

Automating Module 5 Packaging with Customizable Metadata

Dhananjay (DJ) Chhatre, Director, Gilead Sciences

10 March 2020

PhUSE US Connect 2020

DJ and his family 2019





AGENDA

PhUSE US Connect 2020

1

Background

2

Automation Strategy

3

Metadata Driven QC Checks

4

Customizable Metadata for Study Use

Background



FDA Binding Guidance requires Sponsors to submit standardized study data in regulatory submissions.

Providing Regulatory Submissions in Electronic Format — Submissions Under Section 745A(a) of the Federal Food, Drug, and Cosmetic Act

Guidance for Industry

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

December 2014
Electronic Submissions

STUDY DATA TECHNICAL CONFORMANCE GUIDE

Technical Specifications Document

This Document is incorporated by reference into the following
Guidance Document(s):

**Guidance for Industry Providing Regulatory Submissions in Electronic
Format — Standardized Study Data**

For questions regarding this technical specifications document, contact CDER at
cdet.edata@fda.hhs.gov or CBER at cber.cdus@fda.hhs.gov

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

October 2019

Providing Regulatory Submissions In Electronic Format — Standardized Study Data

Guidance for Industry

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

December 2014
Electronic Submissions



CDER 21st Century Review Process

Desk Reference Guide



**Standardized data allows
reviewers to more fully
understand and characterize
efficacy and safety.**



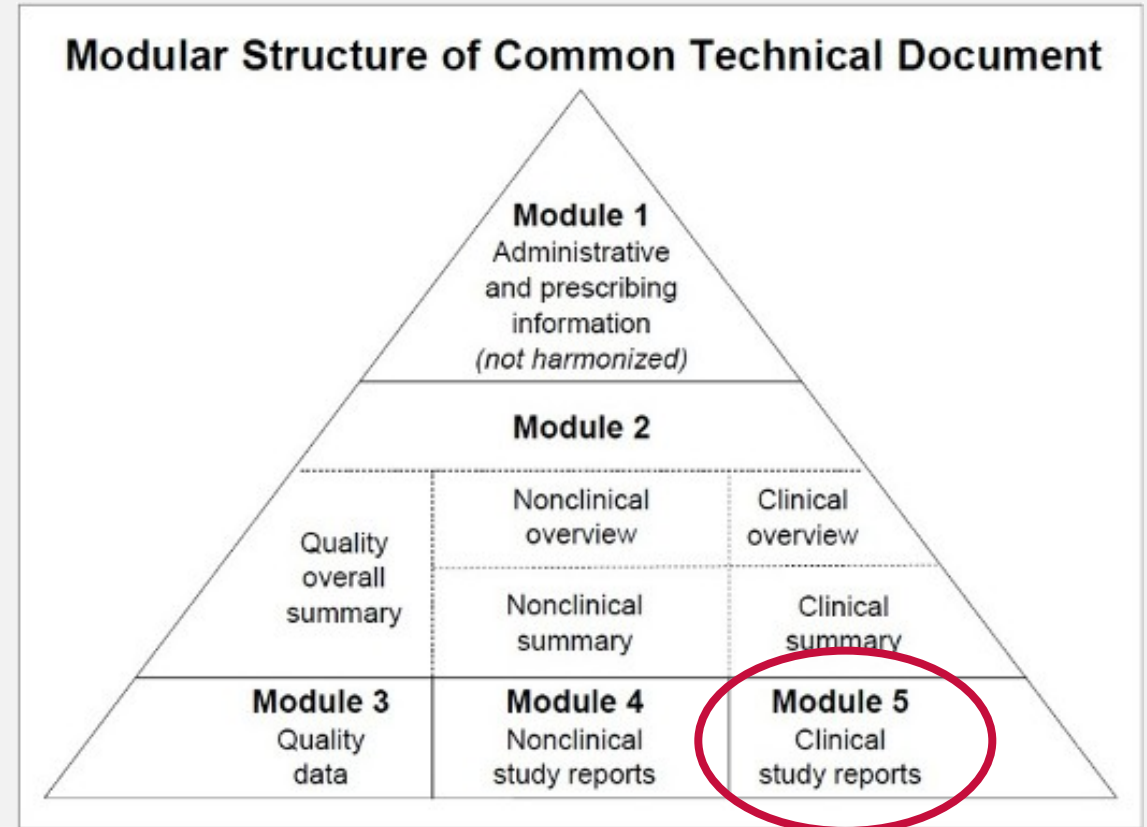
**How do we submit this
study data to FDA?**

eSubmissions Gateway
requiring adherence to
Electronic Common
Technical Document (eCTD)
Standards



Study dataset packages must include relevant documents in Module 5 to facilitate regulatory review

- Tabular Listing of All Clinical Studies
- Study Reports
- Case Report Forms (CRFs)
- Periodic Safety Reports
- IND Safety Reports
- Literature References



DATASETS!

