



Inhalt

1	Vorwort	2
2	Systemvoraussetzungen	3
3	Offene Punkte / Feedback	3
4	Installation.....	4
4.1	Installation von SIARD-Val.....	4
4.2	Installation von Java Version 6	4
5	Konfiguration des SIARD-Val	5
5.1	Bestandteile der Konfigurationsdatei "SIARDVal.conf.xml"	5
6	Validierung starten.....	6
6.1	Validierung starten mittels GUI	6
7	Validierung auswerten	8
8	Beschreibung der Validierungsschritte	9
8.1	Die Validierungsschritte welche zu einem Abbruch führen	9
8.2	Die Validierungsschritte welche zu einem Fehler führen	10
9	Urheberrecht.....	11

1 Vorwort

SIARD-Val ist eine java-basierte Anwendung zur Validierung von Aufbau und Inhalt einer SIARD-Datei (Software Independent Archiving of Relational Databases) zur Archivierung von relationalen Datenbanken. Diese Anwendung steht unter der GPL3+ Lizenz und wird durch die KOST der Öffentlichkeit quelloffen zur Verfügung gestellt. SIARD-Val erfüllt die im Folgenden beschriebenen Anforderungen.

Funktionale Anforderungen:

SIARD-Val liest die SIARD-Datei und validiert die folgenden Punkte aus der SIARD Formatbeschreibung des BAR von September 2008 in der Version 1.0¹ sowie aus der SIARD Spezifikation, eCH-0165 SIARD Formatspezifikation, von 2013 in der Version 1.0²:

ID	Modulname	Wenn Resultat n.i.O.
A	Lesbarkeit	Abbruch
B	primäre Verzeichnisstruktur	Abbruch
C	Header-Validierung	Abbruch
D	Struktur-Validierung	Abbruch
E	Spalten-Validierung	Fehler
F	Zeilen-Validierung	Fehler
G	Tabellen-Validierung	Fehler
H	Content-Validierung	Fehler
I	SIARD-Erkennung	Fehler
J	Zusätzliche Primärdateien	Fehler

Die Resultate (inklusive Meldungen zu Inkonsistenzen oder Fehler) werden pro Schritt ausgegeben und in eine Validierungs-Logdatei geschrieben.

Das Ergebnis der Gesamtvalidierung (korrekte/fehlerhafte SIARD-Datei) wird ebenfalls ausgegeben und im *exit*-Status des Programms sichtbar, so dass die Validierung in eine automatisierte Verarbeitungskette eingebunden werden kann. Der *exit*-Status kann die folgenden Werte annehmen:

- 0 alles OK
- 1 fehlerhafte Angaben
- 2 Validierung nicht bestanden

Die einzelnen Validierungsschritte / Prüfungen werden nacheinander ausgeführt. Wo möglich, wird die Validierung auch bei Fehlern weiter fortgesetzt, um die Anzahl von Korrekturzyklen zu reduzieren. Nur die Prüfungen A) bis D) sind Fehler, die direkt zum Programmabbruch führen.

¹ Die Formatbeschreibung kann von der Seite "Tools und Hilfsmittel zur digitalen Archivierung / Datenbankarchivierung: SIARD Suite" der BAR-Website heruntergeladen werden:
<http://www.bar.admin.ch/dienstleistungen/00823/00825/>.

² Die Spezifikation kann von der eCH Website heruntergeladen werden:
<http://www.ech.ch/vechweb/page?p=dossier&documentNumber=eCH-0165>.

Nichtfunktionale Anforderungen:

Für besondere Aufgaben (ZIP-Entpacken, XML-Validierung) werden externe Programme oder entsprechende Java-Frameworks eingesetzt.

Die Anwendung ist modular aufgebaut, damit ohne viel Aufwand ein oder mehrere weitere Validierungsmodule eingebaut werden können.

Zur Schemavalidierung wird das in der SIARD-Datei hinterlegte Schema verwendet (dynamische Schemavalidierung).

Die Log-/Programmausgabe erlaubt die Verwendung des Tools in einer Prozesskette, d.h. das Ergebnis der einzelnen Validierung kann einfach ermittelt werden.

