



Anwendungshandbuch

Inhalt

| | | |
|-----|--|----|
| 1 | Vorwort | 2 |
| 2 | Systemvoraussetzungen | 3 |
| 3 | Offene Punkte / Feedback | 3 |
| 4 | Installation..... | 4 |
| 4.1 | Installation von SIARD-Val..... | 4 |
| 4.3 | Installation von Java (Version 6 oder neuer) | 5 |
| 5 | Konfiguration des SIARD-Val | 6 |
| 5.1 | Bestandteile der Konfigurationsdatei "SIARDVal.conf.xml" | 6 |
| 6 | cmd_Script-Example..... | 7 |
| 6.1 | "_SIARD-Val_Start.cmd" | 7 |
| 6.2 | "manuelle_Eingabe.cmd" | 7 |
| 7 | Validierung starten..... | 8 |
| 7.1 | Validierung starten mittels Script | 8 |
| 7.2 | Validierung manuell starten | 9 |
| 8 | Validierung auswerten | 11 |
| 9 | Beschreibung der Validierungsschritte | 12 |
| 9.1 | Die Validierungsschritte welche zu einem Abbruch führen | 12 |
| 9.2 | Die Validierungsschritte welche zu einem Fehler führen | 13 |
| 10 | Urheberrecht..... | 14 |

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>

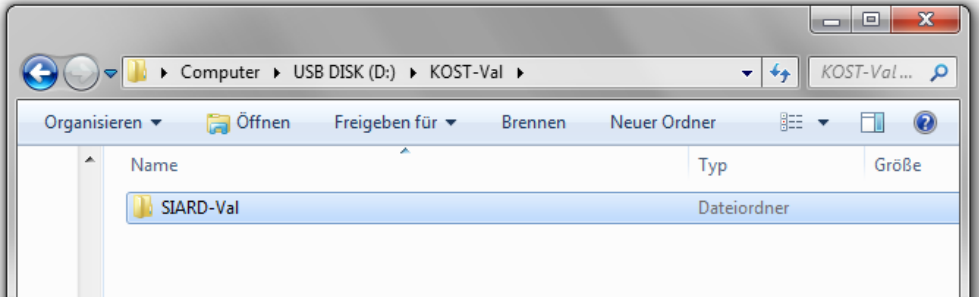
Diese Liste kann und soll durch jedermann erweitert werden und wird durch das Entwicklerteam bearbeitet.

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

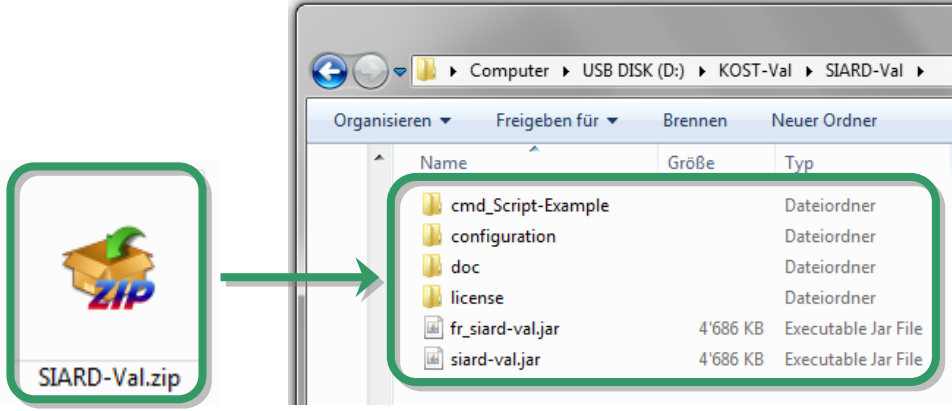
1 SIARD-Val-Verzeichnis erstellen.



