



## Harmonizing Raw Data with CDISC Standards to Streamline SDTMs

*“Apply the 80/20 rule to ensure the Project **automates 80% of the end-to-end metadata and data processing** needed to generate study artifacts suitable for a regulatory submission.” Peter Van Reusel, Sam Hume, CDISC-360 Mission*

Sunil Gupta, [GuptaProgramming@gmail.com](mailto:GuptaProgramming@gmail.com)  
CDISC SME, Trainer and Author

Rahul Madhavan  
[Rahul@PointCross.com](mailto:Rahul@PointCross.com)  
VP – Strategic Programs, PointCross

# About PointCross



Platinum member of CDISC and contributor to PhUSE



Serving 60+ Biotech, CROs & large Pharma clients



Xbiom is our BioPharma solutions platform for:

- ✓ Research insights from nonclinical and clinical studies, assays, biomarkers
- ✓ Validated workflows for regulatory nonclinical and clinical submission preparation
- ✓ Automated Smart Transformation for data curation from data lakes to target model
- ✓ Governance and continuous management of standards, terminology, metadata
- ✓ Trial and study data repository with role based access and controls

# Today's Webinar

---

## Welcome & Introduction

### Sunil Gupta: CDISC 360 Advisor, Trainer, and Author

- Advocate of CDISC automation and standardization, with 25+ years of experience in the pharmaceutical industry
- Teaches practical R for SAS Programmers and a CDISC class at the University of California at San Diego



### Presentation:

- Harmonizing Raw Data with CDISC Standards to Streamline SDTM

---

#### NOTE:

- Questions can be submitted any time via the Chat function, anonymously.
- Written responses will be consolidated and sent to all participants and uploaded to PointCross website (along with the recording of today's webinar)

#### Upcoming:

- PHUSE Webinar Wednesday: Leveraging the Universal Data Model (UDM) to Streamline Pathways to SDTM and ADaM Standards
- January 25, 10:00 AM ET; Registration link in the chat box
- Presenting at PHUSE Connect in Orlando, FL March 5-8, 2023

Contact us: [ask@pointcross.com](mailto:ask@pointcross.com) for a private demo, customer references, and integrations with your workflow

# What is Your SDTM Automation Role and Goals?

## ❑ **Statistical Programming Director – *Better Manage Timelines***

- Project Management, Time, Budget, Compliance, Metadata

## ❑ **Statistical Programmer – *Better Manage Submission Process***

- SDTMs / ADaMs / Define.xml, SDRG / ADRG, Mapping, Specifications

## ❑ **Translational Scientist – *Better Understand Safety Data Issues***

- EDC / Biomarker Raw Staging and Data Processing, Ingestion, Curation, Harmonization

## ❑ **Statistician – *Better Explore and Understand Study Conclusions***

- Analysis, Views, Ad-hoc, Tables, Listings and Figures, SAP, Data Meaning / Exploration, TLFs

# Leverage a Universal Data Model (UDM) to Streamline Pathways to SDTM and ADaM Standards

## ○ Paradigm Shift

- Post CDISC Standards and Experience
- Non-Linear Accretive Processing
- Integrated CDISC Compliance

## ○ Xbiom with Universal Data Model (UDM)

- Low-Code User Interface
- Six Step Visualization Process

## ○ Universal Data Model (UDM) Design

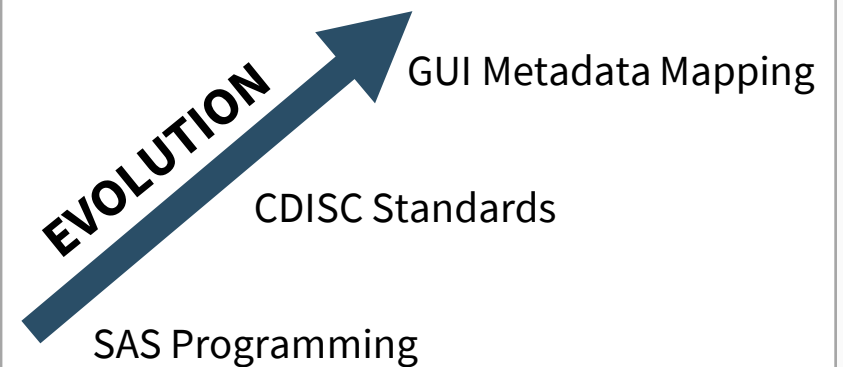
- Reusability and Repeatability
- Direct to Review and Analysis
- Monitoring and Accretive Resolution of Data Issues

## ○ SDTM Generation

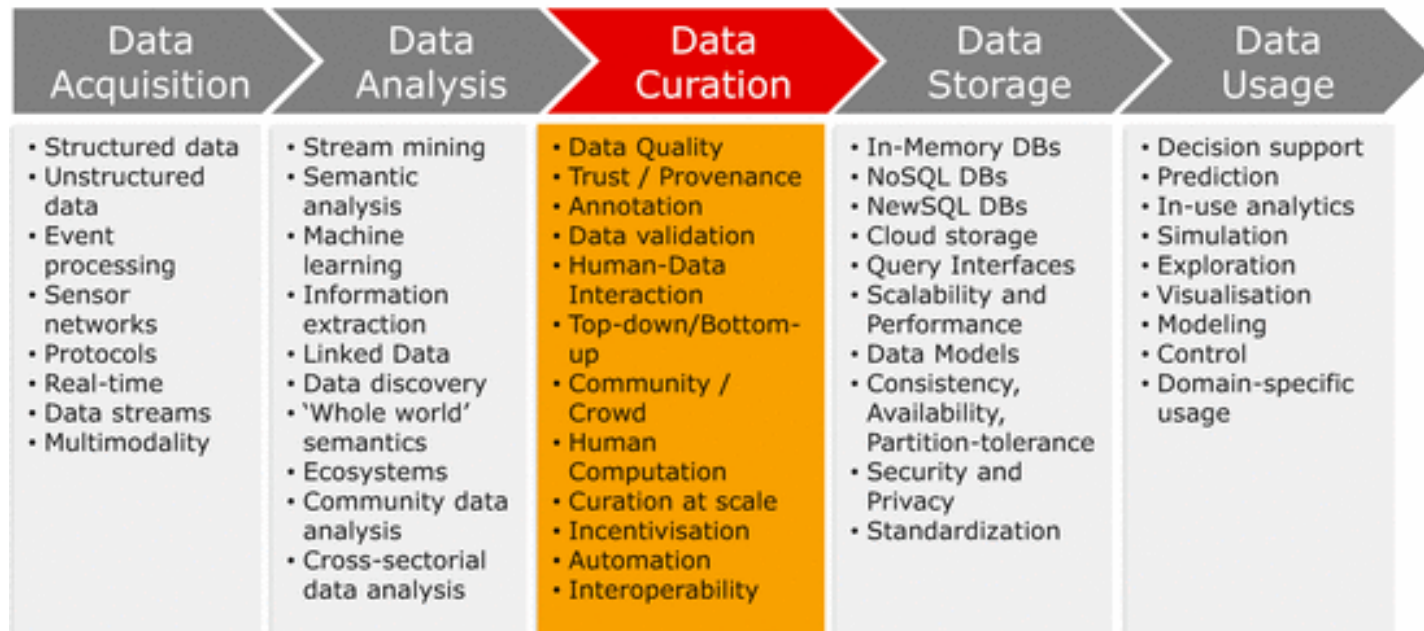
- SDTM IG Specifications, CDISC 360's Mission
- Metadata Repository, Auto-Mapping and User Confirmation
- Continuous Learning Process



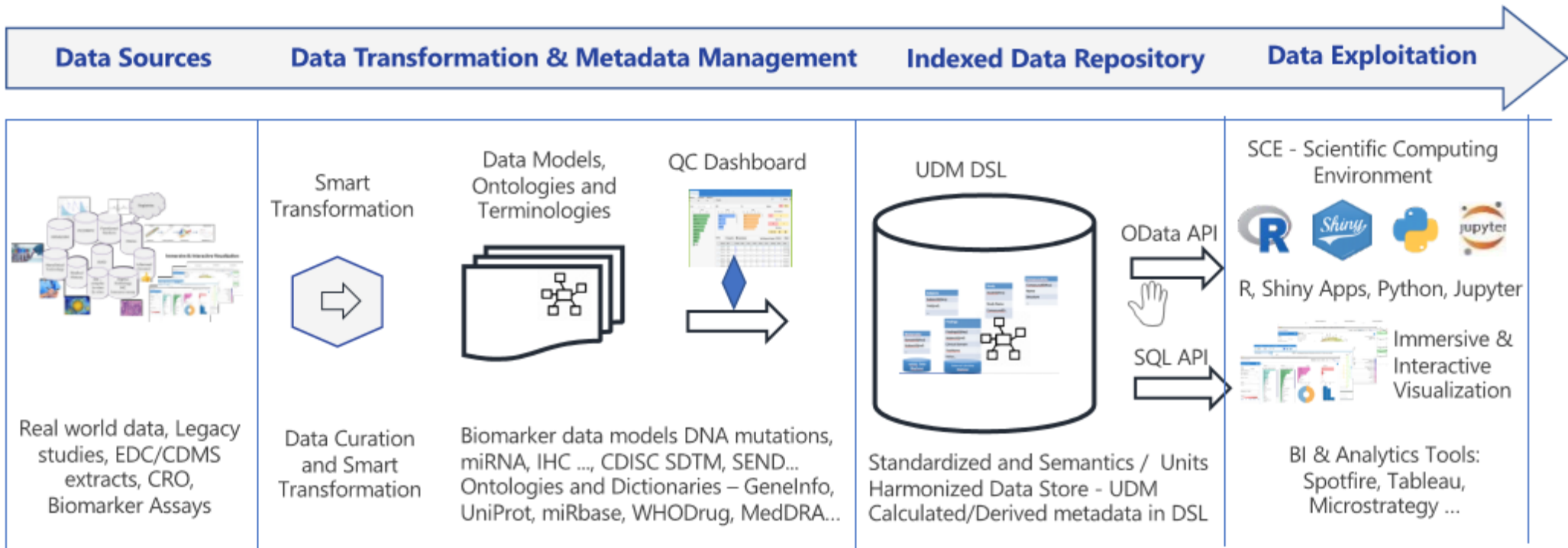
- > Search and Query Interactive Database
- > Ingest, Catalog, Index and Link All Data



# Big Data Value Chain

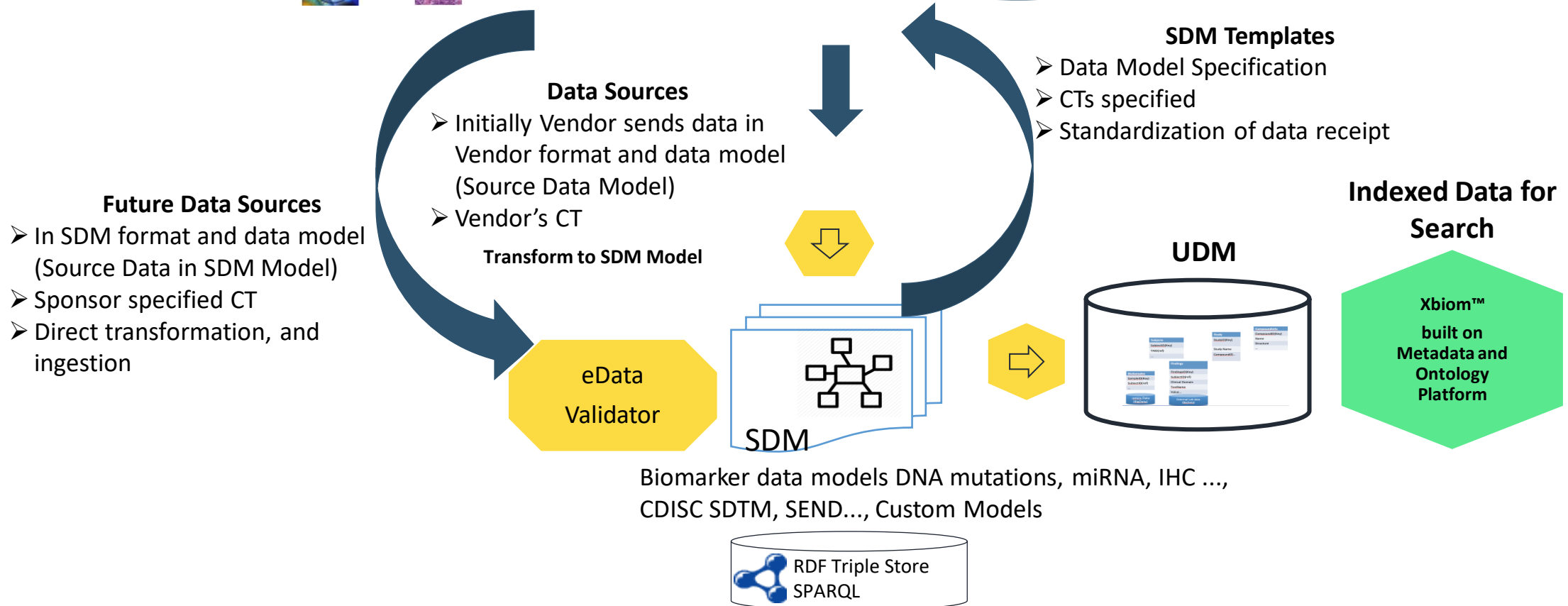
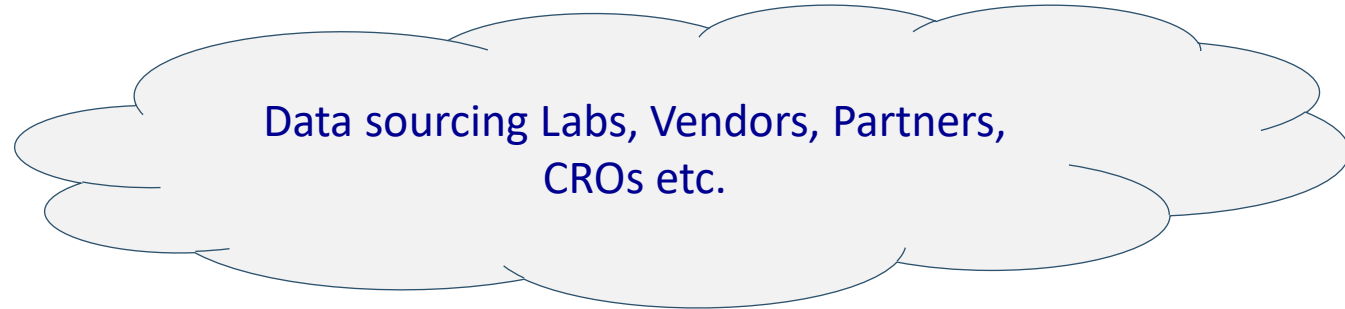
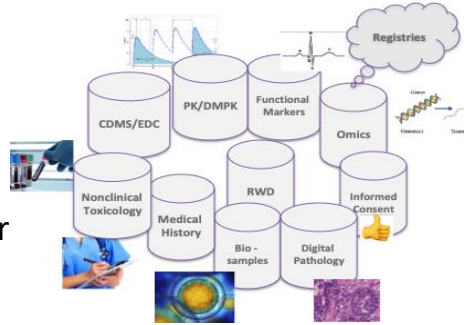


Data Curation is the Repetitive Process to Optimize Data and Metadata to ensure Valuable use of Data.



