

## Auto-mapping of CDISC standards using Agentic AI

Nithyanandan Thangavelu &  
Rohit Bailur  
Lifio.ai

## Introduction

Standardization of clinical trial data into CDISC-compliant **SDTM** and **ADaM** datasets is essential for regulatory submissions, statistical analysis, and cross-study integration.

Mapping raw electronic Case Report Form (eCRF) data to these standards is complex due to variability in CRF structures, metadata, terminology, and derivation requirements.

Traditional manual and rule-based approaches are often resource-intensive and difficult to scale.

**Agentic AI**, based on multiple autonomous but coordinated agents, provides a structured approach to support consistent, explainable, and auditable CDISC mapping workflows.

## AIM

To implement an **Agentic AI-based framework** that:

- Automates extraction of CDISC specifications and study metadata
- Supports syntactic and semantic mapping to **SDTM and ADaM**
- Aligns mappings across **CDASH, SDTM, and ADaM**
- Continuously improves through human feedback and prior study learnings

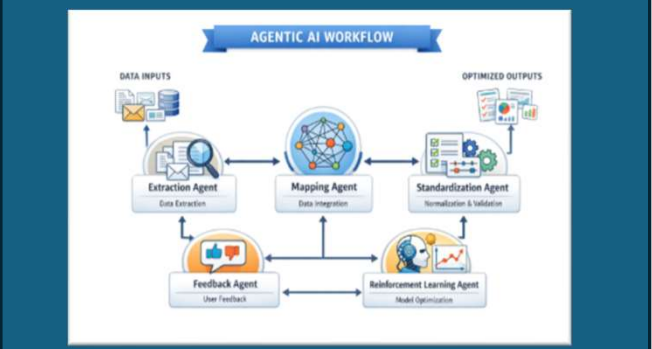
## Results

Initial applications of the framework indicate:

- Reduced manual effort in SDTM and ADaM mapping activities
- Pilot implementations have demonstrated significant gains in mapping speed (up to 60% faster)
- Improved consistency in domain mapping and variable derivations
- Better handling of indirect and derived mappings
- Enhanced traceability of mapping decisions across studies

## Methods

The solution is implemented as a multi-agent system, with each agent responsible for a specific stage of the CDISC mapping lifecycle.



## Agentic AI Framework

- Extracts domains, variables, and specifications from SDTMIG and ADaMIG.
- Extracts CRFs and relevant metadata from study protocols
- Structures source and standards metadata for downstream processing

## 2. Mapping Agent

- Performs syntactic and semantic similarity-based mapping between source variables and CDISC standards
- Supports one-to-one, one-to-many, direct, and indirect mappings
- Identifies cases requiring derived variables when direct mappings are not available
- Triggers mapping algorithms to derive SDTM or ADaM variables as needed

### 3. Standardization Agent

- Aligns mappings consistently across CDASH → SDTM → ADaM
- Ensures harmonized interpretation of variable intent and standard usage

#### 4. Feedback Agent

- Incorporates human reviewer feedback into mapping decisions
- Refines mapping outputs based on corrections and approvals
- Supports controlled human-in-the-loop validation

## 5. Reinforcement Learning Agent

- Learns from previous study mappings and feedback outcomes
- Improves future mapping accuracy and consistency
- Enables continuous system improvement across studies

## Conclusions

- A focused **Agentic AI framework**, built on well-defined and coordinated agents, can effectively support standardized mapping to **CDISC SDTM and ADaM** while maintaining transparency, traceability, and human oversight.
- By integrating structured metadata extraction, syntactic and semantic mapping, standards alignment, and continuous learning from prior studies, the framework complements existing clinical data workflows. This transformative approach supports scalable and inspection-ready CDISC data standardization and has the potential to significantly improve efficiency and consistency in clinical data preparation for regulatory submission.



Connect with us on LinkedIn

