SAS ユーザー総会 2017

申請電子データプロセスの効率化

～メタデータ・社内ツールを中心に～

〇飯島　康浩、梶原　正幹、小南　真由美、
中嶋　優一、北原　孝志

（ノバルティスファーマ株式会社）

Metadata Driven Approach for eSubmission

Yasuhiro Iijima, Yoshimasa Kajiwara, Mayumi Kominami,
Yuichi Nakajima, Takashi Kitahara

Novartis Pharma K.K.

要旨:

本発表では，申請電子データを効率的に作成するためのメタデータ及びメタデータを包括的に管理する社内のガバナンス体制，SASを用いたannotated CRFの自動作成やDefine.xml確認ツールを紹介する。

キーワード: 申請電子データ，メタデータ，annotated CRF，Define.xml

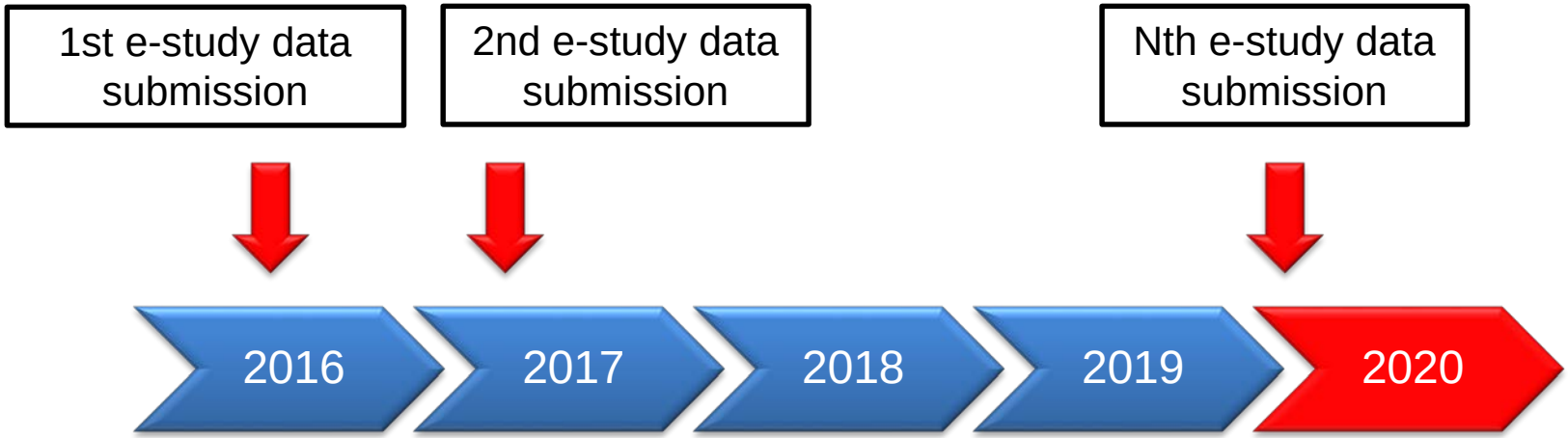
Disclaimer

- The opinions expressed in this presentation and on the following slides are solely those of the presenter and not necessarily those of Novartis. Novartis does not guarantee the accuracy or reliability of the information provided herein.

Agenda

- Introduction
- Part 1: Metadata Management
- Part 2: CRF annotation tool
- Part 3: Automated Define.xml QC Checks
- Summary

Introduction



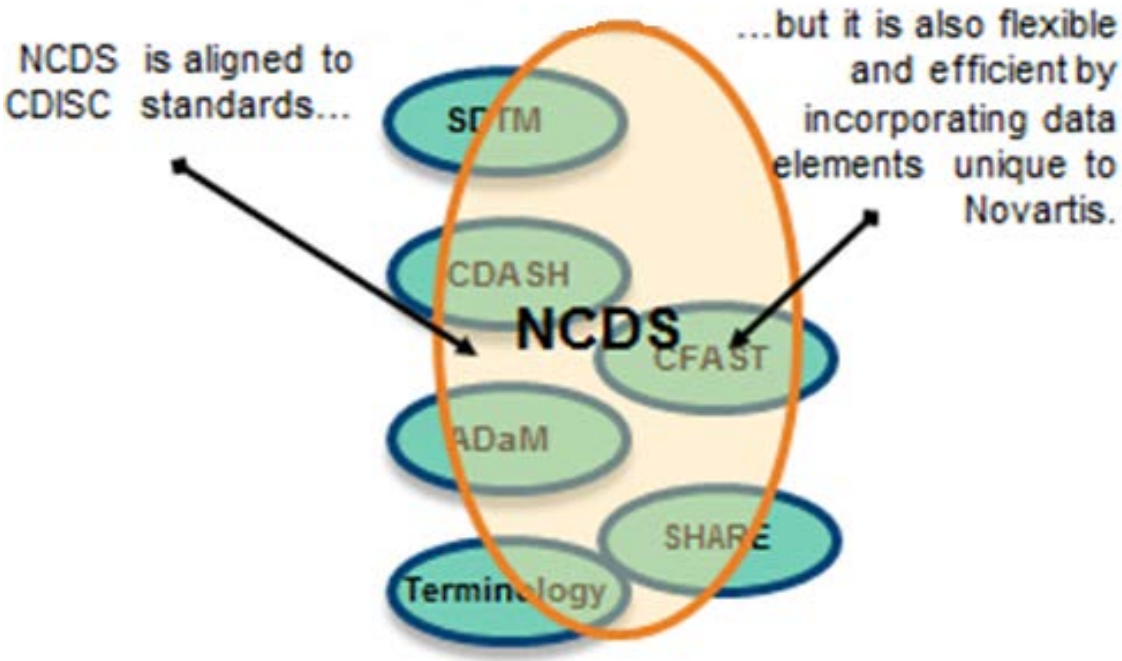
Oct-2016～
Start accepting e-study data

Transitional period (3.5 years)
Partial e-study data submission can be accepted

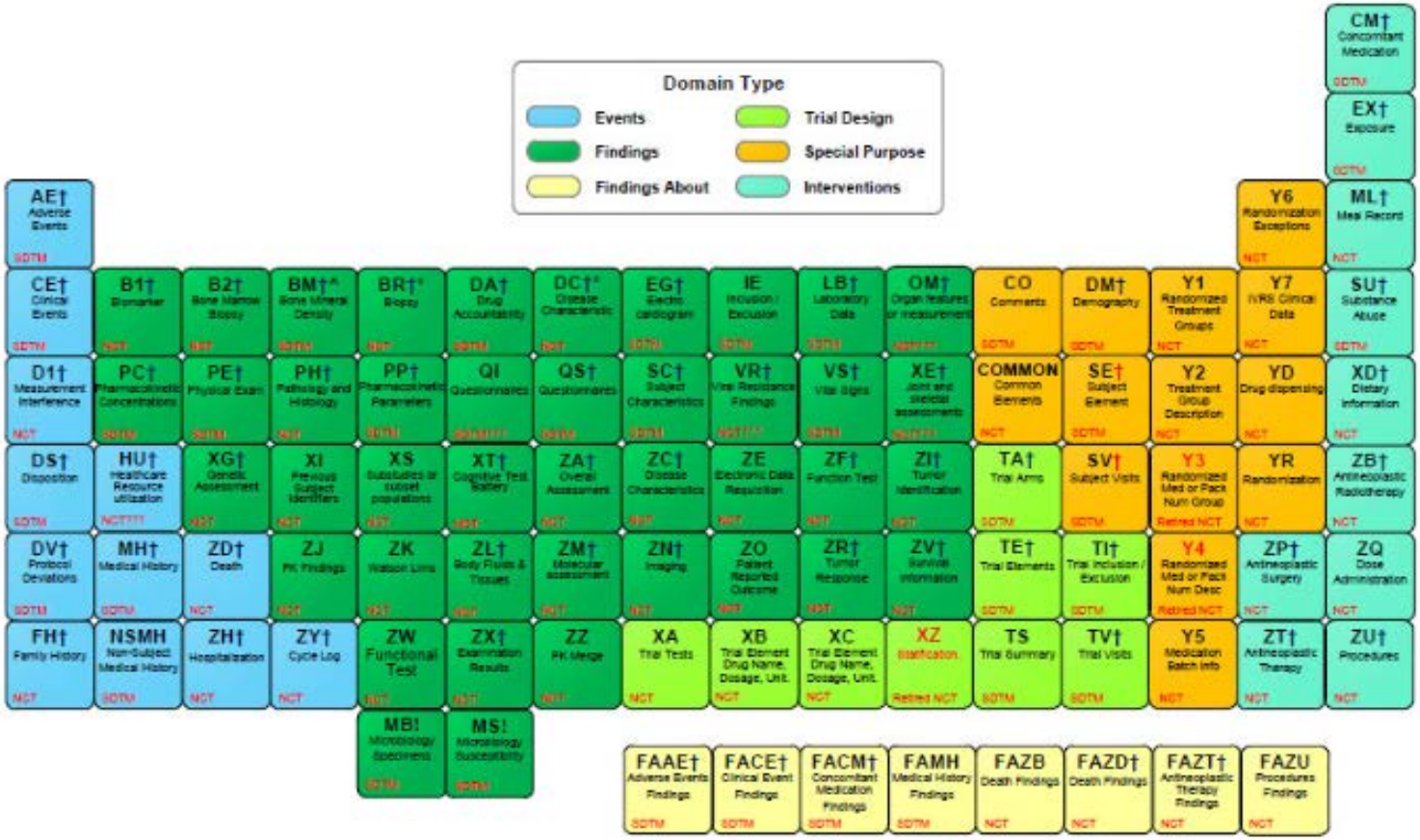
Apr-2020～
e-study data submission will be mandatory.

