cdisc®

2026 Europe Interchange

THE FUTURE IS CONNECTED: STANDARDS AND AI POWERING DIGITAL TRANSFORMATION

LLM-powered end-to-end clinical data automation: standardized eCRFs to SDTM, ADaM, and TLF with full traceability

Anja Jazbec, Standards Developer, Novo Nordisk
Pankaj Attri, Pre-Sales Solutions Architect, Global Health and Life Sciences Advisory, SAS
Anette Jakobsen, Principal Industry Advisor, Global Health and Life Sciences Advisory, SAS

Speakers

Anja Jazbec
Standards Developer
Novo Nordisk

Anette Dalsgaard Jakobsen
Life Science Senior Industry Advisor
SAS

Pankaj Attri
Solutions Architect
SAS

Agenda

1. Introduction

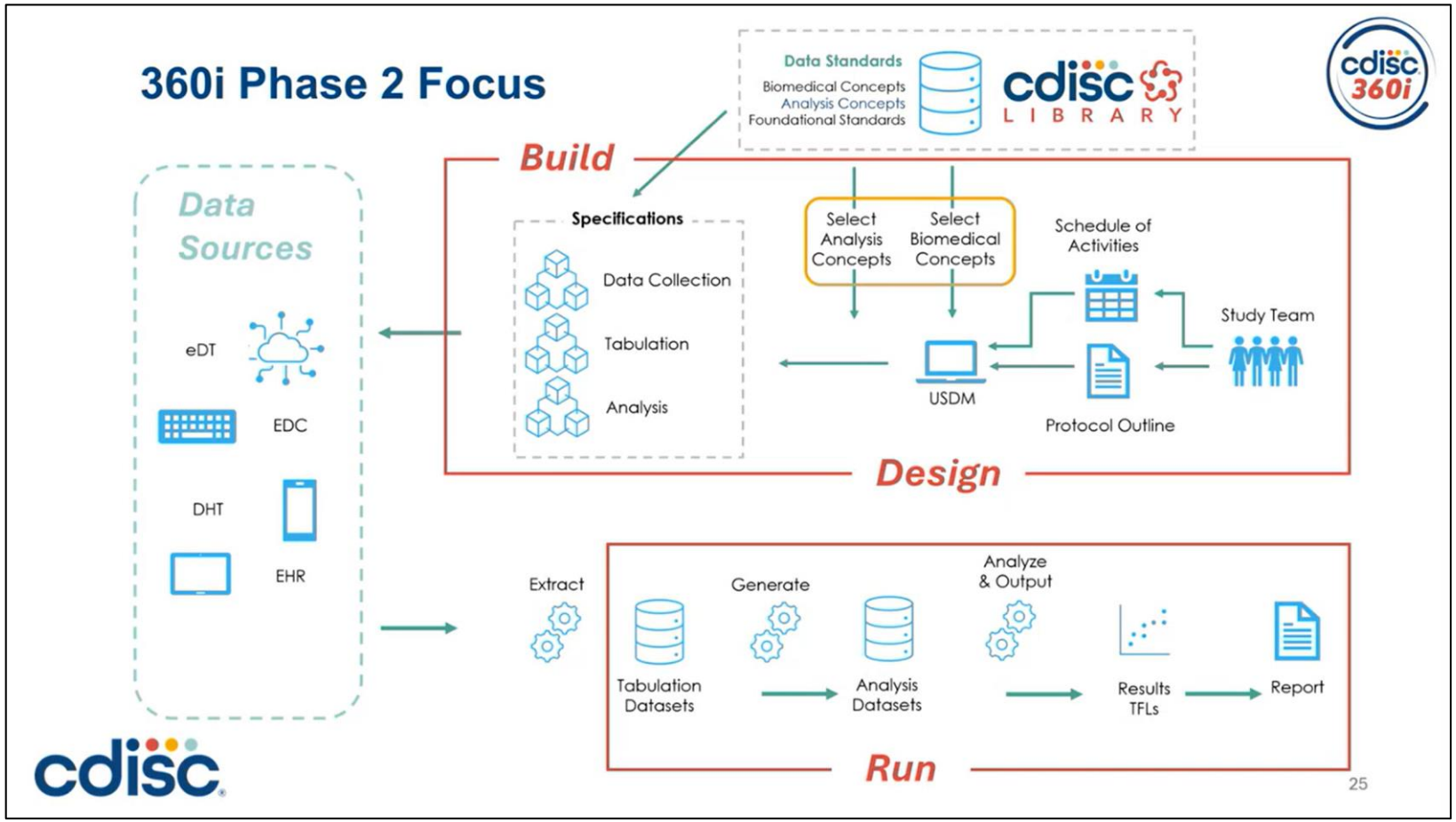
Automation in Clinical Trial workflows

2. Solutions

Generative AI in Clinical Trial workflows

3. Business Impact

Operationalizing 360i vision



Accelerated, standards-driven eCRF design with prequalified metadata and controlled terminology, ensuring data are “SDTM-ready” by design for the RAW→SDTM pipeline.

Automated transformation from RAW to SDTM, ADaM, and TLFs with end-to-end lineage for audit-ready traceability.

