



connect·share·advance

Single Day Events

Trivandrum (Kerala), India
20-Jun-2026

phuse.global



connect·share·advance

AI/ML-Driven Automation of SDRG & ADRG Generation

Using a Human-in-the-Loop
Framework

Balram Chauhan
Cognizant Technology Solution

phuse.global



Will Discuss About

- **Background**
 - SDRG & ADRG in the Define Package
- **The Problem**
 - Manual, Template-Driven, Inconsistent
- **The Opportunity**
 - AL/ML pipeline + Human-in-the-loop
- **Intelli Define Tool**
 - Automated generation of SDRG / ADRG
- **Benefit and Business Impact**
- **Questions**



connect·share·advance

Background

SDRG & ADRG in the Define Package



Background

connect·share·advance

- **SDRG and ADRG are key regulatory documents & submitted with Define package (SDTM & ADaM)**
 - Below details are submitted in SDRG/ADRG:

SDRG

(Study Data Reviewer's Guide)

- Accompanies SDTM datasets
- Protocol & study design summary
- Standards & versions implemented
- Data conformance (Pinnacle 21) summary
- Explains 'Not Submitted' aCRF fields

ADRG

(Analysis Data Reviewer's Guide)

- Accompanies ADaM datasets
- Analysis considerations & conventions
- Key derivation & imputation logic
- Traceability SDTM → ADaM → TLF
- Conformance & known issues.



connect·share·advance

The Problem

Manual, template-driven, Inconsistent

- **Manual authoring is slow and inconsistent**

- Reviewer guides are mostly assembled from information that already exists — in the protocol, the annotated CRF, the metadata and the validation reports. The bottleneck is human transcription.

Time-Intensive

Days to weeks per study, repeated across every submission and amendment.

Inconsistent

Manually copying legacy study templates caused inconsistencies in wording, structure, and completeness.

SDRG/ADRG

Re-Work

Late data refreshes and re-runs of conformance reports force repeated manual edits.

Error-prone

Manually transcribed metadata and P21 Validation reports increase the risk of inconsistencies with Define-XML.



connect·share·advance

The Opportunity

AL/ML pipeline + Human-in-the-loop



The Opportunity

connect·share·advance

- Why AI/ML?

Document Understanding

Modern NLP / LLMs parse protocols and aCRFs to extract titles, design, objectives and field annotations

ML Analysis

Machine Learning and Rule based models analyze structured metadata for datasets summaries

AI/ML

Repeatable & Consistency

Automation engine maps extracted insights to standardized templates, which results in consistency across studies

Repeatable & Auditable

Templates encodes regulatory structure once, the model fills them consistently across all studies

- **A human-in-the-loop generation framework**

- Machines do the heavy lifting of extraction and drafting , subject matter experts reviews AI generated drafts to ensure accuracy and regulatory completeness.

INGEST

- Protocol PDF
- Annotated CRF
- SDTM / ADaM Specs
- Define.XML
- Conformance Reports

GENERATE

- Selection
- Extraction
- Summarisation
- Section Templating
- Citation Tagging

VALIDATION

- SME Review
- Inline edit & accept
- Conformance Check
- Final SDRG /ADRG



connect·share·advance

.Intelli Define Tool

Automated solution generation of SDRG/ADRG

- Structured Sources : P21 Conformance Reports, Metadata
- Unstructured Sources: Study Protocol, Annotated CRF

Intelli Define Tool Balram Chauhan Admin Logout

[← Back to Studies](#)

Study 1
Study 1

- ✓ ADAM
- ✓ ADAM vs Define
- ✓ Define
- Additional SDTM Checks

GENERATOR

- Define.xml
- Excel Spec from XPT
- Excel Spec from Define.xml
- SDRG**
- ADRG

Status New

Delete Project

Protocol Document * (.pdf, .docx)

 [Click to upload or browse](#)
.pdf or .docx files

Annotated CRF Document * (.pdf, .docx)

 [Click to upload or browse](#)
.pdf or .docx files

Metadata Excel * (.xlsx, .xls)

 [Click to upload or browse](#)
.xlsx or .xls files

Validation Report Excel * (.xlsx, .xls)

 [Click to upload or browse](#)
.xlsx or .xls files

Unstructured Data

Structured Data



Methods

connect·share·advance

Sections are auto-populated:

Guide Section	Primary Source	Methods
Protocol title & study design	Study protocol	AI/ML Parsing
Standards & versions implemented	User Configuration Input	Positional placement
Data conformance summary	Conformance reports	Classification & Placement
Derivation descriptions (ADRG)	Specs / SAP notes	AI/ML Parsing & Summarization
'Not Submitted' CRF fields	Annotated CRF	Annotation rules with AI/ML Parsing
Dataset categorization & summarization	SDTM metadata + IG	Summarization



Document Editor

connect·share·advance

Intelli Define Tool

Balram Chauhan Admin

Study 1

Study 1

Project Summary

VALIDATOR

- SDTM
- SDTM vs Define
- ADAM
- ADAM vs Define
- Define
- Additional SDTM Checks

GENERATOR

- Define.xml
- Excel Spec from XPT

Status

New

Delete Project

SDRG Document Editor
No changes

DOCUMENT OUTLINE

Introduction

- Purpose
- Acronyms
- Study Data Standards and D...

