Ein Gemeinschaftsunternehmen von Schweizer Archiven

# Descriptive Information – xlsadg 2.1 Data Dictionary & XML-Schemabschreibung

#### Inhaltsverzeichnis

Einleitung	
Änderungen zur Version 1.6	
Änderungen zur Version 2.0	4
Data Dictionary	5
Zusammenfassung	
Repräsentation in XML	
Attribute	17
Vererbung und Aggregation: obligation	18
(1.3) Entstehungszeitraum / Laufzeit: dates	19
(1.5) Umfang: extentMedium	20
(3.1) Form und Inhalt: scopeContent	
(4.1) Zugangsbestimmungen: accessConditions	
Erweiterung: additionalData	
Das XMI, Schema im Detail	

# **Einleitung**

Das vorliegende Dokument spezifiziert die Schnittstelle für *Descriptive Information*, d.h. für den Import von Verzeichnungsdaten in ein Archivinformationssystem (AIS) oder den Austausch von Findmittelinformationen. Sein Ziel ist primär die weitgehend automatische Befüllung von Archivinformationssystemen unterschiedlicher Hersteller im Übernahmeprozess.

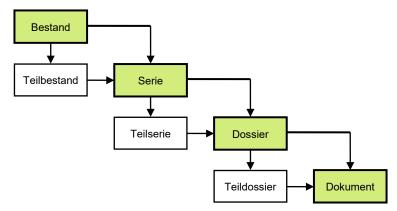
Das untenstehend spezifizierte *Data Dictionary* beschreibt das Datenmodell einer Verzeichnungseinheit, d.h. ein Set von Metadatenelementen, welches einen Verzeichnungsknoten auf einer beliebigen Stufe einer archivischen Tektonik beschreibt. Innerhalb einer Tektonik sind Verzeichnungseinheiten hierarchisch angeordnet<sup>1</sup>. Die gleiche hierarchische Ordnung gilt für die Schnittstelle.<sup>2</sup> Für die Übernahme von *Descriptive Information* müssen deshalb mehrere Verzeichnungseinheiten hierarchisch ineinander verschachtelt werden.<sup>3</sup>

<sup>&</sup>lt;sup>1</sup> Siehe für die hierarchische Abfolge der Verzeichnungsstufen Arbeitsgruppe Normen und Standards des Vereins Schweizerische Archivarinnen und Archivare VSA, Schweizerische Richtlinie für die Umsetzung von ISAD(G) – *International Standard Archival Description (General)*, Zürich/Bern 2009, Abb. 1 S. 8, online unter

http://www.vsa-aas.org/fileadmin/user\_upload/texte/ag\_n\_und\_s/Richtlinien\_ISAD\_G\_VSA\_d.pdf 

2 Je nach verwendeter Ablieferungsschnittstelle sind nicht alle Metadaten auf allen Hierarchieebenen gleichermassen vorhanden (siehe dazu Vererbung & Aggregation) auch können 
ISAD(G)-konforme Metadaten in der SIP-Schnittstelle je nach Hierarchiestufe unterschiedlich 
bezeichnet sein (siehe dazu das jeweilige Metadaten Mapping SIP-Schnittstelle – Descriptive 
Information, z.b. eCH-0160 – xldadg und EAD – xlsadg in "eCH-0160: Archivische Ablieferungsschnittstelle (SIP), Anhang I"

https://www.ech.ch/vechweb/page?p=dossier&documentNumber=eCH-0160&documentVersion=1.1 
<sup>3</sup> Es ist zwar theoretisch möglich, nur jeweils eine einzige Verzeichnungseinheit zu übernehmen, aber in der Regel umfasst eine Übernahme mehrere hierarchisch angeordnete Verzeichnungseinheiten.



Verzeichnungsstufen in ihrer hierarchischen Abfolge⁴

Für die einzelnen Elemente der Schnittstelle für *Descriptive Information* sind nachstehend *Nummer, Name, Zweck, Herkunft, Wertebereich, Vorkommen, Abbildung in XML* und *Kommentar* vermerkt.

**Nummer, Name und Zweck** sind (wo möglich) zitiert aus der Definition von ISAD(G). **Herkunft** verweist je nach betrachtetem Prozess auf unterschiedliche Quellen, es kann Information aus dem SIP sein, Information sein die beim *Ingest* Prozess generiert wird; es kann Metainformation sein die beim *Ingest* implizit zur Verfügung steht, aber nicht im SIP verzeichnet ist oder auch Information sein die erst später bei der Erschliessung ergänzt werden. Konkret ist hier für Herkunft, wo immer möglich, das Metadatum innerhalb eines SIPs angegeben<sup>5</sup>, aus dem die jeweilige Information gewonnen wird (explizite Herkunft). *Implizit* bedeutet in diesem Zusammenhang, dass sich der Wert aus *Aggregation* oder *Vererbung* ergibt. *Manuell verzeichnet* bedeutet, dass es unmöglich ist, das Element automatisch zu befüllen, es muss also beim *Ingest* oder bei der Erschiessung erfasst werden.

Der Wertebereich wird wo nötig und möglich spezifiziert.

Das **Vorkommen** von Elementen kann *obligatorisch*, *fakultativ*, *vererbt* oder *aggregiert* sein. Die beiden letzteren Möglichkeiten sind spezielle Ausprägungen von *obligatorisch*, die im Kontext der Verzeichnungshierarchie interpretiert werden müssen. *Vererbt* bedeutet, dass das Element auf mindestens einer Hierarchiestufe vorhanden sein muss und von dort aus auf die darunter liegenden Stufen vererbt wird; auf den darüber liegenden Hierarchiestufen ist das Element fakultativ. (*Fakultative* Metadaten werden ebenfalls vererbt, wenn sie vorhanden sind.) *Aggregiert*<sup>6</sup> bedeutet, dass das Element auf mindestens einer Hierarchiestufe vorhanden sein muss und die Werte auf die jeweils darüber liegenden Stufen logisch zusammengeführt werden; auf den darunter liegenden Hierarchiestufen ist das Element fakultativ.

**Abbildung in XML** definiert wie die einzelnen Elemente der Schnittstelle im xlsadg XML Schema abgebildet werden sollen. Im Falle, dass eine einfache Abbildung nicht möglich ist (d.i. 1.3 Entstehungszeitraum / Laufzeit ,1.5 Umfang, 3.1 Form und Inhalt und 4.1 Zugangsbestimmungen) wird auf ein entsprechendes Kapitel in *Repräsentation in XML* verwiesen.

