Sycamore Informatics

SPA MDR CDR SCE

Engineering the Digital Thread:
Provenance and Impact Analysis in Metadata Management

Mukul Goyal
Associate Director – Development, Sycamore Informatics

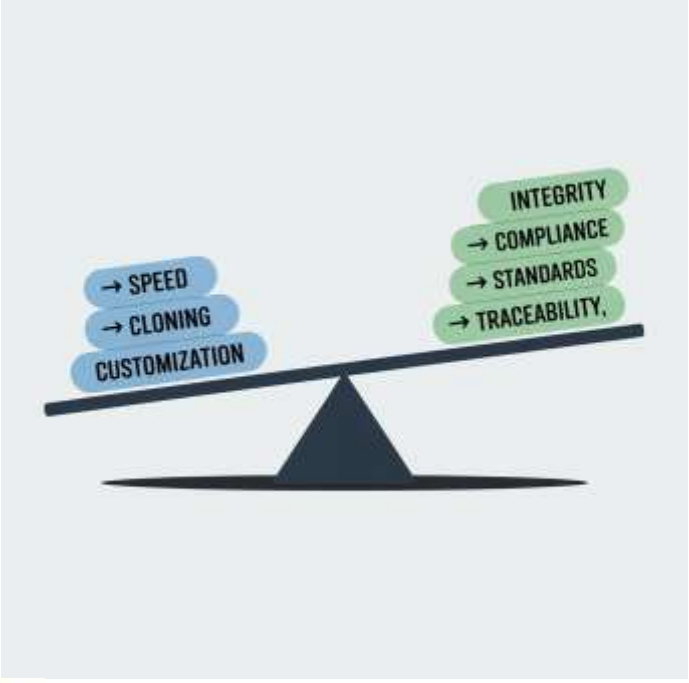
Pharma SUG SDE
APRIL 11, 2026

Agenda

- **The Metadata Paradox:** Balancing Innovation with Governance
- **Defining the Digital Thread:** Beyond Simple Cloning
- **Core Pillar I:** Provenance to CDISC Foundation Standards
- **Core Pillar II:** Provenance to Company & TA Libraries
- **Solving the Lineage Problem**: The Metadata Archetype
- **Solving the Mapping Problem:** Automated Impact Analysis & Relationships.
- **Enforced Integrity:** Reducing the Manual Validation Burden
- Summary & Q&A

PharmaSUG SDE APRIL 11, 2026

The Metadata Paradox: Innovation with Integrity



- Study teams need to build specifications rapidly.
- Governance mandates require strict adherence to standards.
- **The Lineage Problem:** Traditional "copy-paste" cloning breaks the link to the original source.
- **The Mapping Problem:** Assessing how a change in a single variable ripples across CRFs, SDTM mappings, and ADaM specifications is manual, error-prone, and slow.

What is the "Digital Thread"?

- It is an architecturally enforced connection.
- It tracks metadata from its "Archetype" (origin) to its final adoption.
- It ensures that the "Source of Truth" is never lost during customization.

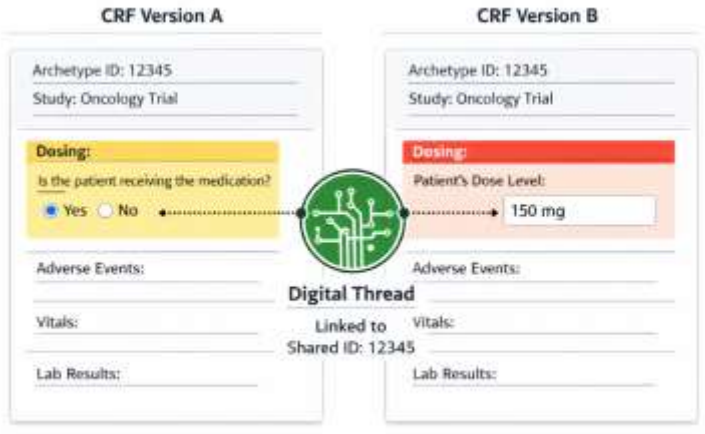
Pillar I: The CDISC Foundation & The Archetype

- **Visibility:** Display the specific IG and Version for Data Models like SDTM, ADaM, and Controlled Terminology.
- **The Archetype:** Defining the immutable origin of metadata.
- **Multi-Layer Persistence:** Relationships remain intact across infinite layers of cloning.
- **1:1 Traceability:** Immediate audit trail from any study variable back to the CDISC source.

Pillar II: Global Libraries & TA-Specific Adoptions

- **Global Library Management:** Reusable “Global Standard" CRFs.
- **TA-Specific Adoptions:** Global Library are reused across multiple Therapeutic Area (TA) Libraries, these CRFs have TA-specific adoptions.
- **Provenance**: When used across studies, provenance is maintained to TA and Global Libraries
- **Variant Identification:** System automatically identifies which version/variant is in use.
- **Automated Lineage:** Maintaining the link between the Study, the TA Library, and the Global Library.

Solving the Lineage Problem: The Metadata Archetype



- **Vertical Lineage in Action**: Every study-level object maintains a persistent, automated link back to its parent in the TA or Global Library.
- **The Archetype ID:** A shared, immutable ID (e.g., 12345) that serves as the permanent "Source of Truth".
- **Eliminating Drift:** Provenance is preserved across infinite layers of cloning and customization.

Solving the Mapping Problem: Automated Impact Analysis

- **The Horizontal Challenge**: Manual assessment of cross-lifecycle variable changes is slow and error-prone.
- **The Governance Engine**: Implements **Horizontal Traceability** via programmatic traversal of metadata relationships.
- **Relationship Mapping**: The system "crawls" the metadata graph to identify Form, Variable, and Mapping dependencies.

Visualizing Horizontal Traceability

- **End-to-End Connectivity**: Automated flow from **Standard Variable** -> **Study CRF** -> **SDTM** -> **ADaM**.
- **Chain of Integrity**: Seamless relationship tracing from initial collection to final analysis.
- **Real-Time Governance**: Instant, data-driven impact reports for Standards Governance boards.
- **System Enforcement**: Replaces manual oversight with architecturally enforced connectivity.

Shifting the Burden: Architecturally Enforced Integrity

Before	After
Manual Oversight	System Enforcement
Excel trackers	Digital Thread
Validation meetings	Automated Compliance

- Study teams focus on **Deliverables** (Specs), not **Maintenance**.
- Reduced manual validation burden.
- Compliance is a byproduct of the workflow, not an afterthought.

Summary: Innovation with Integrity

- **Vertical Lineage**: **Immutable Archetypes** eliminate metadata drift, providing a 1:1 trace from study variables to CDISC/Global sources.
- **Horizontal Traceability**: **Automated Mapping** enables real-time impact analysis from Data Collection (CRF) to Analysis (ADaM).
- **Agile & Compliant:** Reduced validation burden leads to faster study startup.

Contact Information

Name: Mukul Goyal

Title: Associate Director - Development

Organization: Sycamore Informatics

E-mail: mukul.goyal@sycamoreinformatics.com

Sycamore Informatics

SPA MDR CDR SCE

Thank You

PharmaSUG SDE
APRIL 11, 2026