



PhUSE SDE Basel

Metadata driven process from global definition to analysis specifics

Aurélien Guillouche, Marjorie Nowak

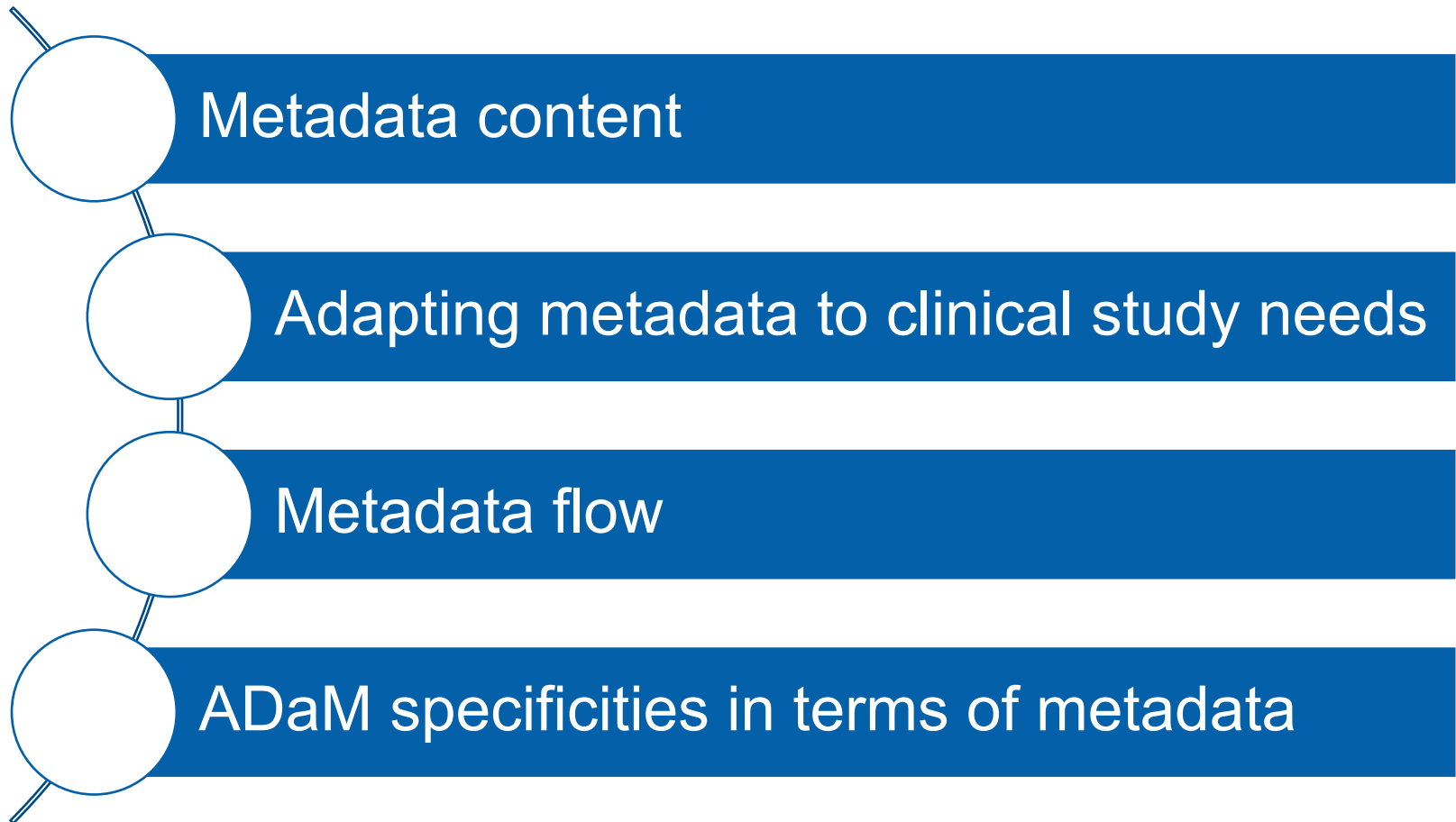
PhUSE Single Day Event
Novartis Campus Basel
May 16th, 2017

Disclaimer

The opinions expressed in this presentation and on the following slides are solely those of the presenters and not necessarily those of Novartis.

Novartis does not guarantee the accuracy or reliability of the information provided herein.

