

Implementing the **CDISC Library API** in software applications: first experiences



Jozef Aerts
XML4Pharma

Phuse Connect Amsterdam

November 2019

What is the CDISC Library API ?

- It is a huge metadata repository (MDR) ultimately containing almost all of the CDISC standards
- Machine-readable
- Queryable using RESTful web services

What are RESTful Web Services ?

What is an API ?

- RESTful web services are services running over the internet (usually HTTP/HTTPS)
- GET / POST / PUT / POST / DELETE
- Return structured information formatted as XML, JSON, RDF-Turtle, ...
- Can usually be tested in the browser, but are meant to be used by applications
- The API describes the available methods and how to run them

Example in browser

- <https://library.cdisc.org/api/mdr/sdtmig/3-3/datasets/DM/variables/RFSTDTC>

```
<datasetVariable>
  <ordinal>5</ordinal>
  <name>RFSTDTC</name>
  <label>Subject Reference Start Date/Time</label>
  <description>
    Reference Start Date/time for the subject in ISO 8601 character format. Usually equivalent to date/time when subject was first
    exposed to study treatment. See Assumption 9 for additional detail on when RFSTDTC may be null.
  </description>
  <role>Record Qualifier</role>
  <simpleDatatype>Char</simpleDatatype>
  <core>Exp</core>
  <describedValueDomain>ISO 8601</describedValueDomain>
  <_links>
    <self>
      <href>/mdr/sdtmig/3-3/datasets/DM/variables/RFSTDTC</href>
      <title>Subject Reference Start Date/Time</title>
      <type>SDTM Dataset Variable</type>
    </self>
    <modelDatasetVariable>
      <href>/mdr/sdtm/1-7/datasets/DM/variables/RFSTDTC</href>
      <title>Subject Reference Start Date/Time</title>
      <type>SDTM Dataset Variable</type>
    </modelDatasetVariable>
  </_links>
  <parentProduct>
    <href>/mdr/sdtmig/3-3</href>
  </parentProduct>
</datasetVariable>
```

HATEAOS principles

- Response contains links to parent and children and related information

Example: LB - **LBTESTCD** (SDTM-IG 3.3)

```
<self>  
  <href>/mdr/sdtmig/3-3/datasets/LB/variables/LBTESTCD</href>  
  <title>Lab Test or Examination Short Name.</title>  
  <type>SDTM Dataset Variable</type>  
</self>  
<codelist>  
  <href>/mdr/root/ct/sdtmct/codelists/C65047</href>  
  <title>  
    Version-agnostic anchor resource for codelist C65047  
  </title>  
  <type>Root Value Domain</type>  
</codelist>
```

```
<priorVersion>  
  <href>/mdr/sdtmig/3-2/datasets/LB/variables/LBTESTCD</href>  
  <title>Lab Test or Examination Short Name</title>  
  <type>SDTM Dataset Variable</type>  
</priorVersion>
```

HATEAOS principles

LBTESTCD Codelist

- CodeLists have ... versions

[illegible]

HATEAOS principles

LBTESTCD Codelist 2018-12-21

- CodeLists versions have ... terms

```
<term>
  <conceptId>C100461</conceptId>
  <submissionValue>A1MICG</submissionValue>
  <definition>
    A measurement of the alpha-1 microglobulin in a biological specimen.
  </definition>
  <preferredTerm>Alpha-1 Microglobulin Measurement</preferredTerm>
  <synonyms>Alpha-1 Microglobulin; Protein HC</synonyms>
</term>
```