Die Konsolenausgabe begrenzt sich auf die Bezeichnung der Validierungsschritte (inklusive zugehöriges Testmodul), die Ergebnisse "valid" oder "invalid", das Gesamtergebnis sowie die Pfade zur SIARD- und zur generierten Log-Datei. Alle zusätzlichen Informationen werden in der Log-Datei aufgeführt.

2 Systemvoraussetzungen

Microsoft Windows 98, ME, NT, 2000, XP, VISTA oder Windows 7

Mindestens 128 MB RAM (512 MB oder mehr empfohlen)

Mindestens 20 GB Festplattenspeicher

Java Runtime Environment (JRE) Version 6 oder neuer³ (Download: <http://www.java.com>)

3 Offene Punkte / Feedback

Die offenen Punkte, von Bugs respektive Fehler bis hin zu Ergänzungswünschen und Fragen sind auf der Entwicklungsplattform GitHub unter Issues ersichtlich:

<https://github.com/KOST-CECO/SIARD-Val/issues>

Diese Liste kann und soll durch jedermann erweitert werden und wird durch das Entwicklerteam bearbeitet.

³ Die aktuell installierte Version lässt sich überprüfen mit der Eingabe von `java -version` im cmd-Fenster.

4 Installation

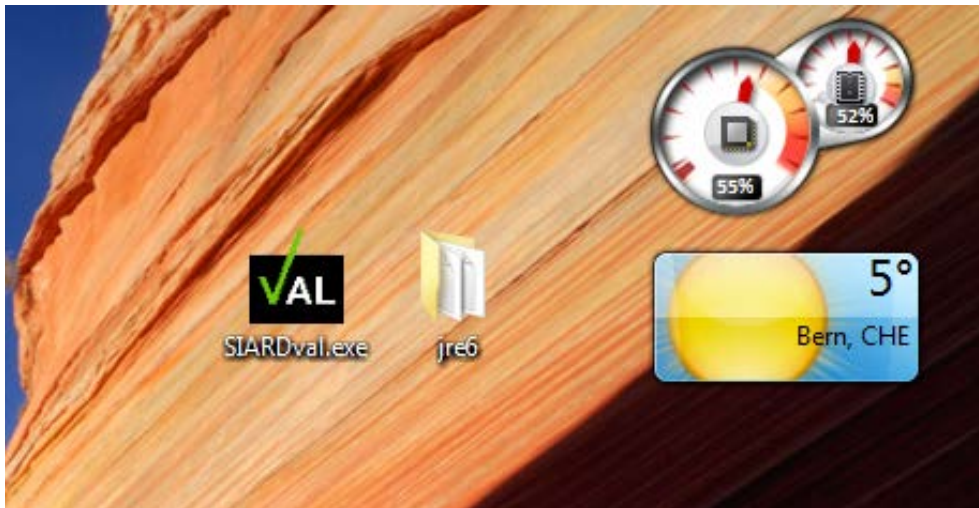
4.1 Installation von SIARD-Val

- 1 SIARDval.exe herunterladen und an den gewünschten Ort kopieren zum Beispiel auf den Desktop.



4.2 Installation von Java Version 6

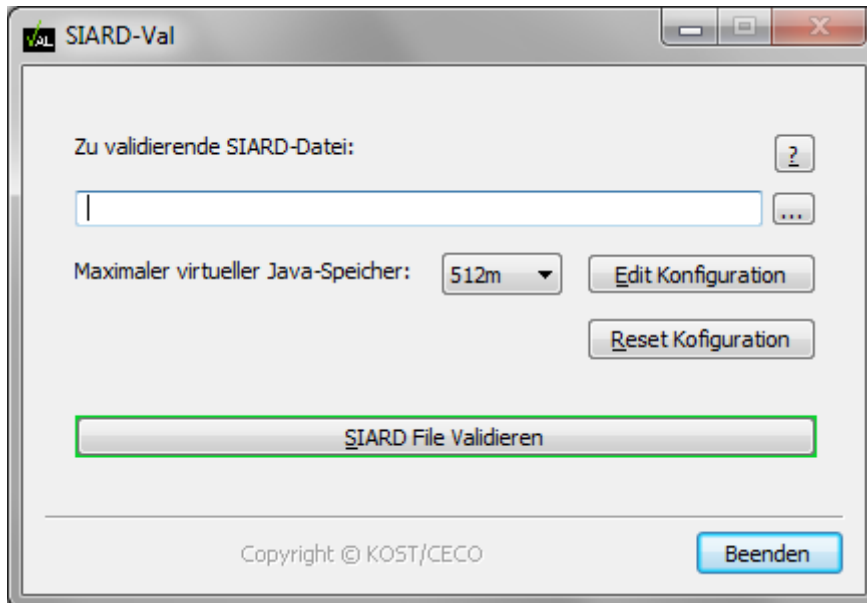
- 1 Sollte Java Runtime Environment (JRE) nicht bereits in Version 6 oder neuer⁴ installiert sein, muss jre6 unter <http://www.java.com/> heruntergeladen und am gleichen Ort wie SIARDval.exe abgelegt werden.



