



AI Code Generator: Automating TFL Programming with ARS Metadata

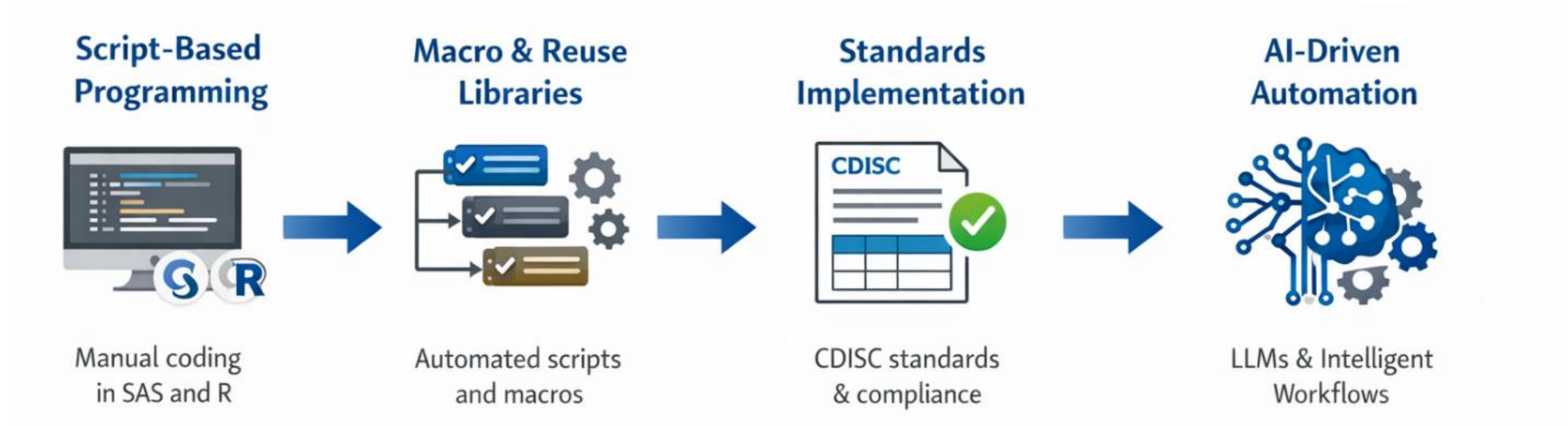
Paper & Presentation # SDo4

PHUSE US Connect 2026 @ Austin, Texas

Navin Dedhia, Head of Software Engineering, Clymb Clinical

Bhavin Busa, Principal & Co-Founder, Clymb Clinical

Evolution of Statistical Programming



**AI is the new
superpower, not
the replacement.**



Key Pillars of the AI-Enabled Statistical Programming Framework

From Code-Centric
to Model-Centric
Programming

CDISC as the Enabler,
Not the Constraint

Where LLMs Fit,
and Where They
Do Not

A Practical
Architecture for
SAS, R, and AI
Co-Existence

Organizational
and Skill Challenges

The Evolving Role
of the Statistical
Programmer

Code-Centric vs. Model-Centric Statistical Programming

Code-Centric Programming



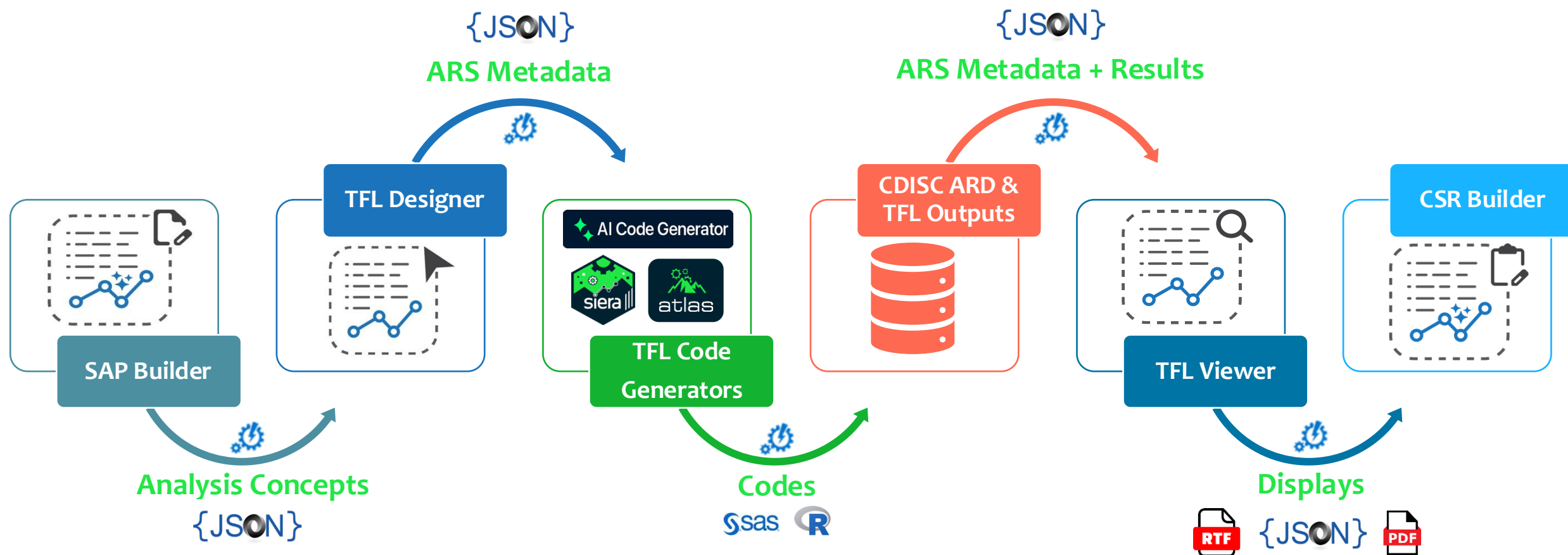
- Extensive manual coding
- Script-heavy, error-prone
- Difficult to scale
- Limited traceability & auditability

