



A-UpS! Werkstattbericht Archivierung von Unterlagen aus privaten Systemen

Was erwartet Sie in den nächsten 20 Minuten?

1. Was ist das AfZ?
2. Ausgangslage für A-UpS!
3. Workflow für die digitale Archivierung
4. Massnahmen zur Vertrauenswürdigkeit
5. Fragen und Diskussionspunkte

Was ist das AfZ?

1. **Sicherung** privater Nachlässe und institutioneller Archive mit Schrift-, Ton- und Bildquellen zur Geschichte der Schweiz vom späten 19. Jahrhundert bis in die Gegenwart
2. **Vermittlung** des historischen Kulturguts zur schweizerischen Zeitgeschichte an Lehre, Forschung und interessierte Öffentlichkeit in der Schweiz und im Ausland



Arten von Archivalien im AfZ

- **Archivalienarten** (Medientypen)
 - Schriftgut
 - Fotografien / Bilder / Plakate
 - Video / Film / Ton
 - Gegenstände
- **Trägermedien**
 - Papier
 - Film (Negative / Durchlicht)
 - Mikroformen (Fichen, Mikrofilme)
 - Magnetbänder (Ton, Video)
 - Textilien / Holz / Metall etc.
 - Magnetische, magnetoptische, optische und elektronische Datenträger (Festplatten, CD, DVD, BD, Memorysticks, SD-Cards usw.)



Digitalisierung

document type	Target file	Target file format	Finalized end of 2016	Pages, meters, hours
written	300 dpi greyscale	pdf/a 2b (j2k, 90) / pdf	41'256 dossiers	4.003 Mio pages 520 meters
graphic art	450 dpi	tif / pdf	4414 / 4536 dossiers	111'956 files
photos	25 MB / 75 MB	tif / pdf	8294 / 10'663 dossiers	
objects	75 MB	tif / pdf	26 / 164 objects	
maps	450 dpi	pdf/a 2b	9 / 103 dossiers	
posters	75 MB	tif / pdf	289 / 446 dossiers	
sound	24bit / 48-96 KhZ	wav / mp3	969 / 1392 dossiers	
video	1920x1080 (HD) 720x540 (SD)	mkv (ffv1) lossless / mp4	679 / 1125 dossiers	~ 410 hours
film	1440/1080	mkv (ffv1) lossless / mp4	202 / 338 dossiers	~ 60 hours
total			56'138 dossiers (28%)	

A-UpS!?: Ausgangslage

Herausforderungen als Spezialarchiv:

- Privatprovenienzen:
 - Teilweise unbekannte juristische Ablieferer
 - (derzeit) gänzlich unbekannte natürliche Ablieferer
- Einfluss auf Informationsmanagement abgesehen von Beratung beschränkt
- Keine Policies zu Informationsmanagement oder Formaten durchsetzbar
- Flexibilität im Umfang mit verschiedenen Ablagepraxen und –logiken nötig
- Fallspezifische Abwägungen zu Archivierbarkeit und Bearbeitungsaufwand nötig



A-UpS!?: Ausgangslage

Wildwuchs in der digitalen Archivierung:

- Akzessionen unter Zeitdruck (z.B. nach Ableben eines Bestandsbildners)
- fehlende Sensibilität betr. digitaler Unterlagen bei Akzessionen
 - Grosse, hybride, unstrukturierte und lückenhafte Ablieferungen
 - Doppel, Trippel und mehr
 - Problematische Praxen (Verschlüsselungen, mangelnde Dokumentation)
- Abschieben der Nachbearbeitung, Bewertung und Strukturierung an die IT



Vorarchiv und vorarchivische Beratung

- Frühzeitige Kontaktaufnahme mit den Anbietern
- Empfehlungen zum Informationsmanagement für institutionelle Anbieter und Privatpersonen
- Sichtung vor Ort mit Checkliste
- Falls möglich Mitarbeit der abgebenden Stelle bei Bereinigung, Metadatierung und Zusammenstellung des Materials