Protocol Description

- Protocol Number and Title
- Protocol Design
- Trial Design Datasets
 - 2.3.1 TA – Trial Arms
 - 2.3.2 TE – Trial Elements
 - 2.3.3 TV – Trial Visits
 - 2.3.4 TI – Trial Inclusion/E...
 - 2.3.5 TS – Trial Visits

Add Paragraph

Add Table

Save Changes

Export DOCX

Study Data Standards and Dictionary Inventory

Standard or Dictionary	Versions Used	
SDTM	3.4	×
Controlled Terminology	2024-09-27	×
Data Definitions	define.xml v2.1	×
Medications Dictionary	Placeholder	×
Medical Events Dictionary	Placeholder	×
+ Add Row		

Protocol Description

Protocol Number and Title

Protocol Number: CTJ301UC201

Protocol Title: A Phase II, Randomized, Double-blind, Placebo-controlled Study to Evaluate the Safety and Efficacy of TJ301 (FE 999301) Administered Intravenously in



connect·share·advance

Benefits and Business Impact



Benefits

connect·share·advance

Increasing Organizational compliance over time

Efficiency Gains

AI/ML automation reduces time upto 70%

Enhanced Quality

Better Define.xml alignment lower regulatory risk, shorten review time, and improves inspection readiness

Benefits

Consistency

Standardized generation logic ensure consistent output across studies, minimizing variability from different authors

Audit Trails & Version Control

Every Validation reports and document generated link back to its results and Historical Navigation



Business Impact

connect·share·advance

Faster, more Consistent Reviewer Guides



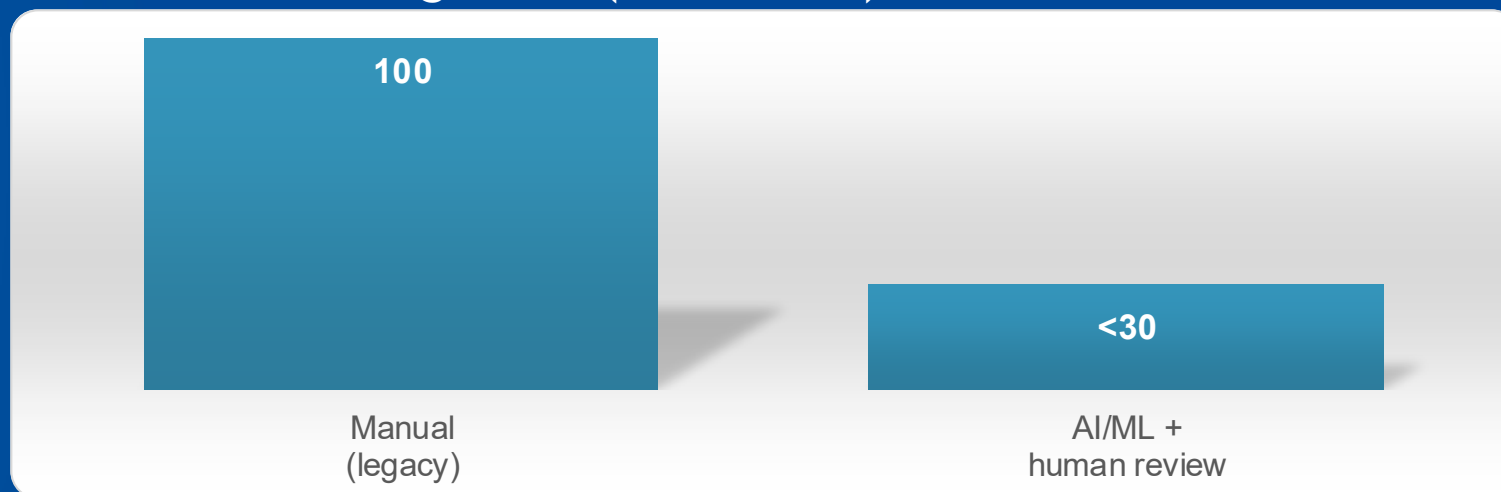
up to

70-85%

reduction in preparation time

for the SDRG / ADRG authoring cycle

Relative authoring effort (illustrative)



Consistency

Uniform structure and wording across studies.



Repeatability

Re-runs after data refresh in minutes, not days.



Traceability

Source-linked content speeds reviewer Q&A.



connect·share·advance

Questions & Discussion

connect·share·advance



THANK YOU!

connect·share·advance



The Global Healthcare Data Science Community

Contact Channels

✉ office@phuse.global
🌐 phuse.global

Social Media

🌐 [/company/phuse](https://www.linkedin.com/company/phuse)
▶ [/phusetube](https://www.youtube.com/channel/UCphuse)
✕ [/phuse_global](https://twitter.com/phuse_global)
f [/phusebook](https://www.facebook.com/phusebook)