**Standardized data allows
reviewers to more fully
understand and characterize
efficacy and safety.**



Metadata Driven Automation

>

Manual Preparation & QC

- Reduce resource burdens
- Improve efficiencies
- Improve quality



Automation Strategy

IDEA

SCIENCE

THOUGHT

INTENSITY

SHARE

KNOWLEDGE

BRAIN

INNOVATION

CREATIVITY

TEAM

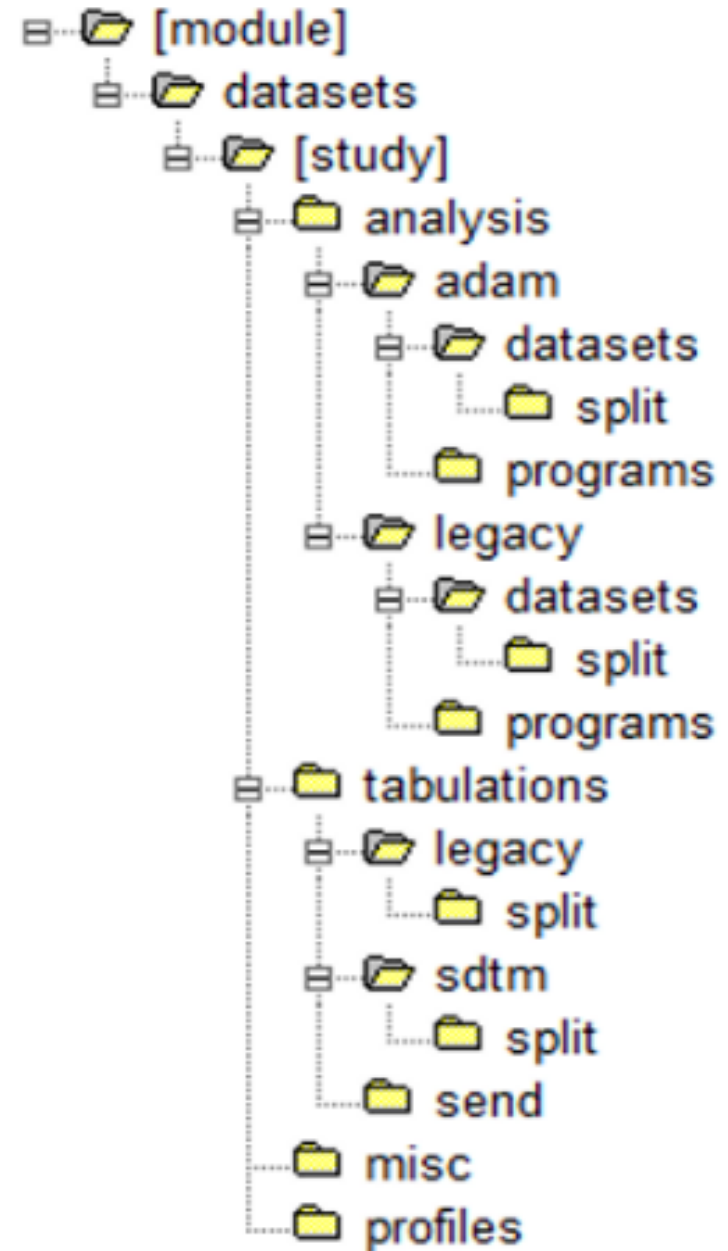
POWER

ILLUMINATION

LIGHTING

ENGINEERING

The eCTD directory structure allows for easy mapping of study working area to target M5 folders



Metadata files drive automation and enable traceability, while allowing for study customization