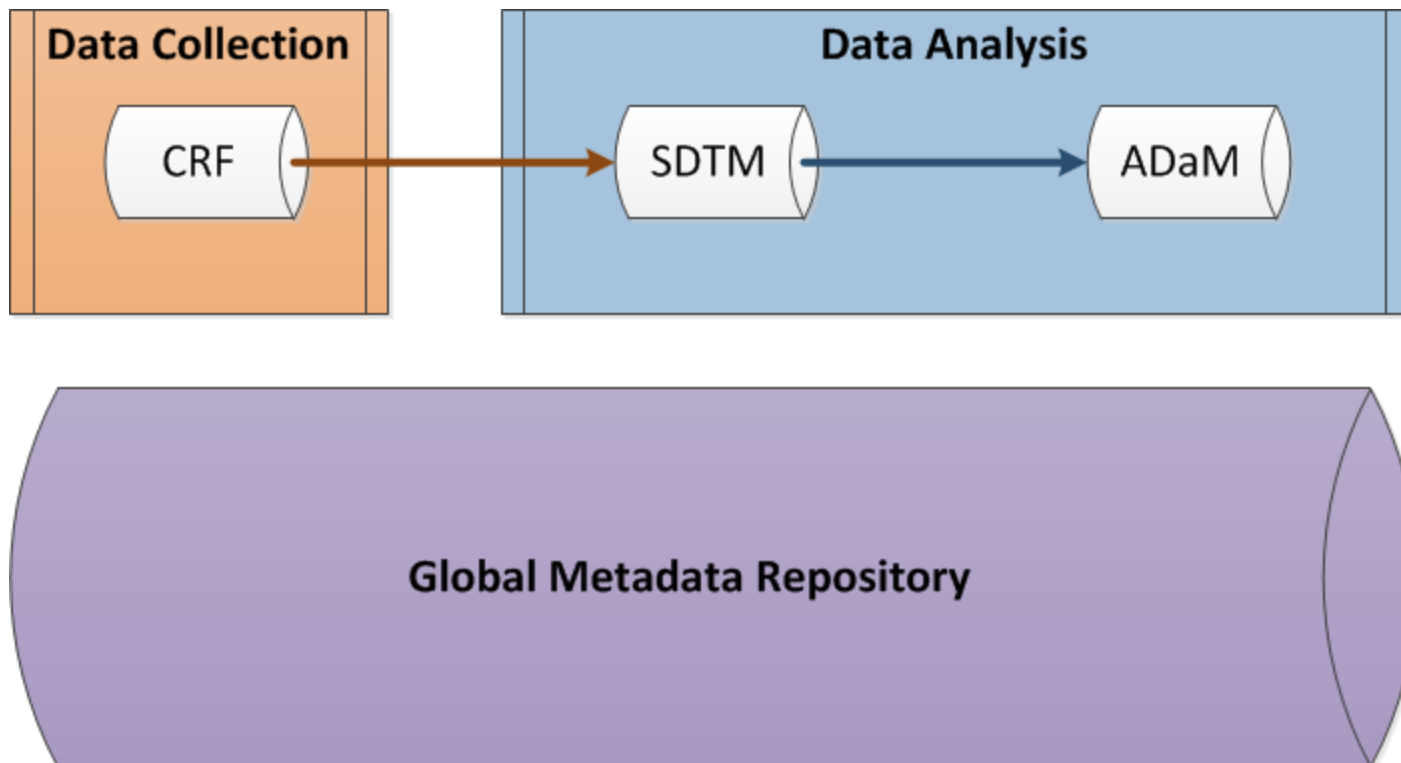
Agenda



Metadata Content

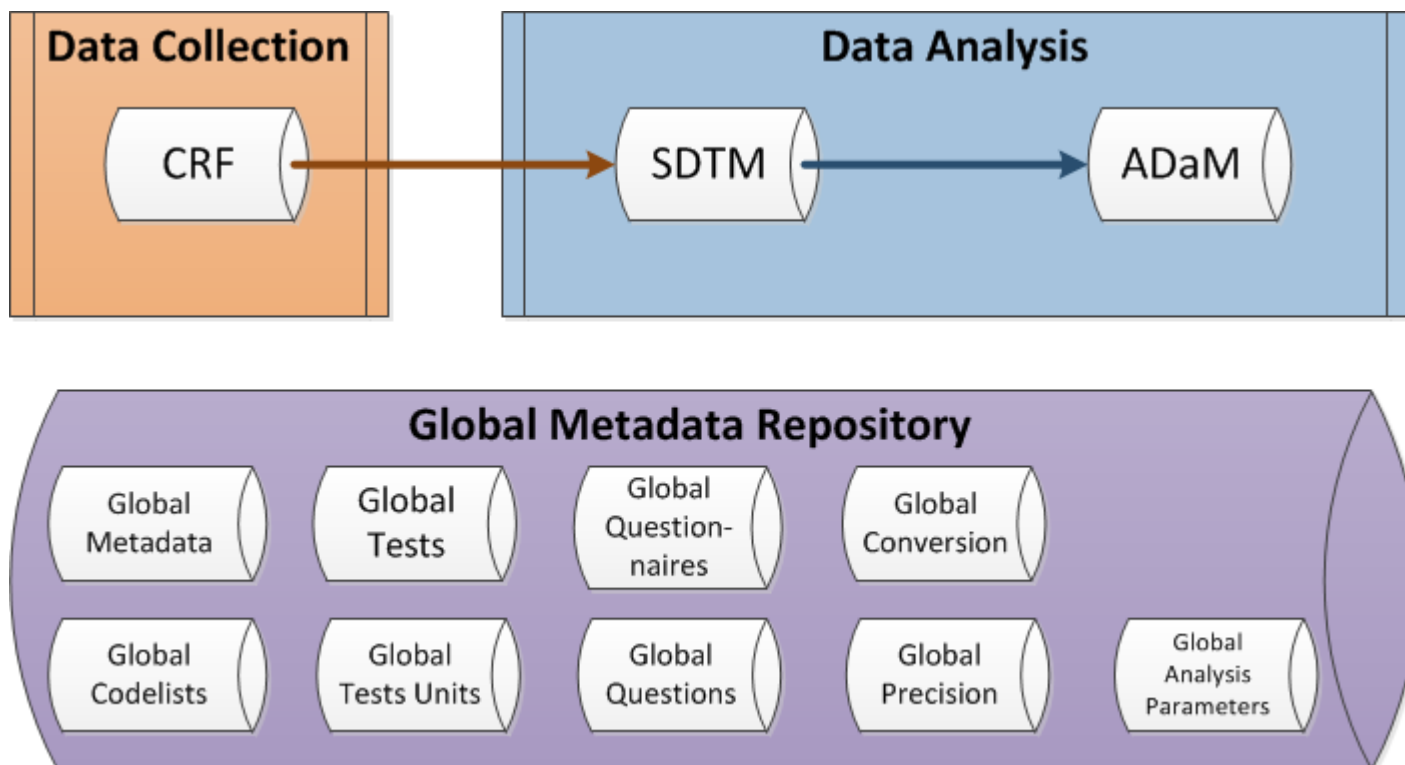
Metadata Content

Overview : from collection to analysis



Metadata Content

Overview : Global Metadata Repository