<sup>&</sup>lt;sup>4</sup> Siehe Schweizerische Richtlinie für die Umsetzung von ISAD(G), Abb.1 (oben Anm. 1)

<sup>&</sup>lt;sup>5</sup> Ursprünglich bezog sich "Herkunft" auf das im *bentō* Projekt 2010 entwickelte AIP/SIP Modell, <a href="http://kost-ceco.ch/cms/index.php?bento\_de">http://kost-ceco.ch/cms/index.php?bento\_de</a> in der jetzigen Fassung auf die Archivische Ablieferungsschnittstelle (SIP) eCH-0160.

<sup>&</sup>lt;sup>6</sup> Aggregation: lateinisch: aggregatio: Anhäufung, Vereinigung.

Abschliessend sind jeweils **Kommentare** zum Gebrauch des Elements im Kontext digitaler Unterlagen aufgeführt.

# Änderungen zur Version 1.6

Im Rahmen des *bentō*-Projekts ist 2010 dieses *Data Dictionary* und dazu das Schema xlsadg 1.6 zur Beschreibung von *Descriptive Information* entstanden. Die Bedürfnisse in der konkreten Anwendung haben dann zu mehr oder weniger koordinierten Änderungen und Erweiterung bis xlsadg 1.9 geführt<sup>7</sup>. Das Stadtarchiv Zürich hat sich zusammen mit den Staatsarchiven Appenzell Ausserrhoden, Basel-Stadt, Bern, Luzern, St. Gallen und Thurgau und der KOST-Geschäftsstelle darum bemüht, diese neuen Anforderungen und Wünsche zusammenzuführen und xlsadg 2.0 zu spezifizieren.

Das neue Schema ist weitestgehend rückwärtskompatibel, die Änderungen zur Version 1.6 sind hier zusammengefasst und in den folgenden Text integriert:

- Bezug genommen wird neu wo möglich auf eCH-0160 SIP statt auf bentō-SIP/AIP.
- Elemente, die auf der gleichen Ebene mehrfach in unterschiedlicher Ausprägung vorkommen können, sind im Schema wiederholbar (z.B. Entstehungszeitraum / Laufzeit, isad:dates)
- Herkunft, isad:anyElement@origin, bezeichnet neu nicht nur die Informationsquelle "Archiv" oder "SIP" sondern kann auch den genauen xPath-Verweis in das Quellpaket enthalten. [Das Attribut @archivalOrigin (boolean) soll nicht weiter benutzt werden]
- Zeitraum "von bis", genaues Datum und circa-Datum werden im Schema genauer formuliert: isad:fromDate, isad:toDate, isad:pointofTime und isad:dates@circa. [die Datumselemente isad:fromYear und isad:toYear sollen nicht weiter benutzt werden]
- Metadaten als Schlüssel-Werte-Paare: isad:additionalData/isad:mdWrap ist neu wiederholbar und enthält den Schlüssel als Attribut isad:property@key und den Wert als Elementinhalt. [Das Attribut @value soll nicht weiter benutzt werden]
- additionalData Schlüssel-Werte-Paare isad:property sind zudem auch rekursiv wiederholbar, so dass mehrdimensionale Tabellen abgebildet werden können.
- Wertelisten können neu auch durch Freitext erweitert werden: isad:scope und isad:physTech.
- Neu kann //isad:extentMedium/isad:description als Text Element für unstrukturierte Angaben zu Umfang (1.5 Menge und Abmessung) verwendet werden.
- Die Zugangsbestimmungen //isad:accessConditions werden mit drei weiteren Unterelementen ergänzt, die vor allem im Schweizerischen Bundesarchiv Anwendung finden: //isad:retentionPeriod, //isad:retentionPeriodConditions und //isad:retentionPeriodNotes.
- Nicht rückwärtskompatibel ist die Änderung von isad:description[s]Dates zu isad:descriptionDates, ein Schreibfehler in Schema xlsadg v1.6.xsd.

.

<sup>&</sup>lt;sup>7</sup> Siehe dazu die Darstellung <a href="http://kost-ceco.ch/cms/index.php?xisadg\_de">http://kost-ceco.ch/cms/index.php?xisadg\_de</a> und die detaillierten Versionen unter <a href="https://github.com/KOST-CECO/xlsadg">http://kost-ceco.ch/cms/index.php?xisadg\_de</a> und die detaillierten Versionen unter <a href="https://github.com/KOST-CECO/xlsadg">https://github.com/KOST-CECO/xlsadg</a>

# Änderungen zur Version 2.0

Anschliessend an die Veröffentlichung von xlsadg 2.0 Ende 2017 sind noch Änderungswünsche aufgetaucht, die zu einer neuen Version geführt haben. Bei den meisten Änderungen handelt es sich um technische Korrekturen, welche vollständig rückwärtskompatibel sind:

- Das Attribut "origin" bei //isad:referenceCode ist auf Ebene Bestand "archival", bei tieferen Verzeichnungseinheiten hingegen "ingest"
- Die Schemavalidierung von //isad:xmlData ist neu "lax" statt "strict", das heisst eingebettete XML Schema werden bei der Gesamtvalidierung eines Schemas nicht mehr überprüft.
- Bei verschachtelten bzw. mehrdimensionalen Tabellen mit nested //isad:property ist das Attribut "origin" auf jeder Ebene erlaubt. Damit können eCH-0160 v1.1 Vorgang und Aktivität korrekt abgebildet werden.
- Das Attribut "xpointer" in //isad:arrangement erlaubt Verweise auf ein Objekt in der PREMIS Datei. Damit ist es möglich Metainformationen vom Typ 3.3.4 Arrangement auch auf Dossier Ebene mit einzelnen Dateien zu verbinden. Z.B. können in einem Dossier ohne Dokumentebene Original-Dateipfade von Dateien im Dossier festgehalten werden.
- Bei //isad:primaryDataLocator soll das Attribut "uri" direkt auf die betreffende Datei verweisen. Bei //isad:secondaryDataLocator soll das Attribut "xpointer" indirekt, d.i. via PREMIS auf eine Datei verweisen. Ein Verweis via Elementinhalt ist nicht mehr vorgesehen.

Die grösste Änderung betrifft die hinterlegten Wertelisten:

Sämtliche Wertelisten sind neu durchgängig in Englisch formuliert, auch die beiden Bereiche //isad:descriptionLevel und //isad:scopeContent/isad:content@contentQualifier.

Die englischen Begriffe in den Wertelisten sind als Codewerte in der XML Datei für den bezeichneten Inhalt zu verstehen – in einer Anzeige können und sollen dann durchaus die gängigen Deutschen oder Französischen Begriffe Verwendung finden.

