

## eCH-Fachgruppe *Digitale Archivierung* Themengruppe Austauschformat Protokoll der 9. Sitzung

11. April 2013, 14:00-16:10 Uhr  
Bern, Schweizerisches Bundesarchiv, Auditorium

### *Anwesend:*

Dennis Bauer  
Hedi Bruggisser  
Georg Büchler (Protokoll)  
Alain Dubois  
Martin Kaiser (Sitzungsleitung)  
Lambert Kansy  
Tobias Mérinat  
Ursina Rodenkirch  
Tobias Wildi  
Thomas Zürcher Thrier

### *Entschuldigt:*

Felix Akeret  
Niklaus Bütikofer  
Christian Bütler  
David Gubler  
Manuel Kehrl  
Oliver Landolt  
Michel Meyer  
Krystyna Ohnesorge  
Roger Peter  
Peter Rohrbacher  
Stefan Ryter  
Peter Schneider  
Beat Siegrist  
Martin Walder  
Daniel Wittwer

### **Traktanden**

1. Begrüssung
2. Präsentation des Beispiel-SIPs RIS\_3Dossier
3. Crosswalk eCH-160 → xlsadg
4. Crosswalk eCH-160 → EAD
5. Präsentation von Matterhorn METS Profile
6. Weiteres Vorgehen

### **1 Begrüssung**

Martin Kaiser (MK) begrüsst zur Sitzung und stellt die Traktanden vor: Zunächst ein kurzer Rückblick auf die verwendeten Beispiele, danach Präsentation der Crosswalks zu xlsadg und zu METS Matterhorn inkl. Vorstellung des letzteren. Nötig ist danach eine Diskussion des weiteren Vorgehens.

### **2 Beispiel-SIP**

MK präsentiert kurz die Beispiel-SIP für die Crosswalks. Wir verwenden weiterhin 3 Dossiers aus dem Rats-Informationssystem RIS von St. Gallen, die auch auf dem Wiki sowie im Original online verfügbar sind. Sie liegen in drei Varianten vor:

- Original-ZIP mit den 3 Dossiers mit DHTML-Navigation und FileNet-XML-Daten
- Die 3 Dossiers als eCH-0160- SIP
- Das eCH-0160-SIP mit ergänzten Metadaten.

### 3 Crosswalk eCH-0160 → xlsadg

MK hat für die heutige Sitzung den Crosswalk von eCH-0160 zu ISAD(G) / xlsadg vorbereitet, sowohl als Excel-Tabelle als auch als XSLT-Transformation.

Die Excel-Tabelle bietet die konzeptuelle Grundlage. Sie umfasst auf der linken Seite das Data Dictionary von eCH-0160 (unterteilt in vier Spalten: Ablieferung, OS-Position, Dossier und Dokument), rechts die zugeordneten Felder nach ISAD(G) und die entsprechenden Elemente von xlsadg. Die Bahnen von links nach rechts illustrieren Matches und Leerstellen. Gewisse Elemente von ISAD(G) kommen nicht im SIP vor. Es handelt sich dabei idR um Metadaten aus dem archivischen Umfeld, die mithin im SIP nicht vorkommen können.

Zur Umwandlung der Datei metadata.xml aus dem eCH-SIP in eine Metadatendatei gemäss xlsadg benötigen wir eine XSLT-Transformation; genauer sechs Teil-Transformationen. Ausgeführt wird die Transformation zu Testzwecken mit Saxon.

Drei Punkte sind noch nicht überzeugend dargestellt/gelöst:

- *Konvertierung der Daten-Felder*, wenn Anfangs- und Enddatum in der Präzision nicht übereinstimmen. Der Fehler geht auf einen Entscheid im bentō-Projekt zurück, der korrigiert werden muss.
- *Abbildung der Zugriffsrechte*, relativ komplex, aggregiert, bzw. vererbt, deshalb ein externes Template.
- *Umfang (Menge und Abmessung)* wird noch nicht berechnet. Das sollte aggregierend vor sich gehen, was aber mit XSLT schwierig ist.

MK bittet um Feedback und Diskussion für das weitere Vorgehen.

