

Metadata Driven Programming for CDISC SDTM using SAS

**PhUSE Single Day Event – Hyderabad
17th June 2017**

Kishore Shyamala

- Metadata
- XML to SAS
- aCRF creation using XML and SAS

CDISC Example study is used for this presentation.

Why?

- Reduce programming efforts
- Efficiency with simple SAS Macros
- Validate the datasets against metadata
- Use SAS to simplify annotated CRF creation

Metadata

Components of Metadata

- **Dataset-level metadata**
 - Provide basic information about each of the datasets
- **Variable-level metadata**
 - Contain the attributes for each variable within each dataset
- **Value-level metadata**
 - Normalized data structure of the SDTM Findings Class and Supplemental Qualifiers

XML

- An XML document is a string of *characters*
- With **Tags, Element** and **Attribute**
- Both human-readable and machine-readable.
- XMLMAP to read XML into SAS

define.xml

```
<ODM>
```

```
  <Study OID="">
```

```
    <MetaDataVersion OID="">
```

```
      <def:AnnotatedCRF>
```

```
    </def:AnnotatedCRF>
```

```
    <def:SupplementalDoc>
```

```
  </def:SupplementalDoc>
```

```
    <def:ValueListDef OID="">
```

```
  </def:ValueListDef>
```

```
    <def:WhereClauseDef OID="">
```

```
  </def:WhereClauseDef>
```

```
    <ItemGroupDef OID="">
```

```
  </ItemGroupDef>
```

```
    <ItemDef OID="" Name="" DataType="" Length="" SASFieldName="">
```

```
  </ItemDef>
```

Value-level metadata

Dataset-level metadata

Variable-level metadata

XMLMAP

- Lets SAS know how to read XML as a SAS dataset
- Can be created using SAS XML Mapper

```
<?xml version="1.0" encoding="windows-1252"?>

<SXLEMAP name="AUTO_GEN" version="2.1">

  <TABLE description="ItemGroupDef" name="Datasets">
    <TABLE-PATH syntax="XPathENR">/{1}ODM/{1}Study/{1}MetaDataVersion/{1}ItemGroupDef</TABLE-PATH>

    <COLUMN name="ItemGroupDef_SASDatasetName">
      <PATH syntax="XPathENR">/{1}ODM/{1}Study/{1}MetaDataVersion/{1}ItemGroupDef/@SASDatasetName</PATH>
      <TYPE>character</TYPE>
      <DATATYPE>string</DATATYPE>
      <LENGTH>8</LENGTH>
    </COLUMN>

    <COLUMN name="TranslatedText">
      <PATH syntax="XPathENR">/{1}ODM/{1}Study/{1}MetaDataVersion/{1}ItemGroupDef/{1}Description/{1}TranslatedText</PATH>
      <TYPE>character</TYPE>
      <DATATYPE>string</DATATYPE>
      <LENGTH>36</LENGTH>
    </COLUMN>

  </TABLE>
</SXLEMAP>
```


XML to SAS

- Access XML as SAS datasets by using XML engine library and the mapping defined in XMLMAP

```
filename define "&path.\define.xml";  
filename dmap   "&path.\define.map";
```

```
libname define xmlv2 xmlmap=dmap access=readonly;
```

- These datasets are further processed and SAS dataset shells are created for SDTM domains

Dataset-level metadata

ItemGroupDef_ORDINAL	ItemGroupDef_OID	ItemGroupDef_Class	ItemGroupDef_SASDatasetName	TranslatedText
1	IG.TA	TRIAL DESIGN	TA	Trial Arms
2	IG.TE	TRIAL DESIGN	TE	Trial Elements
3	IG.TI	TRIAL DESIGN	TI	Trial Inclusion/Exclusion Criteria
4	IG.TS	TRIAL DESIGN	TS	Trial Summary
5	IG.TV	TRIAL DESIGN	TV	Trial Visits
6	IG.DM	SPECIAL PURPOSE	DM	Demographics
7	IG.SE	SPECIAL PURPOSE	SE	Subject Elements
8	IG.SV	SPECIAL PURPOSE	SV	Subject Visits
9	IG.CM	INTERVENTIONS	CM	Concomitant Medications
10	IG.EX	INTERVENTIONS	EX	Exposure
11	IG.AE	EVENTS	AE	Adverse Events
12	IG.DS	EVENTS	DS	Disposition
13	IG.MH	EVENTS	MH	Medical History
14	IG.DA	FINDINGS	DA	Drug Accountability
15	IG.EG	FINDINGS	EG	ECG Test Results
16	IG.IE	FINDINGS	IE	Inclusion/Exclusion Criteria Not Met
17	IG.LB	FINDINGS	LB	Laboratory Tests Results

