siera: Redefining TFL Automation

Anna Yaggi and Shivani Gupta

October 15th, 2025

Author Bios

Anna Yaggi

Product Manager

Malan Bosman

Lead Automation Engineer

Shivani Gupta

Director of Statistical Programming

CDISC Analysis Results Standard - a Logical Data Model

cdisc
Clear data. Clear impact

Analysis Results Standard (ARS) v1.0

cdisc ARS ANALYSIS RESULTS STANDARD

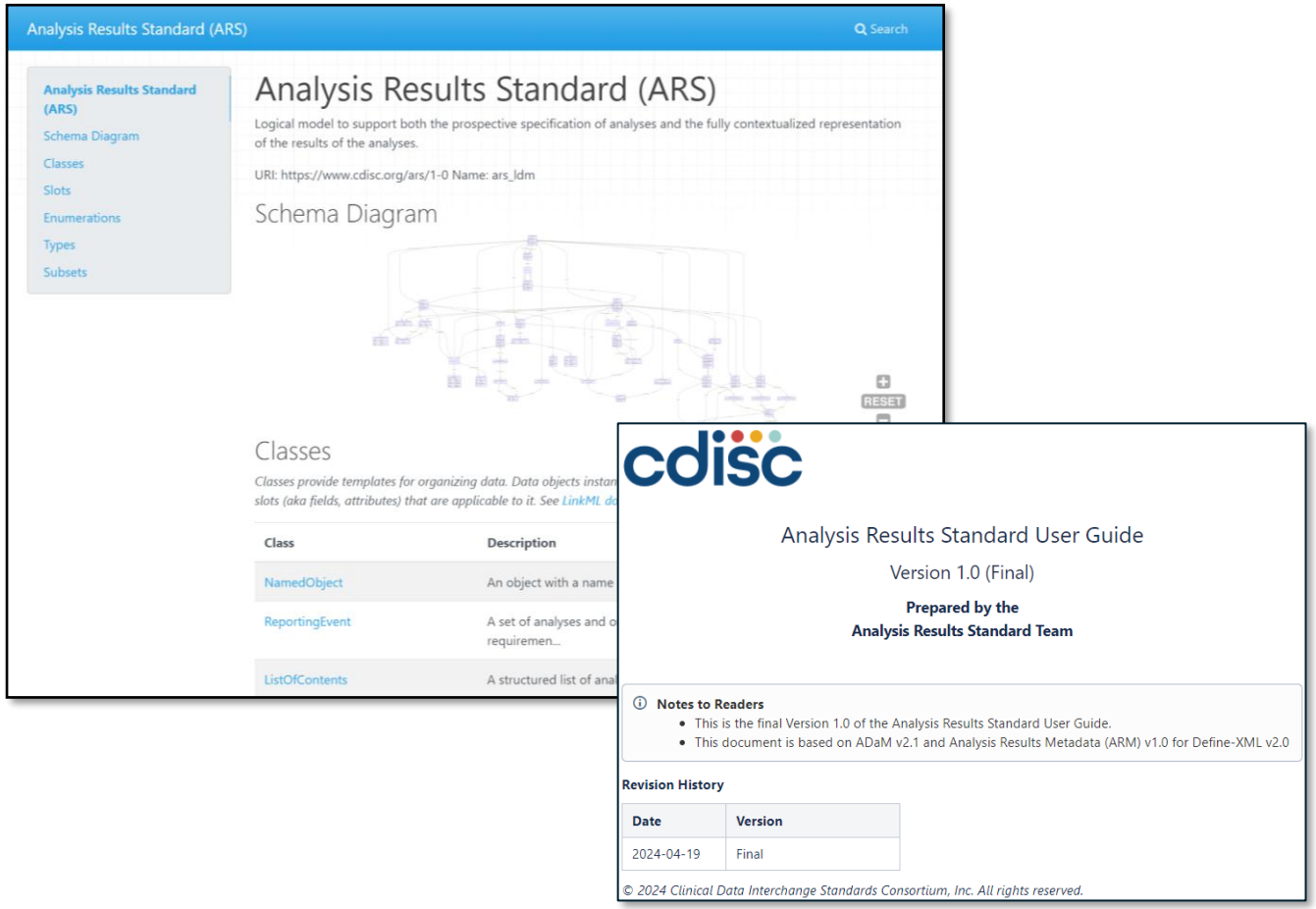
Large trials generate many analysis results in the form of tables, figures, and written reports, yet these results are rarely output in a form that is machine-readable. Previously, there has been no standard way of describing and organizing these results, making it difficult to automate their generation, make them reproducible, trace their origin, or enable them to be reused in other outputs.

To address these inefficiencies, CDISC has developed the Analysis Results Standard (ARS), which aim to facilitate automation, reproducibility, reusability, and traceability of analysis results data.

