



# Hands-on Workshop: Deep-dive Review and Testing of CDISC Analysis Results Standards Logical Model and Schema

**PHUSE US Connect: March 06, 2023**

Bhavin Busa, Clinical Data SME, Independent

Bess LeRoy, MPH, Head of Standards Innovation, CDISC

Richard Marshall, Principal Data Modeler, CDISC





# Agenda

1. Background
2. Use Cases for Analysis Results Standards
3. Analysis Results Key Objectives and Key Results
4. Analysis Results Logical Metamodel
  - Representations of model in CMAP, Markdown, and JSON Schema
5. ARS Development on GitHub
6. Review Examples
7. Reference implementation
8. ARS Roadmap
9. Q&A





# Use Cases for Analysis Results Standards

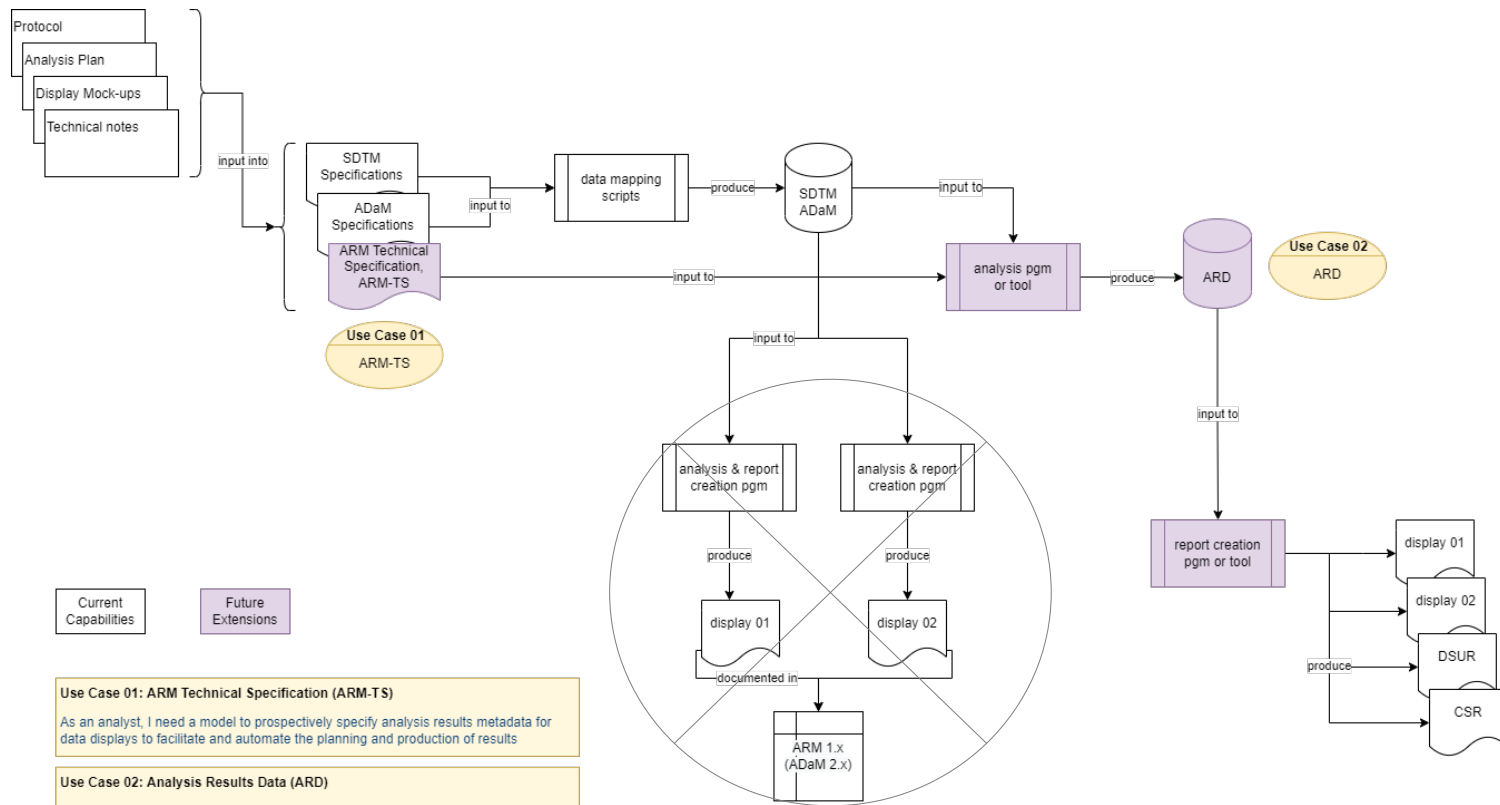
- Use Case 01:

*As an analyst, I need a technical specifications to prospectively specify analysis results metadata for data displays to facilitate and automate the planning and production of results*

- Use Case 02:

*As an analyst, I need a structure to represent analysis results and qualifying metadata to support traceability, reproducibility, reusability and quality*

# Workflow with Future Extensions



Current Capabilities      Future Extensions

**Use Case 01: ARM Technical Specification (ARM-TS)**  
 As an analyst, I need a model to prospectively specify analysis results metadata for data displays to facilitate and automate the planning and production of results

**Use Case 02: Analysis Results Data (ARD)**  
 As an analyst, I need a structure to represent analysis results and qualifying metadata to support traceability, reproducibility, reusability and quality



# Analysis Results Key Objectives

- Use analysis results metadata to drive the automation of results
- Support storage, access, processing and reproducibility of results
- Improved navigation and reusability of analyses and results
- Traceability to Protocol/SAP and to input ADaM data