Variable-level metadata

ItemDef_OID	ItemDef_SASF	TranslatedText	ItemDef_	ItemDef_Len	CodeListRef_Cod	Origin_Typ	PDFPag
IT.AE.AEACN	AEACN	Action Taken with Study Treatment	text	30	CL.ACN	CRF	21
IT.AE.AEBODSYS	AEBODSYS	Body System or Organ Class	text	52	CL.AEDICT_F	Assigned	
IT.AE.AEDECOD	AEDECOD	Dictionary-Derived Term	text	18	CL.AEDICT_F	Assigned	
IT.AE.AEENDTC	AEENDTC	End Date/Time of Adverse Event	date	.		CRF	21
IT.AE.AEENDY	AEENDY	Study Day of End of Adverse Event	integer	3		Derived	
IT.AE.AEENRF	AEENRF	End Relative to Reference Period	text	5	CL.AEENRF	CRF	21
IT.AE.AEMODIFY	AEMODIFY	Modified Reported Term	text	9		Assigned	
IT.AE.AEREL	AEREL	Causality	text	16	CL.AEREL	CRF	21
IT.AE.AESEQ	AESEQ	Sequence Number	integer	1		Derived	
IT.AE.AESER	AESER	Serious Event	text	1	CL.NY	CRF	21
IT.AE.AESEV	AESEV	Severity/Intensity	text	8	CL.AESEV	CRF	21
IT.AE.AESPID	AESPID	Sponsor-Defined Identifier	text	4		CRF	21
IT.AE.AESTDTC	AESTDTC	Start Date/Time of Adverse Event	date	.		CRF	21

XML to SAS

```
proc sql;
  create table varlist as select
    a.ItemGroupDef_SASDatasetName as SASDatasetName,
    a.TranslatedText as SASDatasetLabel,
    b.ItemRef_OrderNumber as OrderNumber,
    b.ItemRef_KeySequence as KeySequence,
    c.ItemDef_SASFieldName as SASFieldName,
    c.TranslatedText as SASFieldLabel,
    c.ItemDef_DataType as Type,
    c.ItemDef_Length as Length,
    c.ItemDef_SignificantDigits as SignificantDigits,
    c.ItemDef_DisplayFormat as DisplayFormat
  from
    datasets as a,
    Variable_order as b,
    Variables c
  where a.ItemGroupDef_OID = b.ItemGroupDef_OID
  and b.ItemRef_ItemOID = c.ItemDef_OID
  order by a.ItemGroupDef_SASDatasetName, OrderNumber;
quit;
```

XML to SAS

SASDatasetName	SASDatasetLabel	OrderNumber	KeySequence	SASFieldName	SASFieldLabel	Type	Length
AE	Adverse Events	1	1	STUDYID	Study Identifier	text	7
AE	Adverse Events	2	.	DOMAIN	Domain Abbreviation	text	2
AE	Adverse Events	3	2	USUBJID	Unique Subject Identifier	text	14
AE	Adverse Events	4	.	AESEQ	Sequence Number	integer	1
AE	Adverse Events	5	.	AESPID	Sponsor-Defined Identifier	text	4
AE	Adverse Events	6	.	AETERM	Reported Term for the Adverse Event	text	25
AE	Adverse Events	7	.	AEMODIFY	Modified Reported Term	text	9
AE	Adverse Events	8	3	AEDECOD	Dictionary-Derived Term	text	18
AE	Adverse Events	9	.	AEBODSYS	Body System or Organ Class	text	52
AE	Adverse Events	10	.	AESEV	Severity/Intensity	text	8
AE	Adverse Events	11	.	AESER	Serious Event	text	1
AE	Adverse Events	12	.	AEACN	Action Taken with Study Treatment	text	30
AE	Adverse Events	13	.	AEREL	Causality	text	16
AE	Adverse Events	14	4	AESTDTC	Start Date/Time of Adverse Event	date	.
AE	Adverse Events	15	.	AEENDTC	End Date/Time of Adverse Event	date	.
AE	Adverse Events	16	.	AESTDY	Study Day of Start of Adverse Event	integer	3
AE	Adverse Events	17	.	AEENDY	Study Day of End of Adverse Event	integer	3
AE	Adverse Events	18	.	AEENRF	End Relative to Reference Period	text	5

SDTM Programming

- Create SAS dataset shells
 - Assign attributes using metadata
- Maintain Order of variables within a dataset
- Check if KEY variables are defined correctly

SDTM Validation

- Validate datasets against metadata
 - Make sure all datasets and variables are present
 - Sorting order is as per the specification
- Identify the data truncation by increasing the length on Validation side

aCRF Creation

- Use metadata to create annotations
 - Dataset information
 - All variables with “CRF” as Origin
 - PDF page numbers from define.xml

Dataset:

ItemGroupDef_SASDatasetName	TranslatedText
SC	Subject Characteristics

Variable:

ItemDef_OID	ItemDef_SASFi	TranslatedText	Origin_Type	PDFPageR
IT.SC.DOMAIN	DOMAIN	Domain Abbreviation	Assigned	
IT.SC.SCDTC	SCDTC	Date/Time of Collection	CRF	6
IT.SC.SCDY	SCDY	Study Day of Examination	Derived	
IT.SC.SCORRES	SCORRES	Result or Finding in Original Units	CRF	3 6
IT.SC.SCSEQ	SCSEQ	Sequence Number	Derived	
IT.SC.SCSTRESC	SCSTRESC	Character Result/Finding in Std Format	Derived	
IT.SC.SCTEST	SCTEST	Subject Characteristic	CRF	6
IT.SC.SCTESTCD	SCTESTCD	Subject Characteristic Short Name	Assigned	

aCRF Creation...

- Process this information and derive annotation text by CRF page number

SASDatasetName	TranslatedText	SASFieldName	text	PageRefs	page
DM	Demographics		DM = Demographics		3
SC	Subject Characteristics		SC = Subject Characteristics		3
DM		SITEID	DM.SITEID	3	3
DM		SUBJID	DM.SUBJID	3	3
SC		SCORRES	SCORRES	3 6	3
DM	Demographics		DM = Demographics		6
DS	Disposition		DS = Disposition		6
SC	Subject Characteristics		SC = Subject Characteristics		6
DM		BRTHDTC	DM.BRTHDTC	6	6
DM		ETHNIC	DM.ETHNIC	6	6
DM		RACE	DM.RACE	6	6
DM		SEX	DM.SEX	6	6
DM		RACEOTH	SUPPDM.QVAL when QNAM = 'RACEOTH'	6	6
DS		DSDECOD	DSDECOD	6 16 18	6
DS		DSSTDTC	DSSTDTC	6 16 18	6
DS		DSTERM	DSTERM	6 16 18	6
SC		SCDTC	SCDTC	6	6
SC		SCTEST	SCTEST	6	6
SC		SCORRES	SCORRES	3 6	6

aCRF Creation...

- Adobe Acrobat uses FDF (Forms Data Format) files to export/import Annotations
 - A one-line **header** - Version of the PDF specification
 - A **body** containing the **objects**
 - A **trailer** giving the location of various objects within the body of the file

FDF (Forms Data Format)

%FDF-1.2

Header

```
1 0 obj<</FDF<</Annots[  
2 0 R  

```

**Object 1: Catalog of all other objects.
Referenced in trailer**

```
2 0 obj<<  
/Type/Annot                /Subtype/FreeText  
/C[0.75 1.0 1.0]          /Contents(DM = Demographics)  
/DS(font: italic bold Arial,sans-serif 14.0pt; text-align:left; color:#000000 )  
/Page 2                    /Rect[4 730 204 750]  
>>endobj
```

trailer

```
<</Root 1 0 R>>
```

```
%%EOF
```

Trailer

aCRF Creation...

- Add the FDF tags to the annotation text
- Create FDF file
- Import into PDF file

REFERENCES:

1. PDF Reference, sixth edition, Adobe Systems Incorporated
http://www.adobe.com/content/dam/Adobe/en/devnet/acrobat/pdfs/pdf_reference_1-7.pdf
2. "Using SAS to Speed up Annotating Case Report Forms in PDF Format" -
Spruck, Dirk & Kawohl, Monika; 2004 Pharmaceutical SAS Users Group. Cary, NC: SAS Institute Inc.
www.lexjansen.com/pharmasug/2004/coderscorner/cc02.pdf
3. "Have SAS Annotate your Blank CRF for you! Plus dynamically add color and style to your annotations."
Steven Black, Agility-Clinical Inc., Carlsbad, CA, 2015 Pharmaceutical SAS Users Group.
www.pharmasug.org/proceedings/2015/AD/PharmaSUG-2015-AD05.pdf

aCRF Creation...

- Annotations are created at the top of the page
- Move and re-size as appropriate
- Create bookmarks and TOC

DM = Demographics		DS = Disposition	SC = Subject Characteristics
DM.BRTHDTC			Assessment Date
DM.ETHNIC	SC01		____/____/____
DM.RACE			
DM.SEX	DEMOGRAPHY		
SUPPDM.QVAL when QNAM = 'RACEOTH'			
DSDECOD	Birth: ____/____/____		
DSSTDTC			
DSTERM			
SCDTC	Gender: ☐ Male ☐ Female		
SCTEST			
SCORRES	☐ Hispanic or Latino ☐ Not Hispanic or Latino		

Thank You!

Kishore Shyamala

Domain Clinical Data Sciences Private Limited

Email: Kishore@domaincds.com