Data Standardization in Novartis

- Novartis Clinical Data Standard
  - The Novartis Clinical Data Standards (NCDS) contain all data elements and attributes needed for data collection, processing, analysis, reporting and submission.



Novartis' domains Overview



Overview of Metadata

Simplification and standardization

Standard governance drives consistency and quality

SDTM /ADAM / Define-XML

- Template programs
- Comprehensive DEFINE Review

aCRF

- Standard CRF
- Annotation tool

SDRG/ADRG

- Template and examples
- Training material and check list

SAS ユーザー総会 2017

Part 1.

Metadata Management

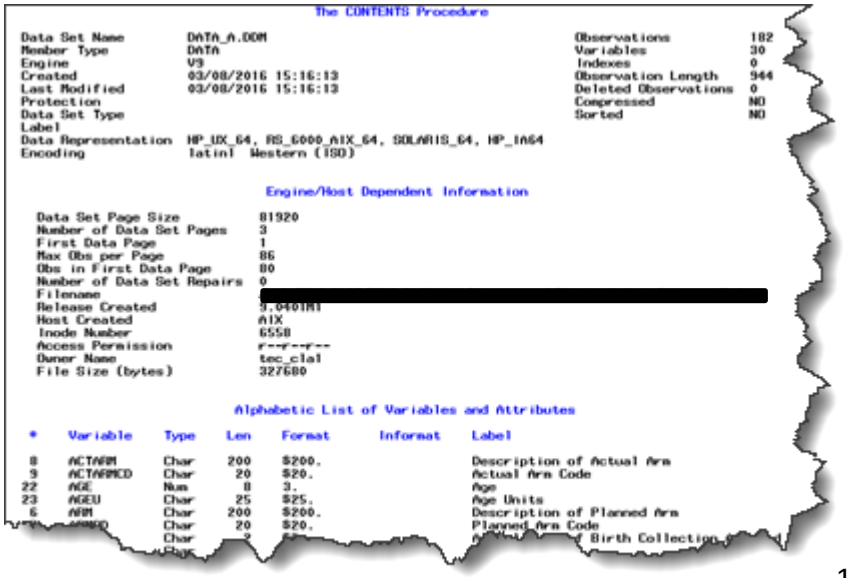
What is Metadata

- Data about data
- Information de scribes a variable
- A proc contents

DATA

	Unique Subject Identifier	Study Identifier	Study Site Identifier	Subject Identifier for the Study	Domain Abbreviation
1	CBYL719C2301_1200001	CBYL719C2301	1200	1200001	DM
2	CBYL719C2301_1200002	CBYL719C2301	1200	1200002	DM
3	CBYL719C2301_1200003	CBYL719C2301	1200	1200003	DM
4	CBYL719C2301_1200004	CBYL719C2301	1200	1200004	DM
5	CBYL719C2301_1201001	CBYL719C2301	1201	1201001	DM
6	CBYL719C2301_1300001	CBYL719C2301	1300	1300001	DM
7	CBYL719C2301_1300002	CBYL719C2301	1300	1300002	DM
8	CBYL719C2301_1301001	CBYL719C2301	1301	1301001	DM
9	CBYL719C2301_1301002	CBYL719C2301	1301	1301002	DM
10	CBYL719C2301_1301003	CBYL719C2301	1301	1301003	DM
11	CBYL719C2301_1301004	CBYL719C2301	1301	1301004	DM
12	CBYL719C2301_1301005	CBYL719C2301	1301	1301005	DM
13	CBYL719C2301_1302001	CBYL719C2301	1302	1302001	DM
14	CBYL719C2301_1303001	CBYL719C2301	1303	1303001	DM
15	CBYL719C2301_1305001	CBYL719C2301	1305	1305001	DM
16	CBYL719C2301_1306001	CBYL719C2301	1306	1306001	DM
[illegible]	[illegible]1306002	CBYL719C2301	1306	1306002	[illegible]
[illegible]	[illegible]1306003	CBYL[illegible]	[illegible]	1306003	[illegible]

METADATA



The CONTENTS Procedure

Data Set Name	DATA_A.DOM	Observations	182
Member Type	DATA	Variables	30
Engine	V9	Indexes	0
Created	03/08/2016 15:16:13	Observation Length	944
Last Modified	03/08/2016 15:16:13	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	HP_IDX_64, RS_6000_AIX_64, SOLARIS_64, HP_IA64		
Encoding	latin1 Western (ISO)		

Engine/Host Dependent Information

Data Set Page Size	81920
Number of Data Set Pages	3
First Data Page	1
Max Obs per Page	86
Obs in First Data Page	80
Number of Data Set Repairs	0
Filename	[redacted]
Release Created	9.0401M1
Host Created	AIX
Inode Number	6558
Access Permission	r--r--r--
Owner Name	tec_cla1
File Size (bytes)	327680

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat	Label
8	ACTARM	Char	200	$200.		Description of Actual Arm
9	ACTARMCD	Char	20	$20.		Actual Arm Code
22	AGE	Num	8	3.		Age
23	AGEU	Char	25	$25.		Age Units
6	ARM	Char	200	$200.		Description of Planned Arm
[illegible]	[illegible]ARMCD	Char	20	$20.		Planned Arm Code
[illegible]	[illegible]	Char	[illegible]	[illegible]		[illegible] of Birth Collection [illegible]
[illegible]	[illegible]	Char	[illegible]	[illegible]		[illegible]

Novartis' domains Overview

Clinical Data Element View

MDR_PRD (Project: Standard Project) (Work Area) Help Switch Sign In Feedbac

View Administer Personalize Favorites

View >>Browse >>MDR Interface Views>>SCR Interface >>CLINICAL_DATA_ELEMENTS Number Of Rows: 20

Display Values View Delivery Options

CLINICAL_DATA_ELEMENTS Rows 1..20 of 17514 1 21 41 61 81 101 121 141 161 181 -> Request

View	1 DATA_DOMAIN	3 DATA_ELEMENT	SCREEN_AND_SAS_LABEL	GENERIC_TYPE	LENGTH	CONTROLLED_TERMINOLOGY	SDTM	2 DOMAIN_TYPE	SDTM_DOMAIN	CRF_DOMAIN	OC_VARIABLE_NAME	SAS_VARIA
	AA	AACAT	Category for Adjudication	text	70	AACAT	Yes	Events	AA	N/A	N/A	AACAT
	AA	AAEVAL	Evaluator	text	40	EVAL	Yes	Events	AA	N/A	N/A	AAEVAL
	AA	AAGRPID	Group ID	text	200	N/A	Yes	Events	AA	N/A	N/A	AAGRPID
	AA	AAOCCUR	Adjudication Occurrence	text	10	NOYES	Yes	Events	AA	N/A	N/A	AAOCCUR
	AA	AAPRESP	Adjudication Prespecified	text	10	NOYES	Yes	Events	AA	N/A	N/A	AAPRESP
	AA	AAREFID	Reference ID	text	30	N/A	Yes	Events	AA	N/A	N/A	AAREFID
	AA	AAREL	Causality	text	40	AAREL	Yes	Events	AA	N/A	N/A	AAREL
	AA	AASCAT	Subcategory for Adjudication	text	70	AASCAT	Yes	Events	AA	N/A	N/A	AASCAT
	AA	AASTDT	Start Date of Event	Date	8	N/A	No	Events	N/A	AA	AASTDT	AASTDT
	AA	AASTDTC	Start Date Time of Event	datetime	20	N/A	Yes	Events	AA	N/A	N/A	AASTDTC
	AA	AASTDTF	F-Start Date of Event	text	200	N/A	No	Events	N/A	AA	AASTDT_FUL	AASTDTF