The screenshot shows a Windows Explorer window with the address bar set to 'Computer > USB DISK (D:) > KOST-Val'. The main pane displays a table with columns 'Name', 'Typ', and 'Größe'. A single entry 'SIARD-Val' is listed with the type 'Dateiordner'.

| Name | Typ | Größe |
|-----------|-------------|-------|
| SIARD-Val | Dateiordner | |

2 SIARD-Val herunterladen und in das SIARD-Val-Verzeichnis entpacken.

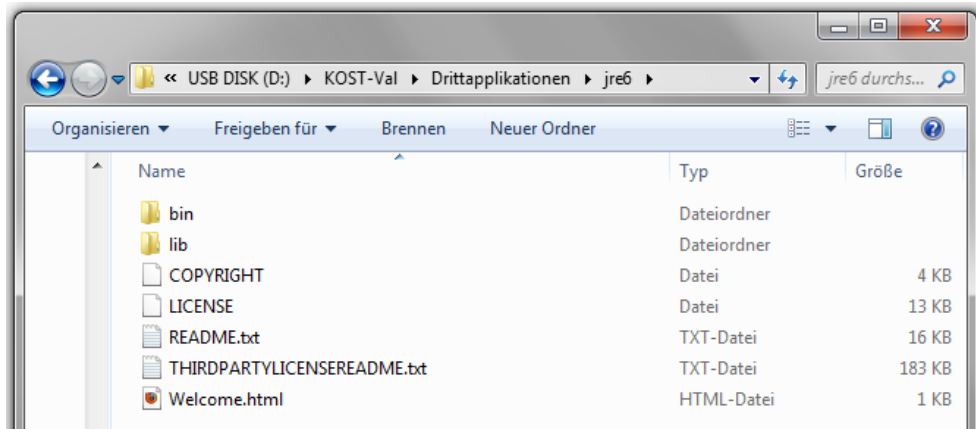


The screenshot shows a Windows Explorer window with the address bar set to 'Computer > USB DISK (D:) > KOST-Val > SIARD-Val'. To the left, a file icon for 'SIARD-Val.zip' is shown with a green arrow pointing to the folder. The main pane displays a table with columns 'Name', 'Größe', and 'Typ'. The contents of the folder are listed below.

| Name | Größe | Typ |
|--------------------|----------|---------------------|
| cmd_Script-Example | | Dateiordner |
| configuration | | Dateiordner |
| doc | | Dateiordner |
| license | | Dateiordner |
| fr_siard-val.jar | 4'686 KB | Executable Jar File |
| siard-val.jar | 4'686 KB | Executable Jar File |

4.3 Installation von Java (Version 6 oder neuer)

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



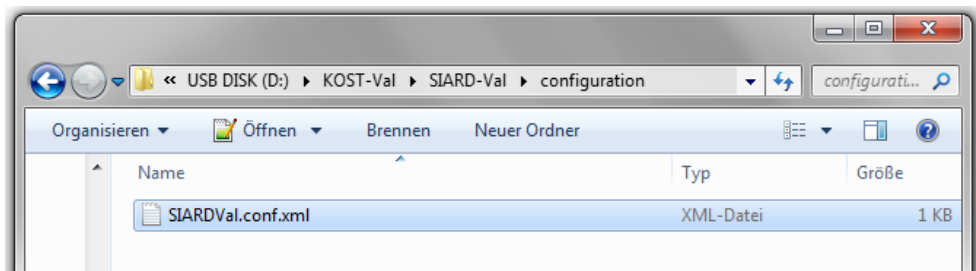
Hinweis:

Ist jre6 oder neuer bereits installiert, kann darauf verzichtet werden diese Dateien erneut abzulegen. Beachten Sie, dass Sie die cmd-Scripts wie in Kapitel 6 beschrieben anpassen müssen.

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 Unterordner "configuration" in der Konfigurationsdatei "SIARDVal.conf.xml" vorgenommen werden.



5.1 Bestandteile der Konfigurationsdatei "SIARDVal.conf.xml"

| xmI-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 **work_SIARD-Val**. Dabei kann der absolute Pfad angegeben werden oder der relative Pfad (aus Sicht von `_SIARD-Val_Start.cmd`). Dieser Ordner muss nicht angelegt werden.