- And terms have a
 - Submission value
 - Definition
 - Preferred term
 - Synonyms ...

```
<codelist>
  <conceptId>C65047</conceptId>
  <extensible>true</extensible>
  <name>Laboratory Test Code</name>
  <submissionValue>LBTESTCD</submissionValue>
  <definition></definition>
  <preferredTerm>CDISC SDTM Laboratory Test Code Terminology</preferredTerm>
  <synonyms>Laboratory Test Code</synonyms>
  <_links>
    <self></self>
    <parentPackage></parentPackage>
    <rootItem></rootItem>
    <priorVersion></priorVersion>
  </_links>
</codelist>

<term>
  <conceptId>C100429</conceptId>
  <submissionValue>A1AGLP</submissionValue>
  <definition>
    A measurement of the alpha-1 acid glycoprotein in a biological specimen.
  </definition>
  <preferredTerm>Alpha-1 Acid Glycoprotein Measurement</preferredTerm>
  <synonyms>Alpha-1 Acid Glycoprotein</synonyms>
</term>

<term>
  <conceptId>C80167</conceptId>
  <submissionValue>A1ANTRYP</submissionValue>
  <definition>
    A measurement of the alpha-1 antitrypsin in a biological specimen.
  </definition>
  <preferredTerm>Alpha-1 Antitrypsin Measurement</preferredTerm>
```

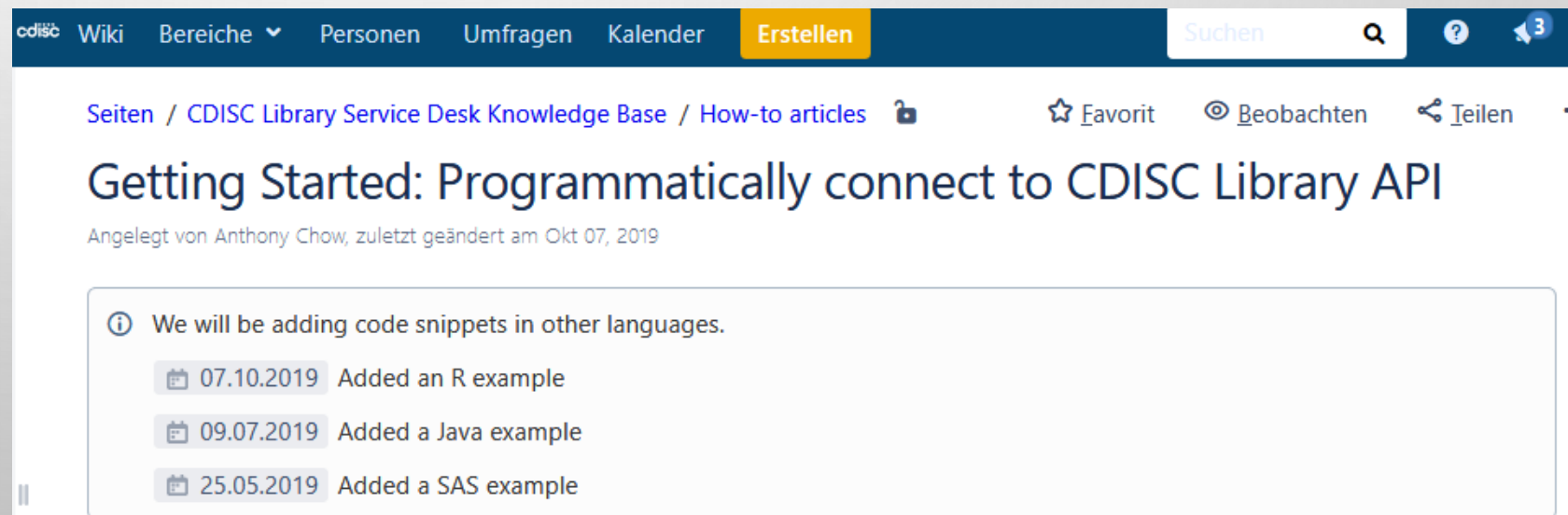
What is currently covered ?

- Controlled Terminology
- CDASH
- CDASH-IG
- SDTM
- SDTM-IG
- SEND-IG
- ADaM

Coming: TAUGs (Therapeutic Area User Guides), Code tables (ValueLists), Questionnaires, Conformance rules ...

