An Ontology Driven Framework for Reliable Generative AI in Safety Engineering

Markus Schweizer, Corporate Research, Robert Bosch GmbH
(Markus.Schweizer@de.bosch.com)
Heiko Lerch, Senior Expert Functional Safety, ITK Engineering GmbH
(Heiko.Lerch@itk-engineering.de)

Abstract

The creation of safety-critical artifacts, as mandated by standards like ISO 26262, is knowledge-intensive, time-consuming and heavily manual process. While generative AI offers significant potential for automation, its inherent limitations regarding reliability, traceability, and consistency currently prevent its direct application in safety-critical domains. Issues such as AI "hallucinations" and the non-deterministic nature of large language models (LLMs) pose unacceptable risks for the development of norm-compliant safety documentation.

This presentation introduces a novel framework designed to overcome these challenges. The framework embeds generative AI within a structure of ontologies and formal process descriptors inspired by category theory. The core of the framework is a safety ontology that formalizes the domain knowledge of functional safety, system architecture, and their relationships into a structured, validated knowledge graph. This graph serves as the single source of truth for the AI, significantly reducing the risk of generating incorrect information.

Formal process descriptors are used to unambiguously describe the individual process steps, such as the transformation of a system architecture description into a Hazard and Risk Analysis (HARA). By providing the AI with precise, enriched prompts based on ontology and process descriptors, the framework guides the generation process. This ensures that the resulting artifacts are not only created efficiently but are also traceable, deterministic, and internally consistent. A subsequent validation module automatically checks the generated artifacts against the rules defined in the knowledge graph, providing a consistency report for expert review.

Beyond the conceptual framework, this talk will bridge theory and practice. We will present the results and share experiences from a recent collaboration where the framework was adopted by an engineering team and applied to a real-world automotive project. This case study will highlight the practical effectiveness of the approach in a commercial setting, share key insights gained, and discuss the lessons learned during its implementation. Attendees will gain insight into a structured, verifiable method for integrating generative AI into safety-critical workflows, backed by tangible results from its industrial application.