Schematron-Validierungen der Fachstelle Bibliothek

Die Fachstelle Bibliothek verwendet zur Erstellung ihrer Analysereports eine Validierung auf Basis von Schematron. Auf dieser Seite stellt sie die entsprechendenen Schematron-Dateien für die verschiedenen Anwendungsprofile zur Verfügung, erläutert ihre Bestandteile und ihren Bezug zum Analyser eport.

Die Anwendung und das Verständnis der Funktionsweise der Schematron-Validierungen setzt Kenntnisse von Schematron und den verwendeten Techniken voraus (XPath, XSLT).

Bei Fragen und Anregungen zu den Schematron-Validierungen wenden Sie sich bitte an die Fachstelle Bibliothek.

- Schematron-Dateien
  - Anwendungsprofil für DDB-METS/MODS
  - Anwendungsprofil für DDB-METS/MODS - Aufbau für Zeitungsausgaben
  - DDB EDM Anwendungsprofil
  - DDB MARC 21 (XML) Anwendungsprofil
- Bestandteile der Schematron-Dateien
  - Versionierung
  - Bezeichnung
  - Meldungstext und Hinweis auf das Anwendungsprofil
  - Verweise auf Beispiele
  - Meldungs-ID
  - Meldungs-Level
  - Properties
    - Allgemein
    - METS/MODS
    - EDM

Schematron-Dateien

Die Schematron-Validierungen der Fachstelle Bibliothek setzen pro Datensatz genau eine Datei voraus. Ein Datensatz muss das jeweilige Wurzelelement enthalten (z. B. <mets:mets>), welches optional in weitere Elemente (z. B. "OAI-Wrapper") eingebettet sein kann.
Die Fachstelle stellt die Validierungen jeweils als Schematron-Datei und als mit der Skeleton Implementierung kompilierte XSL-Datei zur Verfügung. Letztere verwendet sie zusammen mit Saxon-HE zur Erzeugung des Analysereports.

Alle Dateien sowie eine Übersicht der Änderungen stellt die Fachstelle über das GitHub Repository der Deutschen Digitalen Bibliothek zur Verfügung.

Anwendungsprofil für DDB-METS/MODS

- Validierungs Dateien und Übersicht der Änderungen (GitHub)

Diese Validierung prüft Datensätze auf die Anforderungen aus dem Anwendungsprofil für DDB-METS/MODS. Sie ist für Datensätze ausgelegt, die einteilige Dokumente, Teile von mehrteiligen Dokumenten oder Ankersätze beschreiben.

Anwendungsprofil für DDB-METS/MODS - Aufbau für Zeitungsausgaben

- Validierungs Dateien und Übersicht der Änderungen (GitHub)

Diese Validierung prüft Datensätze auf die Anforderungen aus dem Anwendungsprofil für DDB-METS/MODS - Aufbau für Zeitungsausgaben.
Die Validierung ist nur für Datensätze zu Ausgaben ("issue") bzw. Beilagen ("additional") einer Zeitung ausgelegt. Sie kann daher keine Datensätze zu übergeordneten Strukturen (Bände, Zeitungen) prüfen.

DDB EDM Anwendungsprofil

- in Vorbereitung

DDB MARC 21 (XML) Anwendungsprofil

- in Vorbereitung

Bestandteile der Schematron-Dateien

Die Schematron-Dateien enthalten Informationen zur genauen technischen Prüfung und Auswertung. Sie werden im Folgenden beschrieben und stellen - soweit möglich - Bezüge zum Analysereport her.

Versionierung

Die Versionierung erfolgt in der Form des ISO 8601-Zeitstempels im Attribut "shemaVersion" im Wurzelelement "sch:schema".

Beispiel

```
<sch:schema
   ...
   schemaVersion="v2021-07-09T12:13:36"
   ...
   >
   ...
```

Bezeichnung

Die Bezeichnung einer Meldung befindet sich in einem Kommentar über dem Element "sch:assert" bzw. "sch:report".
Sie dient im Analysereport als "Titel" des Meldung.

Beispiel

```
...
<sch:pattern>
   <sch:rule context="mets:xmlData/mods:mods/mods:titleInfo[@type]">
      <!-- Ungültiger Wert im Attribut type von mods:titleInfo -->
      <sch:assert
         id="titleInfo_04"
         ...
      >
    ...
```

Meldungstext und Hinweis auf das Anwendungsprofil

Die Meldung mit dem Hinweis auf die entsprechende Stelle im Anwendungsprofil befindet sich im Elementtext des Elements "sch:assert" bzw. "sch:report".
Beide Elemente können zusätzlich das Unterelement "sch:span" mit dem Attribut "class" mit folgenden Werten enthalten:

- `monotype`
  - Festbreitenschriftart
- `italic`
  - kursiv
- `bold`
  - fett

Diese Angaben verwendet der Analysrerepot für entsprechende Textformatierungen.

Der Hinweis auf das Anwendungsprofil ist in Abhängigkeit von der Meldung ein Text oder eine URL.

