

nestor

# Wege ins Archiv

## Ein Leitfaden für die Informationsübernahme in das digitale Langzeitarchiv

Version 1  
(Entwurf zur öffentlichen Kommentierung)

herausgegeben von der  
nestor-Arbeitsgruppe

Standards für Metadaten, Transfer von  
Objekten in digitale Langzeitarchive und  
Objektzugriff



**Wege ins Archiv: Ein Leitfaden für die Informationsübernahme in das digitale Langzeitarchiv, Version 1 (Entwurf zur öffentlichen Kommentierung), herausgegeben von der nestor-Arbeitsgruppe "Standards für Metadaten, Transfer von Objekten in digitale Langzeitarchive und Objektzugriff"**

**AG Standards für Metadaten, Transfer von Objekten in digitale Langzeitarchive und Objektzugriff :**

c/o Niedersächsische Staats und Universitätsbibliothek Göttingen

**Jens Ludwig**

Papendiek 14

D-37073 Göttingen

Tel.: +49 (0)551-3912121

E-Mail: [<ludwig@sub.uni-goettingen.de>](mailto:ludwig@sub.uni-goettingen.de)

## **ODER**

C/o Bundesarchiv

**Karsten Huth**

Referat B 1a

Potsdamer Straße 1

D-56075 Koblenz

Tel.: +49 (0)261 - 505 466

E-Mail: [<k.huth@barch.bund.de>](mailto:k.huth@barch.bund.de)

## **Autoren des Ratgebers**

**Beinert, Tobias: Bayerische Staatsbibliothek**

**Büchler, Georg: KOST, Schweiz**

**Dr. Graf, Sabine: Niedersächsisches Landesarchiv**

**Huth, Karsten: Bundesarchiv**

**Dr. Keitel, Christian: Landesarchiv Baden-Württemberg**

**Ludwig, Jens: Niedersächsische Staats und Universitätsbibliothek**

**Rödig, Peter: Universität der Bundeswehr München**

**Steinke, Tobias: Deutsche Nationalbibliothek**

## **Weitere Anregungen verdanken wir:**

**Olaf Brandt: SUB Göttingen/KB Niederlande jetzt BStU**

**Enders, Markus: SUB Göttingen jetzt British Library**

**Dr. Keller-Marxer: iKeep AG**

**Dr. Korb, Nikola: Deutsche Nationalbibliothek**

**Schrimpf, Sabine: Deutsche Nationalbibliothek**

**Dr. Wolf-Klostermann: Bayerische Staatsbibliothek**

<b>EINLEITUNG.....</b>	<b>4</b>
<b>1 OBJEKTE.....</b>	<b>6</b>
1.1 Auswahl der zu archivierenden Informationen.....	6
1.2 Auswahl der Metadaten.....	8
1.3 Identifikation von wesentlichen Eigenschaften der Intellektuellen Entitäten und der Informationsobjekte.....	10
<b>2 PROZESSE.....</b>	<b>12</b>
2.1 Definition von Transferpaketen.....	12
2.2 Validierung.....	14
2.3 Transfer von Daten aus dem Produzentensystem.....	16
<b>3 MANAGEMENT.....</b>	<b>18</b>
3.1 Identifikation von rechtlichen und vertraglichen Bedingungen.....	18
3.2 Übernahmevereinbarungen und Dokumentation .....	20
3.3 Qualitäts-, Sicherheits-, Prozess- und Risikomanagement .....	22
<b>GLOSSAR.....</b>	<b>26</b>
<b>LITERATUR.....</b>	<b>28</b>
<b>ÜBERSICHT ÜBER DIE ZIELE DER INFORMATIONENÜBERNAHME.....</b>	<b>30</b>

# Einleitung

Die Langzeitarchivierung digitaler Daten ist ein noch jüngeres Themenfeld als die digitalen Daten selbst und gerade erst gehen verschiedene Projekte in diesem Bereich in den Regelbetrieb über. Wie es für junge Aufgabenbereiche üblich ist, haben sich erst wenige Standards und Verfahren etabliert und von "Best Practice" im eigentlichen Sinne kann man kaum sprechen, denn praktiziert wird die Langzeitarchivierung bisher nur von wenigen Institutionen. Dieser Leitfaden möchte dabei helfen, das erste Hindernis bei der Langzeitarchivierung leichter zu bewältigen: Wie können digitale Informationsobjekte in ein Langzeitarchiv überführt werden, um dort schließlich langfristig sicher gespeichert, verwaltet und bewahrt zu werden?

Die Informationsübernahme in ein digitales Langzeitarchiv ist nicht nur ein technischer Transfer zwischen zwei Systemen, sondern sie ist insbesondere ein Prozess mit vielen organisatorischen Anforderungen, an dessen Ende die Übernahme der Verantwortung durch das digitale Langzeitarchiv steht. Ein Langzeitarchiv ist dabei nicht primär als ein technisches System zu verstehen, sondern als eine organisatorische Einheit oder Institution, die die Aufgabe der Langzeitarchivierung übernimmt und sich idealerweise am OAIS-Referenzmodell orientiert. Viele Anforderungen, die ein Langzeitarchiv auch bei der Informationsübernahme erfüllen muss, sind u.a. im nestor Kriterienkatalog bereits formuliert worden. "Wege ins Archiv" soll helfen, die Ziele und Besonderheiten einer Informationsübernahme zur Langzeitarchivierung deutlich zu machen: Daten müssen aus üblicherweise heterogenen technischen und organisatorischen Kontexten so übernommen werden, das sie trotzdem in ganz anderen, zukünftigen Kontexten verstehbar und nutzbar sein werden.

Die Informationsübernahme stellt aber nicht nur Anforderungen an das Langzeitarchiv, sondern auch an den Produzenten bzw. Lieferanten (wobei es sich dabei auch um andere Langzeitarchive handeln kann) der zu bewahrenden digitalen Objekte. Die besondere Aufwändigkeit, Komplexität und Relevanz rührt zu einem guten Teil daher, dass das Langzeitarchiv und der Produzent zusammen die späteren Nutzungsszenarien und Prozesse schon berücksichtigen müssen: Dürfen aus rechtlichen Gründen überhaupt Veränderungen an den Materialien durchgeführt werden? Welche Eigenschaften müssen an den Materialien definitiv erhalten werden? Und nicht zuletzt müssen die technischen und anderen Qualitäten der Materialien besonders geprüft werden, denn sie bilden die oftmals nicht mehr korrigierbare Grundlage für alle weiteren Maßnahmen.

Ausgangspunkt für die Erarbeitung des Leitfadens war neben der Erfahrung der Beteiligten das OAIS-Referenzmodell und der ergänzende Standard PAIMAS. Diese behandeln die Informationsübernahme, wie sie hier besprochen wird, allerdings entweder primär als Teil eines Langzeitarchivs ("Ingest") oder auf einem Abstraktionsniveau, das als Einstieg wahrscheinlich wenig hilfreich ist. Die Zielgruppen, für die der Leitfaden entworfen wurde, sind:

- Gedächtnisinstitutionen, die gemeinsam mit Informationsproduzenten bzw. den informationsabgebenden Institutionen eine Übernahme realisieren wollen und zum Einstieg eine gemeinsame Basis benötigen.

- Informationsproduzenten bzw. informationsabgebende Institutionen, die ihre Informationsobjekte zur langfristigen Erhaltung an ein digitales Langzeitarchiv übergeben wollen und Informationen über die bevorstehenden Aufgaben suchen.
- Gedächtnisinstitutionen, die über eine Infrastruktur zur Langzeitarchivierung verfügen und zum ersten Mal vor einer Informationsübernahme stehen.
- Gedächtnisinstitutionen, die eine Infrastruktur zur Langzeitarchivierung planen und sich in diesem Rahmen mit der Aufgabe der Informationsübernahme befassen.

# 1 Objekte

## 1.1 Auswahl der zu archivierenden Informationen

Das Digitale Langzeitarchiv muss eine Auswahl der zu übernehmenden Intellektuellen Entitäten (s. Glossar) treffen. Die fachliche, inhaltliche Auswahl der zu archivierenden Information ist bei öffentlichen Institutionen i. d. R. aus dem Gesamtauftrag der Institution abzuleiten. Die notwendigen technischen Entscheidungen bei der Auswahl muss das digitale Langzeitarchiv in Zusammenarbeit mit dem Produzenten treffen. Hierbei geht es vor allem um die Auswahl und Vereinbarung von für die Langzeitarchivierung geeigneten Dateiformaten, die zur Darstellung der Informationsobjekte im System des digitalen Archivs benötigt werden, sowie um die Auswahl der notwendigen Metadatenformate.

