H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, K. Huth (Hrsg.)

nestor Handbuch

Eine kleine Enzyklopädie der digitalen Langzeitarchivierung

Version 2.3

Kapitel 5

Vertrauenswürdigkeit von digitalen Langzeitarchiven

nestor

nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung
hg. v. H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, K. Huth
im Rahmen des Projektes: 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,
Dr. Heike Neuroth, Forschung und Entwicklung, Papendiek 14, 37073 Göttingen

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter http://www.d-nb.de/ abrufbar.

Neben der Online Version 2.3 ist eine Printversion 2.0 beim Verlag Werner Hülsbusch, Boizenburg erschienen.

Die digitale Version 2.3 steht unter folgender Creative-Commons-Lizenz:
„Namensnennung-Keine kommerzielle Nutzung-Weitergabe unter gleichen Bedingungen 3.0 Deutschland“
http://creativecommons.org/licenses/by-nc-sa/3.0/de/

Markenerklärung: Die in diesem Werk wiedergegebenen Gebrauchsnamen, Handelsnamen, Warenzeichen usw. können auch ohne besondere Kennzeichnung geschützte Marken sein und als solche den gesetzlichen Bestimmungen unterliegen.

URL für Kapitel 5 „Vertrauenswürdigkeit von digitalen Langzeitarchiven“ (Version 2.3):
urn:nbn:de:0008-2010062448
http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0008-2010062448

Gewidmet der Erinnerung an Hans Liegmann (†), der als Mitinitiator und früherer Herausgeber des Handbuchs ganz wesentlich an dessen Entstehung beteiligt war.

5 Vertrauenswürdigkeit von digitalen Langzeitarchiven

5.1 Einführung

Susanne Dobratz und Astrid Schoger

Gedächtnisorganisationen wird gemeinhin beim Umgang mit ihren Beständen großes Vertrauen entgegengebracht - insbesondere beim Erhalt des ihnen anvertrauten kulturellen, wissenschaftlichen und gesellschaftlichen Erbes. Wissenschaftliche Erkenntnisse, historische Dokumente und kulturelle Leistungen liegen in stark zunehmendem Maße - und häufig ausschließlich - in digitaler Form vor. Diese spezielle Medienform stellt Gedächtnisorganisationen vor neue Herausforderungen und wirft die berechtigte Frage auf, ob sie auch im digitalen Bereich vertrauenswürdig handeln.

Das Konzept der Vertrauenswürdigkeit digitaler Langzeitarchive, die Kriterienkataloge, Checklisten und Werkzeuge helfen archivierenden Einrichtungen sowie Produzenten und Nutzern die Qualität und Nachhaltigkeit der Langzeitarchivierung zu bewerten und zu verbessern.

5.2 Grundkonzepte der Vertrauenswürdigkeit und Sicherheit

Susanne Dobratz und Astrid Schoger

Der Begriff der Vertrauenswürdigkeit digitaler Langzeitarchive wird von bewährten Konzepten der Vertrauenswürdigkeit von IT-Systemen abgeleitet. Im Internet Security Glossary[1] wird Vertrauenswürdigkeit (engl. trustworthiness) als die Eigenschaft eines Systems definiert, gemäß seinen Zielen und Spezifikationen zu operieren (d.h. es tut genau das, was es zu tun vorgibt) und dies auch in geeigneter Weise glaubhaft zu machen (z.B. durch eine formale Analyse). Die Common Criteria[2] führen Vertrauenswürdigkeit folgendermaßen ein:

„Werte" an deren Erhaltung „Eigentümer" Interesse haben, sind durch „Risiken"[3] bedroht. Zur Minimierung dieser Risiken werden „Gegenmaßnahmen" eingesetzt. Die Prüfung und Bewertung der eingesetzten Maßnahmen erbringt den Nachweis der „Vertrauenswürdigkeit". Vgl. dazu nachfolgende Abbildungen.

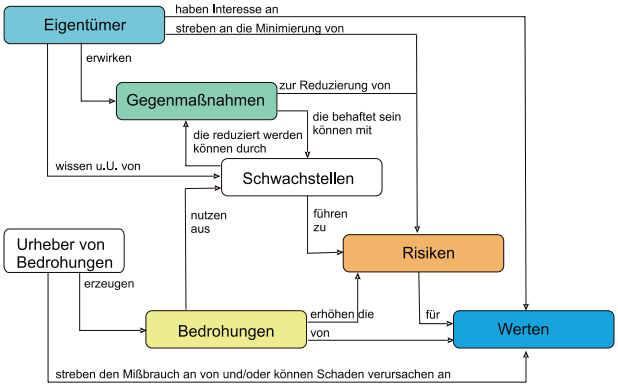

Abbildung 1: Konzept der Bedrohungen und Risiken gemäß Common Criteria (Quelle: BSI 2006)

1 Network Working Group (2007): "trusted system: a system that operates as expected,according to design and policy, doing what is required; trustworthy system: a system that not only is trusted, but also warrants that trust because the system's behavior can be validated in some convincing way, such as through formal analysis or code review." Alle hier aufgeführten URLs wurden im Mai 2010 auf Erreichbarkeit geprüft.

2 BSI (2006)

3 Vgl. dazu auch Howard (1998)

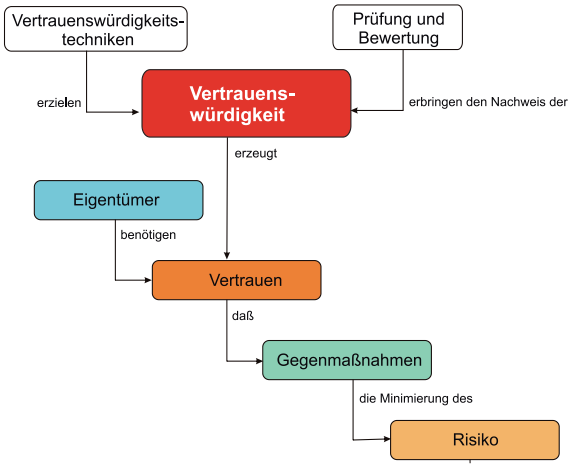

Abbildung 2: Vertrauenswürdigkeitskonzept gemäß den Common Criteria (Quelle: BSI 2006)

Ziel der digitalen Langzeitarchivierung ist der Erhalt der Informationen, die durch digitale Objekte repräsentiert sind.

Im Sinne der Langzeitarchivierung stellen Information den zu erhaltenden „Wert" dar. Informationen, die durch digitale Objekte repräsentiert werden, sind bedroht durch Einbußen in ihrer Integrität, Authentizität und Vertraulichkeit sowie den gänzlichen Verlust der Verfügbarkeit und Nutzbarkeit. Diese Eigenschaften bilden eine Teilmenge des Gesamtkonzeptes Sicherheit in der Informatik, wie sie u. a. in Steinmetz (2002) beschrieben sind:

- Integrität: sagt aus, ob die digitalen Objekte unverändert vorliegen,
- Authentizität: bezieht sich auf die Echtheit der digitalen Objekte, insbesondere den Aspekt der Nachweisbarkeit der Identität des Erstellers (Urhebers, Autors),
- Vertraulichkeit: bezieht sich darauf, dass unberechtigten Dritten kein Zugang zu den digitalen Objekten gewährleistet wird,
- Verfügbarkeit: bezieht sich auf den Aspekt der Zugänglichkeit zum digitalen Objekt.