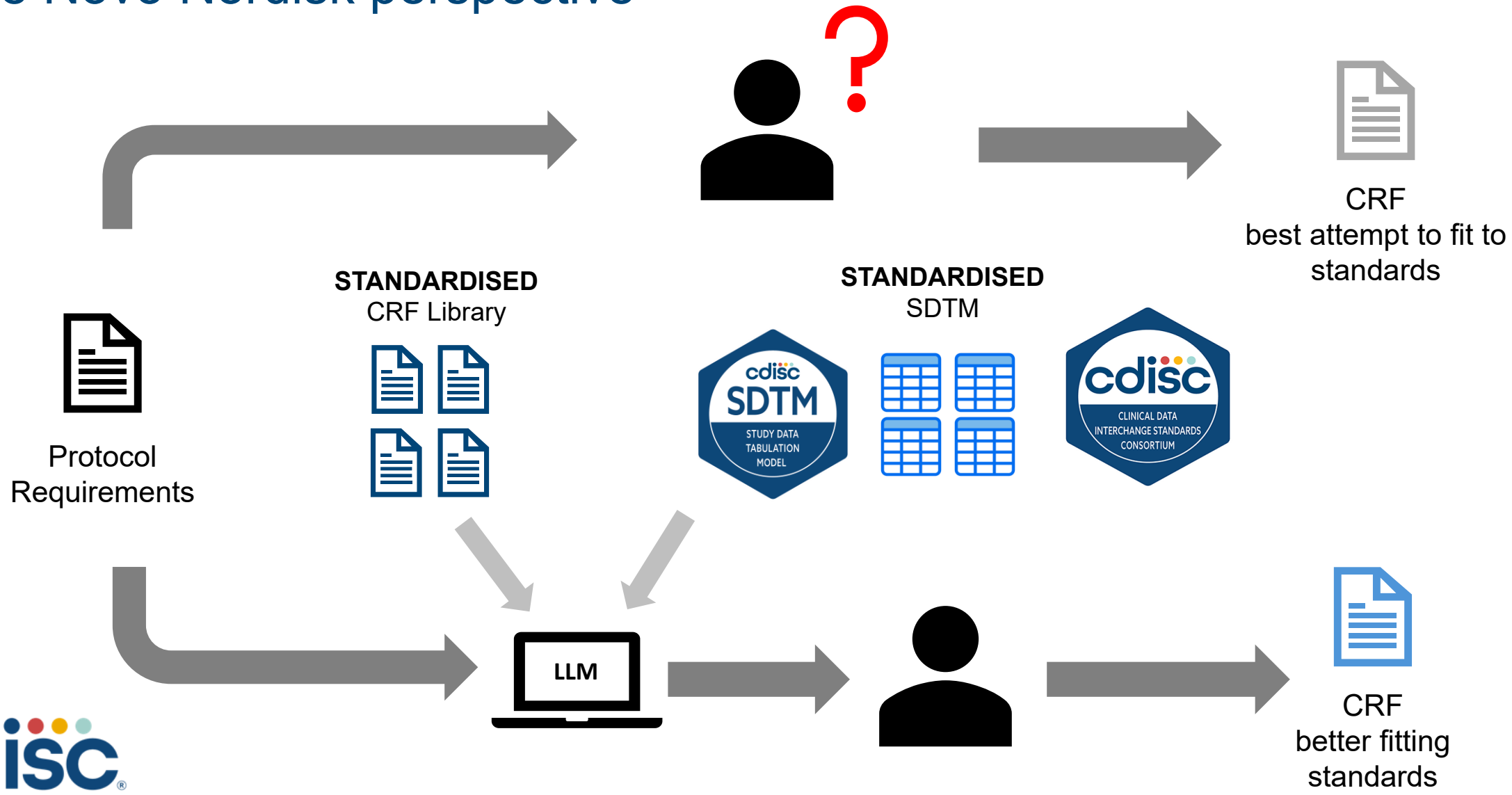
cdisc

LLM Powered eCRF Builder

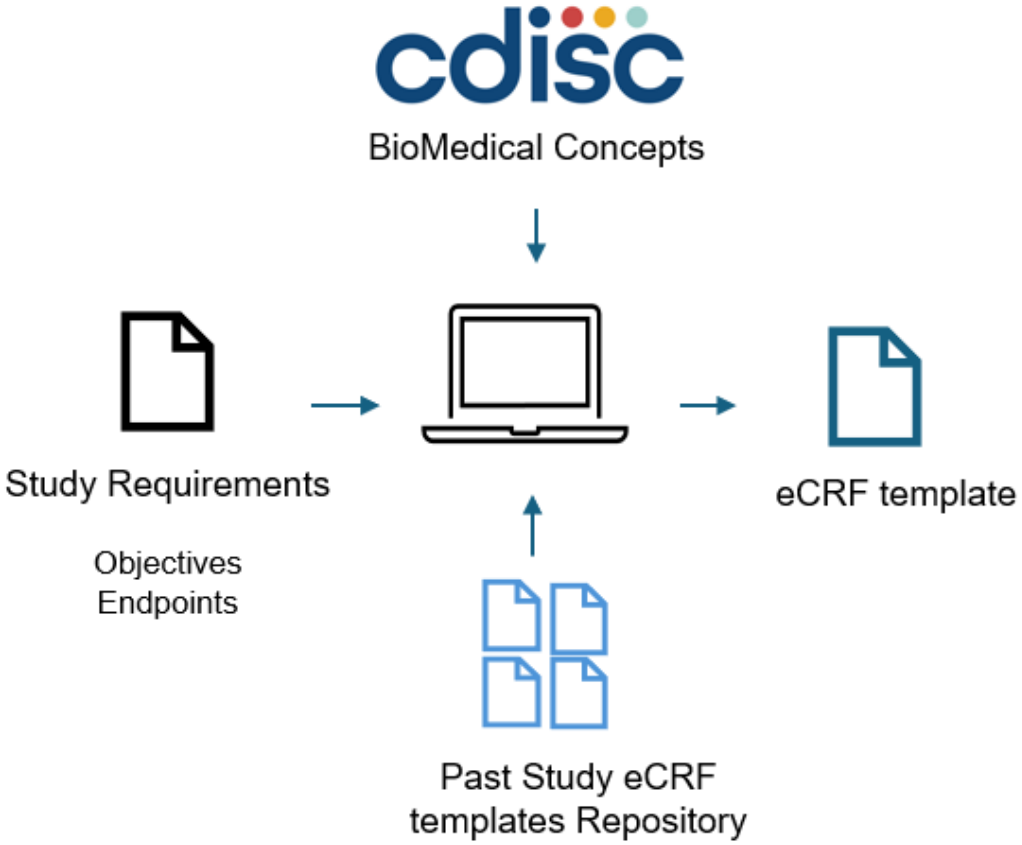
Using eCRF Templates and CDISC BioMedical Concepts Repositories