Derivations/Imputation View

MDR_PRD (Project: Standard Project) (Work Area) Help Switch Sign In Feed

View Administer Personalize Favorites

View >>Browse >>MDR Interface Views>>SCR Interface>>DERIVATIONS_IMPUTATIONS Number Of Rows: 20

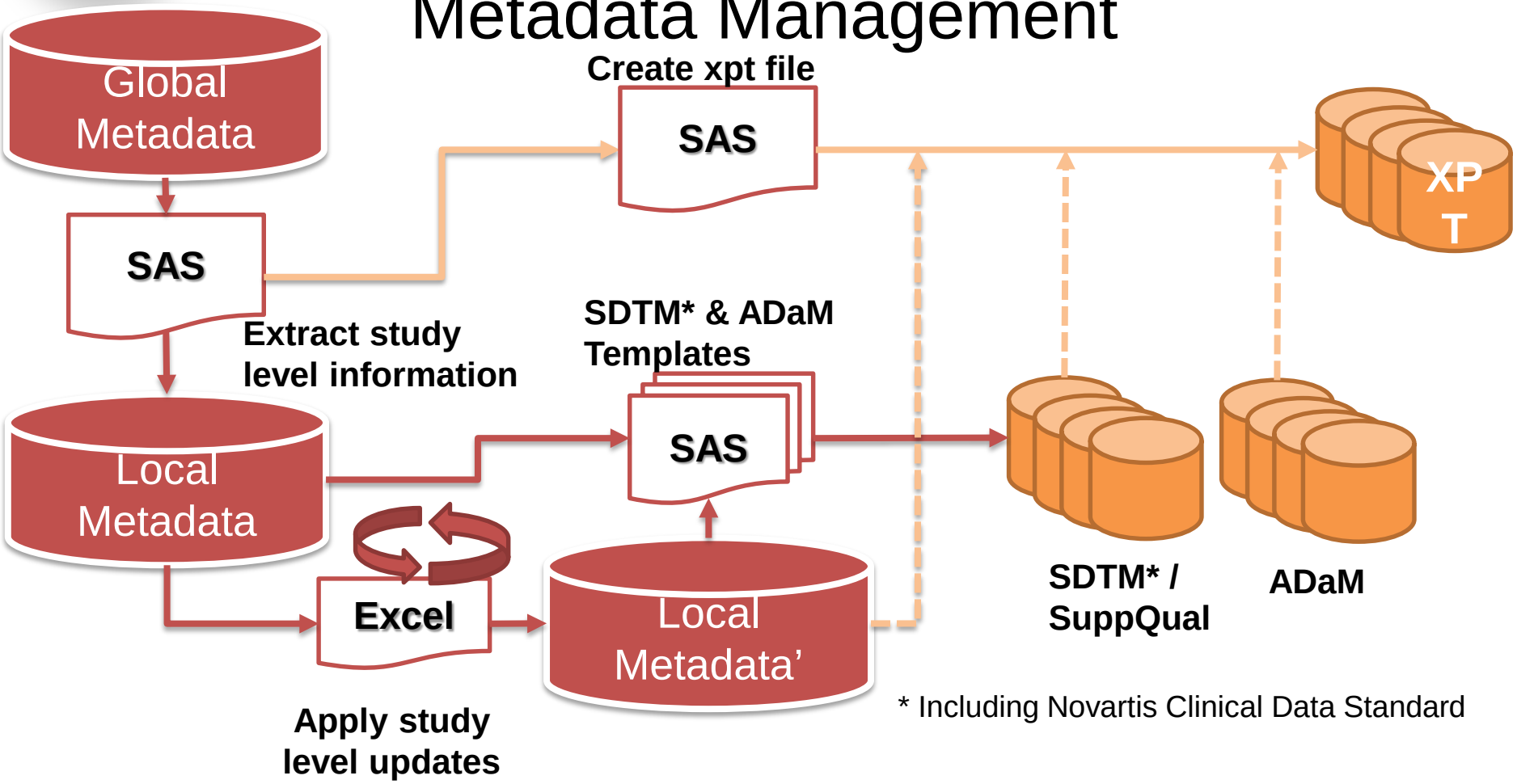
Display Values View Delivery Options

DERIVATIONS_IMPUTATIONS Rows 1..20 of 3625 1 21 41 61 81 101 121 141 161 181 -> Request

View	DATA_DOMAIN	LOGICAL_NAME	DERIVATION_IMPUTATION_NAME	EXECUTION_AREA	OUT_DATA_ELEMENT	FULL_QUAL_DE_NAME	KEEP_FINAL_ELEMENT	CONTRIBUTING_DATA_ELEMENTS	CONTRIBUTING_
	AA		COPY_ELEMENT	CDR-Stage 1	AACAT	S1.AA.AACAT	Yes	S1.AA.CATAA	N/A
	AA		COPY_ELEMENT	CDR-Stage 1	AAEVAL	S1.AA.AAEVAL	Yes	S1.AA.EVALAA	N/A
	AA		COPY_ELEMENT	CDR-Stage 1	AAGRPID	S1.AA.AAGRPID	Yes	S1.AA.GRPIDAA	N/A

Actual contents

Metadata Management



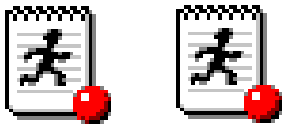
* Including Novartis Clinical Data Standard

Novartis SDTM/ADaM process overview

Metadata Management

- More detail about template program generation process.

Standard macro library



Data domain type (Long text)	Domain name	SAS variable name	Codel used data elem	Derivation contributing elements	Data Element definition	Derivation Name	
SCR DERIVED DATA - DR	LB	LBCAT	LBCAT	DR.LB.LPARM|SR.TESTS.PAR	pe of draw / category / panel name ed to define a category of related cords.	MERGE_ELEMENT	rge input datasets using belo ables: ~{newline DR.LB.LPARM~{newline SR.TESTS.PARM~{newline ~{newline 1}LBCAT= SR.TES
SCR DERIVED DATA - DR	LB	LBCBYR	N/A	RP.LB.LBCBYR	bject birth year used to reconcile rd-party data during loading	COPY_ELEMENT	CBYR = RP.LB.LBCBYR
SCR DERIVED DATA - DR	LB	LBCLSIG	CLSIG	RP.LB.LBCLSIG	nical Significance	COPY_ELEMENT	CLSIG = RP.LB.LBCLSIG
SCR DERIVED DATA - DR	LB	LBCOM	N/A	RP.LB.LBCOM	boratory comments from the central l ovider	COPY_ELEMENT	COM = RP.LB.LBCOM
SCR DERIVED DATA - DR	LB	LBCSEX	SEX	RP.LB.LBCSEX	bject sex used to reconcile third-part ta during loading	COPY_ELEMENT	CSEX = RP.LB.LBCSEX
SCR DERIVED DATA - DR	LB	LBDAT	N/A	RP.LB.LBDAT	art) date of specimen collection	COPY_ELEMENT	DAT = RP.LB.LBDAT
SCR DERIVED DATA - DR	LB	LBDTC	N/A	DR.LB.LBSDT|DR.LB.LBSTM	te of sample collection	DTC	nvert date variable DR.LB.LB d time variable DR.LB.LBST aracter ISO 8601 format.
SCR DERIVED DATA - DR	LB	LBDTN	N/A	DR.LB.LBDTC	te of sample collection	DTN	nvert DR.LB.LBDTC to nume 01 format.
SCR DERIVED DATA - DR	LB	LBDY			en Collection		om a g y 1, the re is no ter date po ortion o DR.L TN the) - (dat ither RFST then LBDY

