

HRA – Visual Resources | Tutorial

Manual for VRA Core 4 Transform Tool

(Version 0.0.7.)

Contents

1. Introduction and project history	3
2. Data preparation	5
2.1. Predefined headers	5
2.2. Repeated values	5
2.3. Display values	6
2.4. Controlled data	7
3. Introducing the template	8
3.1. Downloading the template	8
3.2. The template explained.....	8
3.3. Creating a .csv file	10
4. Accessing the transform tool online	11
4.1. The basic user interface.....	11
5. Uploading a file	12
6. Generating VRA Core 4 XML.....	14
6.1. Selecting a mapping	14
6.2. Selecting a transformation	14
6.3. Defining the records to be processed.....	15
6.4. Generating XML.....	16
6.5. Allow pop-ups	18
7. The interface in Advanced Mode	19
7.1. Applying XSLs	19
7.2. Validating the XML.....	21
8. If validation fails	23
8.1. Example 1 (element type):	23
8.2. Example 2 (element type):	24
8.3. Example 3 (date pattern):	24
8.4. Example 4 (work/image ID):.....	25
9. Generating RDF XML	27
10. Contact.....	29
11. Appendix: Elements, displays and repetitions.....	30
12. Appendix: Columns Full Template.....	32

1. Introduction and project history

The *VRA Core 4 XML Transform Tool* converts descriptive image metadata from flat tables (.csv) to structured VRA Core 4 XML. It makes use of a template with predefined headers. Users may work with the tool in a simple or advanced mode. The transform tool performs XML validation and provides feedback in case of errors.

This manual covers the whole workflow of using the *VRA Core 4 XML Transform Tool* including descriptions of all its features:

- preparation of data
- uploading data
- adjusting usage settings
- data transformation
- data validation
- handling validation feedback
- downloading the XML files

It also covers the (still experimental) transformation to RDF XML and includes a number of appendices for further reference.

Development of the *VRA Core 4 XML Transform Tool* was based on an initiative by Susan Jane Williams and resulted in a successful application by Matthias Arnold for a [Project Grant](#) funded by the [Visual Resources Association Foundation](#) (VRAF) in fall 2014. Additional support was given by the [Heidelberg Research Architecture](#) (HRA) at the [Cluster of Excellence "Asia and Europe in a Global Context"](#), Heidelberg University. The goals of the collaborative project were described in the December 2014 issue of VRAF News:

"Matthias Arnold, Heidelberg Research Architecture, University of Heidelberg, will use the award to support the development of the VRA Core 4 XML Transform Tool. This tool will enable any user who can supply descriptive image metadata in a standardized CSV form (comma separated values, e.g. via Excel) to create validating VRA Core 4 XML. Useful on its own, this XML also represents an important step for further transformations to other XML schemas, like RDF to support output of data as LOD (Linked Open Data). The tool is being developed in consultation with both the VRA Data Standards and the Core OC Committees, will be shared with the image management community during beta-testing for feedback, and will be open-source and freely available upon completion. Susan Jane Williams (Independent Cataloging and Consulting Services) will serve as coordinator for the mapping templates and will help create documentation, demonstrations (sample records and mappings), and further mapping help in use of the tool. We are confident that this tool will have broad value and impact within the visual resources community, and for the VRA Core 4 standard."

["2014-2015 VRA Foundation Project Grant Winners Named"](#), VRAF News & Events, December 19, 2014

The template csv files and the mappings to XML were developed by Susan Jane Williams, Greg Reser, and Matthias Arnold. Implementation and development of the tool was coordinated by the Heidelberg side, programming was conducted by HRA's senior developer Matthias Guth.

The tool is still in BETA version. Any feedback on issues, bugs or problems encountered is very much appreciated and may be directed to the developers (see section "Contact").



The *VRA Core 4 XML Transform Tool* is developed as Open Source, its source code can be found at: <https://github.com/exc-asia-and-europe/csv2xml>

Note: This user manual will be updated with every new feature or functionality. The latest version can be found on GitHub:

<https://github.com/exc-asia-and-europe/csv2xml/tree/master/doc>.

2. Data preparation

The *VRA Core 4 XML Transform Tool* uses **predefined headers** to identify data in the template in order to map the values to the correct elements (or attributes) in XML. These headers can be arranged in templates (for example in .xlsx or .ods format tables).

Any descriptive image metadata can be used for transformation, be it an export from your institution's Digital Asset Management System, metadata embedded in images, or a personal image database. You only need to do two things for preparation:

1. Use the predefined template headers
2. Export data to comma separated value (.csv) file

2.1. Predefined headers

Currently, the template offers 231 columns with fixed headers for your data. To be able to transform your .csv data into XML you have to use these headers.

These include references to your local system, like "IMAGE_Filename" or "IMAGE_LocalRepoName". In the list a number of elements are repeated to allow the transformation of multiple values, for example it offers up to three agents and up to eight subjects for the WORK record. The template also includes many columns for references to controlled vocabularies, like the name and Ref-ID of a subject in a vocabulary, and data types controlled by the VRA Core 4 schema.

For example, each of the eight subjects of a WORK record can be expressed in four columns:

```
WORK_Subject1  
WORK_SubjectType1  
WORK_SubjectVocab1  
WORK_SubjectRefid1
```

The complete list of headers can be found in "*Appendix: Columns Full Template*" below.

IMPORTANT:

You do not need to fill each column with data. The sequence of columns (headers) may be changed and you can even delete columns you will not need.

It is, however, essential **not to change the headers** themselves, i.e. not their "text" or "values".

2.2. Repeated values

For some of the eighteen elementSet's of VRA Core 4 XML for WORK and IMAGE records elements can be repeated. These values usually are combined with other information, like references to authority files.



WORK

Agent:	3
CulturalContext:	2
Date:	2
Location:	2
Material:	4
Measurements:	8 (in 2 sets)
Relation:	2 (plus 1 work-image)
Style/Period:	4
Subject:	8
Technique:	4
Title:	2 (preferred/alternative)
Worktype:	3

IMAGE

Subject:	3
----------	---

An expanded list of repeated elements can be found below, in "*Appendix: Elements, displays and repetitions.*"

2.3. Display values

The template is designed to include data for each of the eighteen elementSet's of VRA Core 4 XML for WORK and most of the IMAGE records. For a large number of elementsSets structured (or "qualified") information can be transformed.

All elementSets may have a display element, with the exception of the following four IMAGE elementSets, which are NOT included in the template: IMAGE_CulturalContext, IMAGE_StateEdition, IMAGE_StylePeriod, IMAGE_Textref.

For more information about the elementSets, which one will only contain display values, and which one will also include other elements and attributes please refer to "*Appendix: Elements, displays and repetitions*" below.

2.3.1. Auto-filling displays

The *VRA Core 4 XML Transform Tool* allows transforming both, structured data for sub-elements and attributes, as well as unstructured data for display values.

If users do NOT have <display> data in their .csv file, *VRA Core 4 XML Transform Tool* will automatically fill the content of the <display> element based on available data from the structured information.

For example, if the template contains

WORK_AgentDisplay	<i>[empty]</i>
WORK_Agent1NameType	personal
WORK_Agent1Name	Michelangelo Buonarroti
WORK_Agent1Role	painter
WORK_Agent1Attribution	school of

The following the display element is created:

<display>school of Michelangelo Buonarroti (painter)</display>

However, if users do provide data for the display element (in the example "WORK_AgentDisplay") these values will be used and NOT overwritten or changed.

2.4. Controlled data

[VRA Core 4.0](#) in the restricted version uses controlled type lists and date formats.

2.4.1. Data type values

For the **type values** please refer to the [VRA Core 4.0 Restricted Schema Type Values](#) document.

2.4.2. Date values

For **date values** there is no individual documentation file, but the [vra-strict.xsd](#) states for dateValueType:

"...Defines a date which follows the ISO 8601 date format, and allows right truncation. [...] In brief, the following formats are allowed:

present	
2006	(2006)
2006-12	(December, 2006)
2006-12-31	(31 December 2006)
-44	(44 BCE)
-44-03	(March, 44 BCE)
-44-03-15	(15 March, 44 BCE)
-10000000	(10 Million Years Ago)
-100000000000	(100 Billion Years Ago)"

Source: <http://loc.gov/standards/vracore/vra-strict.xsd>, last accessed December 12, 2015

2.4.3. Controlled data in the Transform tool

IMPORTANT:

To be able to convert your data into valid VRA Core 4 XML these type values must be used and date format rules must be followed.