In vielen Fällen sind die standardmäßig gebotenen Exportfunktionen der Produzentensysteme nicht in der Lage, die Informationsobjekte in der vom digitalen Langzeitarchiv gewünschten Form zu exportieren. Das Produzentensystem muss dann von seinem Lieferanten um die zusätzlichen Anforderungen erweitert werden, was oftmals sehr kostspielig ist. Der gesamte Übernahmeprozess ist für Produzent und Archiv einfacher und kostengünstiger, wenn taugliche Exportfunktionen bzw. Archivschnittstellen bereits bei der Planung und Beschaffung von elektronischen Systemen berücksichtigt und implementiert werden. Das digitale Langzeitarchiv sollte daher frühzeitig vom Produzenten über die Anschaffung eines neuen EDV-Systems informiert und in dessen Konzeption mit einbezogen werden.

---

### Ziel:

*Das digitale Langzeitarchiv wählt aus dem Material des Produzenten die Intellektuellen Entitäten aus, welche als Archivgut dauerhaft, authentisch und sicher übernommen werden sollen. Diese Auswahl erfolgt nach bestimmten Bewertungskriterien, die sich aus dem gesetzlichen oder vertraglich festgelegten Auftrag eines digitalen Langzeitarchivs ergeben.*

### Verfahren

Digitale Langzeitarchive sammeln für den Menschen les- und interpretierbare Informationen, die in digitaler Form gespeichert werden. Auf dieser sachlich inhaltlichen Ebene ist es zulässig, mit Begriffen zu arbeiten, die aus der alltäglichen Erfahrung im Umgang mit Informationsobjekten stammen (z.B. Dokument, Akte, Film, Foto, Datenbank des Statistischen Bundesamts usw.). Natürlich hat die Festlegung des primären Archivguts Konsequenzen auf der technischen Ebene.

Ein Beispiel für eine standardisierte Auswahl von Intellektuellen Entitäten ist das DOMEA Aussonderungsmodul [Lit. 4], welches Wege zur geregelten Abgabe elektronischer Akten aus Behördensystemen an die zuständigen Archive beschreibt und zwei Verfahren zur Anbietung und Auswahl der abzugebenden Akten kennt. Im ersten Verfahren (4-Stufig) wird dem Archiv eine elektronische Anbieterliste übermittelt, auf der die infragekommenden Objekte mit den relevanten Metadaten verzeichnet sind. Anhand dieser Liste nimmt das Archiv die Bewertung der Archivwürdigkeit der Akten vor. Beim zweiten Verfahren hat das

Archiv einen direkten Zugriff auf die Listen, der zur Aussonderung anstehenden Akten und kann dort seine Bewertung und Auswahl vornehmen.

---

Ziel:

*Das digitale Langzeitarchiv und der Produzent analysieren die Möglichkeiten der Exportschnittstelle des Produzentensystems. Wenn die Daten nach einem Export nicht in der Form vorliegen, in der das digitale Langzeitarchiv diese dauerhaft verfügbar halten kann, werden entsprechende Maßnahmen (z.B. Konvertierungen) geplant.*

Verfahren:

Daten alleine stellen noch keine für den Menschen les- und interpretierbaren Informationsobjekte dar. Sie enthalten codierte Informationen, die von der passenden Hardware/Software-Umgebung korrekt interpretiert und dargestellt werden müssen. Die zur Darstellung notwendigen technischen Mittel können für bestimmte Informationsobjekte sehr aufwendig sein. Dies kann sich auf die Kosten zur Anschaffung der notwendigen technischen Mittel, die Aufrechterhaltung des technischen Betriebs und die zeitlich begrenzte Verfügbarkeit der technischen Mittel beziehen. Sollte das digitale Langzeitarchiv langfristig keine Möglichkeit sehen, die Darstellung der Informationsobjekte in der von der Exportschnittstelle angebotenen Form zu garantieren, müssen Archiv und Produzent über Alternativen bei den Formaten der Informationsobjekte verhandeln [Lit. 11].

---

Ziel:

*Es gibt eine Vereinbarung zwischen Archiv und Produzent über notwendige Anpassungen, wenn die Informationsobjekte in der vorliegenden Form nicht archiviert werden können.*

Verfahren:

Das digitale Langzeitarchiv und der Produzent können verpflichtet sein, die Informationsobjekte anzubieten, zu übernehmen und dauerhaft zu erhalten. Eine generelle Ablehnung der Übernahme bei fehlenden technischen Möglichkeiten des Archivs ist in diesem Fall keine Alternative.

Bei einer Migration werden die Daten verändert. Trotz aller Veränderung müssen jedoch die wesentlichen Eigenschaften der Intellektuellen Entität erhalten bleiben (siehe Abschnitt 1.3). Die Auswahl eines Datenformats, das in der Lage ist, die wesentlichen Eigenschaften aus dem Ausgangsformat zu erhalten, und gleichzeitig zur dauerhaften Erhaltung im Archiv geeignet ist, ist eine der wesentlichen kritischen Entscheidungen der digitalen Langzeitarchivierung. Das Zielformat sollte offen und gut dokumentiert sein. Es sollte mit wesentlich geringeren technischen Mitteln darstellbar sein als das Ausgangsformat [Lit. 11] [Lit 15].

Eine Migration der Daten in ein anderes Format ist ein zusätzlicher Aufwand, der Risiken birgt. Zwischen Produzent und Archiv muss geklärt sein, wer welche Verantwortungen und Lasten trägt. Die Wahl der technischen Mittel zur

Durchführung einer Migration kann beträchtliche Auswirkungen auf die Qualität der Informationsobjekte und die Kosten der Übernahme haben. Konvertierungsprogramme, die das gleiche Zielformat anstreben, müssen nicht zwangsläufig identische Ergebnisse produzieren. Sie können ganz unterschiedliche Techniken (z.B. Komprimierungsalgorithmen) verwenden, die trotzdem alle noch im Rahmen der Formatdefinition liegen, oder proprietäre Techniken verwenden, ohne dies vorher anzugeben. Neben der Migration ist auch die Emulation oder eine Ergänzung der Software/Hardware Umgebung des digitalen Langzeitarchivs eine mögliche Option zur Darstellung der übernommenen Informationsobjekte.

## **1.2 Auswahl der Metadaten**

Nach dem Standard OAIS ist der Ingest dann abgeschlossen, wenn ein vollständiges „Archival Information Package“ (AIP) gebildet wurde. Ein AIP enthält neben dem eigentlichen Inhaltsobjekt hauptsächlich Metadaten. Diese sind insgesamt notwendig um dauerhaft genügend Informationen über das Archivobjekt (AIP) zu haben, also es zu finden, darzustellen, im Kontext mit anderen Archivobjekten auch zu verstehen und zu interpretieren.

Für die Langzeitarchivierung wesentlich sind Informationen, die helfen, aus Informationsobjekten und ihren Daten für den Menschen interpretierbare intellektuelle Entitäten (z.B. Zeitschriftenartikel, Akten, Fotos usw.) herzustellen. Diese Informationen sollten Hinweise auf die notwendigen technischen Umgebungen zur Darstellung enthalten, sollten das Datenformat möglichst eindeutig identifizieren (z.B. Dateiformatname und Version) und zumindest eine Referenz auf eine ausführliche technische Beschreibung des Datenformats (z.B. ISO-Norm, RFC, Dateiformatregistratur) enthalten. Bei komplexen Informationsobjekten bestehend aus mehreren Dateien muss die Struktur des Informationsobjekts nachvollziehbar beschrieben werden. Hinzu kommen Metadaten, die zur technischen Verwaltung der Daten im Speicher notwendig sind, wie Dateinamen, Dateigrößen, Hashwerte zur Validierung der archivierten Daten.

Des weiteren werden Metadaten benötigt, die den Inhalt eines Archivobjekts (AIP) beschreiben (z.B. Autor, Titel bei Publikationen), und Metadaten, welche die Herkunft eines Archivobjekts beschreiben (z.B. die Angaben, welche Behörde wann und durch wen eine elektronische Akte an das Archiv abgegeben hat). Zudem sind für eine verlässliche Langzeitarchivierung Angaben zu den Veränderungen an den Informationsobjekten wichtig, die beim Export aus dem Produzentensystem vorgenommen wurden. Wie detailliert und tiefgehend diese Angaben in den Metadaten verzeichnet werden, hängt von den Ansprüchen des Archivs, des Produzenten und der Art der übernommenen Informationsobjekte ab.

Nicht alle der oben beschriebenen Typen von Metadaten sind bereits nach dem Export eines Informationsobjekts aus einem Produzentensystem vorhanden. Einige Metadaten ergeben sich erst während der Übernahme ins Archiv, einige müssen vom Produzenten erfragt werden. Das Archiv muss deshalb vor der Konzeption eines Transferpakets klären, welche Angaben es aus dem Produ-



zentensystem und vom Produzenten benötigt.

