



cdisc® 2026 Europe Interchange

THE FUTURE IS CONNECTED: STANDARDS AND AI POWERING DIGITAL TRANSFORMATION



MILAN, ITALY | MAIN CONFERENCE: 20-21 MAY | TRAININGS & WORKSHOPS: 18, 19, & 22 MAY

An R Shiny–Based Framework for Executing CDISC Open Rules and Custom Validation of CDISC Datasets

Reshma Radhakrishnan, Principal SAS Programmer, Efficacy Lifescience Analytics
Dineshkumar Subbu, Tech Lead – SAS/R, Efficacy Lifescience Analytics

Speaker



Reshma Radhakrishnan

Principal Statistical Programmer - Efficacy Lifescience Analytics

Reshma Radhakrishnan is a Principal Statistical Programmer with over 16 years of experience in CDISC standards and regulatory submissions.

Co-Author



Dineshkumar Subbu

Tech Lead - Statistical Programmer -Efficacy Lifescience Analytics

Dineshkumar Subbu is a Lead Statistical programmer with 8+ years of experience in R and Shiny. He specializes in developing web applications using R and Shiny, providing robust DevOps support.



Agenda

1. Open CDISC Rules (CORE)
2. Proposed R Shiny Validation Framework
3. Simple vs. Complex Custom Rules
4. Report Consolidation & Visualization
5. Conclusion



CDISC Open Rules

Introduction



CORE is developed and maintained by CDISC.



