



Erhaltungsplanung und Risiko Management

28. November 2022

Voraussetzungen

Neben der Preservation Policy und dem Katalog archivtauglicher Dateiformate sollen auch Risiken für die Erhaltungsplanung im digitalen Langzeitarchiv erhoben werden. Ein Risiko kann den Bitstream betreffen, wenn er nicht mehr intakt ist. Dieses Risiko wird bereits minimiert durch die Strategie der «redundanten Datenhaltung»¹, d.h. die Daten liegen OAIS-konform in mehrfacher Kopie an verschiedenen Standorten vor und es werden regelmässige Integritätschecks vorgenommen. Aber auch die Dateien selbst sind gewissen Risiken ausgesetzt – wenn es keine entsprechende Software mehr gibt, die die Datei lesen kann, oder sie haben abweichende Dateiformatspezifikationen (nicht valide Dateien).

Die Dateiformate im DLZA werden jährlich im Hinblick auf aktuelle Formatentwicklungen geprüft, und es werden mögliche Erhaltungsmassnahmen mit Prioritäten für jedes Format definiert.

Risiken	Massnahmen
Datei entspricht nicht den Formatspezifikationen (nicht valide)	Konvertierung der Datei ins gleiche Dateiformat, das den Spezifikationen entspricht und valide ist
PDF-Dateien: Nicht eingebettete Schriften, Probleme mit Transparenz	Umwandlung in PDF/A (mit Adobe Preflight oder manuell) mit Fokus auf Erhalt der signifikanten Eigenschaften. Die neu erstellte Datei ersetzt das nicht valide PDF.
Dateien mit Extension Mismatch	Die Dateien werden via DROID identifiziert (entweder während des Quality-Assurance-Workflows oder manuell) und vor dem Ingest in das korrekte Dateiformat konvertiert.
Beschädigte oder nicht vollständige Dateien	Abhängig von ihrer Archivwürdigkeit wird die Datei entweder gelöscht oder es wird eine Datenrettung unternommen, indem die Datei im Hex Editor geprüft und falls möglich in eine lesbare Form gebracht wird. Die Datenrettung kann sich auch auf den Erhalt der signifikanten Eigenschaften beschränken (z.B. durch Konvertierung in ein Nur-Textfile).
Passwortgeschützte Dateien	Die Dateien werden falls möglich neu angefordert oder entschlüsselt.
Dateien ohne MIME-Type	Der Medientyp kann mittels Hex Editor in der Signatur der Datei überprüft werden. Die Dateien

¹ Vgl. nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, 12.1 Bitstreamerhaltung. https://nestor.sub.uni-goettingen.de/handbuch/artikel/nestor_handbuch_artikel_163.pdf [Stand 28.11.2022].



	werden übernommen, sofern sie archivwürdig und nicht bereits in einer Repräsentation vorhanden sind.
Unbekannte Dateiformate	Die Dateien werden übernommen, sofern sie archivwürdig und nicht bereits in einer Repräsentation vorhanden sind. Die Lesbarkeit kann nicht garantiert werden.
Dateien mit fehlenden oder falschen technischen Metadaten im DLZA (z.B. keine oder andere PUID)	Die Metadaten werden nachträglich mittels einer Preservation Action im Repository ergänzt.

Kontakt

Inge Moser, UZH Archiv, Universität Zürich

Tel. +41 44 634 23 21

E-Mail: inge.moser@archiv.uzh.ch

www.archiv.uzh.ch