



EDC Raw Data to SDTM Curation, Mapping and Automation with Xbiom Tool

*“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
CDISC SME, Trainer and Author

Rahul Madhavan
Rahul@PointCross.com
VP – Strategic Programs, PointCross

What is Your SDTM Curation and 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 Unified 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 Unified Data Model (UDM)

- Low-Code User Interface
- Six Step Visualization Process

○ Unified Data Model (UDM) Design

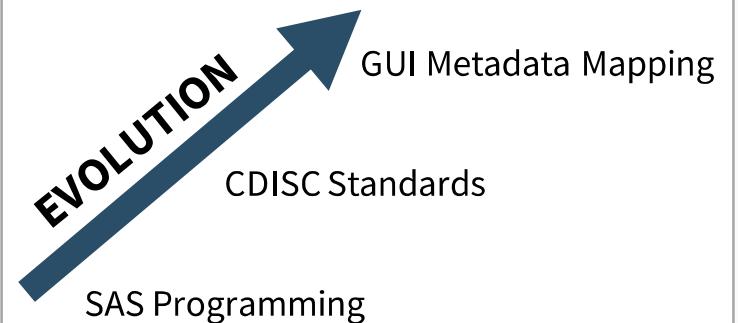
- Reusability and Repeatability
- Direct to Review and Analysis
- Monitoring and Accretive Resolution of Data Issues

○ SDTM Generation

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



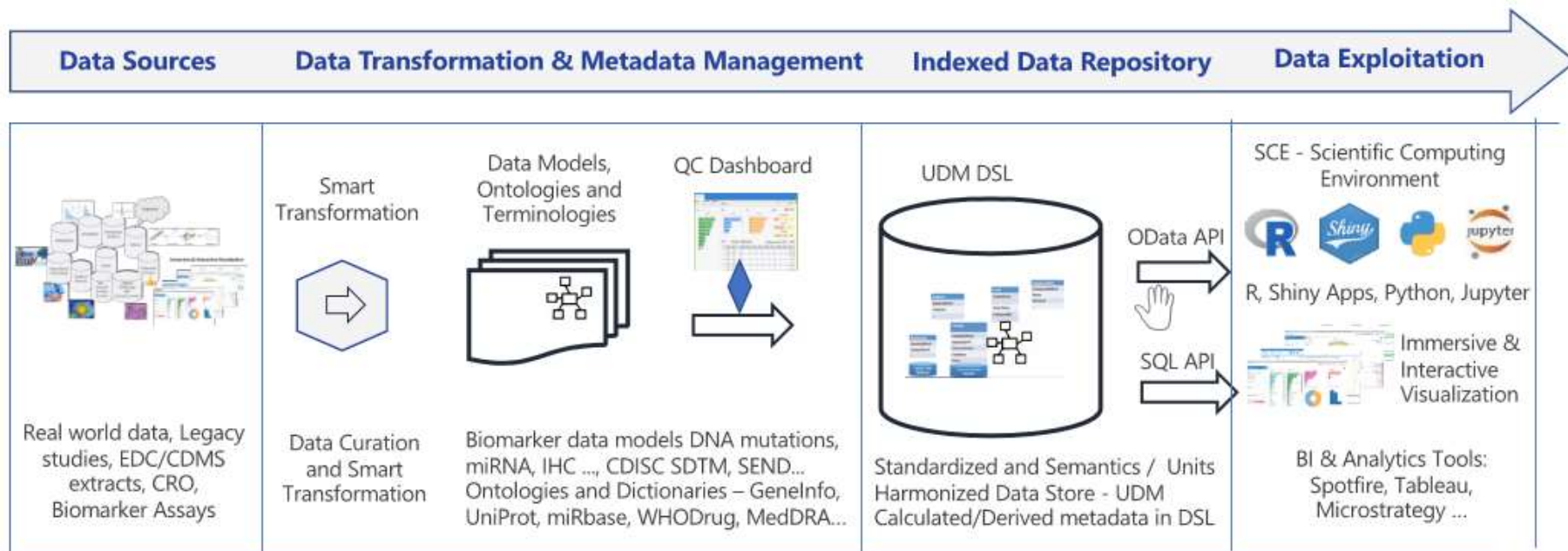
- > Search and Query Interactive Database
- > Ingest, Catalog, Index and Link All Data



Big Data Value Chain



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





CDISC 360 MISSION:

SDTM Design

And Automation

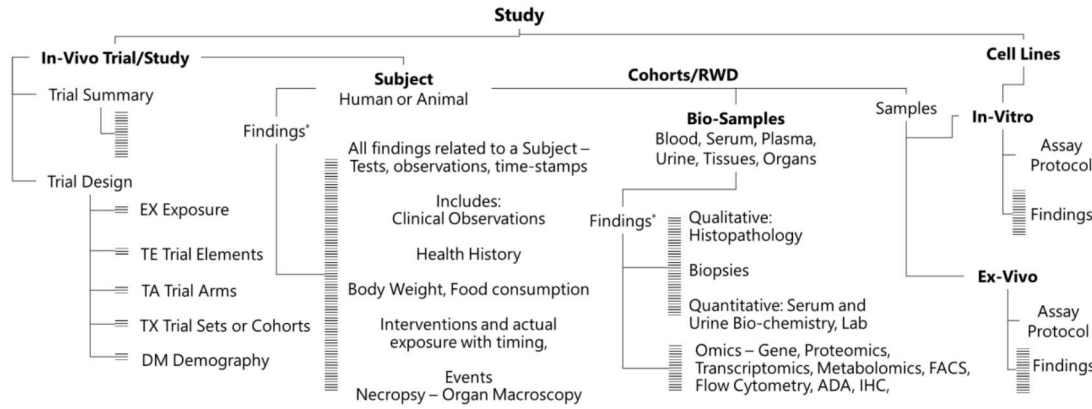
Input: Metadata & Raw Clinical Data

Create End-to-Start Specification

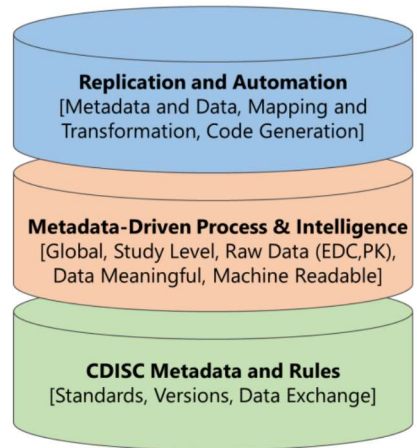
- ☐ Produce a standards-based, machine readable specification