The transform tool will assist you as far as possible.

For example, in <date> the sub-element <latestDate> must not be empty. But if you only entered <earliestDate> the tool will automatically fill <latestDate> with the <earliestDate> value.

If you accidentally mistyped a controlled type value, the validation will point you to this error in the validation results (see chapter "Validation errors" below).

3. Introducing the template

3.1. Downloading the template

The **full template** for *VRA Core 4 XML Transform Tool* is available for download at

<https://github.com/exc-asia-and-europe/csv2xml/tree/master/doc>

in these formats:

- Comma-separated values textfile (.csv)
- Excel (.xlsx)
- OpenDocument spreadsheet (.ods)

3.2. The template explained

The template consists of three worksheets.

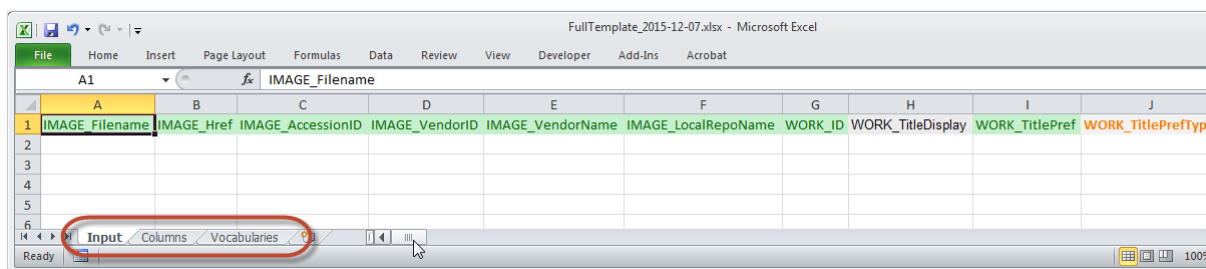


Figure 1: Full template - Input worksheet (in Excel)

The **"Input"** worksheet contains all the columns with their respective header in horizontal layout (colour coded), so that it can be used for editing data.

The **"Columns"** worksheet contains two columns. In the first, all headers are arranged vertically and colour coded, while in the second some hints to the content of the data are given.

The colour codes are included to provide visual feedback in case different types of data are expected for a "field".

Display values, for which the tool will auto-fill content if no data was provided for display are shown in black on grey background.

WORK_TitleDisplay

Type values are shown in orange on light grey. Allowed values are provided in the second column.

IMAGE_RightsType copyrighted, publicDomain, undetermined, other

References to external vocabularies are shown in dark blue on orange background.

WORK_Agent1NameVocab



Special content relevant for validation is shown in black on red background, together with the validation rules. This is used for dates.

WORK_EarliestDate1	VALIDATION <xsd:pattern value="present (-)*[0-9]{1,12}(-[0-9]{2}(-[0-9]{2})*)**"/>
WORK_LatestDate1	VALIDATION <xsd:pattern value="present (-)*[0-9]{1,12}(-[0-9]{2}(-[0-9]{2})*)**"/>

The "normal" columns are shown in dark green on lighter green.

A	B	C	D	E
IMAGE_VendorName	for IMAGE_source			
IMAGE_LocalRepoName	for IMAGE_location			
WORK_ID	for WORK_work			
WORK_TitleDisplay				
WORK_TitlePref				
WORK_TitlePrefType	brandName, cited, creator, descriptive, former, inscribed, owner, popular, repository, translated, other			
WORK_TitleAlt				
WORK_TitleAltType	brandName, cited, creator, descriptive, former, inscribed, owner, popular, repository, translated, other			
WORK_AgentDisplay				
WORK_Agent1NameType	personal, corporate, family, other			
WORK_Agent1Name				
WORK_Agent1NameVocab	Vocabularies!A1			
WORK_Agent1NameRefid				
WORK_Agent1Role				
WORK_Agent1Attribution				
WORK_Agent2NameType	personal, corporate, family, other			
WORK_Agent2Name				
WORK_Agent2NameVocab	Vocabularies!A1			
WORK_Agent2NameRefid				
WORK_Agent2Role				
WORK_Agent2Attribution				
WORK_Agent3NameType	personal, corporate, family, other			
WORK_Agent3Name				
WORK_Agent3NameVocab	Vocabularies!A1			
WORK_Agent3NameRefid				
WORK_Agent3Role				
WORK_Agent3Attribution				
WORK_CulturalContextDisplay				
WORK_CulturalContext1				
WORK_CulturalContext1Vocab	Vocabularies!A1			
WORK_CulturalContext1Refid				
WORK_CulturalContext2				
WORK_CulturalContext2Vocab	Vocabularies!A1			
WORK_CulturalContext2Refid				
WORK_DateDisplay				
WORK_Date1Type	alteration, broadcast, bulk, commission, creation, design, destruction, discovery, exhibition, inclusive, performance, publicat			
WORK_EarliestDate1	VALIDATION <xsd:pattern value="present (-)*[0-9]{1,12}(-[0-9]{2}(-[0-9]{2})*)**"/>			
WORK_LatestDate1	VALIDATION <xsd:pattern value="present (-)*[0-9]{1,12}(-[0-9]{2}(-[0-9]{2})*)**"/>			
WORK_Date2Type	alteration, broadcast, bulk, commission, creation, design, destruction, discovery, exhibition, inclusive, performance, publicat			
WORK_EarliestDate2	VALIDATION <xsd:pattern value="present (-)*[0-9]{1,12}(-[0-9]{2}(-[0-9]{2})*)**"/>			
WORK_LatestDate2	VALIDATION <xsd:pattern value="present (-)*[0-9]{1,12}(-[0-9]{2}(-[0-9]{2})*)**"/>			
WORK_DescriptionDisplay				
WORK_DescriptionSource				

Figure 2: Full template - Columns workbook (in Excel)

The "**Vocabularies**" worksheet contains a list of external authorities or controlled vocabularies. It is by no means comprehensive. The values of the "vocab" attribute are not controlled by the VRA Core 4 schema, but it is strongly recommended to make sure the content is consistent within your data set. This is especially of interest if you intend to transform your data to RDF and use it as Linked (Open) Data.

	A	B	C	D	E	F
1	AAT	http://vocab.getty.edu/aat/				
2	TGN	http://vocab.getty.edu/tgn/				
3	ULAN	http://vocab.getty.edu/ulan/				
4	CONA	http://vocab.getty.edu/cona/				
5	LCNAF	http://id.loc.gov/authorities/names/				
6	LCSH	http://id.loc.gov/authorities/subjects/				
7	LCGFT	http://id.loc.gov/authorities/genreForms/				
8	TGM	http://id.loc.gov/vocabulary/graphicMaterials/				
9	VIAF	http://viaf.org/viaf/				
10	ICONCLASS	http://iconclass.org/				
11						
12						

Figure 3: Full template - Vocabularies workbook (in Excel)

3.3. Creating a .csv file

Export records from your local system directly into .csv files if you can include the template headers.

Alternatively, use your preferred spreadsheet software (e.g. Microsoft Excel, OpenOffice Calc, or LibreOffice Calc) as intermediary to adjust the headers and export from there into .csv files. In the [doc folder on GitHub](#) an Excel (.xlsx) version of the template is available.

IMPORTANT:

Only .csv files can be transformed so make sure you convert spreadsheets (e.g. from Excel or Calc) to .csv text files before transformation.

Use UTF-8 as encoding to conserve possible special characters in your data.

For a guide on how to convert Microsoft Excel files into comma separated value (.csv) files see "Appendix: Converting .xlsx file to .csv using OpenOffice Calc" below.

To be covered:

Other data sources: EMWG export-import tool

4. Accessing the transform tool online

At the moment, the *VRA Core 4 XML Transform Tool* can be accessed online for testing at <http://kjc-ws2.kjc.uni-heidelberg.de:8081/exist/apps/csv2xml/index.xq>

Please note this is a server dedicated to testing software and developments.