Features of ARS v1.0

- A Logical Data Model that describes analysis results and associated metadata.
- A User Guide to illustrate and exercise the model with common safety displays.

https://cdisc-org.github.io/analysis-results-standard/



https://wiki.cdisc.org/display/ARSP/Analysis+Results+Standard+User+Guide+v1.0

Model Components

Analysis Set | Data Subset | Analysis Grouping | Analysis Method | Analysis | Results

Summary of Demographics

Study – CDISC 360 | Page x of y

Table 14.1.1
Summary of Demographics
Safety Population

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)

Output

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 | Page x of y

Table 14.3.1.1
Summary of TEAE by System Organ Class and Preferred Term
Safety Population

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

Source: CDISC

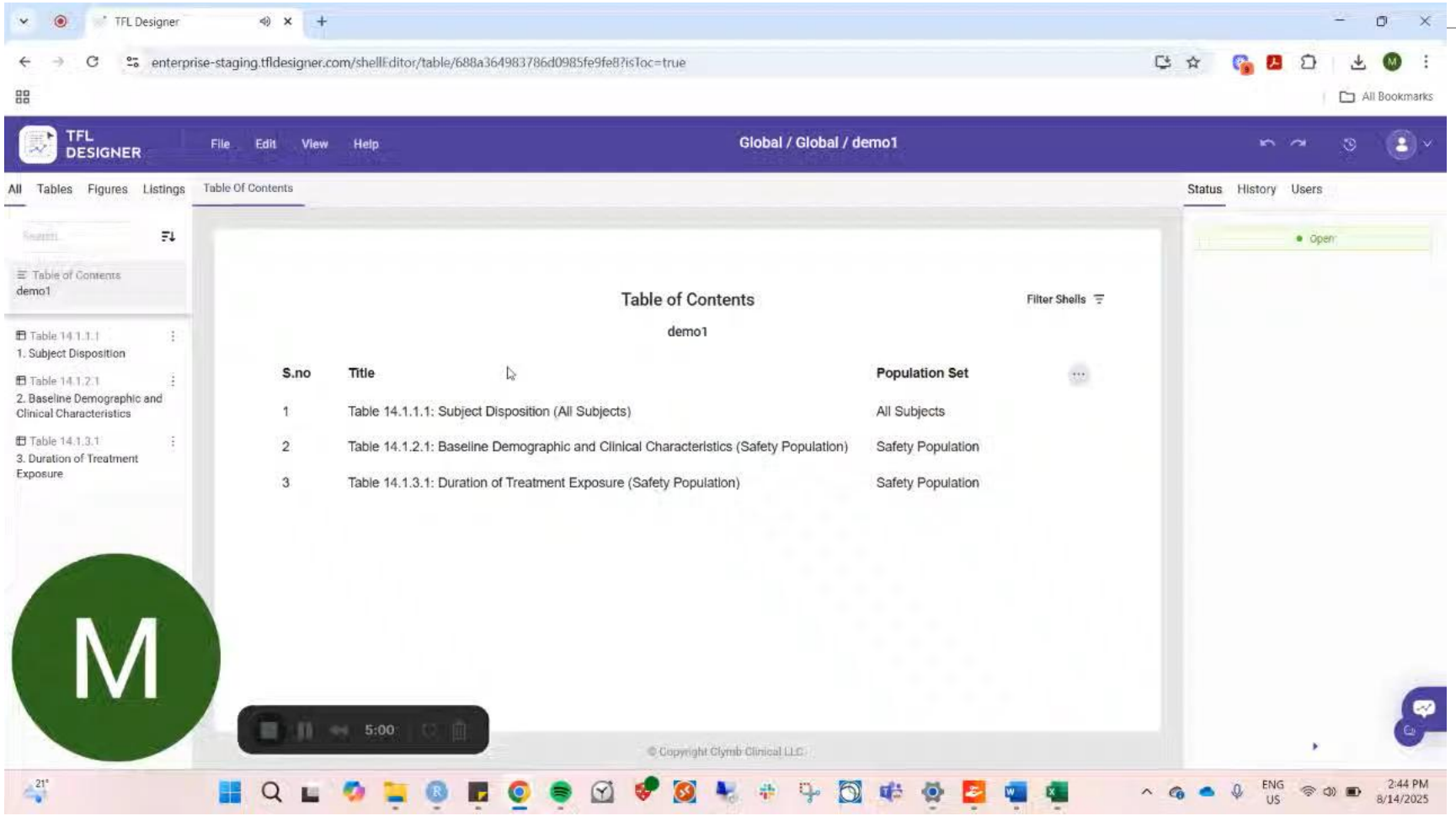
Accessing CDISC ARS

- Download via TFL Designer Community Version (tfldesigner.org)
  - You can build shells and download their full ARS metadata using the community edition of TFL Designer

- Machine-Readable JSON Format
  - CDISC ARS metadata is provided in a structured, machine-readable JSON format.
  - This allows for efficient integration and downstream processing.
- Easily Visualized in Excel

  - The JSON can be converted into an Excel workbook for human-friendly inspection.
  - Each ARS class appears as a separate tab in the spreadsheet.