SDTM Datasets

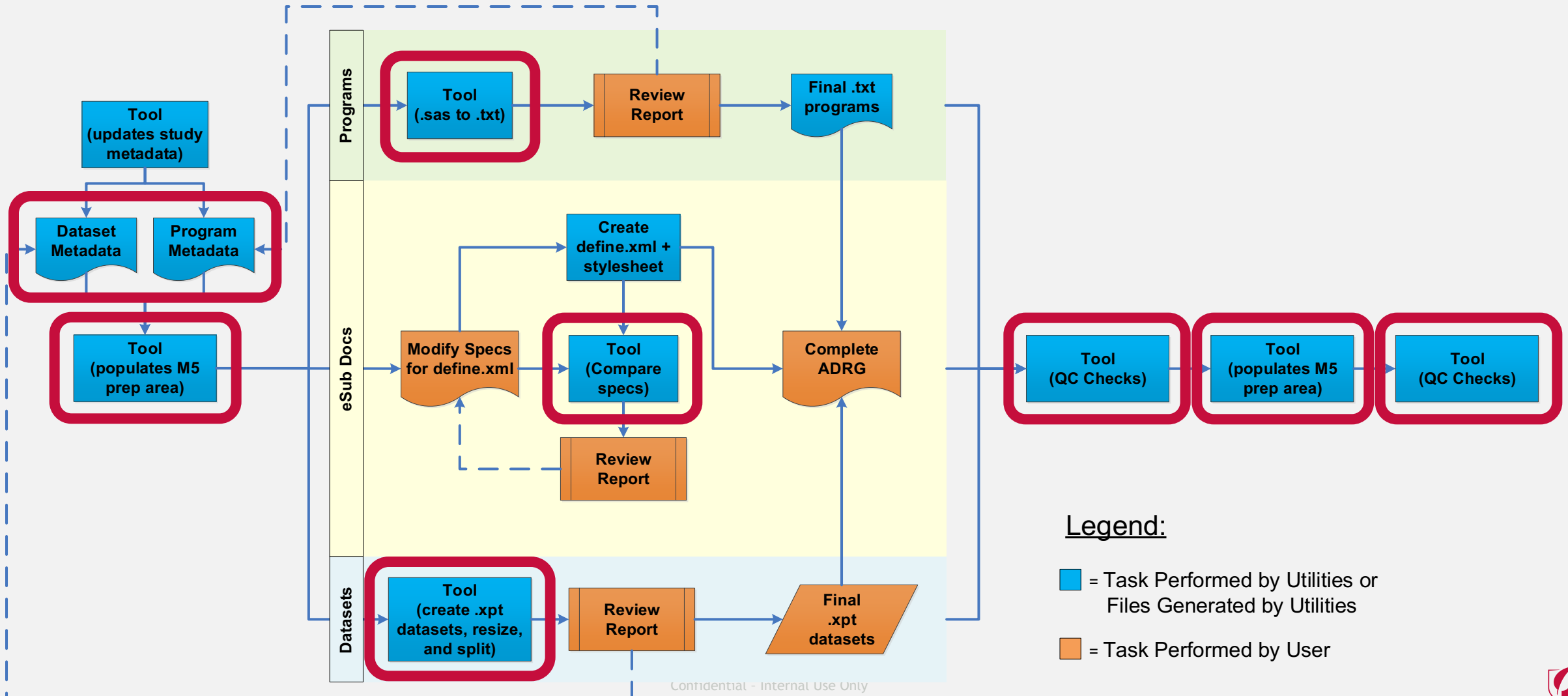
Active	File name	Split	Split name	Where	Source	Destination
Y	ae.sas7bdat	N			\\gasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	cm.sas7bdat	N			\\gasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	sv.sas7bdat	N			\\gasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	dm.sas7bda	N			\\gasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	vs.sas7bdat	N			\\gasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	ds.sas7bdat	N			\\gasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
N	test.sas7bda	N			\\gasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	lb.sas7bdat	Y	lb01	lbcatt = "CHEMISTRY"	\\gasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\split\
Y	lb.sas7bdat	Y	lb02	lbcatt = "MICROBIOLOGY"	\\gasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\split\
Y	lb.sas7bdat	Y	lb03	lbcatt ^in ("CHEMISTRY", "MICROBIOLOGY")	\\gasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\split\

ADaM Programs

Active	File name	Source	Destination
Y	t-demog.sas	\\gasdata\biometrics\projects\p123\s1234567\task\module5\prog\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	l-demog.sas	\\gasdata\biometrics\projects\p123\s1234567\task\module5\prog\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	g-eff.sas	\\gasdata\biometrics\projects\p123\s1234567\task\module5\prog\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
N	header.sas	\\gasdata\biometrics\projects\p123\s1234567\task\module5\prog\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	init.inc	\\gasdata\biometrics\projects\p123\s1234567\task\module5\tools\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	newvisid.sas	\\gasdata\biometrics\projects\p123\s1234567\task\module5\tools\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	a-adsl.sas	\\gasdata\biometrics\projects\p123\s1234567\task\module5\adamprog\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	a-adae.sas	\\gasdata\biometrics\projects\p123\s1234567\task\module5\adamprog\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	global1.sas	\\gasdata\biometrics\projects\p123\s1234567\task\module5\adamprog\	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\



Automation utilities leverage metadata input files to enable a now traceable M5 package generation process



A background image of Scrabble tiles on a wooden surface. Some tiles are arranged in a row to spell out 'LIT Y', while others are scattered around. A dark diagonal band runs across the image, serving as a backdrop for the title text.