# Initial Analysis Results Standards Key Results



Develop a technical specification to prospectively leverage Analysis Results Metadata to drive automation



Develop a structure to represent Analysis Results as data



Illustrate and exercise with a set of common data displays

# Concepts Team Consulted Published Layouts

- PH

Scatterplot and Shift Table Summary of Absolute Lab values – Lab Test 1 Minimum Baseline vs Minimum Post-baseline

- Treatment

- T1 (N = xxx)

3.14安全性の解析（バイタルサイン、身体的所見及び安全性に関連するTable summary of vital signs by visit

- T2 (N = xxx)

<Parameter> *BDS.PARAM*  
 <Visit> *BDS.AVISIT*  
 n  
 Mean (SD)  
 Median *BDS.AVAL*  
 Min - Max

N = number of subjects in each treatment group using the referential demographics.

上記例は、絶対値の集計の場合。  
 バイタルサインのベースラインからの変化量を集計する必要がある場合は *BDS.PCHG* を使用する

- JPI

- Re

- Sul

Table 3. Laboratory Abnormalities that Worsened from Baseline to Grade 3 or 4 Occurring in ≥1% of Patients with dMMR Endometrial Cancer Receiving Product in Study

Laboratory Test	Product N = 104	
	All Grades <sup>a</sup> %	Grade 3 or 4 <sup>a</sup> %
<b>Hematology</b>		
Decreased lymphocytes	37	9
Decreased leukocytes	21	2.9
<b>Chemistry</b>		
Decreased albumin	30	2.9
Increased creatinine	27	2.9
Increased alkaline phosphatase	25	2.9
Increased aspartate aminotransferase	16	1.9
Increased alanine aminotransferase	15	2.9
<b>Electrolytes</b>		
Decreased sodium	26	4.8
Increased calcium	15	1.9
Decreased potassium	15	1.9

<sup>a</sup> Consists of new onset of laboratory abnormality or worsening of baseline laboratory abnormality.



# Focus on Concepts, Not Layout

lays not on subjective

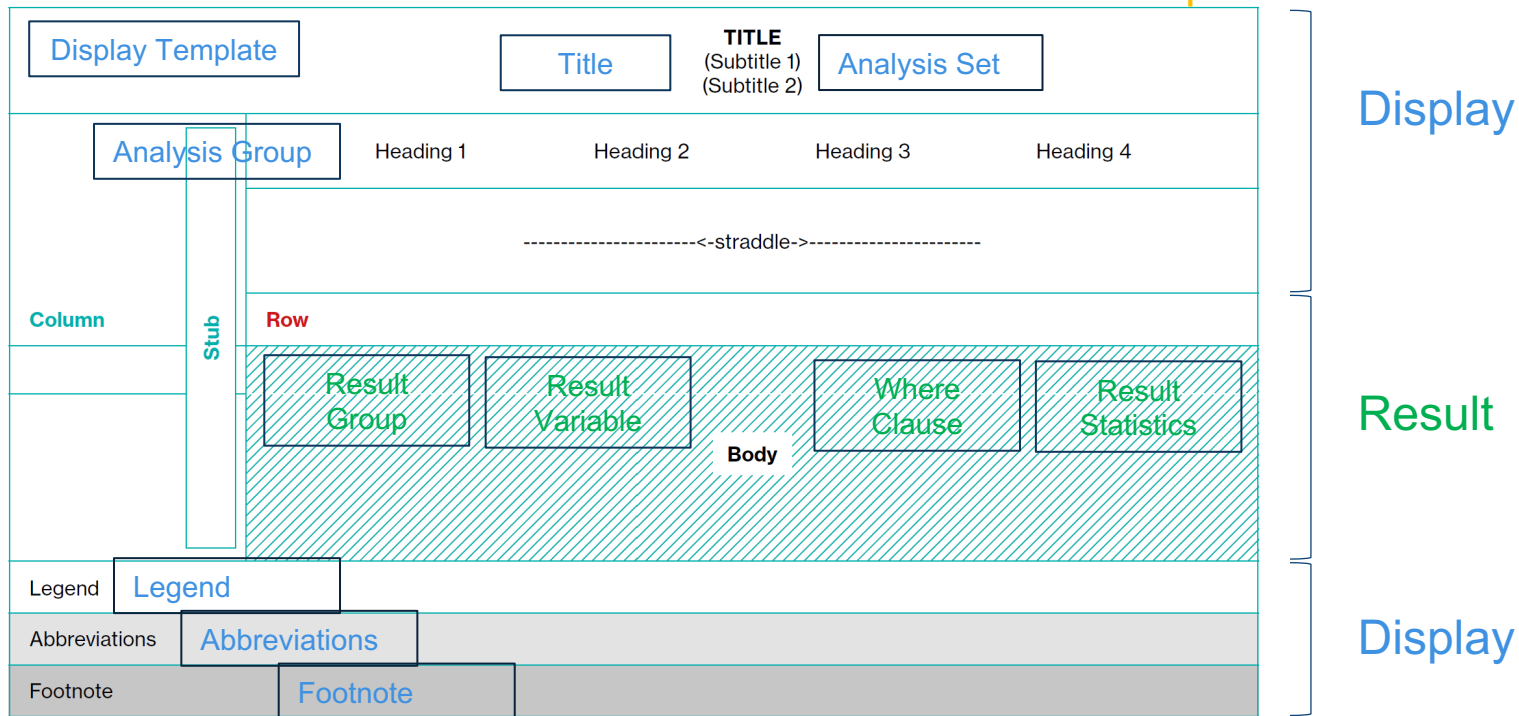
Parameter (Units) Visit	Treatment X (N=XX)	Treatment Y (N=XX)	Total (N=XX)
<Parameter 1> (<unit>)			
Baseline			
n	XX	XX	XX
Mean (SD)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
Median	XX.X	XX.X	XX.X
Q1, Q3	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X
Min, Max	XX, XX	XX, XX	XX, XX
< Visit n >			
n	XX	XX	XX
Mean (SD)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
Median	XX.X	XX.X	XX.X
Q1, Q3	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X
Min, Max	XX, XX	XX, XX	XX, XX
< Visit n Change from Baseline >			
n	XX	XX	XX
Mean (SD)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
Median	XX.X	XX.X	XX.X
Q1, Q3	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X
Min, Max	XX, XX	XX, XX	XX, XX
...			
<Visit n>			
n	XX	XX	XX
Mean (SD)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
Median	XX.X	XX.X	XX.X
Q1, Q3	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X
Min, Max	XX, XX	XX, XX	XX, XX