Generate Start-to-End Metadata

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



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



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

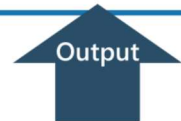
- Use or create utilities to replicate the process: Project Set Up, Mapping Specification, Mapping Creation
- Use analytics tool to identify the areas for replication and automation: Data Profiling & Data Rules for Source Data Review / Edit Checks

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

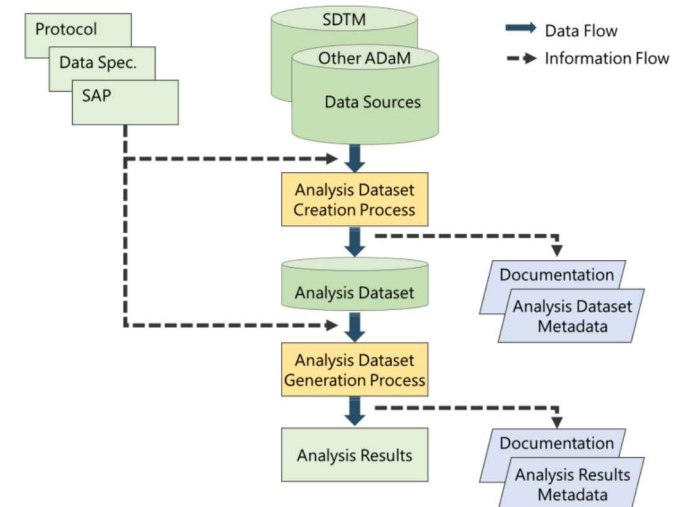
- Train people to understand the standards
- Define standard templates
- Build public libraries for code snippets and public transformation: Custom functions, procedures and packages; public data rules; and public Experts
- Group code snippets and functional transformation into modular mapping and transformation: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