⁴ Die aktuell installierte Version lässt sich überprüfen mit der Eingabe von `java -version` im cmd-Fenster.

5 Konfiguration des SIARD-Val

Sämtliche Konfigurationen des SIARD-Val können im GUI mit der Schaltfläche "Edit Konfiguration" vorgenommen werden.



Mit der Schaltfläche "Reset Konfiguration" werden die Initialwerte gesetzt.

5.1 Bestandteile der Konfigurationsdatei "SIARDVal.conf.xml"

xml-Tag	Beschreibung
<pathtoworkdir>	Pfadangabe zum temporären Arbeitsverzeichnis von SIARD-Val
<table-rows-limit>	Grenze der zu validierenden XML-Tabelle im Modul H

5.1.1 <pathtoworkdir>

Pfadangabe zum temporären Arbeitsverzeichnis von SIARD-Val. Der Initialwert ist das Standard Temporärverzeichnis, welches durch das GUI ermittelt wird. Dabei kann der absolute Pfad angegeben werden oder der relationale Pfad (aus Sicht von SIARDval.exe). Dieser Ordner muss nicht angelegt werden.

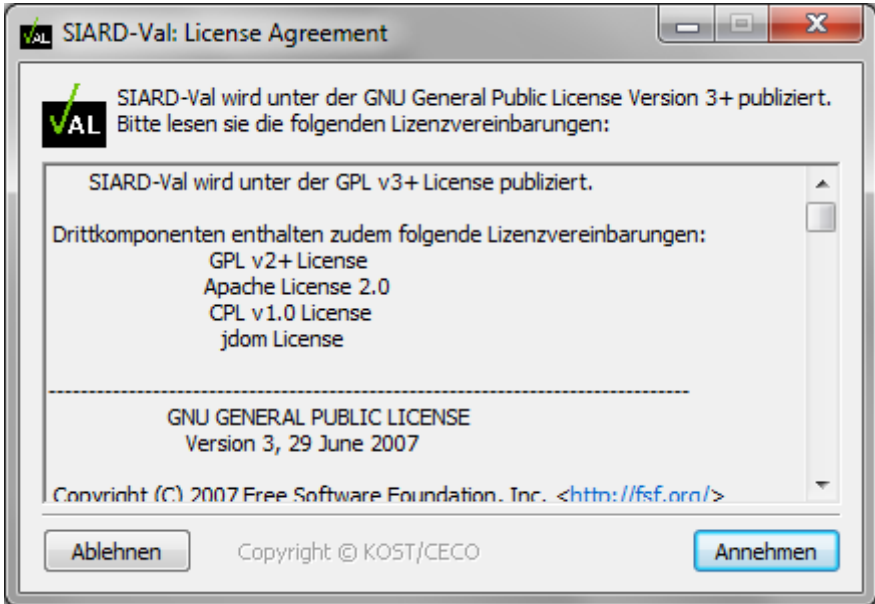
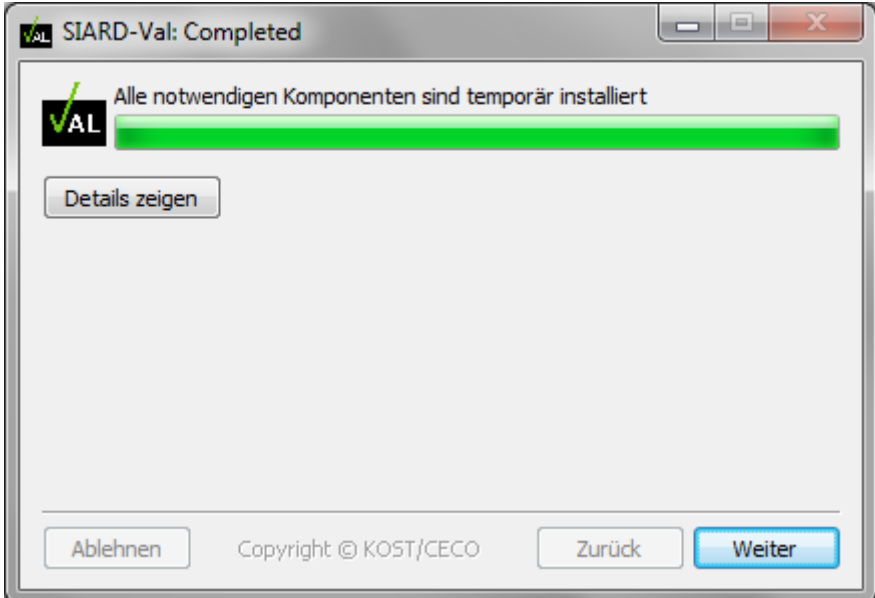
5.1.2 <table-rows-limit>

