Culture Clashes: LLM Support in the Engineering of Safety-Critical Systems

Carsten Thomas, HTW Berlin, Berlin, Germany
Michael Wagner, Edge Case, Pittsburgh (PA), USA

Große Sprachmodelle (LLMs) haben sich als leistungsstarke Werkzeuge in der System- und Softwareentwicklung bewährt – und ihre rasant wachsenden Fähigkeiten machen sie zunehmend attraktiv für den Einsatz in nahezu allen Bereichen des Engineerings komplexer und teilweise auch sicherheitskritischer Systeme. Dennoch bleiben diese Werkzeuge fehleranfällig, und die automatisch generierten Artefakte entsprechen häufig nicht den Qualitätsstandards, die sicherheitskritische Systeme zwingend erfordern.

Wir beleuchten in unserem Beitrag systematisch die vielfältigen Einsatzmöglichkeiten von LLMs entlang des gesamten Engineering-Lebenszyklus sicherheitskritischer Systeme, benennen die damit verbundenen Risiken und arbeiten konkrete Strategien zur Risiko-Beherrschung heraus. Wir definieren mit der IAO-Taxonomie einen neuen Ansatz zur Klassifizierung der LLM-Anwendungsfälle anhand von drei zentralen Dimensionen: den potenziellen Auswirkungen des LLM-Einsatzes, dem Grad der LLM-Autonomie sowie der Beobachtbarkeit der erzeugten Artefakte. Auf dieser Basis vergleichen wir die benannten Strategien zur Risiko-Minimierung mit bewährten Ansätzen aus dem klassischen Systems- und Software-Engineering und ordnen diese ein.

Darüber hinaus untersuchen wir die kulturellen und psychologischen Aspekte, die das Vertrauen in LLM-basierte Engineering-Werkzeuge beeinflussen, sowie die Risiken sowohl übermäßiger Abhängigkeit als auch ungerechtfertigter Ablehnung.

Unsere Analyse zeigt, dass LLMs als Engineering-Unterstützungswerkzeuge erheblichen Nutzen bieten können, jedoch auch signifikante Entwicklungsrisiken verursachen können, wenn sie ohne angemessene Schutzmaßnahmen eingesetzt werden. Auf Grundlage dieser Erkenntnisse schlagen wir praxisorientierte Leitlinien für den verantwortungsvollen Einsatz LLM-basierter Werkzeuge im Engineering sicherheitskritischer Systeme vor.

(Der Beitrag ist eine erweiterte Version eines Vortrags auf der SAE WCX 2026, April 2026)