Akzession

- Hybride Bestände
 - Analogisieren (Ausdrucken)
kleiner digitaler Bestandesteile
- digitale Bestände / Bestandesteile
 - Falls möglich erste Bestandsaufnahme vor der Ablieferung (Checksummen)
 - Transfer in den Akzessionsspeicher AfZ (Checksummen-Abgleich, Dokumentation, Backup)



Bewertung und Erschliessung

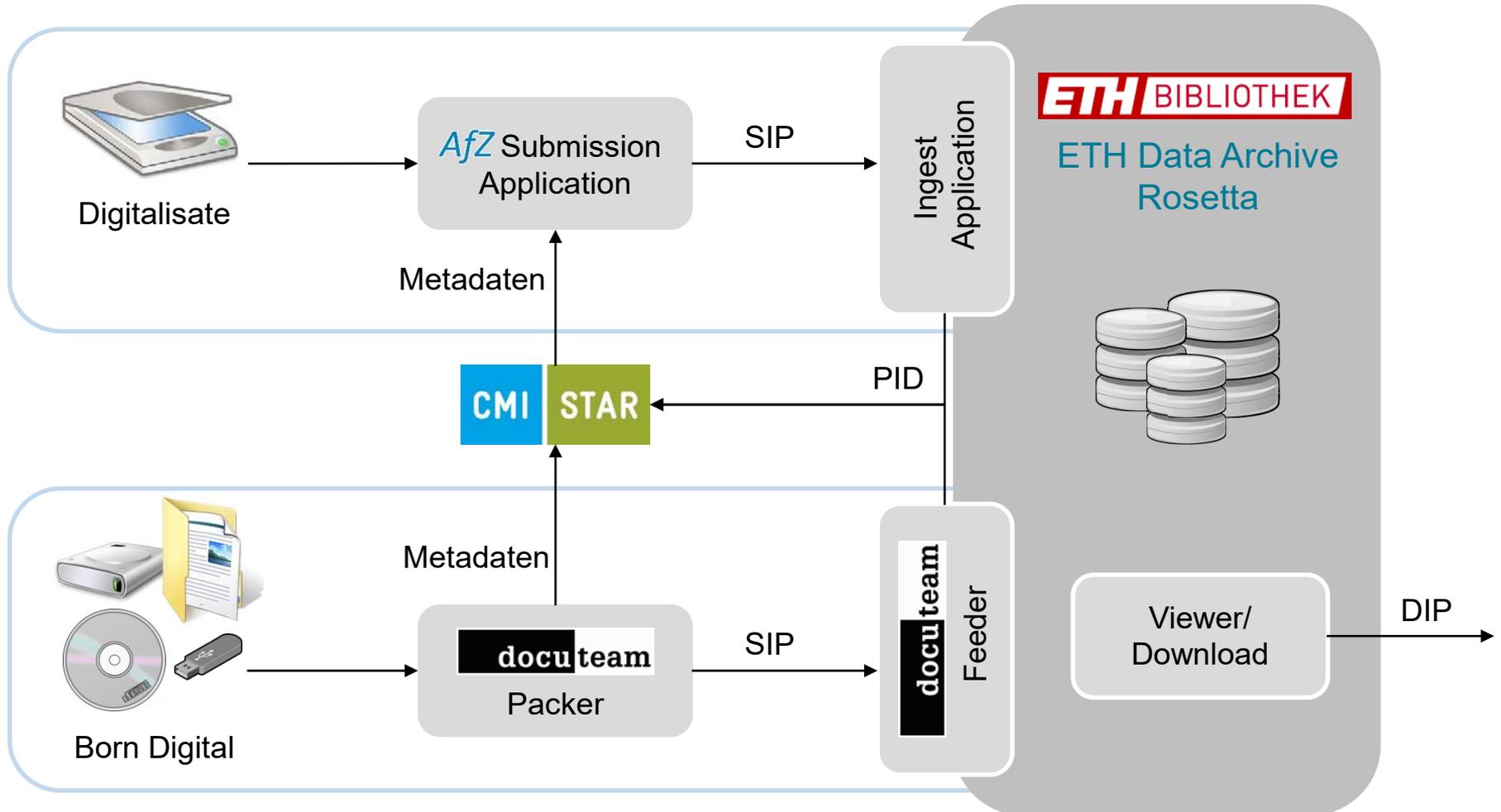
- Technische Massnahmen
 - Normalisierung der Dateinamen
 - Extraktion von beschreibenden und technischen Metadaten
 - DROID-Analyse
 - Duplikate bereinigen
 - Bereitstellen von Arbeitskopien
- Archivische Erschliessung
 - Inhaltliche Bewertung und Kassationsvorschläge
 - Verzeichnung auf Dossierstufe (docuteam Packer)