4.1. The basic user interface

The *VRA Core 4 XML Transform Tool* interface offers a Simple and an Advanced Mode.

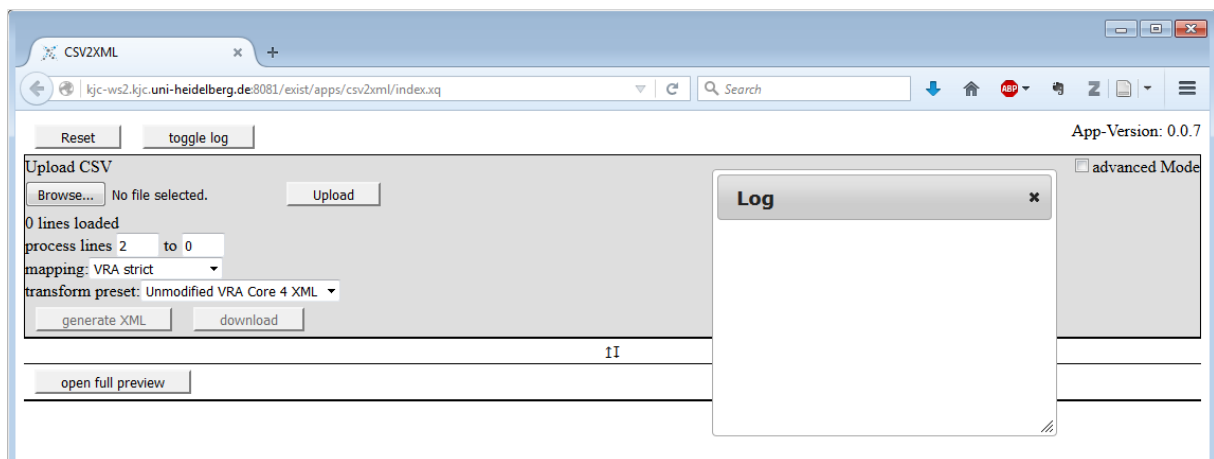


Figure 4: User interface - Simple mode.

In simple mode the XML generation and validation are processed in one step finishing with the optional download of the XML files.

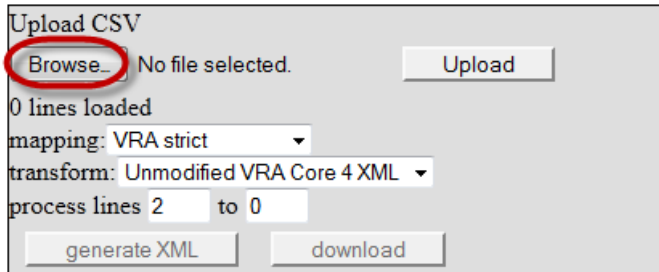
The advanced interface provides additional functionalities for the advanced user including the possibility to apply additional transformations.

Both interfaces offer a "Log" window which is visible by default. While being used, the tool displays messages here, for example which record is currently processed. It can be switched off/on using the "toggle lock" button.

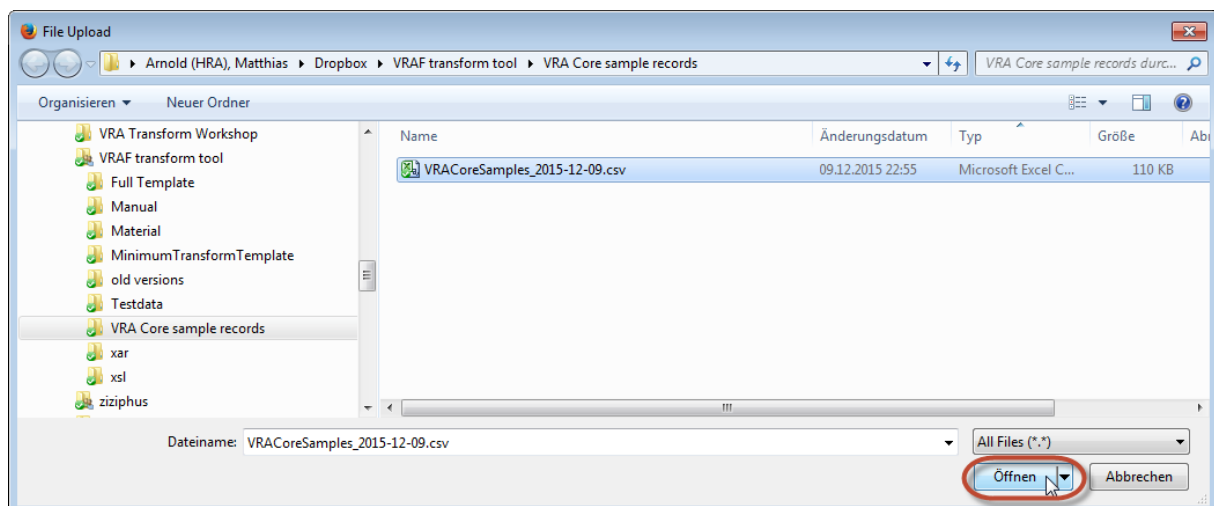


5. Uploading a file

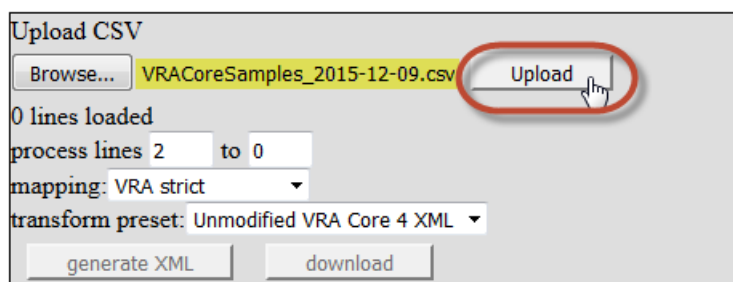
To transform data you need to upload it first. Click the "Browse" button:



Select the .csv file on your computer and click "Open".



The name of the .csv file will be displayed in the interface (in this example we use: "VRACoreSamples_2015-12-09.csv"). Now click on "Upload".



Once the upload is finished you will see an message "CSV upload successful" in the log window.



Reset

toggle log

App-Version: 0.0.7

Upload CSV

Browse...

VRACoreSamples_2015-12-09.csv

Upload

44 lines loaded

process lines to

mapping: VRA strict

transform preset: Unmodified VRA Core 4 XML

generate XML

download

Log

13:50:38: CSV upload successful
13:50:38: Upload successful.
13:50:30: Session resetted.

advanced Mode

11

open full preview

6. Generating VRA Core 4 XML

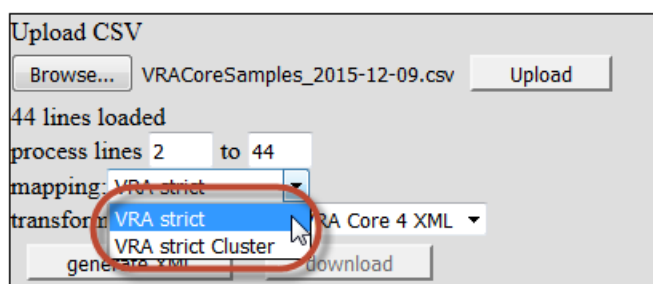
After uploading a .csv file to the *VRA Core 4 XML Transform Tool* you can just hit the "generate XML" button. This will process all records into VRA Core 4 XML using the default settings (transform all records to VRA strict without any modifications).

However, even in Simple Mode you can manipulate how the XML is generated. You have three options:

- Select another mapping
- Select another transform
- Define the range of records to be processed

6.1. Selecting a mapping

The *VRA Core 4 XML Transform Tool* supports generating XML based on different mappings. For VRA Core, two variants are available: "VRA strict" and "VRA strict Cluster".



6.1.1. VRA strict

VRA strict is the **default mapping**. It is based on the VRA Core 4 restricted version (<http://www.loc.gov/standards/vracore/vra-strict.xsd>). It uses controlled values for types and date formats.

For more information please refer to <http://www.loc.gov/standards/vracore/schemas.html>.

