



Advanced EDM Schematron validation in Oxygen XML editor

The XML Schema for the Europeana Data Model (EDM)¹ allows for automatic validation of EDM metadata represented in EDM. However, the XML Schema language has shortcomings: not all EDM constraints can be expressed with it. To capture more validation rules, the EDM XML Schema has been augmented with rules that follow the Schematron² approach. This document explains how to make use of these constraints with the Oxygen XML editor.³

1. Preliminary steps

Download <http://www.w3.org/2001/xml.xsd> to a local folder on your machine. Then create an XML file called `catalog.xml` with the following contents.

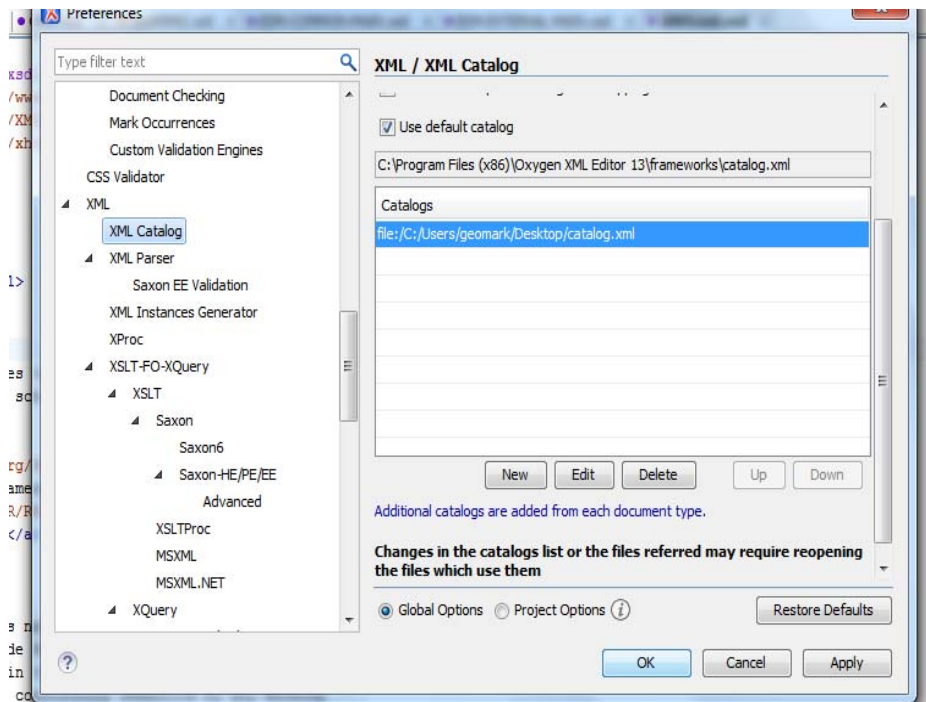
```
<?xml version="1.0"?>
<catalog xmlns="urn:oasis:names:tc:entity:xmlns:xml:catalog">
  <uri name="http://www.w3.org/2001/xml.xsd"
    uri="<path_to_your_local_directory>/xml.xsd"/>
</catalog>
```

Then within Oxygen XML editor go to “Options/Preferences” and then to “XML/XML Catalog” as shown in the image below:

¹ <http://pro.europeana.eu/edm-documentation>

² <http://www.schematron.com/>

³ <http://www.oxygenxml.com/>

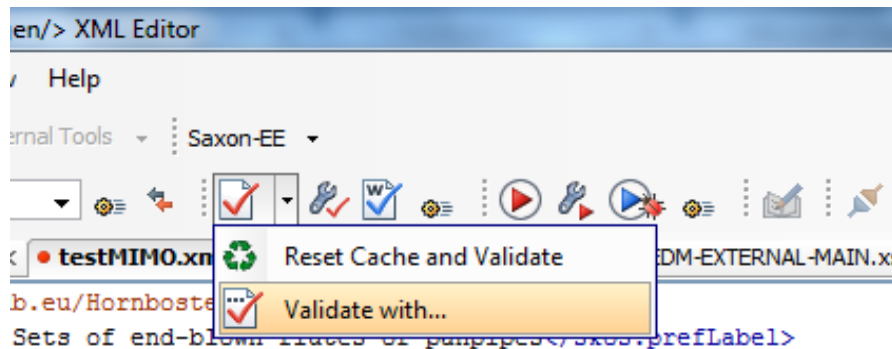


Press New and assign the new catalog to Oxygen.

