



KANTON
NIDWALDEN

STAATSKANZLEI

STAATSARCHIV

Stansstaderstrasse 54, Postfach 1251, 6371 Stans
Telefon 041 618 51 51, www.nw.ch

RICHTLINIEN FÜR DIE ARCHIVIERUNG IM ELEKTRONISCHEN ARCHIV ELAR NW

NW-#418089-v6-Richtlinien_ELAR_NW.docx

Version	Datum	Status	Autor/in	Anmerkung
0.8	09.10.2019	Abgelöst	cb	Ersterstellung für Implementierungsphase 2019
0.9	13.03.2020	Abgelöst	cb	Nachführung nach Implementierungsphase
1.0	09.09.2020	Abgelöst	cb, ew	Nachführung und Genehmigung nach Review durch KOST
1.1	30.04.2021	Abgelöst	cb	Anpassung: DIP-Generierung mit CMI AIS (Kap. 3.4.2)
1.2	08.04.2022	Gültig	cb	Dokumentation neue Workflows, Anpassung Bewertung Tabellenformate

Inhalt

1	Einleitung	5
2	Grundlagen des elektronischen Archivs ELAR NW	6
2.1	Konzeptionelle Grundlagen.....	6
2.2	Systemarchitektur, Softwarekomponenten.....	6
2.3	Systemrollen.....	7
2.4	Prozesse der elektronischen Archivierung	7
2.5	Systemdokumentation	8
3	Anleitung für die elektronische Archivierung	9
3.1	Pre-Ingest.....	9
3.1.1	Akzession erfassen.....	9
3.1.2	Daten bereinigen	9
3.1.3	Analyse mit DROID.....	9
3.1.4	Einsatz von SIARD	9
3.1.5	SIP-Erstellung.....	10
3.2	Ingest	11
3.2.1	Bearbeitung auf der ELAR-Workbench	11
3.2.2	Ingest durchführen.....	12
3.2.3	Spezialfälle bei den Workflows	14
3.2.4	Fehlermeldungen bei Workflows.....	15
3.2.5	Anbindung ans AIS	15
3.2.6	Datenbereinigung nach dem Ingest	16
3.3	Management / Preservation Planning	16
3.4	Datenzugriff / DIP-Generierung	17
3.4.1	DIP-Generierung mit Docuteam Feeder.....	17
3.4.2	DIP-Generierung mit CMI AIS.....	18
4	Allgemeine Grundsätze zur elektronischen Archivierung	19
5	Richtlinien zum Pre-Ingest	20
5.1	Ablieferungsprototypen.....	20
5.2	Dateinamen	20
5.3	Fileablagen	21
5.4	Relationale Datenbanken	21
5.5	Erstellung der SIP.....	21
5.5.1	Benennung der SIP	21
5.5.2	Granularität der SIP	21
5.5.3	Erschließungstiefe in den SIP	21
5.5.4	Minimales Metadaten-set der SIP	22
6	Richtlinien zum Ingest	23
6.1	Archivtaugliche Dateiformate	23
6.2	Archivische Verzeichnung im Archivinformationssystem.....	24
6.2.1	Eingliederung in die Archivstruktur.....	24
6.2.2	Ergänzung von deskriptiven Metadaten im AIS nach dem Import	24
6.2.3	Erstellung von Nutzungskopien	24
7	Richtlinien zu Management und Preservation Planning	25
7.1	Statistik, Planung.....	25

7.2	Beobachtung der Dateiformate	25
8	Richtlinien zum Zugang	26
8.1	Interner Zugang ins Repository	26
8.2	Einsichtnahme in Dokumente	26

1 Einleitung

Die vorliegenden Richtlinien für das elektronische Archiv ELAR dienen verschiedenen Zwecken: Sie dokumentieren die Komponenten des ELAR im Staatsarchiv Nidwalden, sie enthalten die verbindlichen Richtlinien für die elektronische Archivierung und sie sind konkrete Handlungsanleitung für die Arbeitsabläufe im ELAR.

Die Richtlinien ELAR sollen eine einheitliche Praxis in der elektronischen Archivierung, insbesondere im Ingest, sicherstellen. Die aus einer einheitlichen Handhabung resultierende homogene Datenstruktur und Datenqualität erhöht die Zuverlässigkeit und Berechenbarkeit der Überlieferung für die Benutzer. Die Richtlinien sind deshalb verbindlich.

Die Richtlinien zur Bewertung von Archivgut¹ und die Richtlinien zur Erschliessung von Archivgut² bleiben zentrale Grundlagen auch für den digitalen Erschliessungsprozess – sie gelten auch für die Archivierung von digitalen Ablieferungen. Die vorliegenden Richtlinien für das elektronische Archiv ELAR ergänzen die Bewertungs- und Erschliessungsrichtlinien in Bezug auf die elektronische Archivierung.

Da das elektronische Archiv über einen längeren Zeitraum schrittweise in gestaffelten Projekten aufgebaut wird (vgl. Projektplan ELAR), umfassen die vorliegenden Richtlinien vorderhand noch nicht alle Aspekte der elektronischen Archivierung. Sie umfassen vorerst insbesondere den Übernahmeprozess (Ingest). Die übrigen Prozesse der elektronischen Archivierung wie die Erhaltungsplanung (Preservation Planning) und der Zugang zur Benutzung werden soweit heute bereits möglich skizziert, sie sind später auszuformulieren. Die Richtlinien ELAR bestehen aus drei Teilen: Im ersten Teil werden kurz die Systemgrundlagen beschrieben und die entsprechenden Dokumentationen angeführt (Kap. 2). Der zweite Teil enthält eine Übersicht über die Prozesse und Arbeitsschritte sowie konkrete Schritt-für-Schritt-Anleitungen für die Arbeit im ELAR, vorderhand insbesondere für den Ingest-Prozess (Kap. 3). Der dritte Teil enthält die geltenden Richtlinien, die bei der Arbeit zu beachten sind, auch hier vorderhand insbesondere für den Ingest. (Kap. 4 bis 8).

Das ELAR-Basissystem wurde 2018 zusammen mit dem Staatsarchiv Obwalden im ILZ aufgebaut. Die vorliegenden Richtlinien wurden für Nidwalden 2019 in Version 0.9 verabschiedet und dienten als Grundlage für die Implementierungsphase im Jahr 2019. Nach Abschluss der Implementierungsphase wurden die Richtlinien ELAR aufgrund der Rückmeldungen der Mitarbeitenden und eines Reviews durch die Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen KOST überarbeitet und als Version 1.0 in Kraft gesetzt.

Normen und Standards der elektronischen Archivierung unterliegen (immer noch) einem stetigen Wandel. Es ist davon auszugehen, dass an diesem Dokument in regelmässigen Abständen auch weiterhin Anpassungen vorgenommen werden müssen.

¹ Dok. #386029.

² Dok. #343983.

2 Grundlagen des elektronischen Archivs ELAR NW

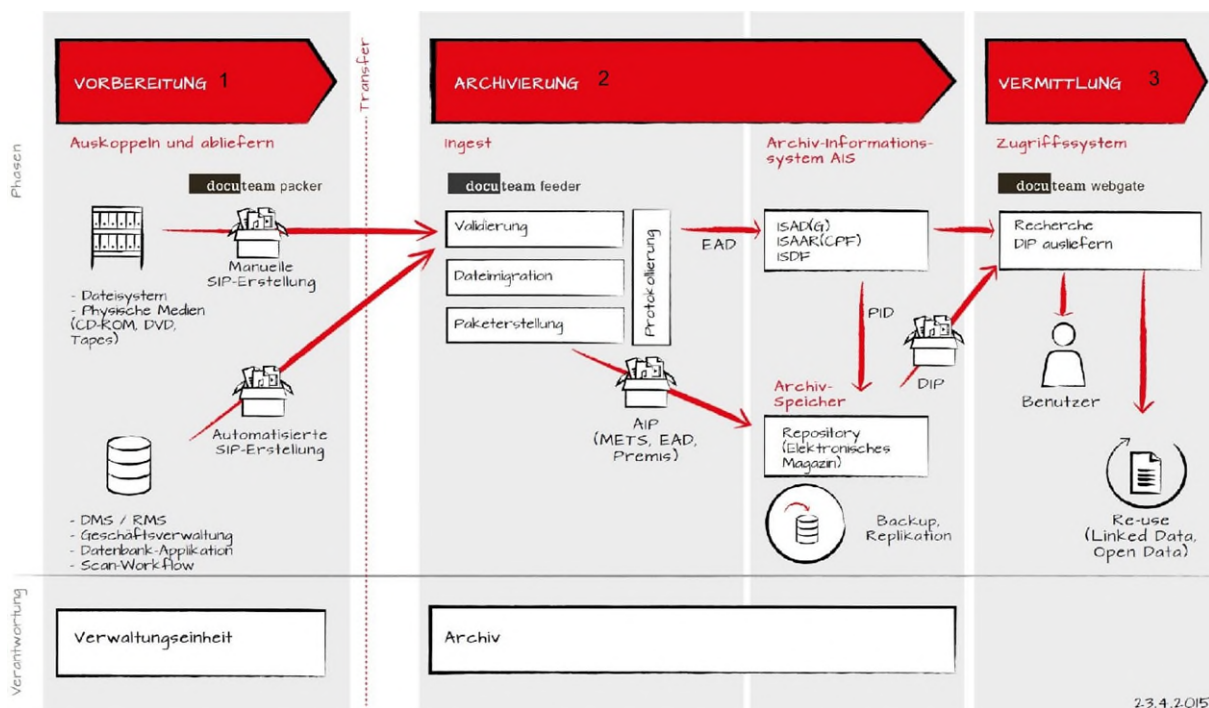
2.1 Konzeptionelle Grundlagen

Das elektronische Archiv ELAR basiert grundsätzlich auf dem OAIS-Referenzmodell³ und auf den Minimalanforderungen der Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen KOST an ein digitales Langzeitarchiv⁴.