# Streamlining Data Interchange with External Sources

Ongoing Studies, Real world data, Legacy studies, EDC/CDMS extracts, CRO, Biomarker Assays





# CDISC 360 MISSION:

## SDTM Design

## And Automation

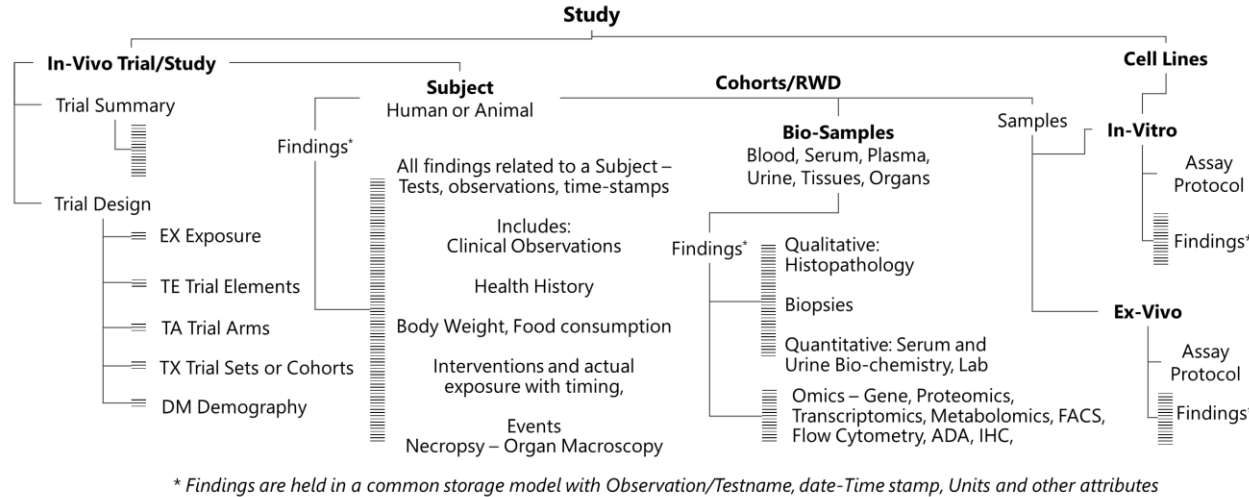
### Input: Metadata & Raw Clinical Data

#### Create End-to-Start Specification

- ☐ Produce a standards-based, machine readable specification

#### Generate Start-to-End Metadata

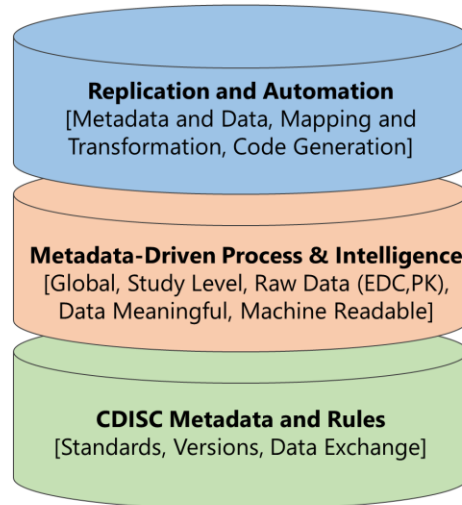
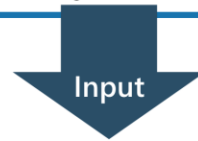
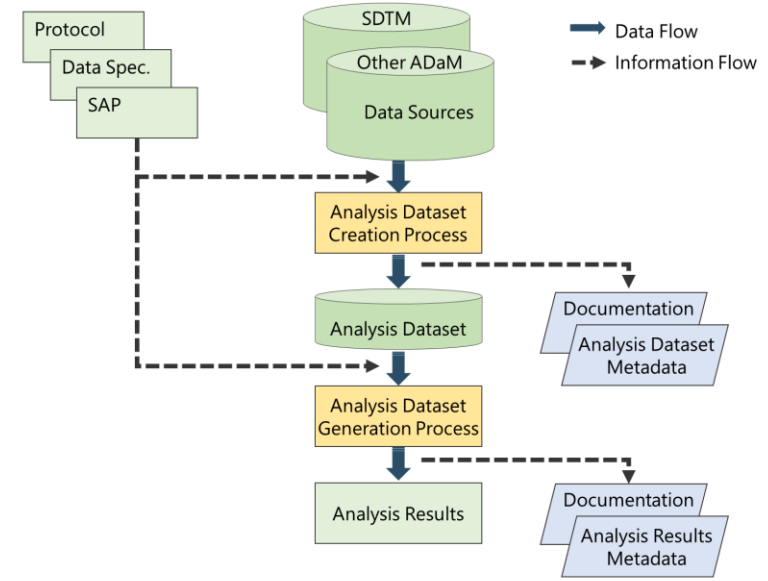
- ☐ Use standards specification to generate study metadata artifacts
- ☐ Demonstrate the ability to generate study metadata given a specification



### Output: SDTMs, ADaMs, Define.xml & TFLs

#### Data Curation

- ☐ Repetitive Process to Optimize Data and Metadata to ensure Valuable use of Data



### Transformation and Automation: Reusability & Repeatability

#### Transformation Data Start-to-End

- ☐ Use machine-readable metadata to generate study data artifacts
- ☐ Demonstrate the ability execute data transformations given the study

#### Replication and automation are the focuses

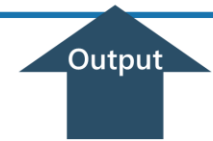
- Use or create utilities to replicate the process: Project Set Up, Mapping Specification, Mapping Creation
- Use analytics tool to identify the areas for replication and automation: Data Profiling & Data Rules for Source Data Review / Edit Checks

#### Metadata-driven process is the key for automation

- Metadata makes data meaningful
- Metadata is machine readable
- Metadata is the base for automation

#### Standard adoption is the key for code reusability

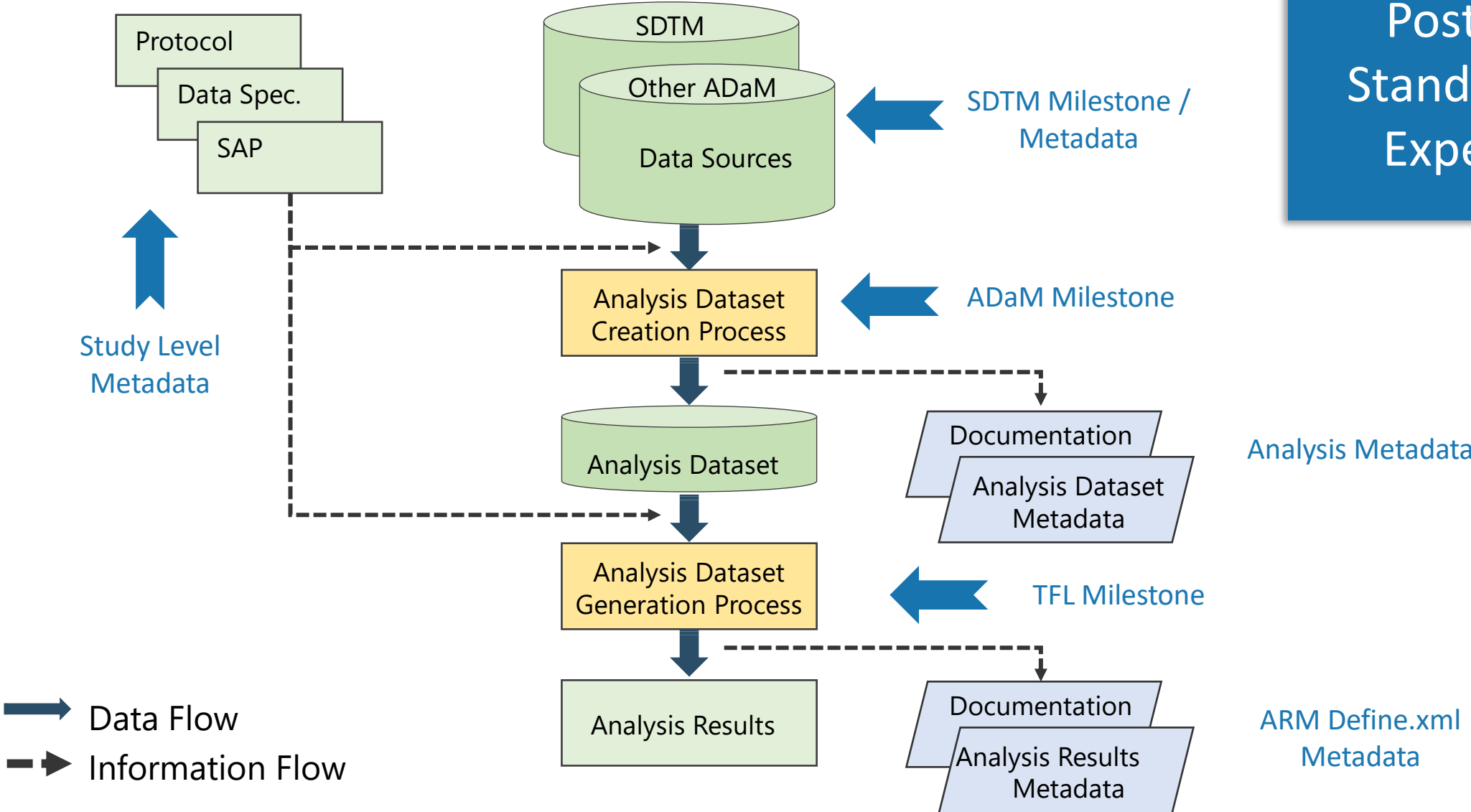
- Train people to understand the standards
- Define standard templates
- Build public libraries for code snippets and public transformation: Custom functions, procedures and packages; public data rules; and public Experts
- Group code snippets and functional transformation into modular mapping and transformation: pluggable maps
- Define workflow to govern the process: Workflow Manager and Process Flows



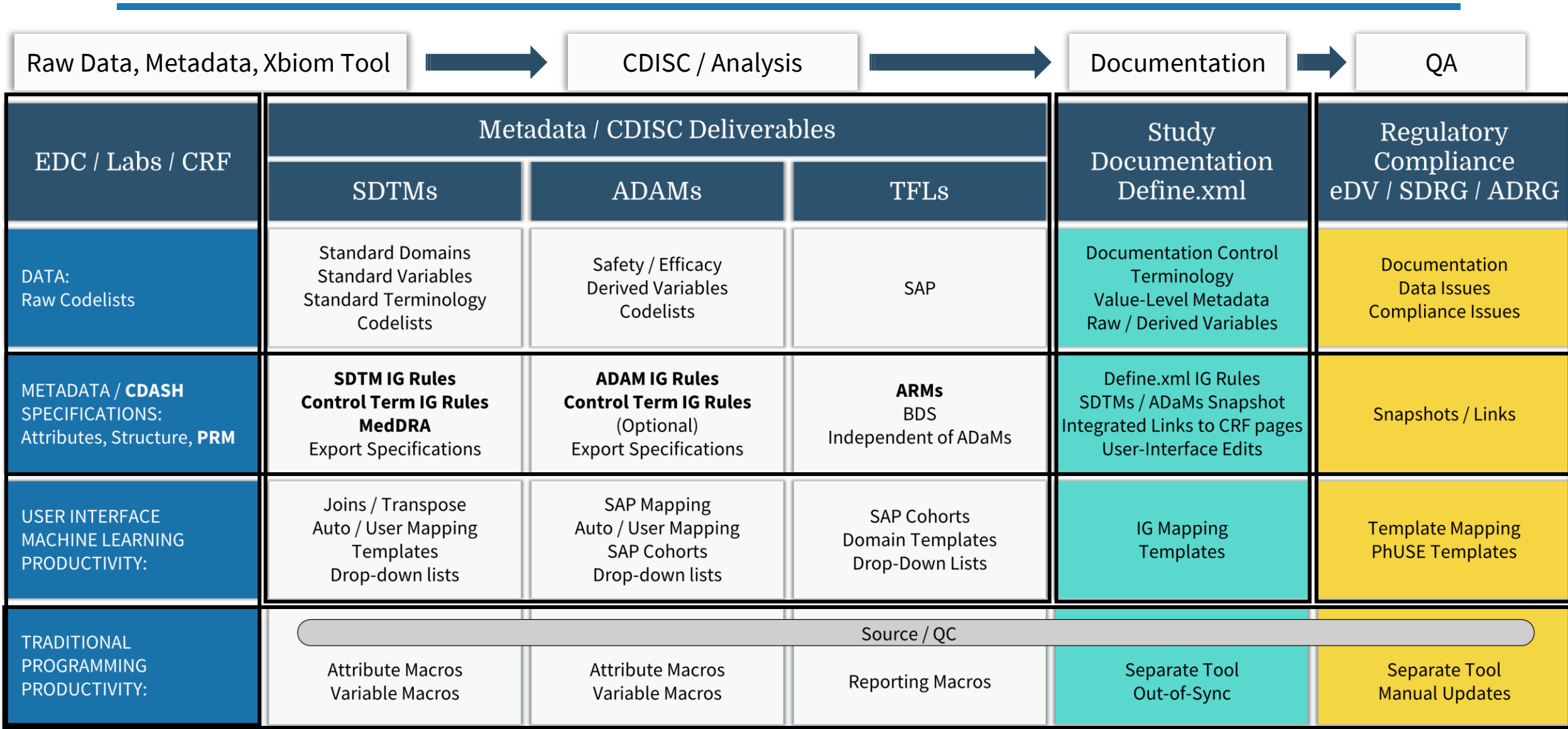


# Paradigm Shift: Post CDISC Standards and Experience

## CDISC Standards / Metadata



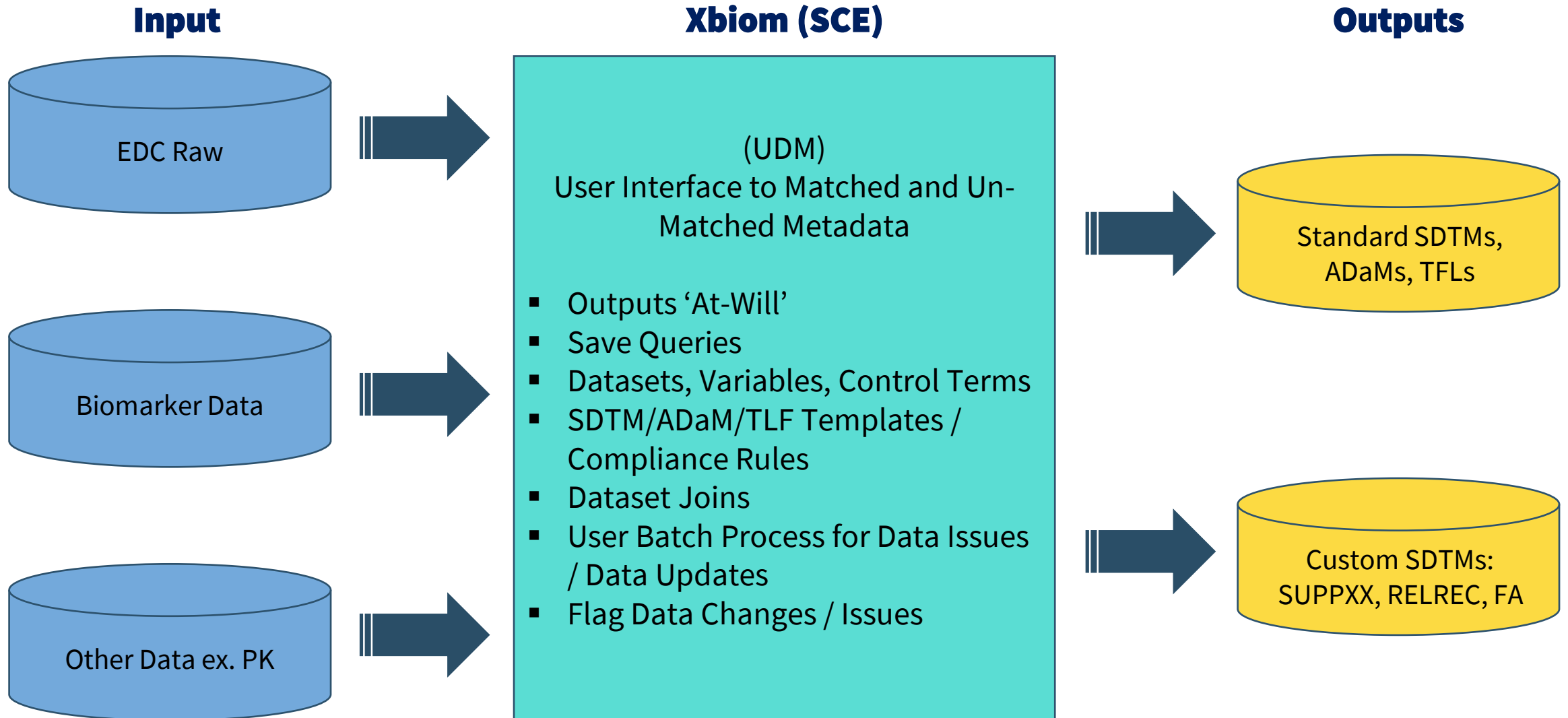
# End-to-End Clinical Study MetaData-Driven Process and Intelligence CDISC and Submission Flow