Implementation in Software

- RESTful web services are extremely easy to implement in modern software
- "How to" examples for different computer languages on the Wiki:
<https://wiki.cdisc.org/display/LIBSUPRT/Getting+Started%3A+Programmatically+connect+to+CDISC+Library+API>
 - Python
 - SAS & R
 - Java
 - Coming: XSLT, XQuery



The screenshot shows the CDISC Wiki interface. The top navigation bar includes links for Wiki, Bereiche, Personen, Umfragen, Kalender, and a yellow 'Erstellen' button. A search bar with the text 'Suchen' and a magnifying glass icon is on the right. Below the navigation bar, the breadcrumb trail reads 'Seiten / CDISC Library Service Desk Knowledge Base / How-to articles'. The main title of the page is 'Getting Started: Programmatically connect to CDISC Library API', with a subtitle 'Angelegt von Anthony Chow, zuletzt geändert am Okt 07, 2019'. A blue information box contains the text 'We will be adding code snippets in other languages.' followed by a list of updates: '07.10.2019 Added an R example', '09.07.2019 Added a Java example', and '25.05.2019 Added a SAS example'. On the right side of the page, there are icons for 'Favorit', 'Beobachten', and 'Teilen'.

Implementation in Software: Java (Using Jersey-2 libraries)

- CDISC Library requires "Basic Authentication"
 - This needs to be taken into account in applications

```
private String base = "https://library.cdisc.org/api"; // the base of the service
private String username = "my.name@mycompany.com"; // your user name
private String pass = "xxxxxxxxxxx"; // your passkey

private Client client;

public BasicCDISCLibraryClient() {
    ClientConfig clientConfig = new ClientConfig();
    // see e.g. https://jersey.github.io/documentation/latest/logging_chapter.html for how
    //clientConfig.property(LoggingFeature.LOGGING_FEATURE_VERBOSITY_CLIENT, LoggingFeatur
    client = ClientBuilder.newClient(clientConfig);
    // Basic Authentication
    HttpAuthenticationFeature feature = HttpAuthenticationFeature.basic(username,pass);
    client.register(feature);
}
```

Implementation in Software: Java (Using Jersey-2 libraries)

- Example: get SDTM Variable Properties, with parameters "IG-version", "domain", "variable"

```
// Get the properties of the provided SDTM-IG variable in the provided domain and for the provided IG version
public String getSDTMVariableProperties(String igVersion, String domain, String variable) {
    WebTarget webTarget = client.target(base).path("mdr").path("sdtmig").path(igVersion).path("datasets").path(domain).path("variables").path(variable);
    Invocation.Builder invocationBuilder = webTarget.request(MediaType.APPLICATION_XML);
    // Response requires javax.annotation-api-1.2.jar, hk2-api-2.5.0-b42.jar, hk2-locator-2.5.0-b42.jar, hk2-utils-2.5.0-b42.jar
    // and javax.inject-1.jar, javax.inject-2.5.0-b42.jar
    // and jersey-media-json-binding.jar
    // from the Jersey-2 library - see https://jersey.github.io/
    Response response = invocationBuilder.get();
    System.out.println("HTTP Response Status = " + response.getStatus());
    String xml = response.readEntity(String.class);
    //System.out.println(xml);
    return xml;
}
```

Implementation in Open Source Software:

Java example: "Smart Submission Dataset Viewer"