Da das elektronische Archiv über einen längeren Zeitraum schrittweise in gestaffelten Projekten aufgebaut wird (vgl. Projektplan ELAR), sind momentan noch nicht alle Grundanforderungen vollständig erfüllt. Dies wird in späteren Teilprojekten nachgeholt.

2.2 Systemarchitektur, Softwarekomponenten

Die Prozesse der elektronischen Archivierung, der Systemaufbau und die eingesetzten Komponenten lassen sich wie folgt schematisch illustrieren.



Vorbereitung (Pre-Ingest)

- **Docuteam Packer:** bereitet Ablieferungen aus Dateisystemen für die Übernahme vor. Digitale Ablieferungspakete (SIP = Submission Information Packages) werden erzeugt.
- **DROID (Digital Record and Object Identification):** Open Source-Tool zur automatischen Identifikation von Dateiformaten und Doppelerkennung von Dateien. DROID wird als Hilfsmittel für die Bewertung eingesetzt.
- **SIARD Suite:** SIARD (Software Independent Archiving of Relational Databases) ist eine vom Bundesarchiv entwickelte Software zur Archivierung von relationalen Datenbanken.

Übernahme (Ingest) und Archiv (Repository)

- **Docuteam Feeder:** serverbasiertes System für den digitalen Ingest-Prozess mit Übernahme, Validierung / Verifizierung, Migration und Metadaten-Anreicherung von Daten in das elektronische Magazin (Repository). Im Repository werden aus den Ablieferungspaketen (SIP) Archivpakete (AIP = Archival Information Packages).

³ ISO 14721:2012.

⁴ https://kost-ceco.ch/cms/minimal_specifications_de.html.

- **Workbench:** Fileshare auf dem Applikationslaufwerk "L". Dient als Zwischenablage.
- **Fedora Commons (Repository):** Zur Verwaltung des digitalen Archivs wird die Repository-Software Fedora Commons verwendet. Sie steuert Aktionen im Archivspeicher und überwacht die Integrität der Daten.
- **CMI AIS:** Für die Verwaltung der deskriptiven Metadaten wird das Archivinformationssystem (AIS) CMI AIS verwendet. Für den Metadatenimport in CMI AIS wird die Schnittstelle Docuteam-CMI AIS verwendet.

Vermittlung und Zugriff

- **CMI AIS:** Abweichend von der Grafik erfolgt der Zugriff auf die digitalen Dokumente nicht über Docuteam Webgate sondern direkt über CMI AIS oder mittels definiertem Workflow in **Docuteam Feeder**. Archivmitarbeitende generieren aus einem AIP ein Auslieferungspaket (DIP = Dissemination Information Package), das dem Benutzer zur Verfügung gestellt werden kann.

2.3 Systemrollen

Für die Benutzung der ELAR-Infrastruktur werden im Staatsarchiv Nidwalden zwei funktionale Rollen definiert: ELAR Admin und ELAR Benutzer.

ELAR Admin (Administratorenrechte)

- Mitglieder: Staatsarchivar, Wissenschaftliche Mitarbeitende
- Aufgaben: Benutzer- und Prozessverwaltung
- Administratorenrechte in Docuteam Feeder

ELAR Benutzer (Superuserrechte)

- Mitglieder: Alle Mitarbeitenden
- Aufgaben: Analyse von Fileablagen, SIP erstellen, Ingests durchführen, Metadaten in AIS importieren, Fehleranalyse und -Bereinigung
- Docuteam Packer Installation
- DROID Installation
- SIARD Installation
- Zugriff auf ELAR Workbench
- Zugang zu Docuteam Feeder
- Benutzerrechte in Docuteam Feeder
- Zugang zu Fedora (DIP Generierung).

2.4 Prozesse der elektronischen Archivierung

Das elektronische Archiv ELAR umfasst vier fachliche Kernprozesse (vgl. Grafik in Kap. 2.2), übergeordnet laufen Prozesse des Gesamt-Systembetriebs.

1. **Systembetrieb:** Übergeordnete Planungs- und Koordinationsprozesse zwischen den Staatsarchiven und dem ILZ zur Pflege und Weiterentwicklung des ELAR
2. **Pre-Ingest:** Vorbereitung der Archivierung, Aufbereitung der Daten
3. **Ingest:** Übernahme der bereinigten Daten ins Repository und Übernahme der Verzeichnismetadaten ins AIS bzw. Anlegen der Verzeichnismetadaten im AIS
4. **Management, Preservation Planning:** Verwaltung und Kontrolle der archivierten Daten sowie Planung und Durchführung von Erhaltungsmaßnahmen
5. **Zugang (Vermittlung):** Ausziehen von Daten aus dem Repository zur Einsicht / Nutzung

Systembetrieb

Der Betrieb des elektronischen Archivs wird in einer gemeinsamen Betriebsorganisation der beiden Staatsarchive Obwalden und Nidwalden sowie des ILZ gewährleistet. Die Betriebsorganisation sieht zwei Gremien vor: den Betriebsausschuss als strategische Entscheidungsebene und die Arbeitsgruppe als operative Betriebsleitung. Dazu kommen als Rollen der technische

Systembetrieb im ILZ und die ELAR-Verantwortlichen als Fachverantwortliche in den Staatsarchiven.

Pre-Ingest

Der Pre-Ingest umfasst einerseits die herkömmlichen Prozesse der Akzession (Datenübernahme, Dokumentation der Akzession). Zusätzlich kommt im ELAR-Prozess der Bereinigung der übernommenen elektronischen Daten eine grosse Bedeutung. Die Datenbereinigung umfasst die inhaltliche Bewertung der Dateien, das Ausscheiden von Kopien und Vorversionen, die technische Kontrolle (Dateiformate, geschützte / gesperrte Dateien) und die Herstellung einer ISAD(G)-konformen Struktur. Ziel des Pre-Ingests ist es, möglichst authentische, einheitliche und strukturierte Daten und Metadaten für den Ingest bereitzustellen.

Die Bereinigung der Daten ist vorderhand besonders wichtig und aufwändig, weil elektronische Daten vor allem aus unstrukturierten oder wenig strukturierten Ablagen übernommen werden. Mit dem zunehmenden Einsatz von strukturierten Fachanwendungen sollte der Aufwand für den Pre-Ingest abnehmen.

Wenn möglich, soll die Provenienzstelle die Daten aufbereiten. In der Praxis wird die Datenaufbereitung (bis auf Weiteres) wohl vor allem im Staatsarchiv geschehen. Dennoch ist die Provenienzstelle insbesondere in die Bewertung miteinzubeziehen. Hilfsmittel ist das [Merkblatt für die Kantonsverwaltung zu Ablieferungen aus Papierablagen und aus digitalen Filesystemen](#) mit der Positivliste für die Bewertung.

Ingest

Im Ingest werden die Daten aus dem Pre-Ingest archiviert. Dies umfasst einerseits die Aufbereitung (Authentizitäts- und Virenprüfung, Formatumwandlung) und Übernahme der Daten ins Repository, andererseits die Übernahme der Verzeichnismetadaten ins AIS bzw. das Anlegen von Verzeichnismetadaten im AIS. Abschliessend werden Daten und Verzeichniseinheit eindeutig miteinander verbunden und ggf. Nutzungskopien im AIS hinterlegt.

Management, Preservation Planning

Management und Preservation Planning umfassen hauptsächlich bereits bekannte Prozesse der Erhaltung und wenden sie auf digitale Daten an: regelmässiger Review der Datensicherung (Konzept und Hardware), periodische Kontrolle der Datenintegrität, Monitoring der Dateiformate und allfällige Planung und Durchführung von Formatmigrationen sowie die Bewirtschaftung des Repositories. Im Bereich der Verzeichnismetadaten finden regelmässige Prozesse der Qualitätskontrolle, der Publikation im Online-Verzeichnis sowie allfällige Datenbereinigungen statt. Diese Prozesse sind vorläufig erst summarisch geregelt und werden später ausdifferenziert.

Zugang (Vermittlung)

Der Zugang umfasst die Abfrage von Metadaten im AIS (Suche) sowie die Bereitstellung der zugehörigen Daten (Primär- und / oder Nutzungsdaten) für die Benutzenden. Diese Prozesse sind vorläufig erst summarisch geregelt und werden später ausdifferenziert.

2.5 Systemdokumentation

Das elektronische Archiv ELAR ist in den folgenden Dokumenten verbindlich beschrieben:

- Betriebshandbuch: Betriebsorganisation Obwalden, Nidwalden, ILZ⁵
- Richtlinien ELAR: Prozessrichtlinien für das ELAR im Staatsarchiv Nidwalden
- Technische Systemdokumentation des ILZ⁶
- Wartungsvertrag mit der Docuteam GmbH⁷.

⁵ Dok. #404794.

⁶ Dok. #473376

⁷ Dok. #596375

3 Anleitung für die elektronische Archivierung

Dieses Kapitel bietet eine Schritt-für-Schritt-Anleitung für einen Musterprozess von der Akzession, über den Ingest bis zum Import ins AIS. Es dient als Handlungsanleitung für die Bearbeitung einfacher elektronischer Ablieferungen. Komplexere Ablieferungen, Fehler und Probleme sind mit dem Fachverantwortlichen ELAR (FVELAR) und / oder der für die Erschliessung fachverantwortlichen Archivarin (FVE) zu bearbeiten bzw. zu bereinigen.

3.1 Pre-Ingest

3.1.1 Akzession erfassen

Die Ablieferung wird im AIS (CMI AIS) als Akzession erfasst und erhält dadurch eine eindeutige Akzessionsnummer.

