

DIN 31645**DIN**

ICS 01.140.20; 35.240.30

**Information und Dokumentation –
Leitfaden zur Informationsübernahme in digitale Langzeitarchive**

Information and documentation –

Guide to the transfer of information objects into digital long-term archives

Information et documentation –

Guide pour la prise en charge d'information dans les archives électroniques à long terme

Gesamtumfang 25 Seiten

Normenausschuss Bibliotheks- und Dokumentationswesen (NABD) im DIN

Inhalt

	Seite
Vorwort	4
Einleitung.....	5
1 Anwendungsbereich	6
2 Normative Verweisungen	6
3 Begriffe	6
4 Objekte	9
4.1 Auswahl der zu archivierenden Informationen.....	9
4.1.1 Allgemeines	9
4.1.2 Auswahl der Informationsobjekte	9
4.1.3 Analyse des Datenexports	9
4.1.4 Problemlösung „Anpassung der Informationsobjekte“	10
4.2 Auswahl der Metadaten.....	10
4.2.1 Allgemeines	10
4.2.2 Definition der Metadaten	11
4.2.3 Verantwortlichkeit für die Metadaten	11
4.3 Identifikation von signifikanten Eigenschaften der Informationsobjekte.....	12
4.3.1 Allgemeines	12
4.3.2 Definition der signifikanten Eigenschaften.....	12
4.3.3 Verzeichnung und Dokumentation der signifikanten Eigenschaften.....	12
5 Prozesse	13
5.1 Definition von Transferpaketen.....	13
5.1.1 Allgemeines	13
5.1.2 Festlegung der Verhältnisse.....	13
5.1.3 Struktur der Paketinhalte	13
5.1.4 Identifizierung des Pakets	14
5.2 Validierung	14
5.2.1 Allgemeines	14
5.2.2 Definition der Validierungsprozesse und ihrer Erfüllungsgrade	14
5.2.3 Beteiligte Personen und Hilfsmittel	14
5.2.4 Zeitliche Abfolge der Validierungsprozesse.....	14
5.3 Transfer von Daten aus dem Produzentensystem	15
5.3.1 Allgemeines	15
5.3.2 Analyse der Rahmenbedingungen.....	15
5.3.3 Prüfung der Infrastrukturen	15
5.3.4 Implementierung des Transferprozesses	15
6 Management.....	16
6.1 Identifikation der rechtlichen und vertraglichen Bedingungen	16
6.1.1 Allgemeines	16
6.1.2 Identifizierung der Akteure aus rechtlicher Sicht.....	16
6.1.3 Bestimmung der Art der rechtlichen Beziehungen.....	16
6.1.4 Bestimmung des Inhalts der rechtlichen Regelungen.....	16
6.1.5 Bestimmung des urheberrechtlichen Status der Informationsobjekte.....	17
6.1.6 Regelung von Haftung und Gewährleistung.....	17
6.2 Übernahmevereinbarung und Dokumentation	17
6.2.1 Abschluss einer Übernahmevereinbarung	17
6.2.2 Protokoll der Übernahme	18
6.3 Kosten und Personal	18
6.3.1 Allgemeines	18
6.3.2 Kostenabschätzung.....	18
6.3.3 Personalabschätzung.....	18

Anhang A (normativ) **Übersicht über die Planungsschritte der Informationsübernahme** 19

Anhang B (informativ) **Beispiele zu Metadatenstandards** 22

B.1 **Inhaltsbeschreibende Metadaten (Beispiele zu 4.2.2)** 22

B.2 **Technische- und strukturbeschreibende Metadaten (Beispiele zu 4.2.2)** 22

B.3 **Klassifizierung signifikanter Eigenschaften (Beispiel zu 4.3)** 23

Literaturhinweise 24

Vorwort

Für das vorliegende Dokument ist der NA 009-00-15 AA „Schriftgutverwaltung und Langzeitverfügbarkeit digitaler Informationsobjekte“ verantwortlich. Es wurde vom zuständigen Arbeitskreis NA 009-00-15-05 AK „Ingest“ erarbeitet.

Es wird auf die Möglichkeit hingewiesen, dass einige Texte dieses Dokuments Patentrechte berühren können. Das DIN [und/oder die DKE] sind nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren.