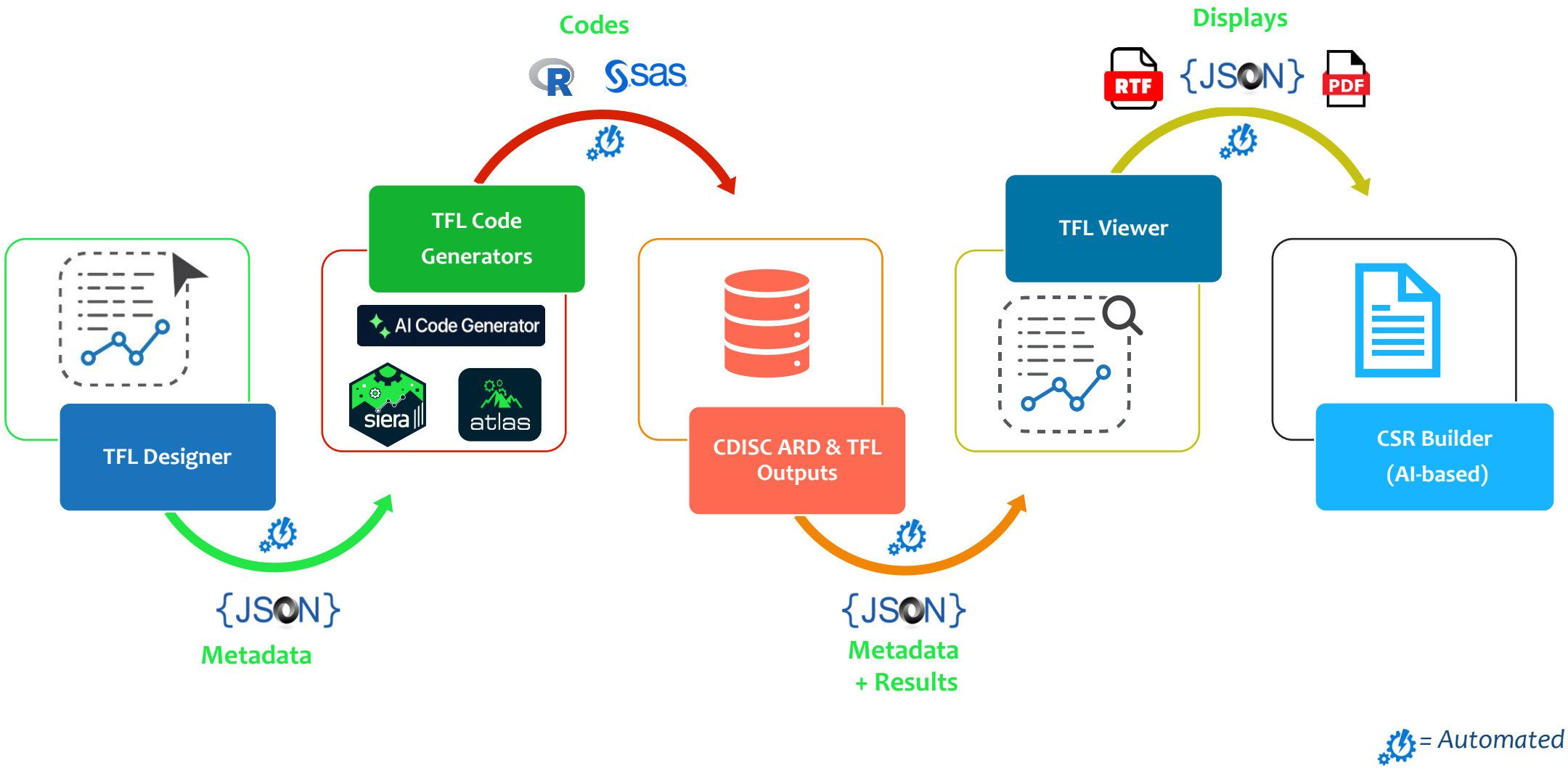
Extracting Metadata from TFL Designer



TFL DESIGNER

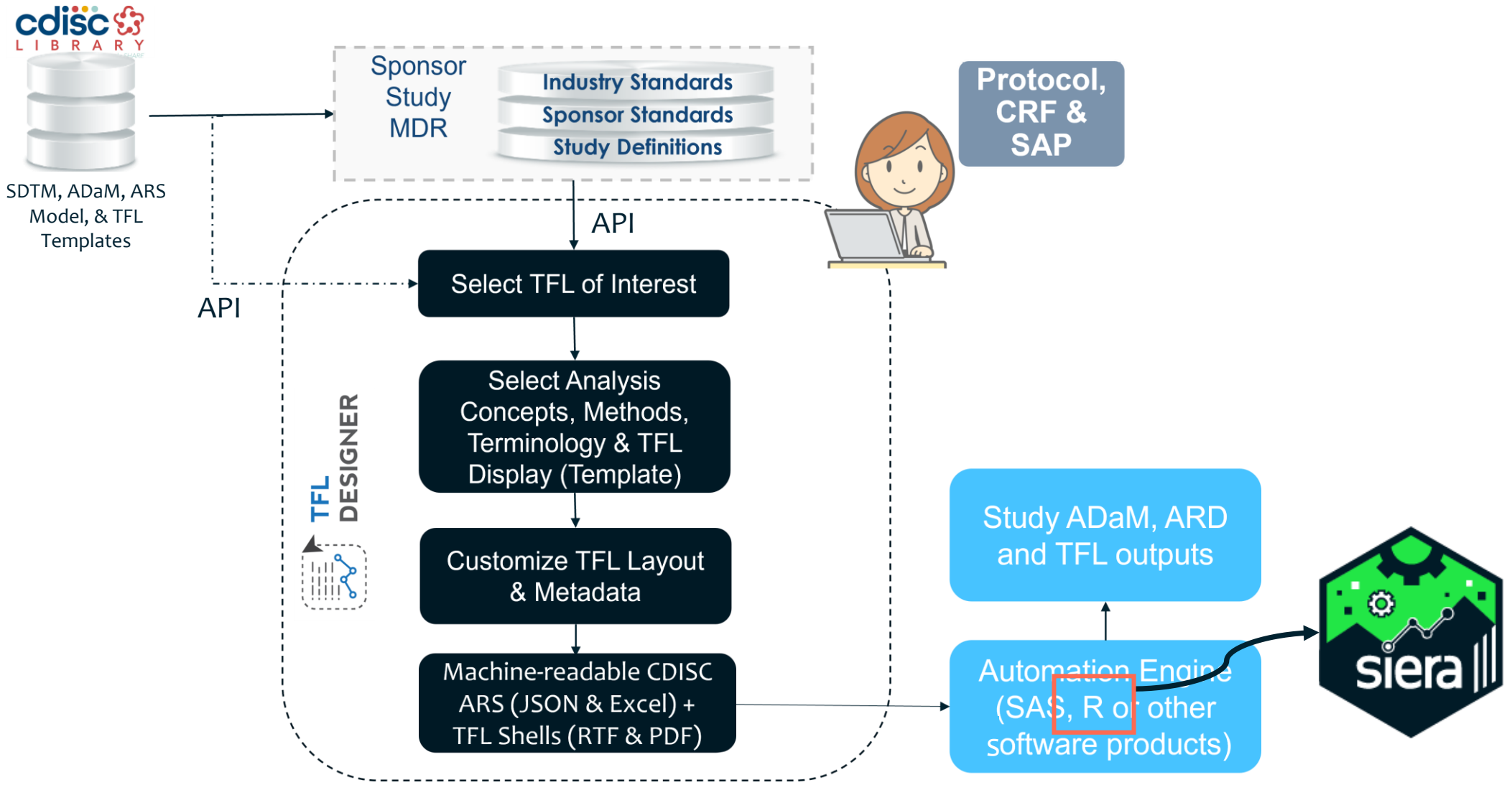
Video link: https://www.loom.com/share/70e39fa1d1824b3c91b2395236e47e55?sid=e8584776-b1a8-4446-a64a-df7d240205ea

Clymb's TFL Automation Solutions & Workflow



TFL Code Generator

Metadata-driven TFL Deliverable Workflow



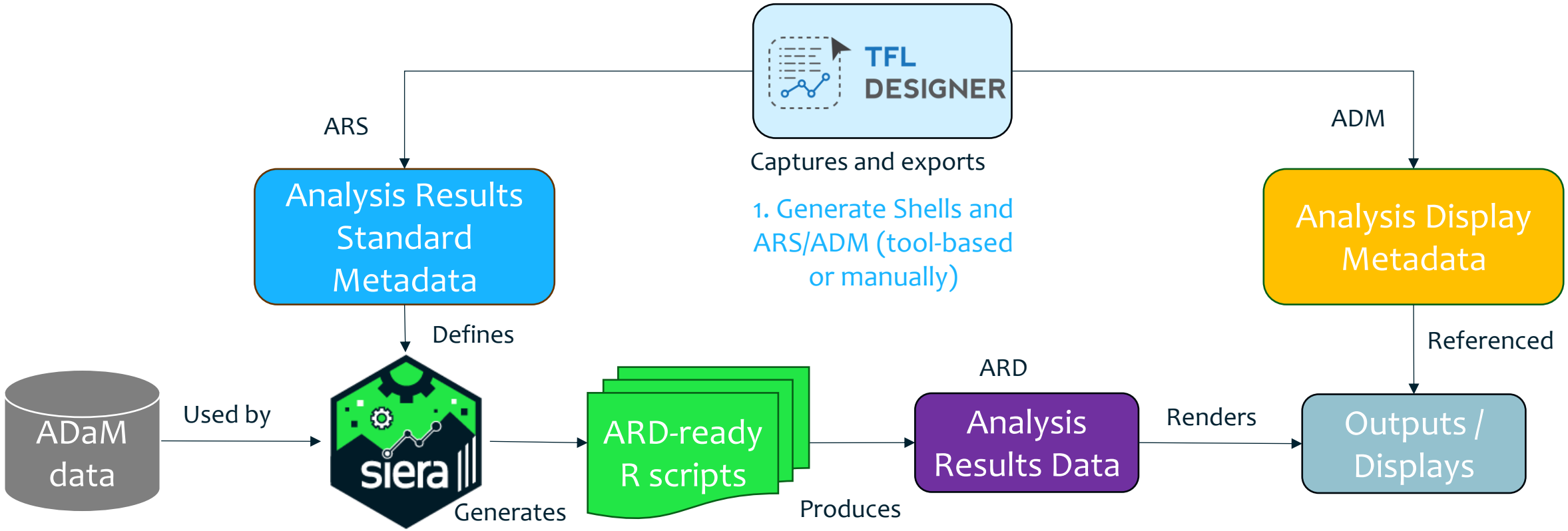
What is siera?