- Die Primärdaten der elektronischen Ablieferung werden pro Akzession in einen Ordner in Laufwerk "R:\Akzessionen ELAR Originaldaten" abgelegt. Der Ordner wird mit Akzessionsnummer beschriftet (bspw.: "2018-19").

3.1.2 Daten bereinigen

- Die Daten sind als erstes zu bewerten (vgl. Bewertungsrichtlinien).
- Die Daten werden aus dem Laufwerk "R:\Akzessionen ELAR Originaldaten" kopiert und unter "R:\Akzessionen ELAR Arbeitskopie" abgelegt. Die Bearbeitung findet ausschliesslich im Arbeitskopie-Ordner statt, die Originaldaten dienen als Sicherungskopie.
- Dokumente, die als archivwürdig taxiert werden, werden nicht umbenannt. Sie behalten den Originaltitel.
- Das vorbereitete Paket wird umbenannt. Der neue Name entspricht dem Titel der neu anzulegenden Verzeichniseinheit in CMI AIS.
- Die Arbeitsschritte und die Bewertung des Pre-Ingests sind direkt im AIS zu dokumentieren (vgl. Erschliessungsrichtlinien).

3.1.3 Analyse mit DROID

In Absprache mit der FVE und dem FVELAR werden die Daten ggf. unter Einsatz von DROID analysiert, falls nötig migriert, bewertet und neu strukturiert (ISAD(G)-konforme Struktur). Eingliederung, Bewertung und Strukturierung richten sich grundsätzlich nach den vorhandenen Richtlinien zur Bewertung und Erschliessung.

DROID ist ein Tool, mit welchem Fileablagen auf die darin befindlichen Dateiformate analysiert werden können. Dies dient hier insbesondere der Doppelerkennung mittels MD5-Hashwert. DROID hat eine ausführliche und gute Anleitung zu seinen Funktionen, die über den Reiter "Help" abrufbar ist.

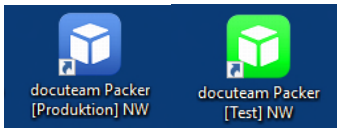
3.1.4 Einsatz von SIARD

SIARD (Software Independent Archiving of Relational Databases) ist eine vom Bundesarchiv entwickelte Software zur Archivierung von relationalen Datenbanken. Sie bildet die Beziehungen zwischen den Tabellen einer relationalen Datenbank in XML ab und ermöglicht es die einzelnen Datensätze zu einem späteren Zeitpunkt anhand der XML-Datei wieder richtig miteinander zu verknüpfen. Das Verfahren ist allerdings sehr komplex, zudem sind im SIARD-Format archivierte Daten nicht schnell und unkompliziert wieder benutzbar. Das SIARD-Format eignet sich deshalb aus Sicht des Staatsarchivs nur bedingt als Format für die Archivierung.

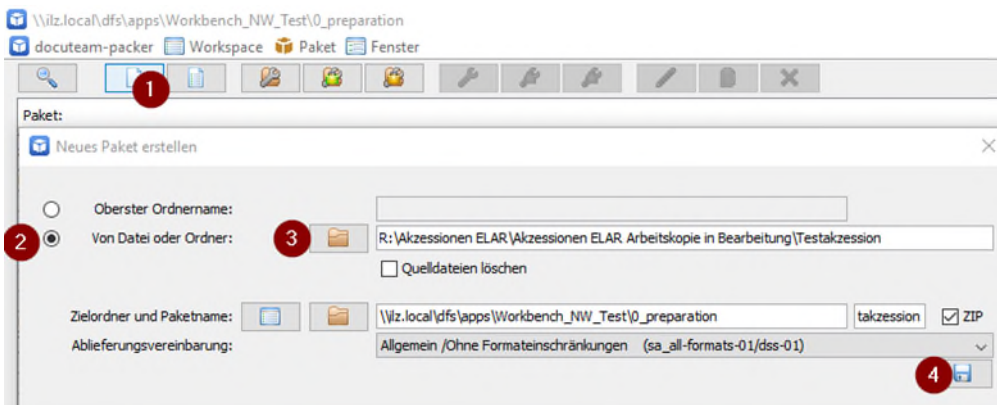
Bei der Archivierung von Datenbanken ist zu überlegen, ob anstatt oder nicht zumindest zusätzlich zur ganzen Datenbank in SIARD nur Reports der als archivwürdig bewerteten Daten aus der Datenbank zu archivieren sind. Der FVELAR entscheidet dies, ggf. in Absprache mit der FVE.

3.1.5 SIP-Erstellung

Mit Doppelklick auf Link auf Desktop Docuteam Packer öffnen:

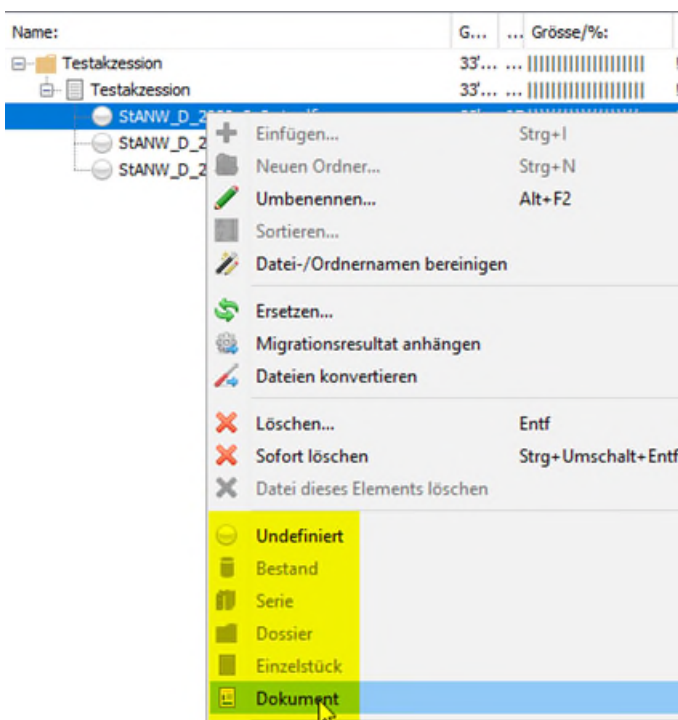


Neues Paket erstellen:



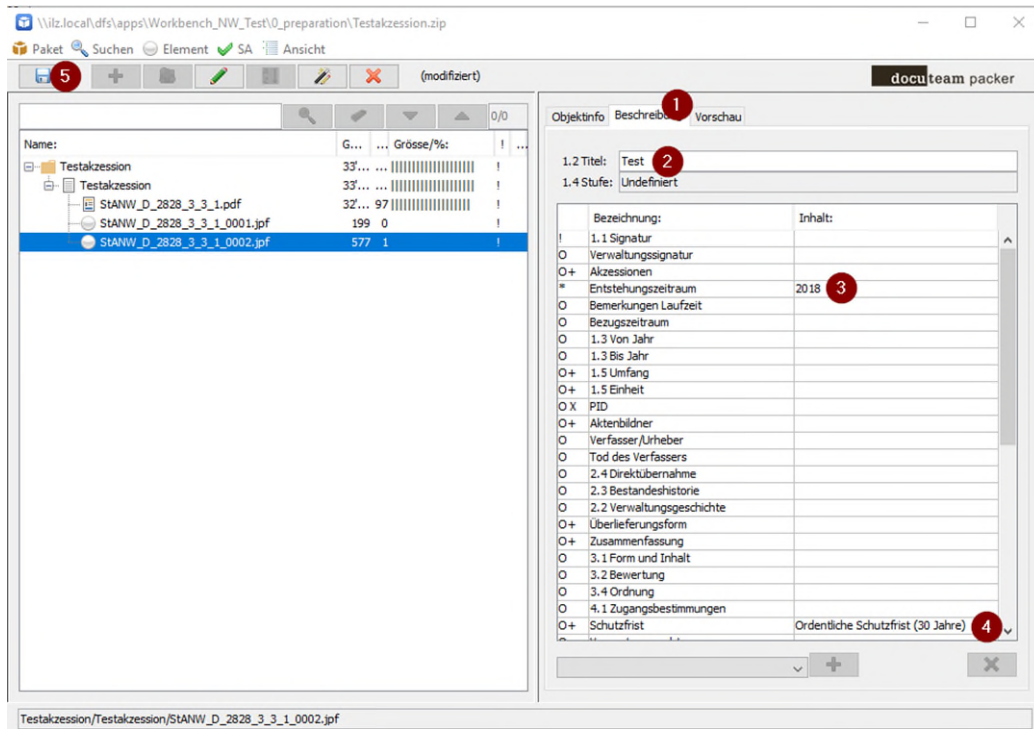
Erschliessungsstufe in Absprache mit FVE zuteilen:

Rechtsklick auf Objekte > Erschliessungsstufe zuteilen ("Bestand", "Serie", "Dossier", "Einzelstück", "Undefiniert"). Ggf. Nutzungskopien/Verzeichnisse als "Dokument" auszeichnen.



Metadaten vergeben, Paket speichern:

Pflichtfelder für die Metadatenvergabe sind **Titel**, **Entstehungszeitraum** und **Schutzfrist**. Diese sind zwingend abzufüllen. Weitere Felder können ggf. zusätzlich abgefüllt werden.