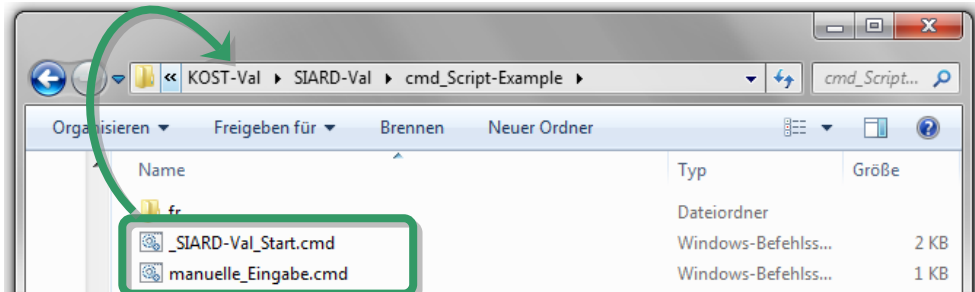
5.1.2 <table-rows-limit>

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 cmd_Script-Example

Im Unterordner "cmd_Script-Example" sind Beispiele von Command Scripts abgelegt, die zur Verwendung in den Ordner hereinkopiert werden können, in welchem sich der Ordner SIARD-Val befindet. In unserem Beispiel ist dies der Ordner "KOST-Val".



Die einzelnen Scripts werden nachfolgend kurz beschrieben.

6.1 "_SIARD-Val_Start.cmd"

Dieses Script validiert alle vorhandenen SIARD-Dateien, welche sich im Ordner "KOST-Val\SIARD" befinden. Die Log-Dateien werden in den Ordner "KOST-Val\logs" geschrieben. Dieses Script verwendet zudem java.exe aus dem Ordner "Drittapplikationen". Falls jre6 auf dem Computer bereits standardmässig installiert ist, kann "Drittapplikationen\jre6\bin\java.exe" durch "java" ersetzt werden.

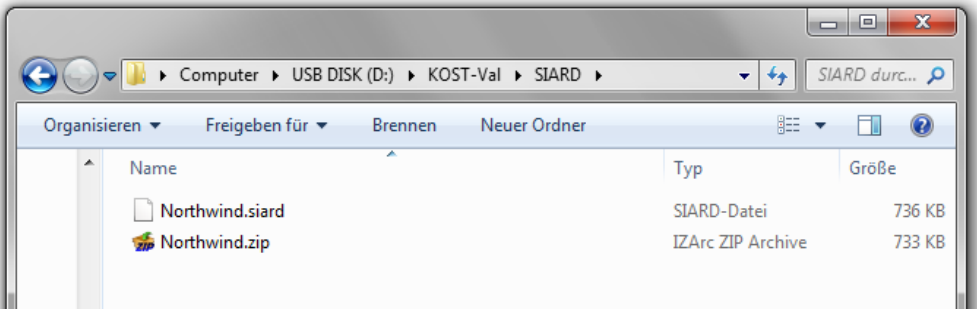
6.2 "manuelle_Eingabe.cmd"


Dieses Script setzt die Java-Settings für das aktuelle cmd-Fenster auf Drittapplikationen\jre6 und zeigt ein Beispiel der manuellen Eingabe an. Letzteres kann als Kopiervorlage dienen.

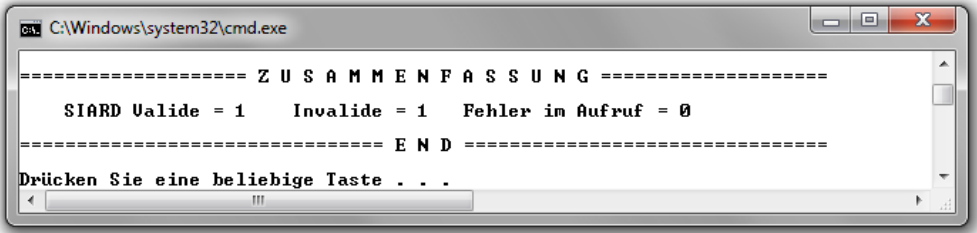
7 Validierung starten

7.1 Validierung starten mittels Script

- 1 Die zu validierenden SIARD-Dateien in den Ordner "KOST-Val\SIARD" kopieren.