Parameter (Units) Visit	Treatment X (N=XX)		Treatment Y (N=XX)		Total (N=XX)	
	Observed	CFB	Observed	CFB	Observed	CFB
<Parameter 1> (<unit>)						
Baseline						
n	XX	XX	XX	XX	XX	XX
Mean (SD)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
Median	XX.X	XX.X	XX.X	XX.X	XX.X	XX.X
Q1, Q3	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X
Min, Max	XX, XX	XX, XX	XX, XX	XX, XX	XX, XX	XX, XX
...						
<Visit n>						
n	XX	XX	XX	XX	XX	XX
Mean (SD)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
Median	XX.X	XX.X	XX.X	XX.X	XX.X	XX.X
Q1, Q3	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X
Min, Max	XX, XX	XX, XX	XX, XX	XX, XX	XX, XX	XX, XX

# Key Metadata Elements of a Table

Output



Reference: PHUSE White Paper "General Output Tips and Considerations", Doc ID: WP-034, Version 1.0, Aug 2020

# Demographics Analysis Results and Metadata

Display Template

Title

Analysis Set

**Table 2. Baseline Demographic and Clinical Characteristics, Safety Population, Pooled Analyses (or Trial X)**

Characteristic	Analysis Group	Drug Name Dosage X N = XXX	Drug Name Dosage Y N = XXX	Placebo N = XXX	Active Control N = XXX	Total Population N = XXX
		n (%)	n (%)	n (%)	n (%)	n (%)
<b>Sex, n (%)</b>		n (%)	n (%)	n (%)	n (%)	n (%)
Male		n (%)	n (%)	n (%)	n (%)	n (%)
Female		n (%)	n (%)	n (%)	n (%)	n (%)
<b>Age, years</b>		X.X (Y.Y)	X.X (Y.Y)	X.X (Y.Y)	X.X (Y.Y)	X.X (Y.Y)
Mean (SD)		X.X (Y.Y)	X.X (Y.Y)	X.X (Y.Y)	X.X (Y.Y)	X.X (Y.Y)
Median (min, max)		X.X (Y.Y, Z.Z)	X.X (Y.Y, Z.Z)	X.X (Y.Y, Z.Z)	X.X (Y.Y, Z.Z)	X.X (Y.Y, Z.Z)
<b>Age groups (years), n (%)</b>		n (%)	n (%)	n (%)	n (%)	n (%)
≥17 to <65	Result Group		Result Variable	Where Clause	Result Statistics	n (%)
≥65						n (%)
≥65 to <75						n (%)
≥75		n (%)	n (%)	n (%)	n (%)	n (%)
<b>Race, n (%)</b>		n (%)	n (%)	n (%)	n (%)	n (%)
American Indian or Alaska Native		n (%)	n (%)	n (%)	n (%)	n (%)
Asian		n (%)	n (%)	n (%)	n (%)	n (%)
Black or African American		n (%)	n (%)	n (%)	n (%)	n (%)
Native Hawaiian or Other Pacific Islander		n (%)	n (%)	n (%)	n (%)	n (%)
White		n (%)	n (%)	n (%)	n (%)	n (%)
Other		n (%)	n (%)	n (%)	n (%)	n (%)

Source: [include Applicant source, datasets and/or software tools used].

<sup>1</sup> Difference is shown between [treatment arms] (e.g., difference is shown between Drug Name dosage X vs. placebo).

Abbreviations: N, number of patients in treatment arm; n, number of patients with given characteristic; SD, standard deviation

Footnote

Abbreviations

Legend

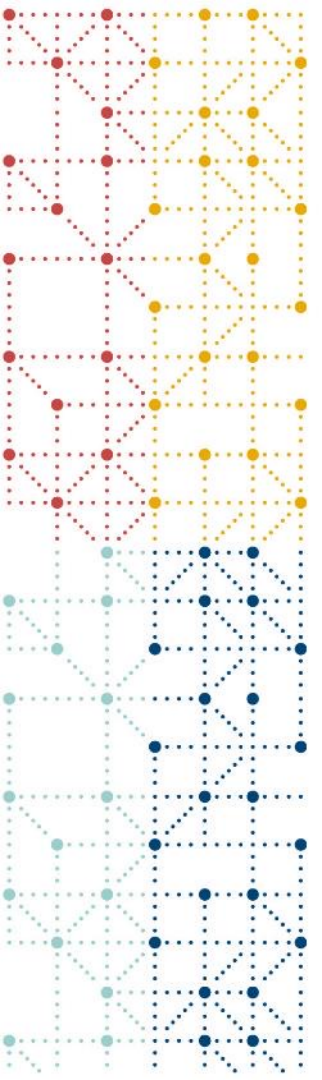
# Analysis Results and Associated Metadata Example