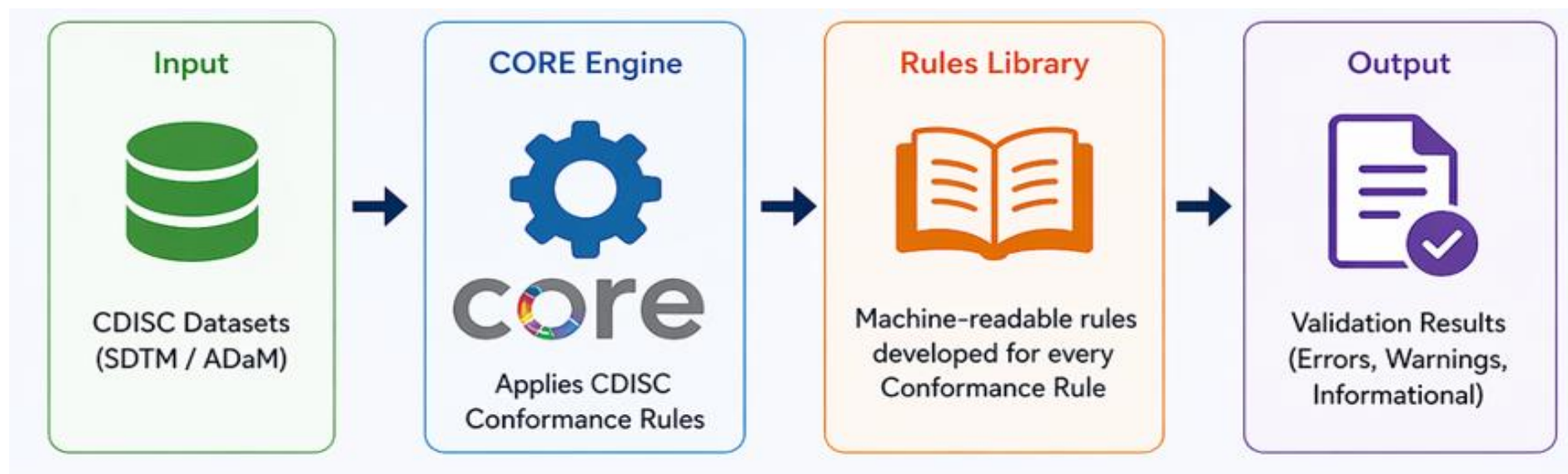
Machine-readable rules



Improves efficiency and data quality



Supports regulatory submissions



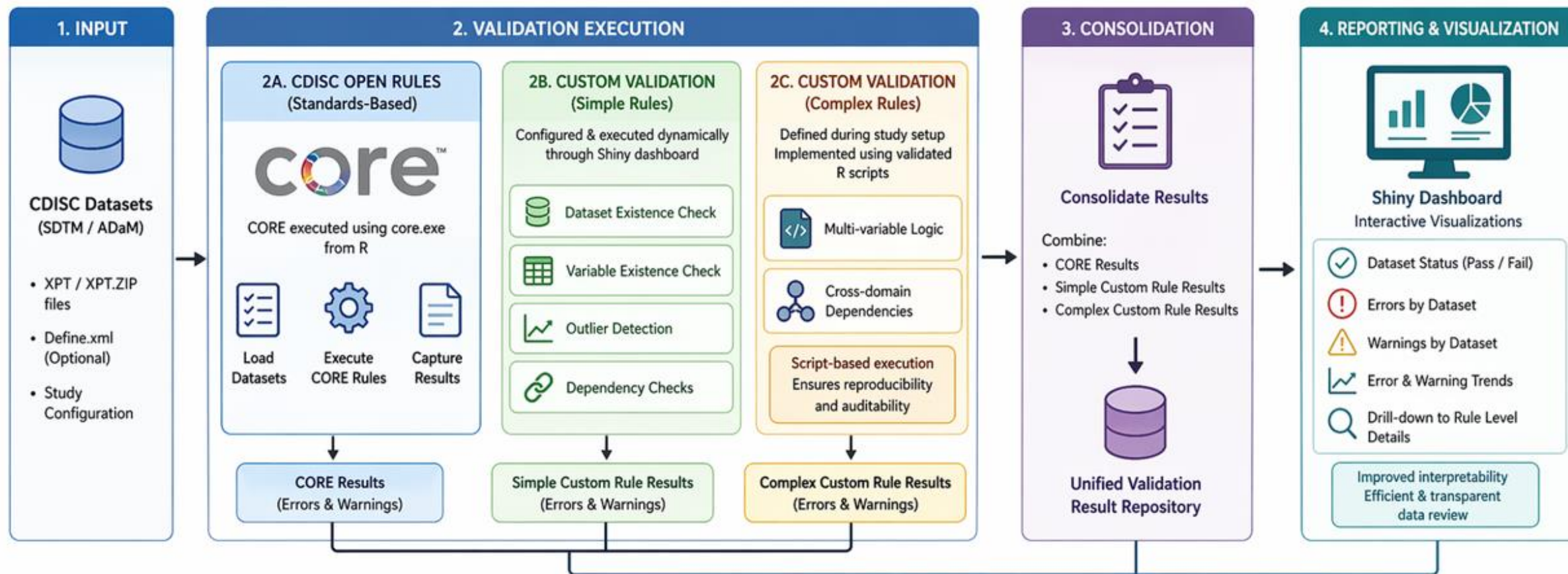
Enhancement

- Study-specific validation needs
- Complex scientific and protocol specific logic
- Standard checks to cover protocol-specific data needs