Model-Centric Programming



- Metadata-driven automation
- Metadata-based, reusable models
- Scalable & consistent
- Structured traceability & audit trail

Clymbr Hub - Automation and AI Solutions from SAP to CSR



Short demo videos: <https://clymb-tfl-automation.carrd.co/>

⚙️ = Automated Process; ARS = Analysis Results Standard; ARD = Analysis Results Dataset



AI Code Generator

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

The screenshot displays the TFL Designer interface, which is used for generating clinical trial reports. The interface is divided into two main sections: a report preview on the left and a JSON export editor on the right.

Report Preview (Left):

- Header:** TFL Designer, File, Edit, View, Help.
- Table of Contents:**
 - 1. Baseline Demographic and Clinical Characteristics
 - 2. Subject Disposition
 - 3. Duration of Treatment Exposure
 - 4. Overview of Adverse Events
 - 5. Subjects With Adverse Events by System Organ Class and...
 - 6. Subjects With Adverse Events Leading to Treatment...
 - 7. Subjects With Serious Adverse Events by System...
 - 8. Deaths
 - 9. Subjects With Common Adverse Events Occurring at...
 - 10. Summary of Observed and Change from Baseline by...
- Report Content:**
 - CDISC - eTFL Portal**
 - Generated using TFL Designer (Community, v1.0)**
 - FDA-DM-T02**
 - Baseline Demographic and Clinical Characteristics**
 - Safety Population**
 - Table:**

Characteristics	Xanomeline Low Dose (N=XX) n (%)	Xanomeline High Dose (N=XX) n (%)	Placebo (N=XX) n (%)
Sex, n (%)			
Male	XX (XX.X)	XX (XX.X)	XX (XX.X)
Female	XX (XX.X)	XX (XX.X)	XX (XX.X)
Intersex	XX (XX.X)	XX (XX.X)	XX (XX.X)
Unknown	XX (XX.X)	XX (XX.X)	XX (XX.X)
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
Min, Max	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X
Age groups (years), n (%)			

JSON Export Editor (Right):

- File:** portal-export_Tables.json
- Code:**

```
{
  "about": {
    "generatedUsing": "TFL Designer Enterprise",
    "version": "1.4.0",
    "note": "The metadata provided herein has been generated using the TFL Designer Enterprise version 1.4.0. The representation of metadata fields av..."
  },
  "studyInfo": {
    "studyId": "CDISC - eTFL Portal",
    "studyTitle": "Safety and Efficacy of the Xanomeline Transdermal Therapeutic System (TTS) in Patients with Mild to Moderate Alzheimer's Disease.",
    "phase": "Phase II",
    "compoundUnderStudy": "Xanomeline Transdermal",
    "description": "",
    "diseaseArea": "Neurology",
    "therapeuticArea": "Alzheimer's"
  },
  "name": "Clinical Study Report",
  "id": "CSR",
  "mainListOfContents": {
    "name": "List of Planned Analyses",
    "label": "LOPA",
    "contentsList": {
      "listItems": [
        {
          "name": "Baseline Demographic and Clinical Characteristics",
          "level": 1,
          "order": 1,
          "outputId": "Out_01",
          "sublist": {
            "listItems": [
              {
                "name": "Summary of Subjects by Treatment",
                "level": 2,
                "order": 1,
                "analysisId": "An_01"
              },
              {
                "name": "Sex, n (%)",
                "level": 2,
                "order": 2,
                "sublist": {
                  "listItems": [
                    {
                      "name": "Summary of Subjects by Treatment and Sex, n (%)",
                      "level": 3,
                      "order": 1,
                      "analysisId": "An_02"
                    }
                  ]
                }
              }
            ]
          }
        }
      ]
    }
  }
}
```