Thomas Zürcher (TZ) sieht die Notwendigkeit für die Verzeichnung von Umfang/Volumen nicht, jedenfalls nicht automatisch, zumal das Volume bei Migrationen ändert. Grundsatz für die deskriptiven Metadaten ist, dass nur verzeichnet werden soll, was wichtig ist.

Lambert Kansy (LK) bezweifelt den praktischen Wert, aber für einen korrekten Crosswalk ist die Übernahme existierender Werte wichtig. MK präzisiert, dass diese Werte nicht im Schema als Werte, sondern implizit in der Struktur enthalten sind und bei der Übernahme expliziert werden müssen. Sinnvoller wäre wohl, auf diesen Aspekt zu verzichten und ihn in eine entsprechende Applikation zu delegieren.

Zusätzlich gibt es MD-Informationen, für die kein passendes ISAD(G)-Feld existiert und die deshalb nicht gemappt werden. TZ: Schutzfristen sollten in „3.4.1 Zugangsbestimmungen“ einfließen. LK: Diese Felder sind in eCH-0160 weniger bedeutend, da die Zugangsbestimmungen de facto aus den anderen Feldern berechnet werden sollen; sie können sich je nach Gesetzeslage ändern. LK: Einige der eCH-0160-Metadatenfelder entsprechen terminologisch/konzeptuell nicht der ISAD(G)-Denkweise und haben deshalb keine Entsprechung.

Weiteres Vorgehen:

- LK: Zunächst müssen wir die offene Fragen lösen wie MD aus dem GEVER-Bereich, die terminologisch/konzeptuell nicht ISAD(G) entsprechen (z.B. Prozessinformationen) übertragen werden oder überhaupt übertragen werden müssen.

- Ursina Rodenkirch (UR): Das Dokument hat auch dokumentarischen/erläuternden Charakter, nicht unähnlich dem KaD; die erwähnten Probleme müssten dort auch erläutert werden.

#### 4 Crosswalk eCH-0160 → EAD

Tobias Wildi (TW) präsentiert den Crosswalk und verdankt die Vorarbeiten zur Analyse von eCH-0160 durch MK. Anlass zu den Arbeiten von Docuteam war das Ablieferungsmodul von CMI zum Export von eCH-0160 aus CMIAXIOMA. Docuteam erstellte für die Uni Zürich eine Transformation von eCH-0160 in Matterhorn METS Profile, mit EAD als Verzeichnungsstandard. Im Ingest-Prozess werden sämtliche MD aus eCH-0160 in ein EAD-File übernommen für den Import in ein AIS → hier war die Vorgabe, potentiell **alle** eCH-0160-MD zu übertragen. 2 Leerstellen bestehen wegen Verständnisproblemen: verwandteSysteme und archivierungsmodusLöschvorschriften. Berechtigungen werden besser mit PREMIS Rights abgebildet, Vorgänge in PREMIS Events. Für die Daten/Zeiträume arbeitet EAD mit Labels (fromYear, toYear, creationPeriod, creationPeriodNotes). Umfang wird nicht als Metadatum hinterlegt, sondern durch die Applikation berechnet; macht Sinn als erste grobe Mengenangabe.

Die Umsetzung wird nicht mittels XSL/T durchgeführt, da Matterhorn METS die physische **und** die logische Struktur übereinstimmt. Es braucht deshalb auch eine Veränderung der Datenstruktur, die durch eine Java-Applikation vorgenommen wird (zusammen mit der Transformation). Docupack wird eine Import-Funktionalität für eCH-0160 bekommen.

Das Mapping beruht auf dem offiziellen Mapping der EAD Tag Library auf ISAD(G). TW regt an, das Mapping genau zu reviewen und danach bei eCH zu veröffentlichen, als Hilfsmittel oder White Paper o.ä. MK regt an, eine XSL-Transformation noch nachzuliefern zum Vergleich und zum besseren Nachvollziehen. TW liefert dazu Beispiele und Unterstützung, MK erarbeitet die Transformation. LK: Wie werden dann die Zugangsbestimmungen umgesetzt? TW: Nicht in PREMIS, sondern analog zu xlsadg in die entsprechenden Elemente, ev. wiederum mit Labels.

#### 5 Matterhorn METS Profile

Alain Dubois (AD) und TW präsentieren das Matterhorn METS Profile.