Metadata Content

What is in the Global Metadata?

Proc Contents

Domain

Element

Type

Length

...

Additional information

Derivation

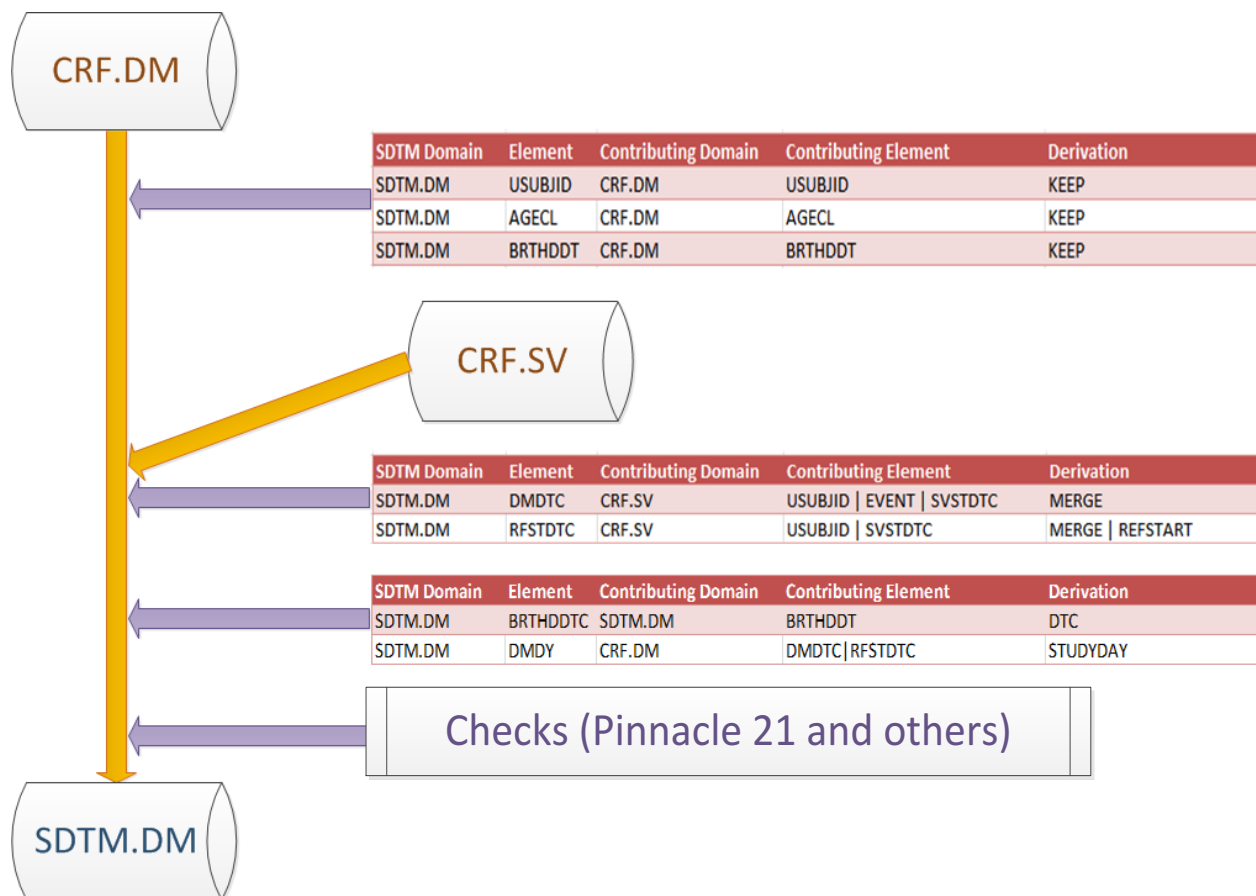
Contributing
Elements

Permissible
Element

...