Gemäß OAIS[4] wird unter einem digitalen Langzeitarchiv 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

4 Vgl. dazu CCSDS (2002) sowie Kapitel 4 dieses Handbuchs

Berücksichtigung möglicher Änderungen der Zielgruppe.

Aus dem Ziel der Langzeitarchivierung lassen sich folgende zentrale Aufgaben eines digitalen Langzeitarchivs ableiten: Aufnahme (Ingest), Archivablage (Archival Storage), Nutzung (Access); ferner unterstützende Aufgaben wie das Datenmanagement und die Administration des Gesamtsystems. Besondere Bedeutung kommt der strategischen Planung (Preservation Planning) und Durchführung der Langzeiterhaltungsmaßnahmen[5], die die Langzeitverfügbarkeit und Interpretierbarkeit (d.h. der Rekonstruierbarkeit der darin enthaltenen Informationen) sicherstellen, zu. [6]

Diese Aufgaben stellen die Grundlage für die Spezifikation von Anforderungen an digitale Langzeitarchive dar, wie bereits 2002 im Bericht[7] der RLG/OCLC Working Group on Digital Archive Attributes aufgeführt. Die RLG-NARA Task Force on Digital Repository Certification hat 2007 (als Entwurf zur öffentlichen Kommentierung bereits 2006) eine Liste von Kriterien „Trustworthy Repositories Audit and Certification: Criteria and Checklist (TRAC)“[8] erarbeitet, die ein vertrauenswürdiges digitales Langzeitarchiv erfüllen muss. Diese Liste dient der Orientierung beim Auf- und Ausbau digitaler Langzeitarchive und kann als Checkliste auch zur Selbstevaluierung sowie zum externen Audit eingesetzt werden.

nestor hat unter Berücksichtigung nationaler Ansätze und Arbeitsergebnisse wie des „DINI-Zertifikats für Dokumenten- und Publikationsserver“[9] sowie den oben genannten internationalen Arbeiten Kriterien entwickelt, die den speziellen Bedürfnissen der deutschen Gedächtnisorganisationen Rechnung tragen. Diese wurden zunächst im Sommer 2006 als Entwurf zur öffentlichen Kommentierung publiziert und dank der vielfältigen Rückmeldungen der Anwender gründlich überarbeitet und liegen nun in Version 2 vor.[10]

Eine Prüfung und Bewertung digitaler Langzeitarchive gemäß dieser Kriterienkataloge kann somit den Nachweis der Vertrauenswürdigkeit erbringen. Die Grundprinzipien der Kriterienkataloge sowie deren Inhalte werden in Kapitel 5.4 anhand des nestor-Kriterienkataloges genauer erörtert.

Im Rahmen des EU-Projektes DigitalPreservationEurope (DPE) in Zusammenarbeit mit dem Digital Curation Centre (DCC) wurde das Tool „Digital

5 Vgl. dazu Kapitel 8 dieses Handbuchs

6 Vgl. dazu das Funktionsmodell des OAIS-Referenzmodells

7 RLG/OCLC (2002)

8 RLG/NARA (2007)

9 DINI (2007)

10 nestor AG Vertrauenswürdige Archive - Zertifizierung (2008)

Repository Audit Method based on Risk Assessment (DRAMBORA)"[11] zur Selbstevaluierung entwickelt, das die Risikoanalyse als Methode einsetzt. Ausgehend von den Zielen eines digitalen Langzeitarchivs müssen zunächst die Aktivitäten spezifiziert und die damit verbundenen Werte identifiziert werden. In einem weiteren Schritt werden dann die Risiken aufgedeckt und die zu deren Minimierung eingesetzten Maßnahmen bewertet.

Somit wird ein anderer Weg zum Nachweis der Vertrauenswürdigkeit beschrieben.

Internationale Kooperation, Standardisierung und Zertifizierung – 10 gemeinsame Prinzipien

Bevor ein international abgestimmtes Zertifizierungsverfahren für digitale Langzeitarchive entwickelt werden kann, ist es zunächst wichtig, einen internationalen Konsens über die Evaluierungskriterien zu finden. Ferner müssen aus den Erfahrungen mit der Anwendung der Kriterienkataloge und Evaluierungstools Bewertungsmaßstäbe für unterschiedliche Typen von digitalen Langzeitarchiven ausgearbeitet werden.

Wesentliche Vertreter des Themas Vertrauenswürdigkeit auf internationaler Ebene - Center for Research Libraries (CRL), Digital Curation Centre (DCC), Projekt DigitalPreservationEurope (DPE) sowie nestor haben 10 gemeinsame Prinzipien[12] herausgearbeitet, die den oben genannten Kriterienkatalogen und Audit Checklisten zu Grunde liegen. Diese stellen die Grundlage der weiteren inhaltlichen Zusammenarbeit dar. Die 10 Kriterien lauten wie folgt[13]:

1. Das digitale Langzeitarchiv übernimmt die Verantwortung für die dauerhafte Erhaltung und kontinuierliche Pflege der digitalen Objekte für die identifizierten Zielgruppen.
2. Das digitale Langzeitarchiv belegt die organisatorische Beständigkeit (auch in den Bereichen Finanzierung, Personalausstattung, Prozesse), um seine Verantwortung zu erfüllen.
3. Das digitale Langzeitarchiv verfügt über die erforderlichen Rechte (per Vertrag oder Gesetz), um seine Verantwortung zu erfüllen.
4. Das digitale Langzeitarchiv besitzt ein effektives und effizientes Geflecht von Grundsätzen (policy).
5. Das digitale Langzeitarchiv erwirbt und übernimmt digitale Objekte auf der Grundlage definierter Kriterien gemäß seinen Verpflichtungen und

11 DCC/DPE (2008)
12 CRL/DCC/DPE/nestor (2007)
13 nestor-Übersetzung

Fähigkeiten.

6. Das digitale Langzeitarchiv stellt die Integrität, Authentizität und Nutzbarkeit der dauerhaft aufbewahrten Objekte sicher.
7. Das digitale Langzeitarchiv dokumentiert alle Maßnahmen, die während des gesamten Lebenszyklus auf die digitalen Objekte angewendet werden, durch angemessene Metadaten.
8. Das digitale Langzeitarchiv übernimmt die Bereitstellung der digitalen Objekte.
9. Das digitale Langzeitarchiv verfolgt eine Strategie zur Planung und Durchführung von Langzeiterhaltungsmaßnahmen.
10. Das digitale Langzeitarchiv besitzt eine angemessene technische Infrastruktur zur dauerhaften Erhaltung und Sicherung der digitalen Objekte.

Sowohl die Kriterienkataloge als auch das Verfahren DRAMBORA werden zur Zeit der nationalen sowie internationalen Standardisierung zugeführt. Im Rahmen des DIN kümmert sich der neu gegründete Arbeitskreis „Vertrauenswürdige digitale Archive“ im Rahmen des NABD15 „Schriftgutverwaltung und Langzeitverfügbarkeit digitaler Informationsobjekte“ um die Vorbereitung einer deutschen Norm[14] für diesen Bereich. Dieser Arbeitskreis arbeitet eng zusammen mit den entsprechenden Ausschüssen der ISO und zwar dem TC46/SC11, der mit dem Entwurf „Risk assessment for records systems” die Normungsarbeit an DRAMBORA übernommen hat sowie dem TC20/SC13, der „TRAC: Trustworthy Repositories Audit and Certification: Criteria and Checklist“ einer Normierung zuführt.