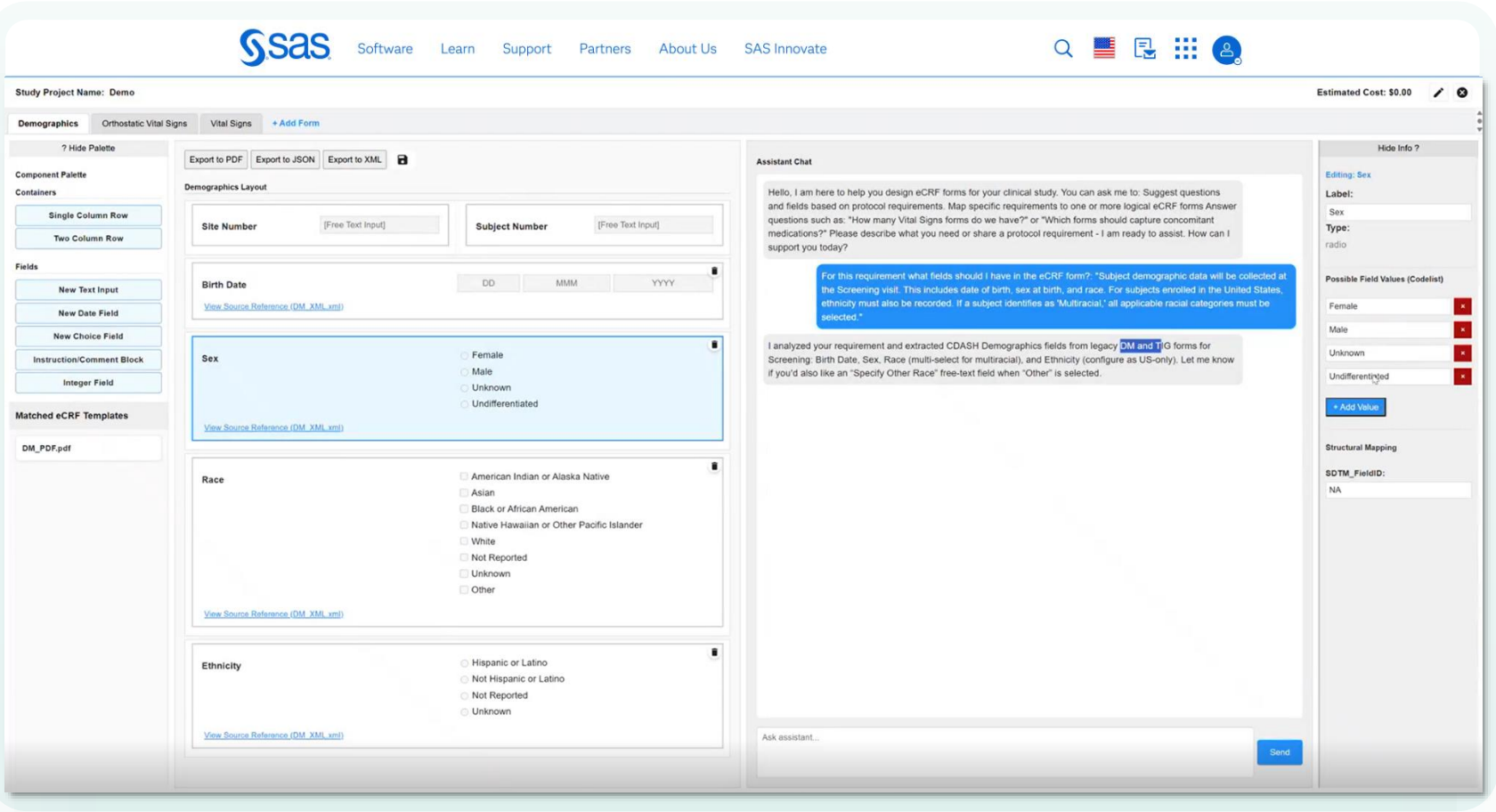
Challenges of standardisation

The Novo Nordisk perspective



AI Assisted eCRF form Creation





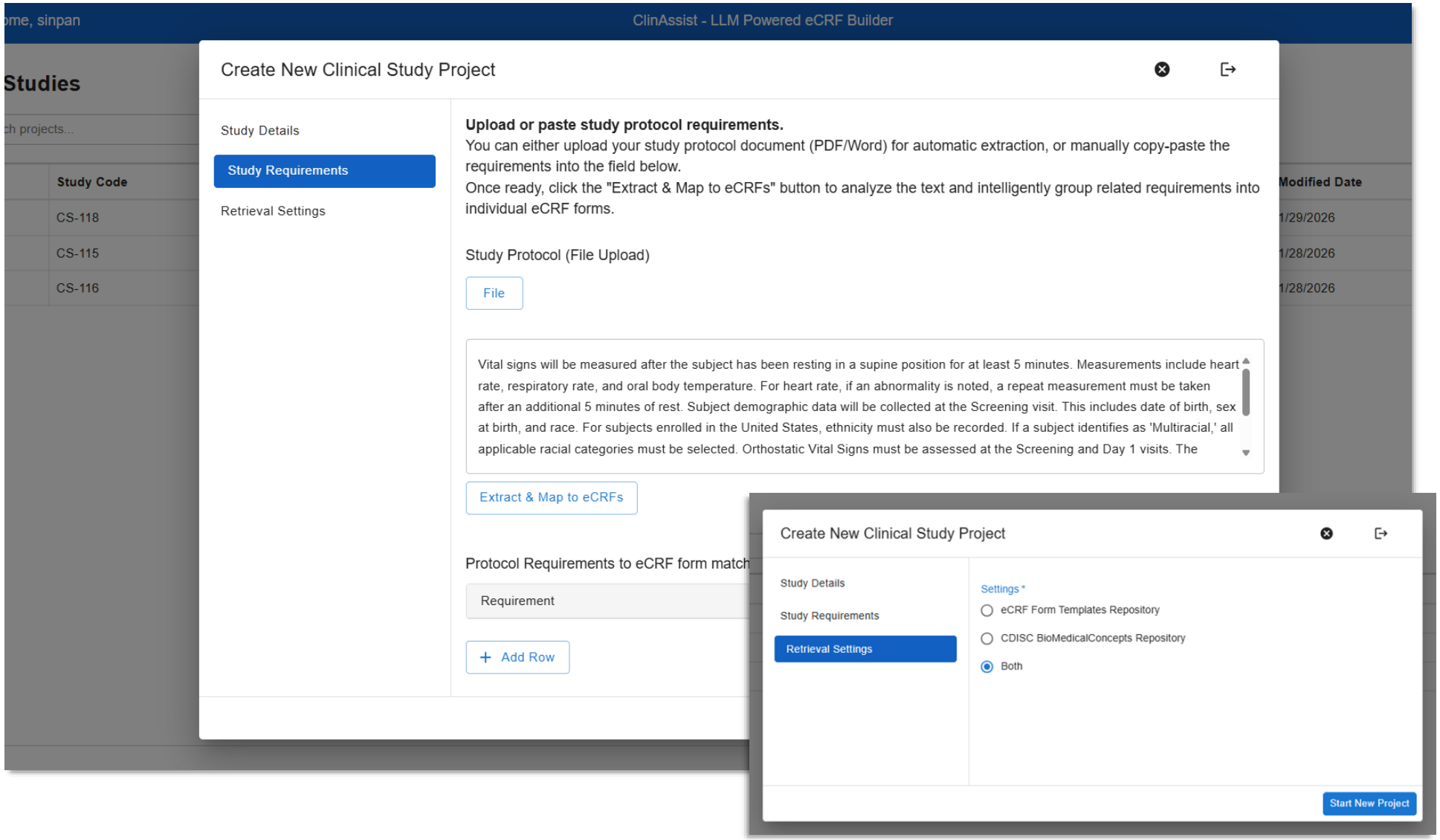
AI Assisted eCRF Creation - Prototype

Clinical Trials eCRF Build Automation

1

Map Requirements to eCRFs

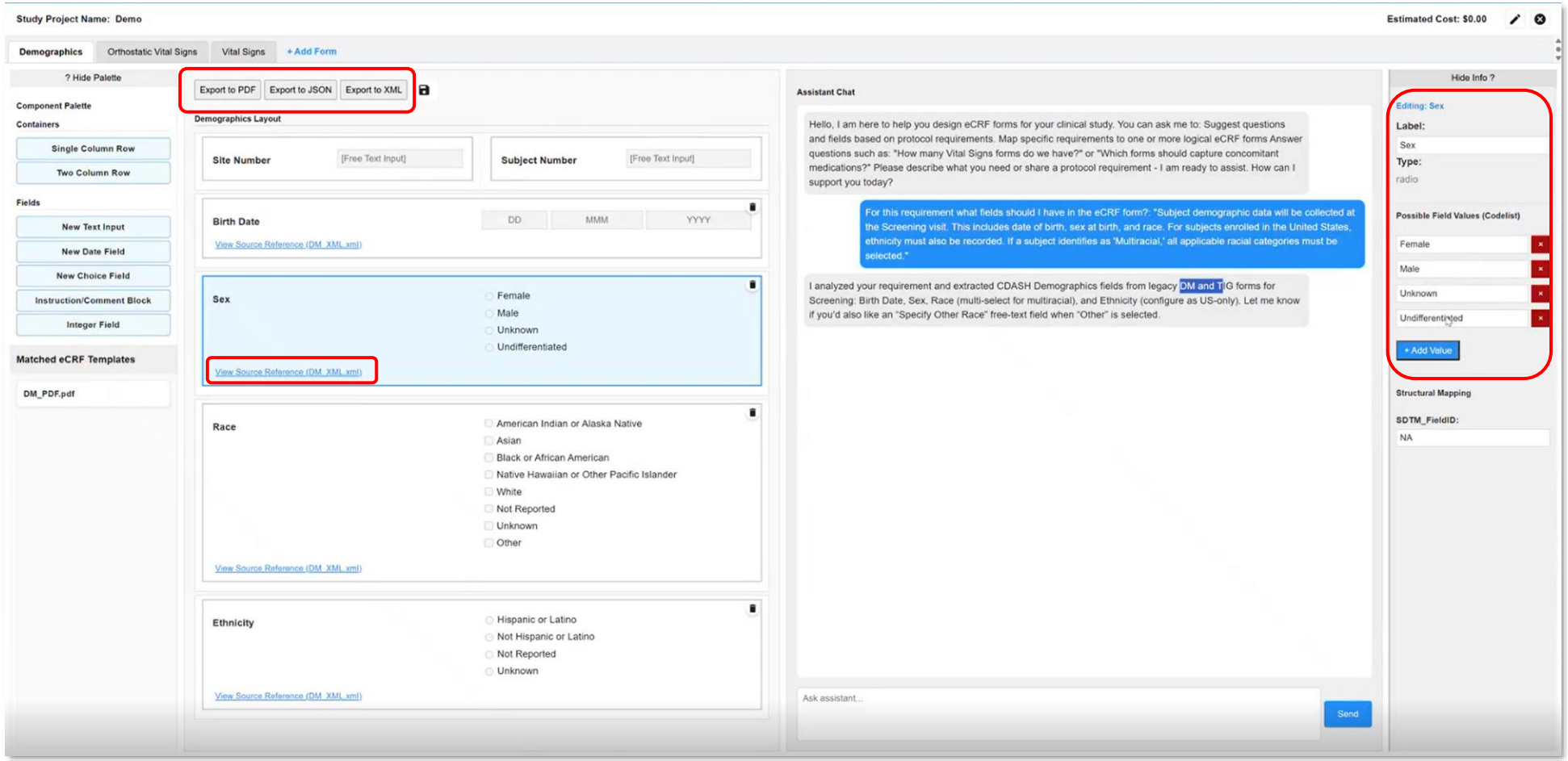
Upload Protocol, other digital documents, OR copy/paste study requirements.



Clinical Trials eCRF Build Automation

2

Edit and Export eCRFs
Edit metadata, add/edit new fields to eCRF form, export to PDF/JSON/XML.



Introducing Generative AI in Clinical Data Flow

Challenges in Current State of Clinical Trial Biometrics

Complex Clinical trials