Grenze der zu validierenden XML-Tabelle im Modul H. Der Initialwert ist **20000**⁵. Diese Grenze wird nur verwendet, sollte die dazugehörige XSD-Datei die genaue Anzahl Datenzeilen der XML-Tabelle enthalten. Sind mehr Datenzeilen in der Tabelle enthalten als in der Konfigurationsdatei eingegrenzt, wird diese einzelnen Tabelle nicht validiert, damit einen entsprechenden Out-of-Memory-Fehler verhindert werden kann und die restlichen Tabellen und Module validiert werden können.

⁵ Dieser Initialwert ist ein erster Erfahrungswert der KOST-Geschäftsstelle. Er kann von System zu System unterschiedlich sein.

6 Validierung starten

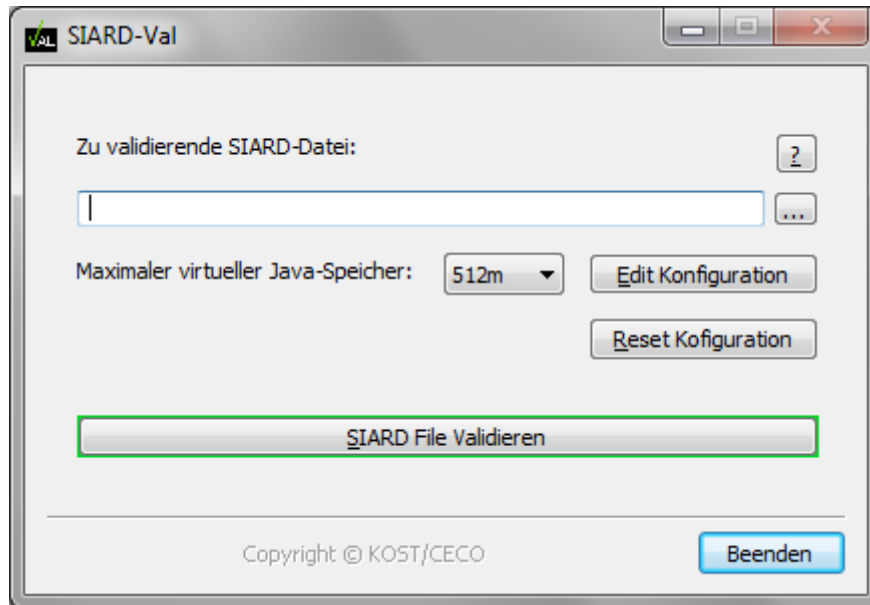
6.1 Validierung starten mittels GUI

1	Die SIARDval.exe-Dateien mittels Doppelklick ausführen.
2	<p>Damit eine Validierung stattfinden kann, müssen sämtliche Lizenzvereinbarungen angenommen werden.</p> 
3	<p>Alle notwendigen Komponenten werden temporär ins Standard Arbeitsverzeichnis abgelegt.</p> 

4

Die zu validierende SIARD-Datei wird mittels der Schaltfläche "..." ausgewählt. Dabei wird einerseits bereits geprüft, ob die Datei eine ZIP-Datei mit der Dateierweiterung ".siard" ist.

