



Turning Data into Deliverables: Automating Biometrics from Metadata to TFLs

Presenter Name: Inderjeet Arora

Email: Inderjeet.a222@geninvo.com

www.geninvo.com

North America | UK | Europe | India



Introduction

- TLF Importance: TLF generation is a key part of clinical trial analysis and reporting.
- Current Limitations: Existing macros and rule-based tools streamline parts of the process but leave the overall data flow fragmented and manual.
- Challenges: Multiple manual steps across SDTM and ADaM transformation lead to inefficiency, communication delays, and potential errors.
- Need for Integration: A unified, end-to-end automation platform can simplify data flow from raw data to TLFs.
- AI/ML Advantage: Intelligent automation enables smarter data mapping, faster processing, and reduced manual effort.
- Compliance Focus: Audit trails, governance, and access control ensure traceability and regulatory compliance.
- Goal: To demonstrate how AI-driven automation enhances efficiency, quality, and collaboration in clinical trial data workflows.



www.geninvo.com
North America | UK | Europe | India

Components

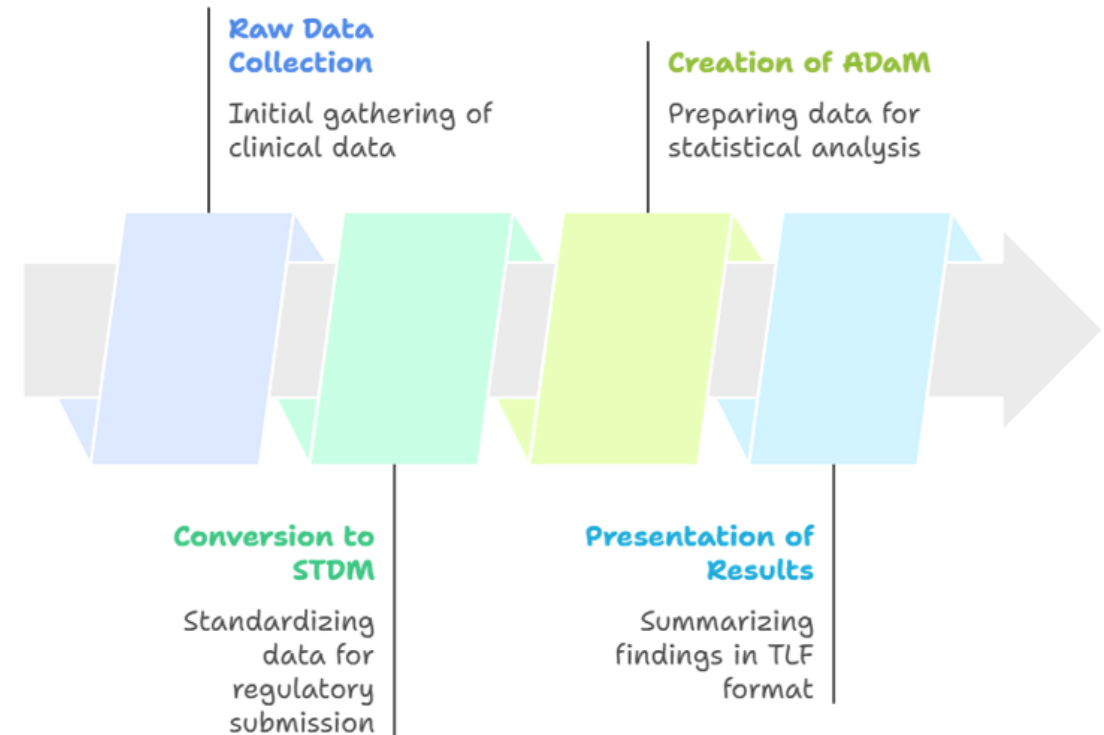
- Raw Data
- CDISC
 - SDTM
 - ADaM
- MDR
- SAP
- Protocol
- Mock Shells



www.geninvo.com
North America | UK | Europe | India

Traditional Approach

- Manual Programming
- Sequential Workflow
- High Dependence on Expertise
- Time-Intensive Process
- Limited Reusability
- Compliance Through Manual Validation