VSTESTCD	VSTEST	VSPOS	VSORRES	VSORRESU	VSSTRESC	VSSTRESN	VSSTRESU
DIABP	Diastolic Bl...	SUPINE	64	mmHg	64	64	mmHg
DIABP	Diastolic Bl...	STANDING	83	mmHg			
DIABP	Diastolic Bl...	STANDING	57	mmHg			
DIABP	Diastolic Bl...	SUPINE	68	mmHg			
DIABP	Diastolic Bl...	STANDING	59	mmHg			
DIABP	Diastolic Bl...	STANDING	71	mmHg			
DIABP	Diastolic Bl...	SUPINE	56	mmHg			
DIABP	Diastolic Bl...	STANDING	51	mmHg			
DIABP	Diastolic Bl...	STANDING	61	mmHg			
DIABP	Diastolic Bl...	SUPINE	67	mmHg			
DIABP	Diastolic Bl...	STANDING	61	mmHg			
DIABP	Diastolic Bl...	STANDING	65	mmHg			
DIABP	Diastolic Bl...	SUPINE	56	mmHg			
DIABP	Diastolic Bl...	STANDING	50	mmHg			
DIABP	Diastolic Bl...	STANDING	54	mmHg			
DIABP	Diastolic Bl...	SUPINE	64	mmHg			
DIABP	Diastolic Bl...	STANDING	55	mmHg			
DIABP	Diastolic Bl...	STANDING	53	mmHg			
DIABP	Diastolic Bl...	SUPINE	72	mmHg	72	72	mmHg
DIABP	Diastolic Bl...	STANDING	61	mmHg	61	61	mmHg
DIABP	Diastolic Bl...	STANDING	53	mmHg	53	53	mmHg

Define-XML metadata:

Name: VSORRESU

Label: Original Units

Mandatory: No

Datatype: text

Length: 200

CodeList: VSUNIT

CodeList NCI Code: C71620

CDISC Library information:

Name: VSORRESU

Label: Original Units

Core: Exp

XPT simple data type: Char

CodeList NCI Code: C66770

Discrepancies with CDISC Library:

- Incorrect CDISC/NCI codelist assigned

Expected: C66770

[UNIT] codelist was assigned instead of [VSRESU] codelist

Implementation in the "Smart Submission DatasetViewer"

Is the following OK?

VSTESTCD	VSTEST	VSPOS	VSORRES	VSORRESU
DIABP	Diastolic BI...	SUPINE	67	mmHg
DIABP	Diastolic BI...	STANDING	61	mmHg
DIABP	Diastolic BI...	RUNNING	65	mmHg
DIABP	Diastolic BI...	SUPINE	56	mmHg

Depends on what
the define.xml says ...

**"RUNNING" was not
defined as an
"extended value"
in the define.xml**

VSPOS	VSORRES	VSORRESU	VSSTRESC	VSSTRESN	VSST
SUPINE	67	mmHg	67	67	mmHg
STANDING					
RUNNING					
SUPINE					
STANDING					
STANDING					
SUPINE					
STANDING					
STANDING					
SUPINE					
STANDING					
STANDING					
SUPINE					
STANDING					
STANDING					
SUPINE					
STANDING					
STANDING					
SUPINE					
STANDING					

Define-XML metadata:
Name: VSPOS
Label: Vital Signs Position of Subject
Mandatory: No
Datatype: text
Length: 8
CodeList: VSPOS
CodeList NCI Code: C71148
CDISC Library information:
Name: VSPOS
Label: Vital Signs Position of Subject
Core: Perm
XPT simple data type: Char
CodeList NCI Code: C71148
Discrepancies with CDISC Library:
- Define.xml codeList C71148 contains the value:
[RUNNING]
that could not be found in the CDISC Library CodeList
and that was not defined as an extended value in the define.xml

Implementation in the "Smart Submission DatasetViewer"

File Tools Search Options

DM VS

STUDYID	DOMAIN	USUBJID	VSSEQ	VSTESTCD	VSTEST	VSPOS	VSORRES	VSORRESU	VSSTRESC	VSSTRESN	VSSTRESU	VSSTAT	VSLOC	VSBLFL	VISITNUM	VIS
CDISCIL...	VS	01-701-1015	1	DIABP	Diastolic Bl...	SUPINE	64	mmHg	64	64	mmHg				1	SCREE
CDISCIL...	VS	01-701-1015	2	DIABP	Diastolic Bl...	STANDING	83	mmHg	83	83	mmHg				1	SCREE
							57	mmHg	57	57	mmHg				1	SCREE
							68	mmHg	68	68	mmHg				2	SCREE
							59	mmHg	59	59	mmHg				2	SCREE
							71	mmHg	71	71	mmHg				2	SCREE
							56	mmHg	56	56	mmHg		Y		3	BASEL
							51	mmHg	51	51	mmHg		Y		3	BASEL