Mittels der Schaltfläche "SIARD File Validieren" wird SIARD-Val.jar ausgeführt. Dabei wird automatisch während der laufenden Validierung eine Konsolenfenster geöffnet und nach Fertigstellung der Validierung auch wieder geschlossen.



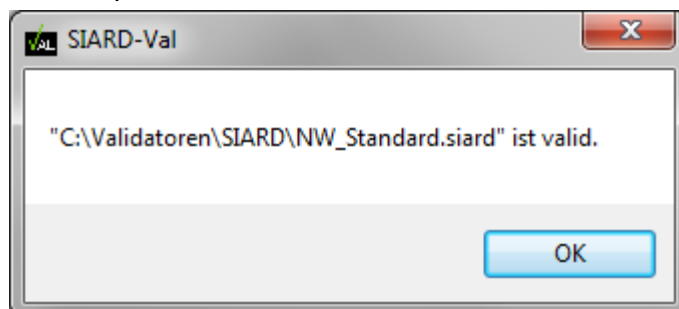
Wenn gewünscht kann die Konfigurationsdatei wie im Kapitel 0 beschrieben angepasst werden.

Bei SIARD-Dateien mit grossen XML-Tabellen kann es nötig sein, den maximalen virtuellen Java-Speicher auf 1024m zu setzen, damit der benötigte Speicherplatz für die Validierung ausreicht.

Das Handbuch wird mittels der Schaltfläche "?" geöffnet.

5

Sobald die Validierung beendet ist erscheint bei validen SIARD-Dateien die Meldung, dass die Datei valid ist. Mit OK gelangt man wieder in das eigentliche Hauptfenster.



Sollte jedoch die SIARD-Datei nicht valid sein, wird die entsprechende Logdatei mit der Fehlermeldung geöffnet. Diese kann bei Bedarf mit "speichern unter..." gesichert werden.

SIARD-Val v1.0.0, Copyright (C) 2012-2013 Claire Roethlisberger (KOST-CECO), Christian Eugster, Olivier Debenath, Peter Schneider (Staatsarchiv Aargau)
 This program comes with ABSOLUTELY NO WARRANTY.
 This is free software, and you are welcome to redistribute it under certain conditions; see GPL-3.0_COPYING.txt for details.

SIARD Validierung -> Accounting_A_32c.siard

[A] = Invalid (Test "Lesbarkeit" nicht bestanden)
 [A] -- Die SIARD-Datei ist komprimiert.

TOTAL = Invalid

SIARD = C:\Validatoren\SIARD\Accounting_A_32c.siard
 Log = C:\Users\U80809-1\AppData\Local\Temp\Accounting_A_32c.siard.validationlog.log

Validierung abgebrochen.

N length: 657 lines: 21 Ln: 21 Col: 1 Sel: 0 UNIX ANSI INS

Sobald man den Editor schliesst, gelangt man wieder in das eigentliche Hauptfenster.

6 Bei Bedarf können weitere SIARD-Dateien validiert werden. SIARD-Val GUI wird mittels der Schaltfläche "Beenden" geschlossen und alle temporären Dateien werden gelöscht.