Alle Wertelisten sind, obwohl Enumeration Listen, sind für künftige Anwendungen frei

Alle Wertelisten sind, obwohl *Enumeration* Listen, sind für künftige Anwendungen frei erweiterbar.

# **Data Dictionary**

Die einzelnen Elemente einer Verzeichnungsebene:

Nummer 1.1

Name Signatur

Zweck Identifizierung des Archivs und Verknüpfung des Archivguts mit der zuge-

hörigen Verzeichnung.

Herkunft Metadatum "Archivsignatur"

Wertebereich Freitext Vorkommen vererbt

Abb. in XML //isad:referenceCode

Kommentar Gemeint ist selbstredend die Archivsignatur (nach den Vorgaben des ent-

sprechenden Archivs), nicht das Aktenzeichen im originalen System. Im Ingest Fall ist das die Einfügeposition in die Archivtektonik und wird ma-

nuell erfasst und anschliessend vererbt.

Nummer 1.2
Name Titel

Zweck Benennung der Verzeichnungseinheit

Herkunft SIP-Metadatum "Titel"

Wertebereich Freitext Vorkommen obligatorisch Abb. in XML //isad:title

Kommentar Keine Präzisierung notwendig.

Nummer 1.3

Name Entstehungszeitraum / Laufzeit

Zweck Identifizierung des Entstehungsdatums bzw. -zeitraums der in der Ver-

zeichnungseinheit enthaltenen Unterlagen.

Herkunft SIP-Metadatum "Zeitraum" auf Stufe Ablieferung, Dossier oder Dokument

Wertebereich Datum gemäss ISO 8601<sup>8</sup> Vorkommen aggregiert oder vererbt

Abb. in XML //isad:dates siehe dazu: (1.3) Entstehungszeitraum / Laufzeit: dates

Kommentar "Die Entstehungsdaten der in der Verzeichnungseinheit enthaltenen Un-

terlagen sind als Einzeldatum oder Zeitraum anzugeben" (ISAD[G]). Das Element muss in der Schnittstelle entweder in der Präzision "Datum", "Jahr" oder "Jahrhundert" jeweils mit Start- und Endpunkt angegeben werden, allenfalls ergänzt mit dem Hinweis "circa". Für Einzeldaten fallen Start- und Endpunkt zusammen. Zu jedem Dokument muss im SIP ein Datum übernommen werden (wenn möglich das Abschlussdatum). Das

Element "Entstehungszeitraum/Laufzeit" auf den höheren Verzeichnungsstufen wird aus den Daten der hierarchisch untergeordneten Dokumenten

oder Verzeichnungsstufen aggregiert.

xlsadg\_DataDictionary\_v2.1.doc 01.03.2018

<sup>&</sup>lt;sup>8</sup> Siehe <a href="http://de.wikipedia.org/wiki/ISO\_8601">http://de.wikipedia.org/wiki/ISO\_8601</a> und <a href="http://www.cl.cam.ac.uk/~mgk25/iso-time.html">http://www.cl.cam.ac.uk/~mgk25/iso-time.html</a>.

Nummer 1.4

Name Verzeichnungsstufe

Zweck Identifizierung der Verzeichnungsstufe

Herkunft explizit wird Provenienz zu Bestand oder Teilbestand und Dossiers zu

Dossier. *Implizit* werden aus den dazwischen liegenden Stufen der Ordnungssystemhierarchie im SIP Serien, bzw. Fond, Serie und Teilserie, je

nach Tektonik im Archiv.

Wertebereich Werteliste für die Verzeichnungsstufe (fonds, sub-fonds, series, sub-

series, file, sub-file, item und zusätzlich representation)

[die englischen Begriffe sind hierbei Platzhalter für Bestand, Teilbestand,

Serie, Teilserie, Dossier, Teildossier und Dokument]

Vorkommen obligatorisch und vererbt Abb. in XML //isad:descriptionLevel

Kommentar Gemeint ist selbstredend die archivische Verzeichnungsstufe, nicht die

Registraturstufe.

Nummer 1.5

Name Umfang (Menge und Abmessung)

Zweck Identifizierung (a) des physischen Umfangs und (b) der Archivalienart der

Verzeichnungseinheit.

Herkunft SIP-Metadatum "Umfang" bzw. Implizit im SIP vorhanden (d.i. aus dem

Inhalt zu berechnen).

Wertebereich Zahl für Umfang mit Werteliste für die Einheit (kB, MB, GB, TB, Ifm, kg,

m3) und Freitext für die Archivalienart

Vorkommen aggregiert

Abb. in XML //isad:extentMedium siehe dazu: (1.5) Umfang: extentMedium

Kommentar Dieses Element kann sowohl das Datenvolumen bezeichnen (für digitale

Unterlagen als eine Anzahl kB, MB, GB oder TB) als auch die Anzahl Objekte einer Archivalienart (für digitale Unterlagen wird die Archivalienart

"Dateien" standardmässig vorausgesetzt).

Nummer 2.1

Name der Provenienzstelle

Zweck Identifizierung der Provenienzstelle(n), bei der (denen) die Verzeich-

nungseinheit entstanden ist.

Herkunft SIP-Metadatum "Aktenbildner Name"

Wertebereich Freitext, wenn möglich gemäss kontrolliertem Vokabular der Provenienz-

stellen des Archivsprengels

Vorkommen vererbt

Abb. in XML //isad:creator

Kommentar Keine Präzisierung notwendig

Nummer 2.2

Name Verwaltungsgeschichte/Biografische Angaben

Zweck Information über Verwaltungsgeschichte der Provenienzstelle bzw. die

Biographie, wenn es sich um natürliche Personen handelt, zum besseren

Verständnis des zur Verzeichnungseinheit gehörenden Kontextes.

Herkunft SIP-Metadatum "Geschichte Aktenbildner" Wertebereich Freitext

Vorkommen fakultativ

Abb. in XML //isad:adminBioHistory

Kommentar keine Präzisierung notwendig

Nummer 2.3

Name Bestandsgeschichte

Zweck Information über den Wechsel der Eigentums- und Besitzverhältnisse der

Verzeichnungseinheit, die für deren Authentizität, Integrität, Vollständig-

keit und Interpretation von wesentlicher Bedeutung sind.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:archivalHistory

Kommentar keine Präzisierung notwendig

Nummer 2.4

Name Abgebende Stelle

Zweck Darstellung der Umstände, die mit der direkten Übernahme der Verzeich-

nungseinheit von der abgebenden Stelle verbunden sind.