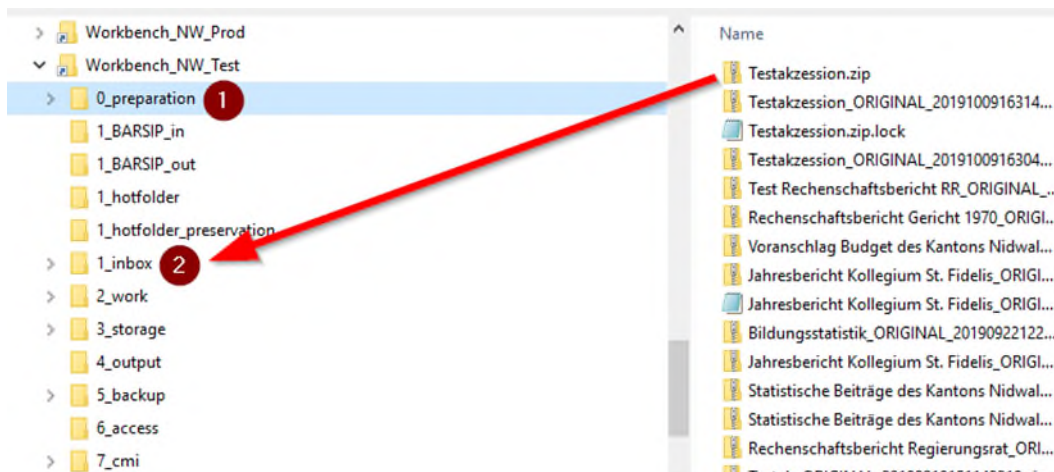
Paket wird geschrieben; Fehlermeldung "Es gibt Elemente mit nicht ausgefüllten Pflichtfeldern" ignorieren.

3.2 Ingest

3.2.1 Bearbeitung auf der ELAR-Workbench

Das mit Dokuteam Packer erstellte Ablieferungspaket wird automatisch in der sog. Workbench in Laufwerk L gespeichert. Die ELAR-Workbench dient als Zwischenablage für sämtliche ELAR-Prozesse. Es gibt eine Workbench für das Testsystem und eine für das Produktivsystem.

Paket (zip-Ordner) in "Workbench_NW_Prod" resp. "Workbench_NW_Test" aus "0_preparation" (1) nach "1_inbox" (2) verschieben. zip-Ordner "[...]_ORIGINAL_[...]" sowie lock-File "[...]_zip.lock" stehen lassen.



Die Workbench ist in verschiedene durch Docuteam vorgegebene Ordner aufgeteilt, welche einzelne Schritte des ELAR-Prozesses repräsentieren:

- **0_preparation:** SIP werden nach ihrer Erstellung mit Docuteam Packer automatisch in diesem Ordner abgelegt.
- **1_BARSIP_in:** Ablageort für eCH-0160-Pakete im "unbehandelten" eCH-0160-Standard.
- **1_BARSIP_out:** Ablageort der mit dem Feeder-Workflow "BARSIP Converter" in METS umgewandelten eCH-0160-Pakete.
- **1_hotfolder:** Zwischenspeicher für gewisse Workflows. Die entsprechenden Workflows werden im Moment nicht verwendet.
- **1_hotfolder_preservation:** Output-Ordner für den Workflow "Preservation: Deliver DIP by PUID". Im Produktivsystem wird der Inhalt des Ordners automatisch vom Workflow "Preservation: Update Fedora Objects" angesteuert und die entsprechenden Dateien in einem neuen Format als Version zum bereits bestehenden Fedora-Objekt dazugeschrieben.
- **1_inbox:** Ordner aus dem die SIP nach der Vorbereitung mit dem Packer vom Feeder angesteuert und die Verarbeitung der SIP gestartet wird. SIP müssen nach der Erstellung und Abspeicherung im "0_preparation"-Folder manuell in den "1_inbox"-Ordner verschoben werden. Erst dann können sie über Docuteam Feeder angesteuert werden.
- **2_work:** Zwischenspeicher für vom Feeder in Arbeit befindliche SIP. Kann ein Workflow nicht erfolgreich beendet werden, liegen hier Daten drin, die nur über den Workflow "Cleanup" entfernt werden können.
- **3_storage:** Hier werden zu jedem erfolgreich verarbeiteten SIP die PID in einem txt.-File abgespeichert. Können nicht alle Dateien in einem SIP in Fedora gespeichert werden, liegen hier auch Kopien der Dateien, die nicht in Fedora gespeichert werden konnten.
- **4_output:** Wird im Moment nicht verwendet.
- **5_backup:** Hier wird nach erfolgreichem Ingest automatisch eine Kopie des SIP abgelegt.
- **6_access:** Im Ordner "6_Access" werden die mit dem Feeder mit dem Workflow "Access: Deliver DIP by PID" generierten DIP abgespeichert.
- **7_cmi:** Im Ordner "7_cmi" werden nach einem erfolgreichen Ingest die deskriptiven Metadaten aus einem SIP als EAD-xml ausgegeben. Aus diesem Ordner erfolgt die Weiterverarbeitung mittels Schnittstelle Docuteam/CMI AIS.
- **8_sip_aus_cmi:** Hier werden SIPs aus Axioma (Lifecyclemodul Ablieferung) abgelegt. Die Weiterverarbeitung erfolgt mit Docuteam Feeder.
- **20_Analyse Docuteam:** Standard-Ordner für Docuteam; wird von uns nicht verwendet.

3.2.2 Ingest durchführen

Der Ingest erfolgt im Webdienst Docuteam Feeder.

Der Ingest-Prozess benötigt viel Zeit und absorbiert grosse Rechenleistung. Deshalb kann es sich lohnen, diesen über Nacht laufen zu lassen. Sperrzeiten für den Ingest-Prozess sind die Service-Fenster des ILZ am zweiten und vierten Mittwoch im Monat ab 17 Uhr. An diesen Abenden dürfen keine Ingest-Prozesse laufen!

Workflow starten

Unter dem Reiter "Workflows" (1), Standard-Workflow "Quality Assurance, Migration and Storage" starten (2), im Dropdown-Menü gewünschtes SIP auswählen (3), "Ausführung erstellen" (4) anklicken:

docu team feeder NW ▾ Ausführungen Workflows **1** Cockpit Benutzer Admin ▾

Workflows

ID	Name	starten	verwalten
12	Administration		
13	Run		
14	Quality Assurance		
15	Check SA and Migration		
16	Quality Assurance, Migration and Storage 2		
17	Quality Assurance and Storage		

Neue Ausführung für Quality Assurance, Migration and Storage

Objekt-ID SIP 3

oder

Manuelle Eingabe

4

Der Workflow startet, alle Schritte müssen vom System mit dem Status: **"beendet mit Code 0"** quittiert werden. Ansonsten liegt ein Fehler vor, der mit dem FVELAR analysiert werden muss.

Die im Feeder vorhandenen Workflows wurden von Docuteam eingerichtet. Der Administrator kann sie bearbeiten oder eigene Workflows hinzufügen. Die Feeder-Workflows und ihre genauen Funktionen sind auf dem Docuteam-Wiki ausführlich beschrieben.⁸ Hier folgen nur Kurzbeschreibungen der im Einsatz stehenden Workflows:

ID	Workflow	Beschreibung Verwendung NW
12	Administration	Wird nicht verwendet
13	Run	Wird nicht verwendet
14	Quality Assurance	Wird nicht verwendet
15	Check SA and Migration	Wird nicht verwendet
16	Quality Assurance, Migration and Storage	Standardworkflow um SIP, die auch nicht archivtaugliche Dateiformate enthalten, ins Repository zu speichern. > Führt Formatmigration durch.
17	Quality Assurance and Storage	Standardworkflow um SIP, die ausschliesslich archivtaugliche Dateiformate enthalten, ins Repository zu speichern. > führt keine Formatmigration durch.
18	Cleanup	Entfernt fehlgeschlagene Workflows und löscht die steckengebliebenen SIP aus dem Ordner "2_work".
19	Check Checksums	Wird nicht verwendet
20	Preservation: Deliver DIP by PUID	Wird zusammen mit Workflow 21 verwendet. Holt ein spezifisches Dateiformat aus dem Repository und speichert sämtliche Dateien in "1_hotfolder_preservation".

⁸ https://wiki.docuteam.ch/doku.php?id=docuteam:feeder-steps_340.

21	Preservation: Update Fedora Objects	Holt sämtliche Dateien aus "1_hotfolder_preservation" und führt an ihnen eine Formatmigration in ein bei der File Migration definiertes Format durch. Speichert die Files als neue Version wieder im Repository.
22	Access: Deliver DIP by PID	Wird zur Erzeugung von DIP mit dem Feeder verwendet.
25	BARSIP Converter	Wandelt eCH-0160-Pakete in "1_BARSIP_in" in METS Pakete um und speichert diese in "2_BARSIP_out". Pakete müssen dort mit dem Packer geöffnet und per "Save As" als .zip-Container in "1_inbox" gespeichert werden (vgl. unten).
27	Submission: Create-SIPFromExcel	Der Workflow erlaubt den Import von deskriptiven Metadaten aus einem Excel in ein SIP nach bestimmten Regeln (vgl. unten).
29	Quality Assurance, Migration and Storage without virus check	Gleicher Workflow wie der Standardworkflow 16 aber ohne im Ingest einen Viruscheck durchzuführen. Der Workflow ist vor allem für grosse Video-/Filmdateien vorgesehen, die denen der Viruscheck eine Fehlermeldung produziert.
31	Purge Fedora Objects	Workflow zur gezielten Löschung von AIP (PID und der damit verbundenen Daten) aus Fedora (vgl. unten).
35	Quality Assurance, Migration and Storage (Originale behalten)	Standardworkflow um SIP, die auch nicht archivtaugliche Dateiformate enthalten, ins Repository zu speichern. > Führt Formatmigration durch/Microsoft Office-Dokumente werden zusätzlich im Originalformat belassen.
37	Quality Assurance, Migration and Storage without virus check (Originale behalten)	Standardworkflow um SIP, die auch nicht archivtaugliche Dateiformate enthalten, ins Repository zu speichern ohne im Ingest einen Viruscheck durchzuführen. > Führt Formatmigration durch/Microsoft Office-Dokumente werden zusätzlich im Originalformat belassen.