7 Validierung auswerten

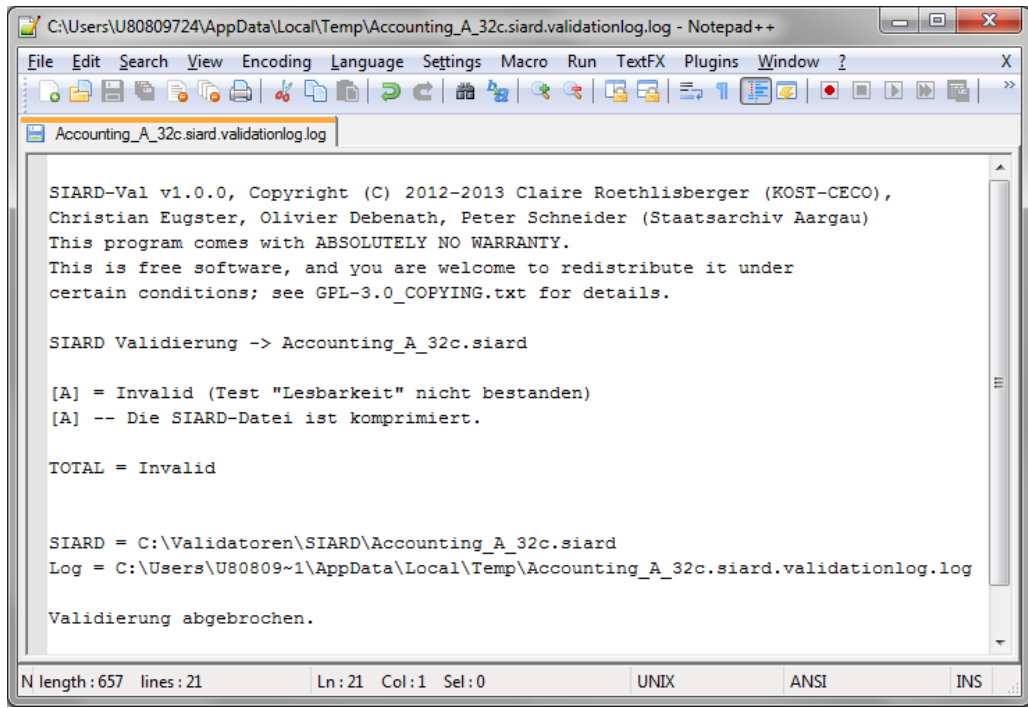
1 Zusammenstellung der Validierungsschritte:

ID	Modulname	Wenn Resultat n.i.O.
A	Lesbarkeit	Abbruch
B	primäre Verzeichnisstruktur	Abbruch
C	Header-Validierung	Abbruch
D	Struktur-Validierung	Abbruch
E	Spalten-Validierung	Fehler
F	Zeilen-Validierung	Fehler
G	Tabellen-Validierung	Fehler
H	Content-Validierung	Fehler
I	SIARD-Erkennung	Fehler
J	Zusätzliche Primärdateien	Fehler

2 Nach dem Start von SIARD-Val werden in der Logdatei die Ergebnisse zu den einzelnen Validierungsschritten gespeichert. Wenn der Validierungsschritt ohne Fehler abgeschlossen wurde, wird "Valid" ausgegeben, z.B. [H] = Valid.
 Sollte ein Validierungsschritt fehlerhaft sein, wird "Invalid" und die Bezeichnung des Validierungsschrittes ausgegeben, z.B. [A] = Invalid (Test "Lesbarkeit" nicht bestanden). Die nächsten Zeilen enthalten die

entsprechende Fehlermeldung z.B. [A] -- Die SIARD-Datei ist komprimiert.

Am Schluss werden das Gesamtergebnis sowie die Pfade zu den einzelnen Dateien angezeigt.



```
SIARD-Val v1.0.0, Copyright (C) 2012-2013 Claire Roethlisberger (KOST-CECO),
Christian Eugster, Olivier Debenath, Peter Schneider (Staatsarchiv Aargau)
This program comes with ABSOLUTELY NO WARRANTY.
This is free software, and you are welcome to redistribute it under
certain conditions; see GPL-3.0_COPYING.txt for details.

SIARD Validierung -> Accounting_A_32c.siard

[A] = Invalid (Test "Lesbarkeit" nicht bestanden)
[A] -- Die SIARD-Datei ist komprimiert.

TOTAL = Invalid

SIARD = C:\Validatoren\SIARD\Accounting_A_32c.siard
Log = C:\Users\U80809~1\AppData\Local\Temp\Accounting_A_32c.siard.validationlog.log

Validierung abgebrochen.
```

8 Beschreibung der Validierungsschritte

Nachfolgend werden die einzelnen Validierungsschritte detaillierter beschrieben; insbesondere welche Kriterien genau überprüft werden und wie diese auszusehen haben. Diese Kriterien sind einen Auszug der Muss-Kriterien aus der SIARD-Spezifikation gemäss Kapitel 1.

8.1 Die Validierungsschritte welche zu einem Abbruch führen

A Lesbarkeit

Kann die SIARD-Datei gelesen werden?