Ausgangspunkt ist das Life-Cycle-Schema der Kantonalen Verwaltung VS, das seit einigen Jahren im Einsatz ist (wenn auch noch nicht perfekt). Auf dieser Basis wurden im Wallis zwei Prozesse digitalisiert und RM in zwei Einheiten eingeführt.

Ziel ist eine dynamische Vision des archivierten Dossiers mit Wiederverwendung von Metadaten aus der aktiven Phase im Archiv. Verzeichnung mittels der drei Normen von ICA, ISAD(G), ISAAR(CPF) und ISDF, jeweils mit ihren XML-Encodings (noch fehlend für letzteres). Dazu wird das OAIS-Informationsmodell in mehrere Standards abgebildet: METS, Formatkatalog, PRONOM, PREMIS, EAD, EAC, ISDF. METS ist sehr flexibel und muss in einem Profil konkretisiert werden: eingebettete Standards, gegenseitige Verbindung und Referenzierung. StAVS und Docuteam haben ein entsprechendes Profil realisiert und registriert. Das Matterhorn METS Profile ist das erste Profil, das EAD nicht nur referenziert, sondern im METS einbettet. Nicht inbegriffen sind EAC und ISDF, die aus den entsprechenden PREMIS-Einträgen heraus referenziert werden.

PREMIS erlaubt die detaillierte Abbildung der mit den Dossiers verbundenen Prozesse, mit Metadaten pro Prozessschritt.

Schlussfolgerung: eCH-0160 ist Ablieferungs-Schnittstelle, METS ist ein Archivformat (AIP-Format). Das eCH-0160 ist relativ einfach zu befüllen für ein GEVER-System und deshalb sinnvoll. METS kann aber auch als SIP verwendet werden, zum Beispiel bei der Forschungsdatenarchivierung an der ETHZ.

MK und LK bemängeln, dass die EAD Tag Library teils relativ offen ist, einzelne EAD-Files nicht unbedingt untereinander kompatibel sind. TW widerspricht, dass der Aufwand für den Transfer zwischen verschiedenen EAD-Flavours aus Erfahrung sehr klein ist (Übersetzungen, Anpassungen) und dass er die EAD Tag Library im Rahmen seiner Arbeit als sehr präzise und ausführlich empfunden hat.

## 6 Weiteres Vorgehen

- An weiteren Mappings besteht vorderhand kein Bedarf.
- TW erwähnt die bevorstehenden Änderungen von EAD. Die schwerfällige hierarchische Struktur wird wegfallen zugunsten einer RDF-inspirierten leichteren Struktur. Momentan liegt eine Beta-Version vor, aber die weitere Zukunft ist absolut unklar (wir müssen deshalb nicht warten).
- In welcher Form sollen die Ergebnisse publiziert werden? MK sieht eher ein Whitepaper, da die Transformationen nur illustrativen Charakter haben. LK, GB plädieren für ein Hilfsmittel (als Ergänzung zu eCH-0160), das die Brücke zwischen eCH-0160 und anderen Standards schlagen kann.
- MK: Gibt es eine Möglichkeit zur Visualisierung der MD nach xlsadg bzw. nach EAD? TW stellt das Docuteam-Stylesheet zu EAD zur Verfügung. LK und MK diskutieren das Problem für xlsadg.
- UR: Was noch fehlt, sind detailliertere Beschreibungen der Crosswalks (was geht, was geht nicht), narrativer Text.
  - ➔ MK setzt EAD in XSL/T um, TW liefert Beispiele und Support.
  - ➔ Die XSL/T-Transformationen werden in der Gruppe verschickt; alle testen und geben Feedback.
  - ➔ Die KOST klärt ab, welche Publikationsmöglichkeiten in Frage kommen und wie sie gestaltet sein sollen.
  - ➔ Die Themengruppe trifft sich am **Montag, 10.06.2013, 13:30 Uhr** (bis längstens 16:30 Uhr) im **Staatsarchiv Bern** zu ihrer nächsten Sitzung. Themen: Review der Transformationen, Entscheid über Publikation, Verteilung der Arbeiten.
  - ➔ Angestrebt wird, den Standardentwurf der Fachgruppe an der Novembersitzung vorzulegen. Dies bedingt einen Versand um Mitte September.

Für das Protokoll:  
*Bern, 11.04.2013*  
*Georg Büchler*