# SDTMs / ADaMs / TLFs

## Integrated Metadata Mapping and Repository

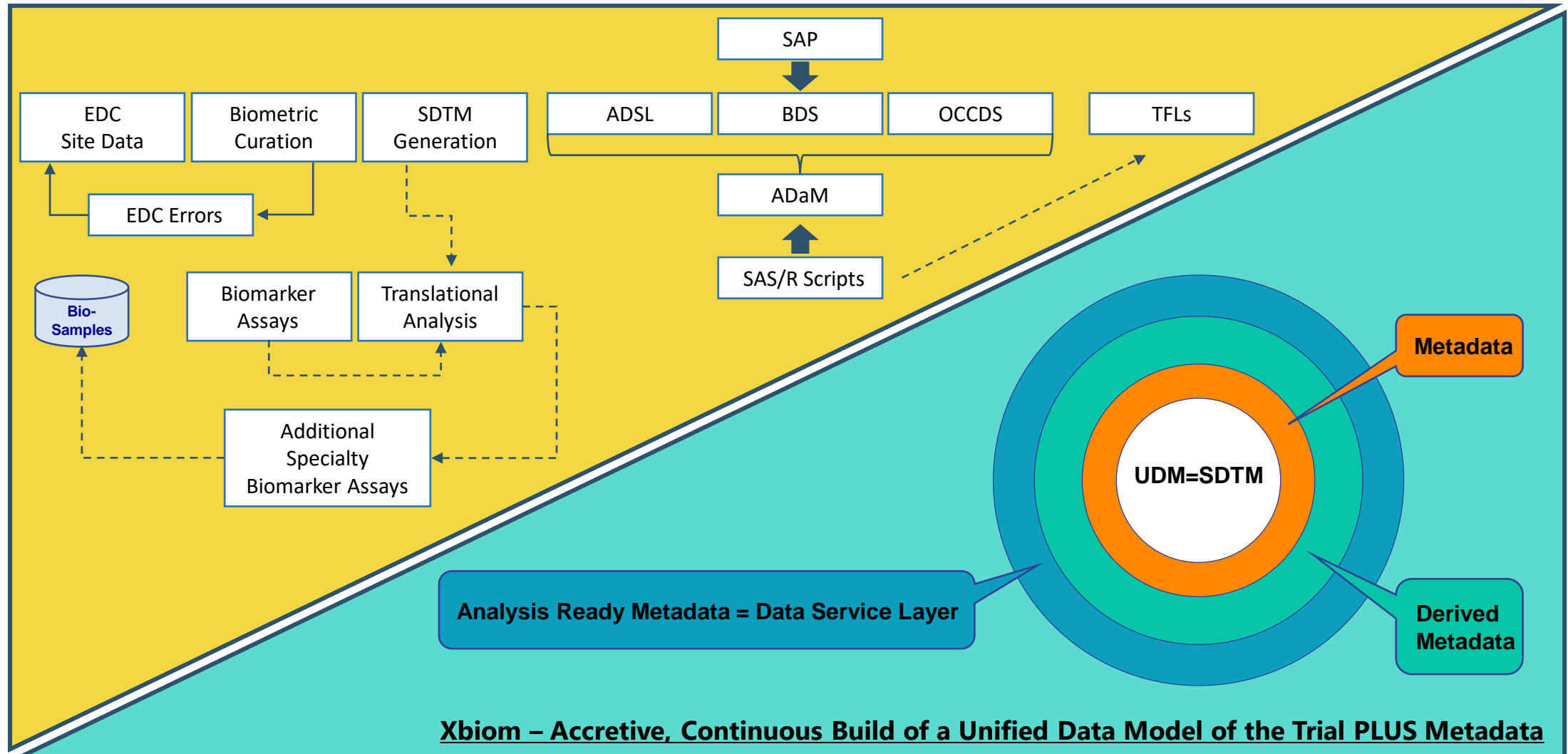
---



# Paradigm Shift: Non-Linear Accretive Processing

From Sequential Processing of Files to Create SDTMs & ADaM for TLFs

To A Single Accretive, Curated Unified Data Model - Select Cohorts, Analyze Data to SAP,  
and Automatically Generate SDTM, ADaM for Submission



**Xbion – Accretive, Continuous Build of a Unified Data Model of the Trial PLUS Metadata**

Universal Data Model (UDM) with Data Service Layer

Cohort Analysis

Cohort Search

Curation, Xform

UDM=SDTM

4

3

2

1

INGISHTS DECISIONS

KNOWLEDGE

INFORMATION & DERIVED METADATA

DATA

Data Exploitation

Search and Query Engine - stratified cohorts

AI & ML Deep Analysis Model

IGO's & TFL - Publishing

Statistical Computing Environment (SCE)

DSL (Data Service Layer)

Medical Annotations and inferences generate by research users

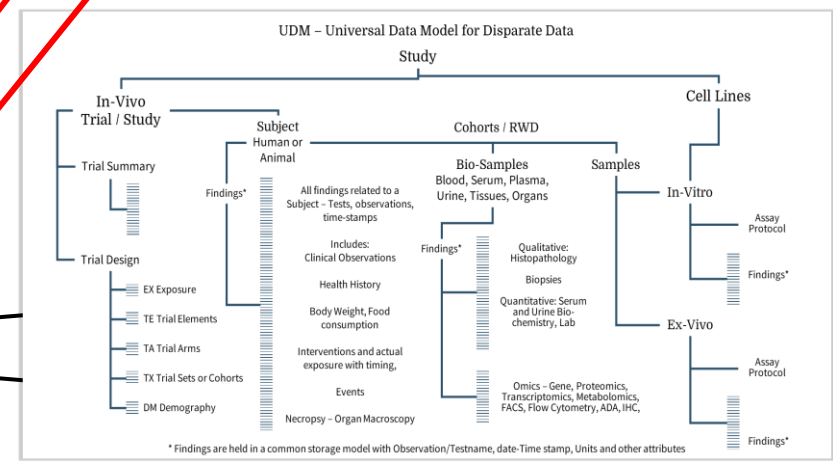
Analysis ready datasets

Calculated / derived data and trends extracted on subject data & cohorts (Statistical algorithms & ML trained models)

Longitudinal integration of data for samples, patient on visit timelines

Curated & Harmonized Data in UDM

Raw Data as Collected



# Xbiom: Low-Code User Interface and Visualization Tool

Xbiom is a one-stop SCE platform designed for integrating Study Data and Submission Packages

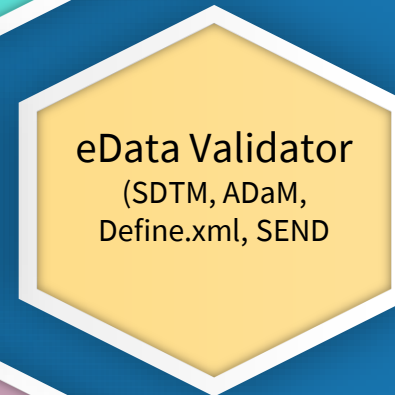
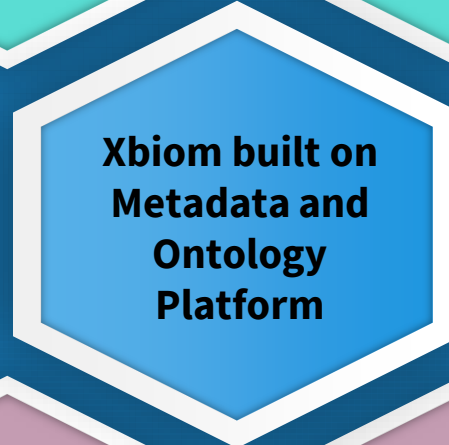
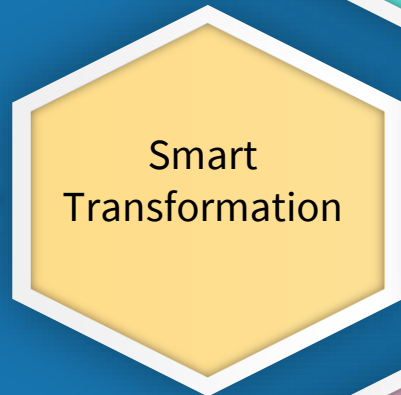


## CAPABILITY

Insights for Translational Research in Monitoring Ongoing Studies / Trials

Data Curation, Transformation, Harmonization, Quality, and Ingestion

Regulatory Data Standardization, Submission Planning and Preparation



## R&D BUSINESS VALUE

Speed to reliable, analyzable, and submittable data

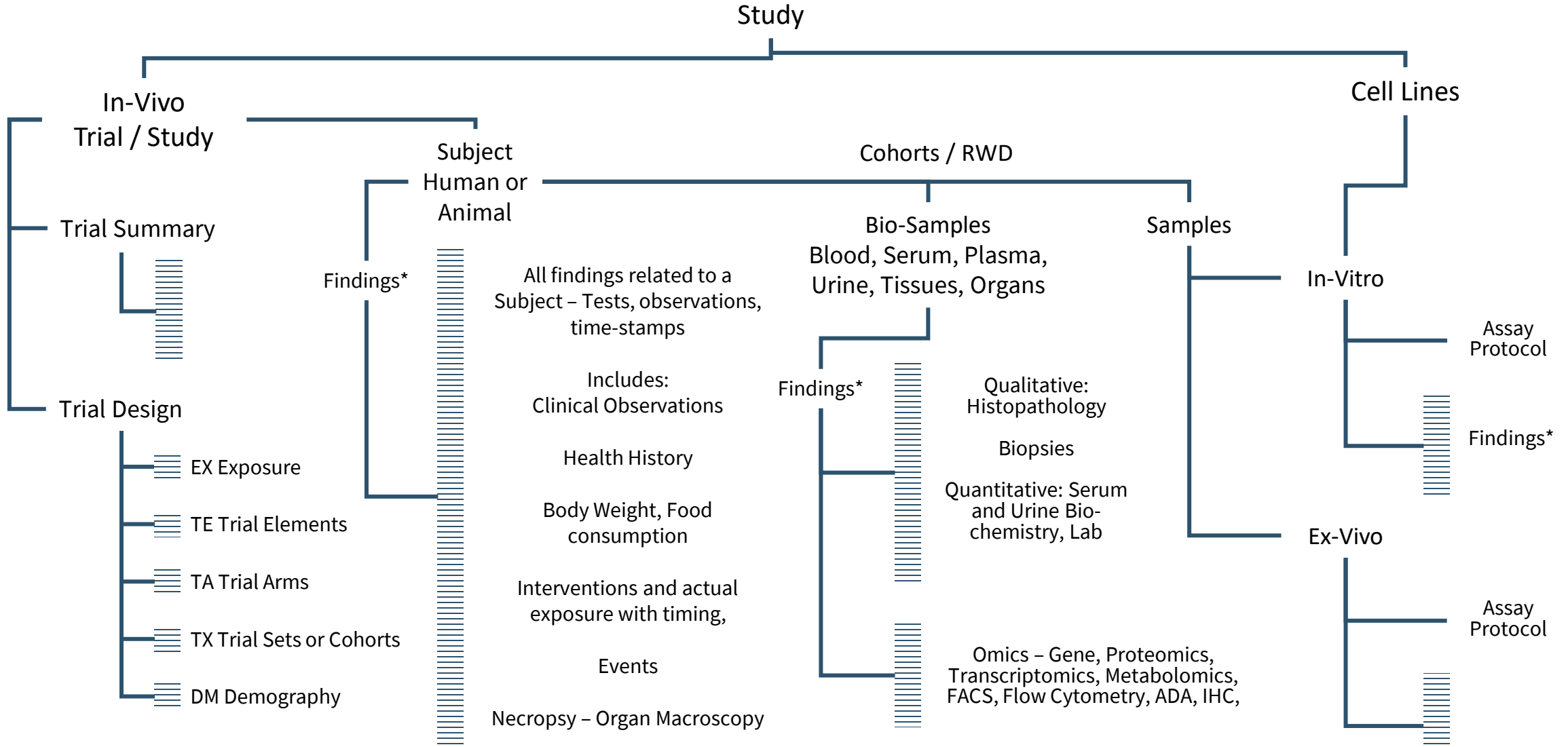
Search, define, and analyze stratified cohorts

Longitudinally integrated subject and patient data

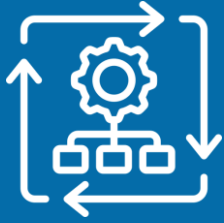
Rapidly generate TFLs for publication and collaboration

# Unified Data Model: Reusability and Repeatability, Direct to Review & Analysis, Monitor and Accretive Resolution of Data Issues

## UDM – Universal Data Model for Disparate Data

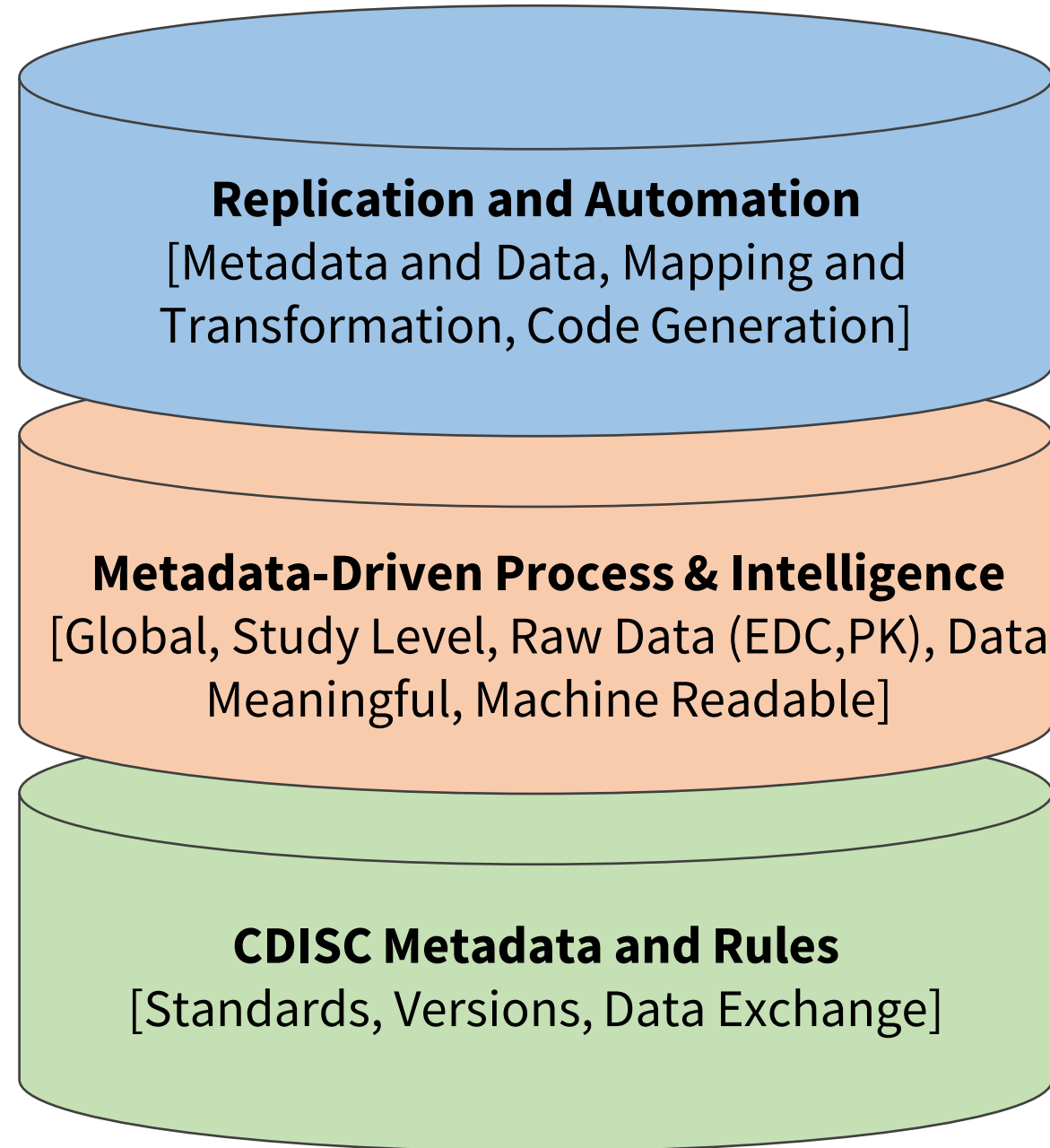


\* Findings are held in a common storage model with Observation/Testname, date-Time stamp, Units and other attributes



## SDTM Automation By Reusability and Repeatability

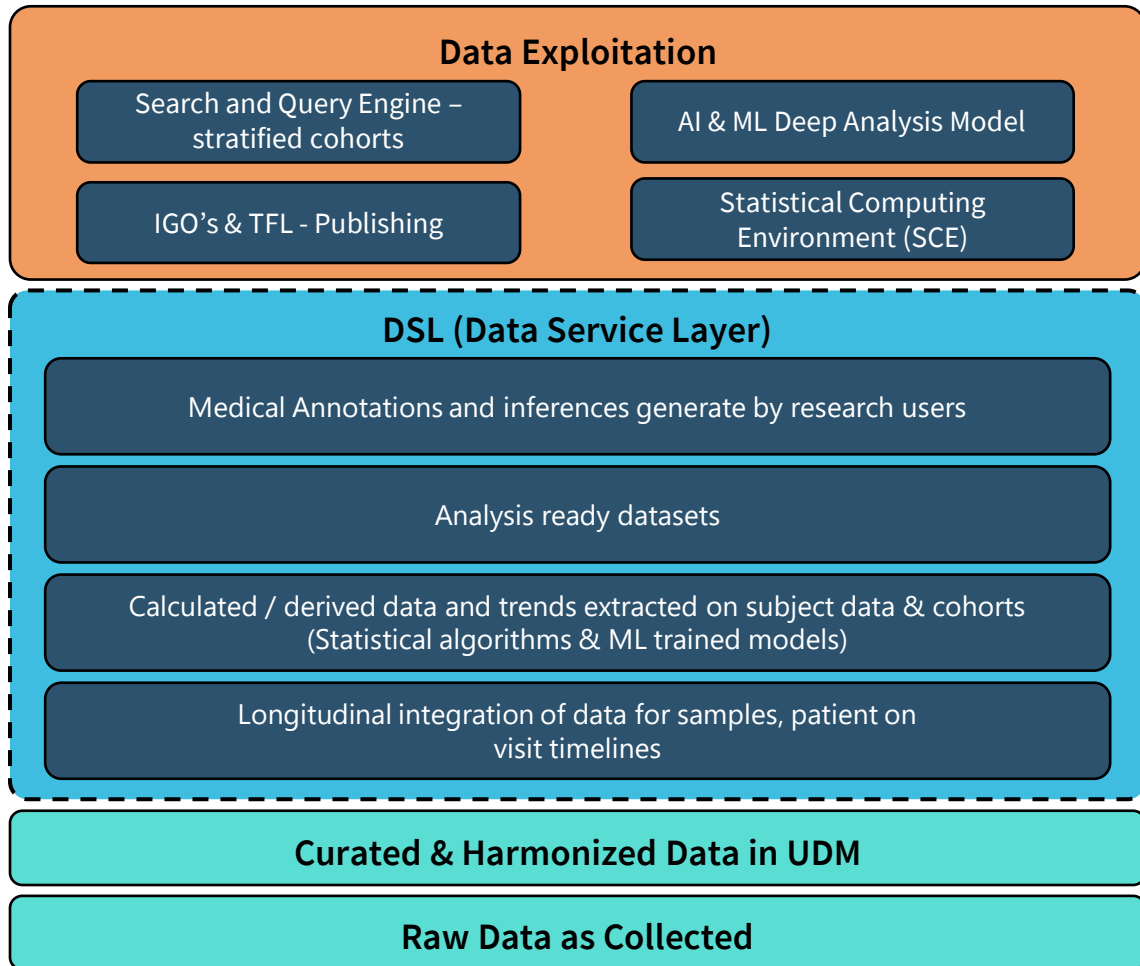
CDISC 360: Seeks to demonstrate how standards enable metadata-driven end-to-end automation.