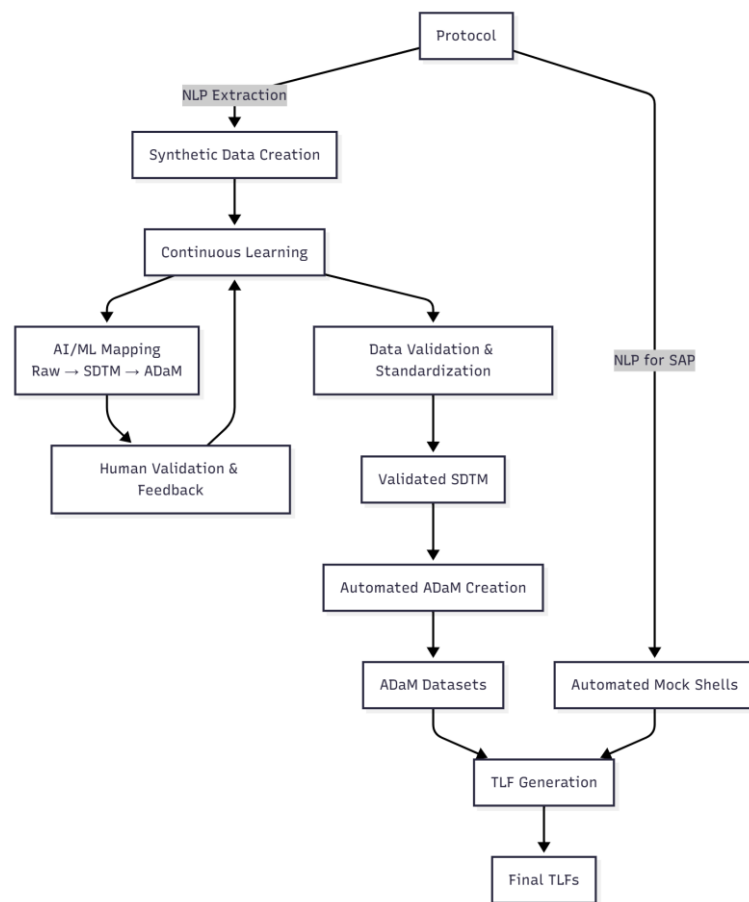
Traditional Approach - Limitations

- High manual effort: Each TLF requires individual programming and validation.
- Time-consuming process: Multiple programming, validation, and review cycles extend study timelines.
- Limited reusability: Programs are often study-specific, reducing standardization and reuse.
- High risk of human error: Manual coding and updates can lead to inconsistencies or inaccuracies.
- Resource intensive: Requires extensive involvement from programmers, statisticians, and reviewers.
- Complex traceability maintenance: Linking protocol, SAP, ADaM, and final TLFs demands meticulous documentation.
- Change management challenges: Any modification to SAP or data structure often requires reprogramming multiple outputs.



www.geninvo.com
North America | UK | Europe | India

Our Approach



www.geninvo.com
North America | UK | Europe | India

Our Approach – Synthetic Data

- AI/NLP extracts key study elements (objectives, endpoints, visits, populations, treatments) directly from the clinical protocol.
- Extracted details are converted into structured metadata for early setup.
- Synthetic data is generated to simulate realistic, privacy-safe raw datasets.
- Enables early metadata design, mapping, and automation testing before real data collection.



www.geninvo.com
North America | UK | Europe | India



Our Approach – Metadata Repository

- System automatically extracts metadata from synthetic raw datasets — including dataset structures, variable details, formats, and derivation rules.
- All extracted metadata is stored in a centralized Metadata Repository (MDR) serving as the core backbone of the process.
- MDR maintains both study-specific and standardized metadata libraries (CDISC SDTM and ADaM).
- Enables traceability, cross-linking, and reuse of mappings and derivation logic across studies.



www.geninvo.com
North America | UK | Europe | India

Our Approach – AI/ML-Powered Metadata Mapping

- AI/ML engine automates mapping from Raw → SDTM → ADaM.
- Predicts variable relationships using pattern recognition and learning models.
- Human-in-the-loop ensures accuracy and compliance.
- Validated mappings stored in MDR for reuse.
- System learns over time, enabling auto-mapping for future studies and faster setup.



www.geninvo.com
North America | UK | Europe | India

Our Approach – SDTM Mapping

Reporting activity: TTCQFT Study: FOAJVG [Next →](#)

Selected Study: FOAJVG

Source: RAW [View All Mapping](#)

Domain: ae [View Source Dataset →](#)

Variable: Choose source variable...

Package/Tags: All Package... [Edit Method](#)

Method: Choose Method... [clear all](#)

Source Values

Mapping: Run Variable Prediction

Target: SDTM

Version: 3.3

Domain: AE

Variable: Choose target variable...

Mapped Elements:

☒ Required/Expected ☐ Permissible

100.00% 23

[Run Variable Prediction](#) » [View Mapping Definition](#) » [Execute](#) » [Review Data](#) » [Generate Output](#) » [View Log](#) » [View Error Log](#)

Variable Prediction Data

#	Raw Variable	Raw Label	Model Value	Probability
1	DOMAIN	Domain Abbreviation	DOMAIN	0.94
2	AESEQ	Sequence Number	AESEQ	0.92
3	AETERM	Reported Term for the Adverse Event	AETERM	0.91
4	AELT	Lowest Level Term	AELT	0.86
5	AELTCD	Lowest Level Term Code	AELTCD	0.86
6	AEDECD	Dictionary-Derived Term	AEDECD	0.9
7	AEPTCD	Preferred Term Code	AEPTCD	0.93
8	AEHLT	High Level Term	AEHLT	0.87
9	AEHLTCD	High Level Term Code	AEHLTCD	0.88
10	AEHLGT	High Level Group Term	AEHLGT	0.98
11	AEHLGTC	High Level Group Term Code	AEHLGTC	0.9
12	AEBOCD	Body System or Organ Class	AEBOCD	0.89
13	AEBOCD	Body System or Organ Class Code	AEBOCD	0.96