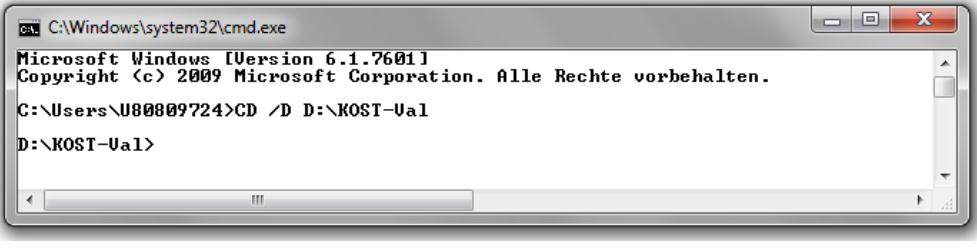
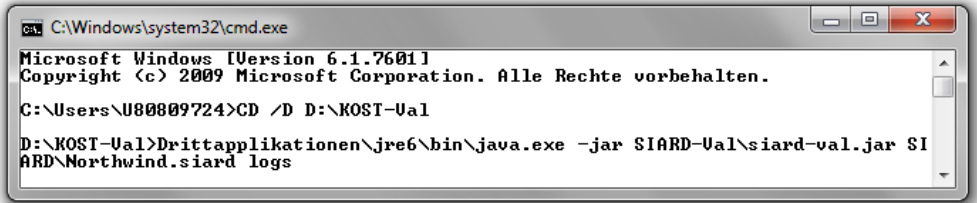

- 2 Starten des Scripts mittels Doppelklick auf "_SIARD-Val_Start.cmd" im Ordner "KOST-Val".


- 3 Alle SIARD-Dateien wurden validiert, sobald die Zusammenfassung im cmd-Fenster ersichtlich ist.



Die detaillierten Resultate sind im Ordner "logs" abgelegt.

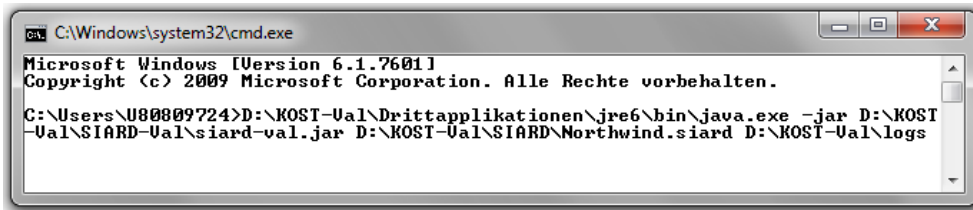
7.2 Validierung manuell starten

| | |
|---|---|
| 1 | <p>Die zu validierende SIARD-Datei bereitstellen.</p>  |
| 2 | <p>Eingabeaufforderung öffnen (Start → Ausführen ... → cmd [OK]) und in das gewünschte Arbeitsverzeichnis wechseln (CD D:\KOST-Val)⁶.</p>  |
| 3 | <p>SIARD-Val-Programmaufruf starten (die einzelnen Eingabebestandteile mit Leerzeichen trennen).</p>  <p>A Drittapplikationen\jre6\bin\java.exe -jar oder java -jar B SIARD-Val\siard-val.jar C SIARD\Northwind.siard D logs</p> <p>A = Java-Befehl (Drittapplikationen\jre6\bin\java.exe -jar oder java -jar⁷), B = relationaler Pfad und die Datei siard-val.jar, C = relationaler Pfad und vollständiger Name der zu validierenden Datei, D = relationaler Pfad und Name des Log-Verzeichnisses.</p> <p><u>Anmerkungen:</u> Wenn ein Eingabebestandteil Leerzeichen enthält, muss dieser in Anführungs- und Schlusszeichen eingegeben werden.</p> |

⁶ Das Laufwerk wird z.B. mit CD /D gewechselt.

⁷ Die Eingabe `java -jar` ist möglich, wenn Java Runtime Environment (JRE) Version 6 die Standardversion ist. Siehe auch Kapitel 3.2.1, Installation von Java (Version 6 oder neuer).