siera* is an open-source **R package** that **ingests ARS metadata** for a reporting event and converts the metadata to **produce executable R scripts**. These R scripts can be run as-is (combined with the relevant ADaM dataset) to **produce each output's ARD**.

https://github.com/clymbclinical/siera

* on CRAN and approved by COSA!

siera Workflow [Design to Execution]



siera Workflow [ARS Metadata to ARD]

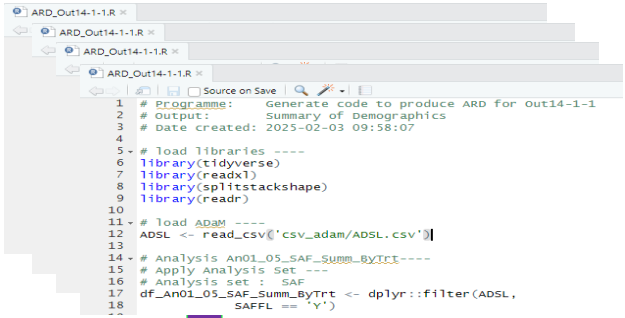
ARS Metadata

siera

Auto-generated R scripts

ADaM Dataset

USUBJID	ARM	AGE	AGEGR1	AGEU	RACE	SEX
01-701-1015	Placebo	63	<65	YEARS	WHITE	F
01-701-1023	Placebo	64	<65	YEARS	WHITE	M
01-701-1028	Xanomeline High Dose	71	65+	YEARS	WHITE	M
01-701-1033	Xanomeline Low Dose	74	65+	YEARS	WHITE	M
01-701-1034	Xanomeline High Dose	77	65+	YEARS	WHITE	F
01-701-1047	Placebo	85	65+	YEARS	WHITE	F

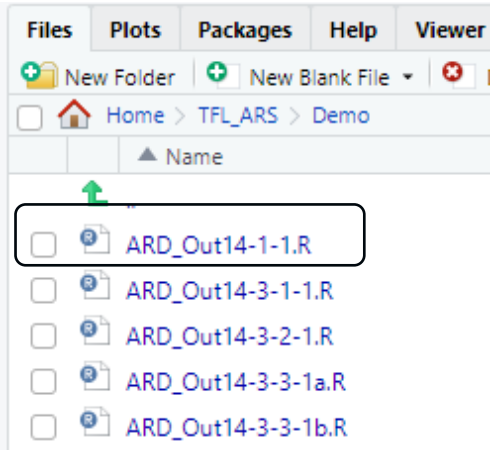


```
# Programme:    Generate code to produce ARD for Out14-1-1
# Output:       Summary of Demographics
# Date created: 2025-02-03 09:58:07

# load libraries ----
library(tidyverse)
library(readxl)
library(splitstackshape)
library(readr)

# load ADaM ----
ADSL <- read_csv('csv_adam/ADSL.csv')

# Analysis An01_05_SAF_Summ_ByTrt----
# Apply Analysis Set ---
# Analysis set :  SAF
df_An01_05_SAF_Summ_ByTrt <- dplyr::filter(ADSL,
                                           SAFFL == 'Y')
```