6.1.2. VRA strict Cluster

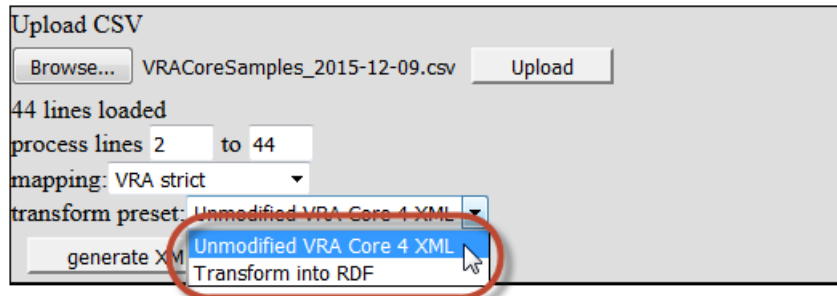
The VRA strict Cluster mapping is an expansion of the restricted VRA Core 4 version developed at the Cluster of Excellence "Asia and Europe in a Global Context", University of Heidelberg. It contains additional elements and attributes, for example, attributes for multilingual data, role attributes for agents, and an element for geo-coordinates (<http://cluster-schemas.uni-hd.de/vra-strictCluster.xsd>).

For more information please refer to the document "[VRA Core – Extensions](#)".

6.2. Selecting a transformation

The tool supports different transformations.

The current version of *VRA Core 4 XML Transform Tool* offers two built-in transformations. Additional .xsl files may be added in the source code. In a future version, this may be changed (e.g. upload transformation stylesheets in the user interface).



Currently, two transformations are implemented:

- Unmodified VRA Core 4 XML
- Transform into RDF (explanations see below)

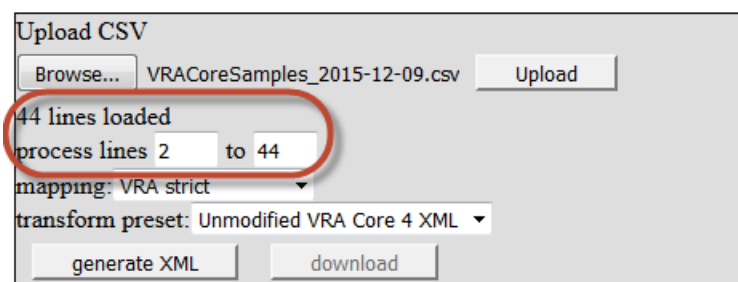
To transform the .csv data into VRA Core 4 XML choose "Unmodified VRA Core 4 XML".

6.3. Defining the records to be processed

By default, all records listed in the uploaded .csv file will be processed.

6.3.1. *Calculating records for processing*

Within the .csv file a record equals a line. In the template, the first record (first line) contains the column headers. Therefore, the *VRA Core 4 XML Transform Tool* will by default process data beginning with line #2 (i.e. record #1).





6.3.2. Defining a range of records

The *VRA Core 4 XML Transform Tool* also allows you to define the range of records to be processed. This is very helpful if you want to check a smaller range of records or even individual ones, look at their xml and test if they validate.

Upload CSV
Browse... VRACoreSamples_2015-12-09.csv Upload
44 lines loaded
process lines 5 to 5
mapping: VRA strict
transform preset: Unmodified VRA Core 4 XML
generate XML download

In this example, only line 5 (i.e. record #4) will be processed.

6.4. Generating XML

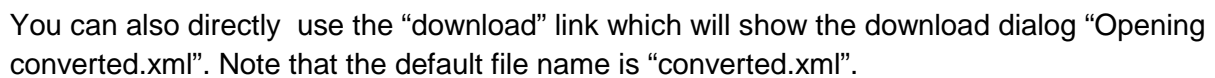
If all settings are made, click the “generate XML” button.

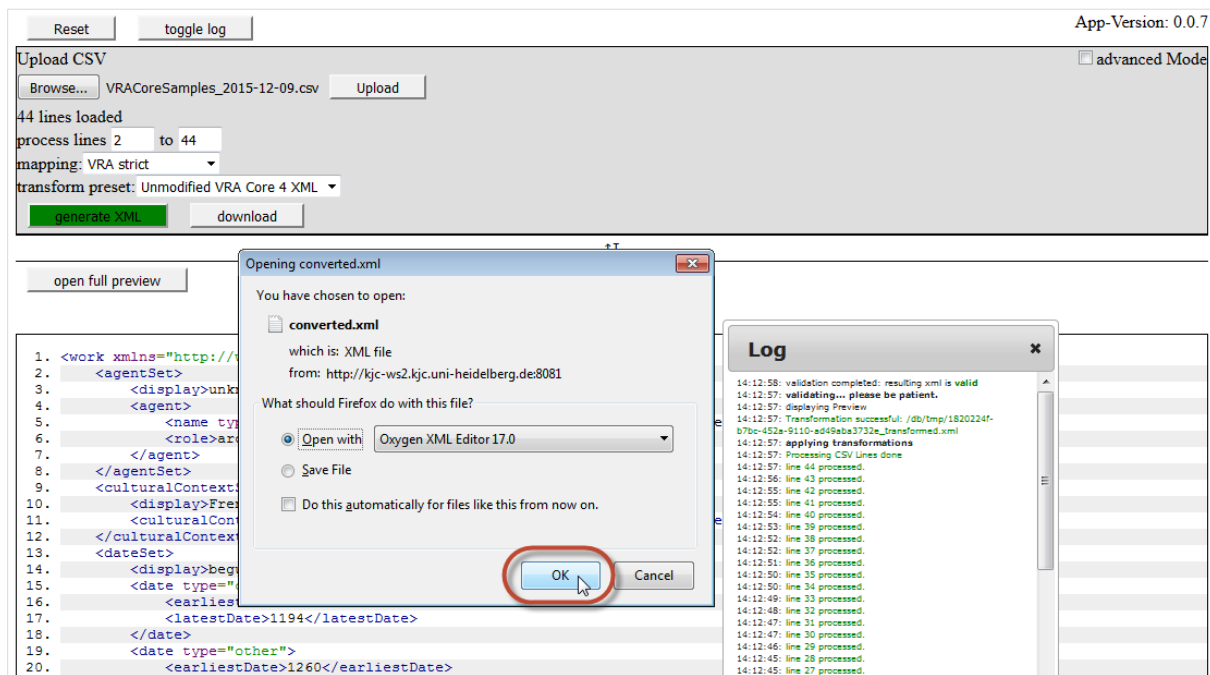
The tool will start generating the XML. You can monitor the progress in the “Log” window.

Upload CSV
Browse... VRACoreSamples_2015-12-09.csv Upload
44 lines loaded
process lines 2 to 44
mapping: VRA strict
transform preset: Unmodified VRA Core 4 XML
generate XML download

Log
14:12:32: line 9 processed.
14:12:31: line 8 processed.
14:12:31: line 7 processed.
14:12:30: line 6 processed.
14:12:29: line 5 processed.
14:12:28: line 4 processed.
14:12:28: line 3 processed.
14:12:27: line 2 processed.
14:12:26: Start processing 43 lines
14:12:26: -> 1820224f-b7bc-452a-9110-ad49eba3732e.xml
14:12:26: Generating parent document...
14:12:26: template loaded: vra-image-template.xml
14:12:26: template loaded: vra-work-template.xml
14:12:26: loadingTemplates
14:12:26: generating...
13:50:38: CSV upload successful
13:50:38: Upload successful.
13:50:30: Session resettled.

After generating XML a preview of the code is displayed while the tool automatically runs a validation. When all is finished the “generate XML” button turns color.

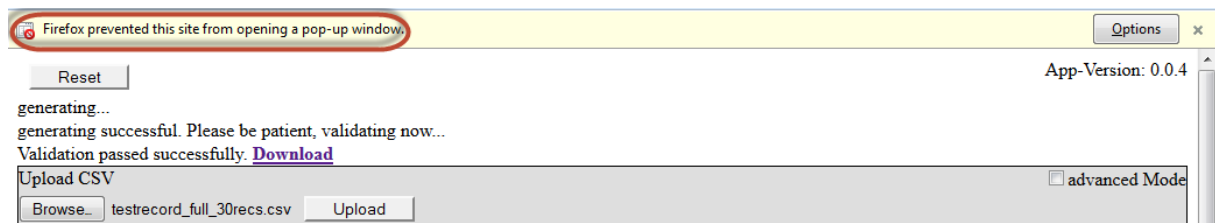




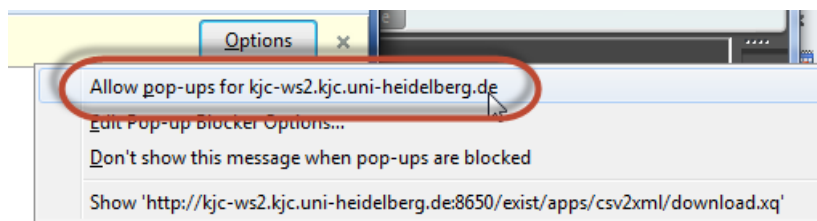
You can save the file on a disk or open it in your preferred XML editor.