Output: SDTMs, ADaMs, Define.xml & TFLs

Data Curation

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



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

Raw Data, Metadata, Xbiom Tool



CDISC / Analysis



Documentation



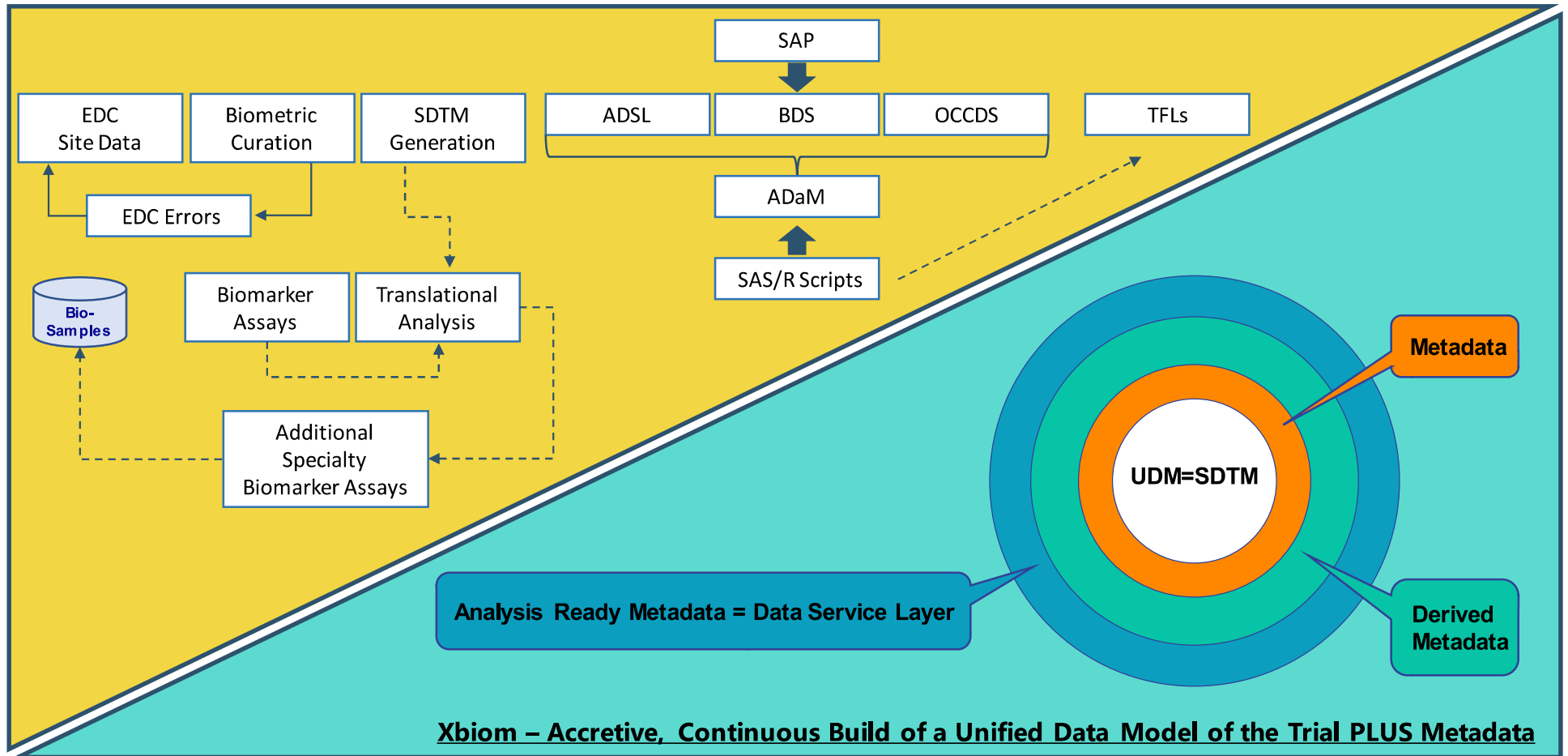
QA

EDC / Labs / CRF	Metadata / CDISC Deliverables			Study Documentation Define.xml	Regulatory Compliance eDV / SDRG / ADRG
	SDTMs	ADAMs	TFLs		
DATA: Raw Codelists	Standard Domains Standard Variables Standard Terminology Codelists	Safety / Efficacy Derived Variables Codelists	SAP	Documentation Control Terminology Value-Level Metadata Raw / Derived Variables	Documentation Data Issues Compliance Issues
METADATA / CDASH SPECIFICATIONS: Attributes, Structure, PRM	SDTM IG Rules Control Term IG Rules MedDRA Export Specifications	ADAM IG Rules Control Term IG Rules (Optional) Export Specifications	ARMs BDS Independent of ADaMs	Define.xml IG Rules SDTMs / ADaMs Snapshot Integrated Links to CRF pages User-Interface Edits	Snapshots / Links
USER INTERFACE MACHINE LEARNING PRODUCTIVITY:	Joins / Transpose Auto / User Mapping Templates Drop-down lists	SAP Mapping Auto / User Mapping SAP Cohorts Drop-down lists	SAP Cohorts Domain Templates Drop-Down Lists	IG Mapping Templates	Template Mapping PhUSE Templates
TRADITIONAL PROGRAMMING PRODUCTIVITY:	Source / QC			Separate Tool Out-of-Sync	Separate Tool Manual Updates
	Attribute Macros Variable Macros	Attribute Macros Variable Macros	Reporting Macros		

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



Unified Data Model (UDM) with Data Service Layer

Cohort Analysis

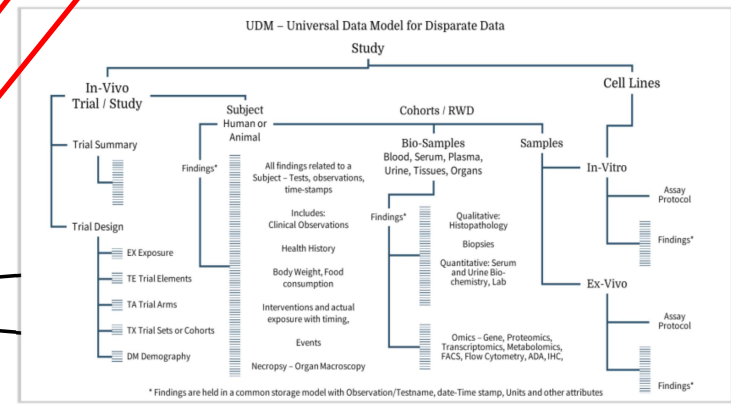
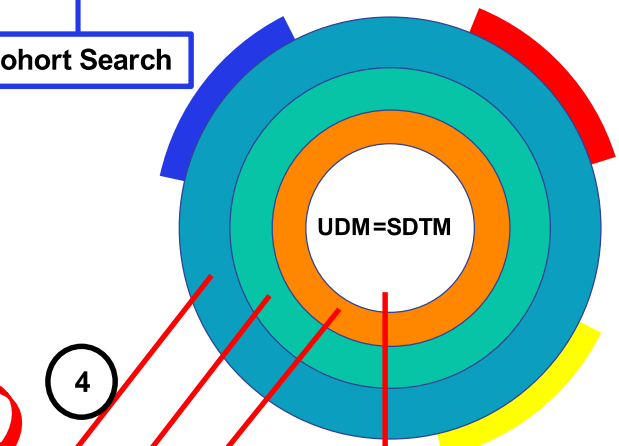
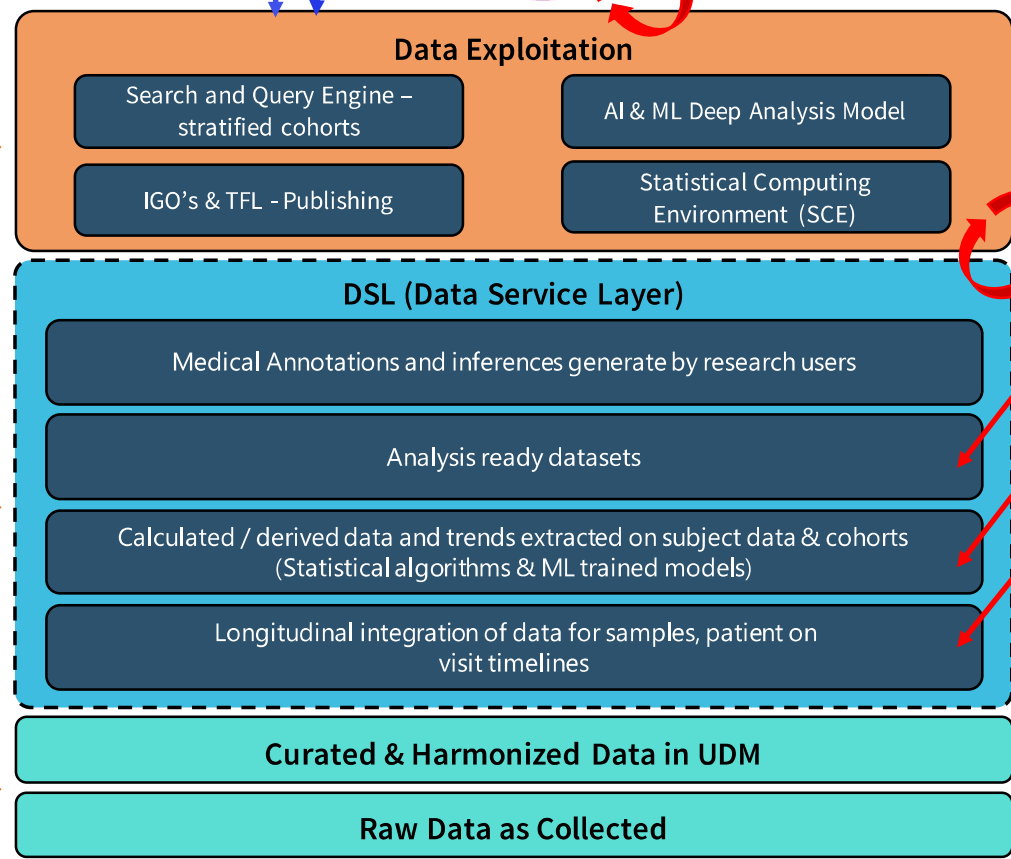
Cohort Search