- More trials globally
- Intricate data sources
- Traceability requirements

Manual data mapping coding

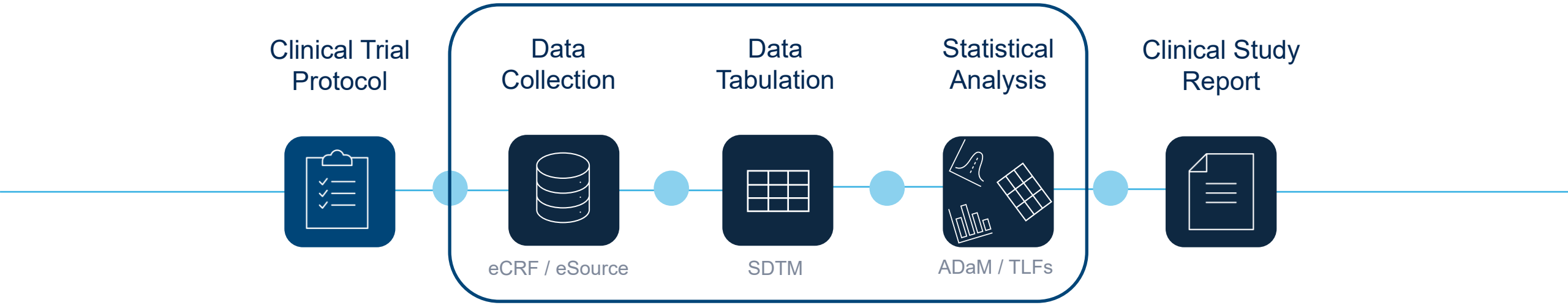
- Even with CDISC standards
- Time consuming
- Resource intensive

Double Programming

- Repetitive coding tasks
- Impacts submission deadlines

cdisc®

End to End Clinical Data Flow



- Raw data to SDTM format
- SDTM to ADaM
- Creation of define.xml
- Creation of TLFs

Clinical Data Flow (CDF) Prototype

- Will be part of SAS Clinical Acceleration

Features

Data Mapping

- Auto generate data transformation rules. (Raw -> SDTM -> ADaM)
- Investigate documents such as Study Protocols, SAP, TLF Shells to generate mapping specifications.