Analysis Results Datasets

id	operation_id	resultGroup1_groupingId	resultGroup1_groupId	resultGroup2_groupingId	resultGroup2_groupId	rawValu	formattedVal
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_1_n	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_1	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_1	14	14
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_1_n	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_1	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_2	72	72
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_1_n	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_2	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_1	8	8
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_1_n	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_2	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_2	76	76
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_1_n	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_3	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_1	11	11
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_1_n	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_3	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_2	73	73
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_2_pct	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_1	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_1	16.27907	(16.3)
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_2_pct	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_1	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_2	83.72093	(83.7)
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_2_pct	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_2	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_1	9.52381	(9.5)
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_2_pct	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_2	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_2	90.47619	(90.5)
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_2_pct	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_3	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_1	13.09524	(13.1)
An03.02_AgeGrp_ByTrt	Mth01_CatVar_ByGrp_2_pct	AnlsGrouping_02_Trt	AnlsGrouping_02_Trt_3	AnlsGrouping_03_AgeGp	AnlsGrouping_03_AgeGp_2	86.90476	(86.9)

siera generated ARD-ready R scripts

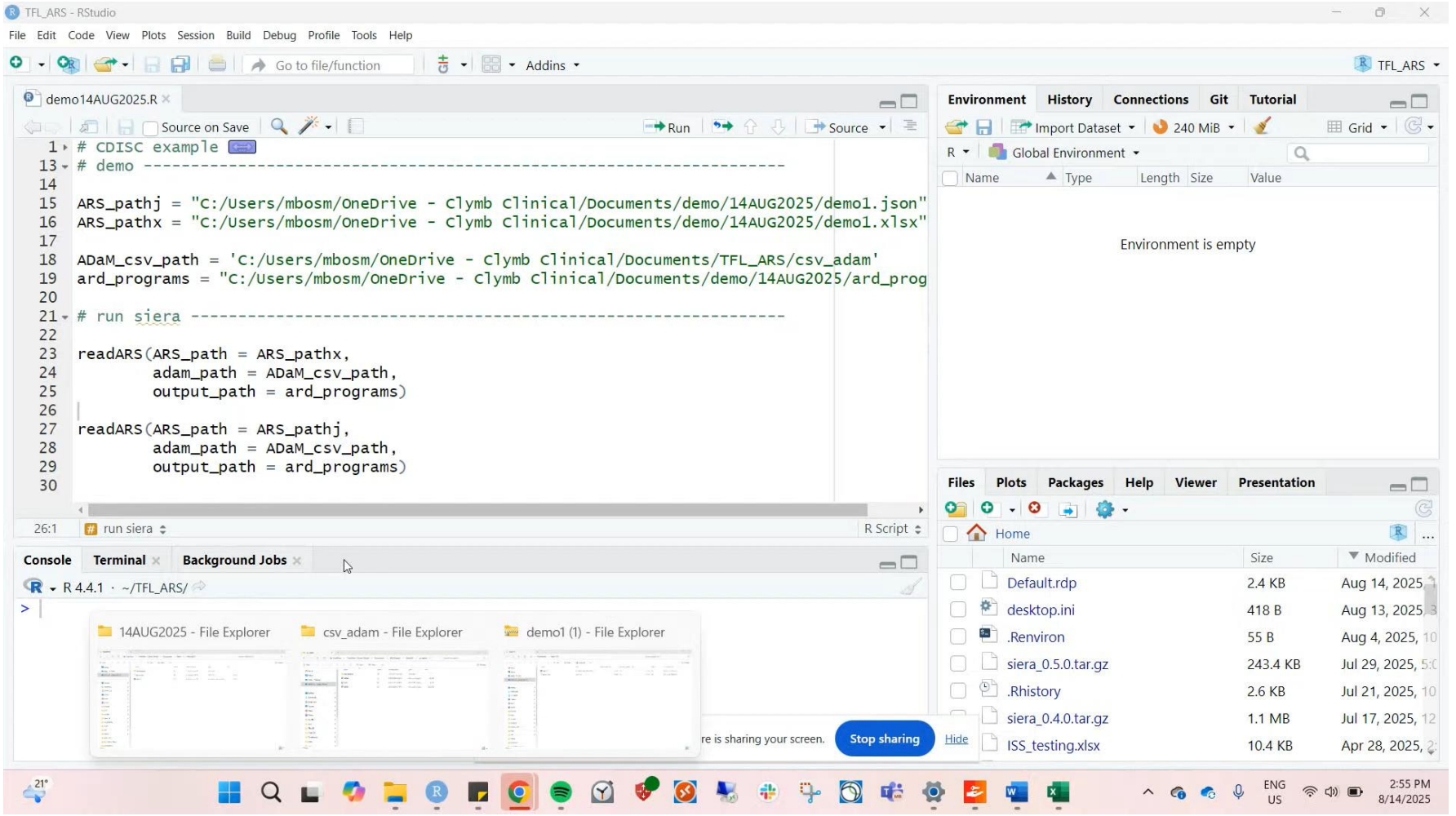


ARD_Out14-1-1.R

Source on Save

```r
# Programme:    Generate code to produce ARD for Out14-1-1
# Output:       Summary of Demographics
# Date created: 2025-02-03 09:58:07

# load libraries ----
library(tidyverse)
library(readxl)
library(splitstackshape)
library(readr)

# load ADaM ----
ADSL <- read_csv('csv_adam/ADSL.csv')

# Analysis An01_05_SAF_Summ_ByTrt----
# Apply Analysis Set ---
# Analysis set :  SAF
df_An01_05_SAF_Summ_ByTrt <- dplyr::filter(ADSL,
            SAFFL == 'Y')



#Apply Analysis Grouping ---
df1_An01_05_SAF_Summ_ByTrt <- df_An01_05_SAF_Summ_ByTrt %>%
          dplyr::group_by(TRT01A)
```

siera Demo Video



Video link: https://www.loom.com/share/4aed3442c335446f83d4ba4caa59faf3?sid=b15d1fa5-bb9f-4857-a754-647de14b6bdb