Identifiers		Analysis Group			Result Variable			Results Statistic		
Name	Title	Dataset	Variable	Value	Variable	Value	Label	Value	Name	Label
Table 2	Baseline Demographics and Clinical Characteristics, Safety Population	ADSL	TR01X	Drug Name Dosage X	SEX	M	Male	53	Count	n
Table 2	Baseline Demographics and Clinical Characteristics, Safety Population	ADSL	TR01X	Drug Name Dosage X	SEX	M	Male	61.6	Percent	%
Table 2	Baseline Demographics and Clinical Characteristics, Safety Population	ADSL	TR01X	Drug Name Dosage X	SEX	F	Female	33	Count	n
Table 2	Baseline Demographics and Clinical Characteristics, Safety Population	ADSL	TR01X	Drug Name Dosage X	SEX	F	Female	38.4	Percent	%



# Moving Towards a Logical Model

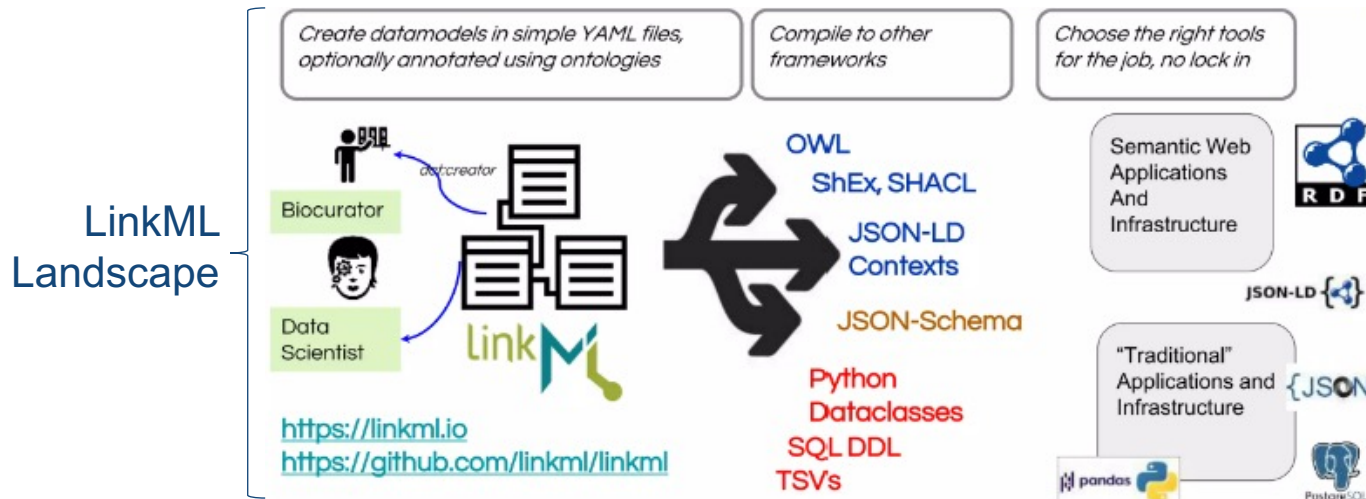
- Logical model that describes analysis results and associated metadata which will support:
  - Analysis Results Metadata Technical Specification (ARM-TS), to support automation, traceability, and creation of data displays
  - Define an Analysis Results Data (ARD) structure, to support reuse, reproducibility, and traceability of results data
- Model definition and documentation
- Illustrate and exercise with a common safety displays
  - Vital signs
  - Demographics
  - Adverse Events



