Out-of-the-box %defineXML
Just a Simple SAS Macro
PhUSE / October 2016 / Katja Glaß
Agenda

- Introduction
- Getting Started
- %DefineXML
- Collaborate
- Summary
“Define-XML transmits metadata for SDTM, SEND and ADaM datasets; … datasets, variables, controlled terms, and other specified metadata …”

CDISC – Define-XML definition
# Define.xml Web View

## Tabulation Datasets for Study CDISC01 (SDTM-IG 3.1.2)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Description</th>
<th>Class</th>
<th>Structure</th>
<th>Purpose</th>
<th>Keys</th>
<th>Location</th>
<th>Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TA</td>
<td>Trial Arms</td>
<td>TRIAL DESIGN</td>
<td>One record per planned Element per Arm</td>
<td>Tabulation</td>
<td>STUDYID, ARMCD, TAETORD</td>
<td>ta.xpt</td>
<td></td>
</tr>
<tr>
<td>TE</td>
<td>Trial Elements</td>
<td>TRIAL DESIGN</td>
<td>One record per planned Element</td>
<td>Tabulation</td>
<td>STUDYID, ETCD</td>
<td>te.xpt</td>
<td></td>
</tr>
<tr>
<td>TI</td>
<td>Trial Inclusion/Exclusion Criteria</td>
<td>TRIAL DESIGN</td>
<td>One record per I/E criterion</td>
<td>Tabulation</td>
<td>STUDYID, ITESTECD</td>
<td>ti.xpt</td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>Trial Summary</td>
<td>TRIAL DESIGN</td>
<td>One record per trial summary parameter value</td>
<td>Tabulation</td>
<td>STUDYID, TSPARMCD, TSSEQ</td>
<td>ts.xpt</td>
<td></td>
</tr>
<tr>
<td>TV</td>
<td>Trial Visits</td>
<td>TRIAL DESIGN</td>
<td>One record per planned Visit per Arm</td>
<td>Tabulation</td>
<td>STUDYID, VISITNUM, ARMCD</td>
<td>tv.xpt</td>
<td></td>
</tr>
<tr>
<td>DM</td>
<td>Demographics</td>
<td>SPECIAL PURPOSE</td>
<td>One record per subject</td>
<td>Tabulation</td>
<td>STUDYID, USUBJID</td>
<td>dm.xpt</td>
<td>See Reviewer’s Guide, Section 2.1 Demographics Reviewers Guide</td>
</tr>
</tbody>
</table>
<Study OID="cdisc01">
  <GlobalVariables>
    <StudyName>CDISC01</StudyName>
    <StudyDescription>CDISC Test Study</StudyDescription>
    <ProtocolName>CDISC01</ProtocolName>
  </GlobalVariables>
  <MetaDataVersion OID="MDV.CDISC01.SDTMIG.3.1.2.SDTM.1.2"
    Name="Study CDISC01, Data Definitions"
    Description="Study CDISC01, Data Definitions"
    def:DefineVersion="2.0.0"
    def:StandardName="SDTM-IG"
    def:StandardVersion="3.1.2">

    <!-- ******************* -->
    <!-- Supporting Documents -->
    <!-- ******************* -->

    <def:AnnotatedCRF>
      <def:DocumentRef leafID="LF.blankcrf"/>
    </def:AnnotatedCRF>

    <def:SupplementalDoc>
      <def:DocumentRef leafID="LF.ReviewersGuide"/>
      <def:DocumentRef leafID="LF.ComplexAlgorithms"/>
    </def:SupplementalDoc>
  </MetaDataVersion>
</Study>
Getting Started

Requirement:

create a submission ready define.xml file

How:

• various tools
• various possible processes

Final Goal:

• complete transparency, automation, regular updates

Solution:

• Do it yourself! Create a SAS macro!
Getting Started

Internet Search

Define-XML-2-0_ReleasePackage20140424

ARM-for-Define-XML

Define Papers

“Creating Define.xml v2 Using SAS® for FDA Submissions”
Getting started

Inputs SDTM

- SDTM data
- SDTM metadata
- Codelists & Thesauri

additional manual input

define.xml

- eCRF / aCRF
- additional mapping specifications
- RAVE loader information

XML

additional manual input
Getting started

Inputs ADAM

ADAM data

ADAM metadata

Codelists & Thesauri

additional manual input

define.xml

SAP, Protocol, TLF spec.

additional mapping specifications
Getting started

Inputs ADAM

- ADAM data
- ADAM metadata
- Codelists & Thesauri

additional manual input

SAP, Protocol, TLF spec.

additional mapping specifications
Getting started

Simplified process

- Step 1: Define XML information
- Step 2: Define XML
Getting Started

Macro Development

1. Prototyp
2. Discuss with Experts (ADAM, SDTM, CDISC)
3. Update prototyp
4. Develop process(es)
Getting started

Long Term Process

- %CreateDefine
- %UpdateDefine
- %ManualUpdate
- %DefineXML
- %CheckStructure
- %CheckContent
- %ReportChanges
- %ReportCheck Changes
%DefineXML

• Scope:
  • Translation of information into XML format according CDISC specifications
  • As interface SAS datasets are used
  • Similar as Pinnacle21 tool, but using SAS and adding functionality
%DefineXML

Interface through datasets (similar as Pinnacle21):

- General
- Datasets
- Variables
- Value Level (including WhereClause)
- Codes & Terms
- Methods
- Comments
- Documents
%DefineXML