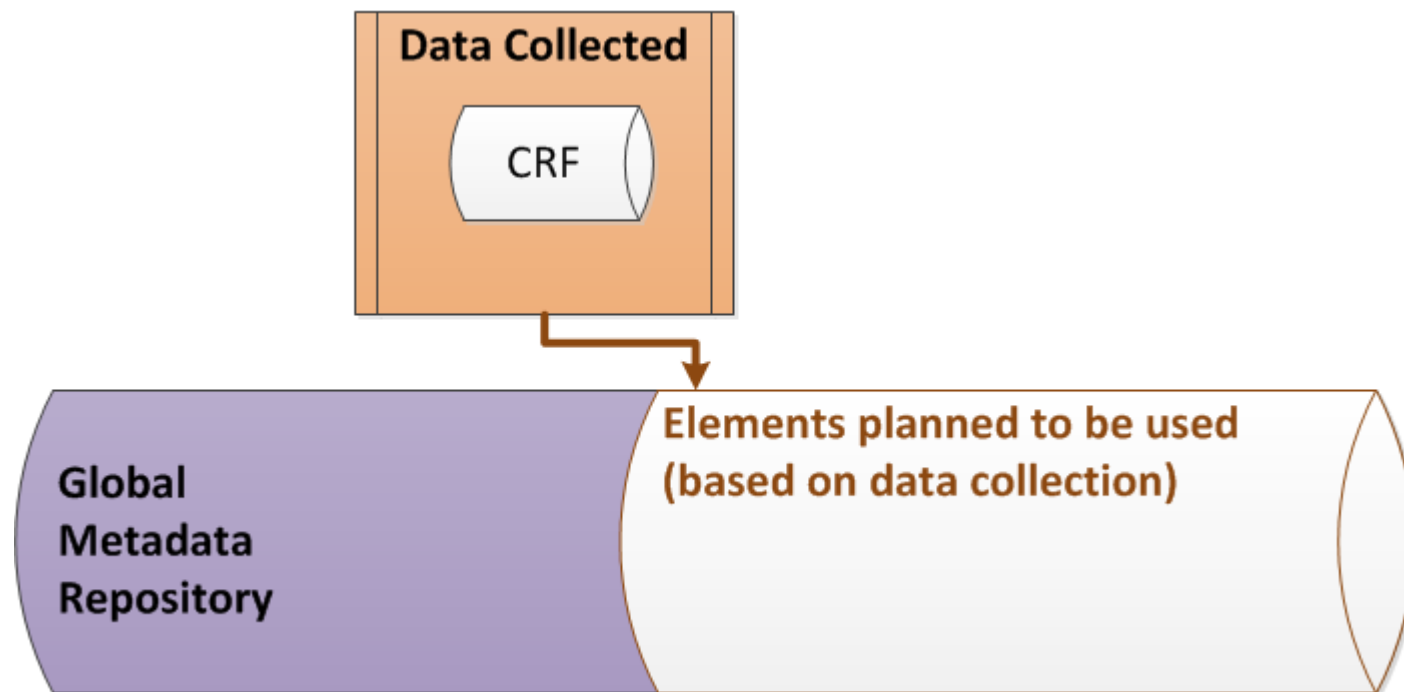
Metadata Content

What are the derivations?