INGISHTS DECISIONS

KNOWLEDGE

INFORMATION & DERIVED METADATA

DATA



Xbiom: Low-Code User Interface and Visualization Tool

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

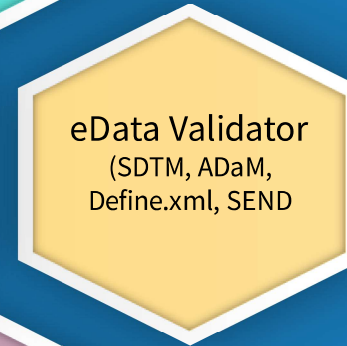
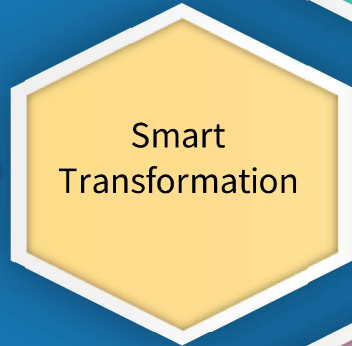


CAPABILITY

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

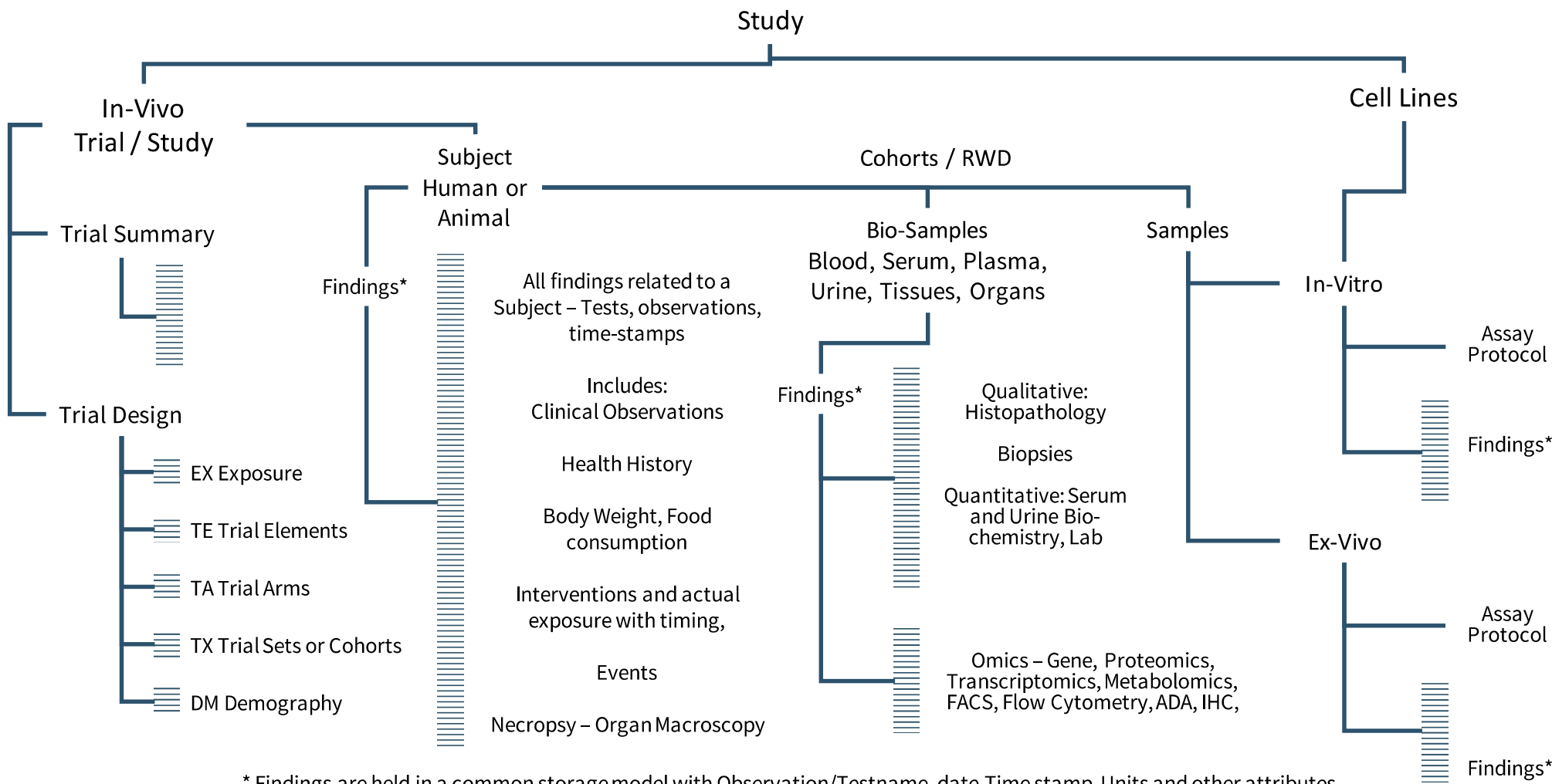
Search, define, and analyze stratified cohorts