---

Ziel:

*Das digitale Langzeitarchiv und der Produzent definieren alle Metadaten, die zu den ausgewählten Informationsobjekten im AIP benötigt werden, Das Produkt der Bemühungen ist eine Auswahl aller Angaben, die zum Erstellen eines ausreichend ausführlichen AIPs nötig sind.*

Verfahren

Zur langfristigen, authentischen Darstellung der Informationsobjekte sind Informationen zu den notwendigen technischen Mitteln und den strukturellen Zusammenhängen der Dateien des Informationsobjekts notwendig. Informationen zur Beschreibung des semantischen Inhalts und Kontextes des Informationsobjektes kommen hinzu. Es sind eine Reihe von Metadatenarten zu erwägen, die vom Produzenten auch nach eigenem Bedarf mitgeliefert werden. Neben den üblichen inhaltsbeschreibenden und den bereits erwähnten strukturellen Metadaten können dies auch technische Metadaten und Verwaltungsmetadaten sein.

Metadatenformate können nach inhaltlichen Kriterien unterschieden werden, z. B. inhaltsbeschreibende Metadaten (im Bibliotheksbereich bibliografische Metadaten), technische Metadaten und rechtliche Metadaten. Für inhaltsbeschreibende Metadaten gibt es spezifische Standards wie MAB2 [Lit. 5] und MARC21 [Lit. 6] für den Bibliothekssektor und übergreifende Standards wie Dublin Core [Lit. 7]. Da Dublin Core in Profilen erweitert werden kann, müssen sich die Partner verständigen, ob sie nur Dublin Core Simple oder ein spezifisches Profil benutzen. Für technische und strukturelle Metadaten gibt es PREMIS [Lit 3], METS [Lit. 8], LMER [Lit. 9].

Im OAIS-Modell [Lit 1] sind für Verwaltungsmetadaten vorgesehen:

- Provenance: Wer hat wann was gemacht? Die Geschichte eines Objekts.
- Context: Paketexterne Beziehungen des Inhalts. Wieso wurde er produziert, wie steht er in Beziehung zu anderen Inhalten und Paketen?
- Reference: Identifier: Numerische oder alphanumerische Zeichenketten, die eindeutig sowohl die Intellektuellen Entitäten, wie auch die dazugehörigen Informationsobjekte referenzieren und im Archiv identifizieren [Lit. 10].
- Fixity: Schutz vor Veränderungen, z.B. Checksummen (siehe auch Kapitel Validierung)

---

Ziel:

*Das digitale Langzeitarchiv und der Produzent treffen eine Vereinbarung, von wem welche der benötigten Metadaten bereit gestellt werden.*

Verfahren

Nicht alle notwendigen Metadaten müssen vom Produzenten bereitgestellt und somit an das digitale Langzeitarchiv übermittelt werden. Die Beschreibung der technischen Mittel zur Darstellung eines Informationsobjekts im Archiv ist eine

Information, die sinnvoller Weise vom Archiv bereitgestellt werden sollte. Wenn beispielsweise der Produzent die Informationsobjekte vor dem Transfer an das Archiv in das Archivformat PDF/A migriert, entstehen Metadaten, welche die Migration beim Produzenten beschreiben und an das Archiv übermittelt werden sollten. Wenn das Archiv dann die eingegangenen PDF/A-Dateien auf ihre Standardkonformität hin überprüft, entstehen wieder Metadaten – diesmal im Archiv -, die in das AIP integriert werden können.

### **1.3 Identifikation von wesentlichen Eigenschaften der Intellektuellen Entitäten und der Informationsobjekte**

Um Intellektuelle Entitäten in digitaler Form über lange Zeiträume verfügbar zu halten, müssen die Informationsobjekte in wechselnden technischen Umgebungen dargestellt werden. Sowohl bei einer auf Emulation als auch bei einer auf Migration basierenden Erhaltungsstrategie werden sich Eigenschaften der Daten zwangsläufig ändern. Als wesentliche Eigenschaften (engl. significant properties) werden diejenigen Eigenschaften bezeichnet, die unter allen Umständen konstant bleiben müssen.

#### Ziel:

*Archiv und Produzent erarbeiten eine Definition der wesentlichen Eigenschaften der ausgewählten Informationsobjekte.*

#### Verfahren:

Das Archiv und der Produzent müssen entscheiden, welche Darstellungsabweichungen noch akzeptabel sind und welche nicht. Ein maßgebliches Kriterium sollten die Bedürfnisse der Nutzer des Archivs sein (Zielgruppe). Das InSPECT-Projekt führt als geläufige Unterteilungen der wesentlichen Eigenschaften an: "

- content, eg. text, image, slides, etc.
  - context, eg. who, when, why.
  - appearance, eg. font and size, colour, layout, etc.
  - structure, eg. embedded files, pagination, headings, etc.
  - behaviour, eg. hypertext links, updating calculations, active links, etc. "
- [Lit. 16]

---

#### Ziel:

*Das Archiv verzeichnet für jede Intellektuelle Entität, die es vom Produzenten übernimmt, entsprechende Eigenschaften, die auf Dauer erhalten bleiben sollen. Anhand dieser Angaben kann später überprüft werden, ob eine anstehende Migration oder ein neuer Emulator für die Langzeiterhaltung der Objekte geeignet ist.*

#### Verfahren:

In welcher Form die wesentlichen Eigenschaften eines Objekts verzeichnet werden, ist Sache des Archivs. Man kann harte technische Werte (z.B. Bildauflö-

sung, Bildbreite, Farbräume usw.) verwenden oder den sensorischen Eindruck einer Intellektuellen Entität beschreiben.

## 2 Prozesse

### 2.1 Definition von Transferpaketen

Nachdem das digitale Langzeitarchiv und der Produzent definiert haben, was archiviert werden soll, muss geklärt werden, in welchen Einheiten das zu Archivierende in das digitale Langzeitarchiv übertragen wird: Die Transferpakete müssen definiert werden. Durch sie wird festgelegt, in welcher Beziehung übertragene Daten und Metadaten der Informationsobjekte zueinander stehen, damit das Informationsobjekt aus den Teilen wiederherstellbar ist. Es kann sich bei den Transferpaketen um Containerformate handeln, die die Bestandteile enthalten, aber auch reine Beschreibungsdateien, die die Daten und Metadaten nur referenzieren und dadurch für das Archiv verfügbar machen, sind möglich. Ein Beispiel für ein Containertransferpaket für Webseiten könnte sein, dass alle Dateien und Metadaten einer Webseite zusammen mit einer beschreibenden XML-Datei in einer ZIP-Datei übermittelt werden.

Innerhalb des OAIS-Modells [Lit. 1] werden die hier ausgeführten Transferpakete als SIPs (Submission Information Packages) bezeichnet und von gegebenenfalls anders aufgebauten Ausgabepaketen (DIPs, Dissemination Information Packages) und intern verwendeten Paketen (AIPs, Archival Information Packages) des Archivs unterschieden. Transferpakete sind auch begrifflich von Paketformaten zu unterscheiden, die das technische Format der Transferpakete darstellen, und von Objektmodellen abzugrenzen, durch die Eigenschaften von Informationsobjekten logisch/konzeptionell abgebildet werden.

Wichtig ist die Festlegung des Transferpakets, weil digitale Langzeitarchive und Produzenten Informationsobjekte intern sehr wahrscheinlich nicht auf die gleiche Weise repräsentieren. Für den Transfer wird deshalb eine gemeinsame Definition dessen benötigt, was eine einzelne Einheit darstellt, die transferiert wird. Häufig haben digitale Langzeitarchive schon Vorgaben bezüglich der Transferpakete.

---

#### Ziel:

*Das Verhältnis zwischen einem Informationsobjekt und einem Paket bzw. mehreren Paketen ist festgelegt.*

#### Verfahren:

Meistens ist es ideal, dass ein Informationsobjekt genau ein einzelnes Paket bildet, weil damit die Komplexität reduziert wird. Die 1:1-Zuordnung kann aber aus technischen oder anderen Gründen nicht sinnvoll oder realisierbar sein, so dass ein Informationsobjekt aus mehreren Paketen besteht oder ein Paket mehrere Informationsobjekte enthält. Es müssen daher ein Verfahren und entsprechende Metadaten definiert werden, wie sich Pakete und Informationsobjekte zueinander verhalten.

- Aufgrund von technischen Größenbegrenzungen können große Informationsobjekte auf mehrere Pakete verteilt werden müssen.
- Eine Aufteilung eines Informationsobjektes auf mehrere Pakete kann

auch sinnvoll sein, wenn die eingelieferten Pakete den späteren Ausgabepaketen ähnlich sind und für relevante Nutzungsszenarien nur ein Zugriff auf einzelne Teile eines Informationsobjektes notwendig ist.

- Wenn bestimmte Daten zu sehr vielen Informationsobjekten gehören, kann aus Effizienzgründen die einmalige Übertragung eines Pakets sinnvoll sein, das von vielen Informationsobjekten und anderen Paketen referenziert wird. Formatvorlagendateien für große Webseiten oder Dokumentkollektionen können ein solcher Fall sein.