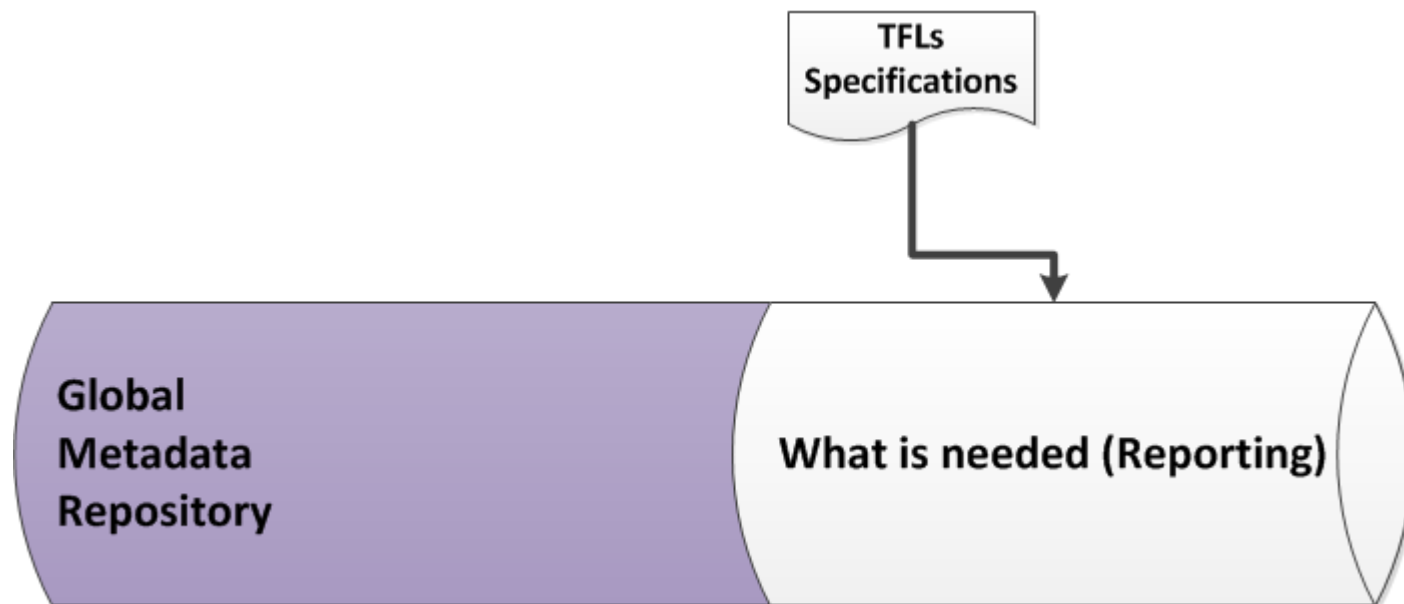


Adapting Metadata to Clinical Study Needs

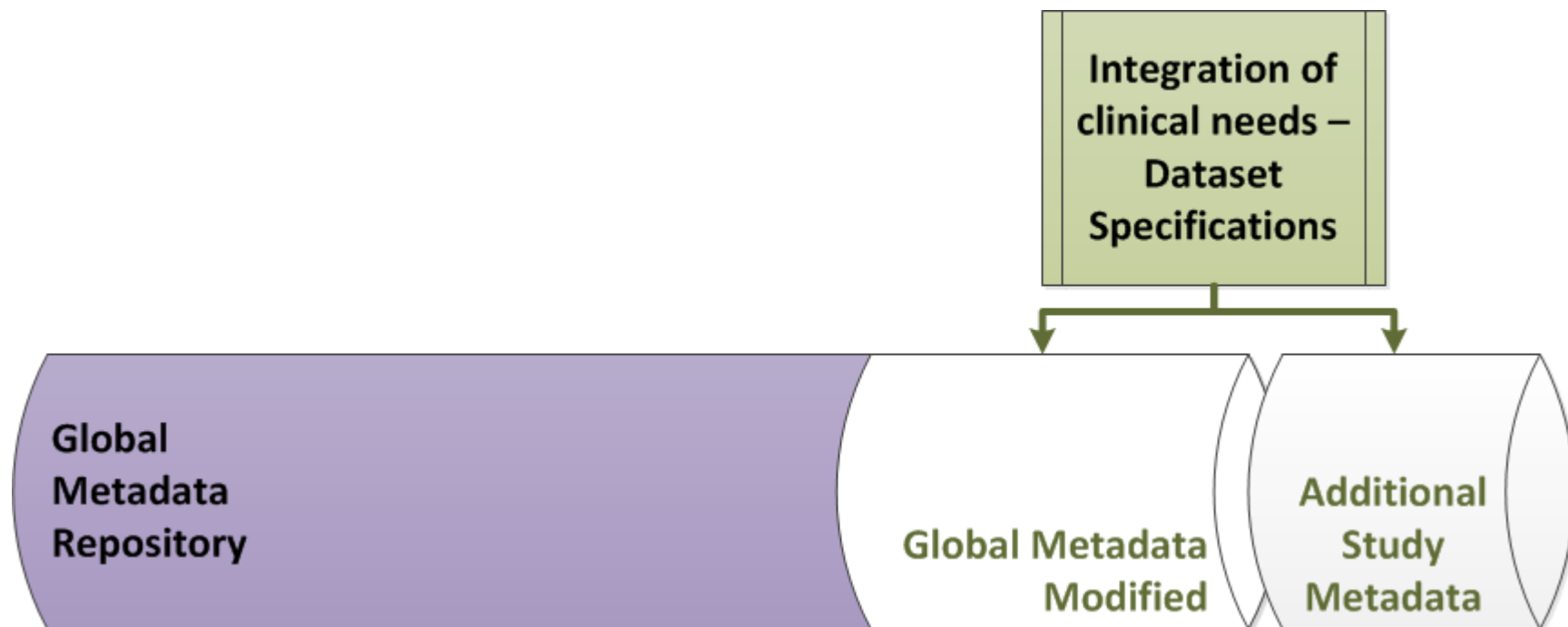
Adapting Metadata to Clinical Study Needs