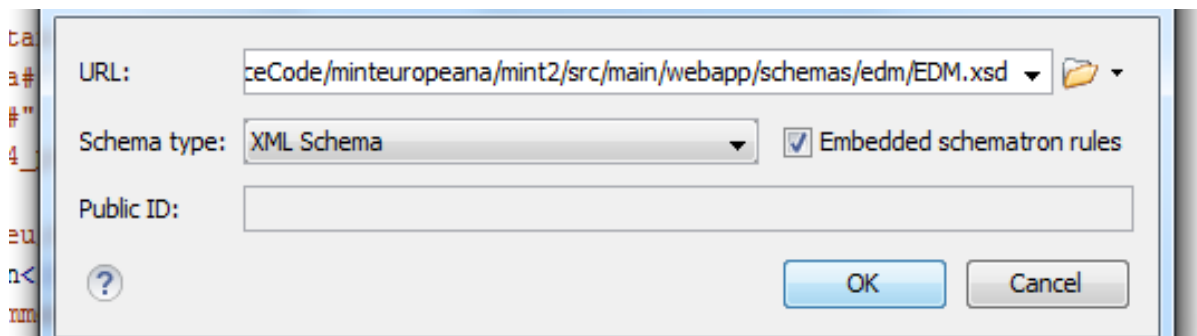
2. Open and validate an XML file in Oxygen

Select File->Open to open the XML file in Oxygen.

Select the “Validate with” icon as shown in the image below:



In the appearing menu set the location of the XSD to EDM.xsd (either locally stored or the http location) as shown below (you shouldn't forget to check the “Embedded Schematron Rules” box):





After pressing OK, the validation results will appear on the bottom of the screen as shown below. These will include both structural XSD validation and Schematron validation messages.

The screenshot displays the XML Editor interface with the following components:

- XML Code (Lines 37-57):**

```
<edm:hasView rdf:resource="http://www.mimo-db.eu/media/GNM/IMAGE/MIR161_1296203446687_2.jpg" />
<edm:isShownAt rdf:resource="http://www.mimo-db.eu/media/GNM/IMAGE/MIR161_1296203446687_2.jpg" />
<edm:isShownBy rdf:resource="http://www.mimo-db.eu/media/GNM/IMAGE/MIR161_1296203446687_2.jpg" />
<edm:object rdf:resource="http://www.mimo-db.eu/media/GNM/IMAGE/MIR161_1296203446687_2.jpg" />
<edm:provider>MIMO - Musical Instrument Museums Online</edm:provider>
<edm:rights rdf:resource="http://creativecommons.org/licenses/by-nc-sa/3.0/" />
</ore:Aggregation>
<edm:Proxy rdf:about="#GNM:740090">
  <dc:coverage rdf:resource="http://sws.geonames.org/2867714/" />
  <dc:creator rdf:resource="http://www.mimo-db.eu/InstrumentMaker/Person/412" />
  <dc:creator rdf:resource="">Bruggen</dc:creator>
  <dc:date>(um) 1900</dc:date>
  <dc:description>>15 Rohrpfleifen</dc:description>
  <dc:identifier>#GNM:740090</dc:identifier>
  <dc:title xml:lang="de">Rohrpfleife</dc:title>
  <dc:type rdf:resource="http://www.mimo-db.eu/InstrumentsKeywords/4024" />
  <dc:type rdf:resource="http://www.mimo-db.eu/HornboitelundSachs/258" />
  <edm:type>TEXT</edm:type>
</edm:Proxy>
<edm:EuropeanAggregation rdf:about="#GNM:740090">
```
- Attributes Table:**

Attribute	Value
rdfabout	#GNM:740090
- Validation Errors (Bottom):**
 - ISO Schematron (NSLT 2.0) id: #GNM:740090 - Within a Proxy context, dclanguage is mandatory when dclanguage has the value 'TEXT'. (not(edm:type='TEXT') or (edm:type='TEXT' and exists(dclanguage))) [assert]
 - testMIMO.xml (1 item) E [Xerces] cvc-complex-type.2.4.a: Invalid content was found starting with element 'edmidataProvider'. One of: '{http://www.europeana.eu/schemas/edm/ aggregatedCHO}' is expected.
 - testMIMO.xml (1 item) E ISO Schematron (NSLT 2.0) id: #GNM:740090 - Within a Proxy context, dclanguage is mandatory when dclanguage has the value 'TEXT'. (not(edm:type='TEXT') or (edm:type='TEXT' and exists(dclanguage))) [assert]
- Errors Panel (Bottom Right):** Validation failed.



Acknowledgements, contact

The first version of this document has been created by Georgios Markakis, Europeana Office. For more information please refer to the EDM documentation pages at <http://pro.europeana.eu/edm-documentation>.

Thanks to Kilian Schmidtner (DM2E - Digitised Manuscripts to Europeana, State Library Berlin) for testing the schema, feedback and inspiration for using Schematron.