Req/Exp: 22% Perm: 0% [View Mapping Definition](#) [Reset](#)



www.geninvo.com
North America | UK | Europe | India

Our Approach – Automated Mock Shell Creation

- Runs in parallel with data mapping process
- Uses NLP to extract key specifications from SAP and Protocol
- Identifies:
 - Analysis parameters (e.g., endpoints, change-from-baseline)
 - Analysis populations (e.g., Safety, Efficacy, ITT)
 - Visit/timepoint details
 - Statistical methods and output formats
- Enables auto-generation of mock shells and TLF templates



www.geninvo.com
North America | UK | Europe | India

Our Approach – Mock Shell Generator

Create New Table

Name *

Enter new TFL name

Table Type *

Demographic Table

Protocol: *

-- select --

Scrap Data

Artifact:

Treatment Variable

Value:

Please select artifact from list

No Additional Column ☒ | Add **Total** Column ☐ | Add **Overall** Column ☐

Create Table

Please right click on the items to see more options

Page 1 of 1

GENINVO				
Head				
Study Name : GeninvoStudy				
Title	Analysis Set			
	Active Intervention: 150 Mg Bid	Active Intervention: 75 Mg Bid	Active Intervention: 25 Mg Bid	Placebo
AGE5				
n	XX	XX	XX	XX
Mean (SD)	XX.X (X.XX)	XX.X (X.XX)	XX.X (X.XX)	XX.X (X.XX)
Median	XX.X	XX.X	XX.X	XX.X
Min, Max	XX, XX	XX, XX	XX, XX	XX, XX
ETHNICITY				
Hispanic	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Not Hispanic	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Not Reported	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
Unknown	XX (XX.X)	XX (XX.X)	XX (XX.X)	XX (XX.X)
HEIGHT				
n	XX	XX	XX	XX
Mean (SD)	XX.X (X.XX)	XX.X (X.XX)	XX.X (X.XX)	XX.X (X.XX)
Median	XX.X	XX.X	XX.X	XX.X
Min, Max	XX, XX	XX, XX	XX, XX	XX, XX



www.geninvo.com
North America | UK | Europe | India

Our Approach – Automated TLF Generation

- ADaM datasets feed directly into automated TLF generation
- Statistical computations (mean, SD, change from baseline) applied automatically
- Outputs generated in multiple formats — RTF, PDF, Excel, or dashboards
- Ensures accuracy, consistency, and end-to-end traceability



www.geninvo.com
North America | UK | Europe | India



Our Approach – TLF Generator

Report Generator1.0

Select Table
Demographic Table

Select Mockshells
-- Select --
Only exported mockshells will be displayed here for reference

Select Dataset
adslxpt

Select Population * ⓘ
Safety

Template View Dataset View

+ Apply filters for dataset

Character Variables ⓘ
RACE x SEX x

Numeric Variables ⓘ
AGE x

Choose Stats * ⓘ
n x Mean (SD) x Median x Min, Max x

Treatment Variables * ⓘ
ARM

☐ Additional Grouping

☒ Total Summary Column ⓘ

☒ N Labels ⓘ

Display Demographic Table

1 of 2

Test

Table 1.0: DM Safety

	100		
	Active intervention: 25 mg BID N = 72	Active intervention: 75 mg BID N = 58	150 N = 58
AGE			
n	72	58	58
Mean (SD)	46.3 (11.57)	54.9 (12.75)	55.9 (13.75)
Median	59.0	59.0	59.0
Min, Max	18, 78	18, 78	18, 78
RACE - n (%)			
ASIAN	4 (5.6)	1 (1.7)	1 (1.7)
BLACK OR AFRICAN AMERICAN	13 (18.1)	3 (5.2)	5 (8.6)
OTHER	0	1 (1.7)	1 (1.7)
WHITE	55 (76.4)	53 (91.4)	51 (87.9)
SEX - n (%)			
F	72 (100.0)	5 (8.6)	5 (8.6)
M	0	53 (91.4)	53 (91.4)

Table

Output Path : ttf\tf-data\S1\Output\1.0.pdf
Program Path : app\ttf\src\dm\report_generation_dm.R
Date and Time program was run: 19-02-2025 10:25:48 AM



www.geninvo.com
North America | UK | Europe | India

Our Approach – Continuous Learning And Reusability

- System acts as a self-learning ecosystem
- Reviewed mappings, derivations, and TLF definitions are fed back into AI and stored in the MDR
- Enables the system to:
 - Recognize similar data structures in new studies
 - Reuse validated transformations
 - Improve mapping accuracy over time
 - Shorten setup and validation cycles
- Builds a rich, reusable metadata repository for future studies



www.geninvo.com
North America | UK | Europe | India



Ideas
Delivered
Intuitively

Thank You!

+1-833-GEN-INVO (436-4686) | info@geninvo.com | www.geninvo.com