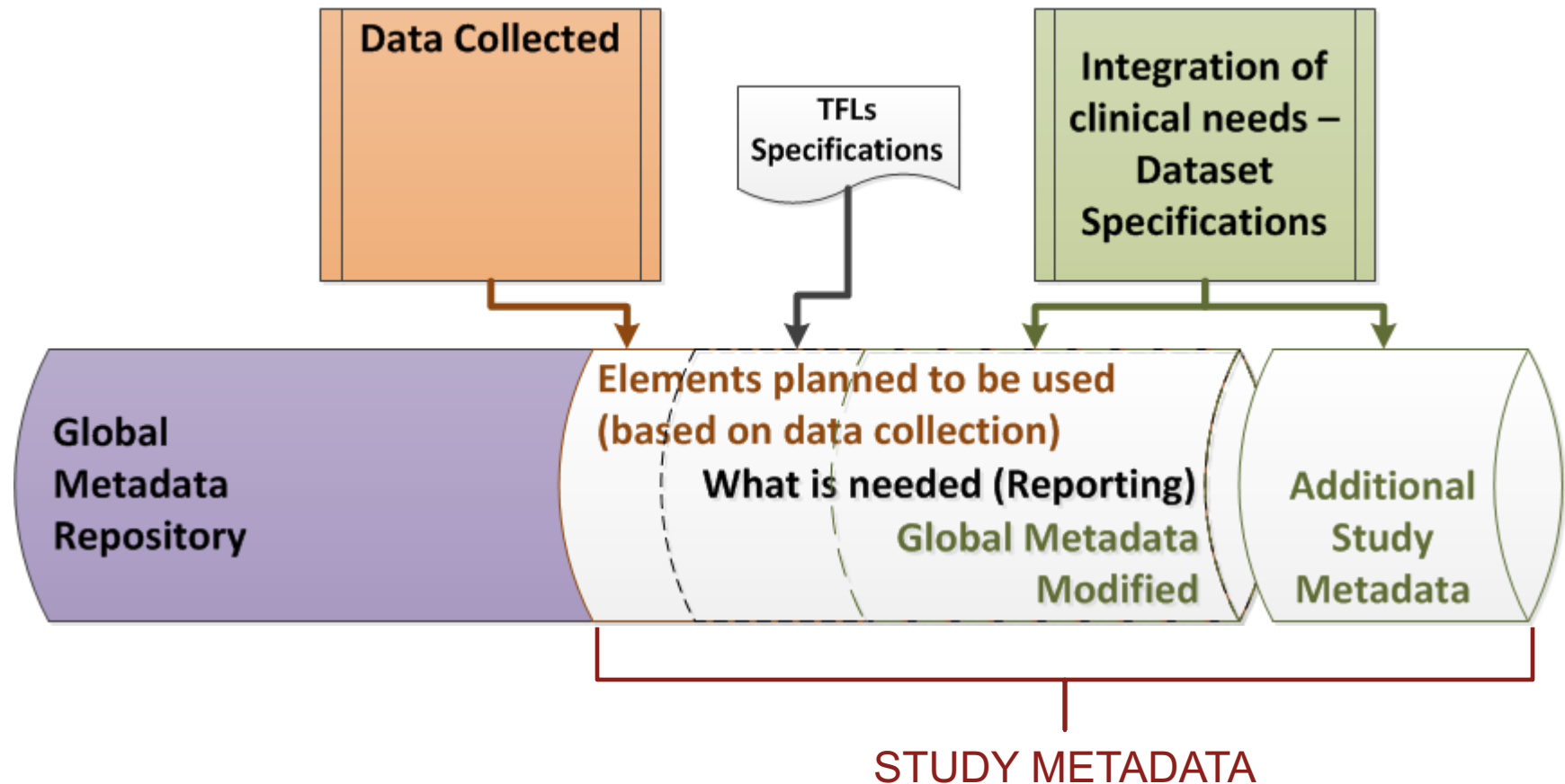
Adapting Metadata to Clinical Study Needs



Adapting Metadata to Clinical Study Needs



Adapting Metadata to Clinical Study Needs



Metadata Flow

Metadata Flow

From study metadata to the study deliverables

Define.xml contains structured metadata on different levels and covers SDTM and ADaM

- Dataset definitions
- Datasets variables definitions
- Value Level Metadata definitions
- Controlled Terminology Definitions
- Computational Methods

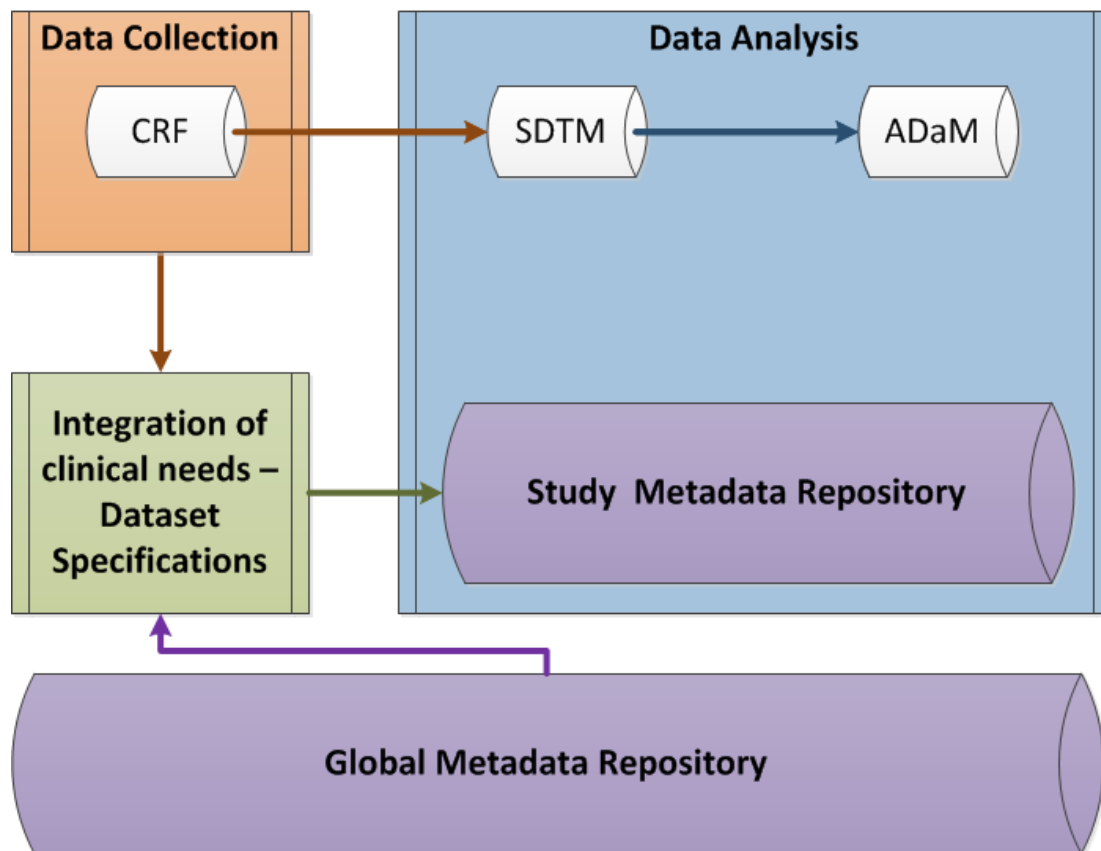
**Same content as our
metadata repository**