Longitudinally integrated subject and patient data

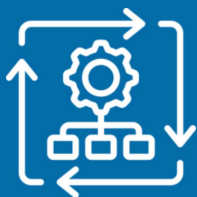
Rapidly generate TFLs for publication and collaboration

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

UDM – Unified Data Model for Disparate Data

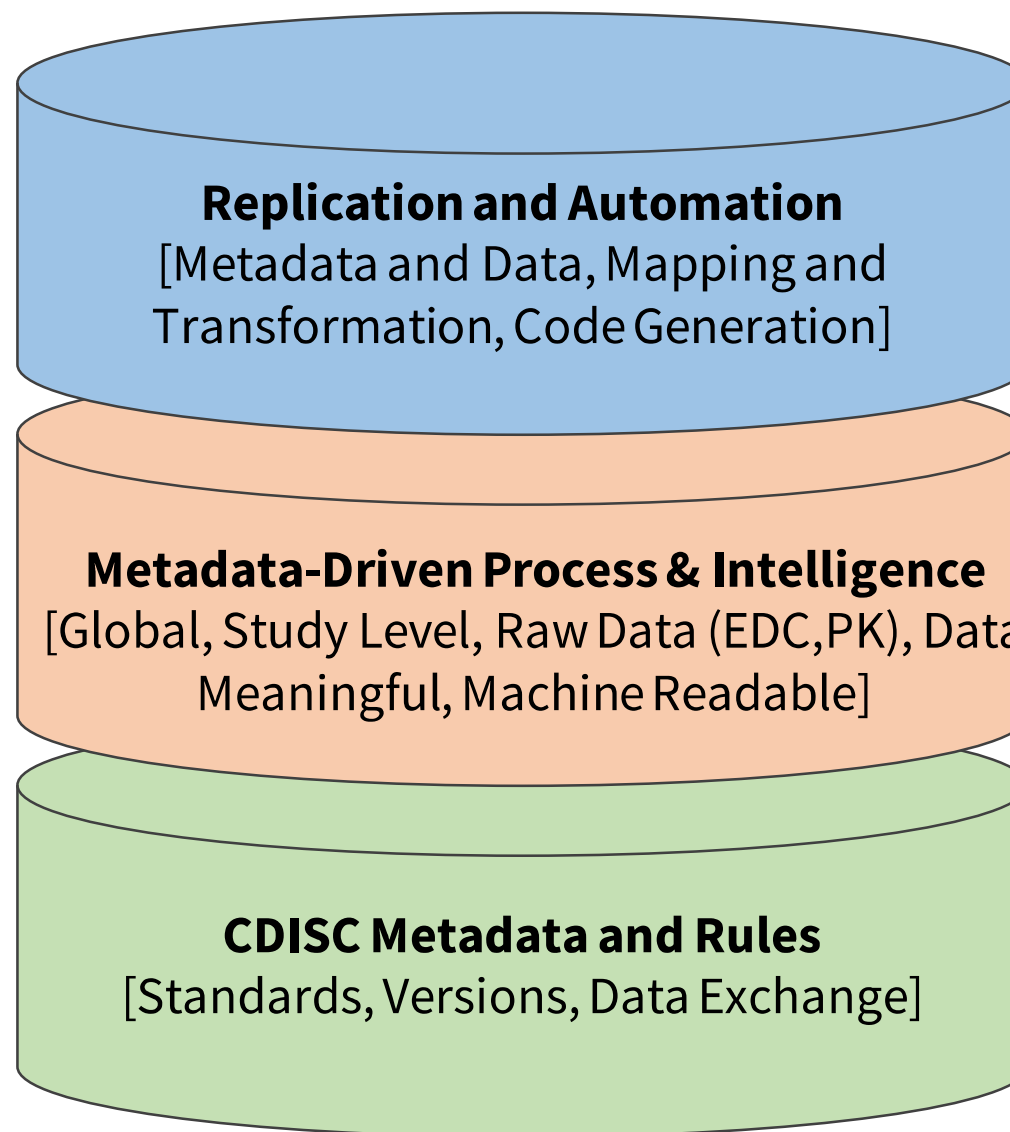


* Findings are held in a common storage model with Observation/Testname, date-Time stamp, Units and other attributes