Herkunft SIP-Metadatum "Abliefernde Stelle"

Wertebereich Freitext, wenn möglich gemäss kontrolliertem Vokabular

Vorkommen vererbt

Abb. in XML //isad:acqInfo

Kommentar Dieses Element wird als notwendig erachtet, weil davon auszugehen ist,

dass in der digitalen Welt die abgebende Stelle häufiger als in der analo-

gen nicht mit der Provenienz identisch ist.

Nummer 3.1

Name Form und Inhalt

Zweck Feststellung von Hauptgegenstand und Form der Verzeichnungseinheit,

um Benutzern eine Beurteilung ihrer Relevanz zu ermöglichen.

Herkunft SIP-Metadatum "Form Inhalt" und SIP-Metadatum "Inhalt".

Wertebereich Werteliste für Form (Text, Bild, Audio, Video, strukturierte Daten [cf.

KaD<sup>9</sup>]) oder Freitext Freitext für Inhalt.

Vorkommen aggregiert oder manuell verzeichnet

Abb. in XML //isad:scopeContent siehe dazu: (3.1) Form und Inhalt: scopeContent

Kommentar Form soll sich im digitalen Bereich auf die Formatkategorie der Unterlagen

beschränken. Da Formatkategorien und Dateiformate in einem SIP nicht obligatorisch verzeichnet sein müssen, gibt es jedoch keine Möglichkeit,

-

<sup>&</sup>lt;sup>9</sup> Katalog archivischer Dateiformate der KOST, http://www.kost-ceco.ch/wiki/whelp/KaD.

dieses Element automatisch zu befüllen<sup>10</sup>. Die Aggregierung auf höhere Ebenen muss relativ zu den unterliegenden Formatkategorien und Dateiformaten erfolgen.

Inhalt kann aus dem SIP übernommen werden oder wird bei der Erschliessung manuell erfasst.

Nummer 3.2

Name Bewertung und Kassation

Zweck Bereitstellung von Informationen über jede vorgenommene Bewertung

und Kassation.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:appraisalDestruction

Kommentar keine Präzisierung notwendig.

Nummer 3.3

Name **Neuzugänge** 

Zweck Angaben für den Benutzer über mögliche Veränderungen im Umfang der

Verzeichnungseinheit.

Herkunft manuell verzeichnet

Wertebereich Freitext
Vorkommen fakultativ
Abb. in XML //isad:accruals

Kommentar keine Präzisierung notwendig.

Nummer 3.4

Name Ordnung und Klassifikation

Zweck Bereitstellung von Informationen über die Ordnung und Klassifikation der

Verzeichnungseinheit.

Herkunft SIP Metadatum "Ordnungssystem Name" oder manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:arrangement

Kommentar Hier kann auch Information über die den Aufbau der Ablage, Pfad und Da-

teiname, der Originalablage erfasst werden. 11

xlsadg\_DataDictionary\_v2.1.doc 01.03.2018

<sup>&</sup>lt;sup>10</sup> Falls das Dateiformat im SIP verzeichnet ist, könnte die Formatkategorie mittels eines Mappings (PRONOM-ID →) MIME-Type → Formatkategorie ermittelt werden.

<sup>&</sup>lt;sup>11</sup> Siehe dazu die Englische Erläuterung zu ISAD(G) 3.3.4:

To provide information on the internal structure, the order and/or the system of classification of the unit of description. Specify the internal structure, order and/or the system of classification of the unit of description. Note how these have been treated by the archivist. For electronic records, record or reference information on system design.

https://www.ica.org/sites/default/files/CBPS 2000 Guidelines ISAD%28G%29 Secondedition\_EN.pdf

4.1 Nummer

Name Zugangsbestimmungen

Angabe derjenigen Bestimmungen, die den Zugang zur Verzeichnungs-Zweck

einheit einschränken oder beeinflussen.

SIP-Metadaten "klassifizierungskategorie", "datenschutz", "oeffentlich-Herkunft

keitsstatus", "schutzfristenkategorie" und "schutzfrist".

Wertebereich Je nach Archiv unterschiedliche Werteliste (z.B.: öffentlich zugänglich,

normale Schutzfrist, verlängerte Schutzfrist, ausserordentliche Schutzfrist)

Vorkommen aggregiert

Abb. in XML //isad:accessConditions

siehe dazu: <isad:arrangement xpoin-

ter="premis.xml#xpointer(//xmlID('object-id\_22')" origin="ingest">

//Meier/Briefablage/Übersetzung Einladung ws1.pdf</isad:arrangement>

<isad:arrangement xpointer="premis.xml#xpointer(//xmlID('object-id 23')" origin="ingest"> //Meier/Briefablage/Übersetzung Protokoll ws1.pdf</isad:arrangement>

In diesem Beispiel wird auf der Ebene Dossier für zwei Dateien die Metainformation originaler Dateipfad in der Ablage vor der Ablieferung angegeben, «origin» der Information ist in diesem Fall «ingest».

Es ist auch denkbar, dass ein Dokument mit mehreren Dateien einer Ablage verbunden ist (=verschiedene Repräsentationen des gleichen Dokuments) ohen dass eine weiter isad:archivalDescription Ebene vom Typ «representation» existiert, dann kann der selbe Mechanismus verwendet werden.

#### (4.1) Zugangsbestimmungen: accessConditions

Kommentar Aus dem SIP können Angaben zu verschiedenen Faktoren übernommen werden, die auf die Zugangsbestimmungen einen Einfluss haben: Geheimhaltung (bzw. Klassifizierungskategorie), Datenschutzstufe (enthält schützenswerte Personendaten oder nicht), Öffentlichkeitsstatus (Dokument bzw. Rubrik wurde gemäss Öffentlichkeitsgesetz als öffentlich zugänglich bezeichnet). Diese Informationen sind aus den Primärdaten nicht automatisiert ableitbar und müssen deshalb in der DI-Schnittstelle übernommen werden. Die konkrete Festlegung der Zugangsbestimmungen ist von diversen archivspezifischen Vorgaben abhängig.

Es gibt auch Fälle, wo Schutzfristen und Schutzfristenkategorie bereits im SIP eingetragen sind<sup>12</sup> oder als *Descriptive Information* festgehalten werden sollen<sup>13</sup>.

<sup>&</sup>lt;sup>12</sup> Dieser Fall betrifft das Bundesarchiv, wo die Schutzfristen vom abliefernden Amt vorgegeben werden können.

