

Bern, 24. Juni 2015

eCH-Fachgruppe Digitale Archivierung

Sitzung vom 8.6.2014, Bundesarchiv

1 Begrüssung

Entschuldigt: Stefan Bosshard, CMIAG; André Buob, StABS; Anouk Dunant Gonzenbach, StAGE; Jürg Hagmann, Hagmann Consulting; Oliver Landolt, StASZ; Markus Lischer, StALU; Bernhard Rieder, StAZH; Peter Rohrbacher, StASZ; Jean-Daniel Zeller, HCUG

Anwesend: Felix Akeret, thin-k GmbH; Dennis Bauer, Fachlabor Gubler; Hedi Bruggisser, StATG; Jargo De Luca, StadtA Winterthur; Martin Kaiser, KOST (Sitzungsleitung); Lambert Kansy, StABS; Manuel Kehrli, scope solutions AG; Alain Mast, BAR; Martin Lüthi, StASG; Krystyna Ohnesorge, BAR; Claire Röthlisberger, KOST; Stefan Ryter, StABE; Brigitte Sacker, StadtABE; Beat Siegrist, BK; Florian Vionnier, StAVS

Begrüssst wird zudem Alain Mast, neu in der Fachgruppe für das BAR

2 Protokoll der 8. Sitzung vom 19.11.2014

wird genehmigt

3 eCH-0160

Version 1.1

Die Änderungen, welche zur Version 1.1 führen sollten wurden eingegeben. Da Referent Willi Müller, die Änderungen als nicht rückwärtskompatibel einschätzt, muss das weitere Vorgehen beraten werden. Es ist unklar, ob alle oder nur ein einzelnes Kriterium als nicht rückwärtskompatibel eingestuft werden.

Die Änderungen sind:

- zu restriktives wurde aufgelockert
- Vorgang/Aktivität wurde neu reingenommen
- Integration des Whitepapers

Es gibt folgende Möglichkeiten:

1. Version 2.0 akzeptieren
2. Falls nur ein Kriterium den Ausschlag gegeben hat, dieses herausnehmen und Version 1.1 verabschieden
3. Revision aufgeben und in der AG SIP weiterbearbeiten

Version 1.0 würde den Status „abgelöst“ erhalten, was heisst, dass man sie weiterhin verwenden darf.

Beschluss

Da kein Zeitdruck besteht, werden weitere Abklärungen getroffen und die AG informiert. Es wird abgeklärt, welches der nächste Termin für Möglichkeit 2 ist.

Bei der Geschäftsstelle wird beantragt, dass im Moment keine neue Version publiziert wird.

Berichterstattung zu Version 2.0

Zusätzlich zu den Änderungen an Version 1.0 (Version 1.1) hat die Themengruppe an ihrer Sitzung im Dezember 2014 über die Notwendigkeit einer Version 2.0 diskutiert und kam zum Ergebnis, dass der Bedarf schwach, aber absehbar ist. Die Themengruppe favorisiert die Erweiterung mit Drittstandards.

Beschluss

Die Anforderungen werden wieder an der Sitzung der Fachgruppe im November 2015 diskutiert.

4 eCH-0164 Lebenszyklus

Information: französische Version ist jetzt verfügbar

5 eCH-0165 SIARD

Addendum

Mit dem Addendum ist nebst dem unkomprimierten Zip auch die Datenkompression *Deflate* erlaubt. SIARD Suite konnte von Anfang an *deflate* komprimierte SIARD-Dateien lesen. Das Erstellen von *deflate* komprimierten SIARD-Dateien wird jedoch von den Tools noch nicht unterstützt.

Neue Version

Am 19. November 2014 beschloss die Fachgruppe auf diese Sitzung hin alle RFCs für eine Version 2.0 zu sammeln. Zwischenzeitlich wurde auch eine neue Version nötig, um SIARD in das europäische Programm E-ARK und in die Programme von KEEPSolutions einzubinden.

Folgenden *Requests for Change* liegen vor, wobei der RFC 2015-29 aus dem E-ARK Projekt kommt.

RFC 2013-22

Namespace von bar.admin.ch in ech.ch ändern.

Der Namespace muss im Falle, dass SIARD ein ISO-Standard wird, nochmals geändert werden. Deshalb wird vorgeschlagen, bar.admin.ch beizubehalten oder einen rein technischen Namespace zu wählen.