EVOLUTION



# Metadata Categories and Examples Mapped to Xbiom



E  
V  
O  
L  
U  
T  
I  
O  
N

**SAS and R Programs/Applications:** Macro variables, parameters, defaults, programs and libraries, functions, conditional macro calls and assignments (If/Then/Else), operators, templates, styles, applications, data-driven macros, lookup tables, design/logic, code generator, user message, libnames, config file

**CDISC Model and Compliance:** Rules, Events/Findings, structure (BDS/Wide), Order, codelist, define.xml (hierarchy), standard/original values, map/unmap variables

**Dataset and Variable Relationships:** 1:M, Parent: Child, top 10, links

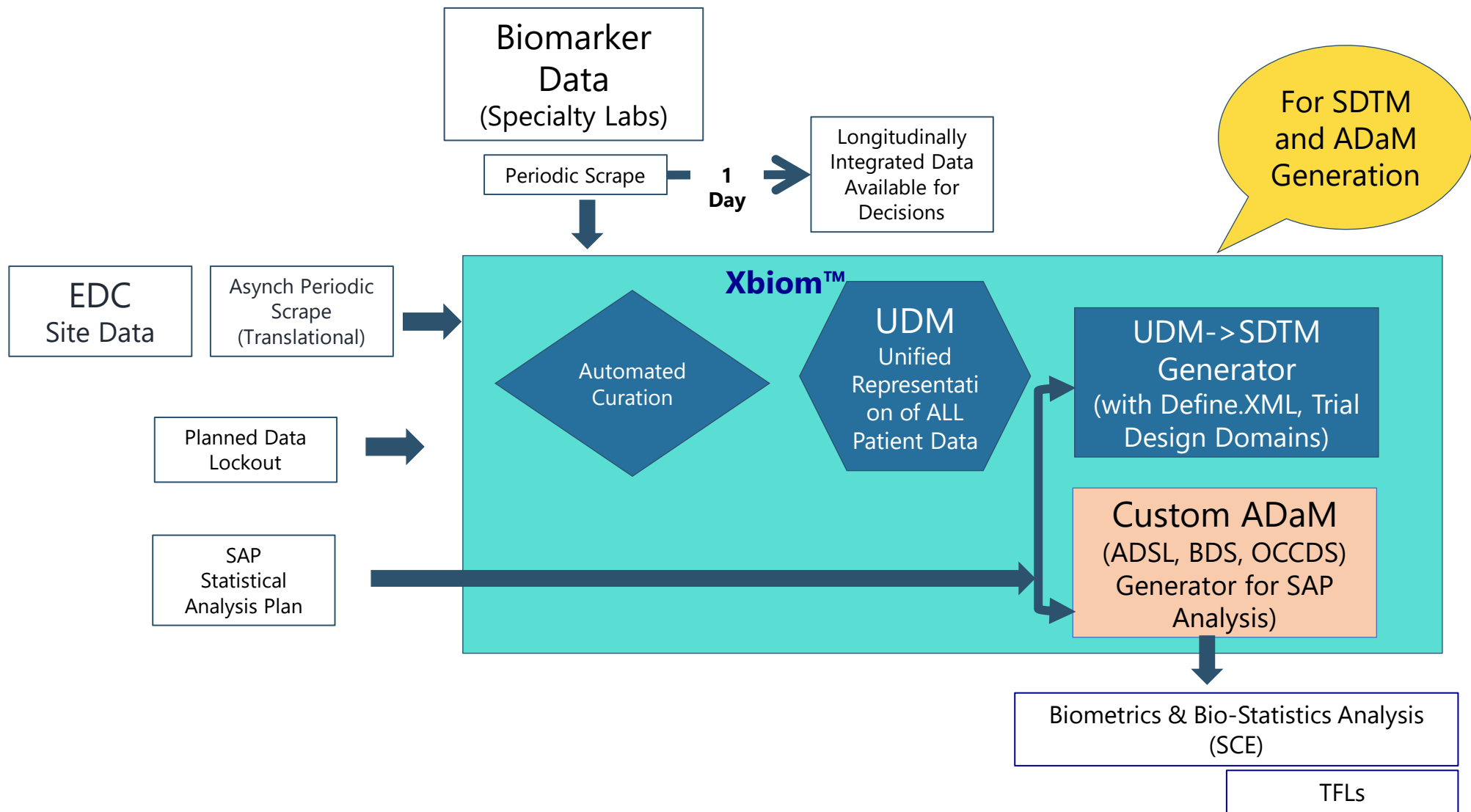
**Folders, Files, Dataset, Variables Attributes:** length, type, format, color, size, weight, access, location, units

**Folders, File, Dataset and Variable Names:** lists, file name, extension type

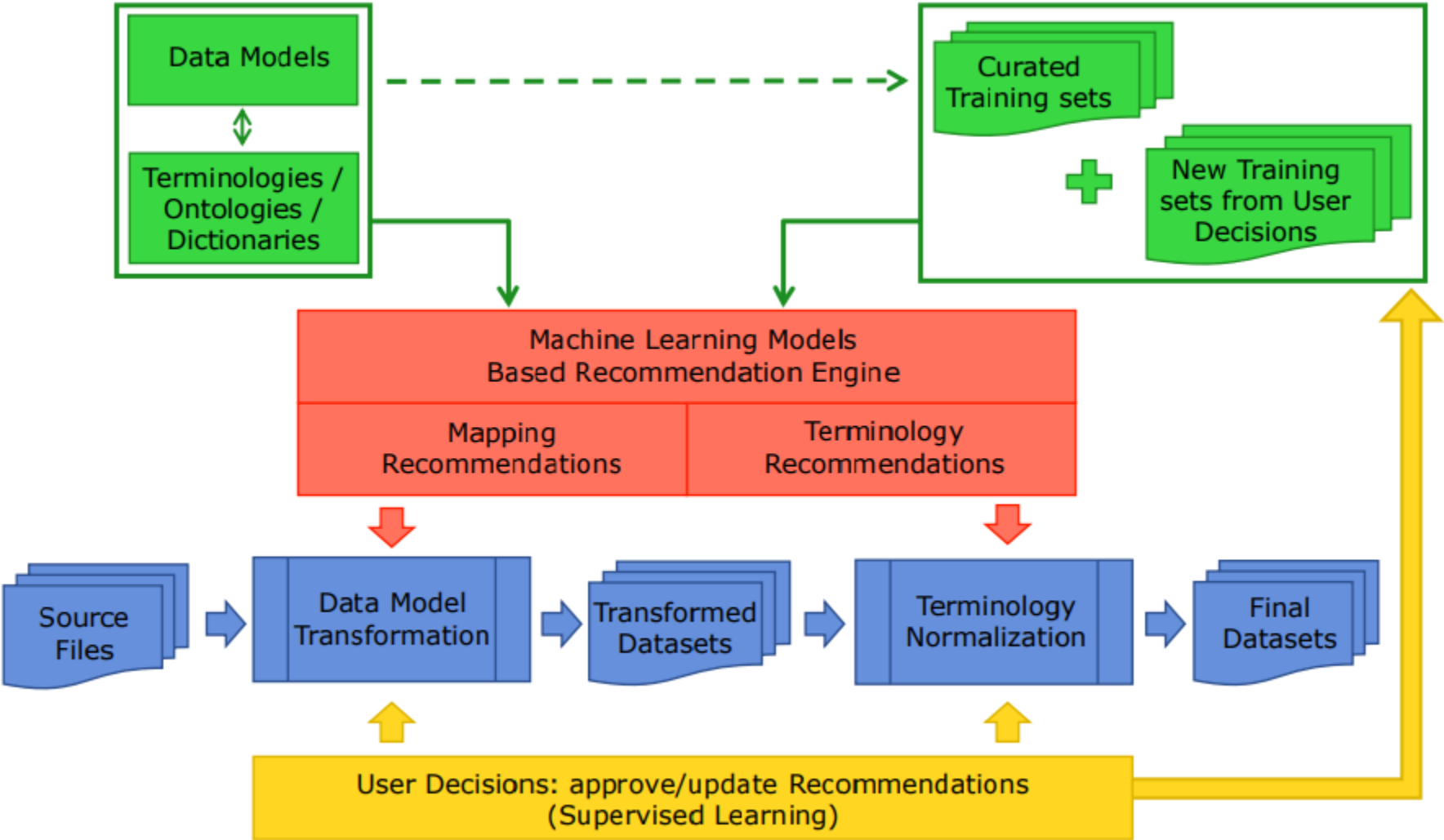
**Variable Values:** valid values, code list, min, max, continuous, categorial, format catalog, statistical analysis, cutoff/flags, new/change, date, times, patterns

**Context:** Meaning, Purpose, mind map

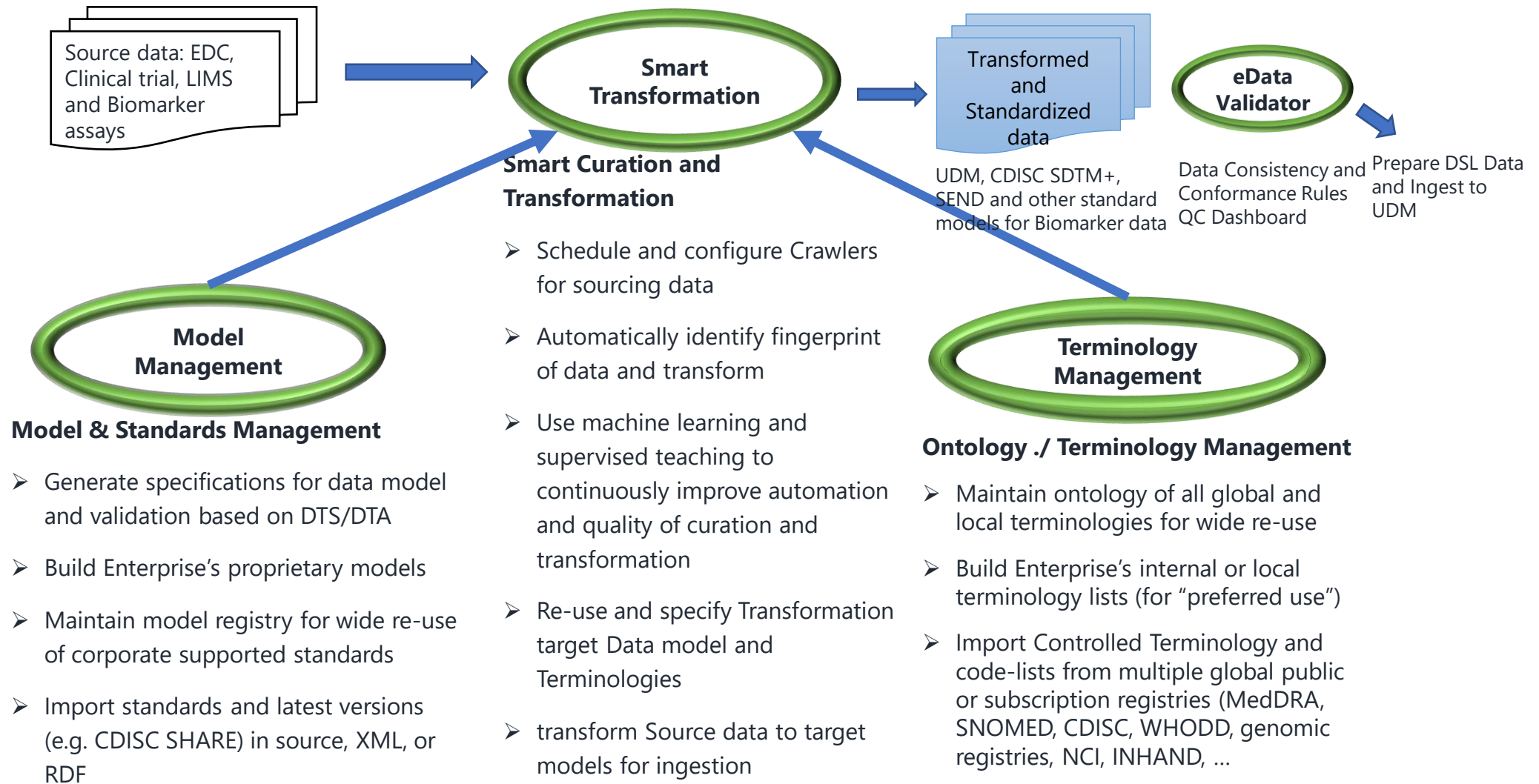
# SDTM: Automapping and User Confirmation, Metadata Repository and Continuous Learning Process



# Machine Learning and Data Transformation Work Flow

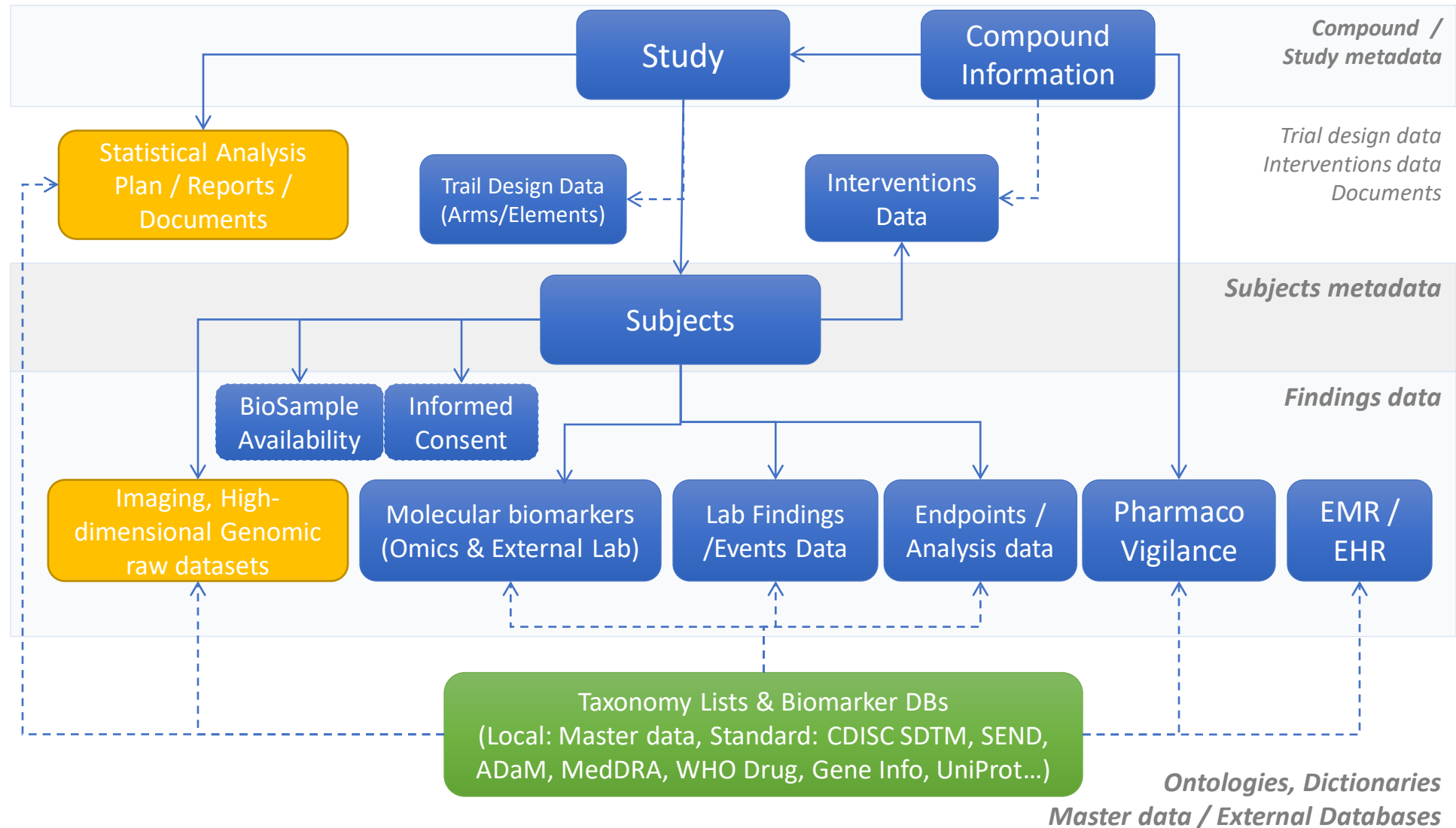


# Smart Curation and Transformation



# Metadata Repository: Conceptual Data Model (UDM)

Simple, Extensible and Ready for analytics



# Metadata Example: Analysis Results

Metadata Field	Metadata
DISPLAY IDENTIFIER	Table 12.3.1.1
DISPLAY NAME	Mean NRS Pain Score Over the Last 5 Days for Overall Pain. Full Analysis Set
RESULT IDENTIFIER	Treatment difference results (Mean, confidence interval, p-value)
PARAM	Overall Pain Score during the 5-day Period
PARAMCD	PLPNOV
ANALYSIS VARIABLE	CHG, BASE, TRT02AN, GEOREGN
REASON	Primary efficacy analysis as pre-specified in protocol
DATASET	ADQS
SELECTION CRITERIA DOCUMENTATION	fas1fl='Y', paramcd='PLPNOV', trt01pn~=., avisit='EoT' See Protocol Section XX for details. Program: program_ex1.sas. NRS scores were analysed using an ANCOVA model which included dose group and region (REG1 and REG2) as fixed factors and baseline NRS pain score of overall pain as a covariate.
PROGRAMMING STATEMENTS	<pre> data pain;   set adam.adqs;   where fas1fl='Y' and paramcd="PLPNOV" and   avisit="EoT"; run;  proc mixed data=pain;   class &amp;trt georegn;   model chg=base &amp;trt georegn;   lsmeans &amp;trt/cl adjust=dunnett;   estimate 'Linear trend' &amp;trt -2 -1 0 1 2;   ods output type3=pvalue;   ods output lsmeans=lsm;   ods output diffs=dif;   ods output estimates=trend; run; </pre>

