

Ein bisschen KI schadet nie? – Wie LLMs das Programmieren in R erleichtern

FDZ Frühjahrsakademie / Spring Academy 2026

- ♦ Dozierender: Dr. Julien Patrick Irmer (Albert Ludwigs Universität Freiburg)
- ♦ Termin:
 - Teil 1: Mittwoch, 11.03.2026, 14:00 – 18:00 h
 - Teil 2: Donnerstag, 12.03.2026, 14:00 – 18:00 h

Abstract

Wie kann Künstliche Intelligenz den R-Workflow sinnvoll erweitern – ohne die notwendige statistische Expertise zu ersetzen?

Dieser Workshop vermittelt praxisnahe Strategien zum Einsatz von Large Language Models (LLMs) in der quantitativen Datenanalyse mit R. Anhand realer Anwendungsfälle lernen die Teilnehmenden, wie KI den Prozess von Datenaufbereitung über Modellierung bis zur Ergebnisdarstellung effizient unterstützen kann – immer mit „Human in the Loop“. Der Fokus liegt auf praktikablem Prompt-Engineering, der kritischen Bewertung KI-generierter Vorschläge und der Integration von KI in bestehende R-Workflows. Ziel ist es, Teilnehmende in die Lage zu versetzen, KI sinnvoll, sicher und reproduzierbar für ihre eigene Forschung oder Lehre einzusetzen.

Inhalte

Nach einer kurzen Einführung in Funktionsweise, Stärken und Grenzen von KI-Systemen wird der klassische Ablauf von Datenimport bis Reporting skizziert und gezeigt, an welchen Stellen KI praktisch unterstützen kann. Ein besonderer Schwerpunkt liegt auf effektivem Prompt-Engineering, der Ideenfindung für Analysen sowie der KI-gestützten Datenaufbereitung und Modellierung. Gleichzeitig wird vermittelt, wie wichtig die kritische Prüfung KI-generierter Vorschläge bleibt („Human in the Loop“) und wie KI beim Schreiben, Dokumentieren und Präsentieren von Ergebnissen bei der Grafikerstellung oder bspw. in RMarkdown oder Quarto unterstützen kann. Abschließend werden zentrale Fallstricke wie Halluzinationen, Bias und Datenschutz sowie pragmatische Best Practices diskutiert.

Wir werden hierzu ChatGPT und GitHub-Copilot verwenden. Das meiste kann allerdings auch einfach mit einem anderen Chat-basiertem KI-System gemacht werden (siehe Software).

Themenschwerpunkte im Überblick:

- Grundlagen von LLMs und typische Grenzen in statistischen Kontexten
- Der klassische R-Workflow und Ansatzpunkte für KI-Unterstützung
- Prompt-Engineering für zuverlässigen Code und Analysevorschlge
- KI in Datenaufbereitung, Modellierung und Ergebnisinterpretation
- Good Practices, Fehlerquellen und verantwortungsvoller KI-Einsatz

Voraussetzungen

- Grundkenntnisse in R (Daten einlesen, einfache Modelle, Packages installieren)
- Basiswissen in statistischer Datenanalyse (Data-Handling in R, Regression, Linear Mixed Effects Models) sowie grafischer Darsetzung (ggplot2)
- Offenheit für KI-gestützte Workflows und Experimentieren

Software

- R in der jeweils aktuellen stabilen Version
- RStudio / Posit Desktop als Entwicklungsumgebung
- Ein Zugang zu einem Large-Language-Model-basierten Chat-Service (z. B. ChatGPT, Gemini)
- Optional: GitHub-Konto und GitHub Copilot zur Unterstützung bei Code-Vervollständigung

Im Kurs werden wir ChatGPT sowie (GitHub)-Copilot verwenden