Impact, Learnings and Next Steps

Impact

- ↑ Metadata setup
- ↓ Manual programming

- ↑ Time spent on complex tasks
- ↓ Time spent on repetitive tasks

- ↑ Quality and efficiency
- ↓ Overall Cost

Learnings

- Need of a design tool to generate both ARM and ADM metadata
- Integrating ARD into the process requires change
- Avoid “black box” solution
- Provide flexibility to the user

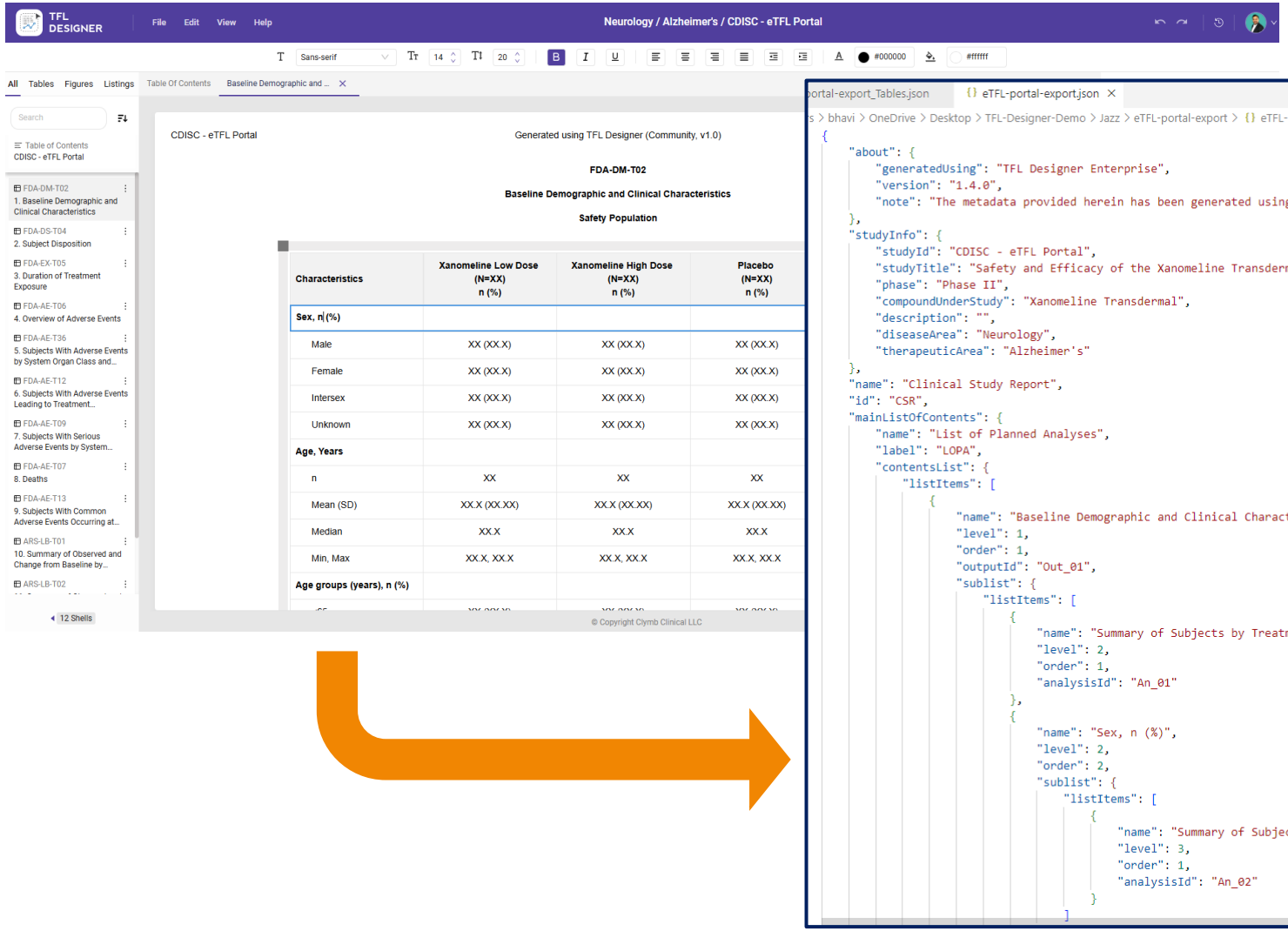
Next Steps

- Integration into workflows

- More feedback, more updates

AI Code Generator

TFL Shells → CDISC ARS Metadata (JSON) from TFL Designer



AI Code Generator for TFL Automation