---

Ziel:

*Die Verknüpfungen der Daten sind anhand der im Paket enthaltenen Strukturmetadaten rekonstruierbar.*

Verfahren:

Daten sind sowohl untereinander als auch mit ihren Metadaten zunächst unverknüpft. Was eine relevante Verknüpfung darstellt, um ein Informationsobjekt zu bilden, ist abhängig von der technischen Umgebung und muss daher explizit benannt werden.

Dateien können in einem gemeinsamen Dateiverzeichnis liegen oder nach einem einheitlichen Schema benannt sein und damit eine technische und logische Beziehung zueinander ausdrücken. Soweit dies nicht durch Formate hinreichend definiert ist, muss ggf. die Zugehörigkeit von Metadaten zu Dokumentdateien oder Abhängigkeiten von Webseitendateien von einer Formatvorlagendatei beschrieben werden. METS [Lit. 8], LMER [Lit. 9] und PREMIS [Lit. 3] z.B. bieten dafür Möglichkeiten.

---

Ziel:

*Die Identifizierung des Pakets durch den Produzenten und das digitale Langzeitarchiv ist gewährleistet.*

Verfahren:

Die Identifizierung kann durch den Identifier des Informationsobjekts oder damit verknüpfte Archivpaketidentifier erfolgen. Es ist aber zu bedenken, dass es keine 1:1-Zuordnung zwischen den Paketen und den Informationsobjekten geben muss und deshalb auch für beide eigene Identifier notwendig sind. Der Identifier sollte sich im Paket befinden, technisch ist aber auch eine Identifizierung auf der Ebene des Transferprotokolls denkbar. Persistente Identifier [Lit. 10] sind zu bevorzugen.

Archivpakete können durch URNs der enthaltenen Dokumente identifiziert werden. Alle bekannten Metadatenstandards bieten dafür Möglichkeiten.

## 2.2 Validierung

Aufgrund der leichten Manipulierbarkeit digitaler Objekte sollte nach dem Transfer überprüft werden, ob diese noch das enthalten, was von ihnen erwartet wird. Auch in den anderen Phasen des Archivierungsprozesses erscheint die Validierung. Letztlich ist sie nach jeder Übertragung - an eine andere Einrichtung, in ein neues Format, auf einen neuen Datenträger - notwendig.

Bei der Validierung handelt es sich stets um einen Vergleich. Er soll die Authentizität des Objekts (d.h. das Objekt entspricht dem, was es vorgibt zu sein) und seine Funktionsfähigkeit belegen. Von dem zu validierenden Objekt ausgehend lassen sich zwei Klassen von Vergleichszielen unterscheiden:

- Das zu validierende Objekt wird anhand seines „Eltern-Objekts“ überprüft (Bsp.: Nach einer Datenträgermigration werden die Hashwerte der Zielfeile und der Ausgangsfeile verglichen).
- Das zu validierende Objekt wird auf formale oder inhaltliche Vorgaben überprüft (Bsp.: Ein Dateiformat wird mit der Beschreibung für dieses Format verglichen).

Anstelle eines einzelnen Objekts kann auch eine ganze Gruppe von Objekten überprüft werden. Aus Gründen der Lesbarkeit wird aber im Folgenden verallgemeinernd nur von "Objekt" gesprochen.

---

### Ziel:

*Definition der einzelnen Validierungsprozesse.*

### Verfahren:

Die Aufgabe der Validierung kann in einzelne Schritte bzw. Prozesse unterteilt werden. Ein Prozess überprüft dann die Einhaltung einer oder mehrerer Eigenschaften des übernommenen Objekts. Er kann automatisiert oder manuell erfolgen. Die Prozesse müssen inhaltlich beschrieben, gegenseitig abgegrenzt und benannt werden. Z.B.:

- Enthält die Lieferung alle vereinbarten Objekte?
- Sind die Objekte integer (entsprechen sie zuvor ermittelten Hashwerten)?
- Sind die Objekte frei von Viren?
- Sind die Dateien fehlerfrei/valide in Bezug auf ihr Dateiformat?

---

### Ziel:

*Das Archiv definiert für jeden einzelnen Validierungsprozess den notwendigen Erfüllungsgrad und die Konsequenzen bei einer Nichterfüllung.*

### Verfahren:

Manche Validierungsprozesse können als Ergebnis nur die Erfüllung bzw. Nichterfüllung einer Eigenschaft erbringen (z.B. Hashwerte). Bei anderen sind graduelle Übergänge denkbar. Häufig entsprechen dann die Ergebnisse nicht mehr vollständig den an sie gerichteten Erwartungen (z.B. farbliche Nuancen). Manchmal können Ausgangs- und Zielobjekt gar nicht vollständig übereinstimmen (z.B. bei Migrationen des Dateiformats). Auch die vollständige Umsetzung einer Norm (z.B. ISO 19005 - PDF/A) ist nur schwer zu realisieren. Für jeden Prozess muss daher ein Erfüllungsgrad definiert werden, der zu erreichen ist. Definiert werden muss auch, was geschieht, wenn dies nicht der Fall ist. Eine Konsequenz könnte die Zurückweisung des Objekts an den Produzenten sein (verbunden mit der Bitte, ein korrektes Objekt zu übermitteln). Eine andere Möglichkeit wäre es, Abweichungen bis zu einem definierten Grad nur in den archivischen Metadaten (in einem Validierungsprotokoll) festzuhalten. PREMIS spricht in diesem Zusammenhang von "Quirks" [Lit. 3, S. 204]. Beispiele für Validierungsprozesse mit graduellen Erfüllungen können sein:

- n% fehlerhafte Eintragungen in einem Datenbankfeld werden noch toleriert.
  - n% undokumentierte Datenbankeigenschaften können noch akzeptiert werden.
- 

### Ziel:

*Für jeden Validierungsprozess sind die beteiligten Personen und Hilfsmittel benannt.*

### Verfahren:

Für jeden Prozess muss zunächst geklärt werden, wer welche Aufgaben übernimmt und welche Hilfsmittel verwendet werden. Darüber hinaus kann es aus methodischen Gründen erforderlich sein, bei der Validierung andere Methoden und Softwaretools zu verwenden als bei der Erstellung des Transferpakets, um methoden- oder toolbedingte Fehler erkennen zu können.

- Wer führt die Validierung aus: Der Produzent oder das Archiv?
  - Werden Dritte (Experten, Vertreter der Designated Community) beteiligt?
  - Welche Tools und Methoden werden für die Validierung eingesetzt?
  - Wo findet die Validierung statt (beim Produzenten oder im Archiv)?
- 

### Ziel:

*Die Validierungsprozesse folgen einer plausiblen zeitlichen Abfolge.*

### Verfahren:

Die Validierungsprozesse werden in einzelnen Phasen während der Übernahme durchgeführt. Die Unterteilung der Phasen und ihre Zahl können sich von Archiv zu Archiv unterscheiden, häufig ist jedoch eine Unterteilung in zwei Phasen. In einer ersten Phase wird dann unmittelbar nach der Übernahme der Objekte ins Archiv geklärt, ob die Objekte grundsätzliche Anforderungen erfüllen. Fallen zu diesem Zeitpunkt Abweichungen vom erwarteten Ergebnis auf, wird

die Übernahme der Objekte abgelehnt. Diesen Eigenschaften kommt somit eine KO-Funktion zu (s.u. die beiden erstgenannten Beispiele). In einer späteren, zweiten Phase werden dann tiefergehende Validierungsprozesse vorgenommen. Erst zu diesem Zeitpunkt spielen auch solche Prozesse eine Rolle, die nicht mehr zu einem eindeutigen Ja/Nein-Ergebnis führen.

## 2.3 Transfer von Daten aus dem Produzentensystem

Wesentlich für die Übernahme der Daten vom Produzenten zum digitalen Langzeitarchiv ist vor allem die korrekte und vollständige Übertragung aller Daten, die im Archiv zur Rekonstruktion der entsprechenden Informationsobjekte und zur dauerhaften Verwaltung, Erhaltung und Darstellbarkeit der Intellektuellen Entitäten benötigt werden. Nach dem Transfer der Daten muss das digitale Langzeitarchiv bei einer eingehenden Lieferung genau nachvollziehen, von wem die Lieferung stammt (Authentizität), in welcher Form Daten und Metadaten geliefert werden müssen (Validität) und welchen Umfang die Datenlieferungen haben müssen (Vollständigkeit). Maßgeblich sind die gesetzlichen bzw. vertraglichen Rahmenbedingungen, die von Fall zu Fall variieren. Die gesetzlichen Vorgaben und die zum Transfer verwendete Technik müssen aufeinander abgestimmt werden.