Beispiel

```
...
<sch:assert
   id="titleInfo_04"
   ...
   >
   Das Attribut <sch:span class="monotype">type</sch:span> im Element <sch:span class="monotype">mods:titleInfo<
/sch:span> darf nur die folgenden Werte enthalten:
   * <sch:span class="monotype">abbreviated</sch:span>
   * <sch:span class="monotype">translated</sch:span>
   * <sch:span class="monotype">alternative</sch:span>
   * <sch:span class="monotype">uniform</sch:span>
  Ist dies nicht der Fall, wird <sch:span class="monotype">mods:titleInfo</sch:span> bei der Bereinigung des
Datensatzes entfernt.Weitere Informationen zu diesem Attribut finden Sie im DDB-METS/MODS-Anwendungsprofil auf
der Seite mods:titleInfo (https://wiki.deutsche-digitale-bibliothek.de/x/xcIeB). </sch:assert>
...
```

Verweise auf Beispiele

Ein Verweis auf die Best Practice-Beispiele zu den für den Meldung relevanten Elementen befindet sich im Attribut "see" des Elements "sch:assert" bzw. "sch:report".
Im Analysereport dient der Verweis als Link für den Text "Beispiele für dieses Element" unterhalb der Meldung.

Beispiel

```
...
<sch:assert
   id="titleInfo_04"
   ...
   see="https://wiki.deutsche-digitale-bibliothek.de/display/DFD/titleInfo"
   >
   ...
</sch:assert>
...
```

Meldungs-ID

Die Meldungs-ID befindet sich im Attribut "id" des Elements "sch:assert" bzw. "sch:report".
Sie dient in der Meldungslistentabelle als Bezeichnung für das entsprechende Tabellenblatt, auf das der Analysereport verweist (z. B. "Die vollständige Liste der Meldungen finden Sie in der Meldungslistentabelle im Tabellenblatt titleInfo_04").

Beispiel

```
...
<sch:assert
   id="titleInfo_04"
   ...
   >
   ...
</sch:assert>
...
```

Meldungs-Level

Das Meldungs-Level befindet sich im Attribut "role" des Elements "sch:assert" bzw. "sch:report".
Es dient zur Einordnung der Schwere des Meldung. Der Analysereport ordnet dem Meldungs-Level eine Farbe und Texte für die Grafik in der Übersicht und den Hinweis am Ende der einzelnen Meldungen zu.

Die Validierungen verwenden die folgenden Werte:

- `info`
  - blau
  - Datensätze mit Elementen, die in der DDB nicht berücksichtigt werden. (Grafik)
  - Diese Elemente werden in der DDB nicht berücksichtigt. (Meldung)
- `warn`
  - gelb
  - Fehlerhafte Datensätze, die unverändert eingespielt werden. (Grafik)
  - Dieser Fehler verhindert nicht das Einspielen der Datensätze. (Meldung)
- `caution`
  - grau
  - Datensätze, bei denen ein Verdacht auf Fehler vorliegt (Grafik)
  - Bei diesen Elementen besteht ein Verdacht auf Fehler (Meldung)
- `error`
  - orange
  - Fehlerhafte Datensätze, die von der Fachstelle korrigiert werden. (Grafik)
  - Dieser Fehler wird von der Fachstelle Bibliothek bereinigt. (Meldung)
- `fatal`
  - rot
  - Fehlerhafte Datensätze, die nicht eingespielt werden. (Grafik)
  - Dieser Fehler verhindert das Einspielen der Datensätze. (Meldung)

Beispiel

```
...
<sch:assert
   id="titleInfo_04"
   role="error"
   ...
   >
   ...
</sch:assert>
...
```

Properties

Die Schematron-Validierungen verwenden in einigen Prüfungen Schematron-Properties (im Attribut "properties" des Elements "sch:assert" bzw. "sch:report").
Sie dienen dazu, relevante Elemente in den Datensätzen einfacher zu lokalisieren (Beispieldatensätze im Analysereport, Spalte "B" in der Meldungslistentabelle) bzw. im Analysereport Übersichten mit ungültigen Werten zu erzeugen.
Es gibt sowohl allgemeine als auch formatspezifische Properties, die im Folgenden beschrieben sind.

Beispiel

```
...
<sch:assert
   id="titleInfo_04"
   ...
   properties="dmd_id type"
   ...
   >
   ...
</sch:assert>
...
```

Allgemein

Die folgenden Properties beziehen sich auf allgemeine XML Aspekte:

- `text`
  - Der Elementinhalt (text()) des relevanten Elements
- `local_name`
  - Der Name des relevanten Elements ohne Präfix (local-name())
- `name`
  - Der Name des relevanten Elements mit Präfix (name())

METS/MODS

Die folgenden Properties finden in den Validierungen für METS/MODS-Anwendungsprofile Verwendung:

- `dmd_id`
  - Der Wert des Attributs ID der mets:dmdSec, in der sich das relevante Element befindet.
- `id`
  - Der Wert des Attributs ID des relevante Elements (z. B. mets:div)
- `type`
  - Der Wert des Attributs type des relevante Elements (z. B. mods:note)
- `TYPE`
  - Der Wert des Attributs TYPE von mets:div Elementen im Element mets:structMap[@TYPE='LOGICAL']
- `source`
  - Der Wert des Attributs source des relevanten Elements (z. B. mods:recordIdentifier)

EDM

Die folgenden Properties finden in den Validierungen für EDM-Anwendungsprofile Verwendung:

- `xpath_parent`
  - Der XPath des relevanten Elements beginnend mit dem Klassenelement
- `about`

    - Der Wert des Attributs rdf:about des relevanten Elements (Klassen-Element, z. B. edm:ProvidedCHO)
- `lang`
    - Der Wert des Attributs xml:lang des relevanten Elements (Property-Element, z. B. skos:prefLabel)