Langzeitspeicherung und Erhaltung

- Langzeitspeicherung im Verbund mit der ETH Bibliothek
 - ETH Data Archive (Rosetta von ExLibris)
 - AfZ-LZA als «deep Archive» ohne Nutzerzugang
 - Benutzerkopien werden unabhängig von der LZA verwaltet
 - Metadaten werden im AIS verwaltet und gepflegt (CMI Star)
- Formatmigration
 - Migration / Normalisierung im AfZ
 - Formatprüfung (docuteam Feeder)
- Erhaltungsplanung
 - Kontrollmöglichkeiten in Rosetta (Parameter)
 - Erhaltungsmaßnahmen werden vom AfZ / ETH Data Archive in Kooperation durchgeführt

Langzeitspeicherung und Erhaltung

Gemeinsame Infrastruktur von AfZ und ETH Bibliothek



Langzeitformate

Dateiart	Beispielformate	Zielformate
Text, E-Mail	.doc, .docx, .msg, .eml, .txt	PDF/A, ODT (?)
Foto, Bild, Grafik	.jpg, .tiff, .gif, .png, .bmp	TIFF, JPEG2000
Video, Film	.mpeg, .avi, .wmv, .mov	FFV1 in MKV
Ton, Audio	.mp3, .mp4, .wav	WAV
Spreadsheets / Tabellen	.xls, .xlsx, .csv	PDF/A, ODS, CSV (?)
Präsentation	.ppt, .pptx	PDF/A, ODP, mkv (?)
Webseiten / Wegpages	.html, .css, diverse Formate	(W)ARC, PDF (?)
Datenbanken	.mdb, .accdb, DB-Dump	SIARD
Applikationen	Code	?
Social Media	Facebook, Twitter, Instagram, YouTube, Blog	?

Automatisierung

- (teil-) automatisierte Prozesse
 - Checksummen
 - Duplikate finden und entfernen
 - Formatanalyse
 - Dateinamen normalisieren
 - Datei-Metadaten auslesen
 - Migration
 - SIP bilden
 - Ingest ins LZA
- Manuelle Prozesse
 - Inhaltliche Bewertung
 - Erschliessung (Strukturierung / Metadatenvergabe)
 - Dokumentation

Vertrauenswürdige Durchführung

- Prozessgestaltung
 - Feste Abläufe und Workflows
 - Sensibilisierung der Mitarbeiter/Innen
 - Bereitstellen von Tools / Schulungen für Mitarbeiter/Innen
 - Dokumentation definierter «Meilensteine»
 - Bereitstellung der Dokumentation im AIS

Probleme und offene Fragen

- Macht es Sinn die Ursprungsdateien zu archivieren?
- Wann ist der optimale Zeitpunkt für die Konvertierung in LZA-Formate?
- Ist ODF eine Alternative / Ergänzung zu PDF/A im Office-Bereich?
- Reicht ein Format pro Archivalienart zur Abbildung aller (signifikanten) Eigenschaften / künftigen Userbedürfnisse?
- Wie granuliert sollen die im Archivierungsprozess getätigten Massnahmen dokumentiert werden?
- Wie weit soll / darf die Bewertung und Erschliessung Maschinen überlassen werden (Data Mining, Entity Extraction)?



Vielen Dank für Ihre Aufmerksamkeit!

Jonas Arnold: arnold@history.gess.ethz.ch

Sonja Vogelsang: vogelsang@history.gess.ethz.ch