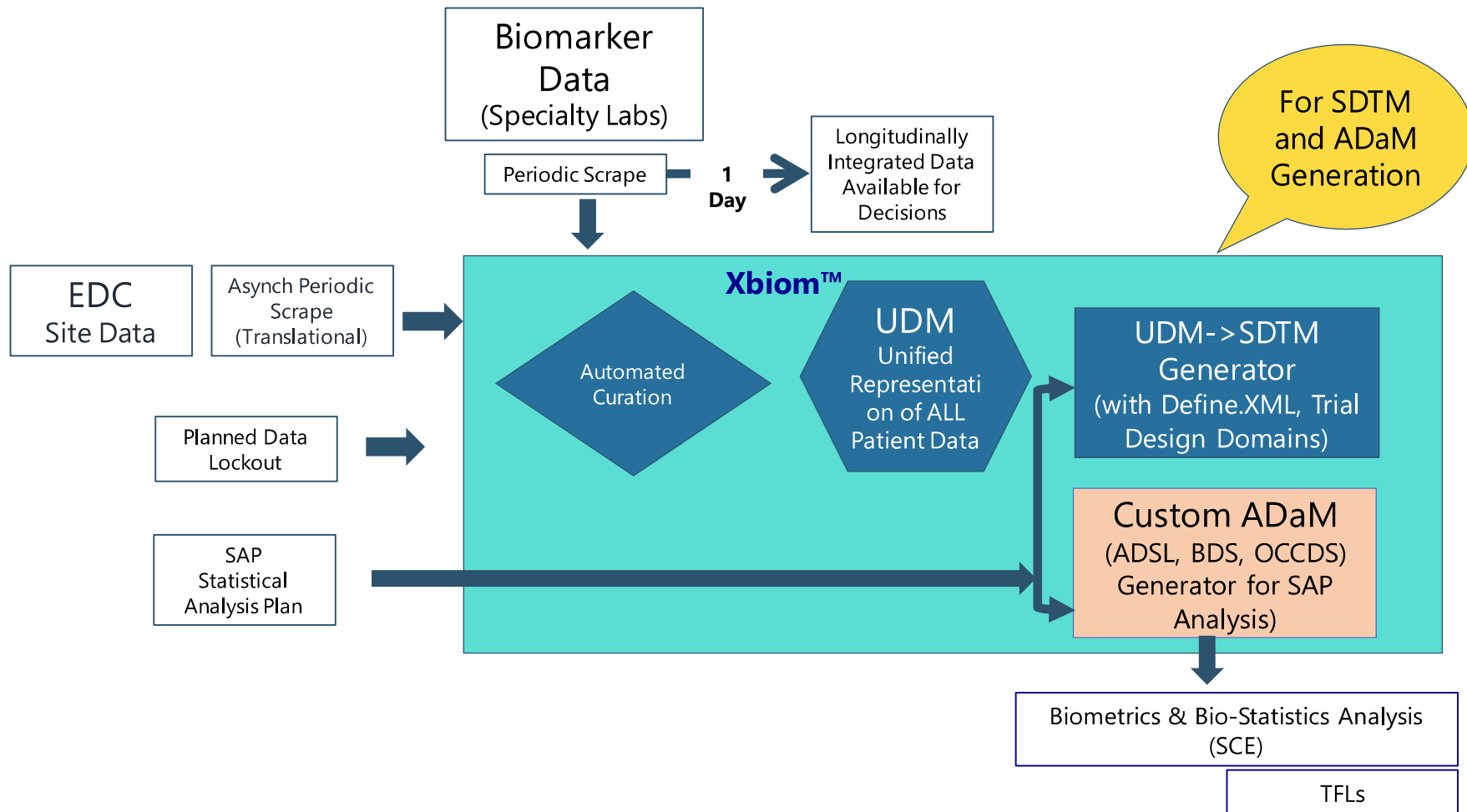
SDTM Automation By Reusability and Repeatability

