

meRlin: A Code Assistant for Clinical Trial Tables

Torben Brehme¹, Steven Brooks², Pietro Mascheroni¹, Rajan Sundar Sri Kannan³, Treesa Vadakkumpadath⁴

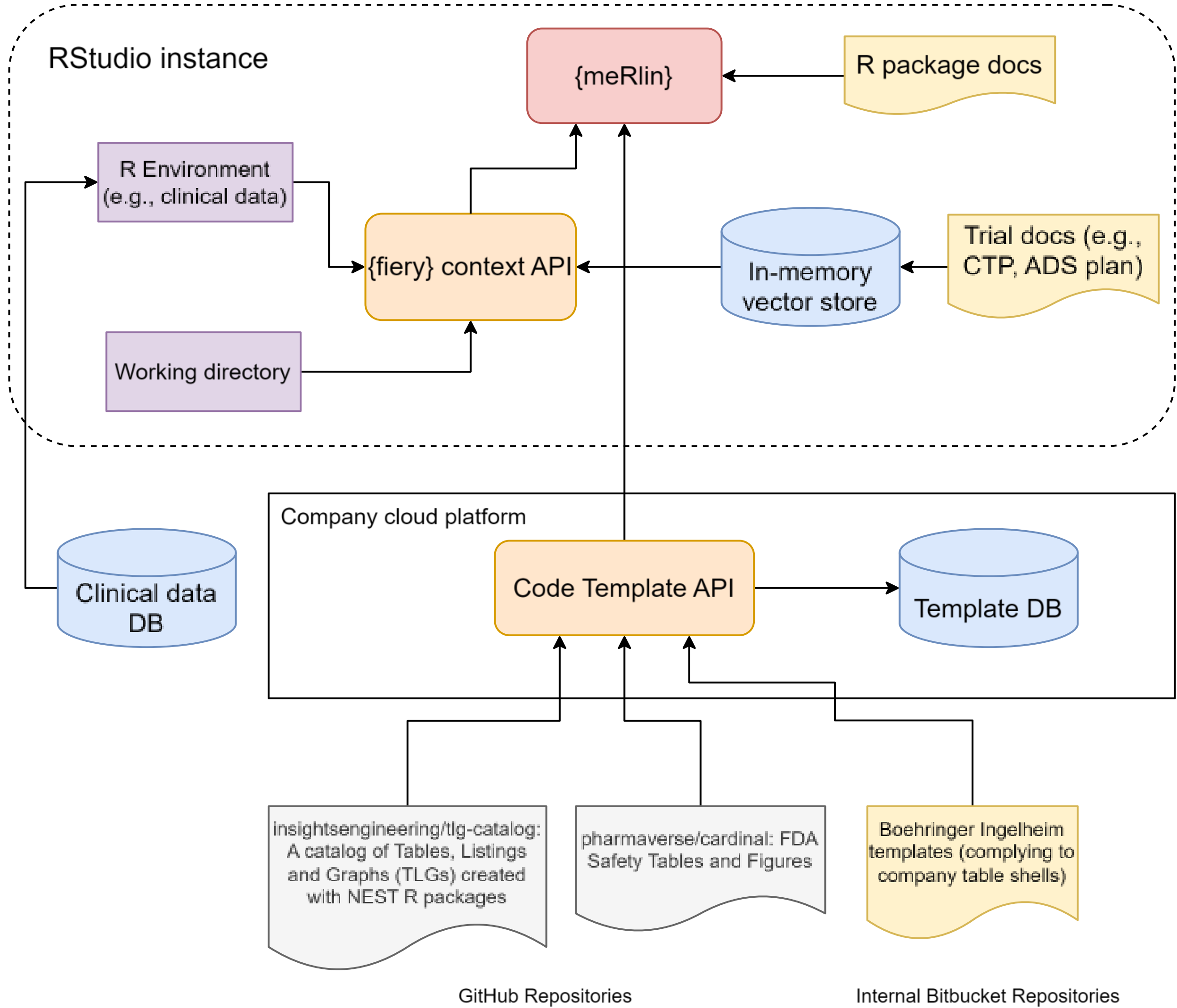
¹Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach an der Riß, Germany; ²Boehringer Ingelheim (China)Investment Co., Ltd., Shanghai, China; ³Technical University of Munich, Garching, Germany; ⁴University of Kassel, Kassel, Germany

Introduction

Tables, listings, and figures (TLFs) are key components in clinical trial reporting, yet their generation remains a time-consuming task. Programmers often face tight deadlines and must ensure outputs are both accurate and traceable. In this work, we present an **agentic workflow** for automating table creation using large language models (LLMs). Our code assistant, **meRlin**, accepts requests for a TLF element from the user, retrieves relevant R code templates, and accesses clinical data and trial documents such as the ADS plan and Clinical Trial Protocol (CTP). Based on these inputs, meRlin adapts the code templates to **generate tailored R code for each TLF element**. Programmers can then review, execute, and debug the code with the assistant’s support. We illustrate the approach with reference table examples and discuss future enhancements.