<sup>&</sup>lt;sup>13</sup> In diesem Fall sollen *Descriptive Information* von einem Findmittel in ein anderes Findmittel übernommen werden.

Nummer 4.2

Name Reproduktionsbestimmungen

Zweck Information über Beschränkungen bei der Reproduktion der Verzeich-

nungseinheit.

Herkunft Unklar, kann nicht automatisch aus dem SIP übernommen werden.

Wertebereich Freitext oder Werteliste

Vorkommen fakultativ

Abb. in XML //isad:reproConditions

Kommentar Gemeint sind laut ISAD(G) insbesondere urheberrechtliche Einschrän-

kungen.

Nummer 4.3

Name Sprache/Schrift

Zweck Identifizierung der in der Verzeichnungseinheit enthaltenen Sprache(n),

Schriftarten und Zeichensysteme.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:languageScripts

Kommentar keine Präzisierung notwendig

Nummer 4.4

Name Physische Beschaffenheit und technische Anforderungen

Zweck Bereitstellung von Informationen über wichtige physische Besonderheiten

oder technische Anforderungen, die die Benutzung der Verzeichnungs-

einheit beeinflussen.

Herkunft SIP-Metadatum "Erscheinungsform"

Wertebereich Werteliste (digital, analog, hybrid) und Freitext

Vorkommen aggregiert
Abb. in XML //isad:physTech

Kommentar Dieses Element soll verwendet werden, um die Unterlagen als digital,

analog oder hybrid zu identifizieren.

Nummer 4.5

Name Findhilfsmittel

Zweck Identifizierung aller für die Verzeichnungseinheit vorhandenen Findhilfs-

mittel.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:findingAids

Kommentar keine Präzisierung notwendig

Nummer 5.1

Name Aufbewahrungsort der Originale

Zweck Nachweise über die aufbewahrende Institution, die Zugänglichkeit oder

die Vernichtung der Originale, falls es sich bei der Verzeichnungseinheit

um eine Reproduktion handelt.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:existenceLocationOriginals

Kommentar Unter "Reproduktion" wird beispielsweise eine Retrodigitalisierung ver-

standen.

Nummer 5.2

Name Kopien bzw. Reproduktionen

Zweck Verweis auf Kopien bzw. Reproduktionen der Verzeichnungseinheit und

ihre Verfügbarkeit.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:existenceLocationCopies
Kommentar keine Präzisierung notwendig

Nummer 5.3

Name Verwandte Verzeichnungseinheiten

Zweck Ermittlung von verwandten Verzeichnungseinheiten im selben Archiv oder

in anderen Archiven.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:relatedUnits

Kommentar keine Präzisierung notwendig

Nummer 5.4

Name Veröffentlichungen

Zweck Ermittlung von Veröffentlichungen, die unter Benutzung oder Auswertung

der Verzeichnungseinheit entstanden sind.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:publication

Kommentar keine Präzisierung notwendig

Nummer 6.1

Name Allgemeine Anmerkungen

Zweck Bereitstellung von Spezialinformationen und Angaben, die in keinem der

anderen Bereiche angebracht werden können.

Herkunft SIP-Metadatum "Bemerkung"

Wertebereich Freitext Vorkommen fakultativ Abb. in XML //isad:notes

Kommentar keine Präzisierung notwendig

Nummer 7.1

Name Informationen des Bearbeiters

Zweck Erläuterungen zur Verzeichnung und über den oder die Bearbeiter.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:archivistsNote

Kommentar keine Präzisierung notwendig

Nummer 7.2

Name Verzeichnungsgrundsätze

Zweck Benennung der bei der Verzeichnung angewandten Normen, Regeln und

Grundsätze.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:rulesConventions

Kommentar keine Präzisierung notwendig

Nummer 7.3

Name Datum oder Zeitraum der Verzeichnung

Zweck Alle Daten, die sich auf die Erstellung oder Veränderung der Verzeich-

nung beziehen, sind anzugeben.

Herkunft manuell verzeichnet

Wertebereich Freitext Vorkommen fakultativ

Abb. in XML //isad:descriptionDates

//isad:mutationDate

Kommentar keine Präzisierung notwendig

Nummer (additionalReference)

Name **Aktenzeichen im Quellsystem (recordReference)**Zweck Identifikator der Verzeichnungseinheit im Quellsystem

Herkunft SIP-Metadatum "Aktenzeichen"

Wertebereich Freitext Vorkommen vererbt

Abb. in XML //isad:recordReference

Kommentar Dieses Element ist notwendig, um eine schnelle Identifikation von archi-

vierten Unterlagen zu gewährleisten, die von der Provenienzstelle zur

Einsicht verlangt werden.

Nummer (additionalReference)

Name Referenz auf das AIP (aipReference)

Zweck Eindeutiger Identifikator des AIP im digitalen Magazin des Archivs

Herkunft wird im *Ingest*-Prozess erzeugt und verzeichnet Wertebereich Identifikator gemäss internen Regeln des Archivs

Vorkommen vererbt

Abb. in XML //isad:aipReference

Kommentar keine Präzisierung notwendig

Nummer (additionalReference)

Name Referenz auf Dateien (primaryDataLocator)

Zweck Referenz auf Primärdateien in Form einer URL.

Herkunft wird im *Ingest*-Prozess erzeugt und verzeichnet

Wertebereich URL / URI / relativer Dateipfad

Vorkommen optional

Abb. in XML //isad:primaryDataLocator

Kommentar keine Präzisierung notwendig

Nummer (additionalReference)

Name Referenz auf technische Metadaten (secondaryDataLocator)

Zweck Referenz in Form eines xpointer-Eintrags auf einen Objekteintrag in einer

premis- oder Imer-Datei, damit von dort auf Primärdateien und deren Ei-

genschaftsattribute zugegriffen werden kann.

Herkunft wird im Ingest-Prozess erzeugt und verzeichnet

Wertebereich xPointer Vorkommen optional

Abb. in XML //isad:secondaryDataLocator
Kommentar keine Präzisierung notwendig

Nummer (additionalData)

Name Metadaten als Schlüssel-Werte-Paare (mdWrap)

Zweck Zusätzliche, nicht ISAD(G)-konforme Metadaten als Schlüssel-Werte-Paar

(key/value pair property list)

Herkunft Zusatzdaten aus dem SIP

Wertebereich Strukturiert nach Vorgabe der abliefernden Quelle

Vorkommen fakultativ

Abb. in XML //isad:mdWrap siehe dazu: Erweiterte Referenzen:

Kommentar Möglichkeit, weitere Metadaten jeglicher Art aus dem Quellsystem via In-

*gest* oder SIP als *Key-Value*-Paare ins Findmittel zu übernehmen. Die Schlüsselbezeichnungen stammen aus dem Quellsystem oder sind im *In-*

gest-Prozess definiert worden.