---

### Ziel:

*Der gesetzliche bzw. vertragliche Rahmen für einen Transfer ist vom Produzent und vom Archiv analysiert und definiert.*

### Verfahren:

Die Bestimmungen für einen Transfer in Punkten wie abgesicherter Datentransfer, Aufrechterhaltung der Gültigkeit von qualifizierten Signaturen usw. [Lit. 12] [Lit. 13] können variieren. Zum Beispiel können Akten einer Behörde unterschiedliche Vertraulichkeitsstufen haben.

---

### Ziel:

*Die technischen und organisatorischen Möglichkeiten für den Transfer sind dem Produzenten und dem Archiv bekannt. Es herrscht bei beiden Klarheit darüber, ob die technischen Möglichkeiten einen bestimmungskonformen Transfer erlauben oder nicht. Gegebenenfalls wird eine Vereinbarung über die notwendigen Anpassungen getroffen.*

### Verfahren:

Beide Seiten müssen ihre technischen Voraussetzungen kennen und prüfen, insbesondere wenn vorher noch nie ein Transfer zwischen den Partnern stattgefunden hat. Die technischen Möglichkeiten von Produzent und Archiv sind, insbesondere vor dem ersten Transfer, möglicherweise nicht konform mit den gesetzlichen Vorschriften. Die Konformität ist genau zu prüfen, da sich IT-Sicherheitsbedingungen auch stetig ändern. Zum Beispiel verlieren Protokolle zur



Verschlüsselung und qualifizierte Signaturen mit der Zeit ihre Wirksamkeit gegenüber böswilliger Manipulation.

---

Ziel:

*Die einzelnen Arbeitsschritte des Transfers sind zwischen Produzent und Archiv genau abgestimmt und getestet.*

Verfahren:

Der Prozess des Transfers ist kritisch für die Authentizität des Informationsobjekts. Die genaue Einhaltung eines vereinbarten Transferprozesses erhöht die Vertrauenswürdigkeit des Archivs. Bei regelmäßigen Transfers, die weitestgehend automatisiert vorgenommen werden, muss der Transferprozess ohnehin technisch in die Systeme des Produzenten und des Archivs implementiert werden. Produzent und Archiv legen die maximale Datengröße des Transfers fest, einigen sich über die technischen Reporte, die erstellt werden, den Transferweg (Datenübertragung mit Protokoll oder auf Datenträger mit Angabe des Datenträgerformats durch Lieferung), den Zeitraum des Transfers, alle notwendigen Identifikatoren und Passwörter sowie die notwendigen Sicherheitsprotokolle.

Die vereinbarten technischen Werkzeuge müssen in die Systeme der Partnerinstitutionen implementiert und aufeinander abgestimmt werden. Zur Sicherheit erstellen der Produzent und das digitale Langzeitarchiv gemeinsam Test-Transferpakete und führen kontrollierte Transfer-Tests durch. Nur unter kontrollierten Bedingungen kann sicher festgestellt werden, dass der vereinbarte und implementierte Transfer korrekt funktioniert.

## 3 Management

### 3.1 Identifikation von rechtlichen und vertraglichen Bedingungen

Bereits vor der Übernahme von digitalen Objekten ins Archiv müssen rechtlichen Regelungen zwischen Produzent und Archiv bestehen bzw. geschaffen werden, um für beiden Seiten langfristig Planungs- und Rechtssicherheit gewährleisten zu können. In diesen Bereich fallen auch rechtliche Fragen, die sich nicht unmittelbar auf die Übernahme beziehen, jedoch bereits bei der Objektaufnahme geklärt sein müssen, um die Dauerhaftigkeit der Archivierung und die Bedingungen für den Umgang mit den archivierten Objekten zu regeln. Die Klärung weitergehender rechtlicher Fragen im Verhältnis Produzent-Archiv ist also Voraussetzung für eine sinnvolle und erfolgreiche Übernahme. Neben der Analyse der Grundlage der rechtlichen Beziehungen zwischen Produzent und Archiv sind urheberrechtliche Fragen der zweite Schwerpunkt in diesem Themenfeld.

---

#### Ziel:

*Alle rechtlich handelnden Parteien und die jeweiligen vertretungsberechtigten Personen sind identifiziert bzw. festgelegt.*

#### Verfahren:

Es ist zu klären, wer genau der Produzent der Informationsobjekte ist und wer über die Rechte an den zu archivierenden Objekten verfügt. Nur so kann sichergestellt werden, dass der Produzent überhaupt in eine Archivierung einwilligen kann bzw. zur Abgabe von Daten zum Zwecke der Archivierung verpflichtet werden kann. Eine rechtliche Regelung zwischen einem nicht über die entsprechenden Rechte verfügenden Produzenten und dem Archiv wäre juristisch wertlos. Die jeweils Vertretungsberechtigten der handelnden Parteien sind ebenfalls zu ermitteln.

---

#### Ziel:

*Die rechtlichen Beziehungen zwischen Produzent und digitalem Langzeitarchiv sind genau bestimmt.*

#### Verfahren:

Die Archivierungstätigkeit bedarf einer rechtlichen Grundlage, um Rechts- und Planungssicherheit sowohl für den Produzenten als auch für das Archiv zu gewährleisten. Es ist zu klären, ob eine solche rechtliche Regelung besteht. Falls diese rechtliche Grundlage besteht, muss deren Natur bestimmt werden. Dabei ist grundsätzlich zu unterscheiden, ob die Archivierungstätigkeit in einem gesetzlichen Auftrag oder durch eine rechtliche Vereinbarung zwischen Archiv und Produzent begründet ist bzw. wird. Falls es keine rechtliche Grundlage in Form einer gesetzlich festgelegten Ablieferungspflicht gibt, muss zumindest für den urheberrechtlich relevanten Bereich eine Übereinkunft (Lizenzvereinbarung) zwischen Produzent und Archiv getroffen werden. Damit das Archiv seine Auf-

gaben wahrnehmen kann, ist es nötig, dass es mittels einer rechtlichen Regelung die nötigen Rechte für die angestrebte Archivierung und Nutzung erwirbt.

---

Ziel:

*Die Verpflichtungen des Archivs und/oder Produzenten bezüglich des Umgangs mit dem zu archivierenden Material sind bekannt.*

Verfahren:

Es ist zu bestimmen, ob sich für das Archiv und/ oder Produzenten aus gesetzlichen Regelungen verpflichtende Vorgaben bezüglich der Aufbewahrung und Nutzung der zu archivierenden Objekten bzw. deren Inhalten ergeben.

Ein Archivgesetz könnte beispielsweise vorschreiben, wie viele Kopien eines Objekts vom Archiv vorzuhalten sind und ob gegebenenfalls Originaldokumente auf Seiten des Produzenten zu löschen sind. Ebenso können möglicherweise Datenschutzbestimmungen einer umfassenden Suchfunktion in den Archivbeständen oder einer Bereitstellung von Archivmaterial an Dritte entgegenstehen.

---

Ziel

*Die urheberrechtlichen Bedingungen des zu archivierenden Materials sind dem Archiv bekannt und dauerhaft hinterlegt.*

Verfahren

Die Rechte des Urhebers am zu archivierenden Objekt sind zu klären. Das zu archivierende Material fällt unter bestimmten Voraussetzungen unter die Bestimmungen von Schutz- und Urheberrechten (Intellectual Property Rights). Sind die entsprechenden Bestimmungen einschlägig, ist eine digitale Archivierung in Deutschland nur in einem äußerst engen Rahmen zulässig, da es sich bei der Archivierung von digitalen Objekten stets um eine Vervielfältigung im Sinne des UrhG handelt. Diese Vervielfältigung muss stets durch eine entsprechende Schrankenbestimmung des UrhG oder durch die Übertragung des entsprechenden Nutzungsrechts vom Rechteinhaber an das Langzeitarchiv gedeckt sein.

In Deutschland sind die Bestimmungen des Urheberrechtsgesetzes (UrhG) nur relevant, wenn ein Werk über eine gewisse Schöpfungshöhe verfügt und es sich im Rahmen der gesetzlich festgelegten Schutzdauer bewegt. Eine Schrankenbestimmung findet sich in § 53 Abs. 2 S. 1 Nr. 2 UrhG, die so genannte Archivschranke, die die Vervielfältigung zum Zwecke der Aufnahme in ein eigenes Archiv erlaubt. Die Archivschranke erlaubt aber lediglich die Vervielfältigung zum Zwecke der Sammlung, Aufbewahrung und Konservierung, nicht jedoch die Nutzung der archivierten Objekte durch Dritte.