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

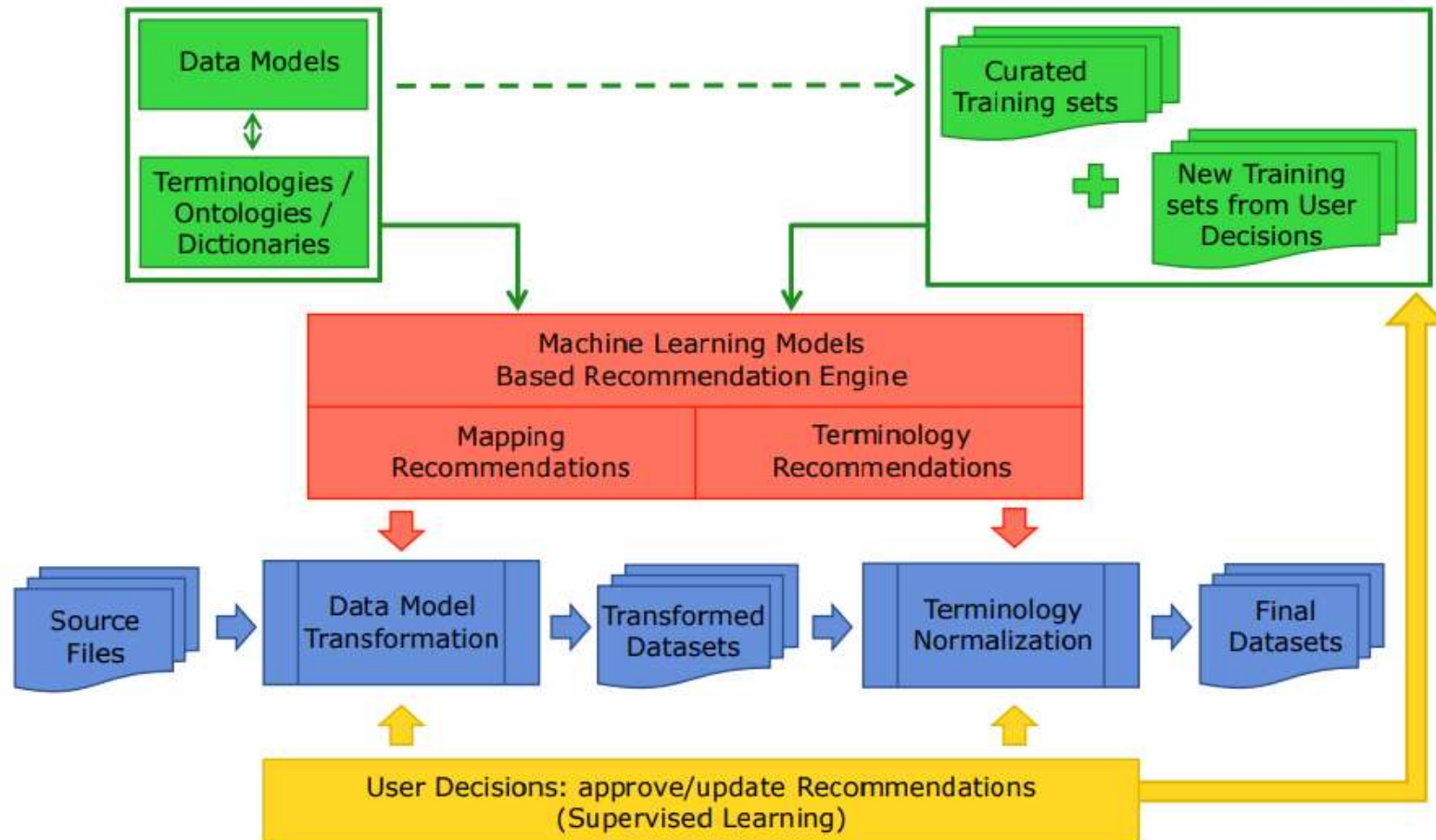


EVOLUTION

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



Machine Learning and Data Transformation Work Flow



SDTM Curation Enables 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

Xbiom Platform Screenshots

Direct to Review Cohorts and Biosamples

The screenshot displays the Xbiom Platform interface for cohort and biosample selection. The top navigation bar includes 'Cohort and Biosample Selection', 'Support', 'I Wish...', and 'Production'. The user is identified as 'Gupta, Sunil' on '29-SEP-2022 10:19:26 (UTC -07:00)'. The interface shows a search bar with 'B1 AND C1' and a breadcrumb trail: 'Study Summary > Subject History-Baseline > Clinical Parameters > Molecular Biomarkers > Biosamples > UDM Variables >'. A sidebar on the left lists various data domains, with 'Vital Signs' highlighted. The main area shows a configuration for a variable named 'Temperature' with a range from 40.02 to 98.92. A yellow callout box on the right states: 'Six Levels to Create Cohorts Variables Useful for grouping or subsetting TLFs'. Another yellow callout box at the bottom center states: '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

Xbiom Platform Screenshots

DataViewer Panel to create Tables for 'End-in-Mind'

DataViewer | *Wish...* | Home

Search Results: 2 of 26 Subjects | Cohort

Study Protocol Summary

STUDYID: SDTM01

Baseline Characteristics

	Count (%)	Mean (SD)	Median (Min-Max)
Number of Participants			
Age			
Sex			
Male			
Female	15 (58%)		
Race			
Asian	1 (4%)		
Black or African American	4 (15%)		
MULTIPLE	1 (4%)		
White	20 (77%)		
ECOG Status			
{not specified}	26 (100%)		
Parameters			
Hemoglobin (Hgb)(mmol/L)		7.19 (0.85)	7.10 (5.50 - 8.80)
Disease Condition			
Treatment History			
Prior Anticancer Therapy	0 (0%)		

Disposition Summary

Trial Arm	Dead	Lost To Follow-Up	Progressive Disease	Withdrawal by Subject
A1:Fixed Dose	7	1	10	
A2:Step Dose-1 Priming Dose			2	2
ΔΔ-NA				1

Study Documents

- Saved TFLs
- Study Documents

Study Information

- Demographics
- Study Drug
- Summary Table

Time Course

- Adverse Events
- Immunogenicity (ADA)

SafetyPharm

- Vital Signs
- ECG Test Results
- ECG Results - Categorical

Laboratory

- Lab Test Results - Quantitative
- Lab Test Results - Categorical

End Point Measurements

- Disease Response
- Tumor Measurements

Medications & Diagnostic Procedures

- Concomitant Medications
- Procedures

Sponsor can Export ADaM Specifications and TLF R programs for Traceability

Scientists and Statisticians can select from a variety of pre-defined Table Templates across all Domains to automatically create TLFs