Contributing Element

Derivation Name

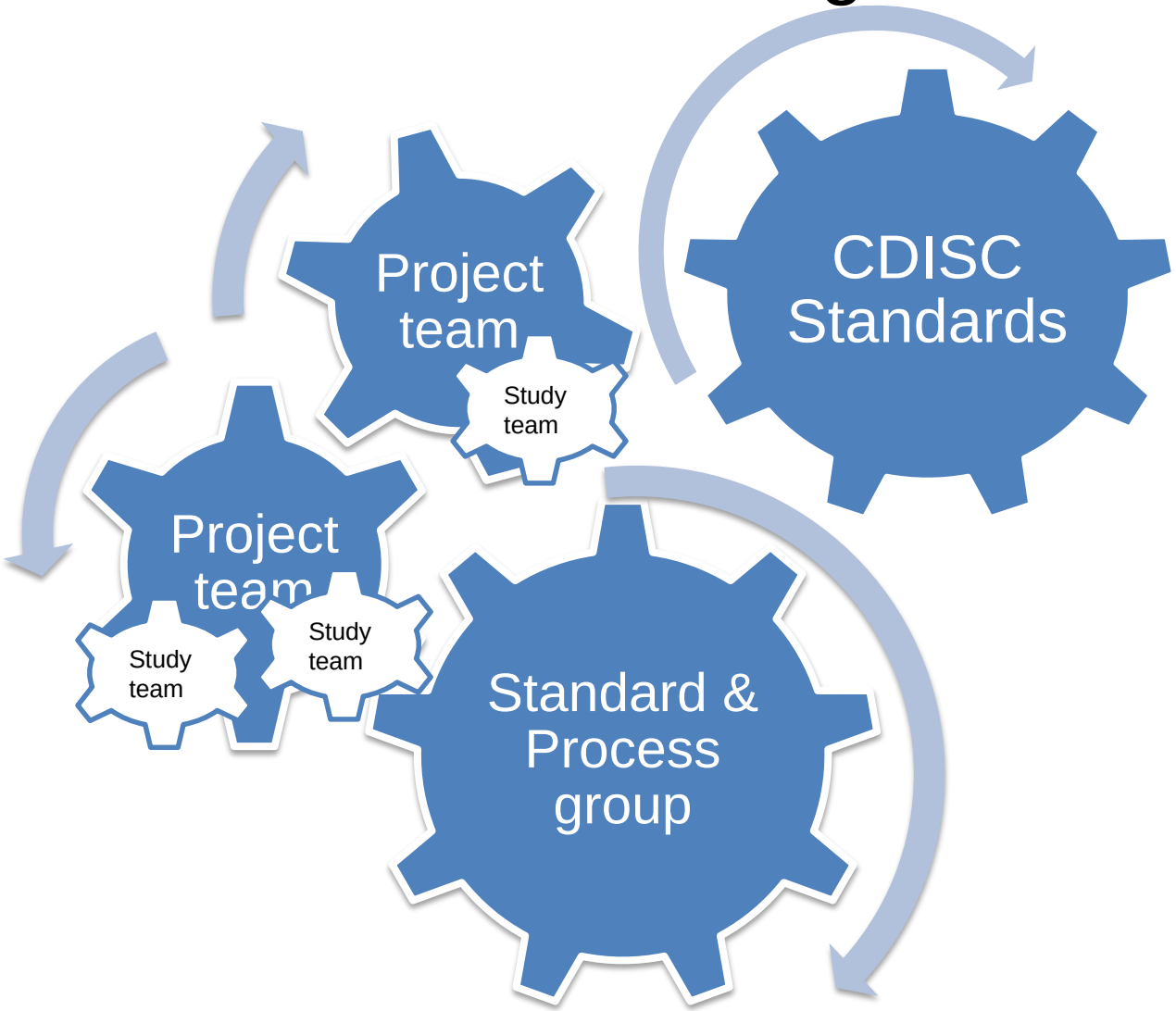
Global / Local Metadata

```
/* Keep statement: Keep variable from M
data _tmplb1;
   %variables_att(&copy_element  )
   *** Read Master input file    ***;
   set &Master_input_file;

    *** Map/setup common variable ***;
   *** Keep only registered variable **
   keep &copy_element &temp_copy_elemen
run;

/* Derive LBBLFL using create_constant */
/* Contributing element(s): N/A */
%create_constant(inds = _tmplb1 ,outds = _tmplb2 ,target = LBBLFL);

/* Derive LBREASND using create_constant */
/* Contributing element(s): N/A */
%create_constant(inds = _tmplb2 ,outds = _tmplb3 ,target = LBREASND);
```

- sdtm_plus_5_01_dta.sas
- sdtm_plus_5_01_dti.sas
- sdtm_plus_5_01_dtv.sas
- sdtm_plus_5_01_dy1.sas
- sdtm_plus_5_01_dy5.sas
- sdtm_plus_5_02_dsv.sas
- sdtm_plus_5_03_dcm.sas
- sdtm_plus_5_03_ddm.sas
- sdtm_plus_5_03_dds.sas
- sdtm_plus_5_03_dex.sas
- sdtm_plus_5_03_die.sas
- sdtm_plus_5_03_dsc.sas

Template programs

Metadata Management



Key Benefits and challenges

Benefits	Challenges
Controlled and Audit trailed environment • **Centralize** all changes made to metadata • **Traceability** of changes	Mapping metadata to the **appropriate version** of the SDTM IG, CDISC Controlled terminology, etc.
Efficacy gained in defining and **maintenance of metadata** e.g. Codelists, lab reference tables, etc	While some of the Controlled terminology values are mapped to NCI values in SDTM/ADaM, it is an on-going **effort to keep our Metadata in alignment with CDISC standards**
Instant access to all end users to browse matadata	**Project level metadata management** is also challenge

Simplification and standardization

Standard governance drives consistency and quality

SDTM /ADAM / Define-XML

- Template programs
- Comprehensive DEFINE Review

aCRF

- Standard CRF
- Annotation tool

SDRG/ADRG

- Template and examples
- Training material and check list

SAS ユーザー総会 2017

Part 2.

Creation of SDTM Annotated CRFs

Advantage of automation tool

Automation

- Re-usable
- Keep quality and consistency

Manual

- Human errors
- Time consuming

Overview of annotation tool

Standard Governance

Develop standard & annotated CRF

Centralize information in Metadata & annotation database

To be used within Individual studies

- Develop global / Therapeutic area standard CRF page
- SDTM annotation is manually setup per individual page.

- Domain / annotation information are loaded to metadata and annotation database respectively.