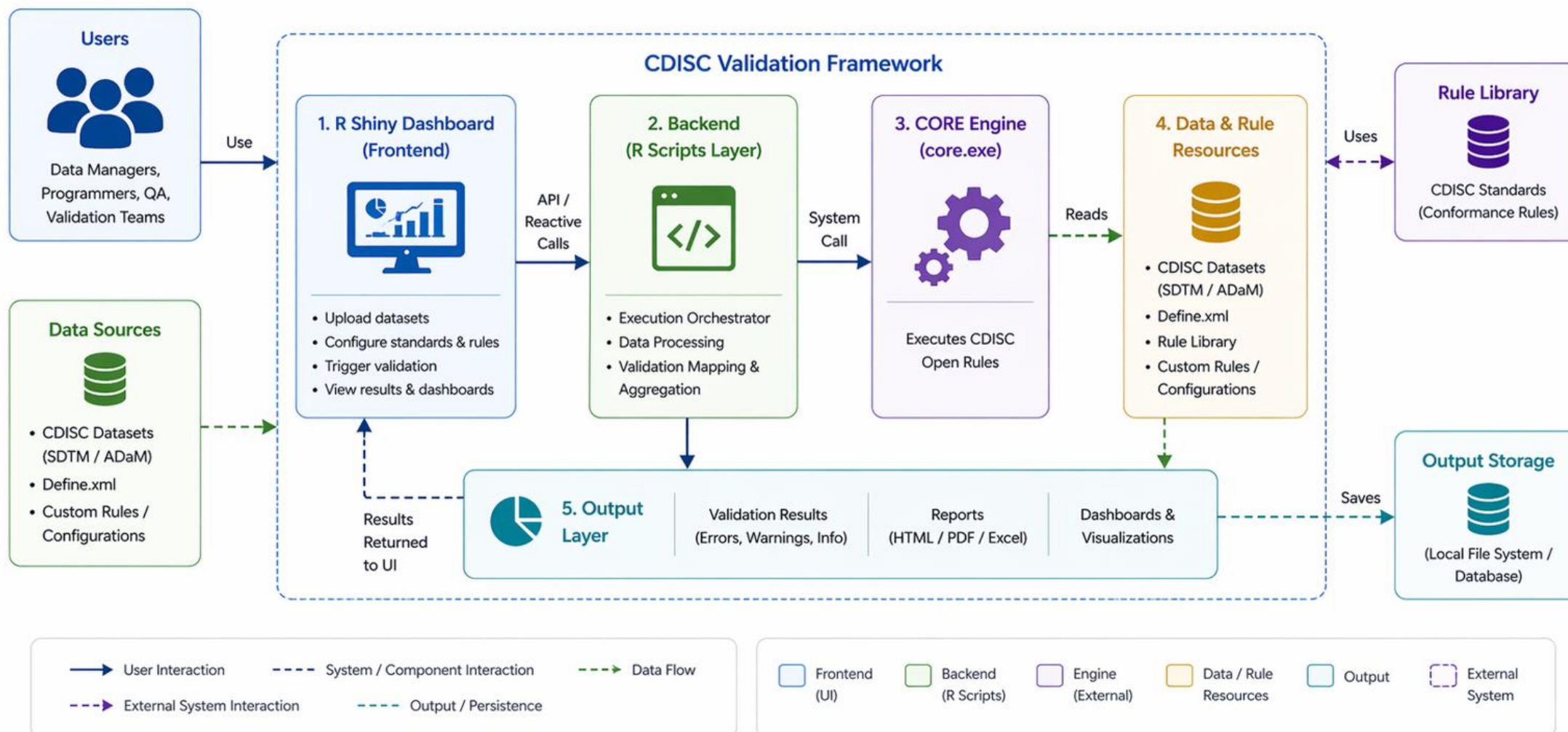


Core Integration with R

Introduction

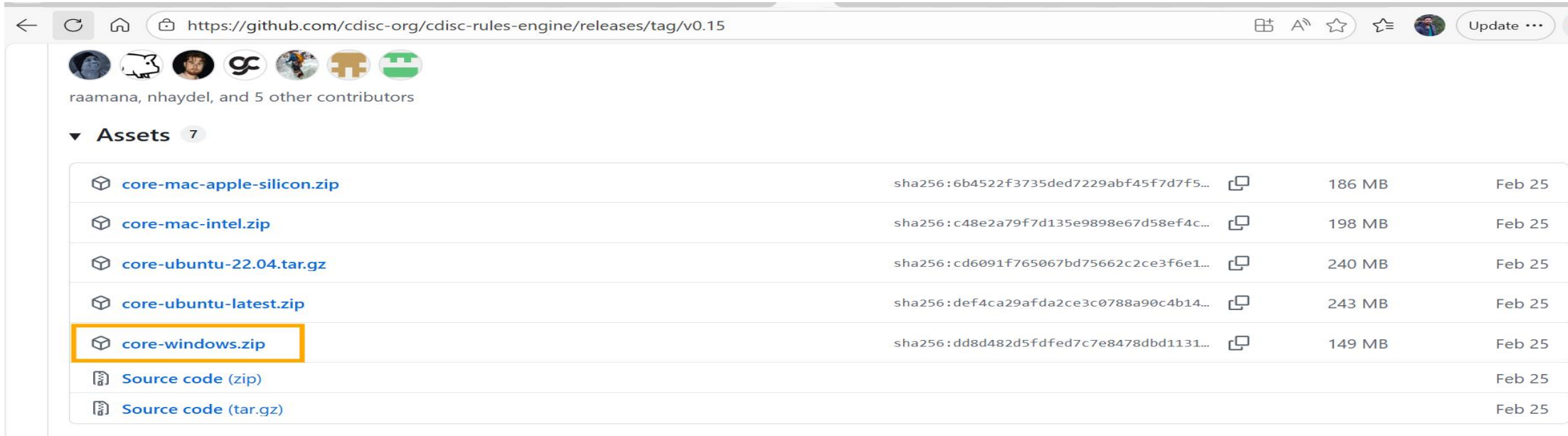


System Architecture



Invoking core using R

<https://github.com/cdisc-org/cdisc-rules-engine/releases/tag/v0.15>



raamana, nhaydel, and 5 other contributors

▼ Assets 7

 core-mac-apple-silicon.zip	sha256:6b4522f3735ded7229abf45f7d7f5...		186 MB	Feb 25
 core-mac-intel.zip	sha256:c48e2a79f7d135e9898e67d58ef4c...		198 MB	Feb 25
 core-ubuntu-22.04.tar.gz	sha256:cd6091f765067bd75662c2ce3f6e1...		240 MB	Feb 25
 core-ubuntu-latest.zip	sha256:def4ca29afda2ce3c0788a90c4b14...		243 MB	Feb 25
 core-windows.zip	sha256:dd8d482d5fdfed7c7e8478dbd1131...		149 MB	Feb 25
 Source code (zip)				Feb 25
 Source code (tar.gz)				Feb 25

```
core_dir <- getwd()
temp_dir <- tempfile(pattern = "engine_data_")
core_exec <- file.path(core_dir, "core", "core.exe")
cdisc_type <- if (input$data_type == "SDTM") "sdtmig" else "adamig"
output_dir <- file.path(core_dir, "output")