Weitere Fachgemeinschaften haben sich auf der Grundlage dieser Vorarbeiten entschieden, eigene, für ihre jeweiligen Fachgebiete angepasste, Kriterien in Form von Katalogen aufzu-stellen, so z.B. die Europäische Ground Segment Coordination Body (GSCB)[15] oder das Data Seal of Approval[16].

Die Anwendung von Konzepten der IT-Sicherheit wie Hashfunktionen, Fingerprintingverfahren und digitalen Signaturen, kann bestimmte Risiken, die den Erhalt digitaler Objekte bedrohen, minimieren, insbesondere jene, welche die Integrität, Authentizität und Vertraulichkeit digitaler Objekte betreffen. Von besonderer Bedeutung für die Langzeitarchivierung ist der „Langzeit“-Aspekt, so dass bei allen eingesetzten Methoden die Nachhaltigkeit besonders geprüft werden muss. Diese Verfahren werden im nachfolgenden Kapitel dargestellt.

---

14 DIN 31644: Information und Dokumentation - Kriterienkatalog für vertrauenswürdige digitale Langzeitarchive

15 GSCB (2009)

16 Data Seal of Approval (2009)

Literatur

Network Working Group (2007): *Internet Security Glossary. Request for Comments: 4949* http://tools.ietf.org/html/rfc4949

BSI Bundesamt für Sicherheit in der Informationstechnik (2006): *Gemeinsame Kriterien für die Prüfung und Bewertung der Sicherheit von Informationstechnik, Common Criteria* V 3.1, https://www.bsi.bund.de/cln_183/ContentBSI/Themen/ZertifizierungundAkkreditierung/ZertifierungnachCCundITSEC/ITSicherheitskriterien/CommonCriteria/cc.html

Howard, John D. / Longstaff, Thomas A. (1998): *A Common Language for Computer Security Incidents*. SANDIA Reports SAND98-8667. Albuquerque, New Mexico : Sandia National Laboratories http://www.cert.org/research/taxonomy_988667.pdf

CCSDS (Consultative Committee for Space Data Systems) (2002*): Reference Model for an Open Archival Information System (OAIS). Blue Book.* http://www.ccsds.org/docu/dscgi/ds.py/Get/File-143/650x0b1.pdf entspricht ISO 14721:2003

RLG/OCLC Working Group on Digital Archive Attributes (2002): *Trusted Digital Repositories: Attributes and Responsibilities*, http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf

DINI Deutsche Initiative für Netzwerkinformation / AG Elektronisches Publizieren (2007): *DINI-Zertifikat für Dokumenten- und Publikationsservice 2007*. DINI-Schriften 3. http://nbn-resolving.de/urn:nbn:de:kobv:11-10079197

nestor Arbeitsgruppe Vertrauenswürdige Archive – Zertifizierung (2008): *nestor-Kriterien: Kriterienkatalog vertrauenswürdige digitale Langzeitarchive.* Version 2. Frankfurt am Main : nestor http://nbn-resolving.de/urn:nbn:de:0008-2008021802

Steinmetz, Ralf (2000) : *Multimedia-Technologie: Grundlagen, Komponenten und Systeme*, 3. Auflage , Berlin, Heidelberg, New York : Springer

DCC Digital Curation Centre / DPE Digital Preservation Europe (2008): *Digital Repository Audit Method Based on Risk Assessment (DRAMBORA), interactive,* http://www.repositoryaudit.eu/

CRL Center for Research Libraries / DCC Digital Curation Centre / DPE Digital Preservation Europe / nestor (2007): *Core Requirements for Digital Archives*, http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/core-re

Data Seal of Approval (2009): Guidelines http://www.datasealofapproval.org/?q=node/35

GSCB Ground Segment Coordination Body (2009): Long Term Preservation of Earth obser-vation Space-Data - European LTDP Common Guidelines http://earth.esa.int/gscb/ltdp/EuropeanLTDPCommonGuidelines_Issue1.0.pdf

5.3 Praktische Sicherheitskonzepte

Siegfried Hackel, Tobias Schäfer und Wolf Zimmer

5.3.1 Hashverfahren und Fingerprinting

Ein wichtiger Bestandteil praktischer Sicherheitskonzepte zum Schutz der Integrität und Vertraulichkeit digitaler Daten sind Verschlüsselungsinfrastrukturen auf der Basis sogenannter kryptographisch sicherer Hashfunktionen. Mit Hilfe kryptographisch sicherer Hashfunktionen werden eindeutige digitale „Fingerabdrücke" von Datenobjekten berechnet und zusammen mit den Objekten versandt oder gesichert. Anhand eines solchen digitalen „Fingerabdrucks" ist der Empfänger oder Nutzer der digitalen Objekte in der Lage, die Integrität eines solchen Objektes zur prüfen, bzw. unautorisierte Modifikationen zu entdecken.

Hashfunktionen werden in der Informatik seit langem eingesetzt, bspw. um im Datenbankumfeld schnelle Such- und Zugriffsverfahren zu realisieren. Eine Hashfunktion ist eine mathematisch oder anderweitig definierte Funktion, die ein Eingabedatum variabler Länge aus einem Urbildbereich (auch als „Universum" bezeichnet) auf ein (in der Regel kürzeres) Ausgabedatum fester Länge (den Hashwert, engl. auch message digest) in einem Bildbereich abbildet. Das Ziel ist, einen „Fingerabdruck" der Eingabe zu erzeugen, die eine Aussage darüber erlaubt, ob eine bestimmte Eingabe aller Wahrscheinlichkeit nach mit dem Original übereinstimmt.

Da der Bildbereich in der Regel sehr viel kleiner ist als das abzubildende „Universum", können so genannte „Kollisionen" nicht ausgeschlossen werden. Eine Kollision wird beobachtet, wenn zwei unterschiedliche Datenobjekte des Universums auf den gleichen Hashwert abgebildet werden.

Für das Ziel, mit einer Hashfunktion einen Wert zu berechnen, der ein Datenobjekt eindeutig charakterisiert und damit die Überprüfung der Integrität von Daten ermöglicht, sind derartige Kollisionen natürlich alles andere als wünschenswert. Kryptographisch sichere Hashfunktionen H, die aus einem beliebig langen Wort M aus dem Universum von H einen Wert H(M), den Hashwert fester Länge erzeugen, sollen daher drei wesentliche Eigenschaften aufweisen:

1. die Hashfunktion besitzt die Eigenschaften einer effizienten Ein-Weg-Funktion, d.h. für alle M aus dem Universum von H ist der Funktionswert h = H(M) effizient berechenbar und es gibt kein effizientes Verfahren, um aus dem Hashwert h die Nachricht zu berechnen[17],
2. es ist - zumindest praktisch - unmöglich zu einem gegebenen Hashwert h = H(M) eine Nachricht M' zu finden, die zu dem gegebenen Hashwert passt (Urbildresistenz),
3. es ist - zumindest praktisch – unmöglich, zwei Nachrichten M und M' zu finden, die denselben Hashwert besitzen (Kollisionsresistenz).

Praktisch unmöglich bedeutet natürlich nicht praktisch ausgeschlossen, sondern bedeutet nicht mehr und nicht weniger, als dass es bspw. sehr schwierig ist, ein effizientes Verfahren zu finden, um zu einer gegebenen Nachricht M eine davon verschiedene Nachricht M' zu konstruieren, die denselben Hashwert liefert. Für digitale Objekte mit binären Zeichenvorräten Z = {0,1} lässt sich zeigen, dass für Hashfunktionen mit einem Wertbereich von 2n verschiedenen Hashwerten, beim zufälligen Ausprobieren von 2n/2 Paaren von verschiedenen Urbildern M und M' die Wahrscheinlichkeit einer Kollision schon größer als 50% ist.

