cdisc®

2026 Europe Interchange

THE FUTURE IS CONNECTED: STANDARDS AND AI POWERING DIGITAL TRANSFORMATION

MILAN, ITALY | MAIN CONFERENCE: 20-21 MAY | TRAININGS & WORKSHOPS: 18, 19, & 22 MAY

Schema-First Metadata Engineering

Roger GAGALI ANGOH, Consultant, QUALIANCE

C1 - Internal use

Speaker

Gagali Angoh Roger

Consultant

Roger Gagali Angoh is a clinical metadata engineer focused on turning CDISC standards into structured, enforceable, and interoperable systems. With over ten years of experience, he combines expertise in SDTM, controlled terminology, and metadata architecture with a strong problem-solving mindset. His work centers on designing metadata structures that enable automation, consistency, and reliable data exchange across clinical environments.

LinkedIn

Agenda

1. The Problem - Metadata Without Structure
2. Clinical Metadata Engineering - A Structural Perspective
3. Schema-First - From Structure to Contract
4. Enforcing Predictable Behavior

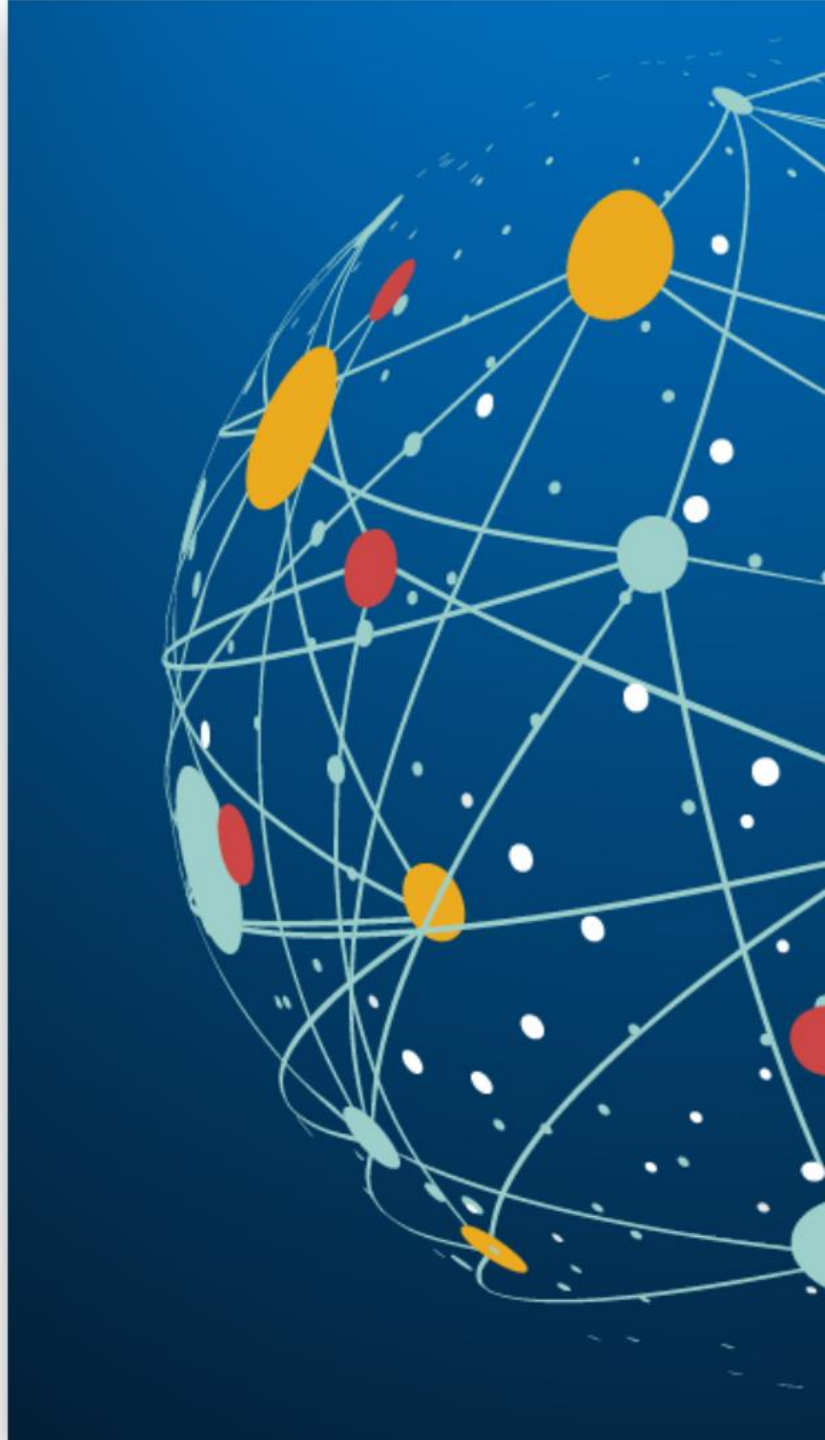
1. The Problem

Metadata Without Structure

We All Work with Metadata

- SDTM, ADaM, Controlled Terminology
- Define.xml, study specifications
- MDRs, templates, Excel files

