CDISC 360 Mission: SDTM Design and Automation
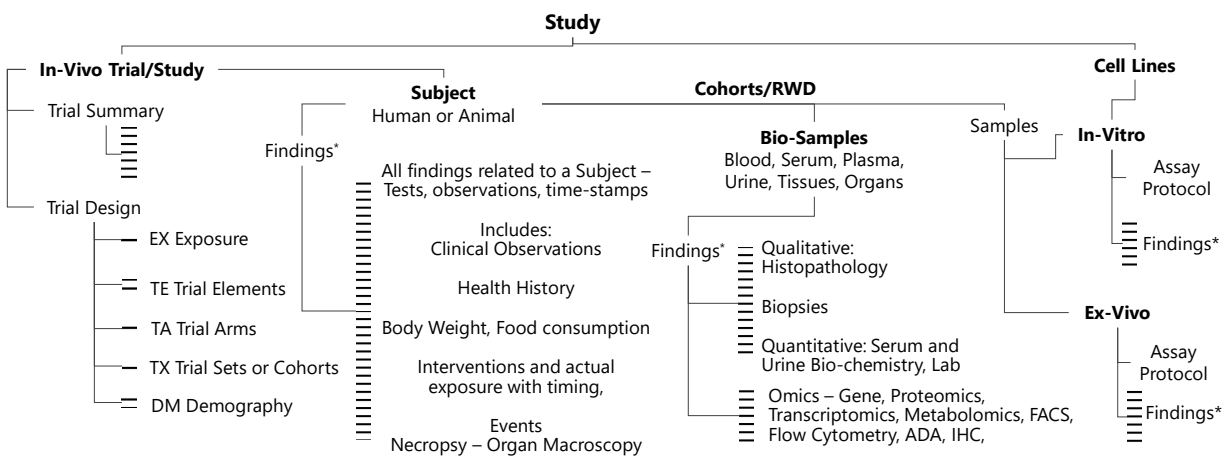
POINT CROSS Life Sciences

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
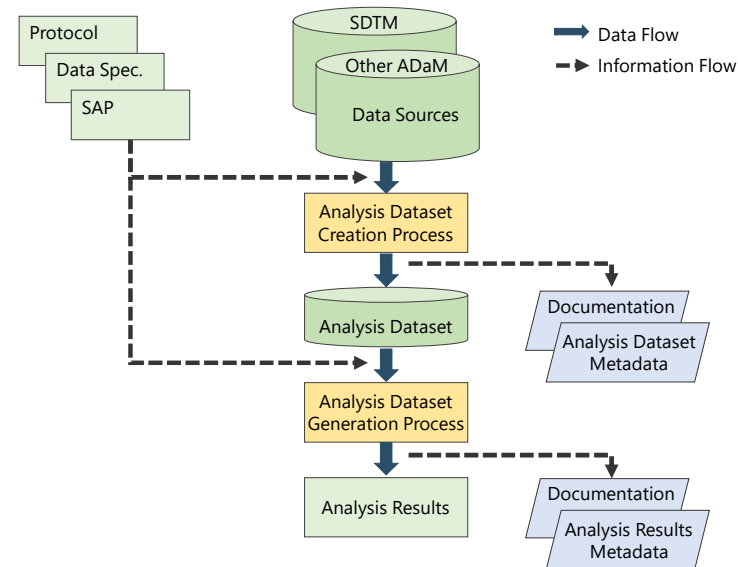
Study

In-Vivo Trial/Study
- Trial Summary
- Trial Design
  - EX Exposure
  - TE Trial Elements
  - TA Trial Arms
  - TX Trial Sets or Cohorts
  - DM Demography

Findings*

Subject
Human or Animal

All findings related to a Subject – Tests, observations, time-stamps

Includes:
Clinical Observations

Health History

Body Weight, Food consumption

Interventions and actual exposure with timing,

Events
Necropsy – Organ Macroscopy

Cohorts/RWD

Bio-Samples
Blood, Serum, Plasma, Urine, Tissues, Organs

Findings*
- Qualitative: Histopathology
- Biopsies
- Quantitative: Serum and Urine Bio-chemistry, Lab
- Omics – Gene, Proteomics, Transcriptomics, Metabolomics, FACS, Flow Cytometry, ADA, IHC,

Samples

Cell Lines

In-Vitro
- Assay Protocol
- Findings*

Ex-Vivo
- Assay Protocol
- Findings*

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

Output: SDTMs, ADaMs, Define.xml & TFLs

Data Curation
❑ Repetitive Process to Optimize Data and Metadata to ensure Valuable use of Data

Protocol → Data Spec. → SAP

SDTM / Other ADaM / Data Sources

→ Data Flow
--► Information Flow

Analysis Dataset Creation Process → Analysis Dataset → Documentation / Analysis Dataset Metadata

