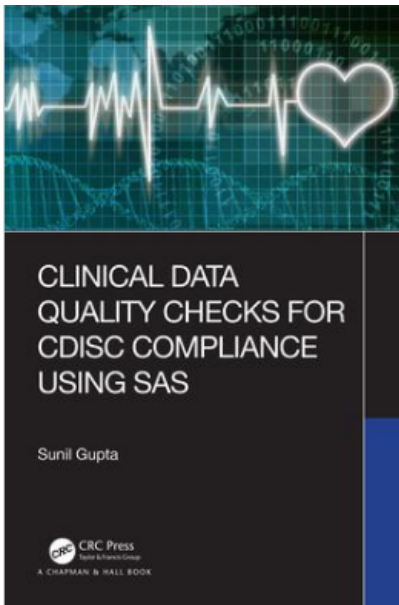


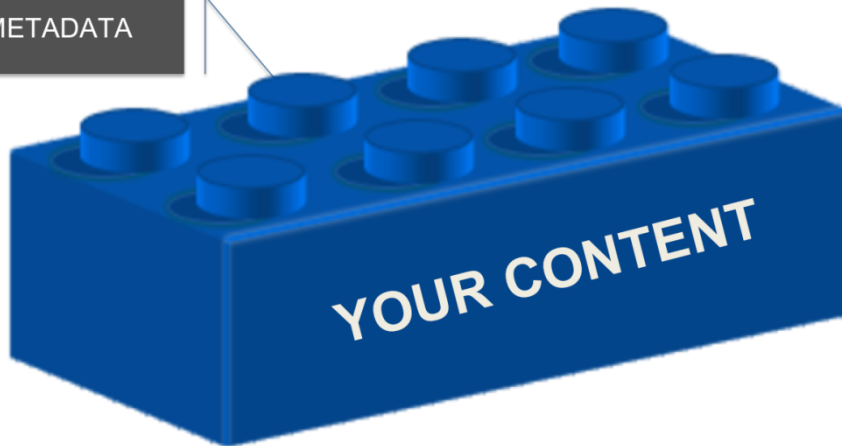
Leveraging Metadata for SDTM QC Automation

Sunil Gupta
Executive Director, Data Sciences
TalentMine



Why is Metadata 'Intelligence' Important?

METADATA



- ✓ Metadata describes and documents the content
- ✓ Metadata is separate from content but within the same file
- ✓ Standard and custom metadata
- ✓ No hard coding of content
- ✓ Standardize process
- ✓ Automate production process

Metadata

The three main types work together.

Attributes/

Elements

Structural

Values

Descriptive or administrative

Metadata example

Structural

Title:

Description:

Creator:

Date:

Type:

Descriptive

'Mad Men' Season 5 plot predictions
The two-hour premiere quickly answered lingering questions from season four.

Administrative

Lucas Shaw

26/03/2012

Article

Metadata Application SAS Examples

Access/Process Metadata information like SAS datasets

**Automatically
Check
DateTime
Sequence of
Source/QC
SAS/Log Files**

**Automatically
Check
DateTime
Sequence of
Datasets
Across
Multiple
Libraries**

**Automatically
Proc Compare
Two SDTM
Libraries**

SAS: Extract intelligence information from metadata and macro processing

Data-Driven process is automatic, quality controlled, transparent and saves time!

INPUT:

Directory of Files (Excel files, Datasets, SAS Programs, Log and Proc Compare.lst)

SAS Tools

- Libnames
- Data Step
- SAS Macro Programming
- SAS & Dataset Functions
- Proc SQL
- Proc Compare
- Proc Means

OUTPUT: Metadata Attributes

- File pathname and names
- # of Files, Datetime stamps
- # and Type of variables
- Macro loop through all files
- Required datasets, variables, etc.
- Codelist dictionary, SDTM/ADaM attributes

Maximum variable lengths

Data cleaning & monitoring of valid variables and special characters

- Compare and contrast previous file
- Descriptive Statistics on categorical and continuous variables
- Search for ERRORS, WARNINGS or Notes in SAS Logs
- Search for QC differences in # of VARs, OBS, attributes and dups

SAS's Four Key Dictionary Tables

Dictionary tables are available once LIBNAME statement is assigned

Category	Description	PROC SQL or DATA Step: SASHelp View
1. Dictionaries	Category of files: Catalogs, etc.	sashelp.vdctnry
2. Members	List of SAS libraries and data sets	sashelp.vmember
3. Tables	List of data sets and columns	sashelp.vtable
4. Columns	List of columns and attributes	sashelp.vcolumn

 dm	8/25/2016 7:40 PM	SAS Data Set	256 KB
 dosing	1/13/2013 1:11 PM	SAS Data Set	17 KB
 formatdata	1/13/2013 1:11 PM	SAS Data Set	17 KB
 labs	1/13/2013 1:11 PM	SAS Data Set	33 KB
 pain	1/13/2013 1:11 PM	SAS Data Set	9 KB
 responders	8/25/2016 7:40 PM	SAS Data Set	256 KB
 sdtm_metadata	8/25/2016 7:08 PM	SAS Data Set	0 KB
 sdtm_metadata1	8/25/2016 7:09 PM	SAS Data Set	0 KB
 sdtm_metadata2	8/25/2016 7:10 PM	SAS Data Set	256 KB
 suppdm	8/25/2016 7:40 PM	SAS Data Set	256 KB
 xp	8/25/2016 7:40 PM	SAS Data Set	256 KB
 adae	9/18/2013 6:51 AM	SAS System Program	2 KB
 adamtrace	1/13/2013 1:11 PM	SAS System Program	13 KB
 adef	9/18/2013 6:50 AM	SAS System Program	2 KB
 adsl	9/18/2013 6:51 AM	SAS System Program	3 KB
 adtte	9/18/2013 6:51 AM	SAS System Program	4 KB