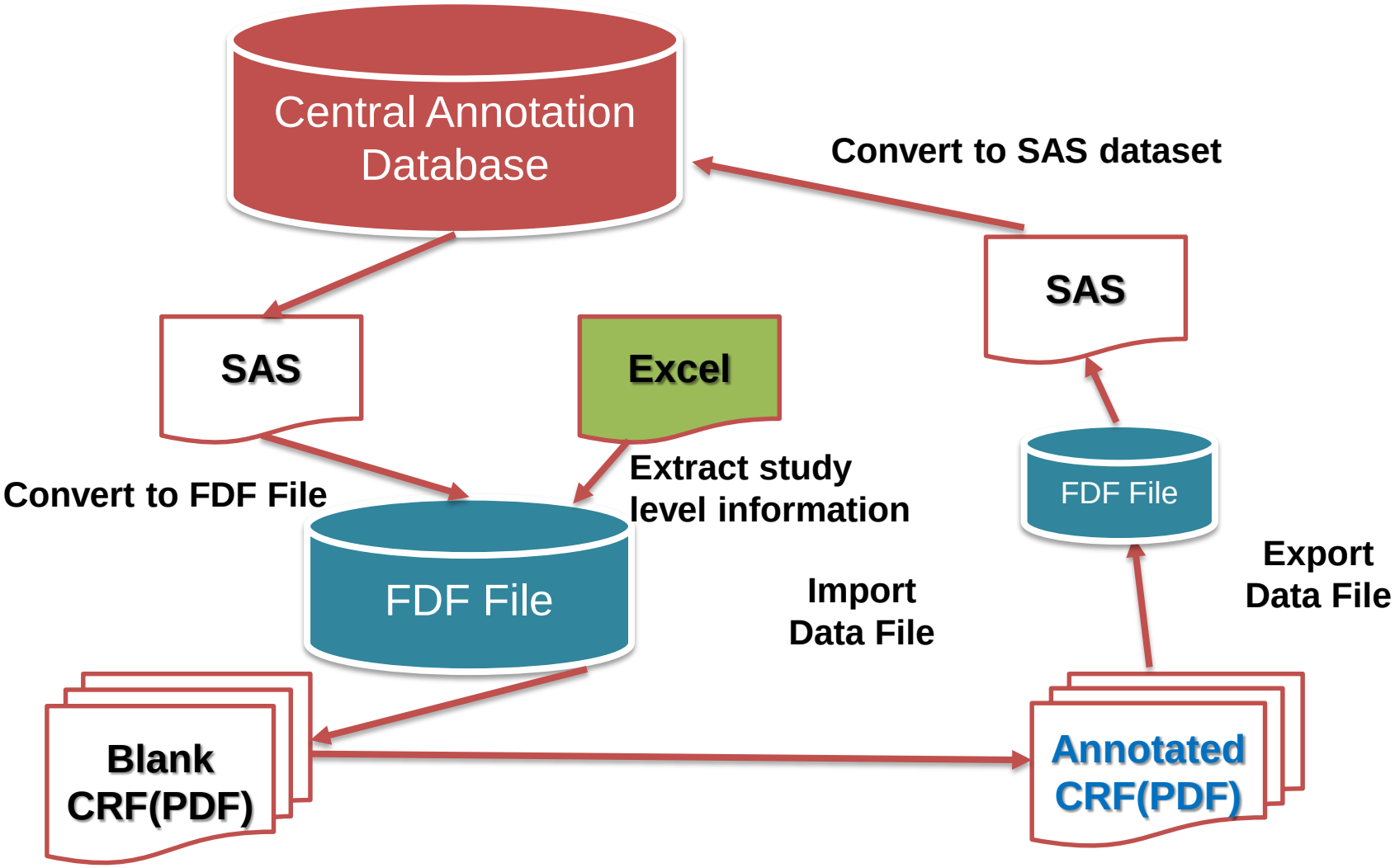
- Get standard information for SDTM annotation from global metadata

ADOBE ACROBAT **FDF** FILE

FDF stands for "Forms Data Format". FDF is a file format for representing form data and annotations that are contained in a PDF form

%FDF-1.2
%âãÏÓ
1 0 boy
<</FDF<</Annots[2 0 R]/F(/C/SUGI/SUGI2017_example_CRF_DM.pdf)/ID[<B8EE73C02F4D874C93AAE44ACF8C2212><C2D8B7C87E8FCE47B3CE92FB8082505F>]/UF(/C/PharmaSUG2014/PharmaSUG2014_example_CRF_DM.pdf)>>/Type/Catalog>>
endobj
2 0 obj
<</BS 3 0 R/C[0.0 1.0 1]/Contents(SEX)/DA(0 0 0 rg /Arial 10)/DS(font: Arial,sans-serif 10.0pt; text-align:left; color:#000000)/M(D:20140406095629-04'00')/NM(7c71d692-03dc-45c4-abd9-4bf182d12a4f)/Page 0/RC(<?xml version="1.0"?><body xmlns="http://www.w3.org/1999/xhtml" xmlns:xfa="http://www.xfa.org/schema/xfa-data/1.0/" xfa:APIVersion="Acrobat:10.1.5" xfa:spec="2.0.2" style="font-size:10.0pt;text-align:left;color:#000000;font-weight:normal;font-style:norma¥
l;font-family:Arial,sans-serif;font-stretch:normal"> SEX </body>)/Rect[407.868 324.866 431.868 338.866]/Subj(TextBox)/Subtype/FreeText/T(NCDS)/Type/Annot>>
endobj
trailer
<</Root 1 0 R>>
%%EOF

CRF Annotation Tool – Process Flow



CRF Annotation Tool
Central Annotation Database

NOVARTIS
Study CTST001A0004

ID
Center No. Subject No.

WEEK 1
Visit Name

Demography

Before entering data, please check CCG for information regarding privacy laws for your country.

Year of birth
Age at first informed consent YEARS
Sex ☐ Male ☐ Female
Is subject of child bearing status
Race
Specify race
Ethnicity

DMG02

DMG02

Epoch Treatment Document Number Subevent# Page Number 141
Verified ☐ Approved ☐ Locked ☐ Frozen ☐ Page Name DM Status Version 1

17	SCORRES when SCTESTCD=SRCRF	527.422 218.31 693.821 230.32
18	SCORRES when SCTESTCD=SBJINIT	387.831 340.796 560.515 355.918
19	SCORRES when SCTESTCD=CHDBER	527.951 299.432 718.581 313.002
20	STUDYID	159.985 554.563 206.077 568.163

crfid	comtype	domlvl	domain
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	1	DM
DMG02	Free Text	2	SC
DMG02	Free Text	2	SC
DMG02	Free Text	2	SC
DMG02	Free Text	2	SC
DMG02	Free Text	2	SC
DMG02	Free Text	2	SC
FHI01	Free Text	1	SC

	1	2
1	CRFID	PAGE
2	SVG01	1
3	DS4G02	2
4	XIA03	3
5	DMG02	4
6	IEG03	5
7	MHG02	6

CRF Annotation Tool – DEMO

Name	Size
master_annotation.sas7bdat	499712

Name	Size
BlankCRF_Study2_Demo.pdf	300317
aCRF_Study1_Demo.pdf	1505181
crfid_pageno.xls	31232
extract_annotation.sas	23561
create_fdf_file.sas	17108

- **master_annotation.sas7bdat**: Central Annotation Database
- **BlankCRF_Study2_Demo.pdf**: Blank CRFs to annotate for Study 2
- **crfid_pageno.xls**: CRF IDs/Page No specification
- **create_fdf_file.sas**: SAS program to create study specific FDF file
- **extraction_annotation.sas**: SAS program to create Secondary Annotation Database(s)

Rational ClearCase Explorer - iijimya1_view (V:\iijimya1_view\CSCRTRUS\CSCRTRUS0016\report\pkpd\pgm_02)

File View Go Tools Environment Help

Name	Size	Kind	Modified	Version
BlankCRF_Study2_Demo.pdf	300317	File El...	02/10/2014 21:06:56	\main\1
aCRF_Study1_Demo.pdf	1505181	File El...	02/10/2014 21:31:45	\main\2
crfid_pageno.xls	31232	File El...	02/10/2014 21:37:33	\main\1
extract_annotation.sas	23561	File El...	28/05/2015 10:33:26	\main\2
create_fdf_file.sas	17108	File El...	27/01/2017 20:25:09	\main\4
Study1.fdf	48076	View-...	06/07/2017 12:17:21	
study1.sas7bdat	2236416	View-...	06/07/2017 12:21:21	
outfdf_demo1.fdf	42675	View-...	06/07/2017 12:30:27	
outfdf_demo2.fdf	58788	View-...	06/07/2017 12:33:30	

Your selection

About | Uses

Name: BlankCRF_Study2_Demo.pdf
View Tag: iijimya1_view
View Path: V:\iijimya1_view\CSCRTRUS\CSCRTRUS0016\report\pkpd\pgm_02
Kind: File Element Version
Modified: 02/10/2014 21:06:56
Version: \main\1
Rule: \main\LATEST

BlankCRF_Study2_Demo.pdf is a versioned file under ClearCase source control.

acrf.pdf - Adobe Acrobat

File Edit View Window Help Novartis Signing

Create

55 / 57 98.9%

Tools Comment

Bookmarks

- Domains
- Visits

VS = Vital Signs

NOVARTIS STUDYID

Study [redacted]

ID USUBJID

Center No. Subject No.

VISIT

Visit Name

Mark if not done ☐ [NOT SUBMITTED]

VSCAT = GENERAL

Vital Signs

Any clinically significant findings present prior to signing informed consent should be recorded on the Medical History page.
Any new or worsening clinically significant finding noted since signing informed consent should be recorded on the Adverse Events page.