Beim heutigen Stand der Technik werden Hashfunktionen mit Hashwerten der Länge n = 160 Bit als hinreichend stark angesehen.[18] Denn, selbst eine Schwäche in der Kollisionsresistenz, wie bereits im Jahre 2005 angekündigt[19], besagt zunächst einmal lediglich, dass ein Angreifer zwei verschiedene Nachrichten erzeugen kann, die denselben Hashwert besitzen. Solange aber keine Schwäche der Urbildresistenz gefunden wird, dürfte es für einen Angreifer mit einem gegebenen Hashwert und passendem Urbild immer noch schwer sein, ein zweites, davon verschiedenes Urbild zu finden, das zu diesem Hashwert passt.

Kern kryptographischer Hashfunktionen sind Folgen gleichartiger Kompressionsfunktionen K, durch die eine Eingabe M blockweise zu einem Has-

17 Obwohl die Ein-Weg-Funktionen in der Kryptographie eine wichtige Rolle spielen, ist nicht bekannt, ob sie im streng mathematischen Sinne eigentlich existieren, ihre Existenz ist schwer zu beweisen. Man begnügt sich daher zumeist mit Kandidaten, für die man die Eigenschaft zwar nicht formal bewiesen hat, für die aber derzeit noch keine effizienten Verfahren zur Berechnung der Umkehrfunktion bekannt sind.

18 Ein Rechner, der in der Lage ist, pro Sekunde den Hashwert zu einer Million Nachrichten zu berechnen, bräuchte 600.000 Jahre, um eine zweite Nachricht zu ermitteln, deren Hashwert mit einem vorgegebenen Hashwert der Länge 64 Bit übereinstimmt. Derselbe Rechner könnte allerdings in etwa einer Stunde irgendein Nachrichtenpaar mit gleichem Hashwert finden.

19 Schneier (2005)

hwert verarbeitet wird. Um Eingaben variabler Länge zu komprimieren, wendet man den Hashalgorithmus f iterierend an. Die Berechnung startet mit einem durch die Spezifikation des Hashalgorithmus festgelegten Initialwert $f(0):=I_0$. Anschließend gilt:

$$f(i) := K\,(f(i-1),M_i) \text{ mit } M = M_1, \ldots, M_n,\ i = 1, \ldots, n$$
$$H(M):= f(n) = h \text{ ist der Hashwert von } M$$

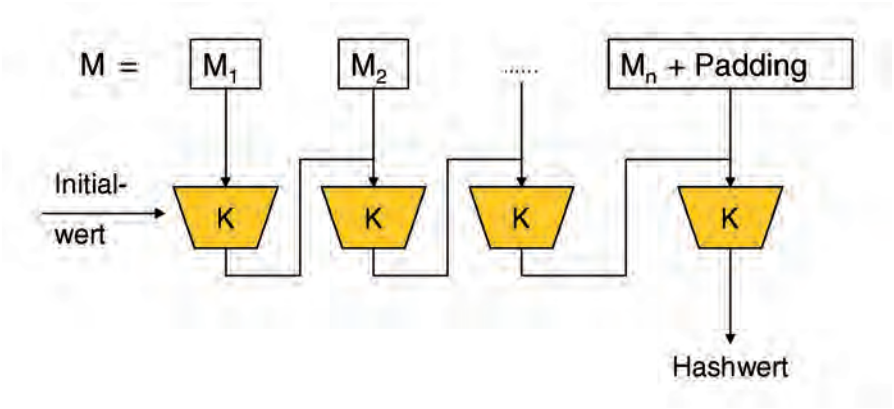

Abbildung 3: Allgemeine Arbeitsweise von Hashfunktionen (nach C. Eckert [20])

Neben auf symmetrischen Blockchiffren, wie dem bereits 1981 durch das American National Standards Institute (ANSI) als Standard für den privaten Sektor anerkannten Data Encryption Standard (DES)[21], finden heute vor allem Hashfunktionen Verwendung, bei denen die Kompressionsfunktionen speziell für die Erzeugung von Hashwerten entwickelt wurden. Der bislang gebräuchlichste Algorithmus ist der Secure Hash Algorithm SHA-1 aus dem Jahre 1993.[22]

Der SHA-1 erzeugt Hashwerte von der Länge 160 Bits[23] und verwendet eine Blockgröße von 512 Bits, d.h. die Nachricht wird immer so aufgefüllt, dass die Länge ein Vielfaches von 512 Bit beträgt. Die Verarbeitung der 512-Bit Ein-

20 Eckert (2001)
21 Vgl. bspw. Schneier (1996)
22 Vgl. bspw. Schneier (1996)
23 Da nicht ausgeschlossen werden kann, dass mit der Entwicklung der Rechentechnik künftig auch Hashwerte von der Länge 160 Bit nicht mehr ausreichend kollisions- und urbildresistent sind, wird heute für sicherheitstechnisch besonders sensible Bereiche bereits der Einsatz der Nachfolger SHA-256, SHA-384 und SHA-512 mit Bit-Längen von jeweils 256, 384 oder 512 Bits empfohlen.

gabeblöcke erfolgt sequentiell, für einen Block benötigt SHA-1 insgesamt 80 Verarbeitungsschritte.

Merkle-Hashwertbäume

In der Kryptographie und Informatik werden Merkle-Bäume[24] eingesetzt, um über große Mengen von Daten oder Dokumenten einen zusammenfassenden Hashwert zu bilden.
Die Blätter eines Merkle-Baums sind die Hashwerte der Dokumente $H_i(d_i)$. Jeder Knoten im Baum wird als Hashwert $H(h_1|h_2)$ seiner Kinder h_1 und h_2 gebildet. Dabei ist h_1 der Hashwert des Dokuments d_1 und h_2 der Hashwert des Dokuments d_2 und „|“ die Verkettung (Konkatenation) der beiden Hashwerte. Die Abbildung 4 zeigt ein Beispiel für einem binären Merkle-Hashbaum.

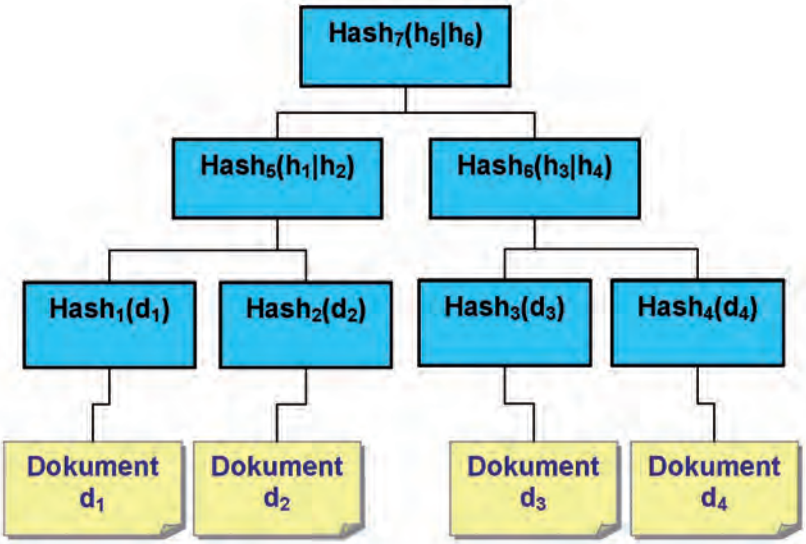

Abbildung 4: Binärer Merkle-Hashbaum.