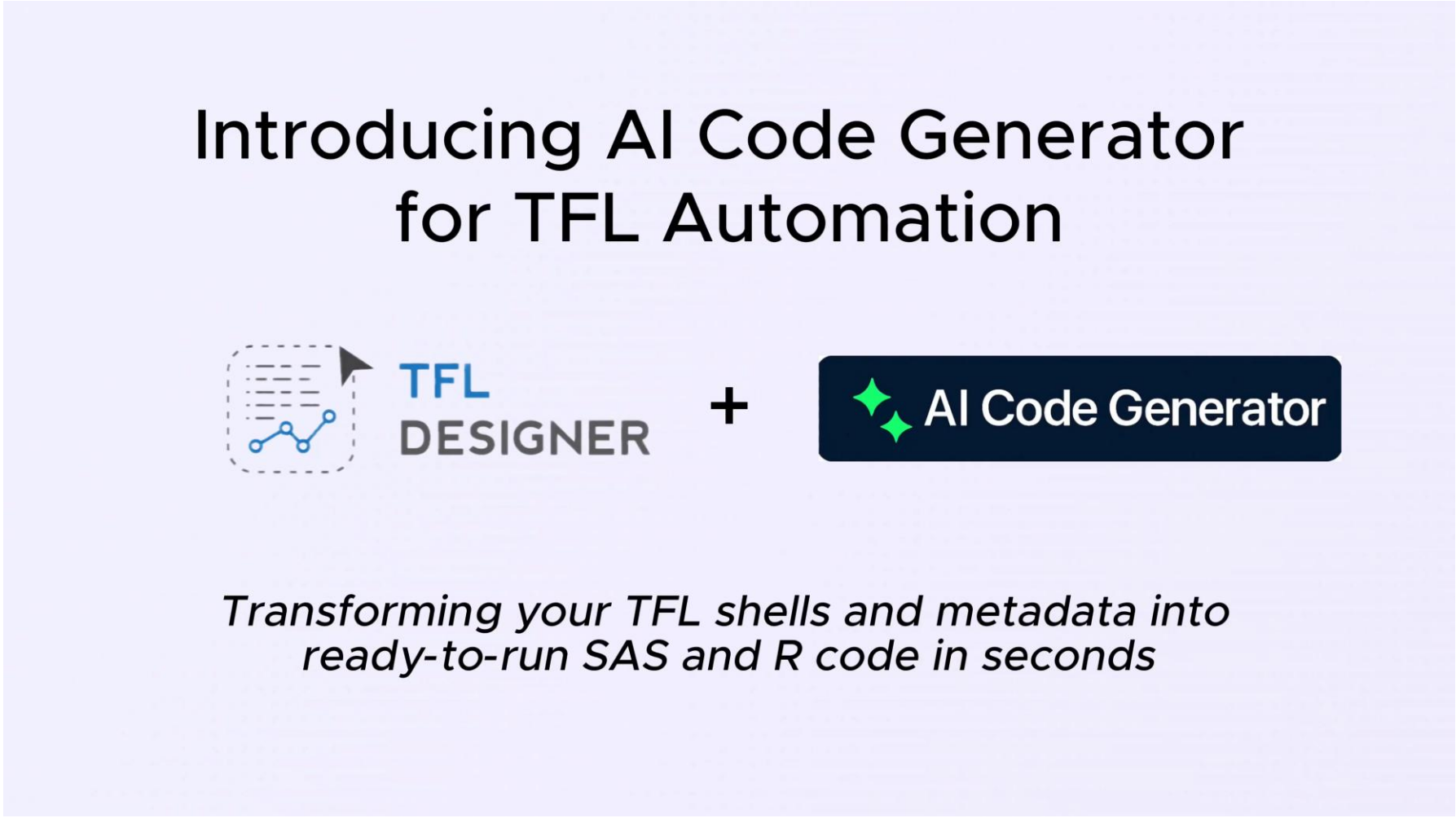
TFL DESIGNER

+

AI Code Generator

Transforming TFL shells and metadata
into ready-to-use SAS and R code

AI Code Generator Demo Video



AI Code Generator

Video link: https://www.loom.com/share/90c1fa1929b54c4b83e1fdf1e578dc04?sid=92589371-8b75-4242-b0e2-bc7c74e3aab5

Q&A

Anna Yaggi
Product Manager
ayaggi@clymbclinical.com

Shivani Gupta
Director of Statistical Programming
sgupta@clymbclinical.com

Malan Bosman
Lead Automation Engineer
mbosman@clymbclinical.com

EMAIL
info@clymbclinical.com

WEBSITE
www.clymbclinical.com

PHONE
781-692-3613