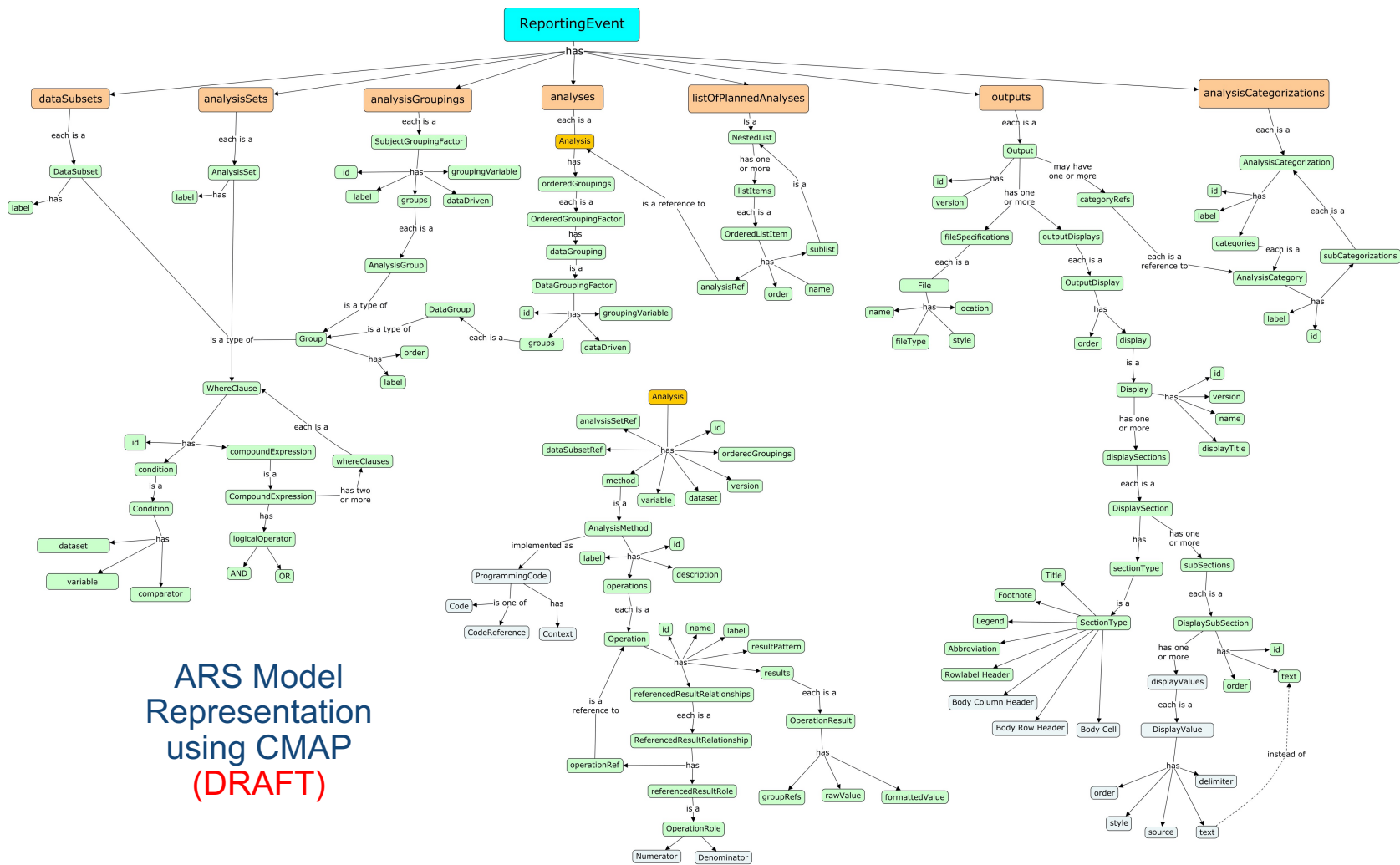
## Q & A

# Using LinkML to Create Analysis Results Model

- LinkML is a general-purpose modeling language that can be used with linked data, JSON, and other formalisms



Reference: <https://www.slideshare.net/cmungall/linkml-intro-july-2022pptx>



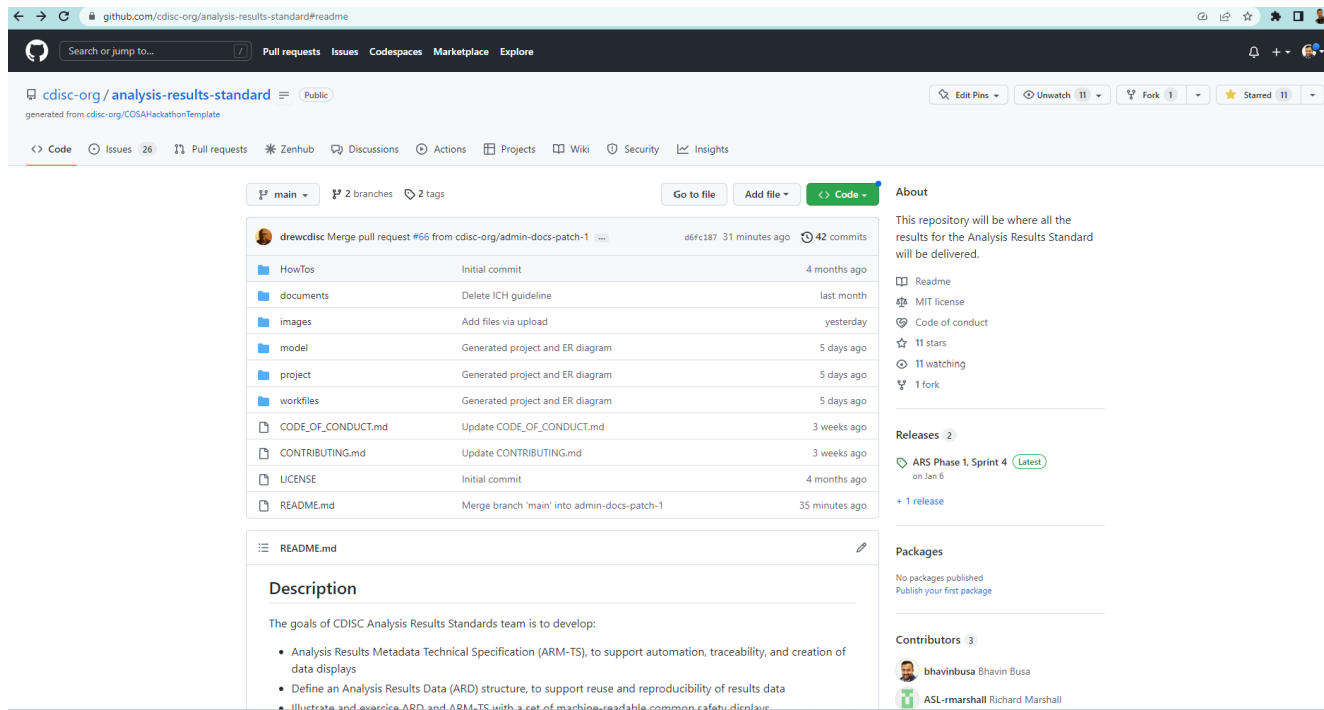
ARS Model  
Representation  
using CMAP  
(DRAFT)