6.5. Allow pop-ups

You can also click on the "open full preview" button to see the whole XML in your browser. This will open a new tab. Depending on your local browser settings, opening the new tab may trigger a pop-up blocking tool. In Firefox it may look like this:



Choose "Allow pop-ups" from the "Options" menu.



7. The interface in Advanced Mode

In the Advanced mode it is possible to

- apply XSL transforms
- choose a schema against which you wish to validate your XML file
- upload your own schema file

It also separates XML generation from validation.

Reset toggle log App-Version: 0.0.7

☒ advanced Mode

Upload CSV

Browse... No file selected. Upload

44 lines loaded

process lines 2 to 44

mapping: VRA strict

generate XML

transform preset: Unmodified VRA Core 4 XML

Apply XSLs (for transformation):

xsl/cleanup-vra.xsl

(modifying the predefined selection may break pagination)

transform XML

Validate against:

<http://www.loc.gov/standards/vracore/vra-strict.xsd>

<http://www.loc.gov/standards/vracore/vra.xsd>

add Schema (.xsd .dtd):

validate download

open full preview

Log

Figure 5: User interface - Advanced mode.

7.1. Applying XSLs

During the default generation of "Unmodified VRA Core 4 XML" a transformation script "cleanup-vra.xsl" will be automatically applied. It removes empty attributes and elements and keeps the resulting XML file slim and clean. It is possible to deselect it here. You should note, however, that without running "cleanup-vra.xsl" the resulting XML code may not validate.

If you choose "Transform into RDF" as transform preset the "VRA2RDF.xsl" will be loaded and applied by default. For more information about RDF transformation see chapter "Generating RDF" below.

It is possible to add other .xsl files in the source code. In a future version of *VRA Core 4 XML Transform Tool* this may be changed.

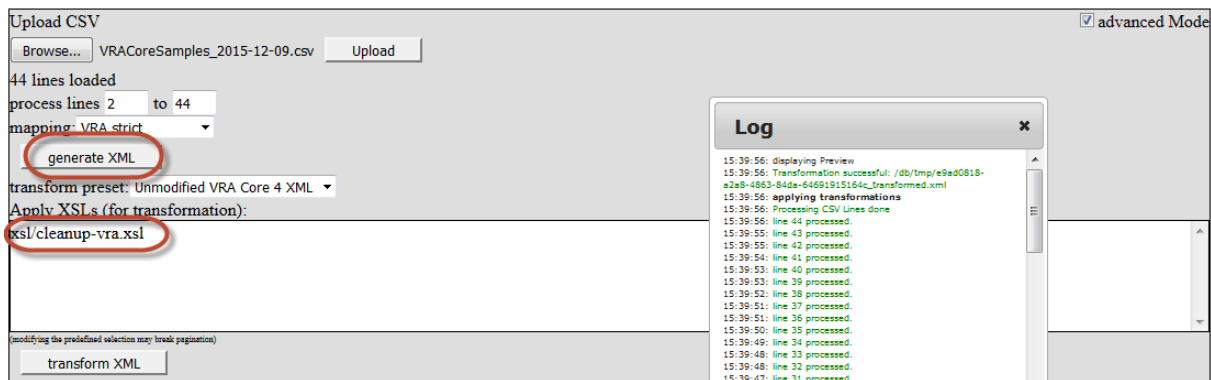


Figure 6: Advanced Mode - XML generation without additional transformation



Figure 7: Preview of XML without clean-up

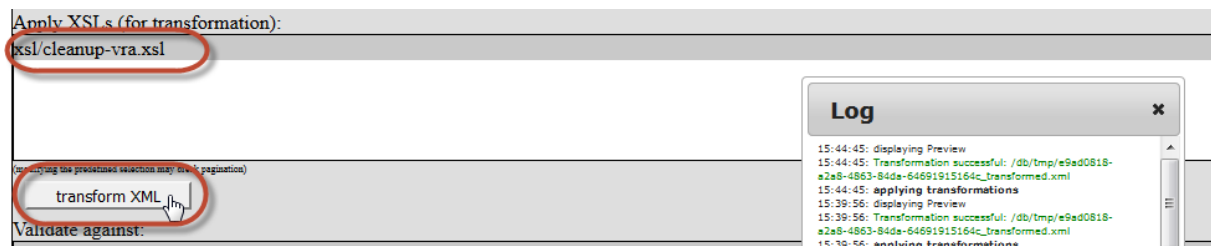


Figure 8: Starting the clean-up manually



1 / 80 ->

```
1. <work xmlns="http://www.vraweb.org/vracore4.htm" id="w_1">
2.   <agentSet>
3.     <display>unknown (French architect)</display>
4.     <agent>
5.       <name type="personal" vocab="ULAN" refid="500125235">unknown French</name>
6.       <role>architect</role>
7.     </agent>
8.   </agentSet>
9.   <culturalContextSet>
10.    <display>French</display>
11.    <culturalContext vocab="AAT" refid="300111188">French</culturalContext>
12.  </culturalContextSet>
13.  <dateSet>
14.    <display>begun 1194 (creation); consecrated 1260 (other)</display>
15.    <date type="creation">
16.      <earliestDate>1194</earliestDate>
17.      <latestDate>1194</latestDate>
18.    </date>
19.    <date type="other">
20.      <earliestDate>1260</earliestDate>
21.      <latestDate>1260</latestDate>
22.    </date>
23.  </dateSet>
24.  <descriptionSet>
25.    <display>The present cathedral was constructed on the foundations of the earlier church; the oldest parts of the cathedral a
26.    <description source="CCO (Cataloging Cultural Objects) Catalog Examples; http://www.vrafoundation.org/ccoweb/cco/examplesind
27.  </descriptionSet>
```

Figure 9: Preview of XML after clean-up

7.2. Validating the XML

Transforming data from .csv to XML does not *per se* guarantee a valid XML result. This is because the restricted VRA Core schema defines a number of fixed values. If the values in the .csv document do not match the values expected by the schema, validation will fail. We therefore strongly recommend validating the generated data in advanced mode.

7.2.1. Choosing a schema for validation

In the "Validate against" box you can choose a schema for validation. Depending on the selected mapping (see above) the respective default validation schema is automatically set.

Validate against:

<http://www.loc.gov/standards/vracore/vra-strict.xsd>

<http://www.loc.gov/standards/vracore/vra.xsd>

add Schema (.xsd .dtd):

For example, if you selected "VRA strict" as mapping, the corresponding schema <http://www.loc.gov/standards/vracore/vra-strict.xsd> will be set.

It is possible to choose different validation schemas. In a future version of the transform tool users may also upload their own schema.

7.2.2. Starting the validation

To start the validation process, click "validate".



transform XML

Validate against:
<http://www.loc.gov/standards/vracore/vra-strict.xsd>
<http://www.loc.gov/standards/vracore/vra.xsd>

add Schema (.xsd .dtd):

validate download

open full preview

Log

15:44:45: displaying Preview
15:44:45: Transformation successful: /db/tmp/e9ad0818-a2a8-4863-84da-64691915164c_transformed.xml
15:44:45: applying transformations
15:39:56: displaying Preview
15:39:56: Transformation successful: /db/tmp/e9ad0818-a2a8-4863-84da-64691915164c_transformed.xml
15:39:56: applying transformations
15:39:56: Processing CSV Lines done
15:39:56: line 44 processed.
15:39:55: line 43 processed.
15:39:55: line 42 processed.
15:39:54: line 41 processed.
15:39:53: line 40 processed.
15:39:53: line 39 processed.
15:39:52: line 38 processed.
15:39:51: line 37 processed.
15:39:51: line 36 processed.
15:39:50: line 35 processed.
15:39:49: line 34 processed.
15:39:48: line 33 processed.