---

#### Ziel:

*Produzent und Archiv haben die Inhalte der urheberrechtlichen Bestimmungen analysiert und falls nötig diesbezüglich Regelungen getroffen.*

#### Verfahren:

Falls eine urheberrechtliche Relevanz der zu archivierenden digitalen Objekte festgestellt wird, sind verschiedene Problemstellungen der Langzeitarchivierung zu berücksichtigen, die entweder gesetzlich geregelt sein müssen oder einer rechtlich wirksamen Vereinbarung zwischen dem ermittelten Inhaber der entsprechenden Rechte und dem Archiv bedürfen. Damit das Archiv seine Aufgaben wahrnehmen kann, ist es nötig, dass es mittels einer rechtlichen Regelung die nötigen Rechte für die angestrebten Formen der Archivierung und Nutzung erwirbt. Die Verantwortlichkeiten für Produzent und Archiv müssen durch gesetzliche Grundlagen und / oder verbindliche Vereinbarungen präzise definiert werden.

Die Migration von digitalen Objekten in andere Dateiformate könnte beispielsweise durch eine spezielle gesetzliche Regelung erlaubt sein. Die Entfernung von technischen Kopierschutzmaßnahmen (DRM) muss zwischen Produzent und Archiv vertraglich geregelt werden, sofern es dafür keine gesetzliche Grundlage gibt. In einer Lizenzvereinbarung können Archiv und Produzent/Rechteinhaber festlegen, dass die archivierten Objekte einem bestimmten Nutzerkreis zugänglich gemacht werden. Zu den rechtlich geregelten Verantwortlichkeiten auf Seiten des Produzenten könnte gehören, dass er verpflichtet ist, dem Archiv die digitalen Objekte anzubieten bzw. sie dort abzuliefern. Auch die Verteilung der Kosten muss rechtlich geregelt sein, darunter fallen unter anderem der Transport ins Archiv, die Kosten für Archivierung und Pflege sowie die Erstellung von Kopien für den Produzenten.

---

#### Ziel

*Zwischen Produzenten und Archiv bestehen Regelungen bezüglich Gewährleistung und Haftung.*

#### Verfahren

Es ist festzulegen, wann und von welcher Seite Schadensersatzansprüche geltend gemacht werden können und welche Sorgfaltspflichten einzuhalten sind. Zudem ist zu klären, ob es Regelungen gibt, wenn die Rechte Dritter durch den Produzenten oder das Archiv verletzt werden.

---

### **3.2 Übernahmevereinbarungen und Dokumentation**

Die Dokumentation der vereinbarten Vorgaben und Eckdaten der Übernahme und die Protokollierung der Übernahmevorgänge ermöglichen die Nachvollzieh-

barkeit eines Teils der Überlieferungsgeschichte der zu archivierenden Objekte. Damit tragen sie zur Gewährleistung der Integrität und Authentizität dieser Objekte bei, wie dies vom Kriterienkatalog vertrauenswürdige digitale Langzeitarchive gefordert wird. An ihre Archivierung werden die gleichen Anforderungen gestellt wie an jene der Primärdaten.

---

Ziel:

*Produzent und Archiv schließen eine Übernahmevereinbarung ab. Das ist eine verbindliche Vereinbarung, welche alle Aspekte der Informationsübernahme regelt.*

Verfahren:

Die Vereinbarung muss vom Produzenten und vom digitalen Langzeitarchiv genehmigt werden. Sie hält die Ergebnisse der Planung der Übernahme fest, wie sie in diesem Dokument detailliert beschrieben sind. Die Vereinbarung dokumentiert die Ergebnisse der Planung der Übernahme und dient als verbindlicher Leitfaden für die eigentliche Übernahme. Allfällige Änderungen der in der Vereinbarung festgelegten Prozesselemente müssen auf geregelte und dokumentierte Art erfolgen. Insbesondere enthält sie:

- die Liste der zu archivierenden Intellektuellen Entitäten inkl. die Definition derer wesentlichen Eigenschaften;
  - die Liste, der diese Intellektuellen Entitäten repräsentierenden Informationsobjekte und Daten, die zu ihrer Archivierung notwendige technische Umgebung und allfällige Migrationsvereinbarungen;
  - die Liste der notwendigen Metadaten mit der Angabe, von wem diese bereitgestellt werden;
  - das Transferpaketformat inklusive der darin benötigten Metadaten, den Identifikator und die Abbildung der Informationsobjekte auf Pakete;
  - der Transfer und seine technische Implementierung;
  - die Definition der einzelnen Validierungsprozesse inkl. deren notwendiger Erfüllungsgrad, die Konsequenzen bei Nichterfüllung, die beteiligten Personen und Hilfsmittel sowie die zeitliche Abfolge;
  - die Angaben, die als Grundlage für die Risikoanalyse dienen; insbesondere die Schätzungen zur Datenmenge, Rechenkapazität und -zeit für die Übernahme sowie die Kostenschätzungen;
  - die rechtlich handelnden Parteien und deren Beziehungen sowie die Regelungen bezüglich Urheberrecht und Haftung;
  - den Zeitplan für die Durchführung der Informationsübernahme.
- 

Ziel:

*Jede Informationsübernahme ist vollständig vom Start bis zur Archivierung protokolliert. Dieses Protokoll sollte ebenso dauerhaft erhalten werden, wie die betreffenden Objekte im Archiv.*

### Verfahren:

Das Protokoll sollte beispielhaft die folgenden Angaben enthalten:

- Liste aller übernommenen Intellektuellen Entitäten und der entsprechenden Informationsobjekte.
- Name des Produzenten.
- Zeit und Datum des Beginns des Transfers
- Zeit und Datum des Eintreffens des Transferpakets beim Archiv.
- Zeit und Datum der Archivierung.
- Vorgenommene Transformationen an den Informationsobjekten.
- Ergebnisse der einzelnen Validierungen.

## **3.3 Qualitäts-, Sicherheits-, Prozess- und Risikomanagement**

Ziele von Managementaktivitäten: Die Übernahme digitaler Information ist ein kritischer Prozess, der die entsprechenden Managementaktivitäten erfordert, um ein gefordertes Qualitätsniveau zu erreichen und Risiken bezüglich Sicherheit und Kosten angemessen zu handhaben. Das Management hat das Gesamtsystem im Blick und sorgt für dessen Schlüssigkeit, d.h. insbesondere, dass Ziele realistisch und entsprechend den Vorgaben des Trägers und des Gesetzgebers definiert werden und dass für die Erreichung der Ziele eine angemessene Organisation und Infrastruktur zur Verfügung stehen sowie alle Prozesse aufeinander abgestimmt sind. Hierzu dient die Einrichtung unterschiedlicher, aber kompatibler Managementsysteme. Eine isolierte Betrachtung der Übernahme mit der Folge unnötiger Schwankungen in der Qualität und Sicherheit sowie lokaler Effizienzbetrachtungen und nicht beachteter Risiken soll somit vermieden werden.

Qualitätsmanagement: Vorrangiges Ziel eines Qualitätsmanagements ist die Kundenzufriedenheit, aber auch die Zufriedenheit weiterer Interessierter wie Aufgabenträger oder die Gesellschaft als Ganzes. Wesentliche Aufgabe eines Qualitätsmanagements ist das Erkennen gegenwärtiger und künftiger Erfordernisse der Kunden und das Umsetzen derer Anforderungen. Hierzu sind die Qualitätspolitik und die Qualitätsziele einer Organisation festzulegen. Um die Qualitätsziele zu erreichen, sind die erforderlichen Prozesse und Verantwortlichkeiten festzulegen. Des Weiteren sind die erforderlichen Ressourcen zu bestimmen und bereitzustellen (vgl. Prozessmanagement). Die Wirksamkeit und die Effizienz der Prozesse sind zu überprüfen und die Ursachen von Fehlern sind zu analysieren und zu beheben. Im Kontext der Übernahme sorgt das Qualitätsmanagement u.a. für:

- grundsätzliche Vorgaben für alle Übernahme- und Validierungsprozesse entsprechend den Qualitätszielen der Organisation
- eine Bereitstellung aller erforderlichen Ressourcen für Qualitätsmanagementprozesse einschl. Validierung

- die Einhaltung der Qualität der erforderlichen Ressourcen
- eine Revision der Qualitätspolitik und Qualitätsziele falls diese nicht erreichbar sind
- eine einheitliche Kontrolle aller Validierungsprozesse
- grundsätzliche Vorgaben für die Dokumentation aller Übernahmeprozesse
- eine Integration ggf. ausgelagerter Prozesse in das Qualitätsmanagement
- die Identifizierung und Anwendung geeigneter Qualitätsstandards.