Metadata allows to

- Publish the information** in several formats
 - Specifications
 - define.xml
- Generate template code** for deliverables
- Maintain traceability**

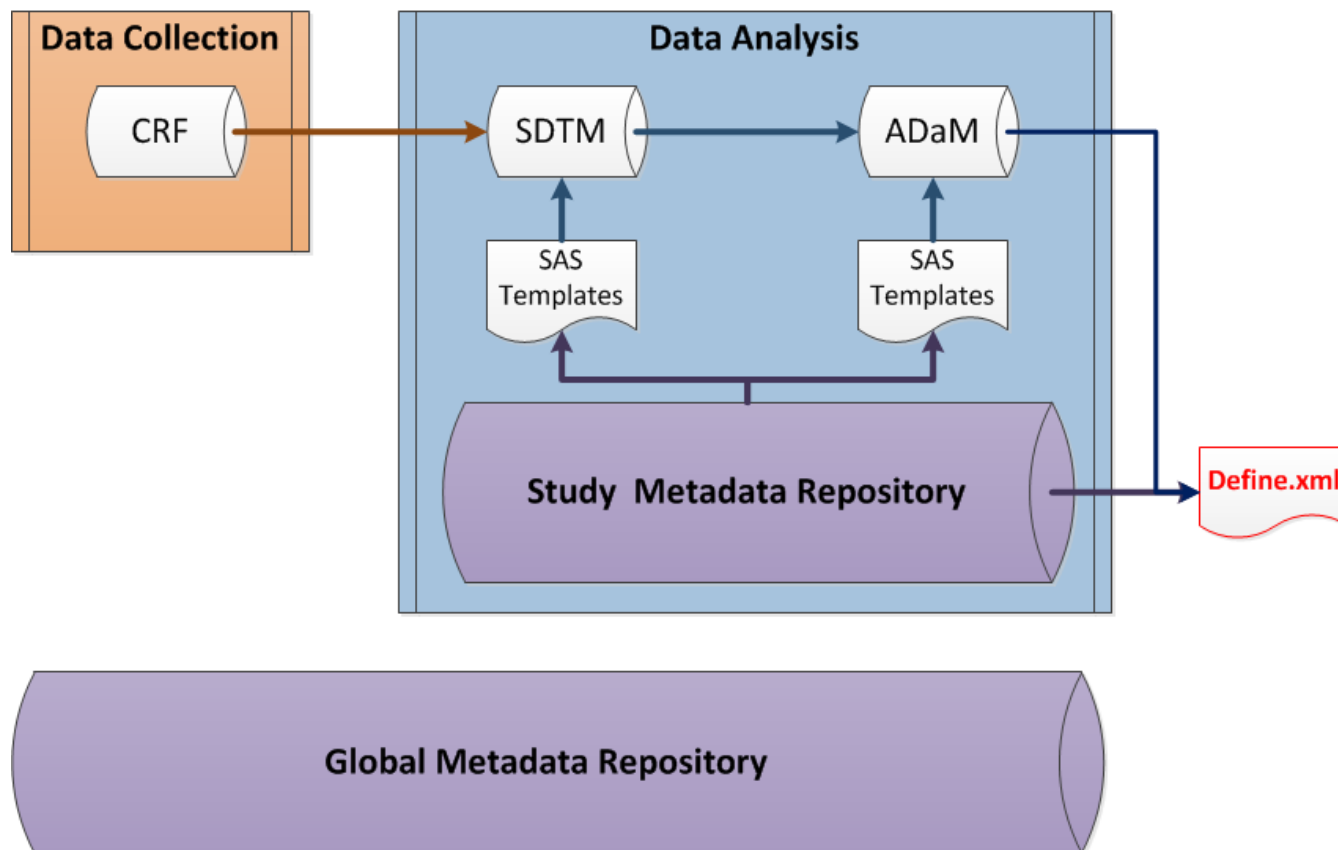
Metadata Flow

From global metadata to study metadata



Metadata Flow

From study metadata to SDTM and ADaM



Metadata Flow

Template code generation

SDTM Domain	Element	Contributing Domain	Contributing Element	Derivation
SDTM.DM	USUBJID	CRF.DM	USUBJID	KEEP
SDTM.DM	AGECL	CRF.DM	AGECL	KEEP
SDTM.DM	BRTHDDT	CRF.DM	BRTHDDT	KEEP