# Analysis Results Standard Repo on GitHub

- <https://github.com/cdisc-org/analysis-results-standard>



github.com/cdisc-org/analysis-results-standard#readme

Search or jump to... Pull requests Issues Codespaces Marketplace Explore

cdisc-org / analysis-results-standard Public

generated from cdisc-org/COSAHackathonTemplate

Edit Pinned Unwatch 11 Fork 1 Starred 11

Code Issues 26 Pull requests Zenhub Discussions Actions Projects Wiki Security Insights

main 2 branches 2 tags Go to file Add file Code

drewcdisc Merge pull request #66 from cdisc-org/admin-docs-patch-1 06fc187 31 minutes ago 42 commits

HowTos	Initial commit	4 months ago
documents	Delete ICH guideline	last month
images	Add files via upload	yesterday
model	Generated project and ER diagram	5 days ago
project	Generated project and ER diagram	5 days ago
workfiles	Generated project and ER diagram	5 days ago
CODE_OF_CONDUCT.md	Update CODE_OF_CONDUCT.md	3 weeks ago
CONTRIBUTING.md	Update CONTRIBUTING.md	3 weeks ago
LICENSE	Initial commit	4 months ago
README.md	Merge branch 'main' into admin-docs-patch-1	35 minutes ago

README.md

### Description

The goals of CDISC Analysis Results Standards team is to develop:

- Analysis Results Metadata Technical Specification (ARM-TS), to support automation, traceability, and creation of data displays
- Define an Analysis Results Data (ARD) structure, to support reuse and reproducibility of results data
- Illustrate and exercise ARD and ARM-TS with a set of machine-readable common safety data files

This repository will be where all the results for the Analysis Results Standard will be delivered.

Readme MIT license Code of conduct 11 stars 11 watching 1 fork

Releases 2

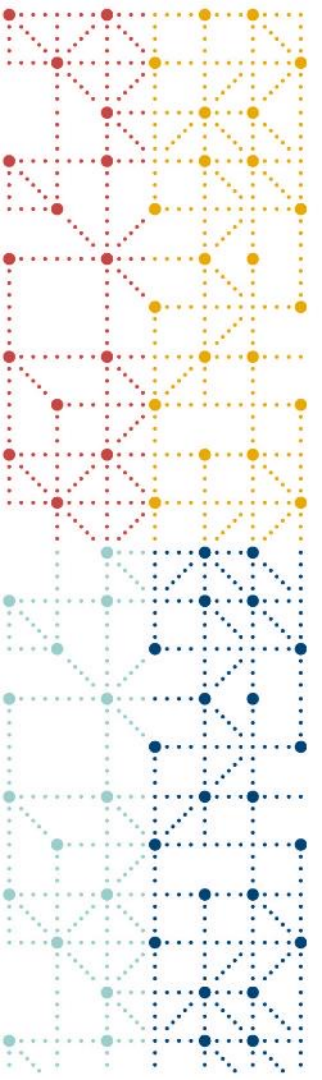
ARS Phase 1, Sprint 4 (Latest) on Jan 6 + 1 release

Packages

No packages published Publish your first package

Contributors 3

bhavinbusa Bhavin Busa ASL-rmarshall Richard Marshall



## Q & A

# Review Examples

## Summary of Demographics

Study - CDISC 360		Table 14.1.1 Summary of Demographics Safety Population			Page x of y
Characteristics	Placebo (N=XX)	Xanomeline Low Dose (N=XX)	Xanomeline High Dose (N=XX)		
Age (years)					
n	XX	XX	XX		
Mean (SD)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)		
Median	XX.X	XX.X	XX.X		
Q1, Q3	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X		
Min, Max	XX, XX	XX, XX	XX, XX		
Age Group, n (%)					
< 65 years	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
≥ 65 years	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
Gender, n (%)					
Male	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
Female	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
Ethnicity, n (%)					
Hispanic or Latino	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
Not Hispanic or Latino	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
Source dataset: adsl, Generated on: DDMONYYYY:HH:MM Program: <pid>.sas, Output: <pid><oid>.rtf, Generated on: DDMONYYYY:HH:MM					