CDISC Library CodeList Details

CDISC CodeList Details for CodeList with NCI code C66770

Short Name: VSRESU

Long Name: Units for Vital Signs Results

Extensible: true

Definition: The unit used to record and describe the result of a test investigating a vital si

Preferred term: CDISC SDTM Unit for Vital Sign Result Terminology

Synonyms: Units for Vital Signs Results

CodeList values:

NCI Code	Submission Value	Definition	Preferred Term	Sy
C25613	%	One hundred times the quotient of one quantity divided by another, with the same units of measurement.	Percentage	Pe
		A unit of temperature of the temperature scale designed so that the freezing point of water is 0 degrees and the boiling point is 100 degrees at standard atmospheric pressure. The current official definition of the Celsius sets 0.01		

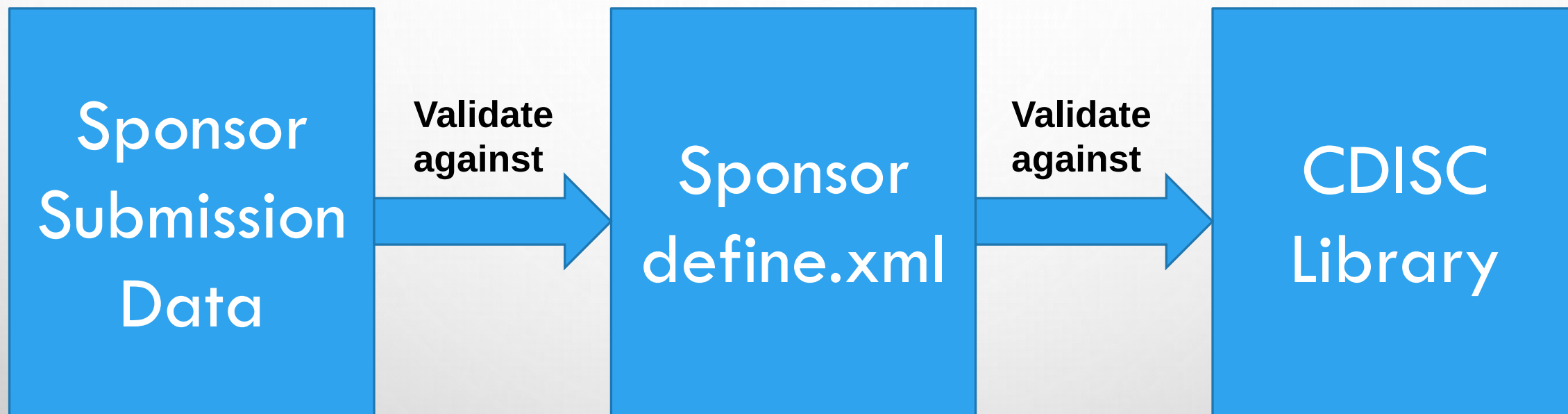
CDISC Library Discrepancies Report

Discrepancies with CDISC Library

Dataset	Domain	Variable	Description	Expected Value
DM	DM	DTHFL	No CDISC/NCI codelist assigned	C66742
DM	DM	ARMCD	Incorrect casing of Label 'Planned arm code'	Planned Arm Code
DM	DM	DMDTC	Incorrect Label 'Date of Collection'	Date/Time of Collection
VS	VS	VSTESTCD	No CDISC/NCI codelist assigned	C66741
VS	VS	VSTEST	No CDISC/NCI codelist assigned	C67153
		VSPOS	Define.xml codeList C71148 contains the value: [RUNNING] that could not be found in the CDISC Library CodeList and that was not defined as an extended value in the define.xml	
VS	VS	VSORRESU	Incorrect CDISC/NCI codelist assigned: C71620	C66770
		VSORRESU	Define.xml codeList C71620 contains values: [BEATS/MIN, IN] that could not be found in the CDISC Library CodeList and that were not defined as extended values in the define.xml	