Date of assessment VSDTC
DD-MON-YYYY

VSTEST	Result	Unit (VSORRESU)	
HEIGHT		Centimeter	VSORRES when VSTESTCD=HEIGHT
WEIGHT		Kilogram	VSORRES when VSTESTCD=WEIGHT

VSORRES when VSTESTCD=SYSBP

VSORRES when VSTESTCD=DIABP

VSPOS SITTING

Pulse VSORRESU BEATS/MIN

VSORRES when VSTESTCD=PULSE

Blood pressure Systolic / Diastolic mmHg VSORRESU

WBG02$

Key Benefits and challenges

- Benefit
  - This process not only helps in reducing the cycle time by a great extent, but also **helps maintain consistency of the annotations across different studies**.
  - FDF file can be easily handled in SAS datasets and easy to extend its usage for other tool e.g. Define-XML.
- Challenge
  - Some manual task such as page mapping, overall QC is still necessary.

Simplification and standardization

Standard governance drives consistency and quality

SDTM /ADAM / Define-XML

- Template programs
- Comprehensive DEFINE Review

aCRF

- Standard CRF
- Annotation tool

SDRG/ADRG

- Template and examples
- Training material and check list

SAS ユーザー総会 2017

Part 3.

SAS Application to Automate a Comprehensive Review of DEFINE and All of Its Components

Problem Statements

The DEFINE package is a large electronic document comprised of many different but interrelated components

The define.xml acts as a road map where embedded hyperlinks allow reviewers to easily navigate between components and understand how data were collected/derived for analysis purpose

It is a massive undertaking to review all the components to ensure **accuracy, completeness, and consistency** once the DEFINE package is created, especially done manually

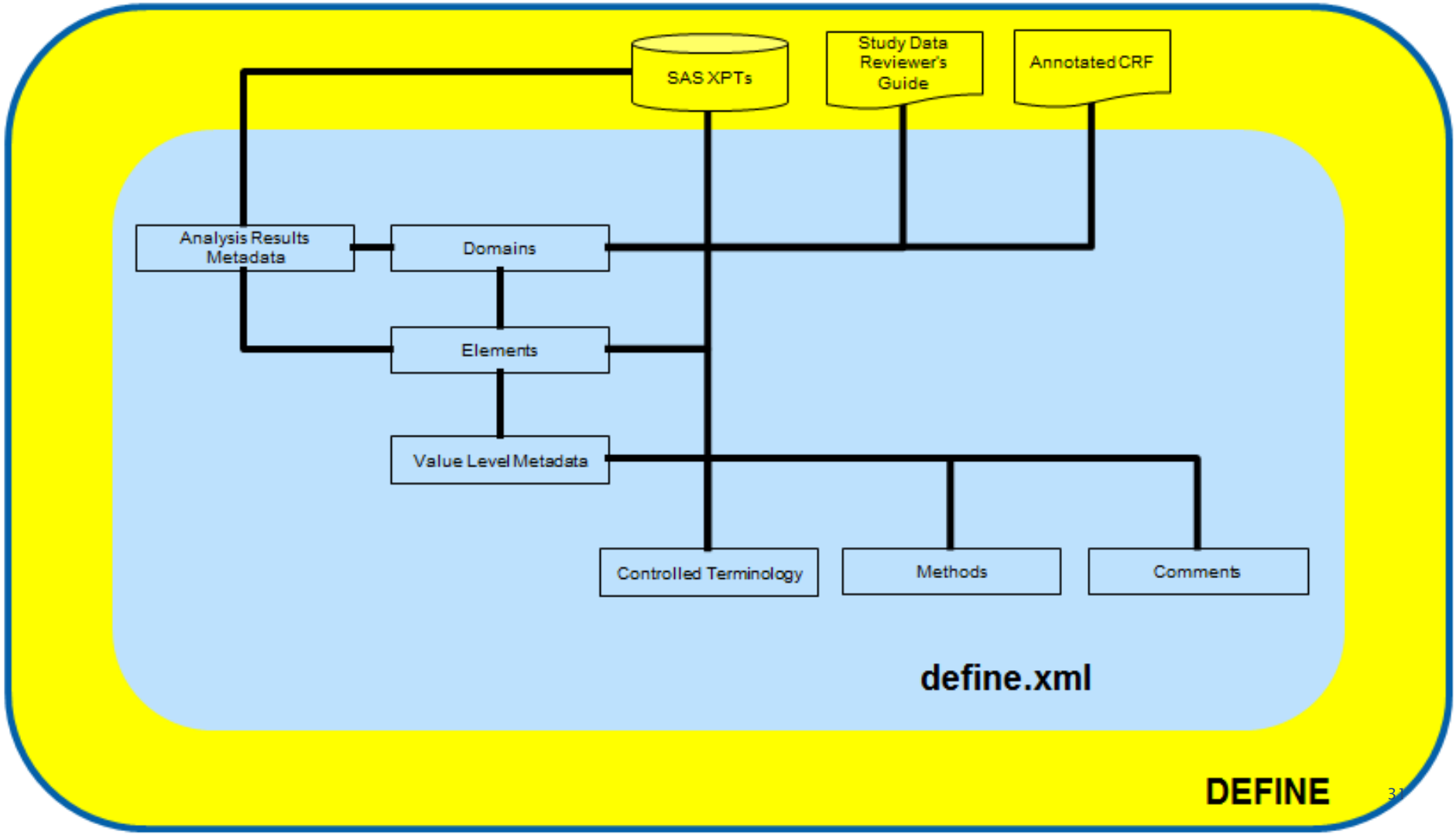
Automation Can Be Achieved once You

are familiar with the DEFINE and all of its distinct but interrelated components and sections

fully understand the scope of what a thorough review of the DEFINE entails

construct a data structure capable of consolidating disparate metadata from each of the DEFINE components

Interrelated DEFINE Components and Structure



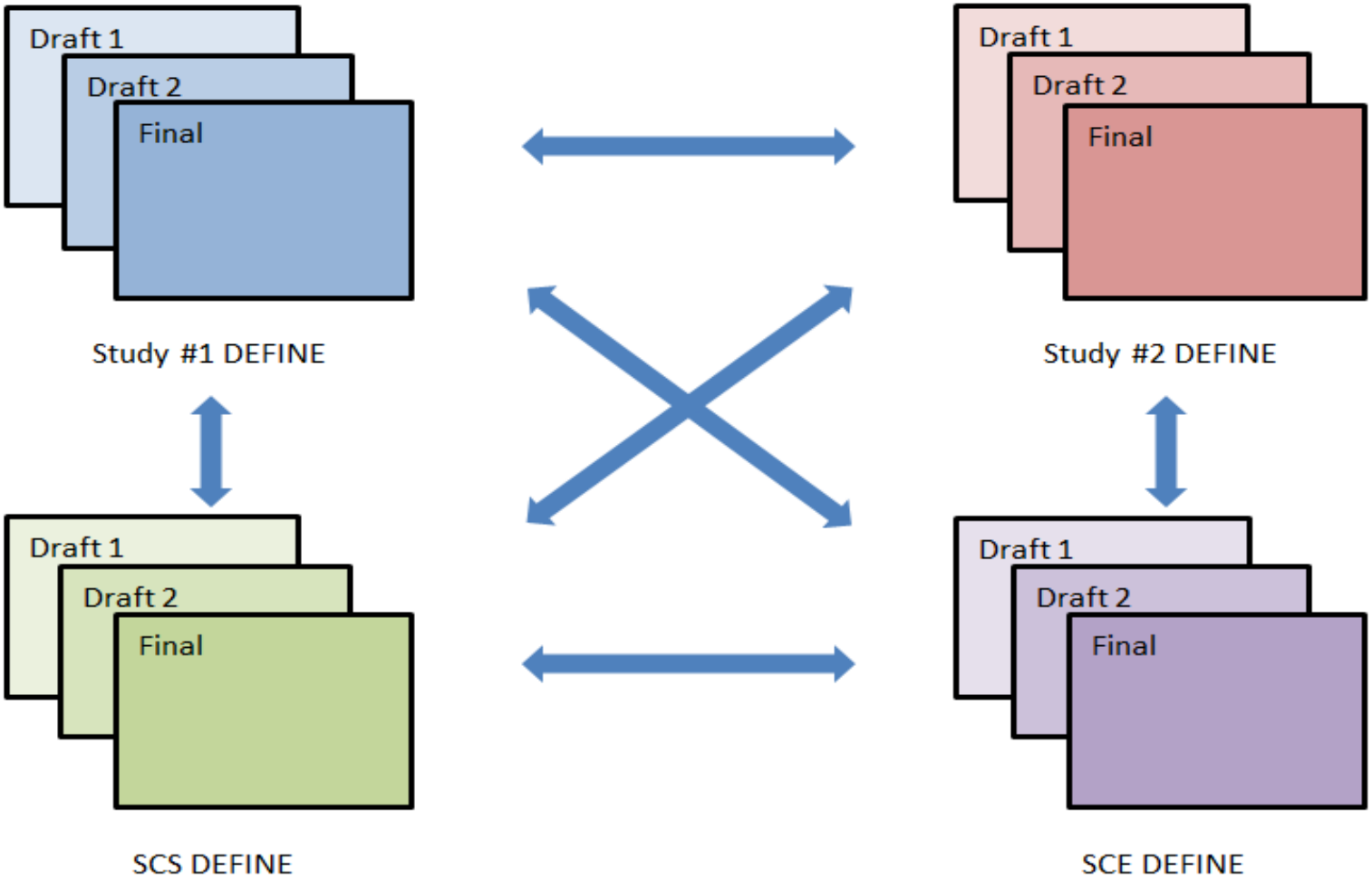
Scope of a Thorough Review within a Single DEFINE