Analysis Dataset Generation Process → Analysis Results → Documentation / Analysis Results Metadata

Transformation and Automation: Reusability & Repeatability

Input ⬇   **Output** ⬆

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
[Metadata and Data, Mapping and Transformation, Code Generation]

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 & Intelligence
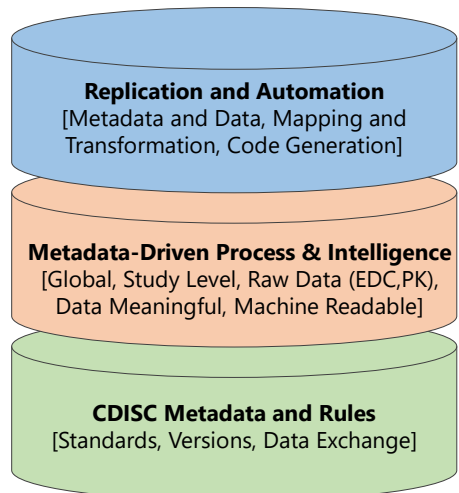[Global, Study Level, Raw Data (EDC,PK), Data Meaningful, Machine Readable]

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

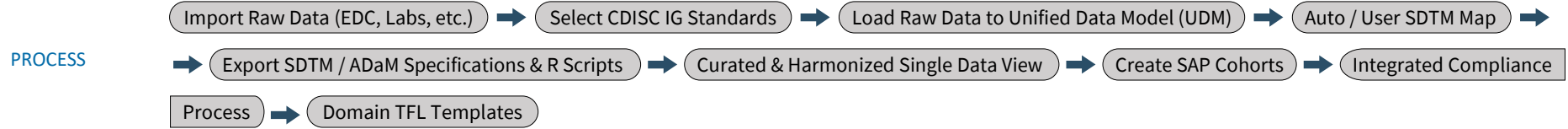**CDISC Metadata and Rules**
[Standards, Versions, Data Exchange]

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

Evolving CDISC to the Next Decades: The CDISC Proof of Concept, Peter Van Reusel, Sam Hume; Efficiency Comes From Reusability and Repeatability, Hanming Tu, Dave Evans

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

|	
CDISC – 360 Metadata Artifacts:	Setup Configuration [Study: Protocol (**PRM**), Design, Workflow, Visits, Parameters], Standards Selection (**CDISC IG**), Data Collection [Raw Data, CRF, **CDASH**], **Control Terminology**, **MedDRA**, Tabulation (**SDTM IG**), Analysis (**ADaM IG**, BDS), **ARM**, **Define.xml**
PROCESS	Import Raw Data (EDC, Labs, etc.) → Select CDISC IG Standards → Load Raw Data to Unified Data Model (UDM) → Auto / User SDTM Map → Export SDTM / ADaM Specifications & R Scripts → Curated & Harmonized Single Data View → Create SAP Cohorts → Integrated Compliance Process → Domain TFL Templates
GOALS	Project Management, Documentation, Standards, Automation (Low Coding), Quality, Productivity (Fewer Resources), Budget (Less Expensive)
REASONS	Project Complexity, CDISC Compliance Requirement, Traceability / Audit Trail, Enforce Hundreds of Rules
INTELLIGENCE DESIGN	Machine Readable / Learning (Decision Tree Process / Pattern Recognition / Consistency)
CROSS-REFERENCE	Directories, Structure / Transpose, Study / Project, Datasets (BDS), Variables, Order, Required, Value-Level, Control Terminology, CDISC IG Rules, Results, Links, Name, Type (Numeric, Character), Length, Label, Format, Units, Descriptive Stats, Codelists (Unique, Missing, Range, Keys, Max Length)

EVOLUTION: SAS Programming → CDISC Standards → GUI Metadata Mapping
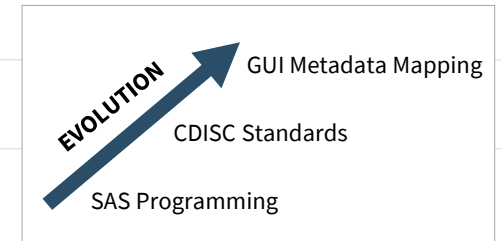
EDC / Labs / CRF	Metadata / CDISC Deliverables			Study Documentation Define.xml	Regulatory Compliance eDV / SDRG / ADRG
	SDTMs	ADAMs	TFLs		
DATA: Raw Codelists	Standard Domains Standard Variables Standard Terminology Codelists	Safety / Efficacy Derived Variables Codelists	SAP	Documentation Control Terminology 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				
	Attribute Macros Variable Macros	Attribute Macros Variable Macros	Reporting Macros	Separate Tool Out-of-Sync	Separate Tool Manual Updates