Einleitung

Die Informationsübernahme in digitale Langzeitarchive ist ein kritischer Prozess innerhalb der gesamten Langzeitarchivierung. Ein Grund ist, dass er üblicherweise besonders kostenintensiv ist, da in ihm erst die Grundlage für eine spätere automatisierte und effiziente Verarbeitung der digitalen Daten gelegt wird. Aus Sicht der Daten sind es besondere Anfangsinvestitionen, die während der Informationsübernahme notwendig werden. Ein weiterer Grund für die Wichtigkeit dieses Prozesses ist, dass er besondere Konsequenzen für die Qualität hat. Ein digitales Langzeitarchiv wird Anfangsfehler meist nur selten verbessern oder rückgängig machen können und sie wirken sich auf den gesamten Archivierungszeitraum aus.

Aus dieser Situation ist ein Bedürfnis nach Information, Anleitung und Standardisierung für die Informationsübernahme in digitale Langzeitarchive erwachsen, auf das diese Norm antwortet. Sie behandelt nicht nur die Anforderungen an das digitale Langzeitarchiv, sondern auch Anforderungen an den Produzenten bzw. Lieferanten der zu bewahrenden Daten. Die Produzenten bzw. Lieferanten spielen eine wichtige Rolle im Übernahmeprozess, verfügen aber diesbezüglich meist über weniger Kenntnisse als die Betreiber von digitalen Langzeitarchiven.

Wesentliche Aufgaben einer Übernahme werden in dieser Norm aufgeführt und anhand von Anmerkungen und Beispielen näher beschrieben und begründet. Die Anwender dieser Norm sind aufgefordert, die Aufgabenliste innerhalb ihres Übernahmeprojekts abzuarbeiten. Zur praktischen Hilfestellung befindet sich eine Kontrollliste im Anhang, in der alle Schritte einer DIN 31645 konformen Übernahme zusammengefasst sind. Bei der Anwendung der Norm ist zu beachten, dass die einzelnen Aufgaben nicht in der aufgeführten Reihenfolge nacheinander abgearbeitet werden müssen.

1 Anwendungsbereich

Diese Norm gibt organisatorische Leitlinien für die Übernahme von Informationen aus einem Informationssystem in ein Langzeitarchiv unter Wahrung der Datenintegrität und Authentizität vor. Anfallende Aufgaben für den Produzenten und das digitale Langzeitarchiv werden in Planungsschritten benannt und in allgemeiner Form beschrieben.

Die Informationsübernahme in ein digitales Langzeitarchiv ist nicht nur ein rein technischer Transfer zwischen zwei Systemen, sondern ein Prozess mit vielen organisatorischen Anforderungen, an dessen Ende die Übernahme der Verantwortung durch das digitale Langzeitarchiv steht. Informationsobjekte müssen aus üblicherweise heterogenen technischen und organisatorischen Kontexten so übernommen werden, dass sie trotzdem in ganz anderen, zukünftigen Kontexten und sich ändernden technischen Umgebungen verstehbar und nutzbar sein werden. Ein Langzeitarchiv ist dabei nicht nur ein technisches System, sondern auch eine organisatorische Einheit oder Institution. Ein digitales Langzeitarchiv kann Bestandteil einer größeren Einrichtung sein, die auch analoge Bestände archiviert. Die Zusammenhänge zwischen digitalen und analogen Beständen sollten gewahrt und bei der Nutzung entsprechend dargestellt werden.

Diese Norm gilt für Vorhaben, in denen Informationsübernahmen in ein digitales Langzeitarchiv geplant und durchgeführt werden. Sie bezieht sich sowohl auf einmalige Informationsübernahmen wie auch auf den Aufbau von Prozessen zur wiederholten, automatisierten Informationsübernahme. Eine Übernahme ist konform zu dieser Norm, wenn die Beteiligten alle Planungsschritte des normativen Teils (Abschnitt 4 bis Abschnitt 6) ausreichend berücksichtigt haben.

2 Normative Verweisungen

Die folgenden zitierten Dokumente sind für die Anwendung dieses Dokuments erforderlich. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

E DIN 31644:2010-08, *Information und Dokumentation — Kriterien für vertrauenswürdige digitale Langzeitarchive*

ISO 14721:2003, *Space data and information transfer systems — Open archival information system — Reference model*

3 Begriffe

Für die Anwendung dieses Dokuments gelten die folgenden Begriffe.

3.1

Archivpaket

(en: Archival Information Package - AIP)

Paket, das im digitalen Langzeitarchiv gespeichert wird

ANMERKUNG Konzeptionelle Einheit aus ISO 14721:2003.

3.2

Authentizität

die zu archivierende Information wurde seit Beginn des Übernahmeverfahrens in das Langzeitarchiv nicht verändert

ANMERKUNG Authentizität beinhaltet, dass nur beabsichtigte, dokumentierte Veränderungen im Sinne der Erhaltungsmaßnahmen am Objekt durchgeführt wurden. Maßstab für die Authentizität sind die signifikanten Eigenschaften des Informationsobjekts.