Hundreds of data elements and values **must be** reviewed for **accuracy, consistency, and traceability**

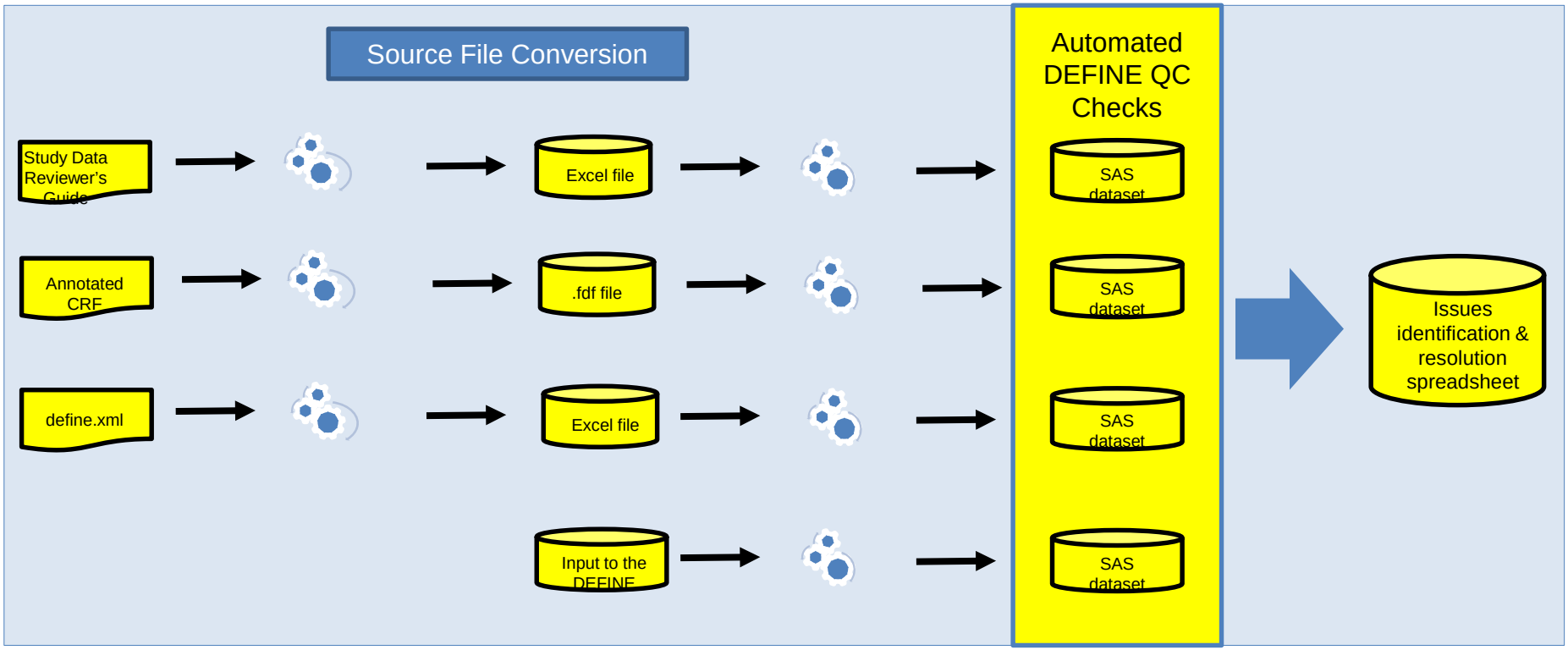
Each of these data points must be re-verified each time a new draft is generated

- Consistency and traceability within define.xml
- Define.xml consistent with **Input Specifications**
- **Define.xml, annotated CRFs, Data Reviewer's Guide consistent with XPT files**
- Annotated CRFs and Data Reviewer's Guide consistent with define.xml

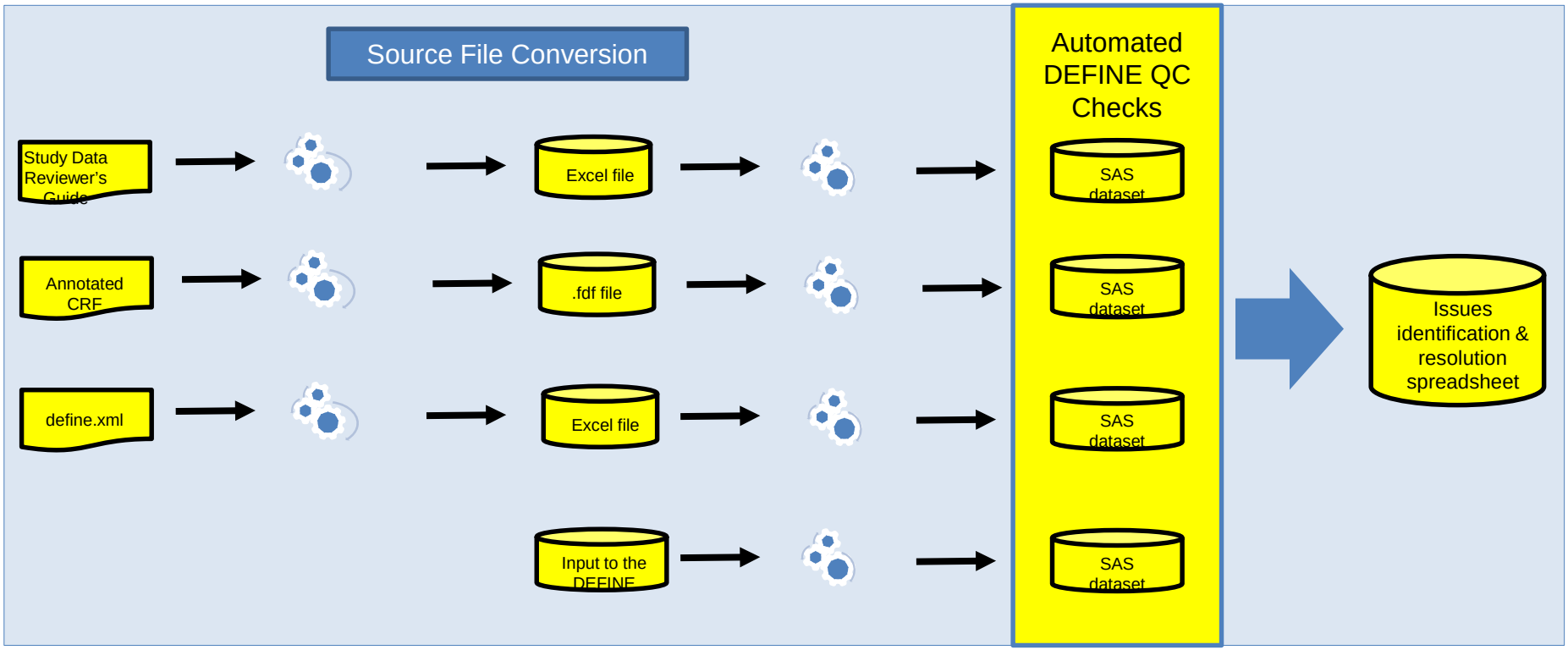
Interrelated DEFINEs within a Submission