Hash-Bäume können zum Nachweis der Integrität beliebiger Daten verwendet werden, die in oder zwischen Rechnern gespeichert, verarbeitet oder übertragen werden.

Die meisten Hash-Baum-Implementierungen sind binär, d.h. jeder Knoten besitzt zwei Kinder, es sind jedoch auch mehrere Kind-Knoten möglich.

24 Im Jahr 1979 von Ralph Merkle erfunden

Holt man für den obersten Hashwert (H_7 in Abbildung 4) einen Zeitstempeln ein, lässt sich beweisen, dass die Dokumente zum aktuellen Zeitpunkt existiert haben und seitdem nicht manipuliert wurden. Dieses Verfahren wurde im ArchiSig-Konzept[25] aufgegriffen und weiterentwickelt, um das Problem der nachlassenden Sicherheitseignung von Hash- und Signaturalgorithmen zu lösen. Dabei wird mit Hilfe von (akkreditierten) Zeitstempeln die vom deutschen Signaturgesetz geforderte Übersignatur von signierten Dokumenten vorgenommen.

5.3.2 Digitale Signatur

Elektronische Signaturen sind „Daten in elektronischer Form, die anderen elektronischen Daten beigefügt oder logisch mit ihnen verknüpft sind und die zur Authentifizierung im elektronischen Rechts- und Geschäftsverkehr dienen. Ihre Aufgabe ist die Identifizierung des Urhebers der Daten, d.h. der Nachweis, dass die Daten tatsächlich vom Urheber herrühren (Echtheitsfunktion) und dies vom Empfänger der Daten auch geprüft werden kann (Verifikationsfunktion). Beides lässt sich nach dem heutigen Stand der Technik zuverlässig am ehesten auf der Grundlage kryptographischer Authentifizierungssysteme, bestehend aus sicheren Verschlüsselungsalgorithmen sowie dazu passenden und personifizierten Verschlüsselungs-Schlüsseln (den so genannten Signaturschlüsseln) realisieren.

Die Rechtswirkungen, die an diese Authentifizierung geknüpft werden, bestimmen sich aus dem Sicherheitsniveau, das bei ihrer Verwendung notwendig vorausgesetzt wird. Dementsprechend unterscheidet das im Jahre 2001 vom deutschen Gesetzgeber veröffentlichte „Gesetz über Rahmenbedingungen für elektronische Signaturen und zur Änderung weiterer Vorschriften“[26], kurz Signaturgesetz (SigG), vier Stufen elektronischer Signaturen:

- „Einfache elektronische Signaturen“ gem. § 2 Nr. 1 SigG,
- „Fortgeschrittene elektronische Signaturen“ gem. § 2 Nr. 2 SigG,
- „Qualifizierte elektronische Signaturen“ gem. § 2 Nr. 3 SigG,
- „Qualifizierte elektronische Signaturen“ mit Anbieter-Akkreditierung gem. § 15 Abs. 1 SigG.

Mit Ausnahme der einfachen elektronischen Signaturen, denen es an einer verlässlichen Sicherheitsvorgabe völlig fehlt, wird das mit der Anwendung

25 Siehe www.archisig.de
26 BGBl I 876; BT-Drs 14/4662 und 14/5324

elektronischer Signaturen angestrebte Sicherheitsniveau grundsätzlich an vier Elementen festgemacht (§ 2 Nr. 2 SigG). Elektronische Signaturen müssen demnach

- ausschließlich dem Signaturschlüssel-Inhaber zugeordnet sein,
- die Identifizierung des Signaturschlüssel-Inhabers ermöglichen,
- mit Mitteln erzeugt werden, die der Signaturschlüssel-Inhaber unter seiner alleinigen Kontrolle halten kann und

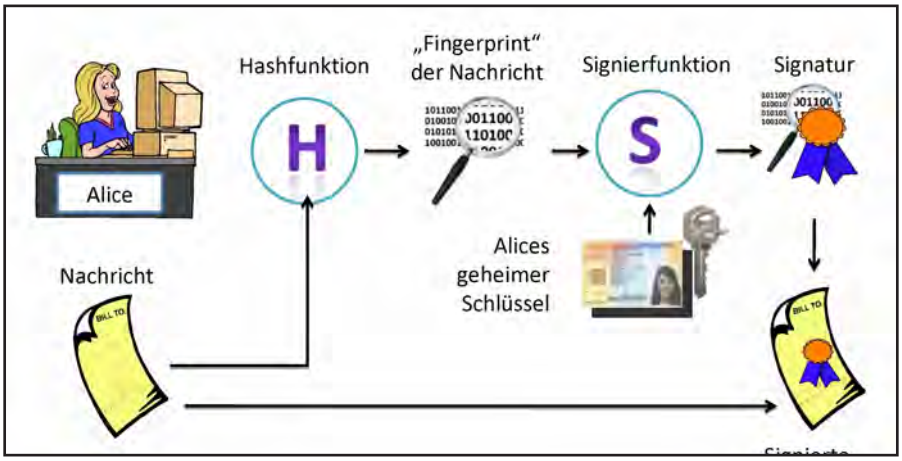

Abbildung 5: Digitale Signatur

- mit den Daten, auf die sie sich beziehen, so verknüpft sein, dass eine nachträgliche Veränderung der Daten erkannt werden kann.

Europaweit als Ersatz für die handschriftliche Unterschrift akzeptiert werden jedoch lediglich qualifizierte elektronische Signaturen. Für sie wird zusätzlich gefordert (§ 2 Nr. 3 SigG), dass sie

- auf einem zum Zeitpunkt ihrer Erzeugung gültigen qualifizierten Zertifikat beruhen und
- mit einer sicheren Signaturerstellungseinheit erzeugt werden.

Das Zertifikat übernimmt in diesem Fall die Authentizitätsfunktion, d.h. es bescheinigt die Identität der elektronisch unterschreibenden Person.[27] Sichere

27 Nach § 2 Nr. 6 SigG sind Zertifikate elektronische Bescheinigungen, mit denen

Signaturerstellungseinheiten sind nach dem Willen des Gesetzgebers Software- oder Hardwareeinheiten, die zur Speicherung und Anwendung des Signaturschlüssels dienen.[28]

Das Verfahren der digitalen Signatur basiert auf so genannten asymmetrischen kryptographischen Authentifizierungssystemen, bei denen jeder Teilnehmer ein kryptographisches Schlüsselpaar besitzt, bestehend aus einem geheimen privaten Schlüssel (private key, Kpriv) und einem öffentlichen Schlüssel (public key, Kpub).

Eine wesentliche Eigenschaft solcher asymmetrischer Authentifizierungssysteme ist, dass es praktisch unmöglich ist, den privaten Schlüssel aus dem öffentlichen Schlüssel herzuleiten, der öffentliche Schlüssel wird durch Anwendung einer sogenannten Einwegfunktion aus dem privaten Schlüssel berech-

Abbildung 6: Hinzufügung der Signaturdaten

net. Der öffentliche Schlüssel kann daher in einem öffentlich zugänglichen Verzeichnis hinterlegt werden, ohne damit den privaten Schlüssel preiszugeben.

Der Urheber, respektive Absender elektronischer Daten „unterschreibt" nun seine Daten, indem er sie mit seinem geheimen, privaten Schlüssel verschlüsselt. Jeder, der die Daten empfängt, kann sie dann mit dem öffentlichen Schlüssel wieder entschlüsseln (s. Abbildung 5).