Available through PhUSE Script Repository

- Macros
- Example
- Documentation

```
docs
  DefineXML - Macro Notes.docx
results
  define2-0-0.xsl
  example_adam.xml

src
  definexml_v1.sas
  example_adam.sas
  definexml_v1_attributes_for.sas
  example_adam.xlsx
```

phuse-scripts / lang / SAS / datahandle / defineXML
DEFINEXML

definexml_v1.sas

```sas
PUT @5 '<ItemDef OID=" itemoid +(-1) '"' /
@7 'Name=" valuename +(-1) '"' /
@7 'SASFieldName=" valuename +(-1) '"' /
@7 'DataType=" type +(-1) '"' /
@7 'Length=" length +(-1) '"' ;
IF significantdigits NE ''
 THEN PUT @7 'SignificantDigits=" significantdigits +(-1) '"';
```
%ELSE %IF %UPCASE(&data.) = DOCUMENTS
%THEN %DO;
  ATTRIB PDFLEAF LENGTH = $200 FORMAT = $200. LABEL = 'PDF Leaf ID';
  ATTRIB PDFLINK LENGTH = $200 FORMAT = $200. LABEL = 'PDF Link';
  ATTRIB PDFTITLE LENGTH = $200 FORMAT = $200. LABEL = 'PDF Title';
  ATTRIB ANNOTATEDCRF LENGTH = 8 FORMAT = 8. LABEL = 'Annotated CRF';
%END;
DATA DOCUMENTS;
  %definexml_v1_attributes_for(documents);
  INPUT PDFLEAF $ 1-21 PDFLINK $ 23-36 PDFTITLE $ 38-66 ANNOTATEDCRF 68-79;
DATALINES4;
LF.AdamReviewersGuide 12345_ADRG.pdf Analysis Data Reviewers Guide .  
LF.SAP 12345_SAP.pdf Statistical Analysis Plan .
; ; ;
RUN;

%definexml_v1(outfile = &path./results/phuse_example.xml);
<table>
<thead>
<tr>
<th>Domain</th>
<th>Description</th>
<th>Class</th>
<th>Structure</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADSL</td>
<td>Subject Level Analysis Dataset</td>
<td>SUBJECT LEVEL ANALYSIS</td>
<td>One record per Subject</td>
<td>Analysis</td>
</tr>
<tr>
<td>ADAQ</td>
<td>Adverse Events Analysis Dataset</td>
<td>OCCURRENCE DATA STRUCTURE</td>
<td>One record per subject</td>
<td>Analysis</td>
</tr>
<tr>
<td>ADEX</td>
<td>Exposure Analysis Dataset</td>
<td>BASIC DATA STRUCTURE</td>
<td>One or more records per Exposure Analysis</td>
<td>Analysis</td>
</tr>
</tbody>
</table>
ADaM-IG 1.0

Analysis Data Rev
Statistical Analysis:
► Analysis Datasets
► Parameter Value
► Controlled Termin
► Analysis Derivatio

Standard
ADaM-IG 1.0

Study Name
Study 12345

Study Description
This is the study description, this study is a dummy study used as example for the define.xml generation.

Protocol Name
Protocol 98765/12345

Metadata Name
Study 12345, Data Definitions

Metadata Description
Study 12345, Data Definitions

Analysis Datasets for Study Study 12345 (ADaM-IG 1.0)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Description</th>
<th>Class</th>
<th>Structure</th>
<th>Purpose</th>
<th>Keys</th>
<th>Location</th>
<th>Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADSL</td>
<td>Subject Level Analysis Dataset</td>
<td>SUBJECT LEVEL ANALYSIS DATASET</td>
<td>One record per Subject</td>
<td>Analysis</td>
<td>STUDYID, USUBJID</td>
<td>adsl.xpt</td>
<td></td>
</tr>
</tbody>
</table>
The following XML Attributes are set for the MethodDef definition:

<table>
<thead>
<tr>
<th>XML Attribute</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>OID</td>
<td>ComputationMethodDefID</td>
</tr>
<tr>
<td>Name</td>
<td>ComputationMethodName</td>
</tr>
<tr>
<td>Type</td>
<td>ComputationMethodType</td>
</tr>
<tr>
<td>Description TranslatedText</td>
<td>Description</td>
</tr>
<tr>
<td>def:DocumentRef leafID</td>
<td>PDFLeaf (if provided)</td>
</tr>
<tr>
<td>def:PDFPageRef PageRefs</td>
<td>PDFPage (if provided)</td>
</tr>
<tr>
<td>Type</td>
<td>“PhysicalRef”</td>
</tr>
<tr>
<td>def:PDFPageRef PageRefs</td>
<td>PDFDestination (if provided and PDFPage not)</td>
</tr>
<tr>
<td>Type</td>
<td>“NamedDestination”</td>
</tr>
</tbody>
</table>
%DefineXML

- Macro is Open Source (MIT-License)

- Validation:
  - Validate Define.XML result!

- Just use it!
Collaboration

Room for improvements:

Define Programming:

- Interface as Pinnacle21 Excel structure (%definexml_v2)
- Analysis Results Level metadata support
- Examples SDTM
- Examples ADAM / Update example

Define-related Programming

- %ReadDefineXML
- %CreateInitialDefineData (proc contents, …)
Collaboration

Room for improvements:

Documentation:

- Programming Specification
- Guidance or How-to Document
- FAQ
Summary

No more manual define.xml updates!

Define.xml creation with correct structure can easily be done by the macro!

Ready for use!
Full Transparency!
Easy to change!

Out of the box!
Questions?
Thank you!