Metadata Driven QC Checks

Metadata can also drive QC checks of M5 package contents

- Gilead Legacy Submission Checklists
- Gilead/Phase I-Specific Checks
- Gilead/Regulatory Operations Checklist
- Checks mentioned in various papers
- FDA Technical Rejection Criteria



Maintenance and updates to checks are streamlined allowing for easier updates and improved planning for new checks



Mock QC report: metadata presented to user to allow for easy understanding of errors & warnings

Status	Source	Tab	Type	Folder Path	File Name	Variable	Value	Check Description
ERROR	tabulations_data.xlsx	Datasets	SDTM	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\module5\tabulations\sdtm\	suppae.xpt			The file is found in corresponding module destination location but is missing from prep destination location.
ERROR	analysis_data.xlsx	Datasets	ADaM	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\module5\analysis\adam\dataset\	adsl.xpt			Subject-level analysis dataset (ADSL) must be submitted.
ERROR	analysis_data.xlsx	eSub Docs	Legacy	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\module5\analysis\legacy\dataset\	define.xml			The file is missing from the specified source location but is found in the specified destination location.
ERROR	analysis_prgm.xlsx	Programs	ADaM	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\module5\analysis\adam\program\	t-demog.sas			Only program files should be in the programs folder and all programs should be TXT files.
ERROR	general	N/A	N/A	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\module5\analysis\	adam			The sdtm and adam folders must be present.
ERROR	general	N/A	N/A	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\module5\tabulations\legacy\	sdrgr.doc			Invalid file type found.
WARNING	tabulations_data.xlsx	Datasets	SDTM	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\module5\tabulations\sdtm\	lb.xpt	LBTESTCD	EGFRMDRD	Variable length for --TESTCD or PARAMCD cannot exceed 8 characters in length. Variable may be truncated.

ADaM Programs



Example 1: Presence/Absence in Metadata and/or M5 folder

A	B	C	D	E	F	G
Active	File name	Split	Split name	Where	Source	Destination
Y	ae.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	cm.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	sv.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	dm.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	vs.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	ds.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
N	test.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	lb.sas7bdat	Y	lb01	lbcat = "CHEMISTRY"	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\split\
Y	lb.sas7bdat	Y	lb02	lbcat = "MICROBIOLOGY"	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\split\
Y	lb.sas7bdat	Y	lb03	lbcat ^in ("CHEMISTRY", "MICROBIOLOGY")	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\split\

Name	Date modified	Type
split	2/12/2020 9:36 AM	File folder
ae.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
cm.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
dm.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
ds.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
lb.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
sv.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
test.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
vs.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...

Missing in Metadata, Present In Folder

Name	Date modified	Type
split	2/12/2020 9:36 AM	File folder
ae.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
cm.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
dm.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
ds.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
lb.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
sv.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...

Present in Metadata, Missing in Folder



Example 2: Date/Time Stamp Mismatch

 tabulations_data.xlsx	1/27/2020 8:45 PM	Microsoft Excel W...	31 KB
---	-------------------	----------------------	-------

Metadata File has newer date/time stamp

Name	Date modified	Type
split	2/12/2020 9:36 AM	File folder
ae.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
cm.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
dm.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
ds.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
lb.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
sv.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
test.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...
vs.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...

Datasets have older date/time stamps



Example 3: Splitting Datasets

A	B	C	D	E	F	G
Active	File name	Split	Split name	Where	Source	Destination
Y	ae.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	cm.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	sv.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	dm.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	vs.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	ds.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
N	test.sas7bdat	N			\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\
Y	lb.sas7bdat	Y	lb01	lbcats = "CHEMISTRY"	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\split\
Y	lb.sas7bdat	Y	lb02	lbcats = "MICROBIOLOGY"	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\split\
Y	lb.sas7bdat	Y	lb03	lbcats ^in ("CHEMISTRY", "MICROBIOLOGY")	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\sdtmdata\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\tabulations\sdtm\split\

Name	Date modified	Type	Size
split	2/12/2020 9:36 AM	File folder	
ae.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...	512 KB
cm.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...	1,216 KB
dm.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...	896 KB
ds.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...	192 KB
lb.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...	8,882,816 KB
sv.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...	1,600 KB
vs.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...	768 KB