Ein Value-Element kann wieder Key-Value-Paare enthalten, sodass auch

mehrdimensionale Tabellen abgebildet werden können.

Nummer (additionalData)

Name Metadaten nach externem Schema (xmlWrap)

Zweck Nach einem externen Metadatenschema organisierte Metadaten: DC

(Dublin Core), EAD (Encoded Archival Description) etc.

Herkunft Metadaten werden vom Quellsystem in einem spezifischen Metadaten-

schema verwaltet und zur Archivierung zur Verfügung gestellt

Wertebereich Externes Metadatenschema (DC, EAD, MARC, MODS, OTHER)

Vorkommen fakultativ

Abb. in XML //isad:xmlWrap siehe dazu: Erweiterte Referenzen:

Kommentar Metadaten, die im Quellsystem in einem eigenen, bekannten Datensche-

ma geführt oder in ein solches exportiert werden, können in diesem

Wrapper übernommen werden.

# Zusammenfassung

Tabellarische Zusammenfassung der Verzeichnungselemente nach ISAD(G) und ihr Vorkommen auf den verschiedenen Verzeichnungsstufen:

	Verzeichnungselement ISAD(G)	Bestand	Serie	Dossier	Dokument
1	Identifikation				
1.1	Signatur				
1.2	Titel				
1.3	Entstehungszeitraum / Laufzeit				
1.4	Verzeichnungsstufe				
1.5	Umfang (Menge und Abmessung)				
2	Kontext				
2.1	Name der Provenienzstelle				
2.2	Verwaltungsgeschichte / Biographische Angaben				
2.3	Bestandesgeschichte				
2.4	Abgebende Stelle				
3	Inhalt und innere Ordnung				
3.1	Form und Inhalt				
3.2	Bewertung und Kassation				
3.3	Neuzugänge				
3.4	Ordnung und Klassifikation				
4	Zugangs- und Benutzungsbedingungen				
4.1	Zugangsbestimmungen				
4.2	Reproduktionsbestimmungen				
4.3	Sprache / Schrift				
4.4	Physische Beschaffenheit und technische Anfor-				
	derungen				
4.5	Findhilfsmittel				
5	Sachverwandte Unterlagen				
5.1	Aufbewahrungsort der Originale				
5.2	Kopien bzw. Reproduktionen				
5.3	Verwandte Verzeichnungseinheiten				
5.4	Veröffentlichungen				
6	Anmerkungen				
6.1	Allgemeine Anmerkungen				
7	Verzeichnungskontrolle				
7.1	Information des Bearbeiters				
	Verzeichnungsgrundsätze				
7.3	Datum oder Zeitraum der Verzeichnung				
	Erweiterte Referenz				
	Aktenzeichen im Quellsystem (recordReference)				
	Referenz auf das AIP (aipReference)				
	Referenz auf Dateien (primaryDataLocator)				
	Referenz auf technische Metadaten				
	(secondaryDataLocator)				
	Erweiterte Metadaten				
	Metadaten als Schlüssel-Werte-Paare (mdWrap)				
	Metadaten nach externem Schema (xmlWrap)				

Legende	obligatorisch	vererbt	aggregiert

Das Vorkommen der Elemente: obligatorisch, vererbt oder aggregiert graphisch dargestellt:

	Verzeichnungsstufe					
	Dokument	Dossier	Serie	Bestand		
obligatorisch						
vererbt	<b>—</b> ○→	-0				
aggregiert				<b>←</b> ○		

# Repräsentation in XML

Für die Kommunikation zwischen *Ingest* und Findmittel wird die *Descriptive Information* durch einen XML-Datenstrom repräsentiert. Das zugrunde liegende Schema ist eine ISAD(G)-basierte Schemadefinition<sup>14</sup>, welche vom belgischen Projekt eDavid<sup>15</sup> erarbeitet und im *bentō*-Projekt erweitert wurde. Es handelt sich im Prinzip um die Abbildung der ISAD(G)-Entitäten auf Elemente mit dem jeweils englischen ISAD(G)-Namen. Die ISAD(G)-Elemente sind unter den jeweiligen Sammelbegriffen *(areas)* zusammengefasst und bilden zusammen die *//isad:archivalDescription*. Diese Verzeichnungseinheiten können beliebig hierarchisch verschachtelt werden.



Ausser dem obligatorischen *isad:identity* können alle anderen Sammelbegriffe (*areas*) in einer Verzeichniseinheit auch fehlen. Die Schemadatei heisst standardmässig *xlsadg v2.0.xsd*, der Namespace lautet *isad:*.

#### **Attribute**

Jedes ISAD(G)-Element hat das Attribut *isadId* für die ISAD(G)-Referenznummer, wenn möglich das Attribut *origin*, das angibt, ob das Archiv die Datenquelle ist oder sonst eine Referenz auf die Quelle im SIP (xPath zum SIP-Element), und das Attribut *obligation* (für das Vorkommen ["obligation" gemäss PREMIS]) mit dem Wertebereich "mandatory", "inherited", "aggregated" und "optional".

<sup>&</sup>lt;sup>14</sup> eDavid XML Schemas: http://www.expertisecentrumdavid.be/xmlschemas/

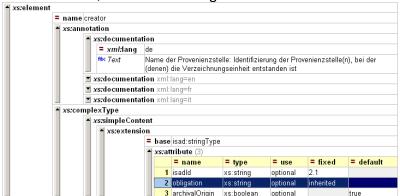
<sup>&</sup>lt;sup>15</sup> Expertisecentrum DAVID vzw is a centre of research and knowledge on digital archiving: <a href="http://www.edavid.be/eng/index.php">http://www.edavid.be/eng/index.php</a>

Das Attribut *archivalOrigin*, "ja/nein" soll nicht weiter gepflegt werden, siehe dazu die Erläuterungen in der Einleitung.

#### Vererbung und Aggregation: obligation

Vererbung und Aggregation sind Varianten zu obligatorisch (=mandatory), d.h. ein Wert ist weiter oben oder weiter unten in der Hierarchie obligatorisch und wird nach unten vererbt oder nach oben aggregiert. Vererbung und Aggregation können mit XML Schema nicht direkt modelliert werden, sie müssen prozedural implementiert werden.