TFL Metadata

TFL Specifications

ADaM Metadata

Protocol / SAP

SAS Code

# Monitoring and Accretive Resolution of Data Issues



## Clinical Data Issues

- Missing Data Values
- Invalid Dates and Data Values
- Character / Numeric Variable Type Conversion
- Zero Records



## Study Protocol Data Issues

- Standardized Terms – no loss of data or context
- Lab Data – identify duplicate records, missing values, invalid units, etc.
- Primary Endpoints – correctly derived
- Survival Analysis – subgroup analysis
- Safety – maximum patients and events
- Deaths – maximum patients
- Related Adverse Events – minimum patients and events
- Protocol Compliance – visit in visit window range

The screenshot displays the Xbiom Platform interface for cohort and biosample selection. The breadcrumb trail at the top indicates the current view: **Study Summary > Subject History-Baseline > Clinical Parameters > Molecular Biomarkers > Biosamples > UDM Variables**. The left sidebar contains a navigation menu with categories such as Study Drug, Events, Findings About Interventions/Events, Test Findings, Tumor Findings, and Interventions. The main content area shows a search criteria input field with 'Temperature' selected. Below the input field, there is a range filter for 'Range in C' with a minimum value of 40.02 and a maximum value of 98.92. A yellow callout box highlights the range filter, and another yellow callout box explains the six levels of cohort variables.

**Six Levels to Create Cohorts Variables**

**Useful for grouping or subsetting TLFs**

**User Defined Temperature Range**

### Create Patient Cohort Query Across Any Domain Variables, Ex. CHOLFLG

- Subject Population: TRT01A='Drug A', SAFFL='Y'
- Safety Population: AEREL='Y', AESER='Y'
- Efficacy Population: PARAMCD='CHOL', VISIT='12 WEEK', PCHG=-30

### Data and Visual Cohort Review and TLFs

- Cohort, Subject, Multiple-Endpoints, Summary, Statistics, Cross-Reference, Comparisons, Safety Vitals, Efficacy Vitals, Demographics





# Auto-Mapping and Continuous Learning Process

**CDISC 360:** Apply the 80/20 rule to ensure the **Project automates 80%** of the end-to-end metadata and data processing needed to generate study artifacts suitable for a regulatory submission.

## Overall Process

- Pre-processing Batch
  - **Variable** Mapping Methods
  - Control Terms Mapping Methods
  - User Approval Methods
  - New Variable Derivations
- Data Update Batch



## **A. Variable Mapping Methods**

1. Direct
2. Transformation, SQL, ex. trim, concatenating
3. Transpose to Vertical Structure
4. One Raw Data to Multiple SDTMs
5. Multiple Raw Data to One SDTM

## **B. Control Terms Mapping Methods**

1. Exact Value Match
2. Approximate Value Match

## **D. 100% User Approval Methods**

1. Machine Recommended\* 
2. Previous Decision\*\* 
3. Preview Raw data and SDTM standard values
4. SUPPXX, RELREC, FA

## **E. New Variable Derivations**

DY, STDY, ENDY, DTC, SDTC, ENDTC, BLFL, VISIT

\* Learn from sample studies, \*\* Learn from clinical studies



# SDTM Mapping Automation Process Examples

Source file metadata

Column Name	Label
PTNO	Patient Number
SEX	Sex
DOB	Date of Birth
INVSITE	Site



Mapping Recommendations



Recommended mappings
PTNO → DM.SUBJID
SEX → DM.SEX
DOB → DM.BRTHDTC
INVSITE → DM.SITEID

Source data file

PARAM	MATRIX
Basophils	Blood
Eosinophils	Blood
Leukocytes	Blood
Glucose	Urine



Mapping Recommendations



Recommended mappings
PARAM → LB.LBTEST
MATRIX → LB.LBSPEC

Transformed dataset

LBSPEC	LBTEST	LBORRESU
Urine	CRE	
Blood	WBC	
	UCREA	
	UGLUCR	
Blood	BASO	G/L
Blood	BASO	%



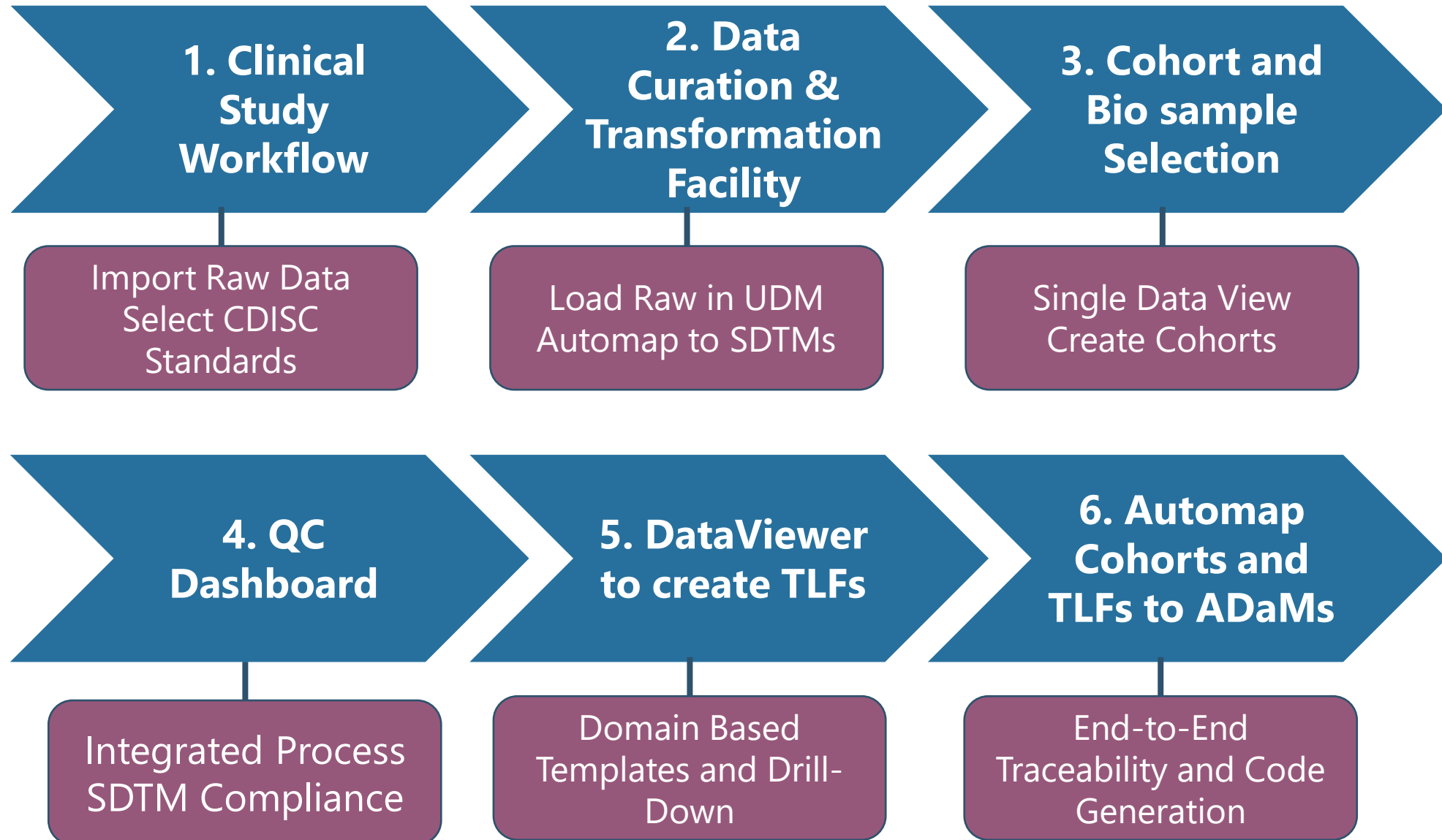
Terminology Recommendations



Recommended Terms

LBSPEC	LBTESTCD	LBTEST
Urine	CREAT	Creatinine
Blood	WBC	Leukocytes
Urine	CREAT	Creatinine
Urine	GLUCCRT	Glucose/Creatinine
Blood	BASO	Basophils
Blood	BASOLE	Basophils/Leukocytes

# The Xbiom Six Step Visualization Process





# Clinical Study Workflow Module

### Cohort and Biosample Selection

Cohort and Biosample Selection



### VDR

is a solution that provides Pharma companies, 's and LIMS systems with secure collaboration and exchange capabilities through hosted data 's while allowing users to work from within their enterprise network firewalls.



### Analysis Workbench

Analysis Workbench



### Folder Templates

mage folder template for Submission and Study data package



### Documents

Search and download the documents across multiple types of information - structured and unstructured. It also helps the user to search the documents pertained to the selected studies.



### Curation & Transformation Facility

Transform clinical, nonclinical and biomarker data from data-lakes to a target model with automation. Recommendation engines using ontologies and vocabularies referenced in the target data model definition harmonizes the transformed data.



### Clinical Study Workflow

Create studies, assign users, import data and statistical analysis deliveries, perform validation checks, acknowledge/reject data through study workflow



### Standard Model Management

Manage standard and sponsor-defined data model

- My Workspace
- All Studies

Production | Bonda,

Study Listing :: All Studies

sh [Create New Study](#) [Add to My Workspace](#) [Data Viewer](#)

Study ID			Project ID/Name	Study Security
<a href="#">CLINICAL_DEMO</a>				Restricted
<a href="#">SDTM01</a>				Restricted
<a href="#">PCSTUDY01</a>				Restricted
<a href="#">TEST</a>				Restricted

**Clinical Study Workflow**

Model :: Tabulation

Study ID: PC202301

Initiate **Data Imp**

- Study Details
- Role Allocation
- Study Specification**
- Study Data Package Listing
- Files and Folders

**Define Version \* :**

2.1

**IG Version \* +**

	IG Name	Default	IG Status	Comments
	SDTMIG 3.4	<input checked="" type="checkbox"/>	Final	
	SDTMIG-PGX 1.0	<input checked="" type="checkbox"/>	Final	
	SDTMIG-MD 1.0	<input checked="" type="checkbox"/>	Final	

**CT Version \* +**

	CT Name	Default	CT Status	Comments
	SDTM Terminology 2022-09-30	<input checked="" type="checkbox"/>	Final	

**Validation Rules**

PointCross Data Validator Rules \* :

FDA Validator Rules 1.5 x PMDA Rules for SDTM 3.0 x

PointCross Define Validator Rules:

CDISC Define Conformance Rules x

The screenshot displays the Xbiom Platform interface. On the left, a file explorer shows a tree view under 'Study Data Package'. The 'Raw Data' folder is expanded, and the 'EDC' folder is selected. On the right, a detailed view of the 'EDC' folder shows a list of CSV files. The table below represents the data shown in the right-hand pane.

	File Name	File Size	Version
<input type="checkbox"/>	<a href="#">AE.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">AE_YN.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">BMBLL1.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">BT.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">BT_YN.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">CM.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">CM2.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">CM2_YN.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">CM3.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">CM3_YN.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">CM_YN.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">CRS_IRR.csv</a>	231.62 KB	1.00
<input type="checkbox"/>	<a href="#">DM.csv</a>	231.62 KB	1.00

- Import EDC, PK and other Data Files into Metadata Repository System
- Automatic Conversions
- XPTs, CSV files
- One excel file with multiple sheets into unique raw data files





# Smart Transformation Module

- ✓ Access Xbiom Metadata Repository System
- ✓ Access Raw EDC/Biomarker Data Source using UDM

Batch Details

**Batch Details:** Provide input (source) and output (target model and CT)

Auto Generate

**Auto Generate:** Used if STUDYID, DOMAIN, SEQ (Sequence) values to be automatically generated by system.

Dataset Transformations

**Scripts for any Transformations:** 3 sections (Dataset Transformations, Additional Transformations, Additional Scripts) are provided to write any custom scripts in SQL, Python or PySpark languages for file processing or for any data derivations or corrections.

Metadata Mappings

**Metadata Mappings:** Used to map source data structure to target model domains and columns. System recommends mappings based on training sets and users' previous decisions. User can approve or modify the recommended mappings.

Terminology Normalization

**Terminology Normalization:** Used to map the source terms to target terms. External dictionaries like MedDRA, NCBI Gene Info, UniProt, mirBase, HMDB also supported. System recommends mappings to target terms based on Xbiom global CT and loaded external dictionaries. User can approve or modify the recommended mappings.

Controlled Terminology

MedDRA

NCBI Gene Info

UniProt

mirBase

HMDB

Additional Transformations

Derivations

**Derivations:** to derives the data, if missed to collect in source systems.

Additional Scripts

Data Updates

**Data Updates:** To perform custom data updates.

Output

Batch Details

Auto Generate

Dataset Transformations

Metadata Mappings

Terminology Normalization

Controlled Terminology

MedDRA

NCBI Gene Info

UniProt

mirBase

HMDB

Additional Transformations

Derivations

Additional Scripts

Data Updates

Output

Batch Identifier \* :

EDC to UDM

Input Data Source \* :

Clinical Study Data Repository

Data Package \* :

Study Data

Study List:

PCSTUDY01 x

Input folder Path:

Auto x

## Batch Output Details

Output Data Location \* :

Clinical Study Data Repository

Output folder Path \* :

Auto

CT Version \* :

GLOBAL SDTM Terminologies x

Output Model Name \* :

UDM x

Batch Details

Auto Generate

Dataset Transformations

Metadata Mappings

Terminology Normalization

Controlled Terminology

MedDRA

NCBI Gene Info

### Auto Generate

Auto Generate

STUDYID

DOMAIN

--SEQ

Batch Details

Auto Generate

**Dataset Transformations**

Metadata Mappings

Terminology Normalization

Controlled Terminology

MedDRA

NCBI Gene Info

UniProt

mirBase

HMDB

Additional Transformations

Derivations

Additional Scripts

Data Updates

Output

New Transformation +

- SQL Query
- Python Script
- PySpark

SV\_Derivation

tl2

nml2

SQL Query DS\_EOT

Select Files: ds x eot x

```
1 select
eot.*,d.DSSTDAT,d.DSSTDAT_RAW,d.DSSTDAT_INT,d.DSSTDAT_YYYY,d.DSSTDAT_MM,d.DSSTDAT_DD,d.DSDHDAT,d.DSDHDAT_YYYY,d.DSDHDAT_MM,d.DSDHDAT_DD,d.DSTERM,d.DSTERM_STD,d.DSDHCAUS,d.DSDHCAUS_STD,d.DSOTHSP from EOT left join (select
SUBJECT,DSSTDAT,DSSTDAT_RAW,DSSTDAT_INT,DSSTDAT_YYYY,DSSTDAT_MM,DSSTDAT_DD,DSTERM,DSTERM_STD,DSDHDAT,DSDHDAT_YYYY,
DSDHDAT_MM,DSDHDAT_DD,DSDHCAUS,DSDHCAUS_STD,DSOTHSP from ds) d on eot.SUBJECT=d.SUBJECT
```

Script

Output File Name: DS\_EOT

Preview Choose Study: SDTM01 Number of Records 10, Displaying 10 Records

SITENUMBER	SITEGROUP	INSTANCEID	INSTANCENAME	INSTANCEREPEATNUMBER	FOLDERID	FOLDER	FOLDERNAME
3	World	5099	End of Treatment Visit (1)	0	13161	EOT	End of Treatment V
42	World	6985	End of Treatment Visit (1)	0	13161	EOT	End of Treatment V
32	World	6875	End of Treatment Visit (1)	0	13161	EOT	End of Treatment V
3	World	7480	End of Treatment Visit (1)	0	13161	EOT	End of Treatment V
42	World	8477	End of Treatment Visit (1)	0	13161	EOT	End of Treatment V