Das Bundesarchiv hält die Marke auf SIARD in Europa und den USA, es könnte die Domain siard.org registrieren lassen und für die Verwendung als Namespace freigeben.

Beschluss

Es wird abgeklärt, was die Vorgaben von ISO sind.

RFC 2013-23

Soll eine Kann-Anforderung als leeres Feld programmiert werden müssen oder kann das Feld ganz weggelassen werden?

Beschluss

Wird umgesetzt.

RFC 2014-104

Damit SIARD grosse Datenmengen erreicht werden können, ist es nötig, eine Methode zu definieren, mit der Files gesplittet werden sollen.

Beschluss

Die Standard-Funktion von zip verwenden, um die Datei zu splitten.

RFC 2014-106

eCH-0014 macht Vorschläge von Formaten und Verwendungszwecken. eCH-0014 schlug für den Austausch zwischen Datenbanken nur csv vor.

Um kohärent mit eCH-0165 zu bleiben, müsste im Standard erwähnt werden, dass SIARD auch zum Austausch verwendet werden kann, primär aber ein Archivformat ist.

Beschluss

Die Empfehlung als Austauschformat könnte weitere Anforderungen an SIARD nach sich ziehen. Deshalb soll der Hinweis in abgeschwächter Form oder als Fussnote platziert werden.

Es muss darauf geachtet werden, dass es als „vereinfachtes Austauschformat“ verstanden wird, als Vereinfachung in dem Sinne einer bewussten Reduktion der Vielfalt der Datenbank auf eine einfachere Darstellungsweise, welche die IST-Situation abdeckt.

RFC 2014-110

SQL:1999 Kompatibilität

User defined Types, Distinct Type oder *Structured Type* eröffnet der Programmierung eine Möglichkeit, bei der ein Feld nicht nur einen Inhalt, sondern auch eine Funktionalität hat, mit der z.B. weitergerechnet werden kann.

Das Bundesarchiv hat eine Studie dazu in Auftrag gegeben und schlägt eine rückwärtskompatible Lösung vor.

Beschluss

Der RFC wird ausgeführt. Der Inhalt des Papers wird im Standard abgebildet.

RFC 2015-11

Hier handelt es sich um eine reine Korrektur.

RFC 2015-12

Formaterkennung

SIARD im komprimierten Zustand wird nicht erkannt. Mit der Platzierung einer leeren metadata.xsd-Datei im Header der SIARD Datei könnte das Format auch im komprimierten Zustand erkannt werden, weil das ZIP Verzeichnis nicht komprimiert wird. Das SIARD-Erzeugungsprogramm, müsste dahingehend angepasst und das neue Format in PRONOM erfasst werden. Die Verwechslungsgefahr mit dem SIP wird ausgeschlossen, da der Header des SIP kein metadata.xsd enthält sondern ein metadata.xml.

Beschluss

Die Datei soll eine Erklärung über den Sinn und Zweck von sich selbst enthalten.

RFC 2015-13

Bis anhin hat jedes Programm zu Erzeugung von SIARD Dateien ein eigenes Data Type Mapping zwischen dem SQL Dialekt des jeweiligen Quellsystem, SQL:1999 und XML definiert. Wenn dieses Mapping nicht einheitlich definiert ist, können unterschiedliche SIARD Dateien aus der gleichen Quelle, je nach verwendeten Tools, entstehen.

Keine Diskussion

Beschluss

Das Data Type Mapping gehört integral zur Formatspezifikation und muss mit jeder neu unterstützten Datenbank nachgeführt werden.

RFC 2015-14

Fixity

Die Unveränderbarkeit der Datei wird meistens nicht mehr in der Datei selbst festgehalten, sondern vom System gewährleistet. Die Funktion ist überlebt, muss aber nicht verboten werden.

Beschluss

Deprecated

RFC 2015-15

Die in BLOBS erlaubten Binärdateiformate sollen auf archivtaugliche Formate beschränkt werden. Es besteht Handlungsbedarf, denn BLOBS müssen nicht einmal vollständige Formate sein, es kann z.B. ein Datastream sein, dessen Format sogar erst ermittelt werden muss.

Diskussion

Die Einschränkung ist noch nicht praktikabel, denn es gibt noch keine Umsetzung im SIARD-Erzeugungsprogramm. Es müsste abgeklärt werden, ob das SIARD-Erzeugungsprogramm ein Migrationstool bräuchte und ob dieses im Programm oder in der Validation angesiedelt sein sollte. Bei der Validierung und Formaterkennung ist die Frage noch offen, ob auch nicht-erkannte BLOBS mitgenommen werden sollen.