Drill Down from Summary to Details

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

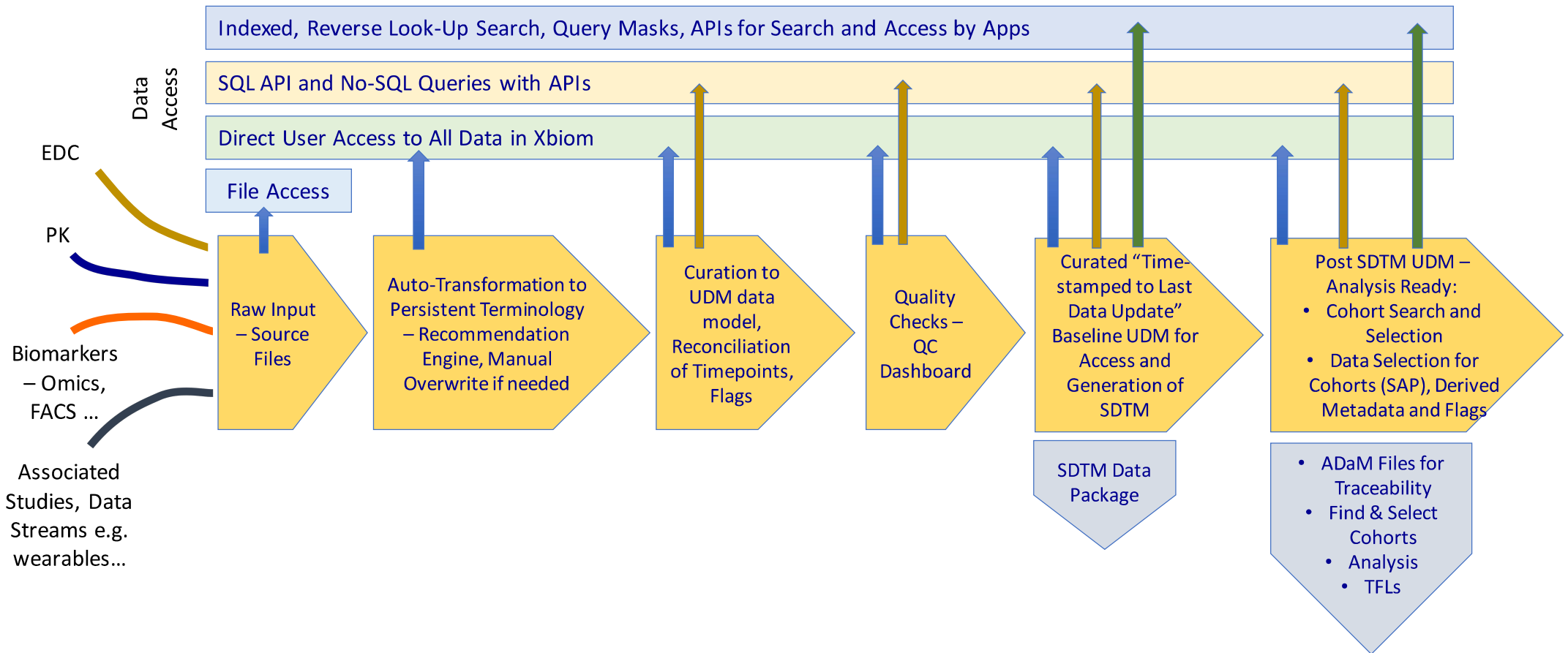
1. Machine Recommended* 
2. Previous Decision** 
3. Preview Raw data and SDTM standard values
4. SUPPXX, RELREC, FA

E. New Variable Derivations

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

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

Data Curation and Standardization Stages to SDTM and Beyond



Clinical Study Workflow

Study ID: PC202301

Initiate **Data Imp**

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

Model :: Tabulation

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

Xbiom Platform Screenshots

Upload or Sync Raw data files into Xbiom - EDC and other sources

The screenshot displays the Xbiom Platform interface. On the left, a file explorer shows a tree structure under 'Study Data Package'. The 'Raw Data' folder is expanded, and the 'EDC' folder is selected. On the right, a table lists the files within the 'EDC' folder, including file names, sizes, and versions.

File Name	File Size	Version
AE.csv	231.62 KB	1.00
AE_YN.csv	231.62 KB	1.00
BMBLL1.csv	231.62 KB	1.00
BT.csv	231.62 KB	1.00
BT_YN.csv	231.62 KB	1.00
CM.csv	231.62 KB	1.00
CM2.csv	231.62 KB	1.00
CM2_YN.csv	231.62 KB	1.00
CM3.csv	231.62 KB	1.00
CM3_YN.csv	231.62 KB	1.00
CM_YN.csv	231.62 KB	1.00
CRS_IRR.csv	231.62 KB	1.00
DM.csv	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

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

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

Scripts for any Transformations: Three 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: 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: 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.

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

Additional Scripts

Data Updates: To perform custom data updates.

Output

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

script SV_Derivation

library t12

SQL Query 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
```

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

Script

Preview of script output

The screenshot displays the 'Metadata Mappings' section of the Xbiom Platform. The main table lists source variables and their corresponding target variables. The 'Recommend...' column contains icons for 'Previous Decision', 'Approved', and 'Machine Recommended'. Below the main table are two smaller tables labeled 'Source' and 'Target' showing detailed mapping data.

#	Source	Mapping	Target ↓	Recommend...
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	⏪ ⏩ ⏴ ⏵

Source					Target				
AESEV	AESER	AEREL	AETERM_PT_CODE	AEOUT	AESEV	AESER	AEREL	AEPTCD	AEOUT
Grade 1 Mild	No	Related	10013911	Not recovered	Grade 1 Mild	No	Related	10013911	Not recovered/not resolved
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 recoverer	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 302 103 1 Target Domain: AE

#	Source	Mapping	Target ↓	Recommend...	N...
44	AESEV	→	AESEV	☑	☒
48	AESER	→	AESER	☑	☒
62	AEREL	→	AEREL	☑	☒
82	AETERM_PT_CODE	→		☑	☒
42	AEOUT	→		☑	☒
80	AETERM_LLTT_CODE	→		☑	☒
79	AETERM_LLTT	→		☑	☒
78	AETERM_HLT_CODE	→		☑	☒
77	AETERM_HLT	→		☑	☒
76	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	Related	10013911	Not recovered/not resolved	
Grade 2 Moderate	No	Related	10001551	Recovered/resolved	Grade 2	Related	10001551	Recovered/resolved	
Grade 1 Mild	No	Related	10001551	Recovered/resolved	Grade 1 Mild	No	10001551	Recovered/resolved	
Grade 2 Moderate	No	Not Related	10028372	Not recovered	Grade 2 Moderate	No	10028372	Not recovered/not resolved	

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

	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

	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

Apply

Cancel

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
OR LATINO	NOT HISPANIC OR LATINO				Y

Target	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			
	NOT HISPANIC OR LATINO				

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

	Source	Target	Expression	Recommendation
+ AE				
+ CM				
- DM				
+ 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

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 Table ...

PointCross's Xbiom Solution for SDTM Automation & Compliance

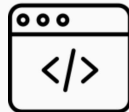
Sunil Gupta
CDISC SME, Trainer & Author
Email: GuptaProgramming@gmail.com

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

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