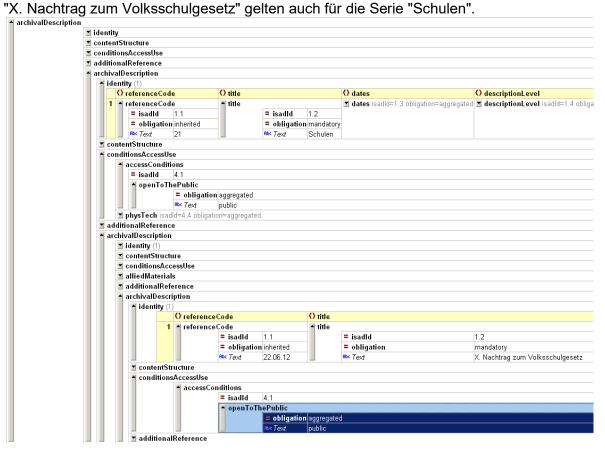
Beispiel für Vererbung: 2.1 Name der Provenienzstelle wird nur in der obersten Verzeichniseinheit erfasst, ist aber dort obligatorisch.



**Beispiel für Aggregation:** Im Gegensatz zu *obligatorisch*, *fakultativ* und *vererbt* ist die **Aggregation** datenabhängig:

- 1.3 Zeitraum wird nach jüngstem und ältestem Datum aggregiert.
- 1.5 Umfang wird aufsummiert oder addiert.
- 3.1 Form und Inhalt: Hier wird der MIME-Type auf Dokument-Ebene nach Vorgabe des KOST Formatkatalogs zu Dokumentklassen auf der nächsten Ebene zusammengefasst. Eine weitere Aggregation ist mit Hilfe eines Schwellwerts denkbar (z.B. sind 90% aller Dokumente in den Dossiers einer Serie Textdokumente, so ist die Serie auch vom Typ "Text").
- 4.1 Zugangsbestimmungen: Innerhalb einer Bestelleinheit (in der Regel ein Dossier) aggregiert die am stärksten einschränkende Bestimmung pro Kategorie, darüber aber, in Serie oder Bestand, aggregiert umgekehrt die schwächste Bestimmung. z.B. erhält ein Dossier mit nur einem Dokument mit personenbezogenen Informationen das Attribut hasPrivacyProtection, die Serie mit diesem Dossier erhält das Attribut aber nicht, wenn nur ein Dossier keine PrivacyProtection hat. (siehe 4.1 Zugangsbestimmungen, S.8).
- 4.1 Zugangsbestimmungen: Hier ist die Aggregation die jeweils am geringsten einschränkende Bestimmung pro Kategorie der darunter liegenden Verzeichnungseinheiten (siehe 4.1 Zugangsbestimmungen, S.8).
- 4.4 Physische Beschaffenheit und technische Anforderungen: Verzeichnungseinheiten von unterschiedlichem Typ aggregieren zum Typ "hybrid".

# Beispiel für Aggregation: 4.1 Zugangsbestimmungen "public" für das Dossier



# (1.3) Entstehungszeitraum / Laufzeit: dates

Zeitraum "von – bis" und ein einzelnes Datum werden im Schema folgendermassen dargestellt. //isad:dates/isad:fromDate, //isad:dates/isad:toDate, beziehungsweise //isad:dates/isad:pointofTime. Jede Zeitangabe kann mit dem Attribut "circa" versehen werden: isad:dates@circa

Die Zeitangaben können folgende Ausprägung annehmen:

```
CC (Century)
YYYY
YYYY-MM
YYYY-MM-DD
YYYY-MM-DDThh:mm
YYYY-MM-DDThh:mm:ss
YYYY-MM-DDThh:mm:ss.mis (Millisekunde)
```

Sind für eine Verzeichnungsstufe mehrere Zeitangaben verfügbar, z.B. Zeitraum und technisches "Creation Date", kann dates mit Angabe von origin wiederholt werden. Jahrangaben als from Year und to Year werden nicht mehr unterstützt.

#### (1.5) Umfang: extentMedium

Wird im Schema in zwei Elemente aufgeteilt:

(a) Physischer Umfang: //isad:extentMedium/isad:extent

Der Umfang wiederum kann Menge //isad:extent/Medium/isad:extent/isad:dataSize mit einem Attribut isad:dataSize@unit, das die Mengeneinheit festlegt,



oder Anzahl Objekte bedeuten //isad:extentMedium/isad:extent/isad:dataQuantity



(b) Archivalienart: //isad:extentMedium/isad:medium als Freitext.



Liegen Angaben zum Umfang nur als Freitext vor, z.B. wie in eCH-0160 das Element Umfang definiert ist, kann //isad:extentMedium/isad:description dafür verwendet werden.



#### (3.1) Form und Inhalt: scopeContent

Wird im Schema in zwei Elemente aufgeteilt:

//isad:scopeContent/isad:scope: Angabe über die hauptsächliche Form der Verzeichnungseinheit mit einer Werteliste: Text, Bild, Audio, Video, strukturierte Daten oder Freitext.

//isad:scopeContent/isad:content: Der "Enthält- oder Darin-Vermerk" wird verwendet, um auf einzelne Teile der Verzeichnungseinheit hinzuweisen, die in formaler Hinsicht als Besonderheit anzusehen sind, z. B. Karten, Pläne, Zeichnungen, Fotos, herausragende Einzelstücke und die nicht aufgrund des Titels zu erwarten sind. Mit einem optionalen Kriterium //isad:scopeContent/isad:content@contentQualifier kann zusätzlich die Art des "Enthält- oder Darin-Vermerke" spezifiziert werden, mit folgender Werteliste: Enthält, Enthält\_unter\_anderem, Enthält\_vor\_allem, Enthält\_auch, Enthält nur (Details dazu auch direkt im Schema)



Die Übersetzung im Gegensatz zum Übersetzungsvorschlag der Archivschule Marburg von Angelika Menne-Haritz, *Schlüsselbegriffe der Archivterminologie*, 1999: <a href="https://internet.archivschule.uni-marburg.de/datiii/germanterms.htm">https://internet.archivschule.uni-marburg.de/datiii/germanterms.htm</a>

ist folgende Übersetzungstabelle hinterlegt:

Darin includes Enthält contains

Enthält\_unter\_anderem contains\_inter\_alia
Enthält\_vor\_allem contains\_above\_all
Enthält\_auch contains\_also
Enthält\_nur contains\_exclusivley