Signaturschlüssel einer Person zugeordnet werden und die Identität einer Person bescheinigt wird. Für die Anwendung von Signaturverfahren von besonderer Bedeutung ist die Feststellung, dass „qualifizierte Zertifikate" nur auf natürliche Personen ausgestellt werden dürfen.

28 Das deutsche Signaturgesetz fordert, § 17 Abs. 1 SigG, dass sichere Signaturerstellungseinheiten vor unberechtigter Nutzung zu schützen sind. Nach § 15 Abs. 1 der Verordnung zur elektronischen Signatur (SigV) ist hierfür eine Identifikation „durch Besitz und Wissen oder durch Besitz und ein oder mehrere biometrische Merkmale" erforderlich. Da bislang keine Implementierungen biometrischer Verfahren bekannt sind, die die Anforderungen des Signaturgesetzes (vgl. Anlage 1 SigV) nachweislich erfüllen, werden für qualifizierte elektronische Signaturen in der Praxis immer Personal Identification Numbers (PIN) als Identifikationsdaten eingesetzt.

Unter der Voraussetzung, dass der öffentliche Schlüssel eindeutig und zuverlässig einer Person zugeordnet werden kann, bezeugt die Signatur folglich die Identität des Unterzeichners. Da die Signatur zudem das Ergebnis einer Verschlüsselungsoperation ist, sind die signierten Daten nachträglich auch nicht mehr veränderbar bzw. eine Änderung ist sofort erkennbar. Die Signatur kann auch nicht unautorisiert weiter verwendet werden, weil das Ergebnis der Verschlüsselungsoperation natürlich abhängig von den Daten ist. Geht man ferner davon aus, dass der private Signaturschlüssel nicht kompromittiert worden ist, kann der Absender der Daten die Urheberschaft auch nicht mehr zurückweisen, weil ausschließlich er selbst über den privaten Signaturschlüssel verfügt.

Technisch wäre natürlich eine Verschlüsselung der gesamten Daten (eines Dokuments oder einer Nachricht) viel zu aufwändig. Aus diesem Grunde wird aus den Daten eine eindeutige Prüfsumme, ein Hashwert (s. dazu auch Kap. 5.3.1) erzeugt, dieser verschlüsselt („unterschrieben") und den Originaldaten beigefügt. Der mit dem geheimen Schlüssel verschlüsselte Hashwert repräsentiert fortan die elektronische Signatur („Unterschrift") der Originaldaten. Der Empfänger seinerseits bildet nach demselben Verfahren, d.h. mit demselben Hash-Algorithmus ebenfalls eine Prüfsumme aus den erhaltenen Daten und vergleicht sie mit der des Absenders. Sind die beiden Prüfsummen identisch, dann sind die Daten unverändert und stammen zuverlässig vom Inhaber des geheimen Schlüssels, denn nur er war in der Lage die Prüfsumme so zu verschlüsseln, dass sie mit dem zugehörigen öffentlichen Schlüssel auch entschlüsselt werden konnte.

Die Hinzufügung der Signaturdaten zu den Originaldaten kann grundsätzlich auf folgende Weise geschehen (s. Abbildung 6):

- Enveloped („eingebettet"): die Signaturdaten sind als Element in den Originaldaten enthalten.
  Dieses Verfahren, auch als so genannte „Inbound-Signatur" bezeichnet, wird vor allem bei der Signatur von PDF-Dokumenten und PDF-Formularen bspw. im Projekt ArchiSafe der Physikalisch-Technischen Bundesanstalt benutzt (s. a. Abb. 7).[29] Dabei werden die binären Signaturdaten direkt in das PDF-Dokument eingebettet und gemeinsam mit den Originaldaten im PDF-Format angezeigt. Mit dem neuen Adobe® Reader® (Version 8) ist der Empfänger der signierten Daten darüber hinaus imstande, unmittelbar eine Überprüfung der Integrität der angezeigten und signierten Daten vorzunehmen.
  Eingebettete Signaturen werden ebenso bei der Signatur von XML-Da-

29 s. http://www.archisafe.de

ten[30] verwendet und sollen zudem nun auch für den neuen XDOMEA Standard 2.0[31] spezifiziert werden. Da die Signatur eine binäre Zahlenfolge ist, lässt sie sich jedoch nicht direkt in ein XML-Dokument einbetten. Man codiert daher die binären Werte im Base64-Format (RFC 1521), um aus ihnen ASCII-lesbare Zeichen zu gewinnen. Die erhaltene Zeichendarstellung der Signatur findet sich schliesslich als SignatureValue in der XML-Signatur wieder[32].

- **Enveloping („umschließend"):** die Signaturdaten „umschließen" die Originaldaten. Diese Methode wird hauptsächlich für die Signatur von E-Mail Nachrichten oder reinen XML-Daten benutzt. Eine S/MIME Client-Anwendung, wie bspw. Microsoft Outlook, bettet in diesem Fall die Nachricht in einen signierten „Umschlag" ein.
- **Detached („getrennt"):** die Signaturdaten befinden sich außerhalb der Originaldaten in einer zusätzlichen, binären Signaturdatei. Diese Form, auch als „Outbound-Signatur" bezeichnet, wird standardmäßig für XML-Signaturen sowie die Signatur binärer Originaldaten eingesetzt. Ein separater Link in den Original-Daten oder zusätzlichen Beschreibungsdaten sorgt dann für die notwendige permanente Verknüpfung der Originaldaten mit den Signaturdaten.

---

30 1999 bis 2002 wurde der W3C-Standard für das Signieren von XML-Dokumenten am Massachusetts Institute of Technology (MIT) entwickelt (XMLDSIG). Die XML Signatur Spezifikation (auch XMLDSig) definiert eine XML Syntax für digitale Signaturen.
In ihrer Funktion ähnelt sie dem PKCS#7 Standard, ist aber leichter zu erweitern und auf das Signieren von XML Dokumenten spezialisiert. Sie findet Einsatz in vielen weiterführenden Web-Standards wie etwa SOAP, SAML oder dem deutschen OSCI.
Mit XML Signaturen können Daten jeden Typs signiert werden. Dabei kann die XMLSignatur Bestandteil des XML Datenpakets sein (enveloped signature), die Daten können aber auch in die XML-Signatur selbst eingebettet sein (enveloping signature) oder mit einer URL adressiert werden (detached signature). Einer XML-Signatur ist immer mindestens eine Ressource zugeordnet, das heisst ein XML-Baum oder beliebige Binärdaten, auf die ein XML-Link verweist. Beim XML-Baum muss sichergestellt sein, dass es zu keinen Mehrdeutigkeiten kommt (zum Beispiel bezüglich der Reihenfolge der Attribute oder des verwendeten Zeichensatzes). Um dies erreichen zu können, ist eine so genannte Kanonisierung des Inhalts erforderlich. Dabei werden nach Maßgabe des Standards alle Elemente in der Reihenfolge ihres Auftretens aneinander gereiht und alle Attribute alphabetisch geordnet, so dass sich ein längerer UTF8-String ergibt (es gibt auch Methoden, die einen UTF16-String erzeugen). Aus diesem wird der eigentliche Hash-Wert gebildet beziehungsweise erzeugt man durch verschlüsseln den Signaturcode. So ist man wieder beim Standard-Verfahren für elektronische Signaturen (RFC 2437).

31 s. http://www.kbst.bund.de

32 Im Rahmen der Struktur eines XML-Dokuments lassen sich Subelemente explizit vom Signieren ausschliessen, so auch die Signatur selbst. Umgekehrt lassen sich beliebig viele Referenzen auflisten, die gemeinsam als Gesamtheit zu signieren sind.