Preview of script output

Batch Details

Auto Generate

Dataset Transformations

**Metadata Mappings**

Terminology Normalization

- Controlled Terminology
- MedDRA
- NCBI Gene Info
- UniProt
- mirBase
- HMDB

Additional Transformations

Derivations

Additional Scripts

Data Updates

Output

ae 302 103 1 Target Domain: AE

#	Source	Mapping	Target ↓	Recommend...	N...
+	44	AESEV	→	AESEV	
+	48	AESER	→	AESER	
+	62	AEREL	→	AEREL	
+	82	AETERM_PT_CODE	→	AEPTCD	
+	42	AEOUT	→	AEOUT	
+	80	AETERM_LL_T_CODE	→	AELLTCD	
+	79	AETERM_LL_T	→	AELLT	
+	78	AETERM_HLT_CODE	→	AEHLTCD	
+	77	AETERM_HLT	→	AEHLT	
+	76	AETERM_HLGT_CODE	→	AEHLGTC	

Preview Choose Study: SDTM01

Source					Target				
AESEV	AESER	AEREL	AETERM_PT_CODE	AEOUT	AESEV	AESER	AEREL	AEPTCD	AEOUT
Grade 1 Mild	No	Related	10013911	Not recovere	Grade 1 Mild	No	Related	10013911	Not recovered/not resolv
Grade 2 Moder...	No	Related	10001551	Recovered/re	Grade 2 Moder...	No	Related	100015...	Recovered/resolved
Grade 1 Mild	No	Related	10001551	Recovered/re	Grade 1 Mild	No	Related	100015...	Recovered/resolved
Grade 2 Moder	No	Not Relat	10028372	Not recovere	Grade 2 Moder	No	Not Relat	100283	Not recovered/not resolv

- User Interface to Confirm Automapping Structures (Variables, Units, Values)
- User to define non-auto mapped Raw Data

- ✓ **ae**

#	Source
+	44 AESEV
+	48 AESER
+	62 AEREL
+	82 AETERM_PT
+	42 AEOUT
+	80 AETERM_LL
+	79 AETERM_LL
+	78 AETERM_HL
+	77 AETERM_HL
+	76 AETERM_HL

AEREL	AEREL_STD	AEACN	AEACN_STD	AECMYN	AECMYN_STD	AECRSIRR	AECRSIR
Related	RELATED	Dose not chang...	DOSE NOT CHANGED	No	N	No	N
Related	RELATED	Dose not chang...	DOSE NOT CHANGED	No	N	No	N
Related	RELATED	Dose not chang...	DOSE NOT CHANGED	No	N	No	N
Not Relat...	NOT RELATED	Not applicable	NOT APPLICABLE	No	N	No	N
Related	RELATED	Dose not chang...	DOSE NOT CHANGED	No	N	No	N
Related	RELATED	Dose not chang...	DOSE NOT CHANGED	No	N	No	N
Not Relat...	NOT RELATED	Not applicable	NOT APPLICABLE	No	N	No	N
Related	RELATED	Dose not chang...	DOSE NOT CHANGED	No	N	No	N
Related	RELATED	Dose not chang...	DOSE NOT CHANGED	No	N	No	N
Related	RELATED	Dose not chang...	DOSE NOT CHANGED	No	N	No	N
Not Relat...	NOT RELATED	Dose not chang...	DOSE NOT CHANGED	No	N	No	N
Related	RELATED	Dose not chang...	DOSE NOT CHANGED	No	N	No	N

> Preview ▶ Choose Study:

+ ✓ **antidrugantibody** 129 35 1 Target Domain:

+ ✓ **cm** 344 91 1 Target Domain:

ae 302 103 1 Target Domain: AE

#	Source	Mapping	Target ↓	Fe
+	44	AESEV	→	AESEV
+	48	AESER	→	AESER
+	62	AEREL	→	AEREL
+	82	AETERM_PT_CODE	→	AEPTCD
+	42	AEOU	→	AEOU
+	80	AETERM_LL	→	AELLTCD
+	79	AETERM_LL	→	AELLT
+	78	AETERM_HLT_CODE	→	AEHLTCD
+	77	AETERM_HLT	→	AEHLT
+	76	AETERM_HLGT_CODE	→	AEHLGTC

AE  
AE  
AG  
AP--  
APDM  
APFA  
APFA--  
APRELSUB  
BE  
BS  
CE

Preview Choose Study: SDTM01

Source	AESEV	AESER	AEREL	AETERM_PT_CODE	AEOU
	Grade 1 Mild	No	Related	10013911	Not recovered
	Grade 2 Moder...	No	Related	10001551	Recovered/re
	Grade 1 Mild	No	Related	10001551	Recovered/re
	Grade 2 Moder...	No	Related	10000270	Not recovered

Target	AESEV	AESER	AEREL	AEPTCD	AEOU
	Grade 1 Mild	No	Related	10013911	Not recovered/not resolv
	Grade 2 Moder...	No	Related	100015...	Recovered/resolved
	Grade 1 Mild	No	Related	100015...	Recovered/resolved
	Grade 2 Moder...	No	Related	100002...	Not recovered/not resolv



ae 302 103 1 Target Domain: AE

#	Source	Mapping	Target ↓	Recommend...	N...
44	AESEV	→	AESEV	📄	✕
48	AESER	→	AESER	📄	✕
62	AEREL	→	AEREL	📄	✕
82	AETERM_PT_CODE	→	AETERM_PT_CODE	📄	✕
42	AEOUT	→	AEOUT	📄	✕
80	AETERM_LLT_CODE	→	AETERM_LLT_CODE	📄	✕
79	AETERM_LLT	→	AETERM_LLT	📄	✕
78	AETERM_HLT_CODE	→	AETERM_HLT_CODE	📄	✕
77	AETERM_HLT	→	AETERM_HLT	📄	✕
76	AETERM_HLGT_CODE	→	AETERM_HLGT_CODE	📄	✕

Preview Choose Study: SDTM01

Source					Target				
AESEV	AESER	AEREL	AETERM_PT_CODE	AEOUT	AESEV	AEREL	AEPTCD	AEOUT	
Grade 1 Mild	No	Related	10013911	Not recovered	Grade 1 Mild	Related	10013911	Not recovered/not resolved	
Grade 2 Moderate	No	Related	10001551	Recovered/resolved	Grade 2 Moderate	Related	10001551	Recovered/resolved	
Grade 1 Mild	No	Related	10001551	Recovered/resolved	Grade 1 Mild	No	Related	10001551	Recovered/resolved
Grade 2 Moderate	No	Not Related	10028372	Not recovered	Grade 2 Moderate	No	Not Related	10028372	Not recovered/not resolved

vs 1194 75 1 Target Domain: [dropdown] [refresh icon]

#	Source	Mapping	Target	Recommend...	N...
37	VSDAT_DD				
38	HEIGHT_VSORRES	✗	VS:VSTESTCD	✓	
	HEIGHT_VSORRES	↶	VS:VSTESTCD	✓	
39	HEIGHT_VSORRES_RAW				
40	HEIGHT_VSORRESU				
41	HEIGHT_VSORRESU_STD				
42	WEIGHT_VSORRES	✗	VS:VSTESTCD	✓	
	WEIGHT_VSORRES	↶	VS:VSTESTCD	✓	
43	WEIGHT_VSORRES_RAW				
44	WEIGHT_VSORRESU				

Preview Choose Study: SDTM01

Source				Target						
HEIGHT_VSORRES	HEIGHT_VSORRES_RAW	HEIGHT_VSORRESU	RESP_VSORRES	VSTESTCD	VSTEST	VSORRES	VSORRESU	VSSTAT		
			17	HEIGHT_VSORRES	Height	63	inches			
			17	HEIGHT_VSORRES	Height	178.6	cm			
	63	inches	16	RESP_VSORRES	Respirator		Breaths/min	Not Done		
	178.6	cm	16	RESP_VSORRES	Respirator	16	Breaths/min			

vs: Variable level mapping Target Domain: VS

Variables

- \*  DATAPAGEID
- DIABP\_VSORRES
- DIABP\_VSORRES\_RAW
- DIABP\_VSORRESU
- DIABP\_VSORRESU\_STD
- ENVIRONMENTNAME
- FILENAME
- FOLDER
- FOLDERID
- FOLDERNAME
- FOLDERPATH
- FOLDERSEQ
- HEIGHT\_VSORRES
- HEIGHT\_VSORRES\_RAW
- HEIGHT\_VSORRESU
- HEIGHT\_VSORRESU\_STD

### Select Variables to Transpose as Rows

<input type="checkbox"/>	Variable	Label	Condition		Map Var Name to	Map Label to	Map Value to	Additional Variables
<input checked="" type="checkbox"/>	HEIGHT_VSORRES	Height			VSTESTCD	VSTEST	VSORRES	HEIGHT_VSORRESU as VSORRESU
<input checked="" type="checkbox"/>	WEIGHT_VSORRES	Weight			VSTESTCD	VSTEST	VSORRES	WEIGHT_VSORRESU as VSORRESU
<input checked="" type="checkbox"/>	SYSBP_VSORRES	Systolic Blood Pressure			VSTESTCD	VSTEST	VSORRES	"mmHg" as VSORRESU
<input checked="" type="checkbox"/>	DIABP_VSORRES	Diastolic Blood Pressure			VSTESTCD	VSTEST	VSORRES	"mmHg" as VSORRESU
<input checked="" type="checkbox"/>	HR_VSORRES	Heart Rate			VSTESTCD	VSTEST	VSORRES	"Beats/min" as VSORRESU
<input checked="" type="checkbox"/>	RESP_VSORRES	Respiratory Rate			VSTESTCD	VSTEST	VSORRES	"Breaths/min" as VSORRESU

### Select Key Variables Add Expression

<input type="checkbox"/>	Variable	Label		Target Variable
<input checked="" type="checkbox"/>	RECORDID	Internal id for th	→	RECORDID
<input checked="" type="checkbox"/>	MINCREATED	Earliest data cre	→	CREATEDON
<input checked="" type="checkbox"/>	MAXUPDATED	Latest data upd	→	UPDATEDON
<input checked="" type="checkbox"/>	DATAPAGENAME	eCRF page nam	→	VSCAT
<input checked="" type="checkbox"/>	INSTANCENAME	Folder instance	→	VISIT

# Mapping source columns data using an "Expression"

ae 302 103 1 Target Domain: AE

#	Source	Mapping	Target	Recommend...	N...
1	PROJECTID				
2	PROJECT				
3	STUDYID	✘		✓	
4	ENVIRONMENTNAME				✘
5	SUBJECTID	→		✓	✘
	CONCAT_WS('-', 'PCSTUDY', SITEID, S...	→	USUBJID	✓	✘
6	STUDYSITEID				✘
7	SUBJECT				✘
8	SITEID				✘
9	SITE				✘

ae: SUBJECTID Sample Data:

```
1 CONCAT_WS('-', 'PCSTUDY', SITEID, SUBJECTID)
```

\*Note: Use Ctrl + Space for auto complete.

Ok Cancel

Preview Choose Study: SDTM01

Source					Target				
SITEID	SITE	SUBJECTID	INSTANCENAME	INS	USUBJID	VISIT	AESPID	AESTDTC	AEENDTC
26	33-Rose Park Labs	792	Common	127	792	PCSTUDY-26-792	Common	PC12468103	2021-11-17
26	33-Rose Park Labs	792	Common	127	792	PCSTUDY-26-792	Common	PC12468109	2021-11-24
26	33-Rose Park Labs	792	Common	127	792	PCSTUDY-26-792	Common	PC12468110	2021-11-29
2	03- Global Research Institute	680	Common	107	680	PCSTUDY-2-680	Common	PC12373588	2021-09-14

dm 26 69 1 Target Domain: DM

#	Source	Mapping	Target	Recommend...	N...
54	ETHNIC STD				
55	RACE1	✗		✓	
	CASE WHEN (RACE2 = " or RACE2 IS...	Y	RACE	✓	✗
	CASE WHEN (RACE2 = " or RACE2 IS...	Y	RACE1	✓	✗
56	RACE2	→	RACE2	✓	✗
57	RACEOTH	→	RACEOTH	✓	✗
58	DMTBIO				✗
59	DMTBIO_STD				✗
60	FOLDERPATH	✗			
61	FILENAME	✗			

Preview Choose Study: SDTM01

Source	ETHNIC_STD	RACE1	RACE2	RACEOTH	DMT
OR LATINO	NOT HISPANIC OR LATINO	Asian			No
OR LATINO	NOT HISPANIC OR LATINO	Asian	White		
ATINO	HISPANIC OR LATINO	White			
OR LATINO	NOT HISPANIC OR LATINO	White			No

Target	ETHNIC	RACE	RACE1	RACE2	RACEOTH
X	ETHNIC	RACE	RACE1	RACE2	RACEOTH
ile	NOT HISPANIC OR LATINO	Asian			
ile	NOT HISPANIC OR LATINO	MULTIPLE	Asian	White	
ile	HISPANIC OR LATINO	White			
male	NOT HISPANIC OR LATINO	White			

The image shows two overlapping SAS windows. The top window, titled 'dm.xpt', displays a table with the following data:

	BRTHDTC	AGE	AGEU	SEX	RACE	ETHNIC	ACTARMCD	ACTARM
5	1971-12-18	49	YEARS	F	Black	NOT HISPANIC ...	A3:SCF	Screen Failure
6	1956-04-10	65	YEARS	M	MULTIPLE	NOT HISPANIC ...	A2:Step Dose-1 ...	Step Dose-1 Prim
7	1948-04-19	73	YEARS	M	WHITE	NOT HISPANIC ...	A1:Fixed Dose	Fixed Dose
8	1961-02-11	60	YEARS					
9	1969-07-22	51	YEARS					
10	1958-02-21	63	YEARS					
11	1949-07-27	71	YEARS					
12	1969-12-02	51	YEARS					
13	1951-11-18	69	YEARS					
14	1960-05-22	61	YEARS					
15	1962-09-04	59	YEARS					
16	1964-05-02	57	YEARS					
17	1959-02-06	62	YEARS					
18	1976-03-22	45	YEARS					
19	1959-09-27	61	YEARS					
20	1958-02-21	63	YEARS					
21	1959-12-09	61	YEARS					
22	1977-08-28	44	YEARS					
23	1952-02-25	69	YEARS					

The bottom window, titled 'SAS Universal Viewer - [suppdm.xpt]', displays a table with the following data:

	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
1	SDTM01	DM	SDTM01-PC044-113			RACE1	Race 1	Asian
2	SDTM01	DM	SDTM01-PC044-113			RACE2	Race 2	White

cm 344 94 1 Target Domain: CM

#	Source	Mapping	Target	Recommend...	N...
27	MAXUPDATED				✗
28	SAVETS				✗
29	STUDYENV SITENUMBER				✗
30	CMINDC	→	CMINDC	✓	✗
31	CMSPID	→	CMSPID	✓	✗
	'CMSPID'	Y	IDVAR	✓	✗
	'AE'	Y	RDOMAIN	✓	✗
	'AESPID'	Y	RVAR	✓	✗
	CMSPID	Y	RVALUE	✓	✗
32	CMINDC_STD			✓	✗

Preview Choose Study: PCSTUDY

ENUMBER	CMINDC	CMSPID	CMINDC
	AE005-OPACIFIATION OF THE BILATERAL MASTO	AE005	AE005-
	AE004-CERUMEN IMPACTION BILATERAL-START	AE004	AE004-
	AE002-PAIN: LEFT SIDE-RIB/STERNUM-START D.	AE002	AE002-
	PROPHYLAXIS		Prophy

CMINDC	CMSPID	IDVAR	RDOMAIN	RVAR	RVALUE
AE005-OPACIFIATION C	AE005	CMSPID	AE	AESPID	AE005
AE004-CERUMEN IMPA	AE004	CMSPID	AE	AESPID	AE004
AE002-PAIN: LEFT SIDE	AE002	CMSPID	AE	AESPID	AE002
PROPHYLAXIS		CMSPID	AE	AESPID	

DOMAIN	USUBJID	AESEC	AESPID	AETERM
AE	PC003-101	11	AE009	THROMBOEMBOLIC EVENT (DVT)
AE	PC003-101	10	AE008	CANCER RELATED PAIN
AE	PC003-108	10	AE007	JEJUNAL HEMORRHAGE
AE	PC003-101	3	AE005	URINARY TRACT INFECTION
AE	PC003-101	7	AE005	URINARY TRACT INFECTION
AE	PC003-108	9	AE005	ANEMIA
AE	PC003-104	4	AE002	HYPONATREMIA
AE	PC003-108			
AE	PC003-108			

AE

DOMAIN	USUBJID	CMSEQ	CMSPID	CMINDC
CM	PC003-101	8	AE008	AE008-CANCER RELATED PAIN -START DATE 19 ...
CM	PC003-101	11	AE008	AE008-CANCER RELATED PAIN -START DATE 19 ...
CM	PC003-101	12	AE008	AE008-CANCER RELATED PAIN -START DATE 19 ...
CM	PC003-101	14	AE008	AE008-CANCER RELATED PAIN -START DATE 19 ...
CM	PC003-101	17	AE011	AE011-CONSTIPATION -START DATE 19 JAN 2021
CM	PC003-101	20	AE009	AE009-THROMBOEMBOLIC EVENT (DVT) -START ...
CM	PC003-101	21	AE009	AE009-THROMBOEMBOLIC EVENT (DVT) -START ...