LB Dataset > 5 gb



> module5 > sdtmdata > split				Search split
Name	Date modified	Type	Size	
lb01.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...	3,512,816 KB	
lb02.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...	1,600,216 KB	
lb03.xpt	5/28/2019 3:58 PM	SAS Xport Transpo...	3,769,784 KB	

Three Split Datasets all <5 gb



Metadata-Driven Execution of QC Checks

Check Type	Check Description	Notes	FDA Severity	PMDA Severity	Prep, M5, or Both?	SDTM?	ADAM?	Misc?	Profiles?
Dataset	If .xpt datasets are greater than 5 gb in size, there should exist child datasets in the /split folder		Error	n/a	Both	Y	N	N	N
Define.xml	Datasets referenced in define.xml are listed in the respective Metadata file	SDTM = Tabulations Metadata ADaM = Analysis Metadata	Error	Error	Both	Y	Y	Y	N
PDFs	Verify that the file name for acrf.pdf is all lower case.	Different name for PMDA	Error	n/a	Both	Y	N	N	N
Programs	All program files have .sas extension	Different name for FDA	n/a	Error	Both	Y	Y	Y	N
Structural	All files in /adam/datasets have Active=Y in Analysis metadata		Error	Error	Both	Y	Y	Y	N



A custom blue and gold motorcycle is displayed in a museum. The motorcycle features intricate gold and blue paintwork, a large chrome headlight, and a prominent chrome engine. It is positioned on a dark, reflective surface, and a metal railing is visible in the foreground. The background shows other motorcycles on display.

Customizable Metadata for Study Use

Use Case 1: Inclusion of Non-Traditional Elements

Active	File name	Source	Destination
Y	t-demog.sas	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\prog\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	l-demog.sas	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\prog\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	g-eff.sas	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\prog\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
N	header.sas	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\prog\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	init.inc	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\tools\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	newvisid.sas	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\tools\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	a-adsl.sas	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\adamprog\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	a-adae.sas	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\adamprog\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\
Y	global1.sas	\\gsasdata\biometrics\projects\p123\s1234567\task\module5\adamprog\	\\gsasdata\biometrics\projects\p123\s1234567\task\esub\prep\analysis\adam\programs\

/tools

Name	Date modified	Type
init.inc	1/25/2017 5:29 PM	INC File
newvisid.sas	1/25/2017 5:29 PM	SAS System Progr...

/prog

Name	Date modified	Type
g-eff.sas	1/25/2017 5:29 PM	SAS System Progr...
header.sas	1/25/2017 5:29 PM	SAS System Progr...
l-demog.sas	1/25/2017 5:29 PM	SAS System Progr...
t-demog.sas	1/25/2017 5:29 PM	SAS System Progr...

/adamprog

Name	Date modified	Type
init.inc	1/25/2017 5:29 PM	INC File
newvisid.sas	1/25/2017 5:29 PM	SAS System Progr...

Documents in Source Areas



Name	Date modified	Type
a-adae.sas	1/25/2017 5:29 PM	SAS System Program
a-adsl.sas	1/25/2017 5:29 PM	SAS System Program
g-eff.sas	1/25/2017 5:29 PM	SAS System Program
global1.sas	1/25/2017 5:29 PM	SAS System Program
init.inc	1/25/2017 5:29 PM	INC File
l-demog.sas	1/25/2017 5:29 PM	SAS System Program
newvisid.sas	1/25/2017 5:29 PM	SAS System Program
t-demog.sas	1/25/2017 5:29 PM	SAS System Program

Documents in Destination



Use Case 2: Adding Datasets to Miscellaneous Folder

Active	File name	Source	Destination
Y	lookup.sas7bdat	\\gasdata\biometrics\projects\p123\s1234567\task\module5\adamdata	\\gasdata\biometrics\projects\p123\s1234567\task\esub\prep\misc

Name	Date modified	Type
 adae.sas7bdat	5/28/2019 3:58 PM	SAS Data Set
 adcm.sas7bdat	5/28/2019 3:58 PM	SAS Data Set
 adeff.sas7bdat	5/28/2019 3:58 PM	SAS Data Set
 adlb.sas7bdat	5/28/2019 3:58 PM	SAS Data Set
 adsl.sas7bdat	5/28/2019 3:58 PM	SAS Data Set
 advs.sas7bdat	5/28/2019 3:58 PM	SAS Data Set
 lookup.sas7bdat	5/28/2019 3:58 PM	SAS Data Set

Documents in Source Areas



Name	Date modified	Type
 lookup.sas7bdat	5/28/2019 3:58 PM	SAS Data Set

Documents in Destination





QUESTIONS?