Daily work is metadata work



The Problem We All See

- Metadata in Excel everywhere
- Multiple versions of the same metadata
- Manual checks and corrections
- Limited reuse and fragile automation across studies
- Business rules held by a limited number of people

"We talk about standards, but not about the structure behind them"

How Inconsistency Appears

- Mixed versions
- Silent decisions becoming defaults
- Misunderstanding of metadata business rules
- Small deviations at study level

Consistency does not disappear suddenly

It fades gradually

Root Cause

- Structure is implicit and left to vendors
- Rules are scattered and interpreted by each company
- Behavior depends on implementation choices

"We don't have a shared database schema for clinical metadata"

2. Clinical Metadata Engineering

A Structural Perspective

Naming the Discipline

"We have been doing this for years but never named it"

"The practice of designing, structuring, and validating metadata so that standards, automation, and clinical systems can work together"

A Structural View of Metadata Layers

LAYER 1	LAYER 2	LAYER 3
DATA	METADATA	METADATA OF METADATA

This is the missing structural layer

cdisc®

Where Engineering Happens

Clinical metadata engineering operates at **Layer 3**

- Defines structure
- Defines relationships
- Defines structural rules

Structure determines behavior

3. Schema-First

From Structure to Contract

The Method - Schema-First

- Design the implementation structure
- Define tables, relationships, constraints
- Before creating metadata content (Layer 2)
- Before tools, documents, or processes

"Define the database schema structure"

Implementation structure first

Quality by Design Starts with Shared Structure

In pharmaceutical industry, quality is built into processes

- Shared tables and relationships
- Shared implementation structure
- Shared validation constraints across systems

Quality is enforced by design, not reconstructed afterwards

cdisc®

The Limitation

A shared database schema alone is not enough

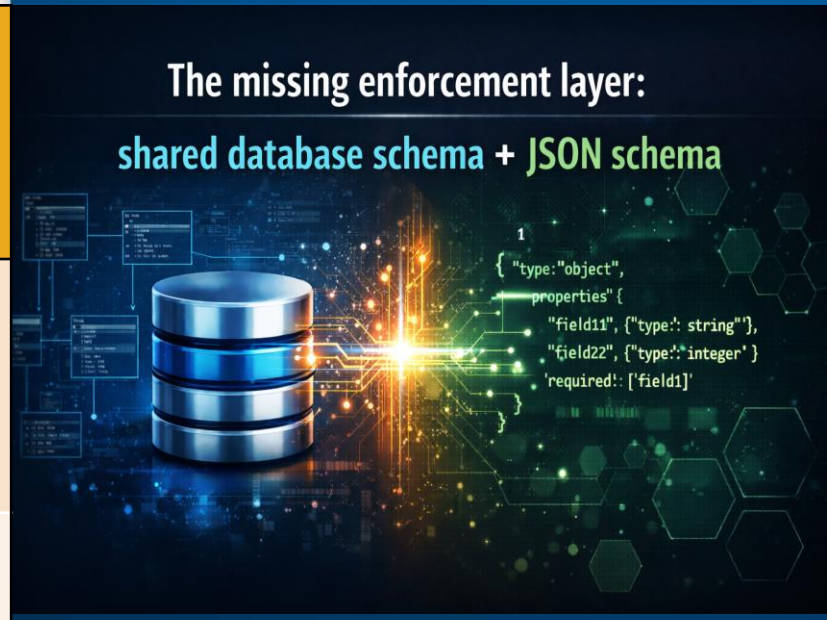
- Some rules are not structural
- Some inconsistencies still pass
- Drift still comes from outside the system

