



# Projektantrag

erstellt von: Alle Beteiligte inkl. Vertreter Geschäftsstelle

geht an: KOST, Steuerungsausschuss

**Betreff:** 20-039: Dateiablage

Stand: 5. Juni 2020 / Version 0.3

---

## 1 Ausgangslage

In der Verwaltung sind unbeachtet von Geschäftsverwaltungssystemen oder Datenbanken viele archivwürdige Informationen in Fileablagen abgelegt und die Inhalte können mittlerweile schon 30 Jahre zurückreichen. Neben Fileablagen aus der Verwaltung kommt der Bereich der Privatarhive, welcher in unterschiedlichem Grade ebenfalls in allen Archiven eine Rolle spielen. Die Struktur der Fileablagen reicht von geordnet bis zum historisch gewachsenen Chaos. Gemeinsame Eigenheiten sämtlicher Ablagen sind die, dass die Anzahl Files rasch sehr gross wird und viele verschiedene Formate existieren. Für die Bereiche Bereitstellung des Aktenangebots, Bewertung, Einlagerung und Erschliessung gibt es kein standardisiertes Produkt, welches zu einem grossen Teil die Aufgaben für die Archive übernimmt. Es gibt verschiedene Tools zu den einzelnen Bereichen. Zwischen verschiedenen Archiven besteht ein Austausch (meistens zu einem konkreten Fall) und Fragen dazu werden sporadisch an Anlässen wie Tagungen diskutiert. Oft ist es aber so, dass die Archive nicht wirklich wissen, wie Berufskolleginnen und -kollegen damit umgehen.

## 2 Projektgruppe und federführendes Archiv

Die Kerngruppe des Projektes setzt sich wie folgt zusammen: Die *Staatsarchive Obwalden, Aargau, Schwyz* und *Zug* das *Archiv für Zeitgeschichte* und das *UZH Archiv*, welches die organisatorische Federführung innehat.

Nach Genehmigung des Projektantrages können weitere Archive sich am Projekt beteiligen. Interesse angemeldet haben die Staatsarchive Thurgau, Schaffhausen und Solothurn, das Stadtarchiv Bern und das Hochschularchiv der ETH Zürich. Ein offizieller Aufruf ist jedoch bis dato noch nicht erfolgt.

Von der KOST Geschäftsstelle werden sich Martin Kaiser und Georg Büchler beteiligen, allenfalls in Stellvertretung.

## 3 Projektziel

Es sollen die Prozesse, Tools und Schnittstellen für den gesamten Prozess der Archivierung von Dateiablagen von der aktenproduzierenden Stelle bis zur Einlagerung ins Repository untersucht werden, um ein Set an Best-Practices und Workflows zusammenzustellen. Auf einem Nebengleis



werden weiterführende Überlegungen angestellt zur Frage, wie der ganze Prozess bis zum Ingest kontrolliert, dokumentiert und automatisiert werden könnte und welche archivfachlichen Überlegungen den jeweiligen Vorschlägen zu Grunde liegen.

#### 4 Projektaufbau

Das Projekt wird in Phasen gegliedert, die sich am Ablauf des Archivierungsprozesses orientieren.

*Erste Phase: Bereitstellung des Aktenangebotes, Aufbereitung und Bewertung und technische Analyse*

*Bereitstellung* behandelt die mit der Ablieferung bzw. Übermittlung einer Dateiablage in Zusammenhang stehenden Fragen.

*Aufbereiten und Bewerten* bei Dateiablagen meint eine inhaltliche Analyse, auf Grund derer dann erste Bewertungsentscheide gefällt werden, nämlich Umstellen der Ablage, Löschen von Teilen der Ablage und Umbenennen der Ordnerstruktur.

Die *technische Analyse* beinhaltet Punkte wie Format- und Formattypenerkennung, Formatvalidierung, De-Duplizierung, Varianten und Versionen erkennen, Volumenanalyse, etc. und die Frage, welche Tools mit Vorteil dafür eingesetzt werden.