7.2.3. Getting the validation result

Once the validation is finished, a new browser tab will be opened with the validation result. In addition, the button's color will change:



Result NOT VALID



Result VALID

Validate against:
<http://www.loc.gov/standards/vracore/vra-strict.xsd>
<http://www.loc.gov/standards/vracore/vra.xsd>

add Schema (.xsd .dtd):

validate download

open full preview

Log

15:50:47: validation completed: resulting xml is valid
15:50:45: validating... please be patient.
15:44:45: displaying Preview
15:44:45: Transformation successful: /db/tmp/e9ad0818-a2a8-4863-84da-64691915164c_transformed.xml
15:44:45: applying transformations
15:39:56: displaying Preview
15:39:56: Transformation successful: /db/tmp/e9ad0818-a2a8-4863-84da-64691915164c_transformed.xml
15:39:56: applying transformations
15:39:56: Processing CSV Lines done
15:39:56: line 44 processed.
15:39:55: line 43 processed.
15:39:55: line 42 processed.
15:39:54: line 41 processed.
15:39:53: line 40 processed.
15:39:53: line 39 processed.
15:39:52: line 38 processed.
15:39:51: line 37 processed.
15:39:51: line 36 processed.
15:39:50: line 35 processed.
15:39:49: line 34 processed.
15:39:48: line 33 processed.
15:39:47: line 32 processed.
15:39:46: line 31 processed.
15:39:46: line 30 processed.
15:39:45: line 29 processed.
15:39:44: line 28 processed.
15:39:44: line 27 processed.
15:39:44: line 26 processed.
15:39:43: line 25 processed.
15:39:42: line 24 processed.

1 / 80 ->

```
1. <work xmlns="http://www.vrweb.org/vracore4.htm" id="w_1">
2.   <agentSet>
3.     <display>unknown (French architect)</display>
4.     <agent>
5.       <name type="personal" vocab="ULAN" refid="500125235">unknown French</name>
6.       <role>architect</role>
7.     </agent>
8.   </agentSet>
9.   <contextContextSet>
10.    <display>French</display>
11.    <contextContext vocab="AAT" refid="300111188">French</contextContext>
12.  </contextContextSet>
13.  <dateSet>
14.    <display>begun 1194 (creation); consecrated 1260 (other)</display>
15.    <date type="creation">
16.      <earliestDate>1194</earliestDate>
17.      <latestDate>1194</latestDate>
18.    </date>
19.    <date type="other">
20.      <earliestDate>1260</earliestDate>
21.      <latestDate>1260</latestDate>
22.    </date>
```

8. If validation fails

VRA Core 4 in the restricted version uses controlled type lists and date formats. If your data does not match the prescribed rules, your XML will not validate.

In case of problems with the validation, the error messages will be displayed in a separate tab. The messages will indicate both line and column of the respective generated XML document. This is displayed in the interface.

Error messages usually come in pairs. The first line will name the invalid value based on an enumeration in the schema (e.g. type values defined by vra-strict). The second line will indicate the parent element.

```
<report>
  <report>
    <status>invalid</status>
    <namespace>http://www.vraweb.org/vracore4.htm</namespace>
    <duration unit="second">24</duration>
    <message level="Error" line="30" column="36">cvc-pattern-valid: Value 'early1' is not facet-valid with respect to pattern 'present{1}[0-9]{1,12}([0-9]{2}([0-9]{2})*)*' for type 'dateValueType'.</message>
    <message level="Error" line="30" column="36">cvc-complex-type.2.2: Element 'date' must have no element (children), and the value must be valid.</message>
  </report>
</report>
```

8.1. Example 1 (element type):

Error message (screenshot):

```
<message level="Error" line="134" column="52">cvc-enumeration-valid: Value 'illustrative' is not facet-valid with respect to enumeration '[brandName, cited, creator, descriptive, former, generalView, inscribed, owner, partialView, popular, repository, translated, other]'. It must be a value from the enumeration.</message>
<message level="Error" line="134" column="52">cvc-attribute.3: The value 'illustrative' of attribute 'type' on element 'title' is not valid with respect to its type, 'titleTypeType'.</message>
<message level="Error" line="134" column="52">cvc-enumeration-valid: Value 'illustrative' is not facet-valid with respect to enumeration '[brandName, cited, creator, descriptive, former, generalView, inscribed, owner, partialView, popular, repository, translated, other]'. It must be a value from the enumeration.</message>
```

Error message (transcribed):

```
<message level="Error" line="134" column="52">cvc-enumeration-valid: Value 'illustrative' is not facet-valid with respect to enumeration '[brandName, cited, creator, descriptive, former, generalView, inscribed, owner, partialView, popular, repository, translated, other]'. It must be a value from the enumeration.</message>
<message level="Error" line="134" column="52">cvc-attribute.3: The value 'illustrative' of attribute 'type' on element 'title' is not valid with respect to its type, 'titleTypeType'.</message>
```

The error message refers to `line="134" column="52"` in the XML:

```
131. </textrefSet>
132. <titleSet>
133. <display predefined work title display/>
134. <title pref="true" type="illustrative">some w_prefTitle</title>
135. <title pref="false" type="translated">some_title</title>
136. </titleSet>
137. <worktypeSet>
```

The first line of the message states that "illustrative" is not valid because it is not included in the list of possible values: "brandName, cited, creator, descriptive, former, generalView, inscribed, owner, partialView, popular, repository, translated, other".

The second line adds that the error occurred within attribute "type" of element <title>, and the provided value for "type" was not valid.

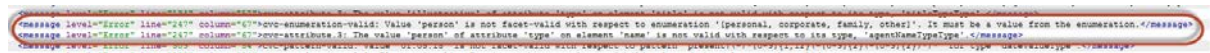
Result:

Values of <title type=""> are controlled by the schema.

You need to delete "illustrative" and use of one of the allowed values, i.e. "brandName, cited, creator, descriptive, former, generalView, inscribed, owner, partialView, popular, repository, translated, other".

8.2. Example 2 (element type):

Error message (screenshot):



Error message (transcribed):

```
<message level="Error" line="247" column="67">cvc-enumeration-valid: Value 'person'
is not facet-valid with respect to enumeration '[personal, corporate, family,
other]'. It must be a value from the enumeration.</message>
```

```
<message level="Error" line="247" column="67">cvc-attribute.3: The value 'person'
of attribute 'type' on element 'name' is not valid with respect to its type,
'agentNameTypeType'.</message>
```

The error message refers to `line="247" column="67"` in the XML:

```
244. <agentSet>
245. <display>school of Michelangelo Buonarroti (painter); formerly attributed to Paul, Priya (collector);
246. <agent>
247. <name type="person" vocab="viaf" refid="24585191">Michelangelo Buonarroti</name>
248. <role>painter</role>
249. <attribution>school of</attribution>
250. </agent>
251. <agent>
252. <name type="personal" vocab="viaf" refid="277848046">Paul, Priya</name>
```

The first line of the message states that "person" is not valid because it is not included in the list of possible values: "personal, corporate, family, other".

The second line adds that the error occurred within attribute "type" of element <name>, and the provided value for "type" was not valid.

Result:

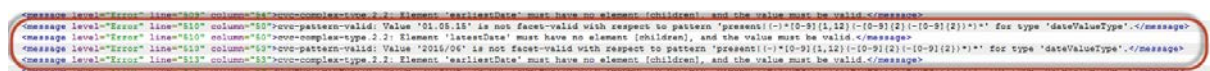
Values of <name type=""> are controlled by the schema.

You need to delete "person" and use of one of the allowed values, i.e. "personal, corporate, family, other".