- lesbare und nicht passwortgeschützte ZIP-Datei oder ZIP64-Datei
- unkomprimierte ZIP-Datei oder ZIP64-Datei

B primäre Verzeichnisstruktur

Besteht eine korrekte primäre Verzeichnisstruktur?

- [Name].siard/header und [Name].siard/content
- [Name].siard/[Name]/header und [Name].siard/[Name]/content
- Andere Ordner oder Dateien sind im Toplevel-Ordner verboten

C Header-Validierung

Ist der header-Ordner valid?

- metadata.xml valid zu original metadata.xsd
(zusätzliche Ordner oder Dateien wie z.B. metadata.xls sind im header-Ordner erlaubt)

D Struktur-Validierung

Stimmt die Struktur aus metadata.xml (folder) mit der Datei-Struktur von content überein?

gleiche Struktur (z.B. schema0/table3 → schema0/table3/tabe3.xsd und table3.xml)

8.2 Die Validierungsschritte welche zu einem Fehler führen

E Spalten-Validierung

Wurden die Angaben aus metadata.xml korrekt in die tableZ.xsd-Dateien übertragen?

gleiche Spaltendefinitionen (Anzahl, Type, Nullable)

F Zeilen-Validierung

Wurden die Angaben aus metadata.xml korrekt in die tableZ.xsd-Dateien übertragen?

gleiche Zeilenzahl (rows in metadata.xml = max = minOccurs in tableZ.xsd)⁶

Wenn 0 bis ∞: Enthält tableZ.xml die gleiche Anzahl Zeilen wie in metadata.xml definiert?

gleiche Zeilenzahl (rows in metadata.xml = Anzahl row in tableZ.xml)

G Tabellen-Validierung

Sind die Spaltennamen innerhalb der Tabelle einmalig?

einmalige Spaltennamen innerhalb der Tabelle

einmalige Tabellennamen innerhalb der Schema(s)

einmalige Schemanamen

H Content-Validierung

Sind die XML-Dateien im content valid zu ihrer Schema-Definition (XSD-Dateien)?

tableZ.xml valid zu tableZ.xsd

I SIARD-Erkennung

Wird die SIARD-Datei als SIARD erkannt?

Extension = .siard

J Zusätzliche Primärdateien

Enthält der content-Ordner Dateien oder Ordner die nicht in metadata.xml beschrieben sind⁷?

Andere Ordner oder Dateien sind im content-Ordner verboten

⁷ Mit Ausnahme der recordZ.txt und recordZ.bin Dateien in den lobZ-Ordnern.

⁷ Mit Ausnahme der recordZ.txt und recordZ.bin Dateien in den lobZ-Ordnern.

9 Urheberrecht

SIARD-Val ist eine Entwicklung der KOST. Alle Rechte liegen bei der KOST. SIARD-Val wurde im 2012 durch die KOST unter der GNU General Public License v3+ veröffentlicht.

Notice:	This product includes software developed by the Apache Software Foundation (http://www.apache.org/).
----------------	---

SIARD-Val stützt sich auf folgende unveränderte Komponenten anderer Hersteller, welche direkt im Quellcode von SIARD-Val eingebunden sind:

Drittprogramm / -Komponente	Version	Lizenz
Apache Commons http://commons.apache.org/ - commons-collections-3.2.1.jar - commons-configuration-1.6.jar - commons-digester-1.8.jar - commons-lang-2.4.jar - commons-logging-1.1.1.jar	3.2.1 1.6 1.8 2.4 1.1.1	Apache License 2.0
Apache log4j http://logging.apache.org/log4j/	1.2.12	Apache License 2.0
Apache Xalan-Java http://xml.apache.org/xalan-j/	2.7.0	Apache License 2.0
Apache Xerces http://xerces.apache.org/	2.7.1	Apache License 2.0
Jdom 2.0.0 http://www.jdom.org/	2.0.0	jdom License
Junit 4.4 http://www.junit.org/	4.4	CPL v1.0
Spring Framework API http://static.springsource.org/spring/docs/3.0.x/api/	3.0.0	Apache License 2.0
zip64 http://sourceforge.net/projects/zip64file/	1.02	GPL v2+ Lizenz

Die Benutzer von SIARD-Val sind gehalten, die Lizenzbestimmungen all dieser Komponenten zu befolgen, welche im SIARD-Val-Verzeichnis license vorliegen.