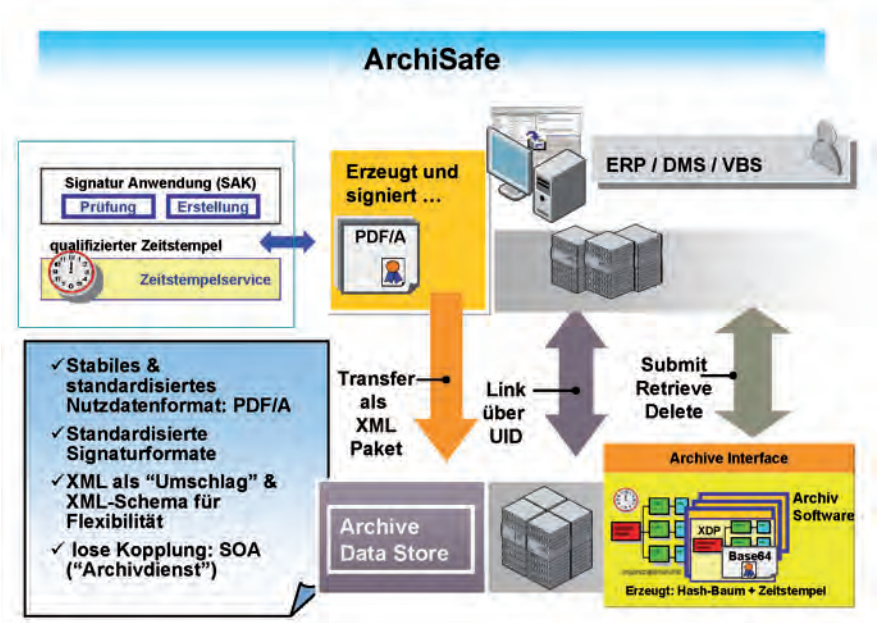

Abbildung 7: ArchiSafe – Rechts- und revisionssichere Langzeitspeicherung elektronischer Dokumente

Die Flexibilität der Hinzufügung von Signaturdaten zu Originaldaten basiert auf der als RFC 3852 – Cryptographic Message Syntax (CMS) im Juli 2004[33] durch die Internet Engineering Task Force (IETF) veröffentlichten Spezifikation sowie dem ursprünglich durch die RSA Laboratories veröffentlichten PKCS#7 (Public Key Cryptography Standard) Dokument in der Version 1.5. In beiden Dokumenten wird eine allgemeine Syntax beschrieben, nach der Daten durch kryptographische Maßnahmen wie digitale Signaturen oder Verschlüsselung geschützt, respektive Signaturdaten über das Internet ausgetauscht werden können. Die Syntax ist rekursiv, so dass Daten und Umschläge verschachtelt oder bereits chiffrierte Daten unterschrieben werden können. Die Syntax ermöglicht zudem, dass weitere Attribute wie z.B. Zeitstempel mit den Daten oder dem Nachrichteninhalt authentifiziert werden können und unterstützt eine Vielzahl von Architekturen für die Schlüsselverwaltung auf der Basis von elektronischen Zertifikaten.

33 Network Working Group (2004)

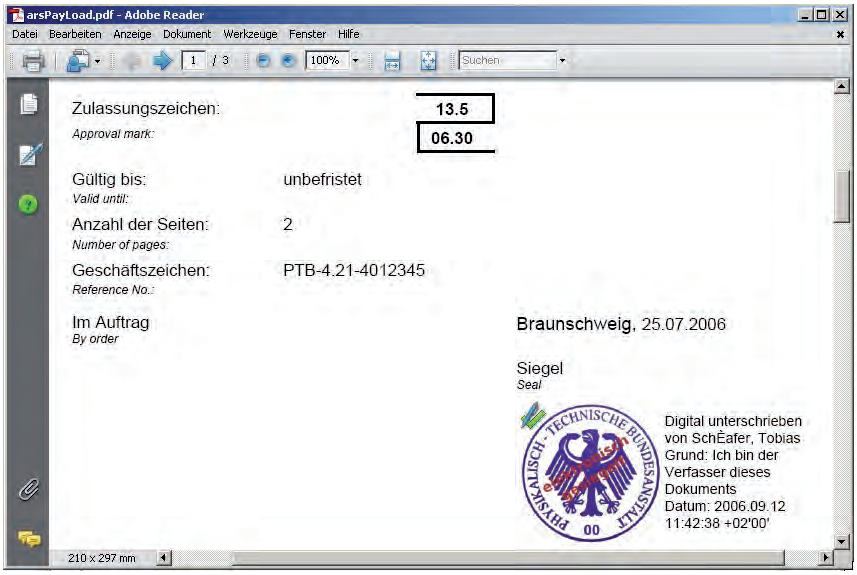

Abbildung 8: Digitale PDF-Signatur

Literatur

Eckert, Claudia (2001): *IT-Sicherheit: Konzepte – Verfahren – Protokolle*. München: Oldenbourg Wissenschaftsverlag

Network Working Group (2004): R. Hously: *Cryptographic Message Syntax (CMS) Request for Comment*s: 3852 http://www.ietf.org/rfc/rfc3852

Schneier, Bruce (1996): *Angewandte Kryptographie*, Bonn u.a.: Addison-Wesley Verl.

Schneier, Bruce (2005): *Schneier on Security. A blog covering security and security technology*. SHA-1 Broken (February 15, 2005) http://www.schneier.com/blog/archives/2005/02/sha1_broken.html

5.4 Kriterienkataloge für vertrauenswürdige digitale Langzeitarchive

Susanne Dobratz und Astrid Schoger

Verschiedene Organisationen und Initiativen (vgl. Kapitel 5.1) haben Anforderungen an vertrauenswürdige digitale Langzeitarchive formuliert. Diese Kriterien betreffen sowohl organisatorische als auch technische Aspekte, die erfüllt werden müssen, um der Aufgabe der Informationserhaltung gerecht werden zu können. Die Grundprinzipien der Kriterienkataloge sowie deren Inhalte werden nun in Kapitel 5.3 anhand des nestor-Kriterienkataloges genauer erörtert.

Grundprinzipien der Kriterienkataloge

Bei der Herleitung und für die Anwendung von Kriterien der Vertrauenswürdigkeit gelten folgende Grundprinzipien, die erstmals von der nestor Arbeitsgruppe „Vertrauenswürdige Archive – Zertifizierung" formuliert wurden:

Abstraktion: Ziel ist es, Kriterien zu formulieren, die für ein breites Spektrum digitaler Langzeitarchive angewendet werden können und über längere Zeit Gültigkeit behalten. Deshalb wird von relativ abstrakten Kriterien ausgegangen.

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 es eine angemessene Dokumentation aller Stufen, die Schlüssigkeit eines digitalen Langzeitarchivs umfassend zu bewerten. Auch alle Qualitäts- und Sicherheitsnormen fordern eine angemessene Dokumentation.

Transparenz: Transparenz wird realisiert durch die Veröffentlichung geeigneter Teile der Dokumentation. 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. 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.

Abbildung 9: Gliederung der Kriterien für vertrauenswürdige digitale Langzeitarchive

Bewertbarkeit: Die Kriterien decken die Aspekte ab, die im Rahmen einer Prüfung der Vertrauenswürdigkeit zu betrachten sind. Sie sind zwar abstrakt formuliert, verlangen aber vom betrachteten digitalen Archiv detaillierte Antworten. Aufgrund dieser Antworten ist eine Bewertung der durch das Archiv eingeleiteten Methoden und Maßnahmen zur Sicherung der digitalen Objekte möglich. Allerdings 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. Nicht objektiv bewertbar ist das Kriterium „Die Organisationsform ist angemessen“.