Der SIARD-Val kann auch von einem beliebigen Ort aus aufgerufen werden. Dies bedingt jedoch die Eingabe von absoluten Pfaden nicht nur im Programmaufruf selber, sondern auch in der Konfigurationsdatei "SIARD-Val.conf.xml".



```
C:\Windows\system32\cmd.exe
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Alle Rechte vorbehalten.
C:\Users\U80809724>D:\KOST-Ua1\Drittapplikationen\jre6\bin\java.exe -jar D:\KOST-Ua1\SIARD-Ua1\siard-val.jar D:\KOST-Ua1\SIARD\Northwind.siard D:\KOST-Ua1\logs
```

- A** D:\KOST-Ua1\Drittapplikationen\jre6\bin\java.exe -jar / java -jar
- B** D:\KOST-Ua1\SIARD-Val\siard-val.jar
- C** D:\KOST-Ua1\SIARD\Northwind.siard
- D** D:\KOST-Ua1\logs

Eine weitere Möglichkeit ist die manuelle Eingabe mit Hilfe des Scripts "manuelle_Eingabe.cmd". Mit diesem Skript kann SIARD-Val etwas einfacher aufgerufen werden:

```
java.exe -jar SIARD-Val\siard-val.jar SIARD\mein_SIARD logs
```



```
C:\Windows\system32\cmd.exe
Mit diesem Skript kann SIARD-Ua1 etwas einfacher aufgerufen werden:
java.exe -jar SIARD-Ua1\siard-val.jar SIARD\mein_SIARD logs
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Alle Rechte vorbehalten.
D:\KOST-Ua1>
```

Dabei kann die 2. Zeile als Kopiervorlage verwendet werden⁸.

- 4** Die SIARD-Datei wurde validiert, sobald "Validierung beendet" oder "Validierung abgebrochen" im cmd-Fenster erscheint.



```
C:\Windows\system32\cmd.exe
[GI] = Valid
[HI] = Valid
[IJ] = Valid
[IJ] = Valid
TOTAL = Valid

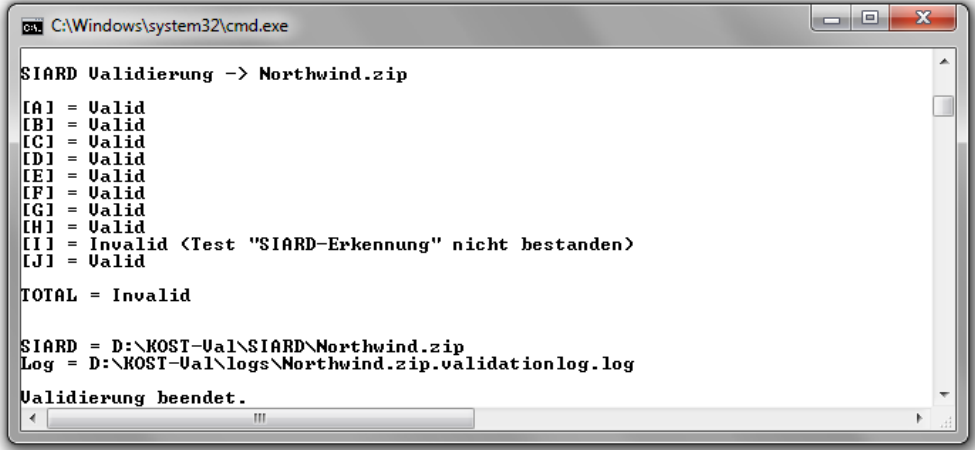
SIARD = D:\KOST-Ua1\SIARD\Northwind.siard
Log = D:\KOST-Ua1\logs\Northwind.siard.validationlog.log

Validierung beendet.
```