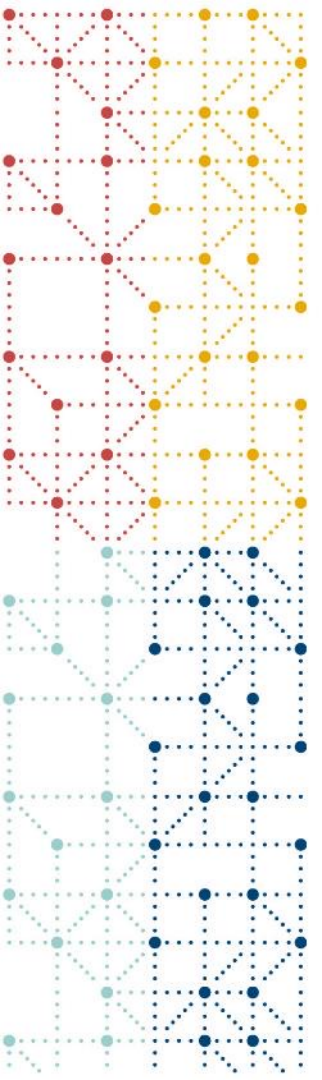
## Summary of TEAE by SOC and PT

Study - CDISC 360		Table 14.3.1.1 Summary of TEAE by System Organ Class and Preferred Term Safety Population			Page x of y
System Organ Class Preferred Term [a], n (%)	Placebo (N=XX)	Xanomeline Low Dose (N=XX)	Xanomeline High Dose (N=XX)		
Number of subjects with at least one event	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
<SOC 1>	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
<Preferred Term 1>	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
...	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
<Preferred Term n>	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
<SOC 2>	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
<Preferred Term 1>	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
...	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
<Preferred Term n>	XX ( XX.X)	XX ( XX.X)	XX ( XX.X)		
Notes: TEAE=Treatment-Emergent Adverse Events. Subjects are counted once within each system organ class and preferred term. [a] All investigators adverse events were coded using MedDRA version xx.x.					
Source dataset: adae, Generated on: DDMONYYYY:HH:MM Program: <pid>.sas, Output: <pid><oid>.rtf, Generated on: DDMONYYYY:HH:MM					

<https://github.com/cdisc-org/analysis-results-standard/tree/main/workfiles/examples/PHUSE%20Connect%20Workshop>



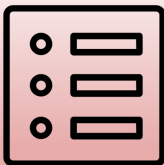
## **ARS model will drive automation and open-source tool development**



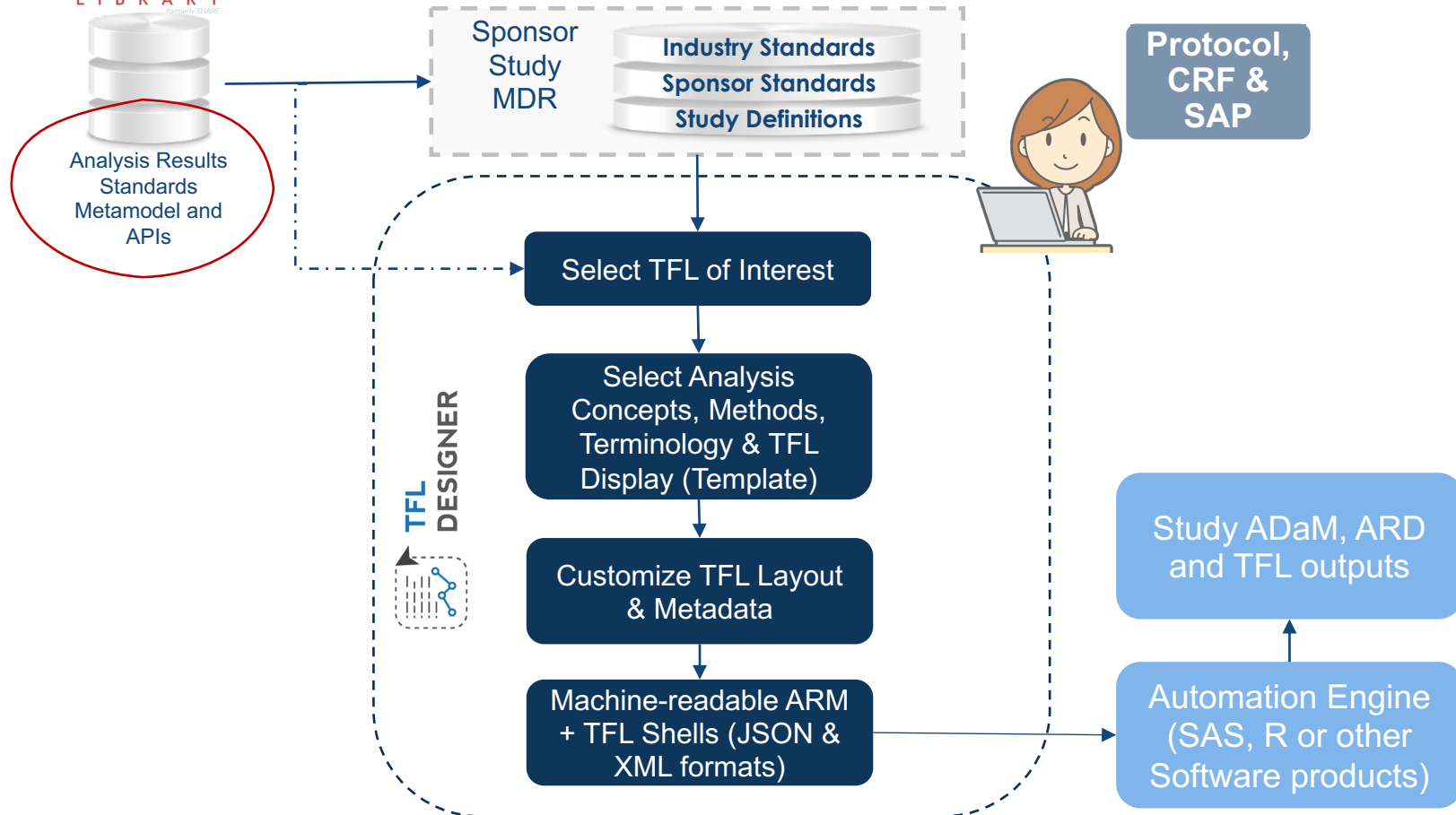
## Reference Implementation Example



# TFL Designer



- An industry leading solution to automate tables, figures, and listings (TFL) design and generation process
- CDISC COSA approved
- Align with CDISC ARS logical model
- Community beta version release (access limited to beta users): Apr 2023