Parallelen zum Web oder PDF Bereich. Prüfen, erkennen, reinschreiben.

Beschluss

Die KOST macht einen Vorschlag. Die Empfehlung könnte sein, dass dokumentiert werden soll, um was es sich handelt und dass auch unerkannte BLOBS mitgenommen und als solche beschrieben werden.

RFC 2015-29

Externer Speicherort für BLOBS und LOBS

Neu soll es auch möglich sein BLOBS und LOBS extern, das heisst ausserhalb der SIARD Datei zu speichern. Die Referenz auf die externe Datei erfolgt mit einer URL oder einer Dateireferenz in table.xml. Diese Lösung erlaubt eine separate Bewirtschaftung der BLOB Dateien (z.B. Office Dokumente oder Bilder) im Archivfindmittel, d.h. diese Dokumente können einzeln im Findmittel erschlossen und zugegriffen werden. Die Formatmigration ist zudem einfacher. Auf die Fileablage kann von SIARD oder vom Findmittel zugegriffen werden.

Der Request ist von dänischer Seite her sehr dringlich. Wenn er von der Fachgruppe nicht zugelassen wird, ist damit zu rechnen, dass es einige Anwender trotzdem machen.

Diskussion

Problematisch ist die Erhaltung die Referenzen von einem System zum andern. Der RFC steht in engem Zusammenhang mit RFC 2015-15. Es ist ein de Facto Bewertungsentscheid, ob man die Formate mit ins Archiv nimmt oder nicht. Das Risiko wird aufgeteilt *pour le bien ou pour le mal*.

Vorschlag und Beschluss

Minimalstandards für die Speicherung der externen Files aufstellen.

Beim SIARD-Validator als erlaubt/nicht erlaubt konfigurieren.

Anmerkung anbringen was BLOBS sind, z.B. .jpg.

RFC 2015-30

Wurde bereits behandelt siehe Namespace

Beschluss über alle RFC

Die KOST formuliert die RFC und integriert sie in die eCH-0165 Spezifikation. Es wird eine Version der eCH-0165 Spezifikation zur Verfügung gestellt in der die Änderungen gut sichtbar sind.

6 eCH-0175

Sehr positives Feedback von eCH. eCH-0175 würde sogar als Best-Practice akzeptiert.

Die Finanzierung der Übersetzung ist noch offen.

7 eCH-0193 Anforderungen DIP

Die Themengruppe dankt der Fachgruppe für die rege Diskussion und die wertvollen Anmerkungen. Im Folgenden sind nur die paar markantesten Diskussionspunkte aufgelistet.

Die Mitteilung von eCH war, dass der ursprüngliche Arbeitstitel keinen Fachterminus enthalten darf.

Diskussion

Titelvorschlag: Anforderungen an das Auslieferungsinformationspaket

Es muss Eindeutigkeit geschaffen werden bei den folgenden zwei Begriffspaaren:
Informationsobjekt – *data object* und Unterlagen – Primärdaten

Es wurde darüber diskutiert, ob man nie direkten Zugriff auf ein AIP haben darf oder ob es mit Systemen mit verbesserten Security Funktionen möglich sein dürfte.

Es wurde präzisiert, dass die Maschinenlesbarkeit nicht vom DIP selbst sichergestellt wird, sondern aus dem AIP übernommen wird respektive erhalten bleibt.

Beschluss

Zurück an die Arbeitsgruppe zur nochmaligen Überarbeitung anhand der Diskussion aus der Fachgruppe. Dabei sollen die deutschen OAIS Begriffe von Nestor verwendet werden und keine Eigenkreationen.

Das überarbeitete Hilfsmittel soll an der nächsten FG-Sitzung nochmals diskutiert werden.

Varia

Nächste Sitzung

Montag 9. November 2015

Mutationen

Ausgetreten sind: Niklaus Bütikofer, Unibe; Stephan Heuscher, Bedag und Eveline Isler, StAZH

Neu eingetreten: Olaf Hamberger, Ruf Informatik AG; Daniel Meier, Ruf Informatik AG; Peter Rohrbacher, StASZ; Martin Schlatter, swisstopo; Florian Vionnier, StAVS; Alain Mast, BAR

Für das Protokoll:

Bern, 24.06.2015

Brigitte Sacker