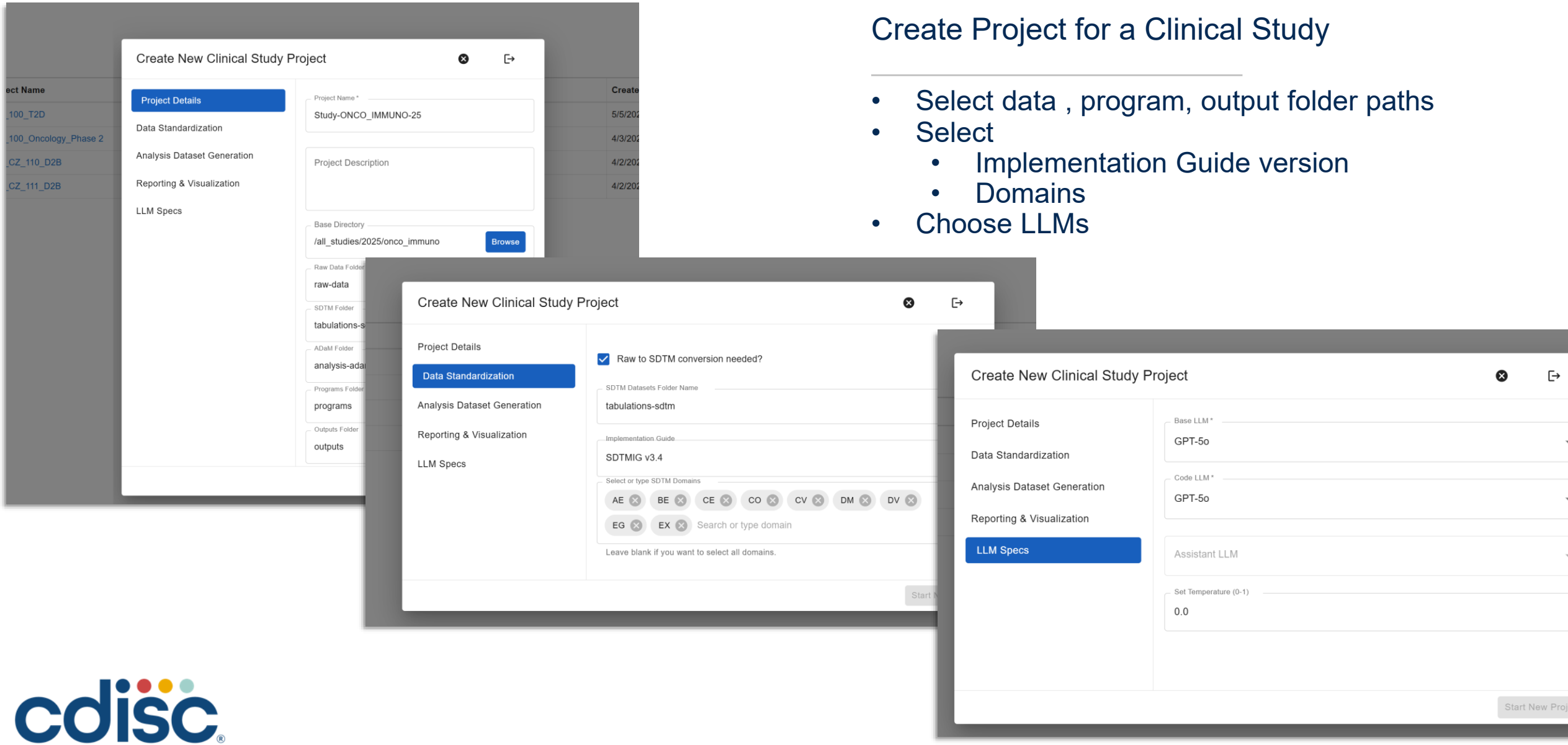
Code Generation

- SAS code generation with error feedback loop.
- Ensure Cyclic dependencies are taken care of when generating datasets.

Traceability

- Lineage view to track source to target data transformation.
- Source data coverage, missed variables reports.

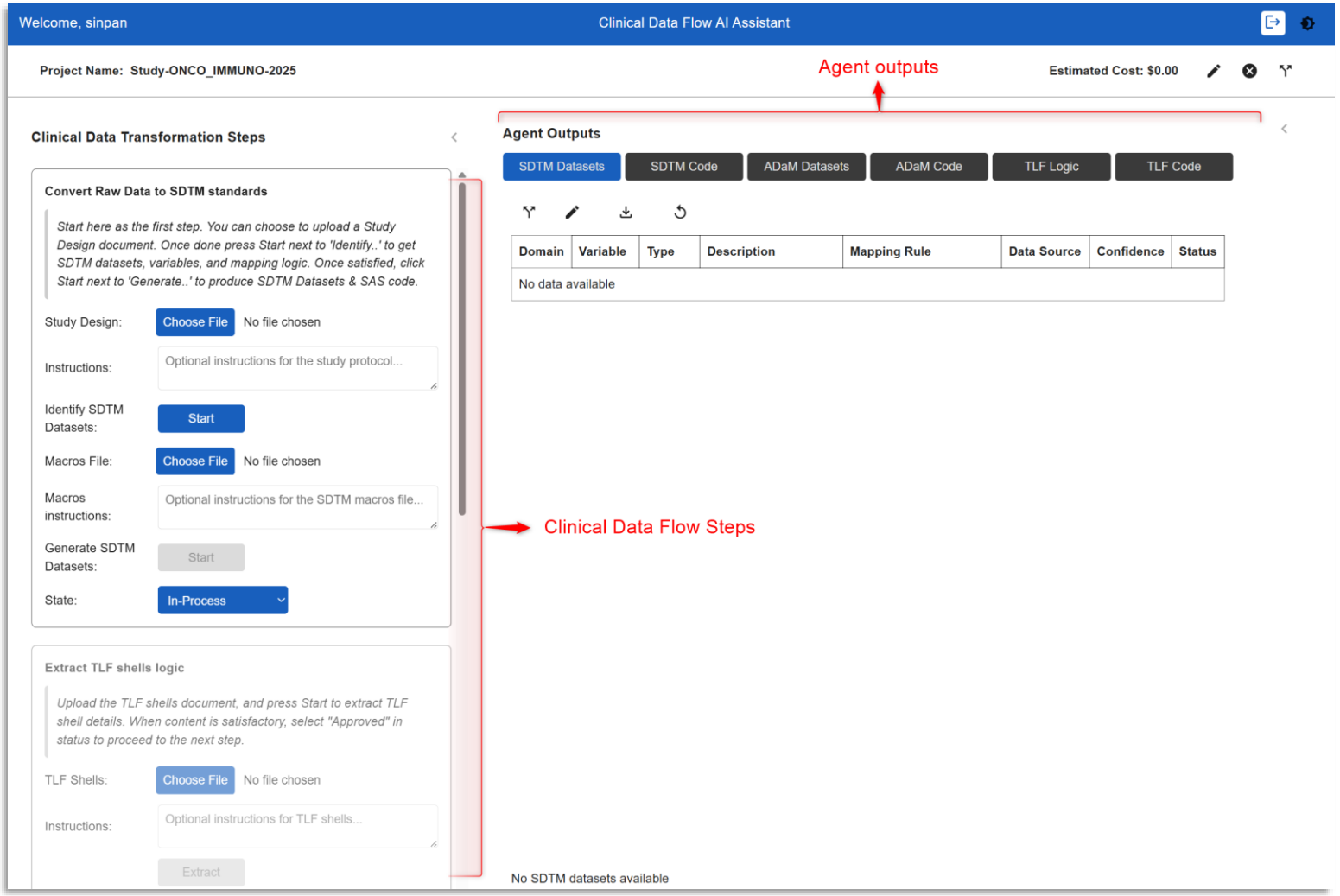
CDF Prototype - Create Project for Study



Create Project for a Clinical Study

- Select data , program, output folder paths
- Select
  - Implementation Guide version
  - Domains
- Choose LLMs