Datasets

**Apply
DO
LOOP
process**

**SAS
Programs**

**Metadata File:
Pathname, Type,
Name, DateTime
Stamp, Size,
Vars, # OBS, etc.**



Automatically Proc Compare Two ADaM Libraries

Source Library

Name	Date modified	Type
 adae	1/13/2013 1:11 PM	SAS Data Set
 adef	1/13/2013 1:11 PM	SAS Data Set
 adsl	1/13/2013 1:11 PM	SAS Data Set
 adtte	1/13/2013 1:11 PM	SAS Data Set

QC Library

Name	Date modified	Type
 adae	1/13/2013 1:11 PM	SAS Data Set
 adef	1/13/2013 1:11 PM	SAS Data Set
 adsl	1/13/2013 1:11 PM	SAS Data Set
 adtte	1/13/2013 1:11 PM	SAS Data Set



Summary Level

- **SASHELP.VTABLE** from two Libnames
- **Metadata: CRDATE, NOBS, NVAR, FILESIZE**

Metadata

- Same SDTMs in both Libraries
- **SASHELP.VCOLUMN** to create Lookup Table of Dataset Names

Detail Level

- **CALL EXECUTE** Proc Compare each Dataset Name
- Extract NOBS from OUTNOEQUAL datasets

Proc Compare SAS Example

```
proc sql;  
create table ds_diff as  
  select unique memname  
  from sashelp.vcolumn where upcase(libname) in ('SOURCE' 'QC');  
quit;
```

**Standard Metadata –
Proc SQL dictionary tables**

```
data ds_diff1;  
set ds_diff;
```

**CALL EXECUTE() executes PROC
COMPARE for each ADaM record in
DS_DIFF for robust processing**

```
  call execute('proc compare data=source.'  
  || memname  
  || ' compare= qc.' || memname || ' listall; run;' );  
run;
```

Example SAS Generated Code repeated for each ADaM

```
proc compare data=source.ADAE compare= qc.ADAE listall;  
run;
```

```
proc compare data=source.ADEF compare= qc.ADEF listall;  
run;
```

SDTM Patient Counts and Dates SAS Example

	1	2	3	4	5	6	7
	 tu_ptc	 tr_ptc	 rs_pct	 tu_tr_rs_pct	 tu_stamp	 tr_stamp	 rs_stamp
1	164	164	147	164	09JUL2018	09JUL2018	09JUL2018

```
proc sql;
create table dc as
select unique
count(unique USUBJID) as tu_ptc label="USUBJID count"
, b.*, c.*, d.*, e.*, f.*, g.*
from sdtm.tu as a,
(select count(unique USUBJID) as tr_ptc from sdtm.tr) as b,
(select count(unique USUBJID) as rs_pct from sdtm.rs) as c,
(select count(unique USUBJID) as tu_tr_rs_pct from tu_tr_rs) as d,
(select unique input(scan(put(crdate, DATETIME16.), 1, ':'), date7.)
format=date9. as tu_stamp from sashelp.vtable
where upcase(libname)="SDTM" and upcase(memname)=upcase("TU")) as e,
(select unique input(scan(put(crdate, DATETIME16.), 1, ':'), date7.)
format=date9. as tr_stamp from sashelp.vtable
where upcase(libname)="SDTM" and upcase(memname)=upcase("TR")) as f,
(select unique input(scan(put(crdate, DATETIME16.), 1, ':'), date7.)
format=date9. as rs_stamp from sashelp.vtable
where upcase(libname)="SDTM" and upcase(memname)=upcase("RS")) as g; quit;
```

**Customized Metadata –
Proc SQL with six subqueries**

Concatenate all new variables

**One subquery for
each SDTM patient
count and date stamp**

Proc SQL dictionary table

Examples of Standard Metadata from SASHELP

Dictionary tables are available once LIBNAME statements are assigned

Standard Metadata for all Domains and Variables

1. VTABLE

- a. Confirm # of datasets, observations and datetime stamps
- b. Confirm required datasets and variables

2. VCOLUMN

- a. Confirm # and type of variables and attributes
- b. Confirm order of variables

Examples of Customized Metadata

Take advantage of metadata to monitor timely completion of primary and secondary endpoints

Proc SQL Custom Metadata on Key Domains and Variables

1. Confirm # of patients
2. Confirm required non-missing variables
3. Confirm not greater than Maximum variable lengths
4. Confirm descriptive statistics on key categorical and continuous variables
5. Confirm compare and contrast previous datasets, attributes (more/less datasets/records/variables)
6. Confirm no duplicate records in each dataset
7. Include codelist dictionary from each dataset, ex. lab units

Leveraging Metadata for SDTM QC Automation Summary

Automate standard process for higher quality and tighter deadlines



- ✓ Proc SQL
 - ✓ Concatenate Variables
- ✓ SAS Generated Program
- ✓ Applicable to ADaM QC
- ✓ Metadata
 - ✓ Standard
 - ✓ Dictionary Tables
 - ✓ Custom
 - ✓ Primary
 - ✓ Secondary
- ✓ Analysis Results (ARM)

Get more information on leveraging metadata webinars @ Sunil.Gupta@TalentMine.net