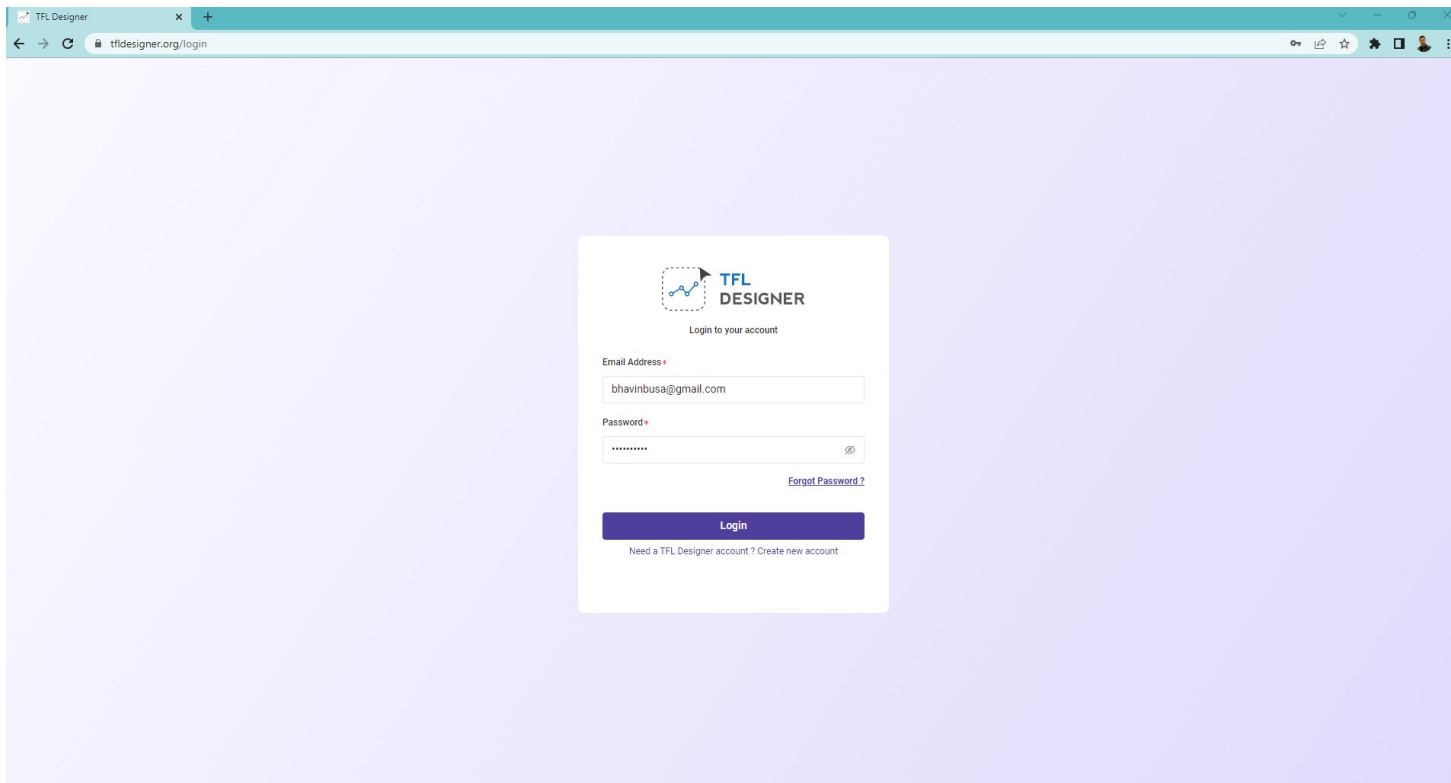


\* Open-Source TFL Designer, Bhavin Busa

Link to GitHub: <https://github.com/bhavinbusa/tfldesigner>



# TFL Designer\* (MVP Demo)





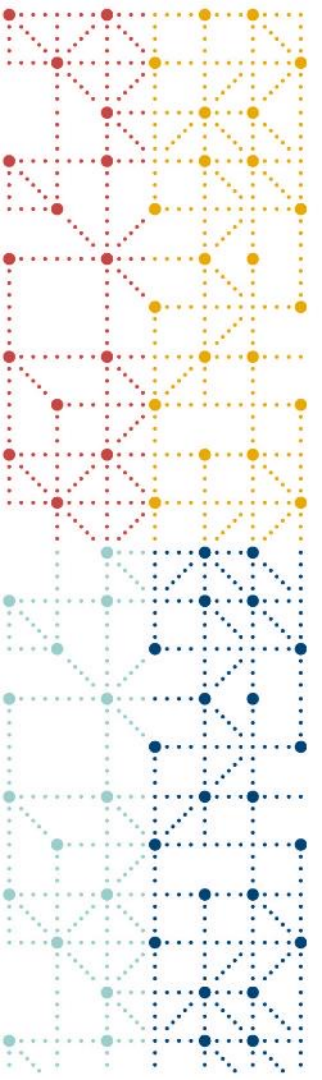
# ARS Roadmap

## MVP for v1.0 (Summer 2023)

- Logical Model of to support ARM TS/ARD
- Four common safety examples based on team developed tables
  - Demographics
  - Adverse Events
  - Vital signs

## Future Development

- Expanded use cases
- Machine readable TFLs available on the CDISC website
- Conformance rules
- Terminology



## Q & A



# Contact Details

**Bhavin Busa**

ARS Product Owner & Co-Lead

[bhavinbusa@gmail.com](mailto:bhavinbusa@gmail.com)

**Bess LeRoy**

ARS Co-Lead

[bleroy@cdisc.org](mailto:bleroy@cdisc.org)

CDISC ARS GitHub Repo:

<https://github.com/cdisc-org/analysis-results-standard>