AI Code Generator for TFL Automation



TFL
DESIGNER

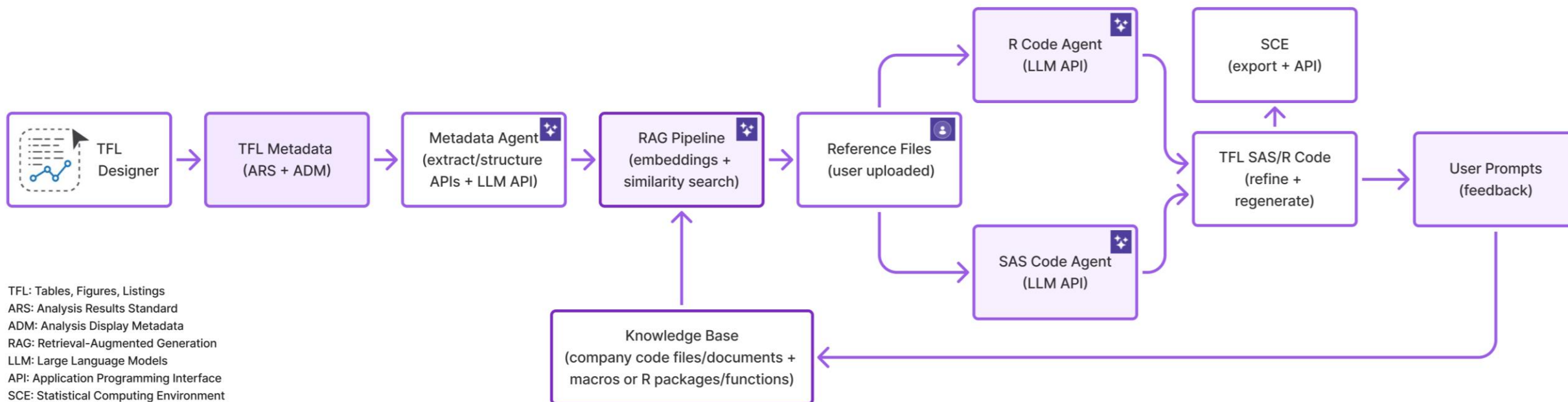
+



AI Code Generator

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

AI Code Generator Architecture



AI Code Generator Demo

TFL DESIGNER

AI Code Generator

Table 02 (Current Shell)
Baseline Demographic and Clinical Characteristics [FDA STF-IG]

SAS Code

GPT-4.1

Add Reference Codes

Regenerate

```

/*****
/* Program Name   : t02_baseline_demo_clin_char.sas          */
/* Purpose        : Generate Table for Baseline Demographic and Clinical Characteristics [FDA STF-IG] */
/* Project        : AGY-10142001                               */
/* Input Datasets : ADSL                                       */
/* Output         : &out./t02_baseline_demo_clin_char.rtf      */
/* Created By     : Anna Yaggi                                  */
/* Date Created   : 30SEP2025                                   */
/* Table ID       : Table 02                                    */
/* Table Title    : Baseline Demographic and Clinical Characteristics */
/* Shell Source   : TFL Designer Enterprise                    */
*****/

/* Change History                                     */
/* -----*/
/* Date      | Author      | Description                                     */
/* -----|-----|-----*/
/* <YYYY-MM-DD> | <Name>      | Initial version                                     */
*****/

```

Copy

Export

Refine the code with custom prompt

Show System Prompt

Clear

Keep

Refine Code

AI Code Generator



Video link: <https://www.loom.com/share/046a583a7c9841489036d8e8238e0bd8?sid=f946f789-8ab8-4c86-a766-1405ddod586a>

The AI Reckoning

Ad hoc Code

Free-Text Specs

Siloed Tools

The Reckoning



Workflow UI



Structured Specifications

Models & Metadata

SAS / R Execution

Human Review

This is the moment where we decide how AI is used.

AI Code Generator – Transforming TFL Automation

How It Works

- Built on TFL shells + JSON metadata (CDISC ARS compliant) from TFL Designer
- Metadata + reference macros → structured prompts for the LLM.
- Generates both SAS and R code (90% ready-to-run → goal is 95%).

Key Benefits

- Predictable & controlled outcomes vs. ad-hoc AI code.
- Handles complexity: safety & efficacy outputs, inferential & comparative stats.
- Efficient debugging & maintenance through integration with existing macro libraries.
- Configurable per Sponsor programming conventions (headers, setup file, macros, proc report logic).
- Learns code styling & approach over time from your standard library.
- Codes can be used for production, validation, or QC purposes.

Impact

- Speeds up statistical programming deliverables.
- Frees programmers for higher-value tasks.
- Brings automation + intelligence into statistical programming.

Thank you!





Navin Dedhia

Head of Software Engineering

ndedhia@clymbclinical.com

Bhavin Busa

Principal & Co-Founder

bhavin@clymbclinical.com



EMAIL

info@clymbclinical.com



WEBSITE

www.clymbclinical.com



PHONE

781-692-3613