3.2.3 Spezialfälle bei den Workflows

BARSIP Converter (25)

- e-CH-0160 Paket in Workbench "1_BARSIP_in" ablegen
- Workflow "25 BARSIP Converter" öffnen
- SIP-Name manuell eingeben bzw. kopieren inklusive Dateiendung .zip
- SIP landet in Workbench-Ordner "1_BARSIP_out"
- SIP mit Docuteam-Packer öffnen und mit "SaveAs" als "zip container" abspeichern oder selbst einfach ein .zip daraus machen (auch wenn am Ende des Dateinamens schon zip steht, es muss ein .zip sein)
- Dieses .zip in die Workbench "1_inbox" kopieren und mit dem Feeder verarbeiten.

Submission: CreateSIPFromExcel (27)

- Fileablage mit DROID öffnen ("add" dann "start")
- Export als CSV -> Import in Excel -> liefert die für den Workflow benötigten Dateipfade
- Excel bearbeiten. Zwingend benötigte Spalten: "path" (Dateipfad); "levelOfDescription" (Stufung in Bestand, Serie usw.) und "unittitle" (Dateinamen)
- Als weitere Spalten sind nur die im levels.xml für die entsprechende Stufe vorgesehenen Metadatenfelder erlaubt
- Gewünschte Spalten im Excel von Hand ergänzen
- Excel-File und Daten für das Paket in Workbench Ordner "0_preparation" (nicht "1_inbox") ablegen (erlaubt Angabe von Pfaden ohne Laufwerk im Excel) und Workflow "CreateSIPFromExcel" starten (Dateinamen ohne oder inklusive Dateiendung .xlsx manuell eingeben).

Purge Fedora Objects (31)

Der Workflow löscht in Fedora vorhandene PID und die dazugehörigen Daten. Die PID werden nicht erneut vergeben und bleiben leer. Dieser Prozess wird nur durch den FVELAR ausgelöst.

- TXT-File mit den aus Fedora zu löschenden PIDs aus dem Workbench Ordner "3_storage" kopieren
- TXT-File mit Excel öffnen (Datenimport) und die zu löschenden PIDs in ein neues Textfile kopieren, Textfile benennen (z.B. PIDsLöschen.txt)
- In die letzte Zeile des Textfiles mit den zu löschenden PIDs ein beliebiges Wort, eine beliebige Buchstabenkombination schreiben (z.B. WegDamit)
- TXT-File in den Ordner "1_inbox" der Workbench ablegen
- Docuteam-Feeder starten und Workflow "28 Purge Fedora Objects" auswählen
- Manuell ins Feld den Namen des Textfiles sowie nach einem Leerschlag das darin auf der letzten Zeile enthaltene Wort schreiben um den Workflow starten zu können (z.B. PIDsLöschen.txt WegDamit)
- Workflow starten und **immer** die vom Feeder ausgeworfene Antwort anschauen. Code 0 (erfolgreiche Durchführung des Workflows) wird beim Purge Workflow auch angezeigt, wenn der Workflow nichts gelöscht hat oder eine Fehlermeldung produziert.

3.2.4 Fehlermeldungen bei Workflows

Kann ein Workflow / Ingest eines SIP nicht vollständig durchgeführt werden, wird vom Feeder angezeigt, bei welchem Schritt der Ingest abgebrochen wurde. Über das Feld "Antwort" lässt sich die ausführliche Fehlermeldung des Schritts anzeigen, die anschliessend analysiert werden muss. Aus der Fehlermeldung lassen sich die problemverursachenden Dateien eruieren. Hilfreich ist die Suche nach den Stichwörtern "warn", "error" oder "invocation".

Die Analyse der vom Feeder produzierten Fehlermeldungen braucht Erfahrung. Sie erfolgt grundsätzlich durch den FVELAR.

Kann nur schon eine Datei nicht migriert oder in Fedora geschrieben werden, bricht der ganze Ingest ab. In ein bereits zum Teil durch die Workflows gelaufenes SIP kann nicht direkt eingegriffen werden. Die Problemdateien müssen aus dem SIP entfernt (oder von Hand bearbeitet) und der Ingest von vorne neu gestartet werden (inklusive vorgängigem Workflow "Cleanup" für den gescheiterten Ingest).

Konnte der Fehler behoben werden, kann der Ingest über "Fehlgeschlagene Schritte neu starten" wieder angestossen werden.

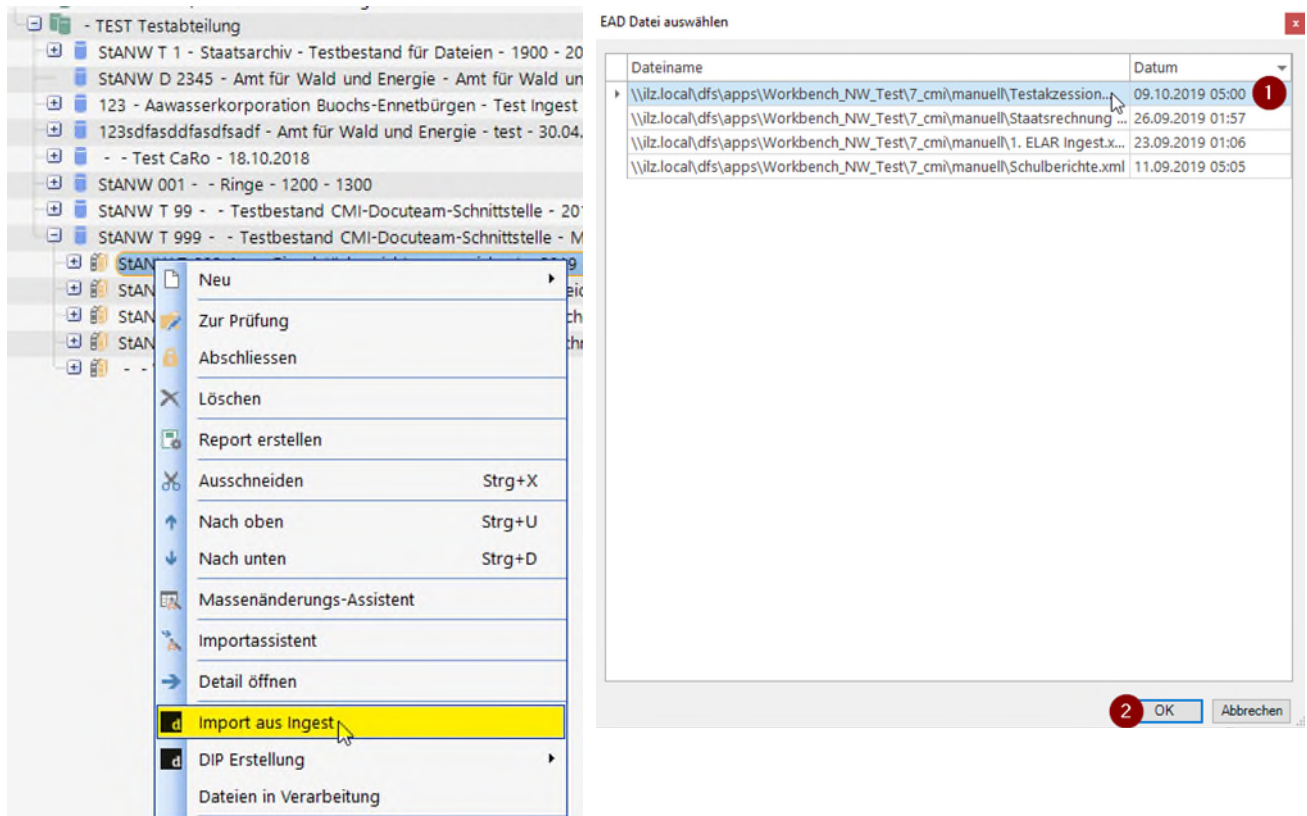
Fehlermeldung beim Schritt "transfer Fedora objects to repository":

Bricht ein Workflow beim Schritt "transfer Fedora objects to repository" mit "could not ingest file" ab, sind die Objekte, die transferiert werden konnten, bereits in Fedora gespeichert. Die bereits transferierten Objekte (Dateileichen) müssen mit dem Purge-Workflow aus Fedora gelöscht werden. Dieser Prozess wird nur durch den FVELAR ausgelöst.

3.2.5 Anbindung ans AIS

Mit dem in CMI AIS integrierten Service "Import aus Ingest" werden die Verzeichnungsstruktur und die deskriptiven Metadaten direkt ins AIS (CMI AIS) übernommen.

In CMI AIS im Baum an entsprechendem Ast mit Rechtsklick Reiter "Import aus Ingest" doppelklicken, EAD-Datei auswählen (1), Mit "OK" bestätigen (2):



Erledigungsmeldung erfolgt vom System per E-Mail.

Mittels Import werden nur bereits vorhandene Metadaten importiert. Alle anderen Pflichtfelder gemäss Erschliessungshandbuch müssen nach dem Import ergänzt werden (nach Möglichkeit mittels Massenänderungsassistent).

3.2.6 Datenbereinigung nach dem Ingest

Der erfolgreiche Abschluss des Ingest-Prozesses ist mittels Stichproben zu prüfen. Dazu werden einzelne DIP erzeugt (siehe: Kap. 3.4).