Informationssicherheitsmanagement: Ziel eines Sicherheitsmanagements ist es, solche Bedrohungen abzuwenden, die eine Gefahr für die Erreichung der Gesamtziele und letztlich für die Vertrauenswürdigkeit der Organisation darstellen. Aufgabe des Sicherheitsmanagements ist die Übernahme der Gesamtverantwortung - hier insbesondere für die Informationssicherheit. Hierzu werden eine Sicherheitspolitik und Sicherheitsziele festgelegt. Für die Umsetzung sind die entsprechenden Prozesse und Verantwortlichkeiten zu bestimmen sowie die erforderlichen Ressourcen bereit zu stellen. Wesentlich sind das Erkennen von Bedrohungen und die Ermittlung des Schadenspotenzials. Des Weiteren ist die Wirksamkeit und die Effizienz der Sicherheitsprozesse zu überprüfen und die Ursachen von Fehlern zu analysieren und zu beheben. Im Kontext der Übernahme sorgt das Informationssicherheitsmanagement u.a. für:

- mit Produzenten bzw. Lieferanten abgestimmte und in Vereinbarungen festgelegte Sicherheitspolitik und Sicherheitsziele
- Vorgaben, die die spezifischen Risiken von Kategorien von Übernahmen berücksichtigen (Selbstarchivierung einer Community, anonymer Zugang, Übernahme von ausführbaren Objekten, eingestuften Materialien oder virtuellen Objekten, u.ä.)
- angemessene organisatorische Rahmenbedingungen (z.B. Bestimmung eines Sicherheitsbeauftragten, Festlegen von Zuständigkeiten für die Vergabe von Passwörtern)
- die (ggf. gemeinsame) Bereitstellung, den Betrieb und die Überwachung einer entsprechenden Sicherheitsinfrastruktur (z.B. Public Key Infrastruktur, Definition geeigneter Metadaten)
- die Identifikation und Anwendung von Sicherheitsstandards bzw. gesetzlicher Vorgaben.

Prozessmanagement: Ziel eines Prozessmanagements ist die effektive und effiziente Umsetzung von Organisationszielen. Es sorgt für interne und externe Transparenz und liefert somit einen Beitrag zur Schlüssigkeit und Vertrauenswürdigkeit einer Organisation. Zu den Aufgaben des Prozessmanagements zählt die Zusammenfassung aller Aktivitäten zu Prozessen unter Berücksichti-

gung zeitlicher, räumlicher und ablauflogischer Abhängigkeiten, so dass betriebswirtschaftlich sinnvolle Aussagen herleitbar sind. Die Ziele der Organisation bilden hierfür die Maßgabe. Prozesse liefern die Basis für die Zuordnung von Ressourcen und Verantwortlichkeiten sowie für die Spezifikation der jeweils zu erbringenden bzw. benötigten Leistung. Organisatorische und technische Schnittstellen können somit genauer festgelegt werden. Zu den Aufgaben eines Prozessmanagements gehört auch die Integration unterschiedlicher Managementprozesse.

Im Kontext der Übernahme sorgt das Prozessmanagement u.a. für

- eine Darstellung aller Auswirkungen von Übernahmen auf nachfolgende Prozesse insbesondere auf die Archivablage und die Auslieferung von Information
- die Vermeidung unausgewogener Auswirkungen auf andere Prozesse durch Maßnahmen zur Qualitätssicherung, Sicherheit oder Effizienzsteigerung von Übernahmen
- eine Betrachtung von der Übernahme vorgelagerten Prozessen. Diesbezügliche Kenntnisse oder möglichst eine Einflussnahme können einen positiven Einfluss auf die Übernahme haben (z.B. Auswahl von Formaten für Inhalte und Metadaten, Kenntnisse über technische und intellektuelle Erstellungskontexte).
- eine Überführung von projekthaft durchgeführten oder vorbereiteten Übernahmen in "Standardprozesse" durch ein entsprechendes Konfigurations- und Änderungsmanagement
- eine Anpassung von bestehenden Übernahmeprozessen an geänderte technische und organisatorische archiv- und kundenseitige Bedingungen durch ein entsprechendes Konfigurations- und Änderungsmanagement.

Risikomanagement: Neben den Risiken für Qualität und Sicherheit bestehen weitere Risiken, die sich durch die Gestaltung von Übernahmen ergeben können. Ziel des Risikomanagements ist es, diese Risiken zu erfassen und zu bewerten, zu minimieren/steuern und kontinuierlich zu überwachen. Grundsätzlich unterscheidet man fünf verschiedene Strategien zur Risikosteuerung: die Vermeidung, Verminderung, Begrenzung, Überwälzung und Akzeptanz von Risiken. Das Risikomanagement ist ein iterativer Prozess und erstreckt sich über den gesamten Ablauf des Übernahmeprozesses. Es involviert sowohl den Produzenten als auch das Archiv. Im Kontext der Übernahme sind besonders die folgenden Risiken in Betracht zu ziehen:

- Finanzielle Risiken: Jeder Übernahmeprozess (und jeder Aspekt des in diesem Leitfaden beschriebenen Übernahmeprozesses) erfordert Ressourcen beim Produzenten und im Archiv und schafft weitere Ressourcenansprüche für die dauerhafte Archivierung der übernommenen Daten. Die detaillierte Planung und das Budget eines Übernahmeprozesses dienen als Grundlage zum Management finanzieller Risiken. Finanzielle Risiken betreffen einerseits personelle, andererseits finanzielle Ressourcen.



- Personelle Ressourcen: Zur Durchführung der Übernahme ist Arbeitszeit von verschieden qualifiziertem Personal erforderlich. Die dauerhafte Archivierung der übernommenen Daten zieht weiteren personellen Ressourcenbedarf nach sich zur allfälligen Migration und zur Aufbereitung für die Benutzung.
- Finanzielle Ressourcen: Der Übernahmeprozess verursacht Kosten für den Datentransfer, die provisorische Speicherung und die benötigte Rechenkapazität. Die dauerhafte Archivierung der übernommenen Daten verursacht Kosten für den Speicherplatz, die Datensicherung und den Datenschutz, für Rechenkapazität und temporären Speicher zur Benutzung sowie für Rechenzeit und temporären Speicher für allfällige Migrationen.
- Rechtliche Risiken: Die Informationsübernahme durch ein Archiv spielt sich im Rahmen verschiedener gesetzlicher und/oder vertraglicher Bestimmungen ab (wie im Kapitel "Identifikation von rechtlichen und vertraglichen Bedingungen" ausgeführt). Das Nichteinhalten dieser Bestimmungen kann verschiedene Sanktionen nach sich ziehen.
- Reputationsrisiken. Neben den gesetzlichen kann ein Archiv auch gewisse moralische Verpflichtungen zur langfristigen Aufbewahrung von digitalen Daten haben. Entsprechend kann die Reputation des Archivs darunter leiden, dass es diese Daten nicht übernimmt, oder darunter, dass es die Daten zwar übernimmt, aber ihre Erhaltung nicht gewährleisten kann.

# Glossar

**Archiv:**

Siehe Digitales Langzeitarchiv.

**Authentizität:**

Das Objekt stellt das dar, was es vorgibt darzustellen. (Übernommen aus: Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Lit. [2] S. 34)

**Daten:**

Digital gespeicherte Bestandteile eines Informationsobjekts (nach PREMIS Files oder Bitstreams). Entspricht dem PREMIS Objekttyp „Bitstream“ (Vgl. PREMIS 2.0, Lit. [3] S. 7)

**Digitales Langzeitarchiv:**

Eine Organisation (bestehend aus Personen und technischen Systemen), die die Verantwortung für den Langzeiterhalt und die Langzeitverfügbarkeit digitaler Daten sowie die Bereitstellung für eine bestimmte Zielgruppe übernommen hat. Dabei bedeutet „Langzeit“: über Veränderungen in der Technik (Soft- und Hardware) und in der Zielgruppe hinweg (z.B. für künftige Generationen, auf unbegrenzte Zeit). Im Text wird Archiv synonym zu digitales Langzeitarchiv verwendet. (Übernommen aus: Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Lit. [2] S. 34)

**Emulation:**

Strategie zur Erhaltung der Langzeitverfügbarkeit digitaler Objekte. Die Strategie sieht vor, dass die Systemvoraussetzungen, die zur Nutzung älterer digitaler Objekte notwendig sind, durch spezialisierte Software auf aktuellen marktgängigen Systemen nachgebildet (emuliert) werden können. Die digitalen Objekte selbst werden dabei möglichst unverändert erhalten.