Bewertungsfragen inhaltlicher Natur sind nicht das Thema. Angesprochen wird aber die Frage, wie konkret eine Bewertung einer sehr grossen Ablage vorgenommen werden kann. In diesem Zusammenhang sollen Erkenntnisse gewonnen werden, in welcher Reihenfolge technische Analyse und Aufbereitung erfolgen sollen und ob es neben den bekannten Analysetools noch andere Möglichkeiten gibt, welche die Bewertungsarbeit erleichtern können.

*Zweite Phase: Paketbildung und Metadatierung*

Die *Paketbildung* und eine erste *Metadatierung* erfolgt in den Archiven mit unterschiedlichen Tools (*docuteam packer*, *BAR Package Handler* und *Apache Spark*, weniger in der Schweiz gebräuchlich: *ByteBarn*, *PreIngest Toolkit PIT* und *DIMAG IngestTool*)

Diese Tools unterscheiden sich stark bezüglich des generierten SIP-Formats und der Möglichkeiten der Metadatenannotation im generierten SIP, aber auch in den Möglichkeiten, das Paket noch zu bearbeiten und diese Veränderungen zu dokumentieren.

Die verschiedenen Ansätze werden untersucht und verglichen.

*Dritte Phase: Abschluss und weiterführende Überlegungen*

Die zusammengetragenen Ergebnisse werden in einem Abschlussbericht zusammengefasst. Die weiterführenden Gedanken beinhalten die Fragen rund um das Thema, wie der ganze Prozess gesteuert, dokumentiert und überwacht werden kann.

#### 5 Arbeitsschritte

*Projekt Kickoff:*

Kennenlernen, konkretes Vorgehen planen, Verteilung von Aufgaben, Festlegung der Kommunikation zwischen den Treffen. Das methodische Vorgehen: (1) Ausgehen von konkreten Ablieferungen oder (2) Analyse von Literatur und Tools, ebenfalls hier beschliessen.

Die Aufgaben werden von den einzelnen Projektteilnehmerinnen und Projektteilnehmer zwischen den Sitzungen in deren Archiven erledigt.



*Maximal sechs gemeinsames Treffen:*

Diskutieren und Bündeln der Ergebnisse, Erteilen neuer Aufgaben:

- die ersten zwei Sitzungen beschäftigen sich mit Phase 1: Bereitstellung, Aufbereitung und technische Analyse
- anschliessend zwei Sitzungen zu Phase 2: Paketbildung und Metadatierung
- eine Sitzung für Abschluss und Bericht
- eine Sitzung bleibt als Reserve

Sitzungsrhythmus: Die Abstände sollten nicht zu lange sein (ein Monat).

*Schlussitzung:*

Diskussion und Finalisierung des Schlussberichts.

*Öffentliche KOST-Veranstaltung:*

Nach Beendigung Durchführung einer öffentlichen Veranstaltung mit dem Ziel, die Ergebnisse zu präsentieren und in einem grösseren Zusammenhang zu diskutieren.

## **6      Arbeitsaufwand und Budget**

Der Arbeitsaufwand lässt sich schwer abschätzen. Die Erfahrungen der beteiligten Archive zum Thema sind verschieden. Zeit lässt sich sicher einsparen, wenn die Archive bei den verschiedenen Analysen schon auf die eigene Erfahrung zurückgreifen können. Finanzielle Kosten entstehen bis auf Reisespesen keine, bzw. sind nur in Form von Arbeitszeit vorhanden. Die Übersetzung des Schlussberichts ins Französische wird von der KOST finanziert.

## **7      Zeitplan (Projektbeginn und -dauer)**

Die Projektdauer ist auf ein Jahr festgelegt.

*Beginn:*

Auf Grund der Covid-19 Pandemie wird das Projekt erst im November 2020 gestartet.

### **Kontakt**

Martin Akeret, UZH Archiv, Universität Zürich

Tel.: +41 44 63 42326 (Mo. – Do.)

E-Mail: martin.akeret@archiv.uzh.ch

www.archiv.uzh.ch