8.3. Example 3 (date pattern):

This example consists of two similar problems:

Error message (screenshot):



Error message (transcribed):

```
<message level="Error" line="510" column="50">cvc-pattern-valid: Value '01.05.15'
is not facet-valid with respect to pattern 'present|(-)*[0-9]{1,12}(-[0-9]{2})(-[0-9]{2})*)*'
for type 'dateValueType'.</message>
```

```
<message level="Error" line="510" column="50">cvc-complex-type.2.2: Element
'latestDate' must have no element [children], and the value must be valid.</message>
```



```
<message level="Error" line="513" column="53">cvc-pattern-valid: Value '2015/06' is
not facet-valid with respect to pattern 'present|(-)*[0-9]{1,12}(-[0-9]{2}(-[0-
9]{2}))*' for type 'dateValueType'.</message>
<message level="Error" line="513" column="53">cvc-complex-type.2.2: Element
'earliestDate' must have no element [children], and the value must be
valid.</message>
```

The error message refers to `line="510" column="50"` and `line="513" column="53"` in the XML:

```
506.     <dateSet>
507.         <display>a predefined work date display</display>
508.         <date type="creation">
509.             <earliestDate>01.05.15</earliestDate>
510.             <latestDate>01.05.15</latestDate>
511.         </date>
512.         <date type="alteration">
513.             <earliestDate>2015/06</earliestDate>
514.             <latestDate>2015/06</latestDate>
515.         </date>
516.     </dateSet>
```

The first lines of each message state that “01.05.15” and “2015/06” are not valid because they do not follow the predefined pattern for dateValues.

The second lines add that the error occurred within element `<latestDate>` (respective `<earliestDate>`), and the element must be valid.

Result:

Values of `<earliestDate>` and `<latestDate>` are pattern-controlled by the schema.

You need to write the dates in the correct format, i.e. “2015-05-01” or “2015-06”.

8.4. Example 4 (work/image ID):

Error message (transcribed):

```
<message level="Error" line="239" column="18"> cvc-datatype-valid.1.2.1: '2' is not
a valid value for 'NCName'.</message>
<message level="Error" line="239" column="18"> cvc-attribute.3: The value '2' of
attribute 'id' on element 'work' is not valid with respect to its type,
'ID'.</message>
```

The error message refers to `line="239" column="18"` in the XML:

```
238. </image>
239.     <work id="2">
240.         <agentSet>
241.             <display>Carlo
```

The first line of the message states that “2” is not a valid NCName, which stands for “Non-colonized” Names (cf. <http://www.w3.org/TR/xmlschema-2/#NCName>).

The second line adds that the error occurred within attribute “id” of element `<work>`, and the provided value was not valid.

Result:



Values of `<work id="">` are regulated by the schema. The "[VRA Core4 Element Description](#)" explains in a note: "The XML *id* attribute must begin with a character. The convention used here prefixes a numeric value with w_ for works, c_ for collections, and i_ for images."

You need to provide a work ID that begins with "w_".



9. Generating RDF XML

In addition to generating VRA Core 4 XML it is possible to transform the data to RDF. At the moment, this feature is still experimental (only basic testing, no output optimisation, etc).

The tool makes use of the XSLT stylesheet as provided by the VRA-RDF-Project <https://github.com/mixerj/VRA-RDF-Project>.

To transform your data to RDF first make sure the .csv data will validate in VRA Core.

Then go back and select "transform into RDF" in the "transform" line. Note that the entry "xsl/VRA2RDF.xsl" appears in the "Applied XSL's" box.

Upload CSV
Browse... VRACoreSamples_2015-12-09.csv Upload

44 lines loaded
process lines 2 to 44
mapping: VRA strict
generate XML
transform preset Transform into RDF
Apply XSLs (for transformation):
xsl/cleanup-vra.xsl
xsl/VRA2RDF.xsl
(modifying the predefined selection may break pagination)
transform XML
Validate against:

Now click "generate XML".

Once generated, the RDF XML will be displayed in the "Result" box and can be downloaded.



1 / 80 ->

```
1. <rdf:Description xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" rdf:about="#w_1">
2.   <rdf:type rdf:resource="http://purl.org/vra/CreativeWork"></rdf:type>
3.   <rdf:type rdf:resource="http://vocab.getty.edu/aat/300007501"></rdf:type>
4.   <rdf:type rdf:resource="http://vocab.getty.edu/aat/300170443"></rdf:type>
5.   <rdf:type rdf:resource="http://vocab.getty.edu/aat/300007466"></rdf:type><!--Skipping over agentSet--><!--Skipping over display-
6.   <vra:creator xmlns:vra="http://purl.org/vra/">
7.     <rdf:Description rdf:about="http://vocab.getty.edu/ulan/500125235">
8.       <rdf:type rdf:resource="http://purl.org/vra/Agent"></rdf:type>
9.       <rdf:type rdf:resource="http://purl.org/vra/Person"></rdf:type>
10.      <vra:name>unknown French</vra:name>
11.    </rdf:Description>
12.  </vra:creator><!--Skipping over culturalContextSet--><!--Skipping over display--><!--Skipping over French-->
13.  <vra:culturalContext xmlns:vra="http://purl.org/vra/">
14.    <rdf:Description rdf:about="http://vocab.getty.edu/aat/French">
15.      <vra:name>French</vra:name>
16.    </rdf:Description>
17.  </vra:culturalContext>
18.  <vra:dateCreated xmlns:vra="http://purl.org/vra/">begun 1194 (creation); consecrated 1260 (other)</vra:dateCreated>
19.  <vra:wasCreated xmlns:vra="http://purl.org/vra/">
20.    <rdf:Description>
21.      <rdf:type rdf:resource="http://purl.org/vra/Event"></rdf:type>
22.      <vra:name>Creation</vra:name>
23.      <vra:startDate>1194</vra:startDate>
24.      <vra:endDate>1194</vra:endDate>
25.    </rdf:Description>
26.  </vra:wasCreated><!--Skipping over descriptionSet--><!--Skipping over display--><!--Skipping over The present cathedral was cons
27.  <vra:description xmlns:vra="http://purl.org/vra/">The present cathedral was constructed on the foundations of the earlier church
28.  <vra:placeOfSite xmlns:vra="http://purl.org/vra/">
29.    <rdf:Description>
30.      <rdf:type rdf:resource="http://purl.org/vra/Place"></rdf:type>
31.      <vra:containedIn>
32.        <rdf:Description rdf:about="http://vocab.getty.edu/tgn/7033010">
33.          <rdf:type rdf:resource="http://purl.org/vra/Place"></rdf:type>
34.          <vra:name>Chartres Cathedral</vra:name>
35.        </rdf:Description>
36.      </vra:containedIn>
37.      <vra:containedIn>
38.        <rdf:Description rdf:about="http://vocab.getty.edu/tgn/7008267">
39.          <rdf:type rdf:resource="http://purl.org/vra/Place"></rdf:type>
40.          <vra:name>Chartres</vra:name>
41.        </rdf:Description>
42.      </vra:containedIn>
43.      <vra:name>Chartres, Centre, France</vra:name>
44.      <vra:description>Eure-et-Loir (department)</vra:description>
45.    </rdf:Description>
```

10. Contact

If you have problems or questions please contact Matthias Arnold at arnold@asia-europe.uni-heidelberg.de.

Contact details

arnold@asia-europe.uni-heidelberg.de

Phone: +49 (0) 6221 - 54 4094

Fax: +49 (0) 6221 - 54 4012

Skype: matz-skype

Office address

Karl Jaspers Centre, Room 005b

MediaLab

Karl Jaspers Centre, Room 005c

Further links

HRA Portal

<http://hra.uni-hd.de/>

MediaLab Tutorials in Sharepoint

<https://sharepoint.urz.uni-heidelberg.de/vjc/kjc/hra/medialab/>

11. Appendix: Elements, displays and repetitions

WORK

[workID]

Agent display

3x nameType – name – nameVocab – nameRefid – role

CulturalContext display

2x culturalContext – vocab – refid

Date display

2x type – earliestDate – latestDate

Description display [display = description]

descriptionSource

Inscription display [display = text]

Location display

Location notes

2x type – Name – NameType – NameVocab – NameRefID – ObjType – ObjRefID – Geo –
GeoVocab – GeoRefid

Material display

4x material – vocab – refid

Measurements display

2x extent – unit – value1-4 – type1-4

Relation display

2x relatedWork – type

[transform adds work-image link]

Rights display [display = text]