Anschliessend wird die Dateiablage im Laufwerk R sowie in der Workbench ELAR bereinigt:

- Die ursprünglichen Primärdaten in Laufwerk "R:\Akzessionen ELAR Arbeitskopie" und "R:\Akzessionen ELAR Originaldateien" werden gelöscht.
- Die in Workbench Ordner "5_backup" abgelegte Kopie des SIP wird bis zur Realisierung eines redundant vorhandenen Archivspeichers als Backup in "R:\Backup ELAR" verschoben.
- In einzelnen Workbench-Ordern bleiben Daten zum Teil liegen. Die Ordner werden kontrolliert und die Daten gelöscht.

3.3 Management / Preservation Planning

Management und Preservation Planning werden in Teilprojekt "ELAR Erhaltungsplanung" (frühestens ab 2023) ausdifferenziert.

Über den Reiter "Cockpit" im Feeder können die im Repository gespeicherten Daten überblickt und analysiert werden. Hier finden sich Angaben zur gespeicherten Datenmenge sowie Detailangaben zu den im Repository vorhandenen Dateiformaten (inklusive PRONOM-Identifizier).

3.4 Datenzugriff / DIP-Generierung

DIP sind aus AIP erstellte Nutzungskopien der archivierten Daten. Die Erstellung von DIP ermöglicht aber indirekt auch eine Kontrolle des Ingest-Prozesses. DIP lassen sich grundsätzlich auf folgende zwei Arten erzeugen.

3.4.1 DIP-Generierung mit Docuteam Feeder

Mit dem **Workflow "Access: Deliver DIP by PID"** im Docuteam Feeder können DIP wie folgt erzeugt werden:

PID aus CMI AIS kopieren:

Dossier:

Home

Identifikation

- Kontext
- Dateien (0)
- Inhalt und innere Ordnung
- Zugangs- und Benutzungsbedi...
- Sachverwandte Unterlagen
- Verzeichnungskontrolle
- Standort (0) und Umfang (0)

Signatur

ID

154638

Alte Archivsignatur

Titel

Test Rechenschaftsbericht RR

Verzeichnungsstufe

Dossier

PID

CH-000101-0:67149

Provenienzsignatur

Archivalienarten

PID in Docuteam Feeder in Feld "Manuelle Eingabe" (1) kopieren; "Ausführung erstellen" (2) anklicken:

Neue Ausführung für Access: Del... x +

ingestownwtest.ilz.local/organizations/2/workflow_executions/new?workflow_id=22

docuteam feeder

NW ▾ Ausführungen Workflows Cockpit Benutz...

Neue Ausführung für Access: Deliver DIP by PID

Objekt-ID

SIP

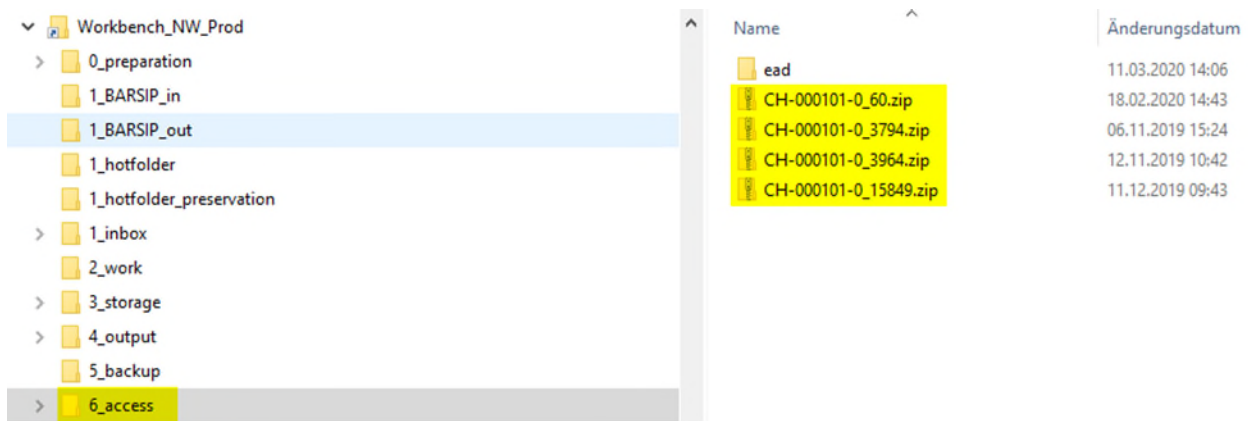
oder

Manuelle Eingabe

CH-000101-0:67149 1

Ausführung erstellen 2

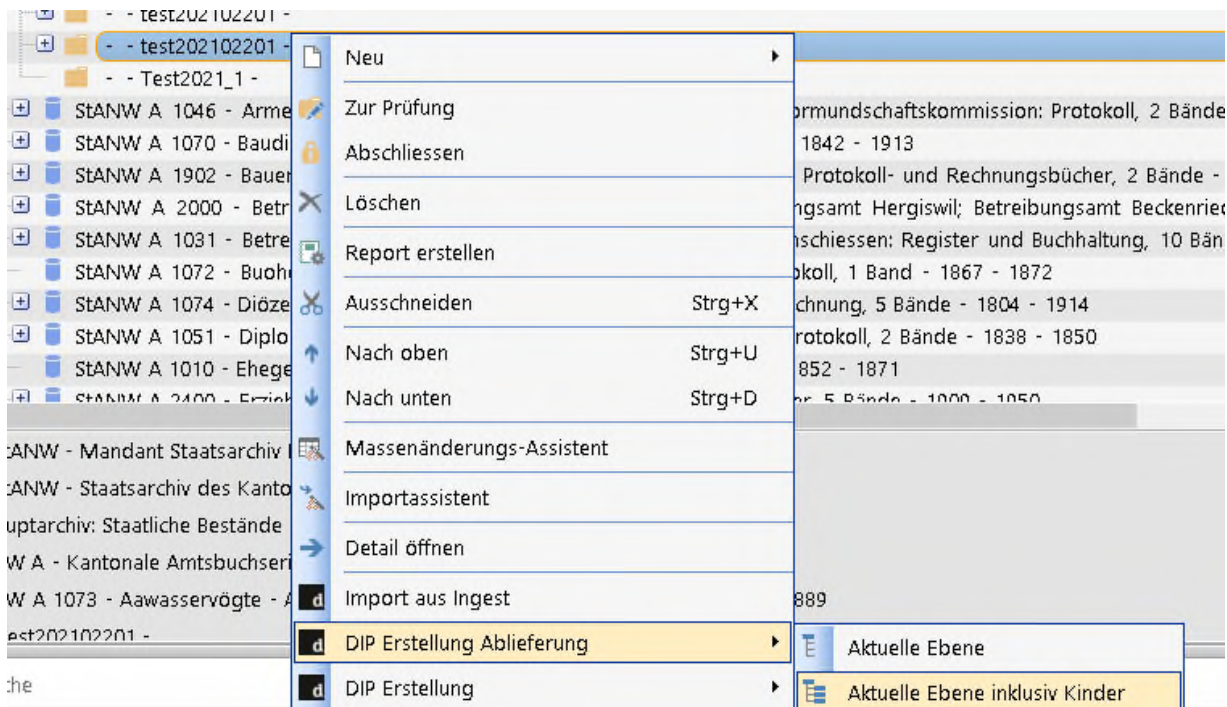
Das DIP wird im Laufwerk "L" in der "Workbench_NW_Prod/Test" im Ordner "6_access" (Ordnername = PID) als Zip-Datei abgelegt. Die DIP werden nach der Benutzung gelöscht.



3.4.2 DIP-Generierung mit CMI AIS

Noch einfacher ist die DIP-Erzeugung im CMI AIS. Mit dem in CMI AIS integrierten Service "DIP Erstellung Ablieferung" kann entweder ein DIP der ausgewählten Verzeichnis-Ebene oder eines der ausgewählten Verzeichnis-Ebene mit allen dazugehörigen Unterstufen ("Kindern") erstellt werden.

In CMI AIS entsprechende Archivplanposition rechtsklicken (1), Reiter "DIP Erstellung Ablieferung" anklicken (2), wählen ob nur die "aktuelle Ebene" oder die "aktuelle Ebene inkl. Kinder" (= untergeordnete Positionen) ausgegeben werden soll (3):



Erledigungsmeldung erfolgt vom System per E-Mail inkl. Downloadlink einer zip-Datei des DIP. (Speicherort DIP: L:\Workbench_NW_Prod\6_access). Die DIP werden nach der Benutzung gelöscht.

4 Allgemeine Grundsätze zur elektronischen Archivierung

Umfang des ELAR

Sämtliches digitales Archivgut (sowohl "digital born" als auch retrodigitalisierte Daten) wird im ELAR abgelegt. Das ELAR ersetzt sämtlichen bisherigen Zwischenablagen für digitales Archivgut.

Erschliessung elektronischer Daten

Elektronische Ablieferungen werden grundsätzlich gleichbehandelt und dokumentiert wie Ablieferungen in Papierform. Die Bewertung und Erschliessung erfolgt gemäss den Richtlinien zur Bewertung von Archivgut sowie den Richtlinien zur Erschliessung von Archivgut. Als Checkliste dient das Laufblatt Akzessionen.⁹ Auf die dort beschriebenen Schritte wird im Folgenden nicht mehr eingegangen.

Archivische Verzeichnismetadaten

Idealerweise werden die archivischen Verzeichnismetadaten sowohl im AIS als auch im AIP hinterlegt. Dies macht gerade vor dem Hintergrund von Linked Open Data Sinn. In der Praxis ist die doppelte Hinterlegung der archivischen Verzeichnismetadaten jedoch problematisch. Sie werden zwar beim Ingest angelegt, ändern aber später (Fehlerbehebungen, neue Erkenntnisse, Nachführung des Verzeichnisses). Gleichzeitig können geänderte Metadaten im Repository (Fedora) heute aber nicht nachgeführt sondern nur erneut importiert werden, was zu einer Diskrepanz zwischen den Metadaten im AIS und im AIP führt. Dies ist zu verhindern.