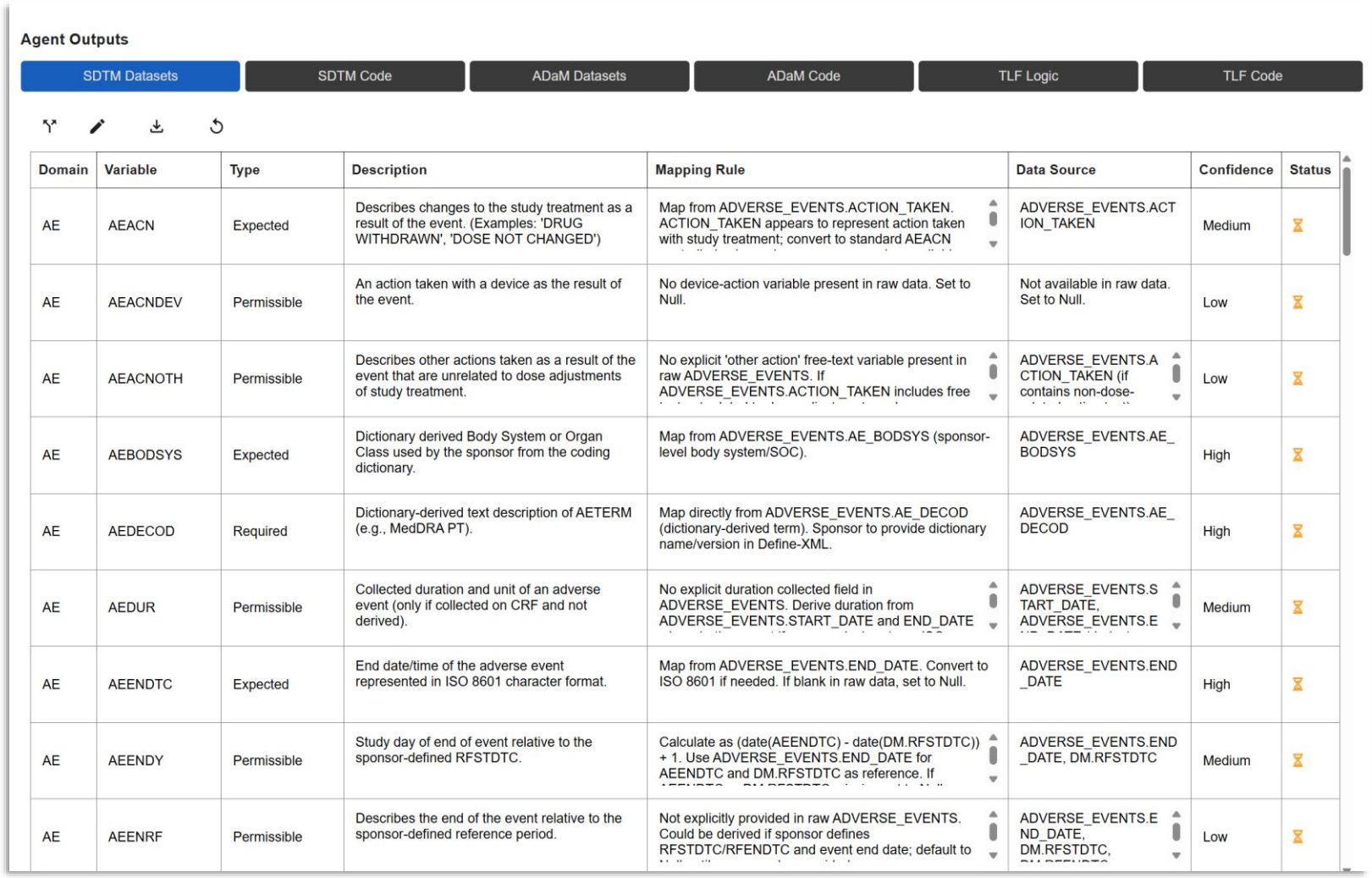
CDF Prototype - Project Details Screen



Project details

- Steps on the Left
- Agent (LLM) outputs on the right

CDF Prototype – SDTM Mapping



Agent Outputs

SDTM Datasets | SDTM Code | ADaM Datasets | ADaM Code | TLF Logic | TLF Code

Domain	Variable	Type	Description	Mapping Rule	Data Source	Confidence	Status
AE	AEACN	Expected	Describes changes to the study treatment as a result of the event. (Examples: 'DRUG WITHDRAWN', 'DOSE NOT CHANGED')	Map from ADVERSE_EVENTS.ACTION_TAKEN. ACTION_TAKEN appears to represent action taken with study treatment; convert to standard AEACN	ADVERSE_EVENTS.ACTION_TAKEN	Medium	
AE	AEACNDEV	Permissible	An action taken with a device as the result of the event.	No device-action variable present in raw data. Set to Null.	Not available in raw data. Set to Null.	Low	
AE	AEACNOTH	Permissible	Describes other actions taken as a result of the event that are unrelated to dose adjustments of study treatment.	No explicit 'other action' free-text variable present in raw ADVERSE_EVENTS. If ADVERSE_EVENTS.ACTION_TAKEN includes free	ADVERSE_EVENTS.ACTION_TAKEN (if contains non-dose-	Low	
AE	AEBODSYS	Expected	Dictionary derived Body System or Organ Class used by the sponsor from the coding dictionary.	Map from ADVERSE_EVENTS.AE_BODSYS (sponsor-level body system/SOC).	ADVERSE_EVENTS.AE_BODSYS	High	
AE	AEDECOD	Required	Dictionary-derived text description of AETERM (e.g., MedDRA PT).	Map directly from ADVERSE_EVENTS.AE_DECOD (dictionary-derived term). Sponsor to provide dictionary name/version in Define-XML.	ADVERSE_EVENTS.AE_DECOD	High	
AE	AEDUR	Permissible	Collected duration and unit of an adverse event (only if collected on CRF and not derived).	No explicit duration collected field in ADVERSE_EVENTS. Derive duration from ADVERSE_EVENTS.START_DATE and END_DATE	ADVERSE_EVENTS.START_DATE, ADVERSE_EVENTS.E	Medium	
AE	AEENDTC	Expected	End date/time of the adverse event represented in ISO 8601 character format.	Map from ADVERSE_EVENTS.END_DATE. Convert to ISO 8601 if needed. If blank in raw data, set to Null.	ADVERSE_EVENTS.END_DATE	High	
AE	AEENDY	Permissible	Study day of end of event relative to the sponsor-defined RFSTDTC.	Calculate as (date(AEENDTC) - date(DM.RFSTDTC)) + 1. Use ADVERSE_EVENTS.END_DATE for AEENDTC and DM.RFSTDTC as reference. If	ADVERSE_EVENTS.END_DATE, DM.RFSTDTC	Medium	
AE	AEENRF	Permissible	Describes the end of the event relative to the sponsor-defined reference period.	Not explicitly provided in raw ADVERSE_EVENTS. Could be derived if sponsor defines RFSTDTC/RFENDTC and event end date; default to	ADVERSE_EVENTS.END_DATE, DM.RFSTDTC,	Low	

SDTM Mappings

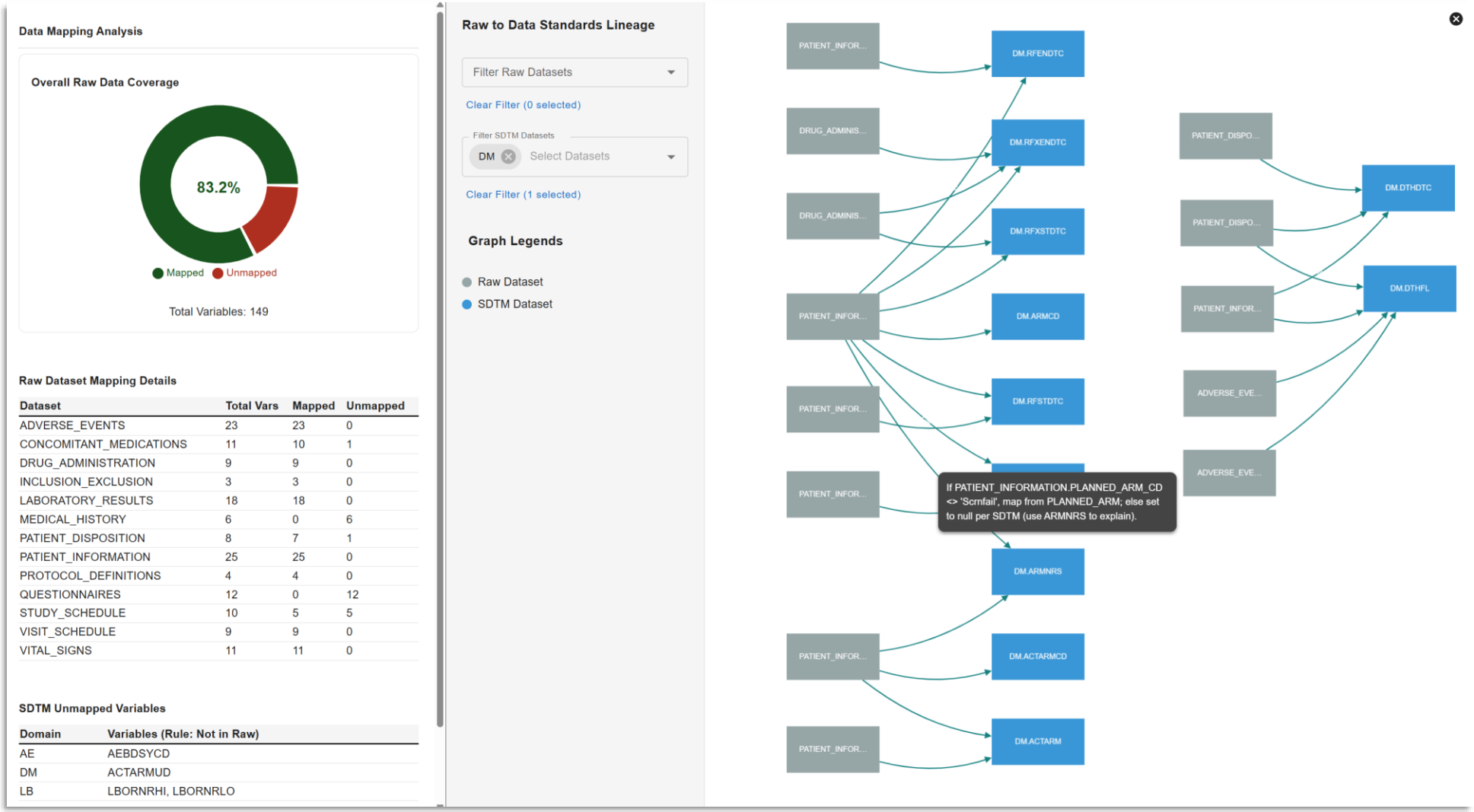
Inputs:

- Raw Data Schema
- Study Design Documents
- SDTMIG/Custom

Outputs:

- Editable Field Level mappings

CDF prototype - Traceability and Data Coverage



Lineage View

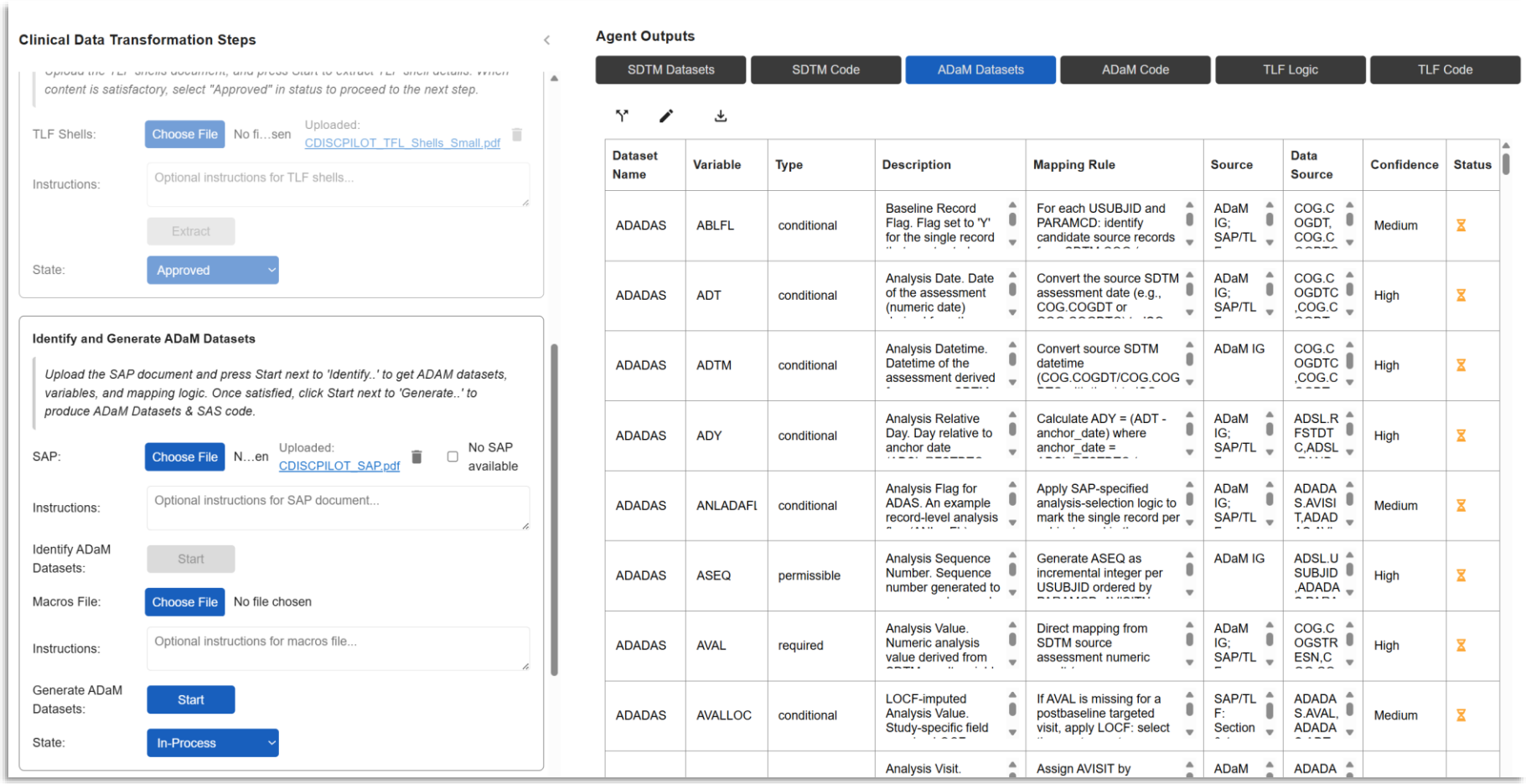
- Raw Data coverage reports.
- Variables with missing mappings.
- Graph view to track lineage.

CDF prototype – SDTM Code

SDTM Code

- Code generated with error feedback loop.
- Edit, re-generate, re-run code form this screen.