CDISCIL... VS 01-701-1015 40 DIABP Diastolic Bl... SUPINE 61

How validation SHOULD work



**The sponsor's
metadata truth**

**The CDISC
metadata truth**

Implementation in Software: SAS

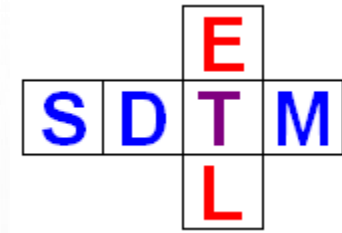
ct_products_full.sas

```
1  /*
2     Example: A GET Request to the CDISC Library That Uses the HEADERS Statement
3     Adjusted by: Marcelina Hungria
4     Date: 2019-04-03
5     Notes: Requires SAS 9.4M3 or later
6  */
7
8  filename response TEMP;
9
10 proc http
11     url="https://library.cdisc.org/api/mdr/products/Terminology"
12     AUTH_BASIC WEBUSERNAME="user" WEBPASSWORD="pass"
13     out=response;
14     headers
15         /* Use "application/xml" instead of "application/json" to get the response in xml format */
16         "Accept"="application/json";
17 run;
18
19 data _null_;
20     infile response;
21     input;
22     put _infile_;
23 run;
```

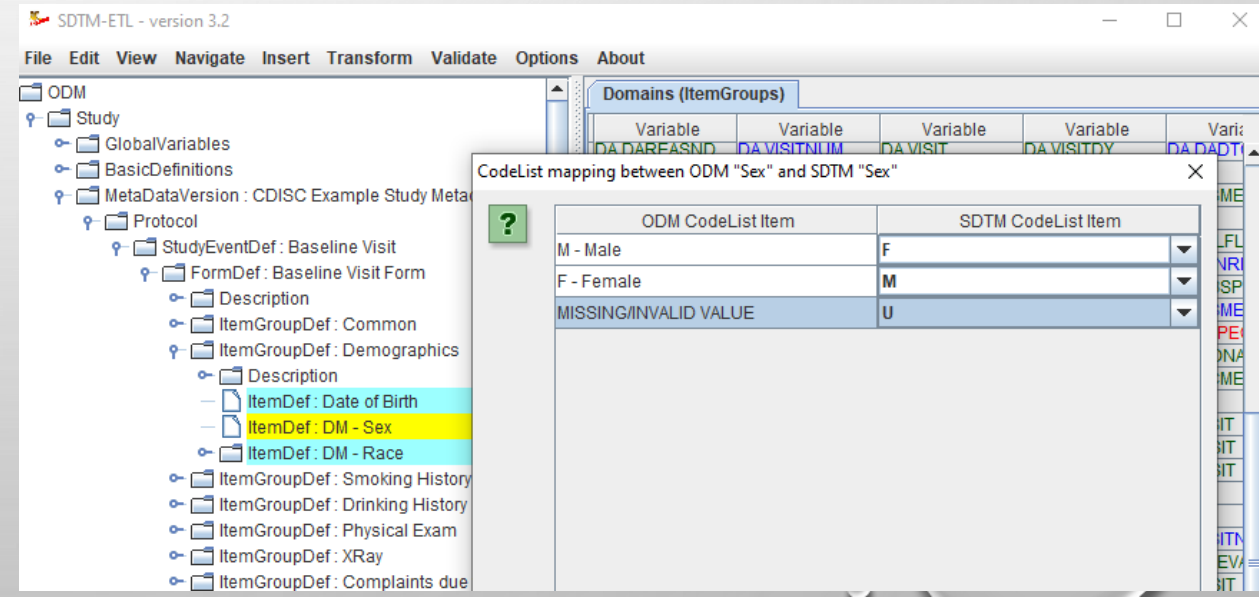
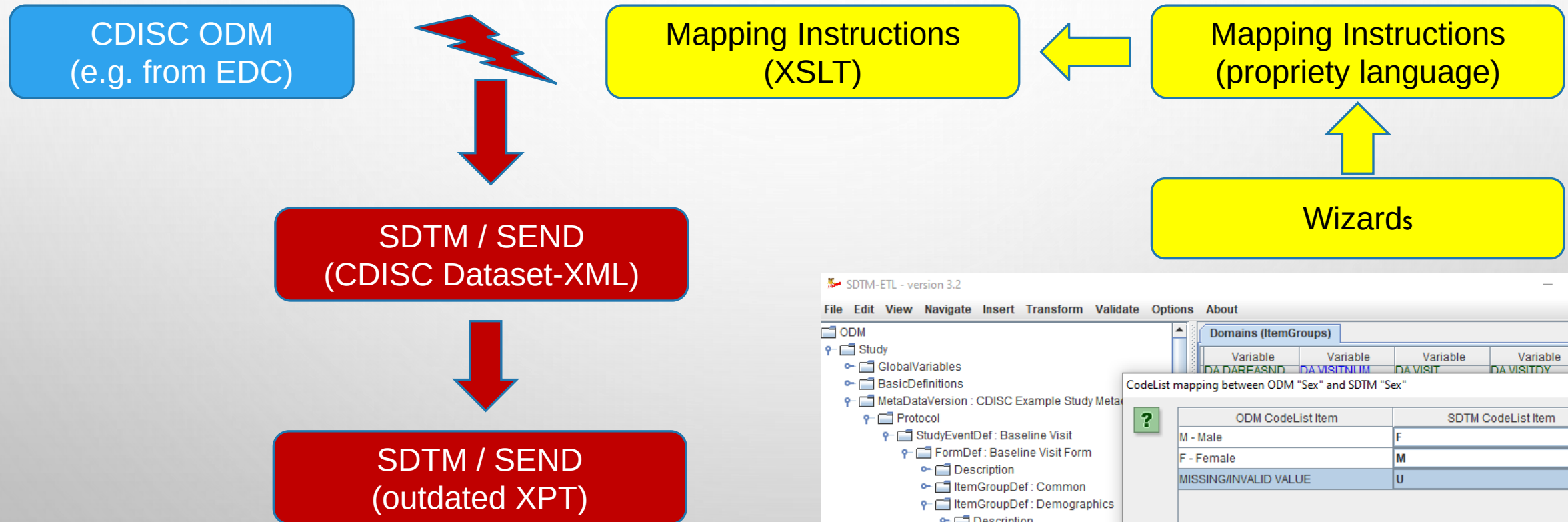
"Search" function in CDISC Library API