Aus diesem Grund werden die archivischen Verzeichnismetadaten vorerst nur im AIS geführt und gepflegt. Im AIP wird nur ein minimales Metadaten set mit Titel, Entstehungszeitraum und Schutzfrist gespeichert (vgl. Kap. 5.5.4).

Originale Dateiformate

Microsoft Office-Dateien werden grundsätzlich doppelt archiviert. Die originalen Dateiformate werden übernommen. Zusätzlich werden, falls die originalen Dateiformate abweichen, Nutzungskopien in archivtauglichen Dateiformaten generiert und archiviert.

⁹ Dok. #386029 und Dok. #343983. Das Laufblatt ist als Vorlage in Officeatwork hinterlegt.

5 Richtlinien zum Pre-Ingest

5.1 Ablieferungsprototypen

Im Idealfall erfolgt die Übertragung der Primär- und Metadaten aus einem Produktiv- ins Archivsystem über Schnittstellen in automatisierter Form. In der Praxis zeigt sich, dass zwischen einem Produktivsystem und dem Archivsystem grosse Hürden zu überwinden sind.

Aus heutiger Sicht lassen sich folgende Ablieferungsprototypen festhalten:

Ablieferungstyp 1: Fileablage auf Laufwerk

Ablieferungen von Unterlagen auf Laufwerken bedingen oft eine aufwendige Vorbereitung (Datenanalyse, manuelle Konvertierungen, Strukturumformungen), weil die Ablagen wenig oder nicht strukturiert und regellos gewachsen sind. Die bereinigten Daten werden mit dem Packer weiterbearbeitet. In Einzelfällen ist es allenfalls möglich Strukturinformation (Tektonik) der Laufwerkablage maschinell auszulesen und im Packer zu verarbeiten (Workflow CreateSIPFromExcel).

Dieser Ablieferungstyp ist heute relativ häufig.

Ablieferungstyp 2: Fachanwendung ohne standardisierte Schnittstelle

Je Ablieferung aus einer Fachanwendungen ohne standardisierte Schnittstelle ist ein eigenständiges Projekt. Zum einen müssen die Geschäftsdaten aus dem System gebracht und zum andern die zugehörigen Files analysiert und allenfalls bearbeitet werden. Ziel muss es sein, die Geschäfts-/Metadaten mit dem Workflow CreateSIPFromExcel in den Packer zu übernehmen und dort weiterzuverarbeiten.

Dieser Ablieferungstyp ist heute selten, dürfte aber mittelfristig wichtig werden (Bsp. RMS).

Ablieferungstyp 3: Fachanwendung mit standardisierter Archivschnittstelle

Das produktive System gibt die zu archivierenden Daten (Primär- und Metadaten) als standardisiertes SIP heraus. Dieses kann nach der eCH-0160-Transformation (Workflow BAR-SIP Converter) direkt dem Ingestprozess übergeben werden.

Dieser Ablieferungstyp ist heute die Ausnahme, wird und muss langfristig aber zunehmen.

Ablieferungstyp 4: Datenbank

Datenbanken können nach der Bearbeitung mit die pr Suite als Datenbanken ins Repository übernommen werden. Ergänzend oder alternativ dazu können aus der Datenbank Reports generiert und (ebenfalls) archiviert werden.

Dieser Ablieferungstyp wird aus heutiger Sicht mit Ausnahme der Archivierung von Reports zumindest vorderhand ohne Bedeutung bleiben.

5.2 Dateinamen

Bei den Dateinamen wird unterschieden zwischen digital entstandenen und übernommenen Dateien ("digital born") und im Nachhinein von analogen Originalen hergestellten Digitalisaten zur Sicherung und / oder Vermittlung via Online-Publikation (Retro-Digitalisate). Es gilt folgende Regel:

- **Digital Born:** Die Dateinamen werden unverändert aus dem Quellsystem (Fachanwendung oder Fileablage) übernommen. Es werden keine Anpassungen durchgeführt.

- **Retro-Digitalisate:** Die Benennung der Dateien richtet sich nach den Regeln für die Integration und Online-Publikation von Dateien in den Erschliessungsrichtlinien des Staatsarchivs.¹⁰

5.3 Fileablagen

Werden dem Staatsarchiv Rohdaten in Form von Fileablagen abgeliefert, sind diese vor der Paketierung zu einem SIP-Paket in Absprache mit der FVE und dem FVELAR zu bereinigen: die Daten werden streng bewertet, falls nötig migriert und/oder neu strukturiert (ISAD(G)-konforme Struktur). Das Ergebnis dieses Prozesses ist gemäss den Bewertungsrichtlinien summarisch im AIS (CMI AIS) zu dokumentieren.

Fileablagen beinhalten in sehr vielen Fällen **Spezialformate**. Diese unbearbeitet durch den Ingest laufen zu lassen, würde zu unzähligen Fehlermeldungen führen. Spezialformate müssen deshalb vor dem Ingest von Hand migriert oder nach einer Bewertung gelöscht werden. Dieser Prozess erfolgt unter Anleitung des FVELAR mittels **DROID**.

5.4 Relationale Datenbanken

Bis dato werden aus Datenbanken in der Regel nur Reports archiviert; in Ausnahmefällen kann die Archivierung im Originalformat und / oder mittels SIARD erfolgen. Der Entscheid erfolgt durch den FVELAR.

Werden Datenbanken als solche ins Langzeitarchiv übernommen, müssen die Datenbanken im Pre-Ingest mit **SIARD Suite** in ein archivtaugliches Format (.siard > xml und db blobs) umgewandelt werden. Die Filemigration erfolgt unter Anleitung des FVELAR.

5.5 Erstellung der SIP

5.5.1 Benennung der SIP

SIP werden einheitlich benannt. Die Benennung der SIP ermöglicht die eindeutige Identifizierung eines SIP. Die Benennung eines SIP erfolgt nach dem folgenden Schema: SIP_[Akzessionsnummer] (Beispiel: SIP_2018-17).

5.5.2 Granularität der SIP

Massgebend für die Granularität sind die Bestände in der Archivstruktur. Für die Granularität eines SIP gilt der Grundsatz, dass für jede Akzession pro (Ziel)-Bestand ein SIP gebildet wird. Wird eine Akzession in einen Bestand übernommen, wird nur ein SIP gebildet. Wird eine Akzession auf mehrere Bestände aufgeteilt, wird pro Zielbestand ein SIP gebildet.

Aus technischen Gründen dürfen SIP **maximal 10'000 Dateien** beinhalten. Enthält eine Akzession mehr als 10'000 Dateien müssen die zu bildenden SIP zwingend aufgeteilt werden. In einem solchen Fall wird vom Grundsatz abgewichen und es ist pro Zielbestand mehr als ein SIP gestattet. Dies ist im Erschliessungsbericht zu dokumentieren.

5.5.3 Erschliessungstiefe in den SIP

Ein SIP ist grundsätzlich in einer ISAD(G)-konformen Hierarchie zu gliedern. Es gibt maximal vier Verzeichnungsstufen (ggf. zuzüglich Teilstufen):

- Bestand (Teilbestand)
- Serie (Teilserie)
- Dossier (Teildossier)
- Einzelstück.

¹⁰ Dok. #343983.

Im Grundsatz wird bis auf Stufe Dossier und nur im Ausnahmefall bis auf Stufe Einzelstück erschlossen. Zu den Ausnahmen zählen bspw. ausgewählte retrodigitalisierte Fotobestände oder ausgewählte retrodigitalisierte Amtsdruckschriften.

Die Stufe Einzelstück wird im Ingest-Prozess nur vergeben, wenn diese später auch ins AIS (CMI AIS) importiert werden soll; heute wird diese Möglichkeit nicht genutzt (vgl. Kap. 4.5.3). Die Dokumente der Stufe Einzelstück werden auf "undefiniert" gelassen (so erhalten sie später keine eigenen Signaturen und werden nicht als Einzelstücke mit importiert).

Die Gliederung in Serien und Dossiers ist in jedem Fall mit der FVE zu definieren.

5.5.4 Minimales Metadatenset der SIP

Welche deskriptiven Metadaten in einem SIP erfasst werden, hängt massgeblich davon ab, wie viele deskriptive Metadaten aus dem Produktivsystem mitkommen. Bei einer Fileablage, die praktisch keine deskriptiven Metadaten enthält, werden im SIP nur die notwendigsten deskriptiven Metadaten erfasst. Bei einer Übernahme aus einem RMS-System oder einer Fachanwendung, wo viele deskriptive Metadaten vorhanden sind, können mehr Metadaten automatisch (Workflow "CreateSIPFromExcel") übernommen werden. Der Entscheid, welche Metadaten im Einzelfall aus RMS- oder Fachanwendungen übernommen werden können, fällt der FVELAR.

Im SIP muss auf allen Verzeichnungsstufen ein Minimum an deskriptiven Metadaten erfasst werden. Dieses besteht aus:

- Titel
- Entstehungszeitraum
- Schutzfrist.

Dieses minimale Metadatenset gilt für die im Packer als Verzeichnungsstufe ausgezeichneten Strukturen. Es gilt nicht für Einzelstücke, die nicht als Verzeichnungseinheiten übernommen und als "undefiniert" ausgezeichnet sind. In der Regel müssen im Packer also deskriptive Metadaten für Bestand, Serien und Dossiers, nicht aber für Einzelstücke angelegt werden.

6 Richtlinien zum Ingest

6.1 Archivtaugliche Dateiformate

Als Grundlage für die Definition der archivtauglichen Dateiformate dient der Katalog archivischer Dateiformate der KOST¹¹. Es ist davon auszugehen, dass es künftig zu Anpassungen der Liste der archivtauglichen Formate kommen wird.