if (!dir.exists(output_dir)) {dir.create(output_dir, recursive = TRUE)}
out_file <- file.path(output_dir, paste0(input$study, "_", input$data_type, "_", timestamp))
version_selected <- if (input$data_type == "SDTM") { input$sdm_version
} else { input$adam_version }

cmd <- sprintf(
  '"%s" validate -s %s -v %s -d "%s" -of XLSX -o "%s"', core_exec, cdisc_type, version_selected, temp_dir, out_file
)
system(cmd, wait = TRUE)
```

R shiny Validation Dashboard

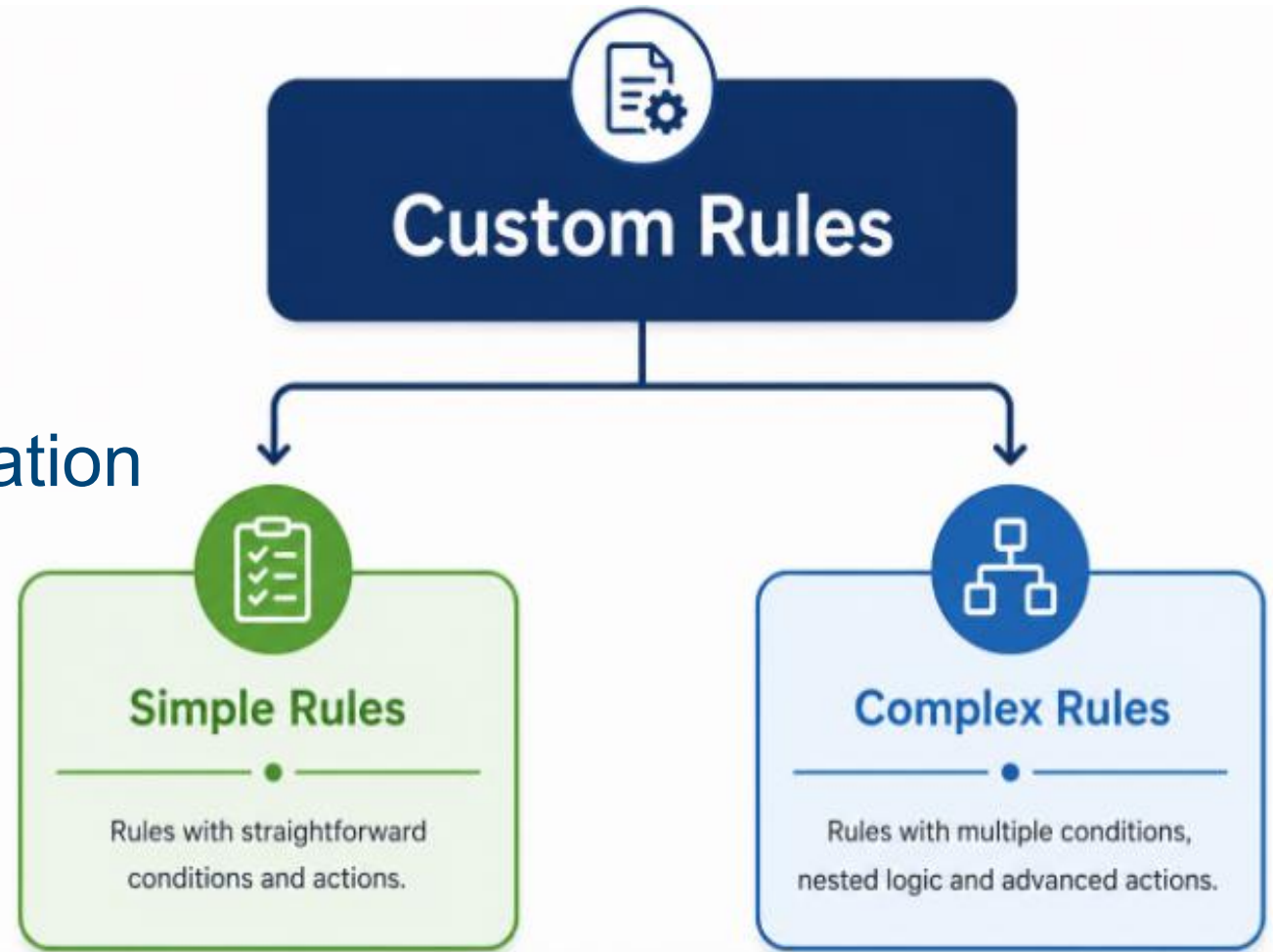
The screenshot displays the CDISC Validation Dashboard interface. On the left, a sidebar contains navigation links for 'Validation Summary' and 'Validation Details'. Below these are configuration options: 'CDISC Data' set to 'SDTM', 'Select SDTMIG Version' set to '3.2', and 'Select Study' set to 'Study1'. The 'Upload Dataset File' section is highlighted with an orange box, showing a 'Browse...' button and a 'No file selected' status. Below this are buttons for 'Custom Rule', 'Run Rule Engine', and 'Download Excel Report'.