Automated DEFINE Review Tool Overview



Automated DEFINE Review Tool Overview



Issues Identification & Resolution Spreadsheet

- All QC issues consolidated into a single Excel file
- Separate tabs used to **store related issues**
- Standard columns / issue messages to **facilitate resolution**

1	Domain	Issue description	DEFINE value	COMPARISON value	Action
7	DS	Different domain keys	STUDYID, USUBJID, DSDECOD, DSCAT, DSSCAT, DSSTDTC, EPOCH	STUDYID, USUBJID, DSDECOD, DSCAT, DSSCAT, EPOCH,DSSTDTC,	
8	DS	Domain (DS) not sorted by keys (STUDYID USUBJID DSDECOD DSCAT DSSCAT DSSTDTC EPOCH)			
9	EX	Different domain keys	STUDYID, USUBJID, EXTRT, EXSTDTC, EXSEQ	STUDYID, USUBJID, EXTRT, EXSTDTC, EPOCH, EXDOSTOT, EXENDTC, EXSEQ	
22	SUPPFAZD	Different domain (xpt) label	Supplemental Qualifiers for SUPPFAZD	Supplemental Qualifiers for FAZD	

Domain checks | Element checks | VLM checks | Controlled Terminology checks | Data Reviewer's Guide checks | Annotated CRF checks

Prior to Implementing the Tool

Manual task

- **Line-by-line visual compare** of all components and data points
- Basic SAS code checks
- Manual documentation of findings in various formats
- Overlapping and/or blind spots by the reviewers

Inefficient review

- **Long time**
- Incomplete
- **Inconsistent**
- More iterations of review

After Implementing the Tool

Spare resources from lengthy, tedious, and repeated manual reviews

Focus valuable **resources on tasks that require a higher level** of functional expertise and knowledge

Eliminate incomplete and inconsistent findings due to factors such as level of experience, fatigue, and time constraints

Improve quality and efficiency of review

What the Tool Does Not Do

- Substitute the review of derivations in Input Specifications and defile.xml, which ensures that the description is clear, complete, and accurate
- Substitute the review of the contents in Data Reviewer's Guide
- Replace the conformance check
- Consistency check within the submission

Project-level Input Specifications

Well-conceived project-level input specifications (a single file which applies to all studies contained in the submission) play a pivotal role

Must be accurate and concise, and **consolidate project-level** data point attributes	Allow for the **existence** of study specific derivations and additions where necessary	Can be used to drive the code used to both create the tabulation and analysis datasets and produce and QC the DEFINE

Summary

Automation eliminates incomplete and inconsistent findings, **reduces the burden on programming and statistical resources**, and greatly improves the quality of the review of the DEFINE

Organized metadata allow us to easily create some kind of tools like the ACRF tool or the DEFINE review tool by SAS and it **could improve our routine work drastically**

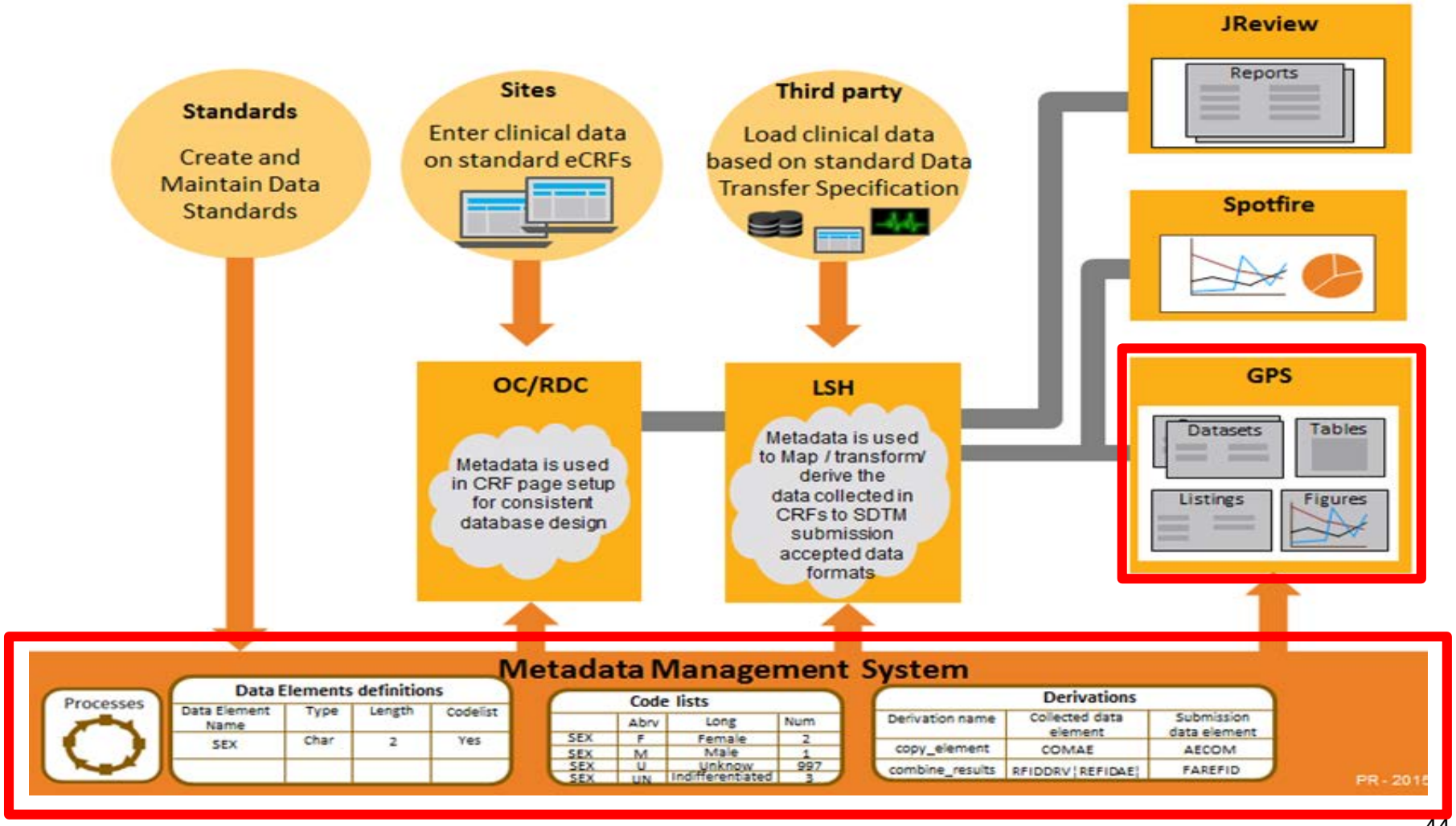
Standardization and Simplification are important elements **for eSubmission**.

Reference

- Archana Bhaskaran, “Implementation and Management of Clinical Metadata in Novartis Metadata Repository (MDR)”, CDISC International Interchange 2015
- Walter Hufford, "Automating Production of the blankcrf.pdf", PharmaSUG 2014 - Paper CC21
- Walter Hufford, Vincent Guo, Mijun Hu, ” SAS Application to Automate a Comprehensive Review of DEFINE and All of Its Components”, PharmaSUG 2017

Back up

Metadata Management



Why Do SDTM Annotations?

- SDTM annotated CRFs is one of the key CDISC format submission deliverables
- This is a blank CRF annotations that document the location of the data with the corresponding names of the datasets and the names of those variables included in the submitted SDTM datasets
- Annotations are meant to help the PMDA/FDA reviewer find the origin of data variables included in the submitted SDTM datasets
- Annotations should be text-based and searchable using standard PDF viewers

CRF Annotation Tool
SAS Programs

- Create_fdf_file.sas (SAS dataset to FDF file)
  - Inputs:
    - Central Annotation Database(SAS dataset)
    - CRF IDs/Page No in a Excel spreadsheet
  - Output: Study specific FDF file that could be imported to the study blank CRFs

- Extract_annotation.sas (FDF file to SAS dataset)
  - Inputs: FDF files exported from already annotated CRF pages (FDF file)
  - Output: Central Annotation Database (SAS dataset)

Savings after Automated DEFINE Review Tool

- An estimate of about **2/3 time reduction**
- **Savings are multiplied** where submissions consist of multiple studies
- Savings remain significant even after considering initial investment in the development of the tool
- The more the tool is used, the more the **savings increase**