SDTM Domain	Element	Contributing Domain	Contributing Element	Derivation
SDTM.DM	DMDTC	CRF.SV	USUBJID EVENT SVSTDTC	MERGE
SDTM.DM	RFSTDTC	CRF.SV	USUBJID SVSTDTC	MERGE REFSTART

SDTM Domain	Element	Contributing Domain	Contributing Element	Derivation
SDTM.DM	BRTHDDTC	SDTM.DM	BRTHDDT	DTC
SDTM.DM	DMDY	CRF.DM	DMDTC RFSTDTC	STUDYDAY

Checks (Pinnacle 21 and others)

```

*****
** FILENAME : <program name> **
** DEVELOPER: <programmer name> **
** **
** REPORTING ACTIVITY : Reporting Activity CSR/DMC... **
** **
** PROTOCOL : <protocol> **
** **
** PROJECT : <project> **
** **
** DESCRIPTION : Create SDTM dm dataset **
** **
*****;
*****;

%setup;

/* Keep variable from main input file */
%Keep(sourcefile=CRF_DM, targetfile=SDTM_DM, keepVar=USUBJID AGECL RACEOR BRTHDDT);

/* Retrieve DMDTC from SV domain by Subjid and Event */
%Merge(sourcefile=SDTM_DM,sourcefile2=CRF_SV,TargetVar=DMDTC);

/* Retrieve RFSTDTC from SV domain by Subjid */
%Merge(sourcefile=SDTM_DM,sourcefile2=CRF_SV,TargetVar=RFSTDTC);

/* Create DTC representation for Birth Date */
%Dtc(sourcefile=SDTM_DM,TargetVar=BRTHDDTC);

/* Calculate Study Day */
%Study_Day(sourcefile=SDTM_DM,TargetVar=DMDY);

/* Run final Checks */
%Check(sourcefile=SDTM_DM);

```

Metadata Flow

Template code capabilities

- Create datasets, variables, tests/parameters
- Copy common variables from one dataset to all other datasets that share primary keys
- Implement database attributes to the data
- Order variables
- Sort observations
- Create user defined format catalogs, analysis parameters, or derivations
- Move variables between datasets and their supqual datasets (simple flag in the variable-level metadata)
- List discrepancies between the data specification and the database, including structure and valid values
- Add integrity constraints to the data
- Compare the consistency of study structure with standards or another study

ADaM Specificities in Terms of Metadata

ADaM Specificities in Terms of Metadata

SDTM

- Few derivations
- Simple derivations
- Mostly standard

ADaM

- All about derivations
- More complex derivations
- Need more flexibility to adapt to analysis needs

Metadata Structure : IDENTICAL

ADaM Specificities in Terms of Metadata

Coverage	Element Level Metadata	Value Level Metadata
Addition of new rows		
Analysis parameters		X
Analysis timepoints		X
Addition of new columns		
Grouping : variable level (ADSL)	X	
Categories		X
Analysis sets	X	X
Analysis flags	X	
Criteria flags		X
Imputations		
Imputations	X	X

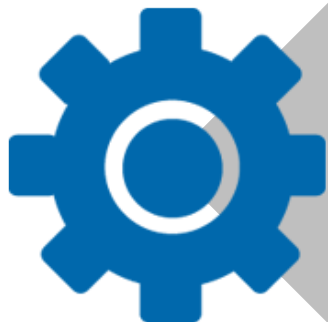
<Non-exhaustive list>

ADaM Specificities in Terms of Metadata

Metadata

Traceability

Automation



Plug user code and specifications in the different levels of metadata

FLEXIBILITY

Consistency between all deliverables

ADaM Specificities in Terms of Metadata



Elements
planned for use
(based on
collection)

What is needed
(based on
analysis needs)

Study specific
updates and
additions

Analysis Dataset

Element/Variable

Controlled Terminology

Value Level Metadata : Analysis parameters, categories, criteria, ...

Derivations incl. Plug user code

ADaM Specificities in Terms of Metadata

ADaM Challenges

- OCCDS, Safety**: standard derivations
- BDS, Efficacy**: Complex and study specific derivations that can not be easily standardized
- standardization of content and parameters with a lot of **analysis driven variability** : complex and/or high time consuming across a large portfolio

Conclusion

Conclusion

Challenges

- Flexibility
- Maintenance
 - New standard = new metadata
 - versionning
 - CDISC compliance
- Standardization highly time consuming
- Training

Value

- Study level programming more efficient
- Deliverables content and workflow
- Structured, consistent and CDISC-compliant deliverables
- Traceability
- Automation

Thank you