3.3**(digitale) Daten**

digitale Repräsentation von Information in einer formalisierten Art, die die Interpretation, Verarbeitung bzw. den Austausch erlaubt

[E DIN 31644:2010-08]

ANMERKUNG Daten können in Form von Dateien technisch selbständige Einheiten bilden oder technisch unselbstständig als Bitsequenzen vorliegen.

3.4**digitales Langzeitarchiv****dLZA**

eine Organisation (bestehend aus Personen und technischen Systemen), die die Verantwortung für den Langzeiterhalt und die Langzeitverfügbarkeit von Information in digitaler Form sowie die Bereitstellung für eine bestimmte Zielgruppe übernommen hat

ANMERKUNG Digitale Langzeitarchive bewahren für den Menschen les- und interpretierbare Informationen, die in digitaler Form gespeichert werden. Im vorliegenden Kontext bedeutet Langzeit, dass Veränderungen in der Technik (Soft- und Hardware) und in der Zielgruppe berücksichtigt werden. Langzeit kann bereits ein Zeitraum von 10 – 30 Jahren sein, oder aber für künftige Generationen, auf unbegrenzte Zeit.

3.5**Emulation**

eine Strategie zur Erhaltung der Langzeitverfügbarkeit von Repräsentationen

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

3.6**Information**

jeder Typ von Wissen, der ausgetauscht werden kann, aus inhaltlicher Sicht

ANMERKUNG Z. B. Werke geistiger Schöpfung, Ergebnisse der Forschung und Entwicklung, Dokumentationen des politischen, sozialen und wirtschaftlichen Handelns. Information an sich ist logisch und zunächst ohne definierte Abgrenzung. Zur Archivierung muss zunächst die interessierende Informationseinheit abgegrenzt, d. h. als logisches Objekt konstituiert werden. Das so gebildete Informationsobjekt muss dann von einem Computer durch greifbare Einheiten repräsentiert werden, die dann ihrerseits durch Repräsentationen begrenzt werden. Dies kann auf verschiedene Weisen, d. h. mit unterschiedlichen Repräsentationen, geschehen.

3.7**Informationsobjekt**

logisch abgegrenzte Informationseinheit

3.8**Integrität**

Vollständigkeit und Unversehrtheit der Daten, d. h. dass diese weder absichtlich noch unabsichtlich, noch durch einen technischen Fehler verändert oder zerstört wurden

3.9**Metadaten**

Informationen zur Beschreibung der Eigenschaften von Informationsobjekten und ihren Repräsentationen

ANMERKUNG In Anlehnung an DIN ISO 15489:2002

3.10

Migration (Dateiformat)

beinhaltet das Erstellen einer neuen Repräsentation in einem neuen Dateiformat unter Zuhilfenahme eines Konvertierungsprogramms

3.11

Migration (Datenträger)

beinhaltet das Kopieren einer Repräsentation auf einen anderen Datenträger

3.12

Produzent

Person, Institution oder System, die/das dem digitalen Langzeitarchiv Repräsentationen von Informationsobjekten zum Zweck der Langzeiterhaltung übergibt

ANMERKUNG Produzenten sind nicht notwendigerweise die Urheber, sondern können auch Lieferanten der Repräsentationen sein.

3.13

Repräsentation

logisch abgegrenzte Einheit digitaler Daten in einem informationstechnischen System, die ein Informationsobjekt repräsentiert

3.14

signifikante Eigenschaften

Eigenschaften der für die Archivierung vorgesehenen Repräsentationen, die für den Erhalt der von ihnen repräsentierten Informationsobjekte im Hinblick auf die zukünftige Nutzung durch die Zielgruppen relevant sind

ANMERKUNG Über lange Zeit können digital gespeicherte Informationen nur erhalten werden, wenn die Repräsentationen selbst oder ihre Umgebung verändert werden. Ziel ist es, die signifikanten Eigenschaften nicht zu verändern. Signifikante Eigenschaften sind daher auf der Ebene der Informationsobjekte zu beschreiben.

3.15

Transferpaket

Paket, das vom Produzenten an das digitale Langzeitarchiv abgegeben wird

ANMERKUNG 1 Die Inhaltsdaten können bereits durch Metadaten angereichert sein.

ANMERKUNG 2 In Anlehnung an ISO 14721:2003 (en: Submission Information Package, SIP).

3.16

Übernahme

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

3.17

Zielgruppe

identifizierbare Gruppe von potenziellen Nutzern mit bestimmten Interessen und Voraussetzungen

ANMERKUNG Dies kann die Allgemeinheit sein sowie eine kleine Gruppe von Experten.