CM

relrec.xpt Show Empty Columns

Study ID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID	
PCSTUDY	AE	PC003-108	AESPID	AE001	ONE	1	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	CM	PC003-108	CMSPID	AE001	ONE	1	1-CONSTIPATION -START DATE 19 JAN 2021
PCSTUDY	CM	PC003-101	CMSPID	AE008	ONE	2	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	AE	PC003-101	AESPID	AE008	MANY	2	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	AE	PC003-101	AESPID	AE005	MANY	3	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	CM	PC003-101	CMSPID	AE005	ONE	3	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	CM	PC003-101	CMSPID	AE009	ONE	4	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	AE	PC003-101	AESPID	AE009	MANY	4	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	CM	PC003-108	CMSPID	AE005	ONE	5	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	AE	PC003-108	AESPID	AE005	ONE	5	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	CM	PC003-104	CMSPID	AE002	ONE	6	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	AE	PC003-104	AESPID	AE002	ONE	6	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	CM	PC003-108	CMSPID	AE007	ONE	7	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	AE	PC003-108	AESPID	AE007	ONE	7	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	CM	PC003-108	CMSPID	AE002	ONE	8	9-THROMBOEMBOLIC EVENT (DVT) -START ...
PCSTUDY	AE	PC003-108	AESPID	AE002	ONE	8	9-THROMBOEMBOLIC EVENT (DVT) -START ...

RELREC



+	45	DSDHCAUS
+	46	DSDHDAT
+	47	DSDHDAT_RAW
+	48	DSDHDAT_INT
+	49	DSDHDAT_YYYY
+	50	DSDHDAT_MM
+	51	DSDHDAT_DD
+	52	<b>DSDHCAUS</b>
+		DSDHCAUS

Target Domain: FA--

Select Variables to Transpose as Rows

Variable	Label	Condition	Map Var Name to	Map Label to	Map Value to	Additional
DSDHCAUS		'DSDHCAUS' IS NOT NULL			FAORRES	

Select Key Variables Add Expression

Variable	Label	Target Variable
RECORDID		FASPID
'DS'		RDOMAIN
'DSSPID'		RVAR
'DEATH'		FAOBJ
regex_replace(SUBJECT,'CLINIC')		USUBJID

FA--:FAORRES

Preview Choose Study: PCSTUDY

Source

SEID	DATAPAGENAME	RECORDID	PAGEREPEATNUMBER	RECORDDATE
	End of Study	PC1295107	0	
	End of Study	PC12213640	0	
	End of Study	PC12220517	0	
	End of Study	PC12463272	0	

Target

DS	FA--	IDVAR	RVAR	RVALUE	FAOBJ	USUBJID
FASPID	RDOMAIN	FASPID	DSSPID	PC1295107	DEATH	PC1295107
PC1295107	DS	FASPID	DSSPID	PC12213640	DEATH	PC12213640
PC12213640	DS	FASPID	DSSPID	PC12220517	DEATH	PC12220517
PC12220517	DS	FASPID	DSSPID	PC12463272	DEATH	PC12463272

USUBJID	DSSEQ	DSSPID	DSTERM	DSDECOD	DSCAT	DSSCAT	DSSTDTC
PC030-001	1	PC12551839	Patient's request (withdrawal of...	WITHDRAWAL ...	END OF TREAT...	DISPOSITION E...	2022-01-19
PC003-104	1	PC12220764	Death	DEATH	END OF TREAT...	DISPOSITION E...	2022-03-12
PC003-104	1	PC12663800	Patient's request (withdrawal of...	WITHDRAWAL ...			
PC003-104	1	PC12213640	Death	DEATH			
PC042-106	1	PC12218105	Lost to follow-up	LOST TO FO...			
PC044-109	1	PC12463272	Death	DEATH			
PC042-103	1	PC12468779	Death	DEATH			
PC003-101	1	PC1295107	Death	DEATH			
PC003-108	1	PC12220517	Death	DEATH			
PC044-110	1	PC12228343	Death	DEATH			

USUBJID	FASPID	FATESTCD	FATEST	FAORES	FASTRES
PC003-101	PC1295107	DSDHCAUS	CAUSE OF DEATH	Disease progressi...	Disease pr...
PC003-104	PC12213640	DSDHCAUS	CAUSE OF DEATH	DEATH	Disease progressi...
PC003-108	PC12220517	DSDHCAUS	CAUSE OF DEATH	DEATH	Disease progressi...
PC042-103	PC12468779	DSDHCAUS	CAUSE OF DEATH	DEATH	Other
PC042-107	PC12220764	DSDHCAUS	CAUSE OF DEATH	DEATH	Disease progressi...
PC044-109	PC12463272	DSDHCAUS	CAUSE OF DEATH	DEATH	Disease progressi...
PC044-110	PC12228343	DSDHCAUS	CAUSE OF DEATH	DEATH	Disease progressi...

DS

FADS

RELREC

PCSTUDY	CM	PC003-101	CMSPID	AE009	ONE	4
PCSTUDY	AE	PC003-101	AESPID	AE009	MANY	4
PCSTUDY	CM	PC003-108	CMSPID	AE005	ONE	5
PCSTUDY	AE	PC003-108	AESPID	AE005	ONE	5
PCSTUDY	CM	PC003-104	CMSPID	AE002	ONE	6
PCSTUDY	AE	PC003-104	AESPID	AE002	ONE	6
PCSTUDY	CM	PC003-108	CMSPID	AE007	ONE	7
PCSTUDY	AE	PC003-108	AESPID	AE007	ONE	7
PCSTUDY	CM	PC003-108	CMSPID	AE002	ONE	8
PCSTUDY	AE	PC003-108	AESPID	AE002	ONE	8
PCSTUDY	FA	PC003-101	FASPID	PC1295107	ONE	9
PCSTUDY	DS	PC003-101	DSSPID	PC1295107	ONE	9
PCSTUDY	DS	PC003-108	DSSPID	PC12220517	ONE	10
PCSTUDY	FA	PC003-108	FASPID	PC12220517	ONE	10
PCSTUDY	FA	PC003-104	FASPID	PC12213640	ONE	11
PCSTUDY	DS	PC003-104	DSSPID	PC12213640	ONE	11
PCSTUDY	FA	PC044-110	FASPID	PC12228343	ONE	12
PCSTUDY	DS	PC044-110	DSSPID	PC12228343	ONE	12

Batch Details

Auto Generate

Dataset Transformations

Metadata Mappings

Terminology Normalization

Controlled Terminology

MedDRA

NCBI Gene Info

UniProt

mirBase

HMDB

Additional Transformations

Derivations

Additional Scripts

Data Updates

Output

## Controlled Terminology



+ AE

+ CM

- DM

	Source	Target	Expression	Recommendation
+ AGEU		(AGEU)		
+ ETHNIC		(ETHNIC)		
- RACE		(RACE)		
+ Asian	Asian	ASIAN		Approved ✓
+ Black	Black	BLACK OR AFRICAN AMERIC...		Approximate Match 🗨️
+ MULTIPLE	MULTIPLE	MULTIPLE		Same term as is →
+ White	White	WHITE		Exact Match 🗨️
+ SEX		(SEX)		

+ DS

+ EG

Batch Details

Auto Generate

Dataset Transformations

Metadata Mappings

Terminology Normalization

Controlled Terminology

MedDRA

NCBI Gene Info

UniProt

mirBase

HMDB

Additional Transformations

**Derivations**

Additional Scripts

## Derivations

Select the variables required for derivation

Reference Day \* :  ⓘ

--DY     --STDY     --ENDY     --DTC     --STDTC     --ENDTC     --BLFL     VISIT

--TPTREF     --RFTDTC    Domains  ⓘ

Populate --STRESN --STRESC --STRESU

Derive using Standard Units

Copy ORRES ORRESU

Batch Details

Auto Generate

Dataset Transformations

Metadata Mappings

Terminology Normalization

Controlled Terminology

MedDRA

NCBI Gene Info

UniProt

mirBase

HMDB

Additional Transformations

Derivations

Additional Scripts

Data Updates

Output

### Output

Original

Converted

Transformed

SDTM01

- [ae.csv](#)
- [ae\\_yn.csv](#)
- [antidrugantibody.csv](#)
- [bmbll1.csv](#)
- [bt.csv](#)
- [bt\\_yn.csv](#)
- [cm.csv](#)
- [cm2.csv](#)
- [cm2\\_yn.csv](#)
- [cm3.csv](#)
- [cm3\\_yn.csv](#)
- [cm\\_yn.csv](#)
- [crs\\_irr.csv](#)

SDTM01

- [ae.xpt](#)
- [cm.xpt](#)
- [co.xpt](#)
- [dm.xpt](#)
- [ds.xpt](#)
- [dv.xpt](#)
- [ec.xpt](#)
- [eg.xpt](#)
- [ex.xpt](#)
- [fa.xpt](#)
- [ho.xpt](#)
- [ie.xpt](#)
- [is.xpt](#)

**Excel file can be read by SAS programs to convert Raw data to SDTMs**

Raw SDTM Datasets

Variable Derivations

SDTM Variables

	A	B	C	D	E	F	G	H
1	Source File Name	Target Domain	Source Column Name	Source Column Label	Mapping	Target Variable	Status	Parent Column
307	cm	CM	INSTANCENAME		Direct	VISIT	Approved	
324	cm	CM	CMINDC		Direct	CMINDC	Recommended (Previous Decision)	
333	cm	CM	CMONGO_STD		Direct	CMMODIFY	Recommended	
341	cm	CM	CMDOSU		Direct	CMDOSU	Recommended (Previous Decision)	
351	cm	CM	CMTRT		Direct	CMTRT	Recommended (Previous Decision)	
356	cm	CM	CMTRT_ATC2		Direct	CMSCAT	Approved	
360	cm	CM	CMTRT_ATC4		Direct	CMCLAS	Approved	
361	cm	CM	CMTRT_ATC4_CODE		Direct	CMCLASCD	Approved	
372	cm	CM	substring(CMINDC,3,3)		Expression	RVALUE	Approved	CMINDC
373	cm	CM	'SPID'		Expression	RVAR	Approved	CMINDC
374	cm	CM	"		Expression	POOLID	Recommended (Previous Decision)	SITEGROUP
375	cm	CM	case when CMDOSFRQ='Other' then CONCAT_WS(':', 'Other', CMFRSPEC) else CMDOSFRQ end		Expression	CMDOSFRQ	Approved	CMDOSFRQ
376	cm	CM	case when CMROUTE='Other' then CONCAT_WS(':', 'Other', CMRTSPEC) else CMROUTE end		Expression	CMROUTE	Approved	CMROUTE
377	cm	CM	case when CMTRT_PRODUCT<>" then CMTRT_PRODUCT else CMTRT end		Expression	CMDECOD	Approved	CMTRT_PRODUCT

Metadata Mappings

Transpose Mapping

Controlled Terminology

MedDRA

Additional Transformation

Additional Script

Data l ...



# Clinical Study Workflow Module

Initiate

**Data Import**

Study Details

Role Allocation




Study Specification

Study Data Package Listing

**Files and Folders**








## Files and Folders

Refresh

Study Data Package		
▼ SDTM01		
▼ <input type="checkbox"/> Study Data		☰
▼ <input type="checkbox"/> Tabulations		☰
<input type="checkbox"/> SDTM		☰  
<input type="checkbox"/> UDM data		☰ 
▼ <input type="checkbox"/> Analyzed Biomarkers		☰
<input type="checkbox"/> DNA		☰
<input type="checkbox"/> RNA		☰

Study Data \ Tabulations \ UDM data



<input type="checkbox"/>	File Name	File Size
<input type="checkbox"/>	 <a href="#">AE.csv</a>	184.43 KB
<input type="checkbox"/>	 <a href="#">CM.csv</a>	386.94 KB
<input type="checkbox"/>	 <a href="#">DM.csv</a>	6.85 KB
<input type="checkbox"/>	 <a href="#">DS.csv</a>	9.76 KB
<input type="checkbox"/>	 <a href="#">DV.csv</a>	331.47 KB
<input type="checkbox"/>	 <a href="#">EC.csv</a>	10.53 KB
<input type="checkbox"/>	 <a href="#">EG.csv</a>	244.07 KB



Study Listing :: All Studies    Study ID: SDTM01

Initiate    **Data Import**

- Study Details
- Role Allocation
- Study Specification
- Study Data Package Listing
- Files and Folders**

### Files and Folders

Refresh

Study Data Package

Study Data \ Tabulations \ SDTM

<input type="checkbox"/>	File Name	File Size
<input type="checkbox"/>	<a href="#">_ae.xpt</a>	182.89 KB
<input type="checkbox"/>	<a href="#">_cm.xpt</a>	855.94 KB
<input type="checkbox"/>	<a href="#">_co.xpt</a>	1.8 KB
<input type="checkbox"/>	<a href="#">_define.pdf</a>	177.6 KB
<input type="checkbox"/>	<a href="#">_define.xml</a>	131.81 KB
<input type="checkbox"/>	<a href="#">_define2-1-0.xsl</a>	183.69 KB
<input type="checkbox"/>	<a href="#">_dm.xpt</a>	7.66 KB

- Standards
- ▼ Datasets
  - CO (Comments)
  - DM (Demographics)
  - SV (Subject Visits)
  - CM (Concomitant/Prior Medications)
  - EC (Exposure as Collected)
  - EX (Exposure)
  - PR (Procedures)
  - AE (Adverse Events)
  - DS (Disposition)
  - DV (Protocol Deviations)
  - HO (Healthcare Encounters)
  - MH (Medical History)
  - EG (ECG Test Results)
  - IE (Inclusion/Exclusion Criteria)
  - IS (Immunogenicity Specifications)
  - LB (Laboratory Test Results)
  - PE (Physical Examination)
  - QS (Questionnaires)
  - RS (Disease Response and Signs)
  - TR (Tumor/Lesion Results)
  - TU (Tumor/Lesion Identification)
  - VS (Vital Signs)
  - FA (Findings About Events)
  - RELREC (Related Records)
  - SUPPDM (Supplemental Data)
- ▼ Controlled Terminology
  - CodeLists

Standard	Type	Status	Documentation
SDTMIG 3.4	IG	Final	
CDISC/NCI SDTM 2022-03-25	CT	Final	

### Datasets

Dataset	Description	Class	Structure	Purpose	Keys	Documentation	Location
<a href="#">CO</a> [SDTMIG 3.4]	Comments	SPECIAL PURPOSE	One record per comment per subject	Tabulation	STUDYID		<a href="#">co.xpt</a>
<a href="#">DM</a> [SDTMIG 3.4]	Demographics	SPECIAL PURPOSE	One record per subject	Tabulation	STUDYID, USUBJID		<a href="#">dm.xpt</a>
<a href="#">SV</a> [SDTMIG 3.4]	Subject Visits	SPECIAL PURPOSE	One record per actual or planned visit per subject	Tabulation	STUDYID, USUBJID		<a href="#">sv.xpt</a>
<a href="#">CM</a> [SDTMIG 3.4]	Concomitant/Prior Medications	INTERVENTIONS	One record per recorded intervention occurrence or constant-dosing interval per subject	Tabulation	STUDYID, USUBJID, CMTRT, CMSTDTC		<a href="#">cm.xpt</a>
<a href="#">EC</a> [SDTMIG 3.4]	Exposure as Collected	INTERVENTIONS	One record per protocol-specified study treatment, collected-dosing interval, per subject, per mood	Tabulation	STUDYID, USUBJID, ECSTDTC		<a href="#">ec.xpt</a>
<a href="#">EX</a> [SDTMIG 3.4]	Exposure	INTERVENTIONS	One record per protocol-specified study treatment, constant-dosing interval, per subject	Tabulation	STUDYID, USUBJID, EXSTDTC		<a href="#">ex.xpt</a>
<a href="#">PR</a> [SDTMIG 3.4]	Procedures	INTERVENTIONS	One record per recorded procedure per occurrence per	Tabulation	STUDYID, USUBJID,		<a href="#">pr.xpt</a>

# Define specification editor (Variable level)

Study ID: SDTM01 Study Model : SDTM01 (SDTM)

Summary Dataset Level **Variable Level** Controlled Terminology Value Level Metadata External Dictionaries Computational Algorithm Test Specification Sup

Edit Refresh Save Copy Merge Undo

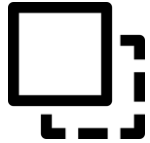
r	Domain*	Variable Name*	Variable Label	Value Coded By	Controlled Terms,Format...	Core	Keys	Data Type	Length
	EC	ECENDTC	End Date/Time of Treatment	Format	<a href="#">ISO 8601 datetime or inter...</a>	Expected		Datetime	
	EG	STUDYID	Study Identifier			Required	1	Text	6
	EG	DOMAIN	Domain Abbreviation	Controlled Terms	< DOMAIN > ["EG"]	Required		Text	2
	EG	USUBJID	Unique Subject Identifier			Required	2	Text	20
	EG	EGSEQ	Sequence Number			Required		Integer	1
	EG	EGTESTCD	ECG Test or Examination Short Name			Required	3	Text	6
	EG	EGTEST	ECG Test or Examination Name			Required		Text	24
	EG	EGCAT	Category for ECG			Permissible		Text	11
	EG	EGORRES	Result or Finding in Original Units			Expected		Text	13
	EG	EGORRESU	Original Units	Controlled Terms	< UNIT > ["beats/min","mins/...	Permissible		Text	9
	EG	EGSTRESC	Character Result/Finding in Std Format			Expected		Text	13
	EG	EGSTRESN	Numeric Result/Finding in Standard ...			Permissible		Float	10
	EG	EGSTRESU	Standard Units	Controlled Terms	< UNIT > ["beats/min","mins/...	Permissible		Text	9
	EG	EGSTAT	Completion Status	Controlled Terms	< ND > ["NOT DONE"]	Permissible		Text	8
	EG	EGMETHOD	Method of Test or Examination	Controlled Terms	EGMETHOD	Permissible		Text	11
	EG	EGBLFL	Baseline Flag	Controlled Terms	NY	Permissible		Text	1
	EG	VISITNUM	Visit Number			Expected	4	Float	5