On the right, the 'Validation Results - Issue Summary' section is visible. An 'Open' file dialog is overlaid on the dashboard, showing the file explorer for the 'cdsic_app > data' directory. The dialog lists several XPT files: 'vs.xpt', 'ex.xpt', 'ec.xpt', 'ds.xpt', 'dm.xpt', and 'ae.xpt'. The 'vs.xpt' file is highlighted with an orange box. The dialog also shows the file name field, the file type set to 'XPT File (*.xpt)', and buttons for 'Upload from mobile', 'Open', and 'Cancel'.

Name	Status	Date modified
vs.xpt	✓	3/16/2022
ex.xpt	✓	3/16/2022
ec.xpt	✓	3/16/2022
ds.xpt	✓	3/16/2022
dm.xpt	✓	4/15/2022
ae.xpt	✓	4/2/2026

Custom Rules

- UI-driven rule setup (Shiny)
- Instant execution during validation
- No coding required
- Dynamic and reusable



Simple Rules : Overview

- **Dataset & Variable Checks**

- Required datasets are present
- Ensure protocol-specific variables exist

- **Outlier Detection**

- Check numeric values against specific values

- **Dependency Checks**

- Validate basic conditional logic within records

Simple Rules: Variable Existence

The screenshot displays the CDISC Validation Dashboard interface. On the left sidebar, the 'Custom Rule' button is highlighted with an orange box. The main panel shows a modal window titled 'Select Custom Validation Rules' with a list of five rules, each with a checkbox and a domain tag (AE, DM, or AE, DM). Below this list, the '+ Add New Rule' button is highlighted with an orange box. The 'Create Custom Rule' section contains three input fields: 'Variable Existence' (a dropdown menu), 'AE' (a dropdown menu), and 'AESEV' (a text input field). The 'Save Rule' button is also highlighted with an orange box. At the bottom right of the modal are 'Cancel' and 'Apply' buttons.

CDISC Validation Dashboard

Validation Summary

Validation Details

CDISC Data

SDTM

Select SDTMIG Version

3.2

Select Study

Study1

Upload Dataset File

Browse... vs.xpt

Upload complete

Uploaded File: vs.xpt

Custom Rule

Run Rule Engine

Download Excel Report

Validation Results - Issues

Select Custom Validation Rules

- ☐ CUSTOM_01 | AEDECOD variable should not have "UNCODED" | AE
- ☐ CUSTOM_02 | AESER must be Y or N | AE
- ☐ CUSTOM_03 | AGE must be populated | DM
- ☐ CUSTOM_04 | AESER = 'Y' requires AGE ≥ 18 in DM | AE, DM
- ☐ CUSTOM_05 | Screen failure subjects USUBJID with DM.ARM = missing but present in AE domain | AE, DM