**Informationsobjekt:**

Entspricht dem PREMIS Objekttyp „Representation“. Wird gebildet aus digital gespeicherten Dateneinheiten (nach PREMIS Files oder Bitstreams) und kann eine Intellektuelle Entität repräsentieren. (Vgl. PREMIS 2.0, Lit. [3] S. 7)

**Intellektuelle Entität:**

Von Menschen interpretierbare logisch abgegrenzte Einheit eines Inhalts, die von Informationsobjekten materiell/physisch dargestellt wird. Entspricht der PREMIS Entität „Intellectual Entity“ (Vgl. PREMIS 2.0, Lit. [3] S. 6)

**Integrität:**

1. die Vollständigkeit der digitalen Objekte, 2. der Ausschluss unbeabsichtigter Modifikationen im Sinne der Erhaltungsregeln. Maßstab für die Integrität sind die als erhaltenswert definierten Eigenschaften eines digitalen Objekts. (Übernommen aus: Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Lit. [2] S. 34)

**Langzeitarchivierung (LZA):**

Langzeitarchivierung digitaler Objekte umfasst alle Maßnahmen, die dazu dienen, digitale Objekte für die Nachwelt dauerhaft zu erhalten. Der Begriff ist eng verwandt mit Langzeitverfügbarkeit, die jedoch die dauerhafte Benutzbarkeit mehr in den Vordergrund stellt. Gängige LZA-Strategien sind Emulation und Migration. (Übernommen aus dem nestor Glossar)

**Metadaten:**

Daten, die Informationen über andere Daten repräsentieren, indem sie z.B. Inhalt, Struktur, Beschaffenheit, Handhabung, Herkunft etc. näher beschreiben. (Übernommen aus: Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Lit. [2] S. 34)

**Migration:**

Dateiformatmigration: Konvertierung des Formats eines Informationsobjekts in ein anderes. Erhaltungsstrategie zur Anpassung an eine veränderte technische Umgebung. Datenträgermigration: Kopieren eines Informationsobjekts auf einen anderen Datenträger.

**Produzent:**

Personen oder Client-Systeme, die dem Digitalen Langzeitarchiv digitale Objekte zum Zweck der Langzeiterhaltung übergeben. Dies sind nicht notwendigerweise die Urheber, sondern können auch Lieferanten der digitalen Objekte sein. (Übernommen aus: Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Lit. [2] S. 35)

**Significant properties:**

Siehe Wesentliche Eigenschaften. (Vgl. PREMIS 2.0, Lit. [3] S. 39)

**Transferpaket:**

Definierte Menge von Informationsobjekten, die vom Produzenten als Einheit an das Langzeitarchiv übertragen wird, z.B. in Form von einer oder mehrerer Dateien.

**Übernahme:**

Beinhaltet die Organisation und Durchführung aller Prozesse, die zur Aufnahme eines Informationsobjekts ins Archiv und zur Übernahme der Verantwortung durch das Archiv führen.

**Wesentliche Eigenschaften:**

Eigenschaften eines Informationsobjekts, welche für eine Zielgruppe als wichtig erachtet werden und erhalten bleiben sollen. (Vgl. PREMIS 2.0, Lit. [3] S. 39)

**Zielgruppe:**

Identifizierbare Gruppe von potentiellen Nutzern mit bestimmten Interessen und Voraussetzungen. Dies kann die Allgemeinheit ebenso sein wie eine Gruppe Fachwissenschaftler. Sie kann heterogen sein und sich aus verschiedenen Nutzergruppen zusammensetzen. (Übernommen aus: Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Lit. [2] S. 35)

# Literatur

## Grundlagen:

**[1]** ISO 14721:2003 - Reference Model for an Open Archival Information System (OAIS) - Space Data and Information Transfer Systems

**[2]** Kriterienkatalog vertrauenswürdige digitale Langzeitarchive (dt./engl) / hrsg. von der nestor-Arbeitsgruppe Vertrauenswürdige Archive - Zertifizierung. Frankfurt am Main : nestor c/o Die Deutsche Bibliothek, 2006. - 40 S. - nestor Materialien 8 - URN: urn:nbn:de:0008-2006060710

**[3]** PREMIS Data Dictionary for Preservation Metadata - version 2.0 - PREMIS Editorial Committee, März 2008

## Erwähnte Literatur und Standards:

**[4]** DOMEA Konzept - Organisationskonzept 2.0 - Erweiterungsmodul zum Organisationskonzept 2.0 Aussonderung und Archivierung elektronischer Akten - Schriftenreihe der KBSt 66, Oktober 2004 - Webseite unter: <[http://www.verwaltung-innovativ.de/cln\\_110/nn\\_1007474/SharedDocs/Publikationen/DE/domea\\_\\_konzept\\_\\_aussonderung\\_\\_und\\_\\_archivierung\\_\\_elektronischer\\_\\_akten.html?\\_\\_nnn=true](http://www.verwaltung-innovativ.de/cln_110/nn_1007474/SharedDocs/Publikationen/DE/domea__konzept__aussonderung__und__archivierung__elektronischer__akten.html?__nnn=true)>

**[5]** [MAB2](#): Maschinelles Austauschformat für Bibliotheken. - Loseblatt-Ausg. - ISSN 0949-5258 - Grundwerk . - 2. Aufl. (auf dem Stand der 1. Erg.-Lfg. Mai 2002). - ISBN: 978-3-933641-00-7

**[6]** MARC Standards - Library of Congress - Network Development and MARC Standards Office - Webseite unter: <<http://www.loc.gov/marc/>>

**[7]** The Dublin Core Metadata Element Set ISO 15836-2003 - Dublin Core Metadata Initiative - Website unter: <<http://dublincore.org/>>

**[8]** METS: Metadata Encoding and Transmission Standard - Library of Congress - Webseite unter: <<http://www.loc.gov/standards/mets/>>

**[9]** LMER: Langzeitarchivierungsmetadaten für elektronische Ressourcen - Version 1.2 - April 2005 - URN: urn:nbn:de:1111-2005041102

**[10]** Persistent Identifier ...eindeutige Bezeichner für digitale Inhalte... - Webseite unter: <<http://www.persistent-identifier.de/?link=900>>

**[11]** SAGA: Standards und Architekturen für E-Government Anwendungen - Version 4.0 - Bundesministerium des Inneren, März 2008 - Webseite unter: <[http://www.kbst.bund.de/cln\\_012/nn\\_837392/SharedDocs/Meldungen/2008/SAGA/saga\\_\\_4\\_\\_0.html](http://www.kbst.bund.de/cln_012/nn_837392/SharedDocs/Meldungen/2008/SAGA/saga__4__0.html)>

**[12]** BSI Standards - Bundesamt für Informationssicherheit - Webseite unter:

<[http://www.bsi.de/literat/bsi\\_standard/index.htm](http://www.bsi.de/literat/bsi_standard/index.htm)>

**[13]** ArchiSafe – Physikalisch-Technische Bundesanstalt – Webseite unter:  
<<http://www.archisafe.de/s/archisafe/index>>

**[14]** ISO 20652:2006 - Producer-Archive Interface Methodology Abstract  
Standard (PAIMAS) - Space Data and Information Transfer Systems

**[15]** ISO-19005-1 - Document management – Electronic document file format  
for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1)

**[16]** Andrew Wilson, Significant Properties Report, 2007  
Webseite unter:  
<[http://www.significantproperties.org.uk/documents/wp22\\_significant\\_properties.pdf](http://www.significantproperties.org.uk/documents/wp22_significant_properties.pdf)>

# Übersicht über die Ziele der Informationsübernahme

## 1. Objekte

- a. Auswahl der zu archivierenden Information
  - i. Intellektuelle Entitäten auswählen
  - ii. Exportform analysieren
  - iii. Notwendige Anpassungen vereinbaren
- b. Auswahl der Metadaten
  - i. Definition der benötigten Metadaten
  - ii. Verantwortung für Bereitstellung der Metadaten klären
- c. Wesentliche Eigenschaften
  - i. Definition der wesentlichen Eigenschaften
  - ii. Erhebung der wesentlichen Eigenschaften

## 2. Prozesse

- a. Transferpakete
  - i. Abbildung der Informationsobjekte auf Transferpakete
  - ii. Rekonstruierbarkeit der Informationsobjekte
  - iii. Identifikation der Transferpakete
- b. Validierung
  - i. Definition der Validierungen
  - ii. Notwendige Erfüllungsgrad und Konsequenzen bei Nichterfüllung
  - iii. Beteiligte Personen und Hilfsmittel
  - iv. Zeitliche Abfolge
- c. Transfer von Daten
  - i. Gesetzliche/vertragliche Rahmen
  - ii. Technische und organisatorische Möglichkeiten
  - iii. Definition und Test der Transferarbeitsschritte

## 3. Management

- a. Recht und Vertrag
  - i. Identifikation juristischer Personen und Akteure
  - ii. Definition der Beziehungen zwischen Produzent und Archiv
  - iii. Verpflichtungen bzgl Archivmaterial sind bekannt
  - iv. Urheberrecht festgestellt
  - v. Regelung bzgl Urheberrecht
  - vi. Gewährleistung und Haftung
- b. Übernahmevereinbarung und Dokumentation
  - i. Verbindliche Dokumentation der Entscheidungen
  - ii. Protokollieren der Übernahmeprozesse
- c. Teilbereiche des Managements
  - i. Qualität
  - ii. Sicherheit
  - iii. Prozesse
  - iv. Kosten und Risiken (Recht, Reputation, Personal, Finanzen)