CDF prototype – ADaM Mappings



ADaM Mappings

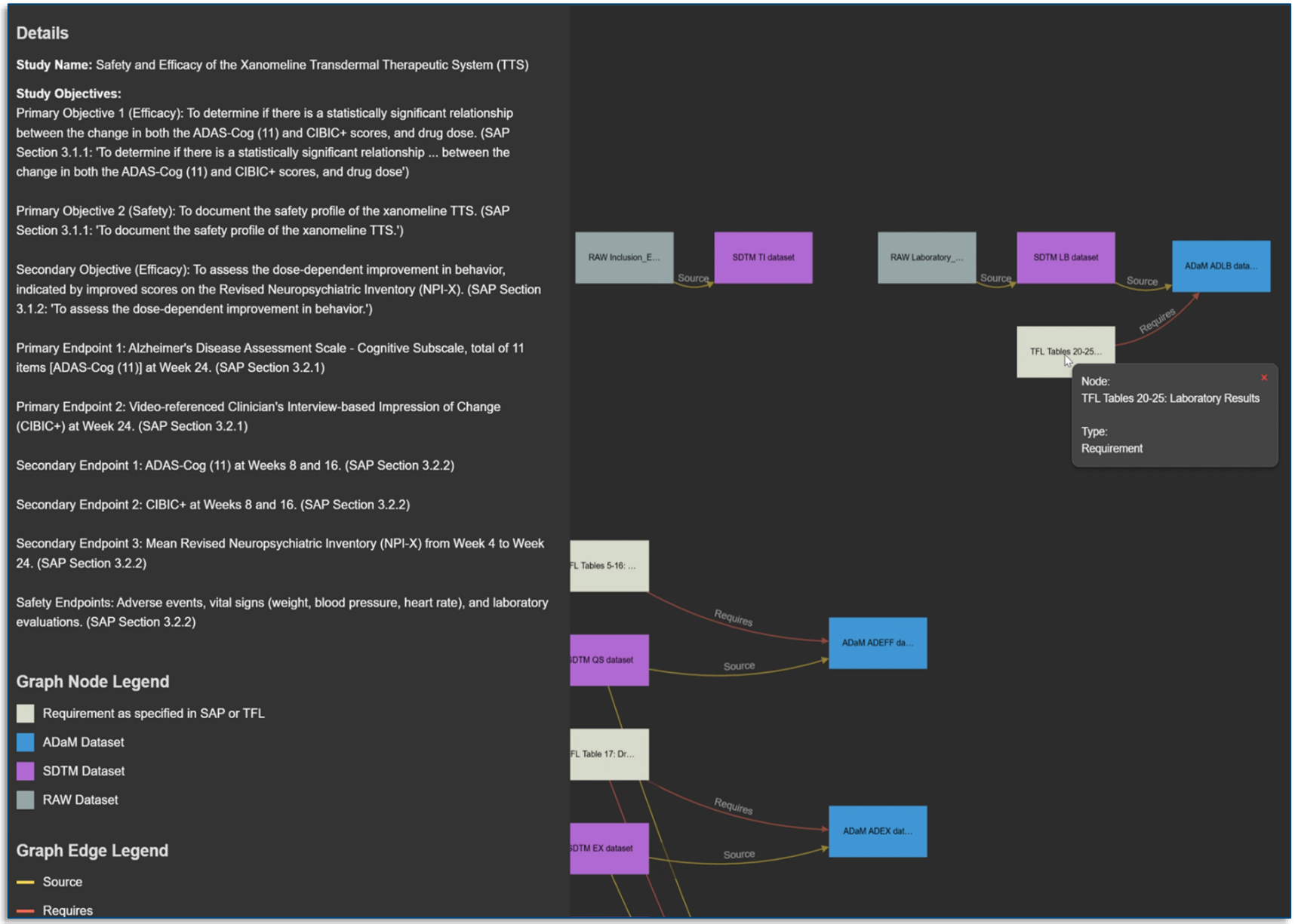
Inputs:

- TLF Shells Logic
- SAP
- SDTM datasets schema

Outputs:

- Editable Field Level mappings

CDF prototype - End to end traceability



Tracing SAP/TLFs requirements from Raw data to SDTM and ADaM datasets

Business Impact

Automated mapping & code generation

Faster builds, lower cost

AI AS ASSISTANT

LLMs enhance rather than replace human programmers, transforming routines into streamlined processes.

COMPRESS TIMELINES

Mapping, SDTM/ADaM, and TLF generation shrink from months to days.

JUMP-START BUT NOT TURNKEY

Automate bulk of the upstream process to create a robust baseline; humans finalize for submission readiness.

REALLOCATE EXPERTISE

Reallocate expertise to high-value work (e.g., complex efficacy datasets), high-level validation, scientific analyses.

cdisc®

Other AI tools supporting Standards Development in Novo Nordisk

Chatbot with domain knowledge

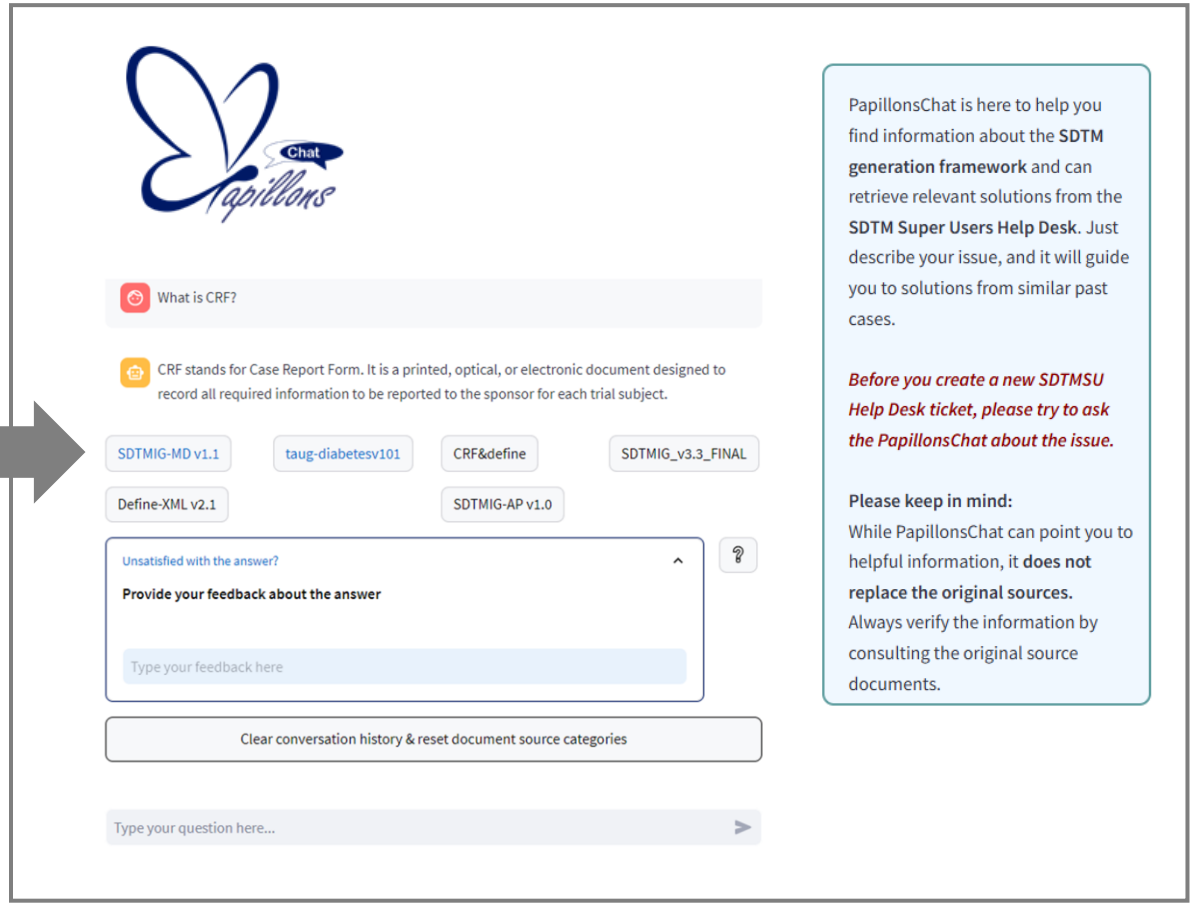
SharePoint lists

Internal documents about SDTM

Programs

SDTM training transcripts

Trusted websites



Anette Jakobsen
Anette.DalsgaardJakobsen@sas.com

Anja Jazbec
ajvr@novonordisk.com