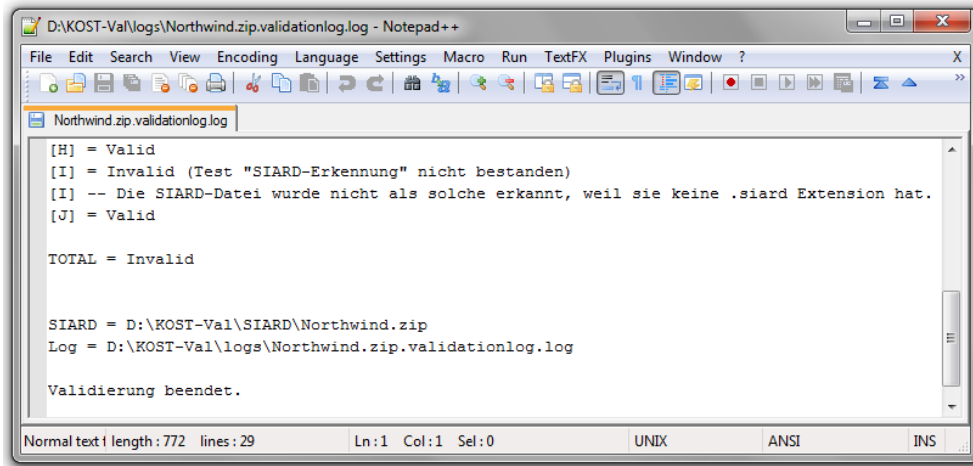
Die detaillierten Resultate sind im Ordner "logs" samt Unterordner abgelegt.

⁸ Zum Kopieren den Text der zweiten Zeile mit der Maus markieren, durch Klicken der rechten Maustaste in den Zwischenspeicher legen und wiederum mit der rechten Maustaste an der Eingabeaufforderung einfügen. Mit der Pfeiltaste zu "mein_SIARD" vorrücken, löschen und gewünschte SIARD-Datei eingeben oder herein kopieren, mit Eingabetaste starten.

8 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 | <p>Nach dem Start von SIARD-Val werden in der Konsolenausgabe die Ergebnisse zu den einzelnen Validierungsschritten ausgegeben. Wenn der Validierungsschritt ohne Fehler abgeschlossen wurde, wird "Valid" ausgegeben, z.B. [H] = Valid.</p> <p>Sollte ein Validierungsschritt fehlerhaft sein, wird "Invalid" und die Bezeichnung des Validierungsschrittes ausgegeben, z.B. [I] = Invalid (Test "SIARD-Erkennung" nicht bestanden).</p> <p>Am Schluss werden das Gesamtergebnis sowie die Pfade zu den einzelnen Dateien angezeigt.</p> | | |
|  <pre> C:\Windows\system32\cmd.exe SIARD Validierung -> Northwind.zip [A] = Valid [B] = Valid [C] = Valid [D] = Valid [E] = Valid [F] = Valid [G] = Valid [H] = Valid [I] = Invalid <Test "SIARD-Erkennung" nicht bestanden> [J] = Valid TOTAL = Invalid SIARD = D:\KOST-Ua1\SIARD\Northwind.zip Log = D:\KOST-Ua1\logs\Northwind.zip.validationlog.log Validierung beendet. </pre> | | | |
| <p>Erfolgt der Aufruf via Script, wird am Schluss noch eine Zusammenfassung aller validierten SIPs ausgegeben.</p> | | | |

- 3 Die Log-Ausgabe führt zusätzliche Detailinformationen über die einzelnen Validierungsschritte auf, insbesondere über die invaliden. Ansonsten ist sie gleich aufgebaut wie die Konsolenausgabe.



```
D:\KOST-Val\logs\Northwind.zip.validationlog.log - Notepad++
File Edit Search View Encoding Language Settings Macro Run TextFX Plugins Window ?
Northwind.zip.validationlog.log
[H] = Valid
[I] = Invalid (Test "SIARD-Erkennung" nicht bestanden)
[I] -- Die SIARD-Datei wurde nicht als solche erkannt, weil sie keine .siard Extension hat.
[J] = Valid

TOTAL = Invalid

SIARD = D:\KOST-Val\SIARD\Northwind.zip
Log = D:\KOST-Val\logs\Northwind.zip.validationlog.log

Validierung beendet.
Normal text | length: 772 lines: 29 | Ln: 1 Col: 1 Sel: 0 | UNIX ANSI INS
```

9 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.

9.1 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)

9.2 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)⁹

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 sind¹⁰?

✗ Andere Ordner oder Dateien sind im content-Ordner verboten

¹⁰ Mit Ausnahme der recordZ.txt und recordZ.bin Dateien in den lobZ-Ordern.

¹⁰ Mit Ausnahme der recordZ.txt und recordZ.bin Dateien in den lobZ-Ordern.

10 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.