Summary		Dataset Level		Variable Level		Controlled Terminology		Value Level Metadata							
+ Add		Edit		Replicate		Delete / Inactive		Refresh		Save		Copy		Merge	
Computatio...	Computational Algorithm			Type	Origin T...	Origin ...	Role								
<input type="checkbox"/>	DY	DY = DTC + DM.RFSTDTC - 1, if DTC >= DM.RFSTDTC else DTC - DM.RFSTDTC		Computation			Record Qualifier								
<input type="checkbox"/>	ENDY	ENDY = ENDTC + DM.RFSTDTC - 1, if ENDTC >= DM.RFSTDTC else ENDTC - DM.RFSTDTC		Computation			Record Qualifier								
<input type="checkbox"/>	STDY	STDY = STDTC + DM.RFSTDTC - 1, if STDTC >= DM.RFSTDTC else STDTC - DM.RFSTDTC		Computation			Record Qualifier								
AE	AETOXGR	Standard Toxicity Grade					Record Qualifier								
AE	AESTDTC	Start Date/Time of Adverse Ev...		Format		<a href="#">ISO 8601</a>	Timing								
AE	AEENDTC	End Date/Time of Adverse Event		Format		<a href="#">ISO 8601</a>	Timing								
AE	AESTDY	Study Day of Start of Adverse ...					STDY	Derived	Vendor	Timing					
AE	AEENDY	Study Day of End of Adverse ...					ENDY	Derived	Vendor	Timing					
AE	AEDUR	Duration of Adverse Event		Format		<a href="#">ISO 8601</a>	Timing								
AE	AEENRF	End Relative to Reference Per...		Controlled Terms		<a href="#">STENRF</a>	Timing								
AE	AEENRT...	End Relative to Reference Tim...		Controlled Terms		<a href="#">STENRF</a>	Timing								
AE	AEENTPT	End Reference Time Point					Timing								

# eDV CDISC Compliance: SDTMs and ADaMs



**Traceability:** SDTMs and ADaMs (variables, value level metadata) traced back to SDTM, raw source data, and mapping decisions.



**Reproducibility:** SDTMs and ADaMs may be reproduced using mapping programs. TLF can be reproduced using TLF programs.



**Conformance:** SDTM and ADaM dataset are CDISC conformant for exchangeability i.e. they meet the specific design specifications as well as the general SDTM and ADaM model specifications



**Completeness:** SDTM are complete, i.e. all raw data items that were designed to be mapped are indeed present in the SDTM datasets – not just EDC data. ADaMs are complete, i.e. all analysis variables are mapped from SDTMs.



**Integrity:** Data integrity is preserved, i.e. data points are not inadvertently affected (e.g. by truncation), no loss of records.

Study Details

Role Allocation

Study Specification

Study Data Package Listing

Files and Folders

Refresh

Study Data Package	PointCross Validator Checks	
	Data Conformance Rules	Define.xml Validation
SDTM01		
Study Data		
Tabulations		
SDTM	<a href="#">FDA : Issues 14...</a> <a href="#">PMDA : Error 3420 R...</a> <a href="#">CDISC : Issue 487</a>	
UDM data		
Analyzed Biomark...		

SDTM 28

**Clinical Study Workflow** Production | Gupta, Sunil | 25-OCT-2022 1:19:50 PM (UTC -07:00) | DONT CROSS Life Sciences

Study Listing :: All Studies | Study ID: TEST | QC Dashboard | Business Base : ClinicalDemo

New Arrival: 29 | Approved: 0 | Rejected: 0 | New Arrival [dropdown] [refresh]

Study [Search for Studies]

SDM [Search for Studies]

File [Search for Studies]

Validation [Error] [Warning]

Data Consistency: 1 Error, 2 Warning

Data Format: 0 Error, 0 Warning

Terminology: 3 Warning

Unit Conversion: 0 Warning

Summary [Compare Files] Showing: 29 records [Refresh Summary] [Approve] [Reject]

<input type="checkbox"/>	Study ID	Data Packag...	SDM	File Name	Model E...	# Colum...	# Rows	Rules F...	Errors	Warnings	Termino...	Status	Notes	Modi
<input type="checkbox"/>	TEST	Study Data	SDTM+	<a href="#">AE</a>	98	18	40	3	2	81	3	New Arrival		29-SE
<input type="checkbox"/>	TEST	Study Data		<a href="#">MYTEST...</a>				0	0	0		New Arrival		29-SE
<input type="checkbox"/>	SDTM01	Study Data		<a href="#">EDC TO ...</a>				0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">LB</a>	62	27	10910	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">TR</a>	32	17	839	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">SUPPDM</a>	10	10	2	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">PE</a>	30	13	2037	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">EX</a>	38	14	335	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">VS</a>	38	19	14234	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">QS</a>	35	15	149	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">EC</a>	45	11	93	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">DS</a>	16	10	65	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">DEFINE</a>				0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">CO</a>	13	8	2	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">DV</a>	16	5	4602	0	0	0		New Arrival		11-00
<input type="checkbox"/>	SDTM01	Study Data	SDTMIG-3.4	<a href="#">AE</a>	60	20	332	0	0	0		New Arrival		11-00

UDM maintains a single view of the trial.

After Generating Automated SDTMs, Immediate check for SDTM Compliance with QC Dashboard.

Similar QC Dashboard for ADaM datasets.

Refresh

Study Data Package

Study Data \ Tabulations \ SDTM

SDTM01

- Study Data
- Tabulations
- SDTM**
- UDM data
- Analyzed Biomarkers

File Name	File Size	Version
<a href="#">ae.xpt</a>	182.89 KB	29.00
<a href="#">cm.xpt</a>	855.94 KB	28.00
<a href="#">co.xpt</a>	2.03 KB	4.00
	160.29 KB	1.00
	125.37 KB	1.00

- Re-Execute PointCross Validator Checks
- Generate define.xml
- Upload Files
- Clear Folder Data
- Download**
  - Download Draft SDRG**
  - Download Files and Folders



The screenshot displays the Xbiom Platform interface. On the left, a table of contents is visible under the 'Headings' tab, listing sections from 1. Introduction to Appendix II. The main area on the right shows a document titled 'Clinical Study Data Reviewer's Guide (cSDRG)' for 'Study SDTM01'. The document is a draft, indicated by the red wavy underline under 'cSDRG'. The page number '1' is visible in the top right corner of the document area.

Search document

Headings Pages Results

- 1. Introduction
  - 1.1. Purpose
  - 1.2. Acronyms
  - 1.3. Study Data Standards and Dicti...
- 2. Protocol Description
  - 2.1. Protocol Number and Title
  - 2.2. Protocol Design
  - 2.3. Trial Design Datasets
- 3. Subject Data Description
  - 3.1. Overview
  - 3.2. Traceability Flow Diagram
  - 3.3. Annotated CRFs
  - 3.4. SDTM Subject Domains
- 4. Data Conformance Summary
  - 4.1. Conformance Inputs
  - 4.2. Issues Summary
    - 4.2.1. FDA Validator Rules 1.5
    - 4.2.2. CDISC Define Conforman...
  - 4.3. Additional Conformance Details
- Appendix I: Inclusion/Exclusion Criteria
- Appendix II: Conformance Issues Det...
  - 1. Purpose
  - 2. Conversion Data Flow

Study SDTM01

Clinical Study Data Reviewer's Guide

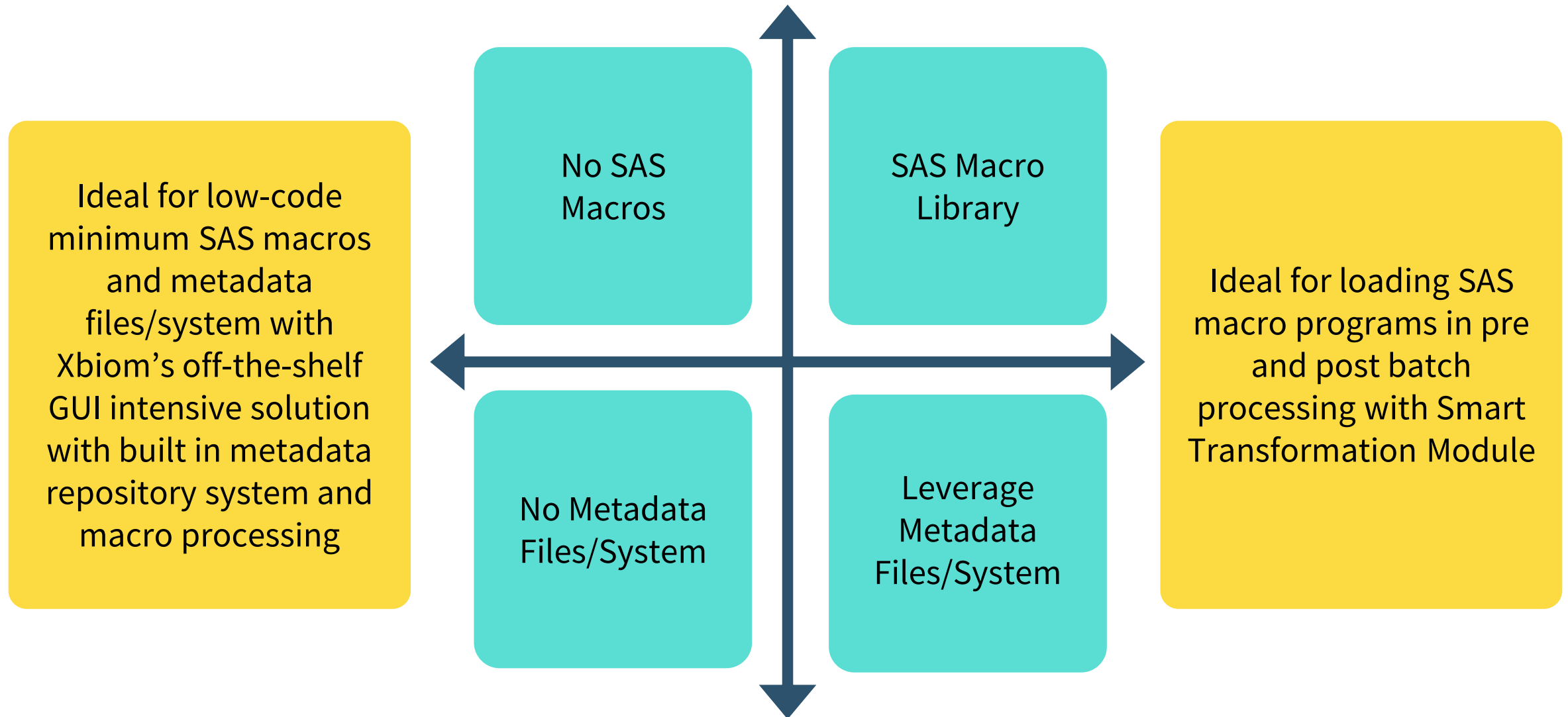
# Clinical Study Data Reviewer's Guide

(cSDRG)

Study SDTM01|

cSDRG Template Version 2023-01-17

# SAS Program Migration Steps to Xbiom Platform



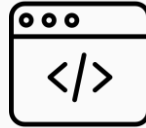
# Xbiom Solution for SDTM Automation & Compliance

## Manage Projects with Automation and Standards



- ✓ One SCE integrate tool for all Submission Deliverables
- ✓ Reduce Time and Budget per Clinical Study

## Manage Submission Process with Low-Code Programming



- ✓ Reduce writing SAS programs and macros
- ✓ Faster SDTMs, Define.xml and SDRG
- ✓ Auto Generate SDTM Mapping Specifications

## Monitor Safety Data Issues with Early Alerts



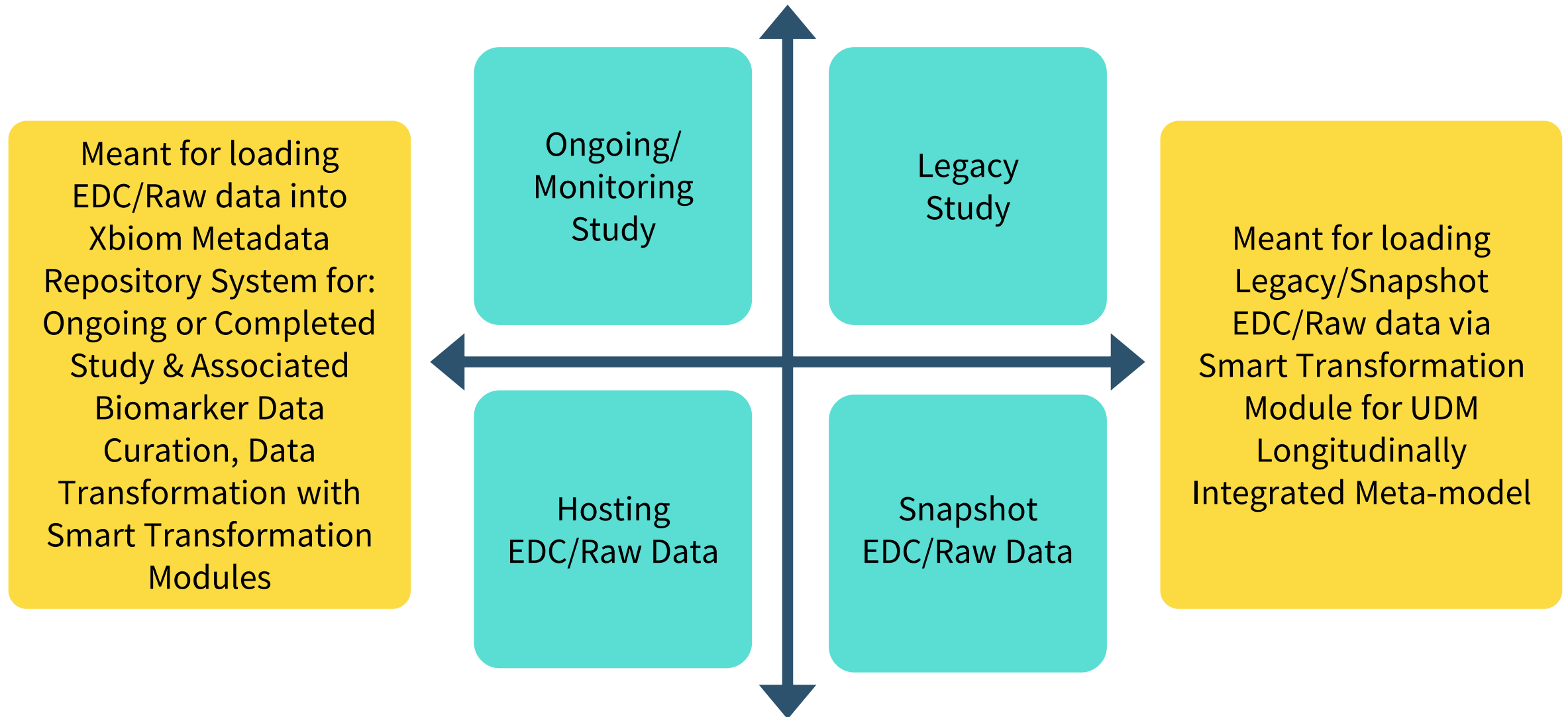
- ✓ Faster Ingestion, Curation and Harmonization
- ✓ User Interface to create SAP Cohorts

## Explore with Pre-defined Templates



- ✓ Reduce Time to Tables, Lists and Figures
- ✓ Drill down from summary to patient level detail

# Xbiom Platform Offers an Easy On and Off Ramp



# Contact Us

[ask@pointcross.com](mailto:ask@pointcross.com)

## Process Flow Charts

- ❖ CDISC-360 Mission: SDTM Design and Automation ([Download PDF](#))
- ❖ End-to-End Clinical Study MetaData-Driven Process ([Download PDF](#))

## Upcoming Events

- ❖ PhUSE Wednesday Webinar, January 25<sup>th</sup> – 10:00 AM ET ([Link](#))
- ❖ PhUSE Connect, Orlando (FL), Software Demonstration, March 5<sup>th</sup> to 8<sup>th</sup>



**Contact Us**  
**Sunil Gupta**  
**CDISC SME, Trainer & Author**  
**Email:** [GuptaProgramming@gmail.com](mailto:GuptaProgramming@gmail.com)



**Contact Us**  
**Rahul Madhavan**  
**VP Strategic Programs**  
**Email:** [Rahul@pointcross.com](mailto:Rahul@pointcross.com)



**Corporate HQ**  
**PointCross LLC**  
**Address:** 1291 E. Hillsdale  
Blvd, Suite 304 Foster City,  
California 94404 USA

[www.pointcrosslifesciences.com](http://www.pointcrosslifesciences.com)

Thank you!