Archivtaugliche Dateiformate im StANW, Stand: April 2022

- PDF/A 2aub (Präferenz: PDF/A 2u)
- TXT
- JP2 (JPEG2000)
- JPG
- WAV
- MP3
- XLSX
- CSV
- SIARD
- XML
- Diverse Videoformate, u.a. MPEG-4, FFV1, MOV, AVI, MKV

Microsoft Office-Dokumente: Von mit Microsoft Office erstellten Dateien (.doc/ .docx, .ppt/ .pptx) wird sowohl das Originalformat als auch eine in PDF/A 2aub konvertierte Version ins ELAR gespeichert.

Ausnahme Excel-Dokumente: .xls-Dateien müssen im Pre-Ingest in xlsx-Dateien umgewandelt werden. .xlsx-Dateien werden im Original ins ELAR übernommen, ohne Migration in ein PDF/A 2aub.

Spezialfall Videoformate: Videoformate können nicht per se als archivtauglich oder nicht archivtauglich klassifiziert werden. Vielmehr ist je nach Entstehungszusammenhang, Archivauftrag und Bewertungsentscheid jeweils ein bestimmtes Format das in diesem Fall am besten geeignete. Der Entscheid liegt beim Fachverantwortlichen für die elektronische Langzeitarchivierung (FVELAR).

Docuteam Feeder führt folgende Formatmigrationen durch:

Ursprungsformat	Migration	Zielformat	Bemerkung
PDF	Ja	PDF/A 2aub	
.ppt / .pptx (Powerpoint)	Ja/Nein	PDF/A 2aub .ppt/ .pptx	
.doc / .docx	Ja/Nein	PDF/A 2aub .doc/ .docx	
.odf	Ja	PDF/A 2aub	
.dotm	Ja	PDF/A 2aub	
.tif	Ja	.jp2	
.png	Ja	.jp2	
.bmp	Ja	.jp2	
.msg	Ja	PDF/A 2aub	
msg Container mit angehängten Files	Ja	PDF/A 2aub	Anhänge werden auch zu PDF/A konvertiert sofern möglich. Alles wird in eine PDF/A-Datei geschrieben.
.rtf	Ja	PDF/A 2aub	

¹¹ https://kost-ceco.ch/cms/kad_main_de.html.

.mp3	Nein	.mp3	
.txt	Nein	.txt	
.xls	Nein	.xlsx	Umwandlung muss im Pre-Ingest manuell durchgeführt werden.
.xlsx	Nein	.xlsx	
.jpg	Nein	.jpg	
.jp2	Nein	.jp2	
.wav	Nein	.wav	
Videoformate	Nein	Ursprungsformat	Im Feeder findet keine Migration statt. Falls nötig, muss diese im Pre-Ingest von Hand durchgeführt werden.
.csv	Nein	.csv	
.siard	Nein	.siard	
.xml	Nein	.xml	

Liegen **andere Dateiformate** vor, muss die Formatmigration in archivtaugliche Dateiformate bereits im Pre-Ingest von Hand durchgeführt werden oder die Formatmigration im Feeder angepasst werden. Ein entsprechender Entscheid trifft der FVELAR.

6.2 Archivische Verzeichnung im Archivinformationssystem

6.2.1 Eingliederung in die Archivstruktur

Die Richtlinien zur Erschliessung von Archivgut gelten grundsätzlich auch für die Erschliessung digitaler Ablieferungen.

Elektronische Bestände werden in den bereits bestehenden Abteilungen erschlossen und nicht speziell gekennzeichnet. Es wird keine eigene Abteilung für elektronische Bestände geschaffen. Anhand der Signatur wird auch nicht ersichtlich, dass eine Verzeichnungseinheit elektronische Dokumente enthält.

6.2.2 Ergänzung von deskriptiven Metadaten im AIS nach dem Import

Mittels Import-Schnittstelle werden nur bereits vorhandene Metadaten importiert. Alle anderen Pflichtfelder müssen gemäss den Richtlinien zur Erschliessung von Archivgut zwingend nach dem Import ergänzt werden (vgl. Kap. 3.5).

Werden ins AIS importierte Metadaten (z.B. Titel) nachträglich geändert, werden diese Änderungen im AIP im Repository nicht übernommen. Ein erneuter Ingest des betreffenden SIP mit den korrigierten deskriptiven Metadaten ist nicht verhältnismässig und wird nicht durchgeführt. Grundsatz: Die deskriptiven Metadaten im AIS sind die massgebenden Metadaten, nicht die im AIP gespeicherten.

6.2.3 Erstellung von Nutzungskopien

Neben den Metadaten können mittels Schnittstelle Docuteam-Tools / CMI AIS auch Nutzungskopien automatisch erstellt und automatisch in CMI AIS importiert werden. Dafür ist im Docuteam Packer die Auszeichnung als "Dokument" notwendig.

Momentan wird diese Möglichkeit nicht verwendet. Die Erschliessung erfolgt nur im Ausnahmefall auf Einzelstückebene. Nutzungskopien werden vorderhand nur im AIS abgelegt, wenn sie im Internet publiziert werden sollen. In diesen Fällen erfolgt die Erstellung der Nutzungskopie gemäss Richtlinien zur Erschliessung von Archivgut (=eigener Prozess) und nicht via Schnittstelle Docuteam-Tools / CMI AIS.

7 Richtlinien zu Management und Preservation Planning

Management und Preservation Planning werden in Teilprojekt "ELAR Erhaltungsplanung" (frühestens ab 2023) ausdifferenziert. Die folgenden Richtlinien sind provisorischer Art.

7.1 Statistik, Planung

Vorderhand werden die folgenden Kennzahlen erhoben:

- Anzahl Ablieferungen (SIPs) pro Jahr
- Anzahl erschlossener/ins ELAR übernommener Dateien pro Jahr
- Dateimenge/Speicherplatz erschlossener/ins ELAR übernommener Dateien pro Jahr.

Der FVELAR überträgt per 31.12. jeden Jahres die Kennzahlen auf das Kontroll-Excel "Laufende Statistik Bestandesumfang".

Die Kennzahl "Dateimenge/Speicherplatz" dient neben der statistischen Erhebung auch der Planung des Speicherplatzbedarfs. Der FVELAR überprüft diesen jährlich für die Budgeteinkauf und legt ihn dem Betriebsausschuss vor. Der Betriebsausschuss entscheidet über den Budgetantrag.

7.2 Beobachtung der Dateiformate

Zeichnet sich die Obsoleszenz eines im Repository abgespeicherten Dateiformats ab, kann der Feeder via Workflow "Deliver DIP by PUID" sämtliche Dokumente eines Dateiformats in den Workbench "1_hotfolder_preservation" auswerfen. Mit dem Workflow "Update Fedora Objects" werden die in "1_hotfolder_preservation" liegenden Dateien automatisch in das neue in der "file migration" definierte Dateiformat umgewandelt. Es lassen sich so z.B. alle .txt-Files aus dem Repository holen, in das neue Dateiformat umwandeln und direkt wieder als Version des Ursprungdokuments ins Repository schreiben. Bei einer allfälligen DIP-Erzeugung wird nur die neueste Version eines Objekts ausgeliefert.

Achtung: Bei Preservation Actions ist zu beachten, dass "gleiche" Formate im Cockpit des Feeders unter verschiedenen Positionen aufgeführt sein können, da bei der Formaterkennung verschiedene PRONOM-Identifizier für das "gleiche" Format vergeben wurden.

Beispiel JPG: JPG werden als fmt/43, fmt/44, fmt/645, x-fmt/391 und eventuell noch mit anderen PRONOM-Identifizierern erkannt. Grund für diese unterschiedlichen PRONOM-Identifizier sind die leicht anderen JPG-Formate der einzelnen Dateien.

Spezialfall EXIF: Im Feeder-Cockpit kann es sich bei EXIF-Files (Exchangeable Image Files) um JPG-Dateien mit EXIF-Daten handeln. EXIF hat gemäss PRONOM-Datenbank höhere Erkennungspriorität als RAW JPG-Streams. Das heisst JPG mit EXIF-Daten werden als EXIF erkannt.¹² Das ist für zukünftige Preservation Actions bei JPG-Dateien wichtig.

¹² Vgl. <https://www.nationalarchives.gov.uk/pronom/fmt/645>.

8 Richtlinien zum Zugang

Der Zugang zu AIP im Repository ist momentan erst provisorisch geregelt. Er erfolgt durch Archivmitarbeitende an deren Arbeitsstationen. Die Weiterentwicklung Richtung digitaler Lesesaal soll in Teilprojekt "ELAR Zugang" frühestens ab 2022 erfolgen.

8.1 Interner Zugang ins Repository

Der direkte Zugriff ins Repository ist dem Staatsarchivar, der FVE und dem FVELAR vorbehalten. Im Regelfall erfolgt der Zugriff auf Daten im Repository über die vordefinierten Prozesse des Feeders oder des AIS (CMI AIS) und nicht über den Direktzugriff.

8.2 Einsichtnahme in Dokumente

Bis zur Entwicklung eines "digitalen Lesesaals" haben Benutzerinnen und Benutzer keinen direkten Zugriff zu Daten aus dem Repository und können selbst keine DIP erzeugen. Elektronische Unterlagen müssen beim Archivpersonal bestellt werden und werden je nach Zugangsbestimmungen und Verwertungsrechten per E-Mail versendet oder sind im Lesesaal einsehbar.

Zu regeln sind in diesem Zusammenhang offene Fragen zum Datenschutz, der Art und Weise der Konsultation im Lesesaal und der Gebühren beim Versand. Wo/solange konkrete Regelungen fehlen, entscheidet der Staatsarchivar analog zu den Regeln für analoge Archivgut.