- New is a "search" function in the CDISC Library API
 - <https://wiki.cdisc.org/display/LIBSUPRT/Search>
- This functionality is extremely powerful, but a bit more complicated to use
- Example: "Search for LBTESTCD" in the scope of the "LB" domain and the IG-version "3.3":
<https://library.cdisc.org/api/mdr/search?q=LBTESTCD&domain=LB&product=SDTMIG%20v3.3>
- Currently only a JSON implementation is available

Implementation in XSLT



- Mapping Software "SDTM-ETL" uses XSLT to execute the mappings



Implementation in Software: XSLT/XQuery

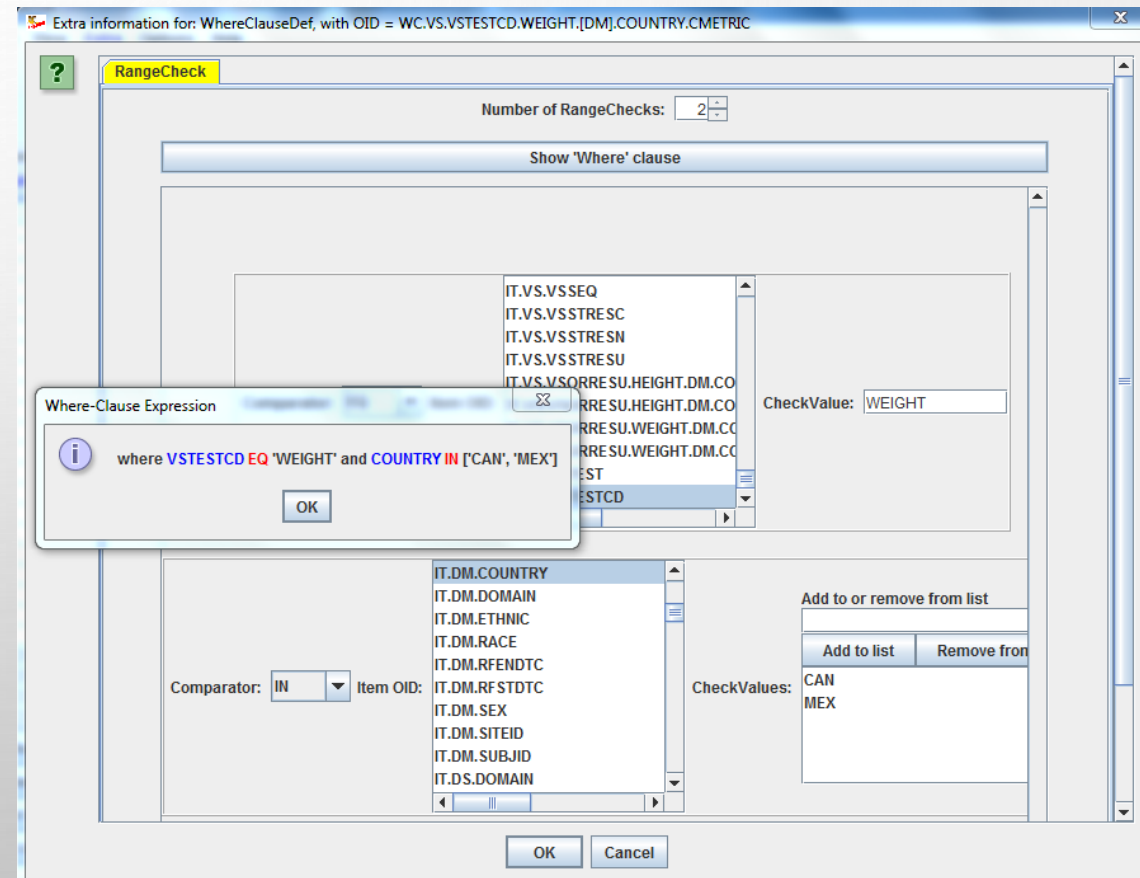
- Problem:
- XSLT/XQuery very well supports RESTful web services
- **But** does not well support "Basic Authentication"
- Most used Java-Library for XSLT "Saxon" however allows to write XSLT extensions so that "Basic Authentication" can be applied
- We are currently implementing this into the "SDTM-ETL" mapping software

Planned further Implementations

- The "**Open Rules for CDISC Standards**" initiative has all CDISC / FDA / PMDA validation rules in a machine- AND human-readable format, and can be executed by ANY modern software
- "**Open Rules for CDISC Standards**" already uses RESTful Web Services for LOINC, UNII, SNOMED-CT and MedDRA lookups
- As **the CDISC Library is "the CDISC single source of truth"**, we will implement the RESTful web services in the "**Open Rules for CDISC Standards**"

Planned further Implementations

- The "**Define-XML Designer**" is an extremely user-friendly software tool to develop and generate define.xml files for SDTM, SEND and ADaM submissions
- The define.xml is "**the sponsor's truth**" about the submission
- But this "truth" also has to comply with CDISC / FDA / PMDA "truth"
 - CDISC "truth": use **CDISC Library API**
 - FDA/PMDA "truth": use "**Open Rules for CDISC Standards**"



CDISC Library API

Summary of first experiences

- Very easy to use and implement in software
- HATEAOS makes "chaining" very easy
- The API is very well documented
- Support over "product inquiry" website is excellent
 - Usually good answers within hours
- Greatest advantages will become eminent with new versions of Standards, Implementation Guides and Controlled Terminology
 - No more "copy and paste" from PDFs or website pages
 - No more futzing around with Excel files

The CDISC Library
The CDISC Single Source of Truth