Sources of drift: sponsor documents, local practices, SME implicit knowledge

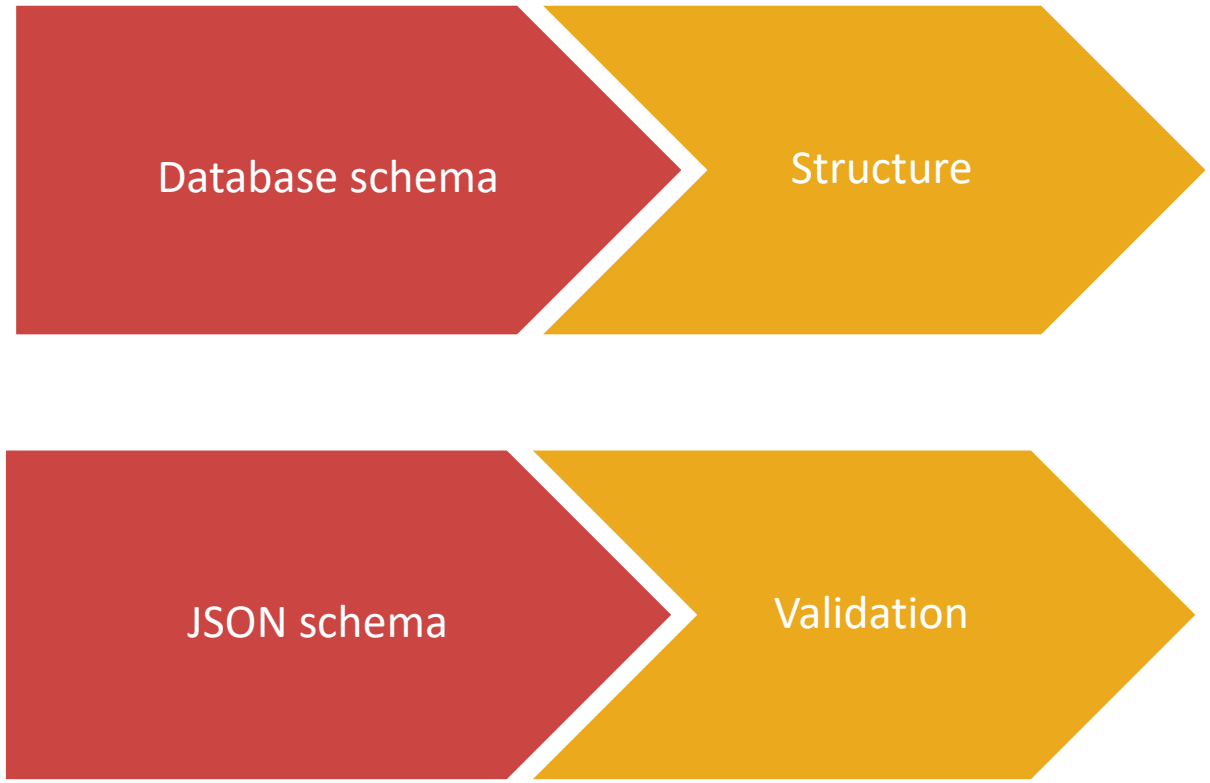
4. Enforcing Predictable Behavior

Two Schemas, Two Roles

Shared Database Schema	JSON Schema
Defines what exists	Defines what is valid
Structural contract	Versioned validation rules
Shared implementation structure	Removes implicit interpretation



The Enforcement Model



Together:

- *rules become explicit*
- *inconsistencies are rejected early*
- *behavior becomes predictable*

Predictable Behavior

Standards provide:

- Predictable meaning
- Consistent structure definition

But not:

- Predictable behavior

"Predictable behavior is a requirement for real interoperability"

Example of Shared Implementation Structure

- **Core versioning Layer**: standards, standard_versions (standard_id)
- **Controlled Terminology Layer**: sponsor_ct (cdisc_ct_version), sponsor_ct_codelists, sponsor_ct_codelist_terms
- **SDTM Versioning Layer**: **sdtm_libraries** (standard_version_id, sponsor_ct_id)
- **SDTM Library Tables**: sdtm_library_domains, sdtm_library_variables, sdtm_library_vlm, sdtm_library_where_clauses, sdtm_library_comments, sdtm_library_methods -> sdtm_library_id

The schema makes the dependencies explicit and enforceable

cdisc®

The Shift

Clinical metadata engineering introduces a structural shift

- Metadata treated as an engineering system
- Structure defined before usage
- Enforcement pushed into a shared database schema
- Behavior made predictable across systems

Final Message

- Standards define semantics
- The industry lacks a structural implementation layer

Clinical metadata engineering fills that gap

Shared database schema defines structure

JSON schema enforces validity

Predictable behavior

across systems

Questions?

Thank You!