+ Add New Rule

Create Custom Rule

Variable Existence

AE

AESEV

Save Rule

Cancel Apply

Simple Rules: Outlier

CDISC Validation Dashboard

Validation Summary

Validation Details

CDISC Data

SDTM

Select SDTMIG Version

3.2

Select Study

Study1

Upload Dataset File

Browse... vs.xpt

Upload complete

Uploaded File: vs.xpt

Custom Rule

Run Rule Engine

Download Excel Report

Validation Results - Iss

Select Custom Validation Rules

☐ CUSTOM_01 | AEDECOD variable should not have "UNCODED" | AE

☐ CUSTOM_02 | AESER must be Y or N | AE

☐ CUSTOM_03 | AGE must be populated | DM

☐ CUSTOM_04 | AESER = 'Y' requires AGE ≥ 18 in DM | AE, DM

☐ CUSTOM_05 | Screen failure subjects USUBJID with DM.ARM = missing but present in AE domain | AE, DM

+ Add New Rule

Create Custom Rule

Outlier

DM

AGE

>

57

Save Rule

Complex Rules: Overview

- Handle advanced logic and multi-step validations
- Developed by R programmers (not UI-driven)
- Integrated as pre – defined code or can be integrated as modular plugins to the validation engine through UI upload feature
- **Examples:**
 - Multi-variable logic
 - Cross-domain dependencies

Complex Rules:

The screenshot displays the CDISC Validation Dashboard interface. On the left, a sidebar contains navigation links: 'Validation Summary' and 'Validation Details'. Below these are sections for 'CDISC Data' (with an 'SDTM' dropdown), 'Select SDTMIG Version' (with a '3.2' dropdown), 'Select Study' (with a 'Study1' dropdown), and 'Upload Dataset File' (with a 'Browse...' button and a 'vs.xpt' file listed). A 'Custom Rule' button is highlighted with an orange box. Below the sidebar, there are buttons for 'Run Rule Engine' and 'Download Excel Report'. The main area shows a 'Validation Results - Issues' section. A modal window titled 'Select Custom Validation Rules' is open, listing four rules. The first three are unchecked, and the fourth is checked. The checked rule is highlighted with an orange box. The modal also features an 'Add New Rule' button and 'Cancel' and 'Apply' buttons at the bottom right, with the 'Apply' button highlighted by an orange box.

CDISC Validation Dashboard

Validation Summary

Validation Details

CDISC Data

SDTM

Select SDTMIG Version

3.2

Select Study

Study1

Upload Dataset File

Browse... vs.xpt

Upload complete

Uploaded File: vs.xpt

Custom Rule

Run Rule Engine

Download Excel Report

Validation Results - Issues

Select Custom Validation Rules

- ☐ CUSTOM_01 | AEDECOD variable should not have "UNCODED" | AE
- ☐ CUSTOM_02 | AESER must be Y or N | AE
- ☐ CUSTOM_03 | AGE must be populated | DM
- ☐ CUSTOM_04 | AESER = 'Y' requires AGE $\geq$ 18 in DM | AE, DM
- ☒ CUSTOM_05 | Screen failure subjects USUBJID with DM.ARM = missing but present in AE domain | AE, DM

+ Add New Rule

Cancel Apply

Validation Results Report

CDISC Validation Dashboard

Validation Summary

Validation Details

CDISC Data

SDTM

Select SDTMIG Version

3.2

Select Study

Study1

Upload Dataset File

Browse... 6 files

Upload complete

Uploaded File: vs.xpt Uploaded File: AE_DM

Custom Rule

Run Rule Engine

Download Excel Report

Validation Results - Issue Summary

Copy CSV Excel

Search:

Dataset	CORE-ID	Message	Issues	Explanation	Executability
ds.xpt	CORE-000364	EPOCH should be null if DSCAT is protocol milestone	18		
ds.xpt	CORE-000586	Multiple records exist within the same EPOCH for a subject when DSCAT = DISPOSITION EVENT	34		
vs.xpt	CORE-000334	At least one expected variable is missing from dataset	1		
vs.xpt	CORE-000550	Variables not listed in the Model List of Allowed Variables for Observation Class should be in SUPPQUAL.	2		
vs.xpt	CORE-000852	Variables are not in the correct order as shown in SDTM for the observation class.	1		
AE_DM	CUSTOM_05	Screen failure subjects USUBJID with DM.ARM = missing but present in AE domain	33		Fully Executable