Methods

meRlin is built as an R Shiny App that can be run within an RStudio instance. It integrates multiple data sources and tools to support table generation workflows.



Support to different LLM providers: the LLM powering meRlin can be chosen among several open source and proprietary alternatives. The integration is done via the `ellmer` package.

Data Access and Environment Integration: meRlin connects directly to the R environment containing clinical datasets through a fiery-based context API. This allows real-time access to dataframes and clinical variables needed for table generation.

Template and Documentation Resources: The agent accesses curated template databases synced to company cloud resources. The current setup includes the repositories:

- `insightsengineering/tlg-catalog`,
- `pharmaverse/cardinal`,
- `Boehringer-specific templates`, adhering to company table shells.

Document Understanding: meRlin utilizes an in-memory vector store to allow semantic searches over user-provided documents, such as trial protocols, statistical analysis plans, and ADS plans.

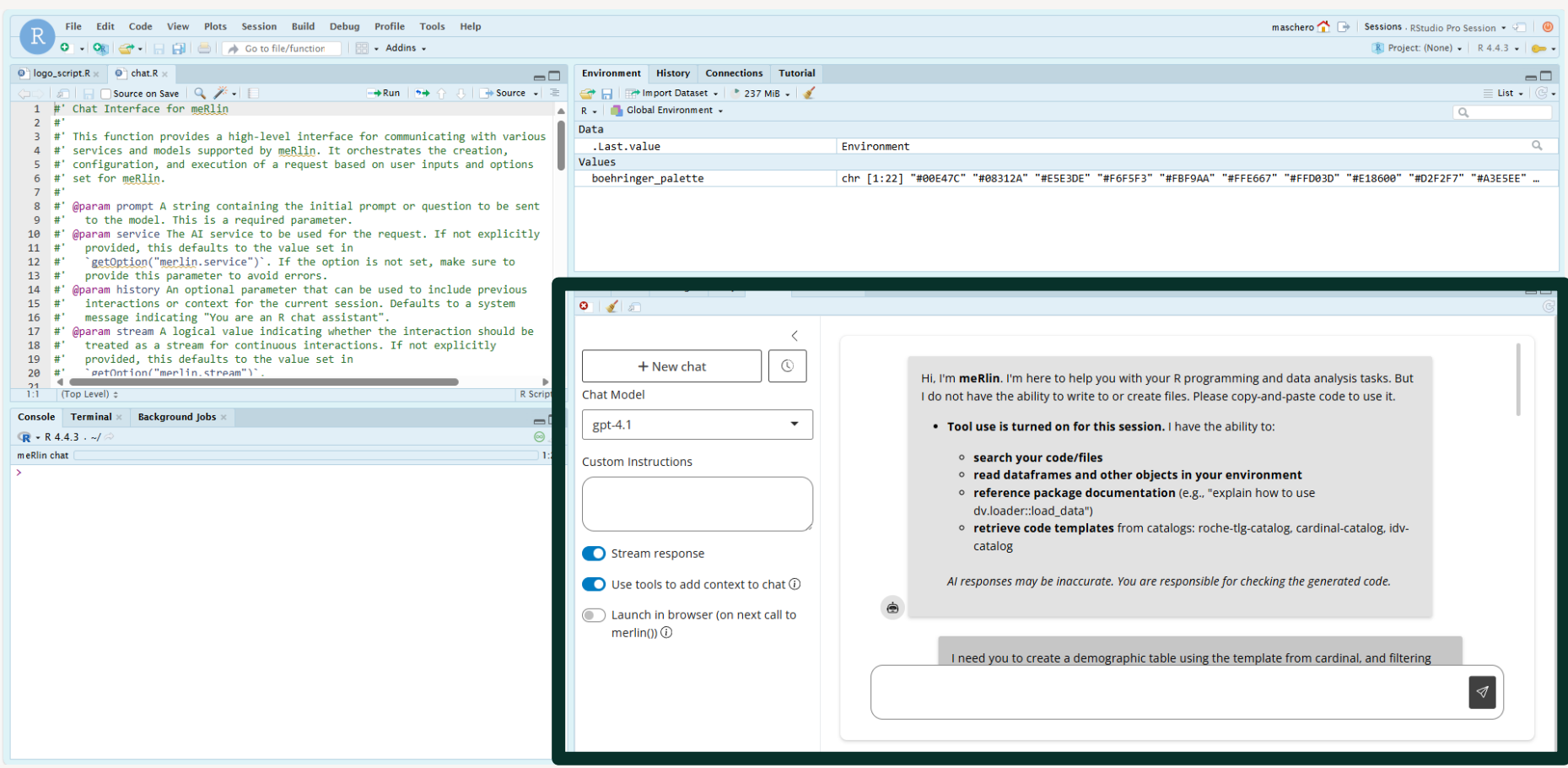
Code Generation Process:

To produce relevant clinical reporting tables, meRlin:

1. Identifies available datasets and variables in the R environment,
2. Retrieves relevant table templates from a structured catalogue,
3. Extracts study-specific logic from protocol documents,
4. Generates R code using packages like `rtables`, `tern`, and `cardinal`,
5. Refines and formats the code to match TLF shell specifications through iterative user interaction.

Results

As a reference, we show how meRlin can generate the code for a demographic table, using the clinical data that are loaded in the current R session and the template retrieved from the `cardinal` repository.



meRlin can be run in the Viewer panel of RStudio, or as a standalone app in the browser.

I need you to create a demographic table using the template from `cardinal`, and filtering patients for being treated and for having a polymorphism type V1/V1. Use the `sample_adsl` data from my R environment and make use of the ads plan to select the proper filters.

Here's how to create a demographic table for patients who are both treated and have SIRPA polymorphism type "V1/V1", using your `sample_adsl` data, the FDA Table 02 template from the `cardinal` catalog, and in line with typical ADS plan filters:

- 1. Filter Logic (from your data and ads plan):**
 - Treated patients: `sample_adsl$TRTFL == "Y"`
 - Polymorphism type: `sample_adsl$SIRPA == "SIRPA V1/V1"`
- 2. Demographic Table Variables (per template):**
 - Typically includes AGE, SEX, RACE, ETHNIC, and derived categories like `AGEGR1`.
- 3. R Code Example:** Below is a reproducible modifying example mixing your data and the `cardinal` template—simply copy, run, and you can further modify for output/format needs:

```
# Load required packages
library(dplyr)
library(gtsummary)

# Filter for Treated AND SIRPA V1/V1
data_demog <- sample_adsl %>%
  filter(TRTFL == "Y", SIRPA == "SIRPA V1/V1")

# (Optional) Derive age group in line with your ADS plan if not already present
data_demog <- data_demog %>%
  mutate(
    AGEGR1 = case_when(
      AGE < 18 ~ "< 18 years",
      AGE >= 18 & AGE < 35 ~ "18 - < 35 years",
      AGE >= 35 & AGE < 65 ~ "35 - < 65 years",
      AGE >= 65 & AGE < 75 ~ "65 - < 75 years",
      ...
    )
  )
```

Conclusions

meRlin demonstrates the potential of AI agents to transform clinical trial reporting by automating TLF generation.

Reduced programming time for routine tables; improved consistency across studies; enhanced accessibility for non-programmers to generate clinical summaries.

Expanding template coverage; enhancing validation workflows; integrating with existing clinical trial management systems.

As regulatory authorities increasingly accept AI-assisted analyses, meRlin positions organizations to leverage these technologies while maintaining the rigor and traceability required for drug submissions.

References

Yang, Yumeng, et al. "Using Large Language Models to Generate Clinical Trial Tables and Figures." arXiv preprint arXiv:2409.12046 (2024).
Minaee, Shervin, et al. "Large language models: A survey." arXiv preprint arXiv:2402.06196 (2024).
Wang, Lei, et al. "A survey on large language model based autonomous agents." Frontiers of Computer Science 18.6 (2024): 186345.

shiny: Web Application Framework for R, <https://CRAN.R-project.org/package=shiny>
ellmer: Chat with Large Language Models, <https://ellmer.tidyverse.org>
fiery: A Lightweight and Flexible Web Framework, <https://fiery.data-imaginist.com>
tern: Create Common TLGs Used in Clinical Trials, <https://insightsengineering.github.io/tern/>
cardinal: FDA Safety Tables and Figures, <https://github.com/pharmaverse/cardinal>
Abbreviations
TLF: Tables, Listings and Figures; LLM: Large Language Model; CTP: Clinical Trial Protocol

Disclaimer
© 2025 Boehringer Ingelheim International GmbH. All rights reserved.
This poster and its contents are property of Boehringer Ingelheim and are, inter alia, protected by copyright law. Complete or partial passing on to third parties as well as copying, reproduction, publication or any other use by third parties is not permitted. The findings and interpretations presented in this poster are solely those of the authors and do not necessarily reflect the views or positions of Boehringer Ingelheim.