Source display

StateEdition display

Style/Period display

4x stylePeriod – vocab – refid

Subject display

8x subject – vocab – refid – type

Technique display

4x technique – vocab – refid

TextRef display

Title display

titlePref, titlePrefType, titleAlt, titleAltType

Worktype display

3x worktype – vocab – refid

IMAGE

[IMAGE_Href (-> href); Filename (-> refid)]

Agent display

1x nameType – name – nameVocab – nameRefid – role

Date display

Description display [display = description]

Inscription display [display = text]

Location display

[localRepoName, accessionID]

Material display

Measurements display

Relation display

[transform adds image-work link]

Rights display

Rights notes

1x type – holder – text

Source display

[VendorName, VendorID]

1x sourceType, sourceValue

Subject display

3x subject – vocab – refid – type

Technique display

1x technique – vocab – refid

[if not provided: <technique vocab="AAT" refid="300237903">digital imaging</technique>]

Title display

1x type, title

WorkType display

1x technique – vocab – refid

[if not provided: <worktype vocab="AAT" refid="300215302">digital images</worktype>]

12. Appendix: Columns Full Template

IMAGE_Filename	WORK_Location1Geo
IMAGE_Href	WORK_Location1GeoVocab
IMAGE_AccessionID	WORK_Location1GeoRefid
IMAGE_VendorID	WORK_Location2Type
IMAGE_VendorName	WORK_Location2Name
IMAGE_LocalRepoName	WORK_Location2NameType
WORK_ID	WORK_Location2NameVocab
WORK_TitleDisplay	WORK_Location2NameRefID
WORK_TitlePref	WORK_Location2ObjType
WORK_TitlePrefType	WORK_Location2ObjRefID
WORK_TitleAlt	WORK_Location2Geo
WORK_TitleAltType	WORK_Location2GeoVocab
WORK_AgentDisplay	WORK_Location2GeoRefid
WORK_Agent1NameType	WORK_MaterialDisplay
WORK_Agent1Name	WORK_Material1
WORK_Agent1NameVocab	WORK_MaterialVocab1
WORK_Agent1NameRefid	WORK_MaterialRefid1
WORK_Agent1Role	WORK_Material2
WORK_Agent1Attribution	WORK_MaterialVocab2
WORK_Agent2NameType	WORK_MaterialRefid2
WORK_Agent2Name	WORK_Material3
WORK_Agent2NameVocab	WORK_MaterialVocab3
WORK_Agent2NameRefid	WORK_MaterialRefid3
WORK_Agent2Role	WORK_Material4
WORK_Agent2Attribution	WORK_MaterialVocab4
WORK_Agent3NameType	WORK_MaterialRefid4
WORK_Agent3Name	WORK_MeasurementsDisplay
WORK_Agent3NameVocab	WORK_Measurements1Extent
WORK_Agent3NameRefid	WORK_Measurements1Unit
WORK_Agent3Role	WORK_Measurements1Value1
WORK_Agent3Attribution	WORK_Measurements1Type1
WORK_CulturalContextDisplay	WORK_Measurements1Value2
WORK_CulturalContext1	WORK_Measurements1Type2
WORK_CulturalContext1Vocab	WORK_Measurements1Value3
WORK_CulturalContext1Refid	WORK_Measurements1Type3
WORK_CulturalContext2	WORK_Measurements1Value4
WORK_CulturalContext2Vocab	WORK_Measurements1Type4
WORK_CulturalContext2Refid	WORK_Measurements2Extent
WORK_DateDisplay	WORK_Measurements2Unit
WORK_Date1Type	WORK_Measurements2Value1
WORK_EarliestDate1	WORK_Measurements2Type1
WORK_LatestDate1	WORK_Measurements2Value2
WORK_Date2Type	WORK_Measurements2Type2
WORK_EarliestDate2	WORK_Measurements2Value3
WORK_LatestDate2	WORK_Measurements2Type3
WORK_DescriptionDisplay	WORK_Measurements2Value4
WORK_DescriptionSource	WORK_Measurements2Type4
WORK_InscriptionDisplay	WORK_RelationDisplay
WORK_LocationDisplay	WORK_RelatedWork1
WORK_LocationNotes	WORK_RelationType1
WORK_Location1Type	WORK_RelatedWork2
WORK_Location1Name	WORK_RelationType2
WORK_Location1NameType	WORK_RightsDisplay
WORK_Location1NameVocab	WORK_SourceDisplay
WORK_Location1NameRefID	WORK_StateEditionDisplay
WORK_Location1ObjType	WORK_StylePeriodDisplay
WORK_Location1ObjRefID	WORK_StylePeriod1



WORK_StylePeriodVocab1	WORK_TextrefNameType
WORK_StylePeriodRefid1	WORK_TextrefRefid
WORK_StylePeriod2	WORK_TextrefRefidType
WORK_StylePeriodVocab2	WORK_WorktypeDisplay
WORK_StylePeriodRefid2	WORK_Worktype1
WORK_StylePeriod3	WORK_WorktypeVocab1
WORK_StylePeriodVocab3	WORK_WorktypeRefid1
WORK_StylePeriodRefid3	WORK_Worktype2
WORK_StylePeriod4	WORK_WorktypeVocab2
WORK_StylePeriodVocab4	WORK_WorktypeRefid2
WORK_StylePeriodRefid4	WORK_Worktype3
WORK_SubjectDisplay	WORK_WorktypeVocab3
WORK_Subject1	WORK_WorktypeRefid3
WORK_SubjectType1	IMAGE_AgentDisplay
WORK_SubjectVocab1	IMAGE_Agent1NameType
WORK_SubjectRefid1	IMAGE_Agent1Name
WORK_Subject2	IMAGE_Agent1NameVocab
WORK_SubjectType2	IMAGE_Agent1NameRefid
WORK_SubjectVocab2	IMAGE_Agent1Role
WORK_SubjectRefid2	IMAGE_DateDisplay
WORK_Subject3	IMAGE_DescriptionDisplay
WORK_SubjectType3	IMAGE_InscriptionDisplay
WORK_SubjectVocab3	IMAGE_LocationDisplay
WORK_SubjectRefid3	IMAGE_MaterialDisplay
WORK_Subject4	IMAGE_MeasurementsDisplay
WORK_SubjectType4	IMAGE_RelationDisplay
WORK_SubjectVocab4	IMAGE_RightsDisplay
WORK_SubjectRefid4	IMAGE_RightsNotes
WORK_Subject5	IMAGE_RightsType
WORK_SubjectType5	IMAGE_RightsHolder
WORK_SubjectVocab5	IMAGE_RightsText
WORK_SubjectRefid5	IMAGE_SourceDisplay
WORK_Subject6	IMAGE_SourceValue
WORK_SubjectType6	IMAGE_SourceType
WORK_SubjectVocab6	IMAGE_SubjectDisplay
WORK_SubjectRefid6	IMAGE_Subject1
WORK_Subject7	IMAGE_SubjectType1
WORK_SubjectType7	IMAGE_SubjectVocab1
WORK_SubjectVocab7	IMAGE_SubjectRefid1
WORK_SubjectRefid7	IMAGE_Subject2
WORK_Subject8	IMAGE_SubjectType2
WORK_SubjectType8	IMAGE_SubjectVocab2
WORK_SubjectVocab8	IMAGE_SubjectRefid2
WORK_SubjectRefid8	IMAGE_Subject3
WORK_TechniqueDisplay	IMAGE_SubjectType3
WORK_Technique1	IMAGE_SubjectVocab3
WORK_TechniqueVocab1	IMAGE_SubjectRefid3
WORK_TechniqueRefid1	IMAGE_TechniqueDisplay
WORK_Technique2	IMAGE_Technique
WORK_TechniqueVocab2	IMAGE_TechniqueVocab
WORK_TechniqueRefid2	IMAGE_TechniqueRefid
WORK_Technique3	IMAGE_TitleDisplay
WORK_TechniqueVocab3	IMAGE_Title
WORK_TechniqueRefid3	IMAGE_TitleType
WORK_Technique4	IMAGE_WorktypeDisplay
WORK_TechniqueVocab4	IMAGE_Worktype
WORK_TechniqueRefid4	IMAGE_WorktypeVocab
WORK_TextrefDisplay	IMAGE_WorktypeRefid
WORK_TextrefName	