#### (3.4) Ordnung und Klassifikation: arrangement

//isad:arrangement besitzt ein Attribut "xpointer". Das Attribut erlaubt direkte Verweise auf Objekte in der PREMIS Datei. Damit können z.B. Metainformationen zu original Ordner und Dateinamen zu einem nicht weiter erschlossenen Digitalen Objekt festgehalten werden, siehe dazu der englische Text ISAD(G) 3.3.4 ...For electronic records, record or reference information on system design. Dieser Fall tritt ein, wenn SIPs bzw. AIPs nur bis Dossier Ebene erschlossen sind (Bsp. eCH-0160 FILES SIP) aber dennoch die Ordner und Dateinamen aus der original Ablage bekannt sind.

```
<isad:arrangement xpointer="premis.xml#xpointer(//xmlID('object-id_22')" origin="ingest">
//Meier/Briefablage/Übersetzung_Einladung_ws1.pdf</isad:arrangement>
<isad:arrangement xpointer="premis.xml#xpointer(//xmlID('object-id_23')" origin="ingest">
//Meier/Briefablage/Übersetzung Protokoll ws1.pdf</isad:arrangement>
```

In diesem Beispiel wird auf der Ebene Dossier für zwei Dateien die Metainformation originaler Dateipfad in der Ablage vor der Ablieferung angegeben, «origin» der Information ist in diesem Fall «ingest».

Es ist auch denkbar, dass ein Dokument mit mehreren Dateien einer Ablage verbunden ist (=verschiedene Repräsentationen des gleichen Dokuments) ohen dass eine weiter *isad:archivalDescription* Ebene vom Typ «representation» existiert, dann kann der selbe Mechanismus verwendet werden.

#### (4.1) Zugangsbestimmungen: accessConditions

Wie im *Data Dictionary* weiter oben schon ausgeführt, besteht Zugangsbestimmungen aus acht Unterelementen, die jeweils einen spezifischen Aspekt abbilden können:

- //isad:accessConditions/isad:hasPrivacyProtection
   Datenschutz: Angabe, ob in der Verzeichnungseinheit besonders schützenswerte
   Personendaten vorhanden sind oder nicht. Wertebereich: true / false
- //isad:accessConditions/isad:openToThePublic
   Öffentlichkeitsstatus: Angabe über die Zugänglichkeit der Verzeichnungseinheit gemäss den geltenden Informationsfreiheits- resp. Öffentlichkeitsgesetzen. Werteliste undefined, public, not\_public
- //isad:accessConditions/isad:classification
   Geheimhaltung: Angabe zur Geheimhaltung/Klassifikation der Verzeichnungseinheit. Werteliste: unclassified, confidential, secret, in house, other
- //isad:accessConditions/isad:otherAccessConditions
   Weitere Zugangs- und Nutzungsbedingungen: Angabe über weitere Zugangs- und Nutzungsbedingungen; NICHT Repro-Bestimmungen
- //isad:accessConditions/isad:accessConditionsNotes
   Bemerkungen zu Zugangs- und Nutzungsbedingungen.

- //isad:accessConditions/isad:retentionPeriod
   Schutzfrist: Dauer der Schutzfrist in Jahren.
- //isad:accessConditions/isad:retentionPeriodConditions
   Schutzfristenkategorie bzw. Artikel der gesetzlichen Grundlage für die Schutzfrist festhält.
- //isad:accessConditions/isad:retentionPeriodNotes
   Weiter Informationen zur Schutzfrist, z.B. Aktenzeichen des Formulars "Meldung von Unterlagen mit besonderer Schutzfrist und öffentlich zugänglichen Unterlagen", in dem die für die Ablieferung vereinbarten Schutzfristen festgehalten sind.



#### Erweiterte Referenzen: primaryDataLocator und secondaryDataLocator

Bei //isad:primaryDataLocator soll das Attribut "uri" direkt auf die betreffende Datei verweisen. Bei //isad:secondaryDataLocator soll das Attribut "xpointer" indirekt, d.i. via PREMIS auf eine Datei verweisen.

<isad:secondaryDataLocator xpointer="premis.xml#xpointer(//xmlID('object-id\_1')"/>

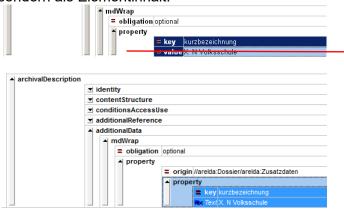
Ein Verweis via Elementinhalt ist aus Gründen der Rückwärtskompatibilität weiterhin erlaubt aber nicht empfohlen bzw. «deprecated»:

<isad:secondaryDataLocator>premis.xml#xpointer(//xmlID('object-id 1')</isad:secondaryDataLocator>

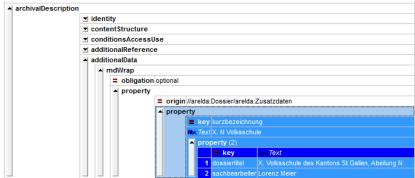
#### Erweiterte Metadaten: additionalData

Diese Schemaerweiterung erlaubt, zusätzlich zu den archivischen (ISAD[G]-konformen) Metadaten weitere Daten aus dem Quellsystem in ein Findmittel zu übernehmen. Das kann vor allem von Interesse sein, wenn das Quellsystem auf anderen, nicht archivischen Standards aufbaut, wie das etwa bei Bilddatenbanken oder Bibliotheksystemen der Fall ist.

Die Erweiterung *mdWrap* erlaubt, strukturierte Metadaten in Form von Schlüssel-Werte-Paaren zu übernehmen. Wert ist neu nicht mehr als Attribut *value* definiert, sondern als Elementinhalt.



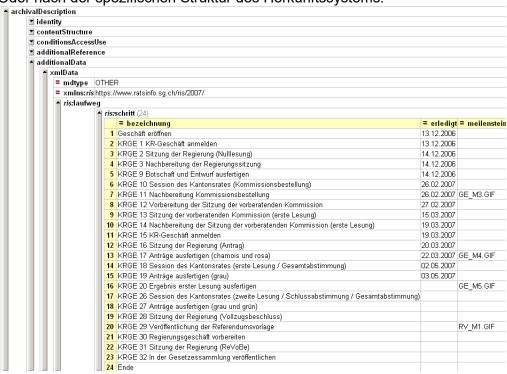
Durch rekursive Schachtelung von *property* können auch mehrdimensionale Tabellen abgebildet werden:



Die Erweiterung **xmlData** erlaubt hingegen, strukturierte Daten nach einem Fremdschema zu übernehmen, hier im Beispiel *Dublin Core*:



Oder nach der spezifischen Struktur des Herkunftssystems:



### Das XML Schema im Detail