Showing 1 to 6 of 6 entries

Previous 1 Next

	A	B	C	D	E	F
1	Dataset	CORE-ID	Message	Issues	Explanation	Executability
2	ds.xpt	CORE-000364	EPOCH should be null if DSCAT is protocol milestone	18		
3	ds.xpt	CORE-000586	Multiple records exist within the same EPOCH for a subject when DSCAT = DISPOSITION EVENT	34		
4	vs.xpt	CORE-000334	At least one expected variable is missing from dataset	1		
5	vs.xpt	CORE-000550	Variables not listed in the Model List of Allowed Variables for Observation Class should be in SUPPQUAL.	2		
6	vs.xpt	CORE-000852	Variables are not in the correct order as shown in SDTM for the observation class.	1		
7	AE_DM	CUSTOM_05	Screen failure subjects USUBJID with DM.ARM = missing but present in AE domain	33		Fully Executable
8						

< >

Conformance Details

Dataset Details

Rules Report

Issue Summary

Issue Detail

...

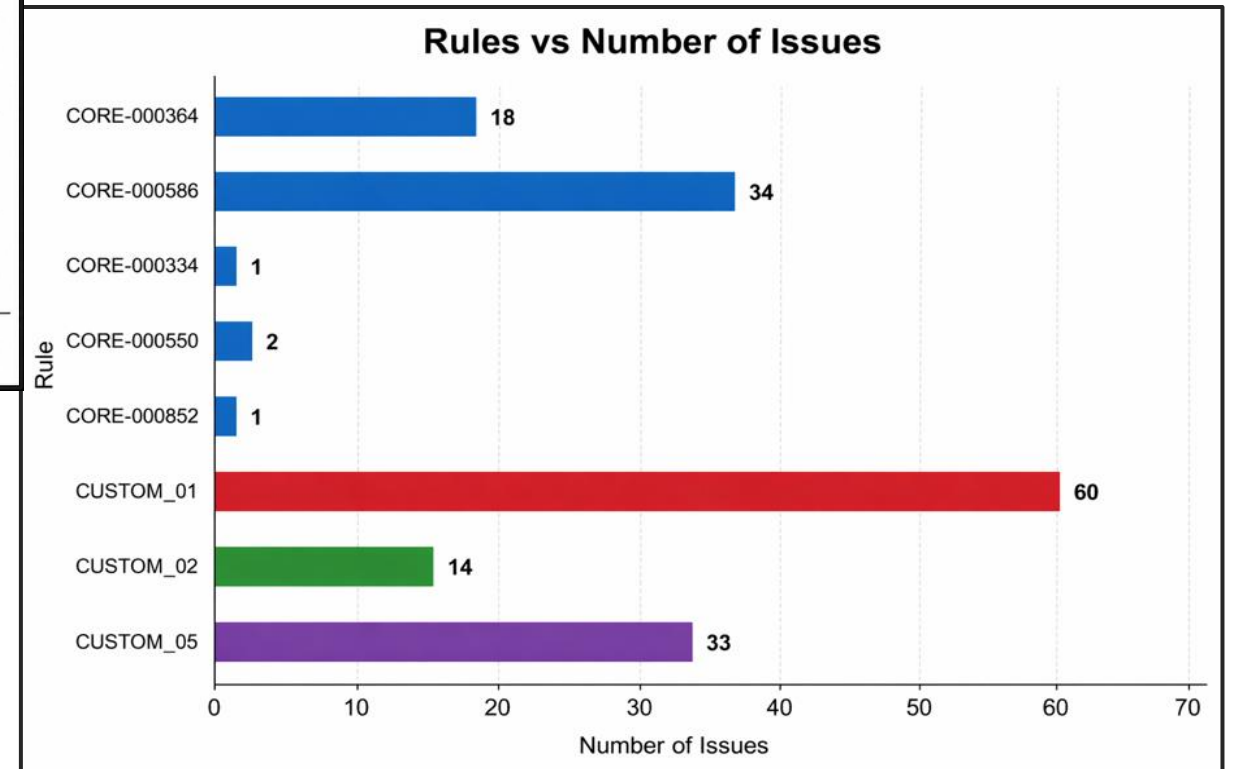
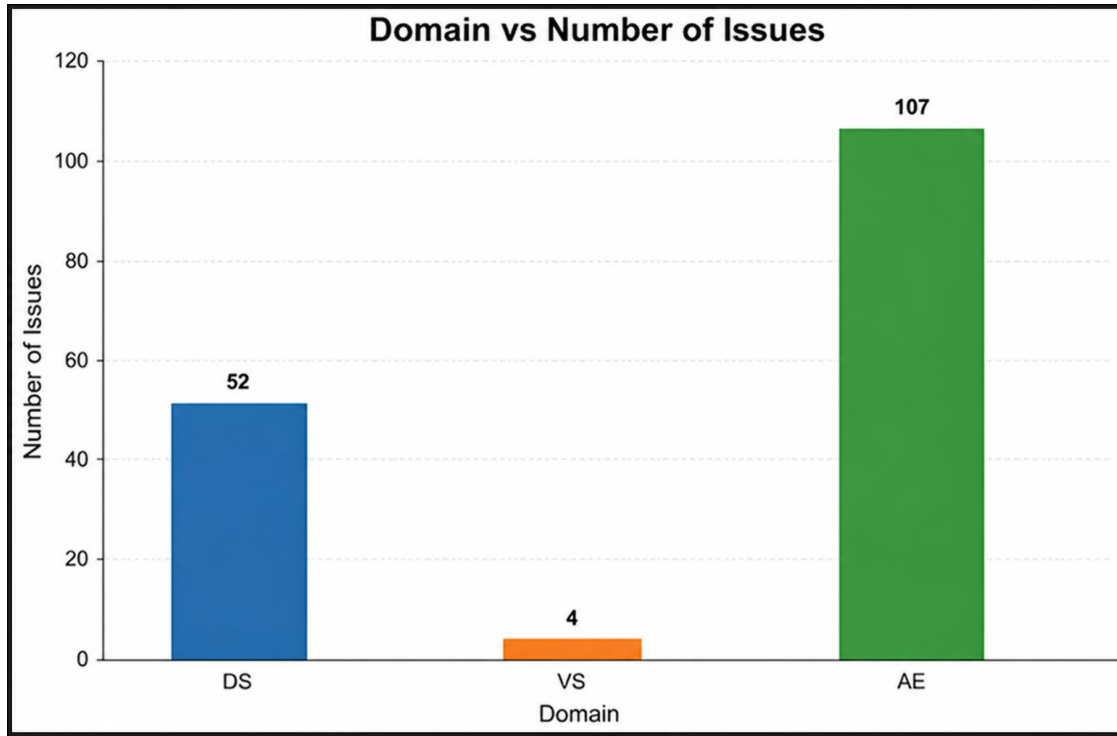
+

