherung und Lesbarkeit der Archivobjekte.
10.4 Das dLZA setzt Strategien zum Langzeiterhalt für jedes Archivobjekt um.
11. Das dLZA ermöglicht die Nutzung der digitalen Objekte nach definierten Vorgaben.
11.1 Das dLZA definiert seine Nutzungsobjekte (Dissemination Information Packages, DIPs).
11.2 Das dLZA gewährleistet eine Transformation der Archivobjekte in Nutzungsobjekte.
12. Das Datenmanagement ist dazu geeignet, die notwendigen Funktionalitäten des digitalen Langzeitarchivs zu gewährleisten.
12.1 Das dLZA identifiziert seine Objekte und deren Beziehungen eindeutig und dauerhaft.
12.2 Das dLZA erhebt in ausreichendem Maße Metadaten für eine formale und inhaltliche Beschreibung und Identifizierung der digitalen Objekte.
12.3 Das dLZA erhebt in ausreichendem Maße Metadaten zur strukturellen Beschreibung der digitalen Objekte.
12.4 Das dLZA erhebt in ausreichendem Maße Metadaten, die die vom Archiv vorgenommenen Veränderungen an den digitalen Objekten verzeichnen.
12.5 Das dLZA erhebt in ausreichendem Maße Metadaten zur technischen Beschreibung der digitalen Objekte.
12.6 Das dLZA erhebt in ausreichendem Maße Metadaten, die die entsprechenden Nutzungsrechte und -bedingungen verzeichnen.
12.7 Die Zuordnung der Metadaten zu den Objekten ist zu jeder Zeit gegeben.

C. Infrastruktur und Sicherheit

13. Die IT-Infrastruktur ist angemessen.
13.1 Die IT-Infrastruktur setzt die Forderungen aus dem Umgang mit Objekten um.
13.2 Die IT-Infrastruktur setzt die Sicherheitsanforderungen des IT-Sicherheitskonzeptes um.
14 Die Infrastruktur gewährleistet den Schutz des digitalen Langzeitarchivs und seiner digitalen Objekte.