Hingegen ist zum Beispiel das Kriterium „Die Objekte sind eindeutig und dauerhaft identifiziert“ bewertbar, da es dazu bereits Standards gibt.

Formaler Aufbau der Kriterienkataloge

Jedem abstrakt formulierten Kriterium ist eine Erläuterung beigegeben, die der Verständlichkeit dient. Beispiele aus unterschiedlichen Anwendungsbereichen sowie spezielle Literaturhinweise tragen zur Anwendbarkeit bei.

Inhaltlicher Aufbau der Kriterienkataloge

Die eigentlichen Kriterien sind in drei Bereiche gegliedert:

A. **Ziele und Organisation (Organisatorischer Rahmen):** Hier werden Anforderungen an den organisatorischen Rahmen gestellt, innerhalb dessen das digitale Langzeitarchiv operiert, d.h. seine Zielsetzung, die rechtlichen Bedingungen ebenso wie die personellen und finanziellen Ressourcen sowie die Organisationsform.

B. **Umgang mit Objekten:** Hier finden sich die Kriterien, die alle objektbezogenen Anforderungen während des Lebenszyklus der Objekte im digitalen Langzeitarchiv spezifizieren. Ausgangspunkt sind die im OAIS-Referenzmodell definierten zentralen Aufgaben: Aufnahme (Ingest), Archivablage (Archival Storage, inklusive Umsetzung der Langzeiterhaltungsmaßnahmen) und Nutzung (Access), die unterstützenden Prozesse Datenmanagement und Planung der Langzeiterhaltungsmaßnahmen (Preservation Planning).

C. **Infrastruktur und Sicherheit**: Die Kriterien in diesem Abschnitt betrachten die technischen Aspekte des Gesamtsystems sowie die Aspekte der Sicherheit.

A - Ziele und Organisation

Um die eingesetzten Verfahren bewerten zu können, ist es notwendig, dass die Betreiberorganisation die Ziele und Rahmenbedingungen für den Betrieb des digitalen Langzeitarchivs spezifiziert, dokumentiert und veröffentlicht hat. Welche Objekte werden gesammelt? Für welche Nutzergruppe wird gesammelt? In welcher Form und unter welchen Konditionen sollen die digitalen Objekte den Nutzern bereitgestellt werden? Der Archivbetreiber muss weiterhin darlegen, auf welcher rechtlichen Basis er operiert, er muss entsprechende rechtliche Regelungen mit seinen Produzenten vereinbaren. Die Organisationsform der Betreiberorganisation und des digitalen Langzeitarchivs muss angemessen sein,

d.h. sie muss für die Erfüllung der Aufgabe geeignet sein und geeignete Entscheidungsstrukturen besitzen. Sowohl eine strategische als auch eine operative Planung sind vorzuweisen, ebenso Pläne zur Reaktion auf substantielle Veränderungen. So ist der Betrieb des digitalen Langzeitarchivs auch im Falle einer Auflösung der Betreiberorganisation sicherzustellen. Der Einsatz entsprechend qualifizierten Personals ist nachzuweisen. Ein weiterer wichtiger Aspekt ist die Durchführung eines Qualitätsmanagements.

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

Der Informationserhalt im Hinblick auf die Nutzung wird als zentrale Aufgabe der Langzeitarchivierung definiert, die in Abhängigkeit von der Zielgruppe und deren Bedürfnissen geleistet werden muss. Nutzung setzt zum einen den Erhalt der Integrität, Authentizität und Vertraulichkeit sowie die Verfügbarkeit der digitalen Objekte voraus, zum anderen die Sicherstellung der Interpretierbarkeit der digitalen Objekte durch die Zielgruppe, um die darin enthaltene Information in einer geeigneten Weise zu rekonstruieren.

Daher werden in diesem Abschnitt zunächst die Anforderungen definiert, die sich auf die Erhaltung der Integrität und Authentizität der digitalen Objekte auf allen Stufen der Verarbeitung konzentrieren. Die Verarbeitungsstufen sind die im OAIS-Modell abgebildeten: Aufnahme, Archivablage und Nutzung sowie zur Sicherung der Interpretierbarkeit die Durchführung der technischen Langzeiterhaltungsmaßnahmen.[34] Die Übernahme der digitalen Objekte von den Produzenten erfolgt nach definierten Vorgaben. Für die Objekte muss spezifiziert sein, wie die Übergabepakete (Submission Information Packages, SIPs), die Archivpakete (Archival Information Packages, AIPs) und die Nutzungspakete (Dissemination Information Packages, DIPs) aussehen. Es muss eine Spezifikation geben, wie die Transformation der Informationspakete untereinander aussieht. Dazu sind die erhaltenswerten Kerneigenschaften zu spezifizieren. Das digitale Langzeitarchiv muss die technische Kontrolle über die digitalen Objekte besitzen, um diese Transformationen sowie speziell die Langzeiterhaltungsmaßnahmen durchführen zu können.

Das Datenmanagement muss dazu geeignet sein, die notwendigen Funktionalitäten des digitalen Langzeitarchivs zu gewährleisten. Die eindeutige und dauerhafte Identifikation[35] der Objekte und deren Beziehungen untereinander ist essentiell für deren Auffindbarkeit und Rekonstruierbarkeit. Das digitale Langzeitarchiv muss in ausreichendem Maße Metadaten[36] für eine formale, inhaltliche, strukturelle sowie für eine technische Beschreibung der digitalen Objekte erheben. Zudem sind Metadaten notwendig, die alle vom digitalen Langzeitarchiv vorgenommenen Veränderungen an den digitalen Objekten beinhalten. Entsprechende Nutzungsrechte und –bedingungen sind ebenfalls in Meta-

34 vgl. dazu auch Kapitel 8 dieses Handbuchs
35 vgl. dazu auch Kapitel 9.4 dieses Handbuchs
36 vgl. dazu auch Kapitel 6.2 dieses Handbuchs

daten zu verzeichnen.

Die nestor-Kriterien lauten im Detail:
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 Lang-zeitarchivierungsmaßnahmen 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

Ein digitales Langzeitarchiv muss eine angemessene IT-Infrastruktur besitzen, die in der Lage ist, die Objekte wie in Abschnitt B – Umgang mit Objekten zu bearbeiten. Es muss ein Sicherheitskonzept existieren und umgesetzt werden, sodass die Infrastruktur den Schutz des digitalen Langzeitarchivs und seiner digitalen Objekte gewährleisten kann.

Die nestor-Kriterien im Detail:
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.

Der Weg zum vertrauenswürdigen digitalen Langzeitarchiv

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 Implementierung
4. Evaluierung

Diese Stufen sind nicht als starres Phasenmodell zu betrachten. Vielmehr müssen sie im Zuge der ständigen Verbesserung regelmäßig wiederholt werden. Das Qualitätsmanagement überwacht diesen Entwicklungsprozess.

Die Kriterienkataloge und Checklisten sowie das Tool DRAMBORA können auf allen Stufen der Entwicklung zur Orientierung und Selbstevaluierung eingesetzt werden.

Darüber hinaus wurden auf der Grundlage von TRAC bereits mehrere externe Audits durchgeführt und DRAMBORA wurde bereits in einigen zum Teil von externen Experten begleiteten Selbstevaluierungen eingesetzt.