:

Reporting Dashboard

Validation Results - Issue Summary					
Copy CSV Excel			Search:		
Dataset	CORE-ID	Message	Issues	Explanation	Executability
ds.xpt	CORE-000364	EPOCH should be null if DSCAT is protocol milestone	18		
ds.xpt	CORE-000586	Multiple records exist within the same EPOCH for a subject when DSCAT = DISPOSITION EVENT	34		
vs.xpt	CORE-000334	At least one expected variable is missing from dataset	1		
vs.xpt	CORE-000550	Variables not listed in the Model List of Allowed Variables for Observation Class should be in SUPPQUAL.	2		
vs.xpt	CORE-000852	Variables are not in the correct order as shown in SDTM for the observation class.	1		
AE	RULE_01	AEDECOD variable should not have "UNCODED"	60		Fully Executable
AE	RULE_02	AESER must be Y or N	14		Fully Executable
AE,DM	RULE_05	Screen failure subjects USUBJID with DM.ARM = missing but present in AE domain	33		Fully Executable
Showing 1 to 8 of 8 entries			Previous 1 Next		

Data Visualization





Conclusion & Roadmap

Conclusion

- Standard checks provide only a partial view of data integrity
- This R Shiny framework unifies standard checks with study-level requirements
- Enables a stronger foundation for study teams
- Supports regulatory submission

Future Outlook & Roadmap

- **Automated Resolution Suggestions:** Providing actionable, rule-based guidance (or AI-driven suggestions) to help users quickly fix errors flagged by each specific rule.
- **Continuous Validation:** Run real-time checks throughout the study
- **Predictive Analytics:** Use past trends to prevent common issues
- **System Integration:** Connect EDC directly to the validation framework

Disclaimer and Disclosures

- The views expressed are those of the author and do not necessarily reflect those of Efficacy Lifescience Analytics.



Thank You!



“It is not the strongest of the species that survives, nor the most intelligent... it is the one most adaptable to change.” -Charles Darwin

Questions?

